P Periodica Polytechnica Transportation Engineering

44(2), pp. 105-114, 2016 DOI: 10.3311/PPtr.8609 Creative Commons Attribution ①

RESEARCH ARTICLE

Heterogeneity of Driving Behaviors in Different Car-Following Conditions

Luigi Pariota^{1*}, Francesco Galante¹, Gennaro Nicola Bifulco¹

Received 26 September 2015; accepted 16 November 2015

Abstract

Many application fields in transportation engineering can benefit from an accurate modelling of car-following behavior. In particular, in recent years, an increased importance is assigned to embed behavioral abilities in ADAS (Advanced Driving Assistance Systems) and in driving automation solutions. However, accurate development of car-following models needs for accounting of the drivers' heterogeneity, which can be easily observed in car-following data. This paper contributes to analyze different sources of heterogeneity with particular focus on three factors: the dispersion over-time of the behavior of a single driver; the heterogeneous behaviors of different drivers; and the possible bias introduced by some oversimplification of the modelling framework, with particular reference to the type of leading vehicle. Our analyses are based on the observation of car-following trajectories collected in a large experiment involving one hundred drivers. Observed behaviors have been interpreted by means of several car-following models proposed in past. The comparison of the values of the parameters identified for the models (versus observed data) is adopted for the analyses. Moreover, directly observed variables (car-following speed and spacing) are adopted to complement and confirm the analyses. Results show that the greater among the sources of dispersion is the across-driver heterogeneity and that by taking into account such an inherent drivers' dispersion of car-following behaviors it is possible to better identify also the effect of the modelling oversimplifications induced by not considering the type of leading vehicle.

Keywords

Driving behavior, car-following behavior, drivers' heterogeneity, ADAS, driving automation

¹Department of Civil Building and Environmental Engineering, University of Naples Federico II, Via Claudio 21, 80125 Napoli, Italy

* Corresponding author, e-mail: luigi.pariota@unina.it

1 Introduction

The importance of car-following (CF) studies is widely recognised, as they contribute to the practical development of different branches of traffic engineering. Early developments were aimed to the microscopic simulation of traffic flows; well-known contributions in this field come from the popular GHR model (Gazis, et al., 1959) and from the Gipps' modelling framework (Gipps, 1981). Since CF applies to the interaction of nearby vehicles in the same lane, it has a significant role in traffic safety studies too (Cheol and Taejin, 2010). Recently, CF studies have gained a renewed attention as they can be adopted in order to interpreter and/or mimic the drivers' behaviour for Advanced Driving Assistance Systems (ADAS), thus allowing for a better (more *human-like*) interaction in some driving automation solution (Bifulco et al., 2013a).

However, the drivers' behaviour in CF can be easily observed to be heterogeneous and a proper analysis (and comprehension) of the sources of this heterogeneity is a key factor for further development in ADAS and/or for enhancing the interaction between human driving logics and automatic-driving systems (Csiszar and Foldes, 2015).

An inherent source of heterogeneity depends on the dispersion of behaviours *across-drivers*, as different drivers react differently to similar stimuli. Moreover, the same driver can exhibit different reactions (to the same stimuli) in different driving sessions or even within the same driving session (*within-driver* heterogeneity). Finally, some of the (apparent) heterogeneity can be explained by some variables often not explicitly considered in CF models (modelling oversimplification), such as the type of leading vehicle.

This paper aims at investigating within-driver and acrossdriver heterogeneity, as well as the impact of oversimplification. Section 2 describes the experiment in which CF observations have been collected, as well as the elaboration of data required for the analyses. Section 3 details the methodology for the analyses, as well as the adopted modelling background on which some of the analyses are based. In Section 4 the main results are presented, while in Section 5 these results are discussed.

2 Data collection and processing

To achieve the paper's purposes, it is needed to observe the behavior of several drivers (in order to understand to role of the across-drivers heterogeneity), for a relevant duration of time (in order to observe within-driver dispersion), in real-world situations and with different types of leading vehicles. To this end, the most appropriate experimental tool is an Instrumented Vehicle (IV). The IV has been equipped at the University of Naples Federico II. Thanks to 2 TRW Autocruise AC20 radars, the IV allows for collecting, among others, data about the relative distance and the relative speed with respect to up to 4 front and/or 4 rear vehicles. In particular, radars can detect and track a target at a distance from 2 to 150 meters and at a speed from 10 to 200 Km/h. In order to analyse the reactions of the same driver as the leading vehicle changes, the IV is employed in the so-called *active* mode, that is data from the front radar are recorded. Some ego-data are recorded too, mainly the cruising speed of the IV itself. All data are collected at a 10 Hz frequency, synchronized and recorded on-board. Raw data are processed off-line in order to smooth the signals and to avoid spikes and bias. They are subject to a Kalman filter procedure, as described in (Bifulco, et al., 2011); this ensures consistent profiles of cruising speed, acceleration, relative speed and spacing with respect to the leading vehicle. Data are complemented with videos taken by a front-oriented camera, aimed to record the type of the leading vehicle.

The dataset was collected within the Italian research project DRIVEIN2 (DRIVEr monitoring: technologies, methodologies, and IN-vehicle INnovative systems); details about the DriveIn2 project can be found in (Bifulco et al., 2012; Bifulco et al., 2015). The experimental test-site as a whole is a large circular ring with a total length of about 80 Km. In particular a long section (16 Km) of this ring has one-lane per direction, a posted speed limit of 60 Km/h and overtaking isn't allowed. Given the characteristics of the test-site, CF conditions are expected to frequently occur, thus the segment has been adopted as the reference test-bed for our analyses. It is also worth noting that on the selected road-segment CF trajectories are more likely to be associated with heavy goods vehicles (HGV) as leading vehicle. This is due to both the not allowed overtaking and the presence of several factories in the zone (and so an high incidence of HGVs). The IV was driven on the testsite by 100 participants, selected in order to match as closely as possible the population of Italian drivers (see first row of Table 1 below). All driving sessions referred to medium-congestion traffic conditions and stop-and-go phenomena were not observed. Moreover, all observations referred to good weather and day light conditions. Indeed, different environmental (and traffic) conditions can influence the car-following behaviour. However, the aim of the paper is to elaborate on within-driver and across-drivers heterogeneity, as well as on the influence of the type of leading vehicle on the car-following behaviour. In

this context, the environmental and traffic conditions can be considered as a bias; thus we have tried to keep them constant during the experiment.

Collected data have been pre-processed according to the video taken by the front camera. As a result of a manually-made visual analysis, each trajectory observed for each driver has been split in several segments. Each segment is characterized by a unique leading vehicle and by uninterrupted CF conditions. As a result, non-CF conditions are discarded and unique vehicle types are associated to segments. Considered vehicle types have been: motorbikes, cars, light commercial vehicles, busses, HGV. The relative incidence of cars and HGVs is dominant as they cover respectively 70 % and 18 % of the sample. Thus, for sake of simplicity, the only considered segments in our analyses are cars and HGVs. This means that after the visual inspection some driving sessions are discarded from the dataset as being not associated (in the considered road-segment) with any CF condition involving cars or HGVs. The composition of the drivers' population after the visual inspection is reported in the second row of Table 1. In total, 84 drivers have been retained in the dataset, of these 54 % were male, 21 % young (less than 25 years old) and 15 % elderly (more than 65 years old).

Table 1 Composition of the drivers' population

	ТОТ	Male	Young (<25 y.o.)	Elderly (>65 y.o.)
Before visual inspection	100	57 %	20 %	20 %
After visual inspection	84	54 %	21 %	15 %

Identified segments have been 263, for a total duration of about 419 minutes of observation. The resulting mean duration of each segment is about 96 seconds. We will employ these segments for both direct and indirect analyses (based in the identification of modelling parameters). In previous studies (e.g. Bifulco et al., 2013a) we have shown that in some cases even 30 seconds are enough to identify behavioural parameters of car-following. The share between cars and HGVs of considered segments is 64 % (for 36 % of the segments the leading vehicle is an HGV). In terms of duration, the segments with a car as the leading vehicle stand for 76 % of the total. The previous figures are summarised in Table 2. In order to investigate behaviour heterogeneity and modelling oversimplification, it is convenient to identify equilibrium car-following conditions. This allows for eliminating from the analyses the bias due to transients in leader-follower adjustment of trajectories or in transition from different leading vehicles. The definition of CF equilibrium and the method adopted to identify these conditions are described in Section 3.1 below. As CF equilibria are the main condition in which we analyse drivers' behaviours, the mechanical characteristics (e.g. the braking system) and the driving style of the leading vehicle should have a negligible

impact, as the leader generally drives at a constant speed, a part random fluctuation that can be incorporated in the experimental bias. Given a CF segment, each sub-trajectory of the segment with uniform and uninterrupted equilibrium condition is defined as a *clip*. The application of the methodology to the collected dataset gives rise to the figures shown in Table 3. Note that the mean number of clips identified in each section is similar in the two sub-datasets (cars and HSGs), while the duration of HGV clips is about 20 % higher.

Table 2 Composition of the car-following segments

Car-following segments								
TOT	Time [min] Num. of records	Num of records	Leader = Car					
		% of Segments	% of Time					
263	418.85	251310	64 %	76 %				

3 Methodology and theoretical background

In order to investigate the drivers' heterogeneity of CF behaviours and the effects of modelling oversimplification, we analyse observed behaviours by adopting two different approaches, both based on the dataset described in Section 2. The first approach analyses different CF clips by directly looking at observed measures (spacing and relative speed at equilibrium) and searching for differences and similarities in clips of different drivers (across-drivers heterogeneity), clips of the same driver (within-driver heterogeneity), and clips with different leading vehicles (modelling oversimplification). The second approach consists of identifying the parameters of a CF model for different clips and searching for differences and similarities with reference to the identified values. In order to enhance the robustness of the model-based analyses, the approach is performed three times, with reference to three different models.

3.1 Car-following equilibrium: definition and identification

Equilibrium CF conditions are characterized by a null relative speed and a constant relative distance (equilibrium spacing) between the leading and the following vehicle. Actually, as well known from experimental evidences and from psychophysiological CF models, the follower oscillates around the equilibrium conditions by continuously adjusting the relative

speed and the relative distance. The result of these oscillations are, once plot in the plane relative speed vs. relative distance (spacing), the well-known CF spirals. One example of car-following spiral is shown in Fig. 1, where the equilibrium point is the center of the oscillation (point A).



Actually, due to changes in the leader's speed or because of other reasons, more than an equilibrium point can be observed in a CF segment (see Fig. 2). Several equilibria have to be duly identified in order to ensure a proper identification of the equilibrium spacing and of the cruising speed at which the equilibrium spacing holds; similarly, the parameters of the CF models we will employ have to be estimated in correspondence to equilibrium conditions.



Fig. 2 Multipla car-following equilibria and clips

Table 3 Car-following clips								
	Segments	nents Clips	Cling per segment	Time [s]		- Driving Sossions		
			Chps per segment	Mean	Std. Dev.	- Driving Sessions	Chps per unving session	
Car	168	378	2.25	44.48	35.6			
HGV	95	222	2.40	54.8	39.1	84	7.14	
TOT	263	600	2.28	46.9	36.6			

As anticipated in Section 2, an uniform and uninterrupted CF trajectory in equilibrium conditions (oscillating around the same equilibrium point) is here called *clip*. The procedure adopted for selecting clips and identifying CF equilibria is based on a machine learning approach, that allows for automating the process, and for applying it to our large amount of observed data.

For each CF segment a vector is defined at each time instant at which the trajectories have been sampled (with 10 Hz sampling rate). The vector consists of two components: the spacing and the follower's cruising speed, as different equilibria hold at different cruising speeds. Moreover, the time elapsed from the beginning of the segment is associated with the two components of the vector in order to define a time-series, as the clip is an uninterrupted part of the segment. A k-means clustering algorithm has been applied to these time-series; it aims to divide the CF section into k clusters in which each observation belongs to the cluster with the nearest mean, that represents the centroid of the cluster. In practice, given the representation of the CF segment as a time series of *n* observations $(x_1, x_2, ..., x_n)$ x_n), where each observation is a two-dimensional real vector, the k-means clustering aims to partition them into $k \leq n$ sets $S = \{S_1, S_2, \dots, S_k\}$ so as to minimize the within-cluster sum of squares. In other words, the objective is to find:

$$\min_{s} \sum_{i=1}^{k} \sum_{x \in S_{i}} \|x - \mu_{i}\|^{2}$$
(1)

Where μ_i is the mean of the points in S_i .

In the clustering technique the number of clusters k has to be a-priori defined, thus a criterion is essential in order to verify that it has been satisfactorily chosen. In our case the *gapstatistic* suggested by (Tibshirani, et al., 2001) has been used. The statistic, given the clusters obtained after having fixed their number, measures the difference of within-cluster dispersion with respect to a reference distribution obtained from the nonclustered observations. The statistic can be used in an automatic process as follows:

- perform several clustering, varying the tentative number of clusters (*k*) from 1 to a predefined max number (*n*_c); for each trial the clusters are constrained to be composed by adjacent observation times;
- the *gap-statistic* is evaluated for each of the clustering trials;
- the actual number of clusters correspond to the trial with maximum value of the *gap-statistic*.

In our procedure the maximum number of clusters n_c has been fixed to 4. It is the best trade-off between the required computational burden of the procedure and the most likely value, as estimated from previous experiences on a pilot dataset, that have suggested the number of equilibria in a single CF segment is less than 4 with a 92 % probability. The outcome of our procedure is comparable to the one that can be obtained adopting a simpler one, based on the exploitation of some peculiar properties of the CF spirals (see (Pariota, 2013) for details). However, the *gap-statistics* approach has been here chosen as it allow for associating both an equilibrium spacing and an equilibrium speed to each clip, as well as for easily ensuring the clips are composed by adjacent samples of the CF trajectory and for easily automatizing the identification procedure for large datasets.

3.2 Adopted car-following models for parametric description of behaviors

In this section some theoretical background on car-following is quickly recalled, as well as three different car-following models developed by the authors in previous works, The identification of the parameters of these models will be adopted in order to reinforce the analyses directly made with respect to the equilibrium spacing and speed. The main theoretical framework to which we refer is the so-called psycho-physiological approach, as framed within the Action Point (AP) theory, initially developed by (Barbosa, 1961; Todosoiev, 1963) and (Wiedemann, 1974). This approach is based on the Michael's theory (Michaels, 1963), represented in Fig. 3.



Fig. 3 Michael's theory

The stimulus to which drivers respond is the angular velocity $(d\vartheta/dt)$ at which the apparent size of the vehicle ahead changes, where the apparent size is defined as the visual angle (ϑ) subtended by the observed leading vehicle, that is proportional to the current relative speed (Δv), the spacing (Δx), and the size (w) of the leader:

$$\frac{d\theta}{dt} = -w \frac{\Delta v'}{\left(\Delta x'\right)^2} \tag{2}$$

As evident from the Fig. 3, the stimulus changes (ceteris paribus) according to the size of the leader (*w*). This justifies the potential effect of the type of leading vehicle on the car-following behavior and suggests that some modelling oversimplification can be induced if this phenomenon is not explicitly considered. Of course, the relevance of the oversimplification (if any) has to be properly estimated. For instance, some researchers (Parker, 1996; Brackstone, et al., 2009) found that the spacing at which drivers tend to follow HGVs is smaller than that for cars, while some others (Aghabayk et al., 2013) observed larger headways and higher reaction time while following HGVs.

Starting from the AP theory, we have proposed in past works (Bifulco et al., 2013b; Pariota and Bifulco, 2015) a specific interpretation based on our original concept of car-following waves. Indeed, if a typical car-following spiral is plot in a different plane, that is with respect to the rate between relative speed and spacing (-DV/DX), some waves are described (Fig. 4). The extremes of these waves correspond to the opening and closing (velocity-based) action points (OPDV and CLDV) of the Wiedemann's theory (Wiedemann, 1974), and they can be found to be located on a segment. Different drivers (and clips) can exhibit different segments that correspond to different behaviors. These behaviours correspond to different values of the model parameters, in particular we refer to the intercept $(a_1 a)$ always very close to zero) and the slope (a_2) of the segment. The slope (a_2) is expected to be different if the model is calibrated with respect to different types of leading vehicles; moreover, once identified/ calibrated against different observed CF segments or clips, the dispersion of the parameter corresponds to the heterogeneity of the drivers' (and clips) behaviors.



Fig. 4 Car-following waves and action points

The second parametric model is a simple time-continuous linear state-space formulation of CF proposed in Pariota et al. (2015). It is developed for conditions close to the equilibrium; even if it shows to be also reasonably applicable far from the equilibrium. The model is based on a Taylor's linear expansion of the (generally non-linear) car-following dynamic process, written at equilibrium conditions. Given the Taylor's approach, the equation of the process is not needed, while it is needed to know the partial derivatives of the process with respect to the independent variables, that are assumed to be the spacing, the relative speed and the cruising speed. Some mild approximation allows to neglect the derivative with respect to the cruising speed and to show that the remaining derivatives to be known can be treated as the parameters (ω_1 and ω_2) of the behavioral model to be identified. Once these parameters are identified, the linearized model can be put in a state-space form and can be shown (Pariota et al., 2015) to fit well the observed data, even far from the equilibrium conditions, thus suggesting that the CF phenomenon is not too much different from its linearization.

The third parametric model is based on a linear formulation too (Bifulco et al, 2013a). It is directly derived from a time-discrete stimulus-response approach and it is specifically developed for ADAS-oriented solutions. Once properly calibrated, it has been shown to be mimic of real car-following behaviors, as it is aimed at enabling an human-like control logic for adaptive cruise control (ACC) systems (Bifulco et al., 2008). The time-discrete linear model uses a polynomial formula to relate the instantaneous speeds of the leader and of the follower (and their relative spacing) with the target spacing the follower applies for the next simulation time-step. The calibration of the model is carried out by adopting the recursive least squares (RLS) algorithm, as described by (Haykin, 2001). Calibration is separately applied to different drivers and, for each driver, to different clips. This allows for the disaggregated identification of the coefficient of the polynomial formula (vector β). These coefficients are assumed to include the behavioral characteristics of any given driver (and clip) and can be compared in order to check for similarities and differences.

4 Analyses and results

One of our first analyses is oriented to understand if the type of vehicle ahead influences the driving behavior in car-following conditions. Figure 5 shows all observed equilibrium conditions; these are plotted considering both the cruising speeds of the vehicles and the inter-vehicular spacing. Actually, the cruising speed of the follower is shown, but in equilibrium conditions the speed of the leader and the one of the follower are very similar. It can be seen in Figure 5 that the point representative of HGVs as leading vehicles are around 18% of the ones that represent cars. This is consistent with the expected composition of the traffic streams in the study area.



Fig. 5 Scatter plot of the equilibrium spacing vs. the cruising speed

Results show, as expected, that the spacing increases according to the cruising speed. This phenomenon is common to both HGVs and cars. However, data seem to be very dispersed and the dispersion seems to increase as the cruising speed increases.

The same analyses can be done with reference to the time headway. The results are shown in Fig. 6 and are similar to the ones in Fig. 5.



Fig. 6 Scatter plot of the equilibrium headway vs. the cruising speed

The distribution of observed cruising speeds (in equilibrium conditions) is adopted in order to identify classes for a further, more detailed, analysis. As it can be seen from Fig. 7, The sentence allow for considering the great part of the observed (equilibrium) cruising speeds. As the observed cruising speeds are not drastically different between HGVs and cars (see Fig. 7 again), the adopted classes are the same for the two types of leading vehicles.



Fig. 7 Distribution of cruising speeds in the dataset

As 12.5 m/s corresponds to 45 km/h and 21.5 m/s to 77 Km/h, the three adopted classes for the cruising speed are consistent with the traffic characteristics of the national road where the observations are collected. All clips where the equilibrium cruising speed is outside of the previous range are discarded from the analyses. The incidence of the discarded clips is 13 % of the total.

The heterogeneity observed in Fig. 5 and 6 accounts for both the *within-driver* (different clips of the same driver) and *across-drivers* (clips of different drivers) phenomena. Table 4 below differentiates the magnitude of the within-driver and the across-drivers dispersions; $E[\sigma_w]$ is the average (over all drivers) of the standard deviations of the equilibrium spacing for all clips of the given driver, it includes the heterogeneity possibly due to the different type of leading vehicle. Similarly, $E[cv_w]$ is the average of the coefficient of variation corresponding to $E[\sigma_w]$. The parameter σ_a is the standard deviation (and cv_a the

coefficient of variation) computed over all the clips and all the drivers; it measures the across-drivers heterogeneity and, given how it is defined, includes the within-driver heterogeneity.

Table 4 Within-driver and across-drivers heterogenity

Datase	t	Within-driver		Across-Drivers		
$E[\sigma_{_{W}}]$		E[cv _w]	σ_a	cv_a		
06	All	9.47	0.33	15.55	0.58	
pacin	Car clips	8.08	0.30	15.20	0.57	
sı	HGV clips	4.94	0.20	10.78	0.44	
Headway	All	0.49	0.31	0.86	0.57	
	Car clips	0.44	0.29	0.88	0.58	
	HGV clips	0.25	0.17	0.57	0.41	

Note (first row of Table 4) that the magnitude of the acrossdrivers heterogeneity is greater (about 1.6 times) than the within-driver one. This is even more remarkable if we consider that the within-driver dispersion is computed for any single driver (before to be averaged over drivers), thus by using much less observations (per driver) than the across-drivers dispersion. Table 4 also shows that the heterogeneity can be reduced if clips with different leading vehicle are considered (rows 2 and 3); this is more evident for the within-driver heterogeneity than for the across-drivers one. The previous means that the across-drivers heterogeneity can induce a remarkable bias to the analysis of the influence of the type of leading vehicle, and that the more convenient way to analyze the influence of the leading vehicle is with reference to the within-driver heterogeneity. It is also worth noting that comparing rows 2 and 3 with reference to the last two columns of Table 4, the evidence is suggested that the behavior of different drivers is more similar in case the leading vehicle, while is more dispersed if the leading vehicle is a car.

Note that analyzing the phenomenon in terms of spacing or headway is quite equivalent. This enables us to look to just the spacing in the follow of the paper.

It has to be noted that, unfortunately, the typical number of clips for each driver is small and often the same driver does not exhibit both cars and HGVs as leading vehicles of his/her clips. This forces to apply all the analyses hereafter carried out to the dataset as a whole (not distinguishing per driver, even if distinction per type of leading vehicle is considered); however, the bias induced by the across-drivers heterogeneity has to be duly taken into account in discussing the obtained results.

According to Table 4, clips have been considered for different ranges of the equilibrium speed in order to build empirical probability density functions (PDF); these have been built with reference to separately considered clips for cars or HGVs and also for the dataset as a whole (not distinguishing the type of leading vehicle). The empirical distributions have been tested to fit a lognormal random variable. The lognormal distributions fit well the observed data, as verified by applying two nonparametric tests, the Kolmogorov-Smirnov and the Chi-square; results have been reported in Table 5, where it can be seen that p-values all greater than 0.05.

 Table 5 Testing the hypotheses on the appropriateness of the lognormal

 distributions for the fitting of the equilibrium points in the different speed classes

Dataset	Speed Class	Kolmogorov- Smirnov	Chi-square
		p-value	p-value
Car	I ([12.5, 15.5] m/s)	0.44	0.18
	II ([15.5, 18.5] m/s)	0.71	0.87
	III ([18.5, 21.5] m/s)	0.68	0.43
	I ([12.5, 15.5] m/s)	0.61	0.60
HGV	II ([15.5, 18.5] m/s)	0.89	0.60
	III ([18.5, 21.5] m/s)	0.86	0.53

The estimated parameters $(k_1 \text{ and } k_2)$ of the lognormal variable are shown in Table 6 for all combinations of leading vehicle and speed class, the corresponding expected value (E), standard deviation (STD) and coefficient of variation (CV) for the equilibrium spacing are reported too. Table 6 also shows that the values for the k, coefficient are different if cars and HGVs are separately considered, the same applies to the expected value of the spacing. Moreover, in particular for the first class of speed (the lower one), the standard deviation for HGVs is remarkably lower than for cars; thus, it seems that the heterogeneity can be partially controlled by avoiding modelling oversimplification and by explicitly taking into account the type of leading vehicle. The previous hypothesis is statistically tested by checking if the null hypothesis holds that the two sample are from the same random distribution. The applied test is the twosample Kolmogorov-Smirnov. Results are reported in Table 7 below, and show that the null hypothesis is never rejected.

The two datasets (car and HGV) have been also compared via the behavioral models introduced in Section 3.2. For what concerns the state-space model, the process of identification was carried out using the Matlab-System Identification Toolbox. With respect to the stimulus-response model and to the one based on waves, a good fitting of the data is ensured by analysis of the r-square statistics; they have not been reported in Table 7 because for all the clips the values assumed were between 0.85 and 1, without any relevant difference between the datasets.

The values assumed by parameters in the two datasets (car and HGV) are reported in Table 8; provided that they have been evaluated for each drivers and for each clip, the statistical distribution of the identified parameters is described in terms of the 25th, 50th, and 75th percentile.

Table 6 Parameters estimated and their statistics

Datasat	Speed	Lognormal pai	Spacing			
Dataset	Class	k ₁	k ₂	Е	STD	CV ₁
	Ι	2.77	0.45	17.65	8.33	0.47
Car	II	3.08	0.47	24.33	12.21	0.50
	III	3.36	0.47	32.10	15.98	0.50
	Ι	2.72	0.27	15.77	4.29	0.27
HGV	II	3.05	0.46	23.58	11.47	0.49
	III	3.34	0.32	29.60	9.79	0.33
	Ι	2.75	0.40	17.00	7.01	0.41
All	II	3.07	0.47	24.18	12.05	0.50
	III	3.35	0.44	31.57	14.72	0.47

Table 7 Testing the hypotheses on the similarity of the distributions in the different classes for speed at car-following equilibrium points

Speed Class	Kolmogorov-Smirnov
Ι	0.76
II	0.78
III	0.34

Table 3	8 Distri	ibution	of	behay	vioural	parameters
rabic	0 Distri	oution	01	ouna	iouiui	parameter

_	ter			Dat	taset		
Mode	Parame		Car	Car			
A		25 th	50 th	75 th	25 th	50 th	75 th
ate ace	ω ₁	0.037	0.085	0.291	0.045	0.105	0.450
St: sp:	ω_2	0.29	0.56	1.56	0.40	0.81	2.41
	β	-0.324	0.414	1.865	-0.113	0.713	2.869
ulus onse	β_1	-0.058	-0.023	-0.008	-0.062	-0.023	-0.006
Stim Resp	β_2	0.494	0.579	0.619	0.455	0.554	0.599
	β_3	-0.062	0.005	0.065	-0.118	-0.017	0.074
ves	a ₁	-0.0010	-0.0002	0.0006	-0.0007	-0.0001	0.0007
Wa	a ₂	-0.059	-0.042	-0.030	-0.059	-0.042	-0.032

With reference to the state-space model, positive values confirm the rational driving behavior hypothesis (Wilson, 2008). In the stimulus-response model, parameters β_1 and β_3 assume quite low value, thus the main stimulus is the one associated with β_2 , the relative speed. In the waves-based model the parameter a_1 should assume a zero value and this is confirmed for both the datasets. For all behavioral models, the values assumed by the parameters in the two datasets (car and HGV) are similar, this is particular evident for the waves-based model where the only relevant parameter a_2 is almost independent on the dataset (car or HGV).

The same distributions are reported in following figures from 8 to 10 for the most relevant of the previous parameters.



Fig. 8 Parameters ω_1 and ω_2 , comparison for car and HGV



Fig. 9 Parameters β_0 and β_2 , comparison for car and HGV



Fig. 10 Parameter a,, comparison for car and HGV

5 Discussion

Results in Tables 5 and 6 show that actually the lognormal distributions of spacing fit well the observed data, and confirm finding of previous studies. Indeed Greenberg (1966) in his early paper introduced the hypothesis of lognormal distribution, even if with reference to the headway (a quantity is related to spacing); this hypothesis has had various confirmation in following years (Piao and McDonald, 2003) up to the recent paper of Jiang and Lu (2015).

Table 5 shows that average values of the equilibrium spacing increase with the speed. This result is somewhat intuitive, and also confirms references from literature. The dispersion of the equilibrium spacing increases with the speed, showing the car-following behaviors are less heterogeneous at lower speed value; this result has been confirmed in Jiang and Lu (2015) too, with reference to a sample composed by Chinese drivers.

Previous considerations hold if the dataset of all clips is split in two by considering the type of leading vehicle (cars and HGVs). With reference to Table 5, a systematic effect can be observed of the type of leading vehicles toward the values of the parameters of the distributions, moreover the dispersion decreases if equilibrium spacing is separately considered for different classes of leading vehicle. Our results about the influence of the type of the leading vehicle seem to partially agree with those in Parker (1996) and Brackstone et al. (2009). Indeed, also in our dataset the average equilibrium spacing at which drivers tend to follow HGVs is smaller than for cars, for all the speed classes.

It is worth considering that our results about the impact of the class of leading vehicle are not statistically significant (Table 6), as also indicated by the distribution of behavioral parameters (Table 7), which are fully comparable in the two datasets, with the partial exception of the parameters ω_2 and β_0 , that are the most different across the two datasets. They exhibit a not negligible systematic difference, and the values related to the HGVs are systematically greater than the ones related to the cars.

Overall speaking, the observed heterogeneity seems to be not associable with statistical significance to the type of leading vehicle. However, this result could have been effected by the different nature of the heterogeneity. The within-driver one should be more sensitive to the type of leading vehicle, but it is dominated (Table 4) by the across-drivers heterogeneity (that also includes the dispersion due to the type of leading vehicle). Unfortunately, given the size of our sample, we are not able to accurately identify the within-driver heterogeneity for different classes of speed. It should also be noted that the driver behaviour (i.e., aggressive driving) and performance (i.e., safety skills) could differ, sometimes substantially, across cultures/countries. The results obtained in this study cannot be extended to other countries without further ad-hoc studies.

Finally, it's worth noting about the results extension that the study were carried out in specific experimental condition: extraurban road context, medium-congestion traffic, middle class vehicle with manual transmission, daylight and good weather.

6 Conclusions

The paper analyzed within-driver and across-driver heterogeneity, as well as the impact of oversimplification, with particular reference to type of leading vehicle, in car-following conditions.

The results of study confirm finding of previous study such as that the lognormal distributions of spacing fit well the observed car-following data, average values of the equilibrium spacing increase with the speed and the average equilibrium spacing at which drivers tend to follow HGVs is smaller than for cars. Regarding the type of leading vehicle, a systematic effect on observed heterogeneity is evident but, is not statistically significant.

In any case a more general conclusion can be drawn, when a car-following model is adopted for (microscopic) traffic simulation it is likely that the overall properties of the traffic stream as a whole are respected also if the effect of the type of the leading vehicle is not taken into account, as the across-drivers dispersion prevails; of course this result concerns the car-following model parameters, because the presence of a significant percentage of HGVs affects the traffic flow for the difference in the acceleration/deceleration rates of the two vehicle classes. On the other hand, when the behavior of a single driver has to be interpreted or integrated by, for instance, an advanced driving assistance system, the within-driver dispersion has to be taken into account and controlled from a modelling point of view. This can be supported by explicitly considering the type of leading vehicle. An extension of this work will be carried out analyzing also car-following behaviors for the motorways observed within the DRIVEIN2 project (Bifulco, et al., 2012), in such a way to explore also the effect of different contests on the car-following behaviors.

Acknowledgement

The research was partially supported by the APP4SAFETY project (B68F12001150005/PON03PE_00159_3) under grant of the Italian Ministry of Research.

References

Aghabayk, K., Sarvi, M., Forouzideh, N., Young, W. (2013) New Car-Following Model Considering Impacts of Multiple Lead Vehicle Types. *Transportation Research Record: Journal of the Transportation Research Board.* 2390, pp. 131-137. DOI: 10.3141/2390-14

Barbosa, L. (1961) Studies on Traffic Flow Models.

- Bifulco, G. N., Pariota, L., Brackstone, M., Mcdonald, M. (2013) Driving behaviour models enabling the simulation of Advanced Driving Assistance Systems: Revisiting the Action Point paradigm. *Transportation Research Part C: Emerging Technologies.* 36, pp. 352-366. DOI: 10.1016/j.trc.2013.09.009
- Bifulco, G. N., Pariota, L., Galante, F., Fiorentino, A. (2012) Coupling Instrumented Vehicles and Driving Simulators: opportunities from the DRIVE IN2 Project. Anchorage, Alaska. pp. 1815-1820.
- Bifulco, G. N., Pariota, L., Simonelli, F., Di Pace, R. (2011) Real time smoothing of car-following data through sensor fusion techniques. *Procedia -Social and Behavioral Sciences*. 20, pp. 524-535. DOI: 10.1016/j.sbspro.2011.08.059
- Bifulco, G. N., Simonelli, F., Di Pace, R. (2008) *Experiments toward a humanlike Adaptive Cruise Control.* Eindhoven. pp. 919-924.
- Bifulco, G., Pariota, L., Simonelli, F., Di Pace, R. (2013) Development and testing of a fully Adaptive Cruise Control system. *Transportation Research Part C: Emerging Technologies*. 29, pp. 156-170. DOI: 10.1016/j.trc.2011.07.001
- Brackstone, M., Waterson, B., McDonald, M. (2009) Determinants of following headway in congested traffic. *Transportation Research F.* 12, pp. 131-142. DOI: 10.1016/j.trf.2008.09.003
- Cheol, O., Taejin, K. (2010) Estimation of rear-end crash potential using vehicle trajectory data. Accident Analysis & Prevention. 46(6), pp. 1888-1893. DOI: 10.1016/j.aap.2010.05.009
- Galante, F., Mauriello, F, Montella, A, Pernetti, M, Aria, M, D'Ambrosio, A. (2010) Traffic calming along rural highways crossing small urban communities: Driving simulator experiment. *Accident Analysis & Prevention.* 42(6), pp. 1585-1594. DOI: 10.1016/j.aap.2010.03.017
- Gazis, D. C., Herman, R., Potts, R. B. (1959) Car-following theory os steadystate traffic flow. *Operations Research*. 7(4), pp. 499-505. DOI: 10.1287/opre.7.4.499
- Gipps, P. G. (1981) A behavioural car-following model for computer simulation. *Transportation Research Part B: Methodological*. 15(2), pp. 105-111. DOI: 10.1016/0191-2615(81)90037-0
- Greenberg, I. (1966) The log normal distribution of headways. Australian Road Research. 2(7), pp. 14-18.
- Haykin, S. (2001) *Adaptive Filter Theory*. IV a cura di Upper Saddle River. Prentice Hall, New Jersey.
- Jiang, J., Lu, J. (2015) *Car-following behavior of chinese drivers*. Washington, D.C.

Michaels, R. M. (1963) Perceptual factors in car following. Paris. pp. 44-59, .

Montella, A., Aria, M., D'Ambrosio, A., Galante, F., Mauriello, F, Pernetti, M. (2011) Simulator evaluation of drivers' speed, deceleration and lateral position at rural intersections in relation to different perceptual cues. *Accident Analysis & Prevention.* 43(6), pp. 2072-2084. DOI: 10.1016/j.aap.2011.05.030

- Pariota, L. (2013) Driving behavior for ADAS: theoretical and experimental analyses. [Online] Available from: http://www.fedoa.unina.it/9074/1/ Pariota_Luigi_25.pdf [Accessed: 1st September 2015]
- Pariota, L., Bifulco, G. N., Brackstone, M. (2015) A Linear Dynamic Model for Driving Behavior in Car Following. *Transportation Science*. DOI: 10.1287/trsc.2015.0622
- Parker, M. T. (1996) The effect of heavy goods vehicles and following behavior on capacity at motorway road work sites. *Traffic Engineering and Control.* 37(9), pp. 524-531.
- Piao, J., McDonald, M. (2003) *Low speed car following behaviour from floating vehicle data*. pp. 462-467.
- Tibshirani, R., Walther, G., Hastie, T. (2001) Estimating the number of clusters in a data set via the gap statistic. *Journal of Royal Statistic Society B*. 63(2) pp. 411-423. DOI: 10.1111/1467-9868.00293

- Todosoiev, E. P. (1963) *The Action Point Model of the Driver Vehicle System, Report No. 202A-3.* The Ohio State University, Engineering Experiment Station, Columbus, Ohio.
- U.S. Department of Transportation (2015) *NGSIM*. [Online]. Available from: http://ops.fhwa.dot.gov/trafficanalysistools/ngsim.htm [Accessed: 1st September 2015]
- Wiedemann, R. (1974) *Simulation des Strassenverkehrsflusses*. (Simulation of Traffic Flow.) Tech. Rep. Institut für Verkehrswesen, Karlsruhe, Universität Karlsruhe Heft 8 der Schriftenreihe des IfV. (in German)
- Wilson, R. E. (2008) Mechanisms for spatiotemporal pattern formation in highway traffic models. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences.* 366 (1872). pp. 2017-2032. DOI: 10.1098/rsta.2008.0018