Australasian Transport Research Forum 2017 Proceedings
27 – 29 November 2017, Auckland, New Zealand
Publication website: http://www.atrf.info

Car-Following Behaviour: Statistical Analysis of Differences Between Drivers

Frank Lehmann, University of Auckland
Prakash Ranjitkar, University of Auckland
Partha S. Roop, University of Auckland

Email for correspondence: fleh889@aucklanduni.ac.nz

This is an abridged version of the paper presented at the conference. The full version is being submitted elsewhere. Details on the full paper can be obtained from the author(s).

1 Abstract

Microscopic traffic models replicate the behaviour of individual drivers and their interactions with each other. As such, the accuracy of simulation predictions is closely tied to a realistic reproduction of actual human driving behaviour. Despite being a well-researched subject, there is not much literature about the differences between drivers; in most simulations dynamics emerge from interactions of homogeneous driver populations or from a few different behavioural types at best. In this paper we investigate and quantify the car-following (CF) differences between drivers. The underlying dataset contains over 700 min of high-precision CF data and has been used in several research papers. Methodologically we focus on analysing feature distributions, visualising correlations of driver acceleration to space gap, headway, relative speed as well as lead vehicle acceleration and calculating the driver reaction times by shifting the correlating fields. The article uses lots of innovative visualisation methods to efficiently summarise the obtained results.

keywords: traffic modelling, BML, Cellular Automata, Timed Automata

2 Introduction

2.1 Motivation and contribution

With world-wide population growth, accelerating urbanisation processes and more people having access to motorised individual transport, understanding jamming behaviour and simulating vehicle movements has become an important branch of civil engineering. Traffic simulations are classified based on their observation and abstraction levels: models involving individually represented vehicles are termed *microscopic*, aggregated distributions of speed, space gap or headways are termed *mesoscopic* and if the individual nature of vehicles dissolves into a stream with fluid-like qualities the term *macroscopic* is used. Due to advancements in computational power, microscopic models have become more accessible and are arguably the largest group. Their popularity is fostered by the wide range of phenomena they exhibit and the possibility to incorporate empirical observations.

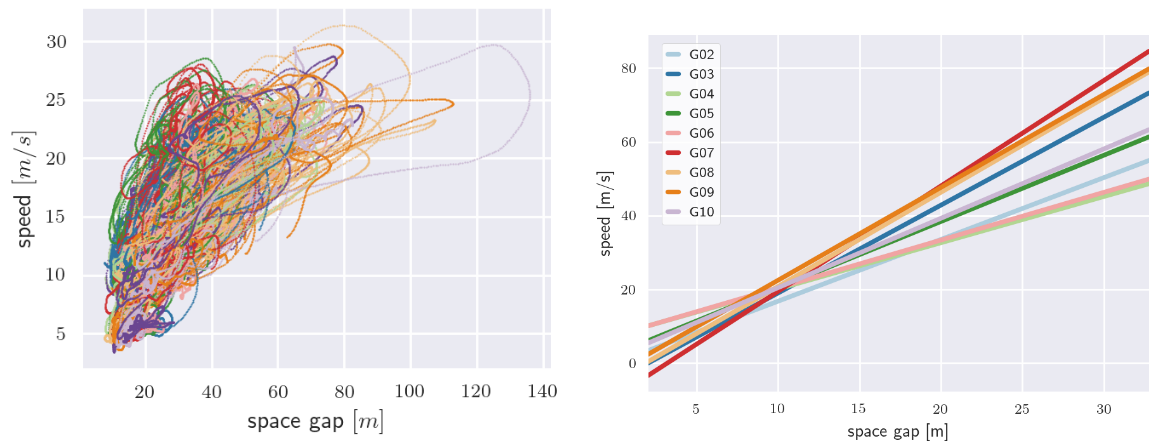
In this publication we carry out a systems analysis in which we statistically investigate driver behaviour from a microscopic perspective. It is based around the the idea that drivers react at diverge times with varying amounts to different stimuli. Our contributions are two-fold: using Spearman, Kendall-Tau and Pearson correlation coefficients, we identify the most relevant stimuli for drivers when following other vehicles. In the course of this statistical analysis we also show which acceleration values drivers select and how high the latency between input stimuli and acceleration/deceleration actions is. Starting with distributions of trajectorial features, we built correlation heatmaps and use time shifting to visualise the spread of reaction times and stimuli responses to car-following input. The central idea of driver differences shows in all figures and runs like a common thread through is work. While similar analyses were conducted individually before, this is the first attempt to capture the whole picture of driver variance. Furthermore, the foundation for this analysis is a high-precision dataset recorded in Tomakomai (Japan) and was used in several publications [16, 15, 17, 18].

2.2 Abbreviations, terminology and related work

CF models treat traffic participants as separate entities whose motions can be described by a vector of state variables. Subscripted with the letter n, the nth vehicle is identified and a vector with variables for speed (v_n), acceleration (a_n) and spatial location (x_n) is defined. A microscopic traffic model is then a set of rules or equations evolving v_n, x_n and a_n over time in dependence on the states of other vehicles. The absence of additional indices already implies that movement are only one-dimensional. With this restriction, there is no difference between *speed* and *velocity*; the terms are used interchangeably unless noted otherwise. Some other symbols used throughout this publication are the front-bumper to front-bumper distance $\Delta\ x_n$, the space a car occupies in a jam (vehicle length plus some additional space) l_{n+1}, the space gap g defined as $g_n := \Delta\ x_n - l_{n+1}$ and the time headway h calculated as $h_n = {}^{\Delta x_n}/_{v_n}$. Relative speed is the difference between two vehicles: $\Delta v := v_{n-1} - v_n$.

Figure 1: Speed vs. space gap and a liner regression.

(a) This plot contains all speed-space gap pairs in the dataset. **(b) Linear regression of the speed-space gap driving relationship grouped by drivers.**



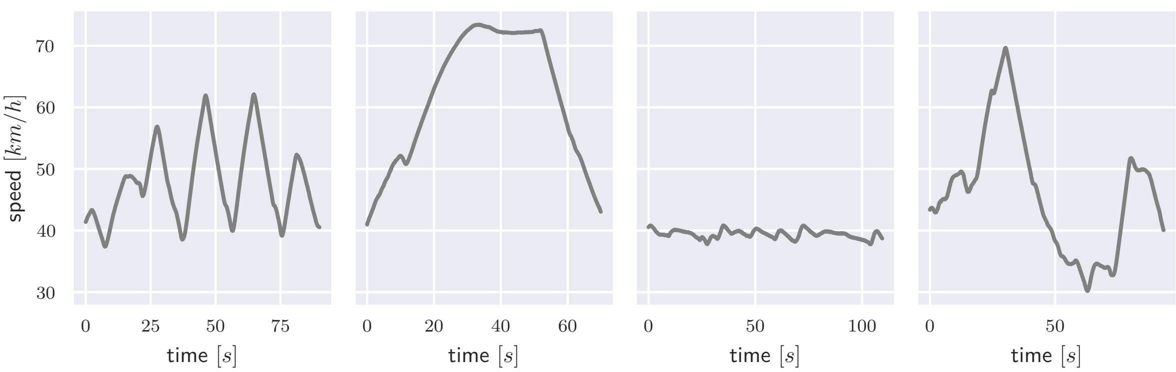
This dependence on other drivers is called *driving relation* and forms the core of all car-following models: it defines a desired velocity as function of distance to the leading vehicle (microscopic) or the vehicle density ρ (macroscopic). The underlying assumption is that drivers tend to keep a individually different but constant time headway, making space gap proportional to speed ($v \propto g$). In order for vehicle n not to crash into the leading vehicle $n+1$, $g_n > vt_{\text{rct}}$, which is equivalent to $h_n > t_{\text{rct}}$. In both relations, t_{rct} denotes the total reaction time and includes mechanical delay as well as anticipation. When g exceeds a certain threshold ($\approx 70\,\text{m}$), the lead vehicle's actions are deemed to become irrelevant for the following vehicle. At this point the driving relation disbonds and drivers adjust their speed freely.

To verify the linear $v-g$ relationship, we plotted all $v-g$ combinations in the dataset against each other in Fig. 1a. Fig. 1b shows the ouput of linear regressions per driver. For space reasons, residuals are not presented, individual drivers can nevertheless be identified easily. Despite the varying slopes, the regression lines roughly meet at $11m < g < 13m$ and $v \approx 20\,\mathrm{m\,s^{-1}}$, a $v-g$ combination

A general description and analysis of traffic flows and individual drivers has been approached from other angles before. In [10] Kerner and Rehborn describe traffic flow's experimental properties of complexity while Lebacque and Lesort studied the order of macroscopic traffic flow models [11]. Chowdhury et al. and Helbing analysed the statistical physics of vehicular traffic as self-driven many-particle systems [3, 9], Nagel et al. investigated road traffic from the perspective of jamming [13]. Daganzo macroscopically described homogeneous, multi-lane motorway traffic as a result of cumulated driver curves/trajectories with different profiles [4]. There have been many attempts at understanding and mimicking human driver behaviour by building psychological models (for an overview, see [2, 7, 12, 20] and references therein). Whilst the majority of these publications are concerned with macroscopic effects and jamming transitions (macroscopic fundamental diagram, perturbations, critical densities, etc.). Our goals are similar those of Todosiev who instructed drivers in a simulator to follow a lead vehicle at what the participants considered to be a minimum safe distance [22].

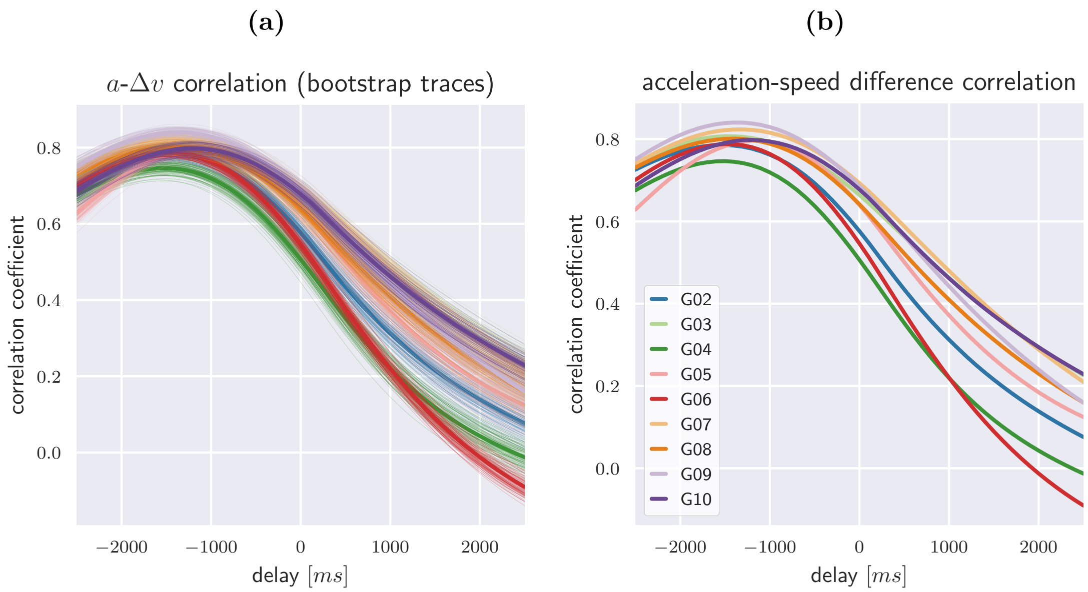
Collecting trajectories under real-world conditions is the preferred way obtain exploitable data about driver behaviour. Unfortunately, the only non-invasive technique to aggregate individual vehicle data are video recordings. From the positions of each road participant, speed and acceleration can be calculated. Because speed/velocity is the first-order antiderivative (primitive function) displacement, and the second-order antiderivative of acceleration. As a result, the noise from recorded location data grows exponentially when subsequently calculating speed and acceleration from position data. The publicly available NGSIM datasets ([23]) are therefore not suited for this analysis. Furthermore, outside controlled experimental conditions drivers might be distracted or base decisions not only on car-following stimuli but on other factors such as opposite and adjacent traffic, curves, or road conditions like potholes, slope or narrow lanes.

Figure 2: Selection of speed patterns as executed by driver $G01$.



To avoid these pitfalls, this CF system analysis is based on experimental data, recorded on a test track located near Tomakomai (Hokkaido province, Japan). Human CF behaviour was observed on two 1200 m straight sections without opposing traffic; this is the optimal case and drivers were probably more attentive here than under real-world traffic conditions. Ten passenger cars were tracked with RTK GPS receivers while following each other. Vehicular speeds were recorded with a Doppler effect-based tool and an accuracy of $0.0\overline{4}\ \mathrm{m\,s^{-1}}$ ($0.16\,\mathrm{km\,h^{-1}}$). All other trajectory features (g, h, x, a) were calculated from the speed measurements. Drivers whose CF behaviour we are interested in were aged from 20 years to 30 years; their years of driv-

Figure 3: Bootstrapped traces for the time-variant $a - \Delta v$ relationship for each of the nine CF vehicles.



ing experience are shown in table 1. To reconstruct real-world driving conditions, driver $D01$ imitated several patterns of acceleration, constant velocity and deceleration (Fig. 2).

Table 1: Summary of drivers participating in the experiment.

driver	driving experience [years]	driver's age	CF time [minutes]
G01	40	60	N/A
G02	10	30	80.8
G03	3	21	80.9
G04	3	25	80.9
G05	5	24	80.8
G06	7	25	80.7
G07	3	23	80.9
G08	2	22	80.7
G09	3	24	69.6
G10	10	30	69.4

3 Methodology And Results

Methodologically, the exploratory data analysis of the >700 min car-following data is based on the visualisation of frequency and distribution for a, g, Δv and other features. Spectra and peculiarities of aggregated and individual driver data are presented using violin plots, level-value plots and box-whisker plots, utilising their respective advantages. Standard deviations on a per-driver basis are calculated based on normalised inputs for the aforementioned trajectory features and their differences classified.

The second major part of the paper deals with other input stimuli being decision factors for a driver to accelerate, decelerate or maintain velocity. This is done by calculating the correlation between a and said factors. Using bivariate Kernel Density Estimations, differences between drivers are elaborated and bounds for selected drivers shown. By shifting the observed data against each other, recalculating correlation coefficients, the time where correlations are highest can be identified. As exemplarily shown in Fig. 3a, said peak is approximately located at $-1\,\mathrm{s}$ for Δv. This trajectory feature shows the highest correlation with driver acceleration behaviour. The importance of Δv and the magnitude of the driver delay confirm previous research. What is new, is the quantification of differences between drivers leading to this aggregated numbers.

4 Conclusion

In this paper we statistically analysed high-precision car-following trajectories covering more than 700 min of nine drivers. Distributions of trajectory features were visualised using violin plots and unravelled that drivers may have a preferred headway they like to maintain but the range differs immensely from driver to driver. Correlations between vehicle acceleration and $a_{n+1}, \Delta v, h$ and g were investigated with scatter plots and bivariate KDE. We quantified how strongly drivers react to car-following input stimuli and found some behavioural commonalities but also individual differences. In the last section we investigated the role of delay by shifting data series against each other. Employing Spearman, Pearson and Kendall Tau correlation coefficients, we found the average total reaction time for drivers to be $\approx 1\,\mathrm{s}$ and showed that Δv is the strongest stimulus when making acceleration decisions. Both results are in agreement with previous research.

Most current traffic simulations generate their macroscopic dynamics from the interactions of homogeneous driver populations. The results for similarly experienced and aged drivers on a test track presented here, indicate that agent-based simulations with higher degrees of variability would replicate reality more accurately. Since the reactions of individual drivers were highly inconsistent consistent, coarse microsimulations like Traffic Cellular Automata do not necessarily yield worse results than more complex simulations. Further research is needed to understand acceleration behaviour in more detail and answer how long drivers accelerate with a constant force, how many discernible acceleration steps they employ and in which combinations drivers switch between these levels.

References

[1] Botev, Z. I., Grotowski, J. F., and Kroese, D. P. (2010). Kernel density estimation via diffusion. *The Annals of Statistics*, 38(5):2916–2957.

[2] Brackstone, M., Sultan, B., and McDonald, M. (2002). Motorway driver behaviour: Studies on car following. *Transportation Research Part F: Traffic Psychology and Behaviour*, 5(1):31–46.

[3] Chowdhury, D., Santen, L., and Schadschneider, A. (2000). Statistical physics of vehicular traffic and some related systems. *Physics Reports*, 329(4–6):199–329.

[4] Daganzo, C. F. (2002). A behavioral theory of multi-lane traffic flow. Part I: Long homogeneous freeway sections. *Transportation Research Part B: Methodological*, 36(2):131–158.

[5] Diwekar, U. and David, A. (2015). Probability Density Functions and Kernel Density Estimation. In *BONUS Algorithm for Large Scale Stochastic Nonlinear Programming Problems*, SpringerBriefs in Optimization, pages 27–34. Springer New York.

[6] Efron, B. (2003). Second Thoughts on the Bootstrap. *Statistical Science*, 18(2):135–140.

[7] Fuller, R. (2005). Towards a general theory of driver behaviour. *Accident Analysis & Prevention*, 37(3):461–472.

[8] Härdle, W. (1991). Kernel Density Estimation. In *Smoothing Techniques*, Springer Series in Statistics, pages 43–84. Springer New York.

[9] Helbing, D. (2001). Traffic and related self-driven many-particle systems. *Reviews of Modern Physics*, 73(4):1067–1141.

[10] Kerner, B. S. and Rehborn, H. (1996). Experimental properties of complexity in traffic flow. *Physical Review E*, 53(5):R4275–R4278.

[11] Lebacque, J. P. and Lesort, J. B. (1999). Macroscopic Traffic Flow Models: A Question of Order. In *Proceedings of the 14th International Symposium on Transportation and Traffic TheoryTransportation Research Institute*, Jerusalem.

[12] Macadam, C. C. (2003). Understanding and Modeling the Human Driver. *Vehicle System Dynamics*, 40(1-3):101–134.

[13] Nagel, K., Wagner, P., and Woesler, R. (2003). Still Flowing: Approaches to Traffic Flow and Traffic Jam Modeling. *Operations Research*, 51(5):681–710.

[14] Parzen, E. (1962). On Estimation of a Probability Density Function and Mode. *The Annals of Mathematical Statistics*, 33(3):1065–1076.

[15] Ranjitkar, P., Nakatsuji, T., and Asano, M. (2004). Performance Evaluation of Microscopic Traffic Flow Models with Test Track Data. *Transportation Research Record: Journal of the Transportation Research Board*, 1876:90–100.

[16] Ranjitkar, P., Nakatsuji, T., Azuta, Y., and Gurusinghe, G. (2003). Stability Analysis Based on Instantaneous Driving Behavior Using Car-Following Data. *Transportation Research Record: Journal of the Transportation Research Board*, 1852:140–151.

[17] Ranjitkar, P., Nakatsuji, T., Gurusinghe, G. S., and Azuta, Y. (2002). Car-Following Experiments Using RTK GPS and Stability Characteristics of Followers in Platoon. In *Applications of Advanced Technologies in Transportation (2002)*, Applications of advanced technology in transportation, pages 608–615. American Society of Civil Engineers.

[18] Ranjitkar, P., Nakatsuji, T., and Kawamua, A. (2005). Car-Following Models: An Experiment Based Benchmarking. *Journal of the Eastern Asia Society for Transportation Studies*, 6:1582–1596.

[19] Rosenblatt, M. (1956). Remarks on Some Nonparametric Estimates of a Density Function. *The Annals of Mathematical Statistics*, 27(3):832–837.

[20] Saifuzzaman, M. and Zheng, Z. (2014). Incorporating human-factors in car-following models: A review of recent developments and research needs. *Transportation Research Part C: Emerging Technologies*, 48:379–403.

[21] Shevlyakov, G. L. and Oja, H. (2016). *Robust Correlation: Theory and Applications*. Wiley Series in Probability and Statistics. John Wiley & Sons, Ltd, Chichester, UK.

[22] Todosiev, E. P. (1963). *The Action Point Model of the Driver-Vehicle System*. PhD thesis, The Ohio State University.

[23] US Department of Transportation (2005). NGSIM - Next Generation Simulation. http://ops.fhwa.dot.gov/trafficanalysistools/ngsim.htm.

[24] Waskom, M., Botvinnik, O., Drewokane, Hobson, P., null, D., Halchenko, Y., Lukauskas, S., Cole, J. B., Warmenhoven, J., Ruiter, J. D., Hoyer, S., Vanderplas, J., Villalba, S., Kunter, G., Quintero, E., Martin, M., Miles, A., Meyer, K., Augspurger, T., Yarkoni, T., Bachant, P., Williams, M., Evans, C., Fitzgerald, C., null, B., Wehner, D., Hitz, G., Ziegler, E., Qalieh, A., and Lee, A. (2016). Seaborn: V0.7.1 (June 2016).