



UNIVERSITY OF LEEDS

This is a repository copy of *From driving simulator experiments to field traffic application: Improving the transferability of car-following models.*

White Rose Research Online URL for this paper:  
<http://eprints.whiterose.ac.uk/164282/>

Version: Accepted Version

---

**Article:**

Paschalidis, E, Choudhury, CF [orcid.org/0000-0002-8886-8976](https://orcid.org/0000-0002-8886-8976) and Hess, S [orcid.org/0000-0002-3650-2518](https://orcid.org/0000-0002-3650-2518) (2021) From driving simulator experiments to field traffic application: Improving the transferability of car-following models. *Journal of Transportation Engineering, Part A: Systems*, 147 (1). ISSN 0733-947X

<https://doi.org/10.1061/JTEPBS.0000468>

---

© 2020 American Society of Civil Engineers. This is an author produced version of a journal article published in *Journal of Transportation Engineering, Part A: Systems*. Uploaded in accordance with the publisher's self-archiving policy. This material may be downloaded for personal use only. Any other use requires prior permission of the American Society of Civil Engineers. This material may be found at <https://doi.org/10.1061/JTEPBS.0000468>.

**Reuse**

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>

# From driving simulator experiments to field traffic application: Improving the transferability of car-following models

Evangelos Paschalidis<sup>1</sup>, Charisma F. Choudhury<sup>2</sup>, Stephane Hess<sup>3</sup>

<sup>1</sup>Research Fellow, Choice Modelling Centre, Institute for Transport Studies, University of Leeds, 36-40 University Rd, Leeds LS2 9JT, UK, Email: E.Paschalidis@leeds.ac.uk

<sup>2</sup>Associate Professor, Choice Modelling Centre, Institute for Transport Studies, University of Leeds, 36-40 University Rd, Leeds LS2 9JT, UK, Email: C.F.Choudhury@leeds.ac.uk

<sup>3</sup>Professor, Choice Modelling Centre, Institute for Transport Studies, University of Leeds, 36-40 University Rd, Leeds LS2 9JT, UK, Email: S.Hess@leeds.ac.uk

## Abstract

Over the last few decades, there have been two main streams of data used for driving behaviour research: trajectory data collected from the field (using video recordings, GPS, etc.) and experimental data from driving simulators (where the behaviours of the drivers are recorded in controlled laboratory conditions). Previous research has shown that the parameters of car-following models developed using simulator data are not directly transferable to the field. In this research, we investigate the differences in detail and compare alternative methods to overcome the problem. Two types of approaches are tested in this regard: 1) econometric approaches for increasing model transferability: Bayesian updating and Combined Transfer Estimation, 2) joint estimation using both data sources simultaneously. Car-following models based on ‘stimulus-response’ framework are developed in this regard, using experimental data collected at the University of Leeds Driving Simulator (UoLDS) and detailed trajectory data collected at the Interstate 80 (I-80), CA, USA and UK Motorway 1 (M1). Estimation results of the initial models show that car-following models using driving simulator data are closer to the UK (M1) data than the I-80, but not directly transferable. Performances of the proposed approaches for improving transferability are evaluated using t-tests for individual parameter equivalence and Transferability Test Statistic (TTS). The results indicate that the transferability can be improved after parameter updating and Combined Transfer Estimation is found to outperform the other approaches. The findings of this study will enable more effective usage of driving simulator data for the estimation of mainstream mathematical models of driving behaviour while the techniques used can be applied to other types of econometric models.

Author keywords: car-following model, driving simulator, trajectory data, transferability, joint estimation, Bayesian updating, Combined Transfer Estimation

## 1. Introduction

Driving decisions and consequently vehicle interactions, are crucial factors for evaluating traffic performance and driving safety. Driving behaviour models, which are mathematical approximations of drivers' decisions regarding longitudinal and lateral movements (e.g. acceleration-deceleration, lane-changing, etc.), have been widely studied in the past few decades (Saifuzzaman & Zheng, 2014; Toledo, 2007; Zheng, 2014). Of particular interest are car-following models, which aim to replicate the accelerations and decelerations of the driver while closely following a lead vehicle in the front. Such models are crucial for increasing the realism of the microsimulation tools as well as safety and emission analyses.

Car-following (and microscopic driving behaviour models in general) are typically developed using two types of data, (a) driving simulator (where drivers drive an instrumented vehicle in a simulated roadway) and (b) road traffic data. Driving simulator data are collected following standardised procedures and are more controllable and reproducible compared to actual road traffic. Furthermore, driving simulators allow researchers to manipulate the surrounding conditions (e.g. geometric layout of the road, number and type of vehicles, level of aggressiveness of other road users, etc.) as well as driver specific conditions (e.g. level of distraction and fatigue). They also allow analysts to run multiple hypothetical scenarios for the same driver and observe driving behaviour for longer time horizons.

The advantages of driving simulators provide the researchers with the opportunity to shift from the development of models completely based on a Newtonian laws of motion approach (i.e. considering only speed, headway etc.) and incorporate further aspects of driving behaviour that can be later applied in microscopic simulation tools. For instance, Saifuzzaman and Zheng (2014) highlighted in their literature paper the need to incorporate human factors in existing car-following model specifications. Also, researchers in psychology (Brackstone & McDonald, 2003; Hancock, 1999; Van Winsum, 1999) have questioned the existing engineering car-following modelling approaches that omit the effects of drivers' characteristics. Along the same direction, (Laagland, 2005) suggested a series of approaches to incorporate drivers' aggression in microscopic driving behaviour models. Apart from traditional microscopic simulation tools, accurate driving behaviour models can be also applied and improve the performance of integrated simulation and hardware-in-the-loop systems (Buse et al., 2018; Chen et al., 2019; Fouladinejad et al., 2011; Y. Zhao et al., 2016; Y. Zhao et al., 2012) where a variety of simulators as driving, traffic and network are combined. It is essential for the human participant of these systems to face realistic and expected behaviour from the surrounding traffic.

Driving simulators offer a research environment where many of the aspects of driving behaviour related to human factors can be investigated, recorded and potentially used in modelling approaches. However, there is scepticism regarding simulator fidelity (physical and behavioural) and how well drivers' behaviour in a simulator matches with their behaviour on real roads (Lee, 2003). On the other hand, traffic data collected from the field best represents true driving behaviour, but have several limitations: short observation time, measurement errors, complex confounding of influencing factors, less control on the external factors and absence of driver characteristics in particular. It may be noted that besides these two sources, naturalistic driving data collected using instrumented vehicles (e.g. UDRIVE, SHRP2 etc.) have also been used in research, but

given the very high costs involved, the availability of these data is still limited. Moreover, like driving simulator data, naturalistic data are likely to be prone to behavioural incongruence; and similar to real road traffic data, the external variables are often not fully controllable, and it is not possible to test the effects of hypothetical scenarios.

Several studies have attempted to investigate the validity of driving simulators concerning drivers' behaviour. Driving simulators' behavioural validity is usually approached in terms of absolute (when the patterns and the magnitude of values are similar to real driving) or relative validity (when the patterns are similar but the magnitudes differ). Godley et al. (2002) investigated behavioural validity in terms of speed. Their research included two types of driving tasks (instrumented vehicle and driving simulator). While their results showed a similar pattern of deceleration in both environments, they noted that drivers adopted faster speed in naturalistic driving conditions and only relative validity prevails. In the same direction, Yan et al. (2008) developed a scenario based on a real signalised intersection and studied simulator validity in terms of speeding and surrogate safety measures. The results showed absolute validity regarding speeding, however, participants adopted riskier behaviours in the driving simulator, thus the safety measures had only relative validity. Bella et al. (2007) reproduced a real two-lane road section composed of 11 parts and tested validity in speed. This study confirmed relative but also absolute validity for most of the examined cases. Risto and Martens (2014) compared the differences in headway choice between an instrumented vehicle and driving simulator without finding significant deviations. Finally, McGehee et al. (2000) compared drivers' reaction times in real and simulated environment and found statistical equivalence between the two cases. In more research studies, Branzi et al. (2017) and Hussain et al. (2019) confirmed relative validity of simulator data while the former study also found evidence of absolute validity in some of the scenarios. With respect to some additional indicators of driving behaviour, Yun et al. (2017) found absolute validity, in terms of lane-changing behaviour while Karimi et al. (2020) reported significant proximity of gap acceptance behaviour during passing manoeuvres between driving simulator and field behaviour.

The development of driving behaviour models based on simulator data has already been reported in literature (Farah et al., 2009; Hou et al., 2014). However, since only relative validity has been established, it remains questionable whether this type of data is suitable for direct use in microsimulation tools for traffic flow and policy analysis. Recent research has shown that the parameters of car-following models developed using simulator data are not directly transferable to the field, although the models as a whole are transferable (Papadimitriou & Choudhury, 2017). However, the Papadimitriou and Choudhury (2017) study acknowledges a major limitation - the model framework used in evaluating transferability ignores reaction time and driver heterogeneity – which have been identified as a crucial factor affecting car-following behaviour (K. I. Ahmed, 1999; Koutsopoulos & Farah, 2012; Toledo, 2007; Van Hinsbergen et al., 2015). Further, the paper does not provide any guidance on how to close the gap between the models developed using the simulator and the real road data.

In this research, we aim to address the research gaps in the previous studies by investigating alternative methods to improve the transferability of car-following models. A better understanding of the differences between the two sources of driving behaviour data (video trajectories and driving simulator) could allow for the estimation

of car-following models from driving simulator data adjusted by real traffic data. This correction could potentially increase the behavioural realism of these models, assuming that the latter data represents the ground truth with respect to drivers' behaviour. At the same time, driving simulator data allows for the implementation of extreme scenarios while information regarding drivers' attributes can also be available as e.g. in Paschalidis et al. (2019) and can be incorporated in the model specifications. In this paper, advanced model structures that incorporate the reaction time (and acknowledges the associated heterogeneity) are used in this regard to address the limitations of the previous study (Papadimitriou & Choudhury, 2017). Experimental data collected at the University of Leeds Driving Simulator (UoLDS) and detailed trajectory data collected at the Interstate 80 (I-80), CA, USA and UK Motorway 1 (M1) are used for this purpose. Based on a review of the literature, two main approaches are tested

- Econometric approaches for improving model transferability,
- Joint Estimation using both data sources simultaneously.

The econometric approaches for improving transferability refer to model parameter updating techniques. These techniques aim to reduce the differences between two contexts. Joint estimation takes models one step further by allowing the specification of scale parameters that account for the differences between two contexts and at the same time consider the effects of variables that are known only for one of the contexts. The key contribution of the paper is the systematic framework: to test the initial transferability of the models estimated using the two kinds of data, implementing the updating techniques, and the subsequent re-evaluation of the different updating techniques. The framework, though demonstrated for specific datasets, can be used in other datasets and other types of driving behaviour models – even though the exact findings may be context-specific.

The remainder of the paper is organised as follows: the next section describes the methodological background. This is followed by the case study description and some preliminary analysis of data. Then, the results of the model estimation are presented which are followed by the transferability and joint estimation results. The paper concludes with a discussion section.

## 2. Background of the analysis

### 2.1 Car-following model

#### Basic structure

The model structure is based on the stimulus-response GM car-following model (Gazis et al., 1961). In the original GM model, acceleration choices for a vehicle are a function of its speed, space headway and relative speed with the lead vehicle. The original specification is as follows

$$\alpha_n(t) = \alpha \frac{V_n(t)^\beta}{\Delta X_n(t)^\gamma} \Delta V_n(t - \tau_n) \quad (1)$$

where  $\Delta X_n$  is the space headway at time  $t$ ,  $V_n$  is the following vehicle speed,  $\Delta V_n$  is the relative speed between the following and the lead vehicle and  $\tau_n$  is the reaction time. Finally,  $\alpha$ ,  $\beta$  and  $\gamma$  are constants.

Based on the GM model, several extensions have been suggested. Herman and Rothery (1965) have been the first to highlight that passenger cars have different acceleration and deceleration capacity. In order to address this shortcoming in the GM model, (K. I. Ahmed, 1999) introduced acceleration-deceleration asymmetry within a stimulus-response framework as presented in Equation 2

$$a_n^{cf,g}(t) = s[X_n^{cf,g}(t-\tau_n)] \times f[\Delta V_n(t-\tau_n)] + \varepsilon_n^{cf,g}(t) \quad (2)$$

where  $s[.]$  represents sensitivity, as a vector of explanatory variables and  $f[.]$  represents the stimulus, given as the relative speed. Also,  $\varepsilon^{cf,g}$  is a normally distributed disturbance term while  $g$  represents the car-following regime (acceleration or deceleration). In the present study, an adaptation of the GM model is applied where the sensitivity and stimulus parts are represented by Equations 3 and 4 respectively

$$s[X_n^{cf,g}(t-\tau_n)] = \alpha^g \frac{1}{\Delta T_n(t)^{\gamma^g}} \quad (3)$$

$$f[\Delta V_n(t-\tau_n)] = \Delta V_n(t-\tau_n)^{\lambda^g} \quad (4)$$

where  $\Delta T_n$  is the time headway,  $\Delta V_n$  is the relative speed between the subject and the lead vehicle and  $\tau_n$  is the reaction time. Finally,  $\alpha^g$ ,  $\gamma^g$  and  $\lambda^g$  are parameters to be estimated and  $g$  indicates the type of regime. It is worth highlighting that instead of applying the original GM model specification, the sensitivity part has been modified in order to consider only time headway, as per the recent literature (Papadimitriou & Choudhury, 2017).

#### The reaction time distribution

The current model specification also allows for the incorporation of reaction time. Following examples in the existing literature (K. I. Ahmed, 1999; Kusuma, 2015), reaction time is assumed to follow a log-normal truncated distribution as presented in Equation 5

$$f(\tau_n) = \begin{cases} \frac{\frac{1}{\tau_n \sigma_\tau} \phi\left(\frac{\ln(\tau_n) - \mu_\tau}{\sigma_\tau}\right)}{\Phi\left(\frac{\ln(\tau^{\max}) - \mu_\tau}{\sigma_\tau}\right) - \Phi\left(\frac{\ln(\tau^{\min}) - \mu_\tau}{\sigma_\tau}\right)} & \text{if } \tau^{\min} < \tau_n \leq \tau^{\max} \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

where  $\phi(.)$  is the standard normal distribution density function,  $\Phi(.)$  is the cumulative normal distribution,  $\tau_n$  is the reaction time of driver  $n$ ,  $\mu_\tau$  is the mean of the distribution of  $\ln(\tau_n)$ ,  $\sigma_\tau$  is the standard deviation and  $\tau^{\max}$ ,  $\tau^{\min}$  are the bounds of truncation. Truncation is required since reaction time is finite. The bounds are set deterministically while the mean and the standard deviation are estimated simultaneously with the rest

of the model parameters. The bounds of reaction time have been set between 0 and 4 seconds (K. I. Ahmed, 1999; Kusuma, 2015).

### Likelihood Function

The assumption of the car-following model is that a driver accelerates if the relative speed is positive and decelerates if it negative. Given this, the distribution of acceleration decisions, conditional on reaction time  $\tau$ , is presented as follows

$$f(a_n^{cf,g}(t)|\tau_n) = f(a_n^{cf,acc}(t)|\tau_n)^{\delta[\Delta V_n(t-\tau_n)]} f(a_n^{cf,dec}(t)|\tau_n)^{(1-\delta[\Delta V_n(t-\tau_n)])} \quad (6)$$

where

$$\delta[\Delta V_n(t-\tau_n)] = \begin{cases} 1 & \text{if } \Delta V_n(t-\tau_n) \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Assuming that the disturbance terms are normally distributed, the acceleration decisions can be expressed as follows

$$f(a_n^{cf,g}(t)|\tau_n) = \frac{1}{\sigma_{e^{cf,g}}} \varphi\left(\frac{a_n^{cf,g}(t) - s[X_n^{cf,g}(t-\tau_n)] \times f[\Delta V_n(t-\tau_n)]}{\sigma_{e^{cf,g}}}\right) \quad (7)$$

where  $g \in (\text{acc}, \text{dec})$ .

In the current specification, the acceleration observations of each driver  $n$  are assumed to be independent while the correlation among the decisions of the same driver (i.e. inter-respondent heterogeneity in driving behaviour) is captured through the reaction time distribution. Thus, the conditional joint density of the acceleration sequential observations, of a driver  $n$ , is the product of the conditional densities of the acceleration decisions is expressed as follows

$$f(a_n(1), a_n(2), \dots, a_n(T_n)|\tau_n) = \prod_{t=1}^{T_n} f(a_n(t)|\tau_n) \quad (8)$$

The unconditional form of the distribution above is expressed in Equation 9

$$f(a_n(1), a_n(2), \dots, a_n(T_n)) = \int_{\tau_{\min}}^{\tau_{\max}} f(a_n(1), a_n(2), \dots, a_n(T_n)|\tau_n) f(\tau_n) d\tau \quad (9)$$

At the final step, the model is estimated by maximizing the log-likelihood function of the acceleration observations as expressed in Equation 10

$$LL = \sum_{n=1}^N \ln[f(a_n(1), a_n(2), \dots, a_n(T_n))] \quad (10)$$

The log-likelihood function has been maximised using the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm implemented in the software R (Team, 2013).

## 2.2 Evaluating Model Performance and Transferability

The basic concept of transferability refers to the transfer of a model estimated in one context to a different one. The main motivation to consider transferability (Rossi et al., 2013) is

- to reduce the efforts in model development (using the same structure of the model previously identified),
- to reduce or eliminate the need for a large data collection in the application context.

In literature, there are limited studies of transferability in the domain of driving behaviour modelling; Papadimitriou and Choudhury (2017) investigated transferability between driving simulator and field data using simple model specifications while in a recent MSc Thesis, Chang (2018) investigated transferability of gap-acceptance models at freeway merging locations for different levels of congestion. However, transferability has been investigated in detail in several other fields of transportation and beyond. The lion's share is dedicated to the investigation of transferability with the application of discrete choice modelling (Atherton & Ben-Akiva, 1976; Ben-Akiva & Bolduc, 1987; Ben-Akiva et al., 1994; Koppelman & Wilmot, 1982), however, other modelling approaches can also be found (Hadayeghi et al., 2006; Wilmot, 1995).

A review of the literature revealed several formal statistical tests of transferability (Sikder et al., 2013) among which the t-tests of individual parameter equivalence and Transferability Test Statistic (TTS) have been found to be most widely used and have been thus selected for this study.

The t-tests of individual parameter equivalence compare parameter estimates of equivalent variables between the two models as e.g. in (Galbraith & Hensher, 1982). The t-stat differences can be expressed as shown in Equation 11

$$t_{\text{diff},k} = \frac{\beta_{\text{est},k} - \beta_{\text{appl},k}}{\sqrt{\left(\frac{\beta_{\text{est},k}}{t_{\text{est},k}}\right)^2 + \left(\frac{\beta_{\text{appl},k}}{t_{\text{appl},k}}\right)^2}} \quad (11)$$

where  $\beta_{\text{est},k}$  is the the parameter estimate of the  $k^{\text{th}}$  parameter of the transferred (simulator data) model and  $t_{\text{est},k}$  is its t-statistic while  $\beta_{\text{appl},k}$  is the the parameter estimate of the  $k^{\text{th}}$  parameter of the application context (video trajectory data) model and  $t_{\text{appl},k}$  is its t-stat. The null hypothesis of parameter equivalence is rejected at the 95% level of confidence if  $|t_{\text{diff},k}| > 1.96$ .

The TTS (Atherton & Ben-Akiva, 1976) assesses whether the null hypothesis of statistical equivalence between the transferred and the application context model is rejected or not. The TTS can be calculated as

$$\text{TTS}_{\text{appl}} = -2 \left[ \text{LL}_{\text{appl}}(\beta_{\text{est}}) - \text{LL}_{\text{appl}}(\beta_{\text{appl}}) \right] \quad (12)$$

where  $LL_{\text{appl}}(\beta_{\text{est}})$  is log-likelihood on the application context data using transferred context parameters and  $LL_{\text{appl}}(\beta_{\text{appl}})$  is the log-likelihood on the application context data using application context parameters, i.e. new estimates. The TTS value follows a chi-squared ( $\chi^2$ ) distribution and the degrees of freedom are equal to the number of model parameters, assuming that the parameters of the transferred model are fixed (Koppelman & Wilmot, 1982). At 95% level of confidence, the models are classified statistically different (i.e. non-transferable) if  $\chi^2 > \chi^2_{\text{critical}}$ .

### 2.3 Methods to Improve Transferability

#### Parameter updating

Findings from previous studies indicate that temporal transferability of a model is improved by updating the model parameters with some information from the application context e.g. Santoso and Tsunokawa (2005). Several updating approaches have been suggested in the literature. Two of the most common techniques are Bayesian updating and Combined Transfer Estimation.

#### Bayesian updating

The Bayesian updating process follows the Bayes theorem in which prior information about the model is combined with a random sample from the application context to obtain updated information that is important in reducing doubt during prediction (Dey & Fricker, 1994). The parameters estimated with the trajectory data can be used as the prior information in this case and the following formula can be used

$$\beta_{\text{upt}} = \left( \frac{\beta_{\text{est}}}{\sigma_{\text{est}}^2} + \frac{\beta_{\text{appl}}}{\sigma_{\text{appl}}^2} \right) \left( \frac{1}{\sigma_{\text{est}}^2} + \frac{1}{\sigma_{\text{appl}}^2} \right)^{-1} \quad (15)$$

where  $\beta_{\text{est}}$  is the parameter of the estimation (driving simulator) context model,  $\sigma_{\text{est}}$  is its standard deviation,  $\beta_{\text{appl}}$  is the parameter of the application (real driving) context model and  $\sigma_{\text{appl}}$  is its standard deviation.

#### Combined Transfer Estimation

The Combined Transfer Estimation method (Ben-Akiva & Bolduc, 1987) can be considered as an extension of Bayesian updating. If the transfer bias (the difference between the real values in the parameter vectors of the estimation and the application context) does not exceed a critical point, Combined Transfer Estimation is a more efficient updating technique, compared to Bayesian updating. The updated parameters are estimated as

$$\beta_{\text{upt}} = \left( \frac{\beta_{\text{est}}}{\sigma_{\text{est}}^2 + \alpha\alpha'} + \frac{\beta_{\text{appl}}}{\sigma_{\text{appl}}^2} \right) \left( \frac{1}{\sigma_{\text{est}}^2 + \alpha\alpha'} + \frac{1}{\sigma_{\text{appl}}^2} \right)^{-1} \quad (16)$$

where  $\alpha = \beta_{\text{est}} - \beta_{\text{appl}}$  and  $\alpha' = \beta_{\text{appl}} - \beta_{\text{est}}$ .

In the existing literature, there are examples where the results of Combined Transfer Estimation outperformed Bayesian updating (Karasmaa, 2007; Santoso & Tsunokawa, 2005, 2010).

### Joint Estimation

The joint estimation of models using various data sources has been introduced in the discrete choice modelling field (Ben-Akiva et al., 1994) and mostly refers to the combination of stated-preference (SP) and revealed-reference (RP) data. The motivation for data combination is the estimation of enhanced models that exploit the advantages of the various data sources while at the same time minimise their shortcomings, by allowing variations in their scales. A basic example regarding the application of this approach could be the reduction of hypothetical bias of a SP survey and improvement of the accuracy of parameter estimates, through joint estimation with RP data. The joint estimation process provides estimates of the common parameters but since the variances of the disturbance terms between SP and RP are likely to be different, an additional scale parameter is introduced to capture this variation. For model identification purposes, the scale of RP is normalised to one while only the scale of SP is estimated. Within a car-following context, S. Hoogendoorn and Hoogendoorn (2010) provided a methodological framework for joint estimation of driving simulator and real traffic data and suggested a weighting correction to account for the differences in sample sizes. However, in their estimation, they considered the contribution of driving simulator data as equal and did not investigate any potential behavioural bias deriving from its hypothetical nature.

## **3. Case study**

The datasets used for the model estimation are described in the present section. Initially, the characteristics and attributes of each site are provided followed by a preliminary descriptive analysis.

### **3.1.1 Data**

#### I-80 trajectory data (USA)

The first of the vehicle trajectories data used in the analysis, has been collected at the Interstate 80 (I-80), CA, USA, within the framework of the Next Generation SIMulation (NGSIM) project (Halkias & Colyar, 2006) and have been extensively used in other studies (Aghabayk et al., 2012; Koutsopoulos & Farah, 2012). The observations have taken place on 13 April 2005. The length of the road segment is approximately 500 meters (1650 feet) and comprises of five lanes plus a high occupancy vehicle (HOV) lane (Figure 1-left). The vehicles' trajectories referring to the observations from 4.00 p.m. to 4.15 p.m. have been further processed by Punzo et al. (2011) and Montanino and Punzo (2013). The final dataset includes information regarding the position, speed, acceleration, lane, size and type of each vehicle.

#### Motorway 1 trajectory data (UK)

The M1 (UK) disaggregate vehicle trajectory data was collected and first introduced by Kusuma (2015). The data was collected between Junction 42 and Junction 43 of the M1 motorway network in the United Kingdom from an overpass located 620 m downstream from Junction 42 (near the city of Leeds) in May 2013. Because of the camera angle and features of the trajectory extraction software, only data from the first 320 m were found to be usable. An overview of the road section is presented in Figure 2. The road section consists of five traffic lanes, three main and two auxiliary lanes.

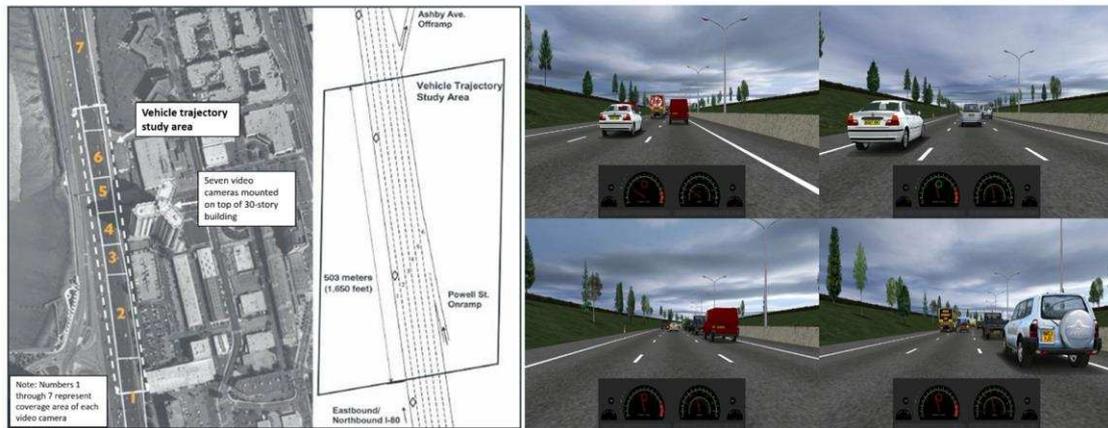


Figure 1: (left) I-80 motorway data collection site, (right) Screenshots of the motorway

### Driving simulator data

The driving simulator data has been collected at the University of Leeds Driving Simulator (UoLDS). The UoLDS is a high fidelity, dynamic simulator (eight degree of freedom motion system), with all driver controls, such as steering wheel and braking pedal, available and fully functional, while there is also a fully operating dashboard. The vehicle is placed in a 4m diameter spherical projection dome. The dome provides fully textured 3-D graphical scene with a horizontal field of view of 250° and 45° vertical. The raw data output consists of observations of 60Hz frequency.

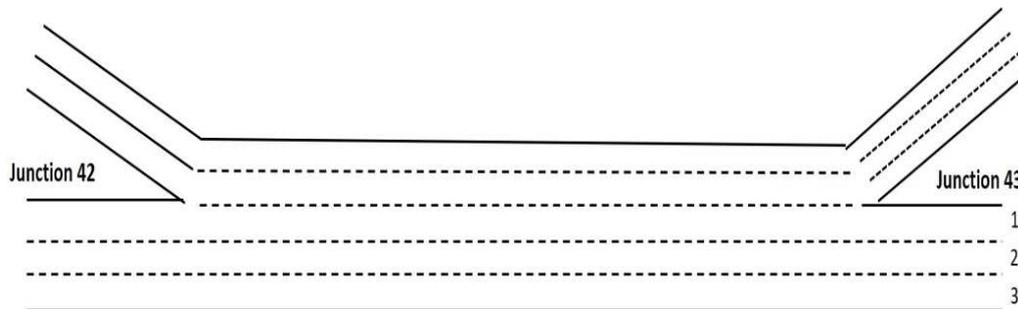


Figure 1: Data collection site at the M1 motorway, UK

The data collection has taken place in the context of the “Next Generation Driving Behaviour Models” project (NG-DBM) that focused on development of driving behaviour models that explicitly account for the effects of observed and unobserved static and dynamic driver characteristics in his/her decisions and the calibration of driving behaviour models combining experimental data collected from the University of Leeds Driving Simulator (UoLDS) and actual traffic data collected using video recordings.

The full data collection process involved around 90 minutes of total driving. Participants have had first a short briefing about the simulator and its operation followed by a practice session of approximately 15 minutes duration to get familiarised with the simulated environment and vehicle dynamics (i.e. motion system). For safety reasons, participants have been accompanied and guided by a researcher in the back seat, during the practice run, as required by the UoLDS protocol. This approach is followed to ensure that participants are familiar with the operation of the vehicle and suffer no ill-effects from simulator exposure, such as virtual environment nausea,

vertigo or visual/vestibular discrepancies. After the practice session, participants started the main driving sessions; an urban and a 3-lane motorway environment, with a short break in between. Both settings were based on UK road environments. In total, 36 drivers (17 females, 19 males) aged from 19 to 57 years old have successfully completed the motorway setting that has been used in the current analysis (Figure 1-right) for the model specification and estimation.

The motorway has been composed of six main sections approximately 6km long each, connected with some shorter road segments specified as intersections. In each of the main road segments, different traffic scenarios have been implemented (e.g. aggressive traffic, slow moving traffic, different levels of time pressure etc.), while the role of intersections has been to provide a smoother transition and also reduce potential residual effects from previous road segments, as no specific events have been planned in these locations. The details of each of the motorway sections is presented in Table 1.

Table 1: The motorway sections in the driving simulator.

Code	Scenario	Time pressure state
M1	No events	Green
M2	Aggressive surrounding traffic	Green
M3	Aggressive surrounding traffic	Amber – Red
M4	Slow traffic	Green – Amber – Red
M5	No events	Amber – Red
M6	No events – Hard braking event	Green – Amber – Red

One of the main objectives of the study has been the investigation of drivers' behaviour under time pressure hence, participants have been deliberately subjected to time pressure. During their briefing session, participants have been instructed that they have had to reach their destination within 35 minutes else, a monetary penalty would be imposed if they are late. An emoji placed on the dashboard of the vehicle (Figure 3) has been used as an indicator of their performance. The emoji could have three different states, namely, green, amber and red. Participants have been instructed that the green state would indicate they have been doing well, in terms of time, while the red would indicate that they have been late. The intermediate amber emoji denoted that they have been marginally fine in terms of time. That is, they would receive a red emoji if they have had further delay in the remaining driving tasks. The introduction of an amber state has been decided to make the shift from green to red emoji more convincing to the participants.



Figure 2: Time pressure emoji

Although participants have been informed that the state of the emoji have been related to their performance, it has been pre-decided in order to induce time pressure in specific road segments. It may be noted that the choice of three different emoji to indicate time

pressure, has been preferred to a conventional countdown timer, since it would be easier to manipulate.

### 3.1.2 Preliminary analysis

All datasets included in the analysis refer to motorway settings however, apart from the different nature of the data (field traffic and driving simulator) there have also been some additional differences that could affect the results of the car-following models. The main differences are outlined below.

- The road environments differ in terms of the surroundings (i.e. landscape) and also infrastructure characteristics, such as the number of lanes.
- Levels of congestion and traffic flow among the field traffic and simulator sites may vary.
- The video datasets include trajectories from several vehicles with different capabilities in terms of acceleration/deceleration while the trajectories of the driving simulator data always refer to the same vehicle.
- The I-80 data has been collected in the USA while the driving simulator data have been collected in the UK hence, cultural differences may exist.
- Vehicle capabilities may vary among the various datasets because of temporal difference in the data collection dates.

Within an effort to reduce the differences among the various data sources, the data has been processed to include sections that are as similar as possible to one another in terms of traffic characteristics. Moreover, data processing has been also applied to consider data that better meet the requirements for the estimation of car-following models. As a first step, relationships regarding the surrounding traffic such as relative speed, acceleration of lead vehicle etc. have been extracted from all datasets. Regarding the I-80 trajectory dataset, only cars that have not attempted a lane-changing manoeuvre during the observation period have been included in the analysis. This approach has not been applied to the M1 (UK) data as the occurrence of lane changes has been frequent and very few observations would have remained. Thus, only the data from the auxiliary lanes has been removed.

With respect to the driving simulator data, a different approach has been followed in order to investigate its similarity to the trajectory data and consider more comparable cases. Each of the six main sections of the simulated motorway has been compared to the other datasets in terms of acceleration, speed, relative speed with the lead vehicle, time and space headway values. The aim has been to select values, that would be later used for the estimation of the car-following model, as close as possible to those observed in the field. After the examination of the descriptive statistics of the aforementioned variables and the visual inspection of their histograms (Figure A.1 - Appendix), the section with the slow-moving surrounding traffic (M4) has been selected, out of the whole motorway. The latter included all levels of time pressure, as shown in Table 1. Given that drivers in real life may be under time pressure, this condition of the simulator scenario is expected to capture some of its effect in the parameter estimates of the car-following model. The histograms of the aforementioned key variables for all three data sets are presented in Figure A.1 of the Appendix.

The considered observation frequency has been 1 observation/sec in all datasets. Also, in order to avoid free-flow observations and following the findings in S. P.

Hoogendoorn (2005), an upper bound of 4s has been applied in the observed time headway; all the values above that threshold, are treated as free-flow and excluded from the analysis. For the final estimation, the I-80 trajectory dataset is composed of 469 individuals and 14,826 observations, the M1 data of 619 individuals and 3,302 observations while the driving simulator dataset 36 individuals and 7,191 observations. Table 2 summarises the descriptive statistics for some main variables across all datasets.

The descriptive statistics indicate that there are differences in the examined variables of the two datasets. These differences are further investigated with an independent samples t-test between the field and driving simulator data. The p-values of the Levene's test are significant for all variables, apart from time headway for the I-80 data, which indicates that the variances of all the variables are different between the video trajectory and the driving simulator datasets. Additionally, the results of the t-test for the equality of means show that the means of all variables are significantly different as well. These findings point out that there are variations in the traffic variables (and thus in traffic conditions) between the video and simulator datasets, which may affect the models' results. An additional remark with respect to congestion levels of the driving simulator data – as potentially reflected by the ranges of speeds and headways – is that it lies between the I-80 and the UK M1 video data. Though these differences impose extra challenge in the transferability of the models, in practical cases, this is very likely to be the reality (i.e. the simulator data being available for a small subset of participants, fixed variations in simulated traffic whereas actual road traffic having larger variability) leading to difficulties in reproducing observed traffic flow patterns in simulator.

#### **4. Estimation results**

The estimation results are summarised in Table 3. After some initial interpretation of the models' parameters, the individual models are further investigated and compared via sensitivity analysis. In the M1 data, the majority of acceleration observations has a negative value (Figure A.1) possibly due to the higher rates of lane-changing. The model estimation hence required constraining the acceleration and deceleration definitions based on the sign of the observed acceleration value as opposed to the sign of the relative speed difference (as proposed by Kusuma 2015).

The estimated reaction time distributions are illustrated in Figure 4. The reaction time distribution from the I-80 data ranges between 0-2s and is mainly centred slightly above 0.5s. On the other hand, the estimated distributions based on the M1 and simulator data cover the whole range of truncation however, the mean of the latter distribution has a higher value with a peak between 1.5s and 2s. This may be an indication of different response patterns for the relative speed stimulus between the field and video contexts. Some further discussion with respect to their differences is presented together with the analysis of the individual models.

##### **4.1.1 Individual Models**

Model 1: Car-following model based on driving simulator data (UK)

The estimated car-following acceleration of Model 1 is shown in Equation 17

Variable	Driving simulator data (1)				I-80 Video data (2)				MI Video data (3)				(1) & (2) t-tests		(1) & (3) t-tests	
													Levene's test for equality of variances	t-test for equality of means	Levene's test for equality of variances	t-test for equality of means
	Min	Mean	Max	sd	Min	Mean	Max	sd	Min	Mean	Max	sd	p-value	p-value	p-value	p-value
Speed (m/s)	7.67	14.71	35.93	4.74	1.52	8.42	26.08	3.58	15.95	26.16	40.77	4.21	0.000	0.000	0.000	0.000
Acceleration (m/s <sup>2</sup> )	-10.04	-0.09	1.90	0.67	-4.73	-0.04	2.90	0.93	-6.73	-1.08	5.88	1.16	0.000	0.000	0.000	0.000
Time headway (s)	0.42	1.82	3.98	0.69	0.58	2.34	4.00	0.68	0.18	1.73	3.99	0.87	0.801	0.000	0.000	0.000
Space headway (m)	5.79	25.93	106.70	14.13	4.24	18.94	80.14	8.70	4.65	45.21	144.12	24.4	0.000	0.000	0.000	0.000
Relative speed (m/s)	-19.23	-0.81	8.98	6.11	-5.80	-0.05	5.40	1.20	-13.59	-1.49	15.47	2.9	0.000	0.000	0.000	0.000

Table 2: Descriptive statistics of the two datasets

**Table 3: Models parameter estimates, t-test of individual parameter equivalence and TTS results**

Variable	Driving simulator data	Video trajectory data - I80		Video trajectory data – M1		t-tests of individual parameter equivalence		
	(Model 1)	Robust	Parameter	Robust	Parameter	Robust	Models 1 & 2	Models 1 & 3
	Parameter estimate	t-ratio	estimate	t-ratio	estimate	t-ratio	t-stat	t-stat
Reaction time distribution								
$\mu_t$	0.664	14.66	-0.3973	-16.42	0.6204	5.72	20.67	0.37
$\sigma_t$	0.3536	2.69	0.3257	66.48	0.6517	4.63	0.21	-1.55
Car-following acceleration								
Constant	0.3506	6.96	0.8304	13.77	0.4283	4.59	-6.11	-0.73
time headway (s)	0.2856	1.85	0.792	8.99	0.0218	0.19	-2.85	1.37
relative speed (m/s)	0.6787	9.71	0.8982	20.58	0.4694	1.37	-2.66	0.60
$\sigma^{acc}$	0.3367	25.26	0.7318	76.71	0.6976	9.07	-24.10	-4.62
Car-following deceleration								
Constant	-0.255	-5.47	-0.5128	-16.03	-0.9206	-25.25	4.56	11.25
time headway (s)	0.4798	2.66	0.1941	2.36	0.2609	8.67	1.44	1.20
relative speed (m/s)	0.7043	9.38	0.928	25.98	0.5195	18.91	-2.69	2.31
$\sigma^{dec}$	0.6893	16.61	0.8007	75.16	0.7545	42.55	-2.60	-1.44
LL = -5610.845		LL = -17240.88		LL = -3857.24				
$\rho^2 = 0.320$		$\rho^2 = 0.138$		$\rho^2 = 0.463$				
Adj. $\rho^2 = 0.319$		Adj. $\rho^2 = 0.137$		Adj. $\rho^2 = 0.461$				
obs = 7191		obs = 14826		obs = 3302				
Transferability Test Statistic (TTS)								
Summary statistics			Simulator to I-80 data transferability			Simulator to UK data transferability		
Degrees of freedom (Dof)			10			10		
LLest ( $\beta_{transf}$ )			-27126.16			-7118.384		
LLapplic( $\beta_{appl}$ )			-17240.88			-3857.238		
-2[LLapplic( $\beta_{transf}$ ) -LLapplic( $\beta_{appl}$ )]			19770.56			6522.29		

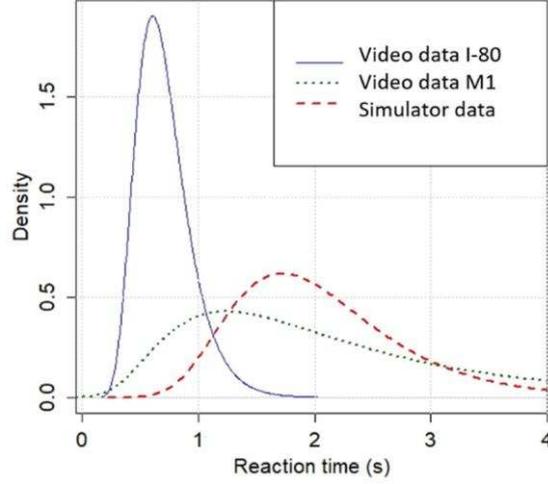


Figure 3: Reaction time distributions of the car-following models

$$a_n^{cf,acc}(t) = 0.3506 \frac{1}{\Delta T_n(t)^{0.2856}} |\Delta V_n(t - \tau_n)|^{0.6787} + \varepsilon_n^{cf,acc}(t) \quad (17)$$

where  $\varepsilon_n^{cf,acc}(t) \sim N(0, 0.3367^2)$ .

In a similar way, Equation 18 presents the deceleration component of Model 1

$$a_n^{cf,dec}(t) = -0.255 \frac{1}{\Delta T_n(t)^{0.4798}} |\Delta V_n(t - \tau_n)|^{0.7043} + \varepsilon_n^{cf,dec}(t) \quad (18)$$

where  $\varepsilon_n^{cf,dec}(t) \sim N(0, 0.6893^2)$ .

The acceleration and deceleration constants both have the expected signs and are statistically significant at 0.05 level. Moreover, the parameters of time headway both have positive signs but the parameter for acceleration regime is significant at the 0.1 level. The positive sign for the time headway parameter of the acceleration regime implies that drivers tend to react less to the leader's speed as time headway increases and they get closer to a free-flow state. Regarding the deceleration regime, the positive sign of the time headway parameter indicates that drivers adopt smaller decelerations at larger headways. The schematic interpretation of the aforementioned parameters and their effects on acceleration/deceleration are illustrated in the next section. Finally, the parameters of relative speed are significant for both acceleration and deceleration regimes. It is worth mentioning that the estimates are in accordance with the a-priori expected values (smaller than 1) as the acceleration or deceleration capabilities of the driver are constrained by the vehicle capability. The impact of each parameter is depicted more explicitly in the sensitivity analyses presented in the next section. Finally, Figure 4 shows the reaction time distribution as expressed by the estimated mean and standard deviation. The distribution extends approximately to the whole 0-4s range and its peak is between 1.5s and 2s. The estimated distribution of reaction time is

$$f(\tau_n) = \begin{cases} \frac{1}{0.3536 \tau_n} \varphi\left(\frac{\ln(\tau_n) - 0.664}{0.3536}\right) & \text{if } 0 < \tau_n \leq 4 \\ 0 & \text{otherwise} \end{cases} \quad (19)$$

Model 2: Car-following model based on the I-80 video data (USA)

The estimated car-following acceleration of Model 2 is presented in Equation 20

$$a_n^{cf,acc}(t) = 0.8304 \frac{1}{\Delta T_n(t)^{0.792}} |\Delta V_n(t - \tau_n)|^{0.8982} + \varepsilon_n^{cf,acc}(t) \quad (20)$$

where  $\varepsilon_n^{cf,acc}(t) \sim N(0, 0.7318^2)$ .

The deceleration component of Model 2 is shown in Equation 21:

$$a_n^{cf,dec}(t) = -0.5128 \frac{1}{\Delta T_n(t)^{0.1941}} |\Delta V_n(t - \tau_n)|^{0.928} + \varepsilon_n^{cf,dec}(t) \quad (21)$$

where  $\varepsilon_n^{cf,dec}(t) \sim N(0, 0.8007^2)$ .

The results of the car-following model estimation based on the I-80 trajectory data are presented in Table 3. All the parameters have expected signs and are significant at 0.05 level. Moreover, the values of relative speed parameters are below 1, as a-priori expected. The reaction time distribution (Figure 4) extends between 0-2s while its peak is approximately after 0.5s. This outcome suggests that drivers' reaction time in real traffic is smaller compared to simulated driving (i.e. the drivers respond faster to the relative speed stimulus in field traffic conditions. This might be a potential indication that drivers perceive changes in traffic conditions differently in the simulator compared to field traffic driving (where a crash occurrence would have genuine consequences). The estimated distribution of reaction time is shown in Equation 22

$$f(\tau_n) = \begin{cases} \frac{1}{0.3257 \tau_n} \varphi\left(\frac{\ln(\tau_n) + 0.3973}{0.3257}\right) & \text{if } 0 < \tau_n \leq 4 \\ 0 & \text{otherwise} \end{cases} \quad (22)$$

Model 3: Car-following model based on the M1 video data (UK)

The estimated car-following acceleration of Model 3 is presented in Equation 23

$$a_n^{cf,acc}(t) = 0.4283 \frac{1}{\Delta T_n(t)^{0.0218}} |\Delta V_n(t - \tau_n)|^{0.4694} + \varepsilon_n^{cf,acc}(t) \quad (23)$$

where  $\varepsilon_n^{cf,acc}(t) \sim N(0, 0.6976^2)$ .

Finally, the deceleration component of Model 3 is shown in Equation 24:

$$a_n^{cf,dec}(t) = -0.9206 \frac{1}{\Delta T_n(t)^{2609}} |\Delta V_n(t - \tau_n)|^{0.5195} + \varepsilon_n^{cf,dec}(t) \quad (24)$$

where  $\varepsilon_n^{cf,dec}(t) \sim N(0, 0.7545^2)$ .

With respect to the acceleration regime, only the constant parameter is significant, which implies that acceleration behaviour at the M1 weaving section is captured to a lesser extent by the car-following model. On the other hand, the parameters related to the deceleration regime are all significant with expected values and signs. It is worth mentioning that the deceleration constant has a higher absolute value compared to the respective parameters of the driving simulator and I-80 models. This finding is an additional indication that the type of the data collection site had an impact on the parameter estimates, as the higher rate of lane-changing may have also resulted in more frequent and higher deceleration. The mean of the reaction time distribution (Figure 4) is between the I-80 and driving simulator data models. This finding may denote higher alertness at the congested driving conditions of the USA data than the UK weaving section. The estimated distribution of reaction time is shown in Equation 25

$$f(\tau_n) = \begin{cases} \frac{1}{0.6517 \tau_n} \varphi\left(\frac{\ln(\tau_n) - 6204}{0.6517}\right) & \text{if } 0 < \tau_n \leq 4 \\ 0 & \text{otherwise} \end{cases} \quad (25)$$

#### 4.1.2 Model comparison and sensitivity analysis

The initial evaluation of the estimated parameters shows that their signs are expected while most of them are significant. The current section investigates the effects of models' variables in the car-following acceleration (deceleration) to further assess the parameter estimates and ultimately compare the extent and nature of differences between the two models. Figure 5 depicts the sensitivity analysis for all parameters while the transferability of the model based on driving simulator data is then compared with the two field traffic data sources.

Focusing on Model 1 (driving simulator), the results indicate that the value of acceleration slightly decreases with the increase of time headway. This pattern reflects drivers' expected behaviour in acceleration state, as also explained earlier. On the other hand, the absolute value of deceleration increases with the decrease of time headway. This outcome meets the expectations, since drivers will decelerate to a higher extent when time headway is short and relative speed is negative (deceleration regime), while it also indicates drivers' safety concerns; as time headway decreases, drivers decelerate to avoid collision. These interpretations also apply for the patterns related to time headway with respect to Model 2 (I-80) and Model 3 (M1). Starting from the latter model, the slope of acceleration is similar to the simulator context however, the values are higher. On the other hand, deceleration values are of considerably higher magnitude both compared to the simulator and the I-80 results. This finding is expected considering the high deceleration constant value of that model and can be an outcome of manoeuvre behaviour that takes place at the weaving section and potentially higher variance in the deceleration behaviour and alertness. Regarding the acceleration-time headway plot of Model 2, the observed steeper slopes, which also result in higher absolute acceleration/deceleration values, may highlight the differences in drivers' sensitivity between simulated and real driving. In the acceleration regime, the

aforementioned differences could indicate a higher variance in traffic conditions – which influences applied acceleration - or suggest that drivers have higher sensitivity in the decrease of time headway during real road traffic conditions. A similar trend is also observed in the deceleration regime of this model - as the plots show that deceleration rate is in general higher - compared to the driving simulator case. This outcome might be an effect of the smaller variance in traffic conditions of the simulated scenario (as happens in the acceleration regime), but given the higher value of the standard deviation parameter in the deceleration regime, and thus a potentially higher variance in observed behaviour, it might also indicate that drivers assess or perceive risk in a different way (e.g. absence of real danger) and this behaviour could influence their deceleration decisions.

Regarding relative speed, acceleration and deceleration reach their maximum absolute values when the former is maximum and minimum, respectively, in both models. The obtained acceleration-regime trends follow the underlying theory of the current car-following model as drivers' acceleration tends to increase while lead vehicle moves faster. In the same way, the increase in absolute deceleration as relative speed declines, is consistent with the safety implications reported about the effect of time headway, as a driver is expected to decelerate to a higher degree when relative speed gets smaller. It should be mentioned that the patterns observed in Model 3 (M1, UK) are closer to Model 1 in terms of acceleration and Model 2 in terms of deceleration. The latter finding may show the difference in drivers' sensitivity and safety concerns between real and simulated traffic. On the other hand, the higher acceleration values related to Model 2 (I-80 data) may be related to lower speeds and more frequent stop and go behaviour that has taken place owing to congestion.

The different patterns in acceleration (deceleration) that the sensitivity analysis revealed, are further investigated through model comparison, in terms of parameter equivalence and model transferability. The results of the t-test of individual parameter equivalence (Table 3) show that for most of the parameters of Models 1 and 2, the null hypothesis of equivalence is rejected ( $|t\text{-value}| > 1.96$ ). The t-stat of the difference is non-significant only for the standard deviation of reaction time distribution and the time headway parameter of the deceleration regime. Moreover, the results of TTS regarding transferability from driving simulator to real driving context, as reflected in the I-80 data, show that the null hypothesis of equivalence between the two models is rejected at 0.05 level ( $\chi^2_{\text{critical}} = 18.31$ ), thus, transferability cannot be validated.

The results of the t-test of individual parameter equivalence between Models 1 & 3 show that apart from the standard deviation, all the other parameters related to the acceleration regime are transferable. This can be an indication that models' results are more transferable between driving simulator and field observations, for data collected in the same country however, it should be also mentioned that most of the acceleration regime parameters of the field M1 data have not been significant. On the other hand, all parameters related to the deceleration regime are significantly different. Moreover, the results of the TTS regarding transferability from driving simulator to the M1 (UK) context, show that the null hypothesis of equivalence between the two models is rejected at 0.05 level ( $\chi^2_{\text{critical}} = 18.31$ ). Thus, transferability from simulated to real traffic driving cannot be validated for both the available video trajectory datasets.

It may be noted that while testing transferability with simpler models (Papadimitriou & Choudhury, 2017), though majority of the parameters have not been found to be transferable, the null hypothesis of equivalence has not been rejected. Albeit the space headway and subjects' speed, of the data used in the current study, have a closer match with the I-80 data in comparison with the sim data used in the previous study, this may indicate that incorporation of reaction time heterogeneity increases the gap between the two sets of models in terms of transferability.

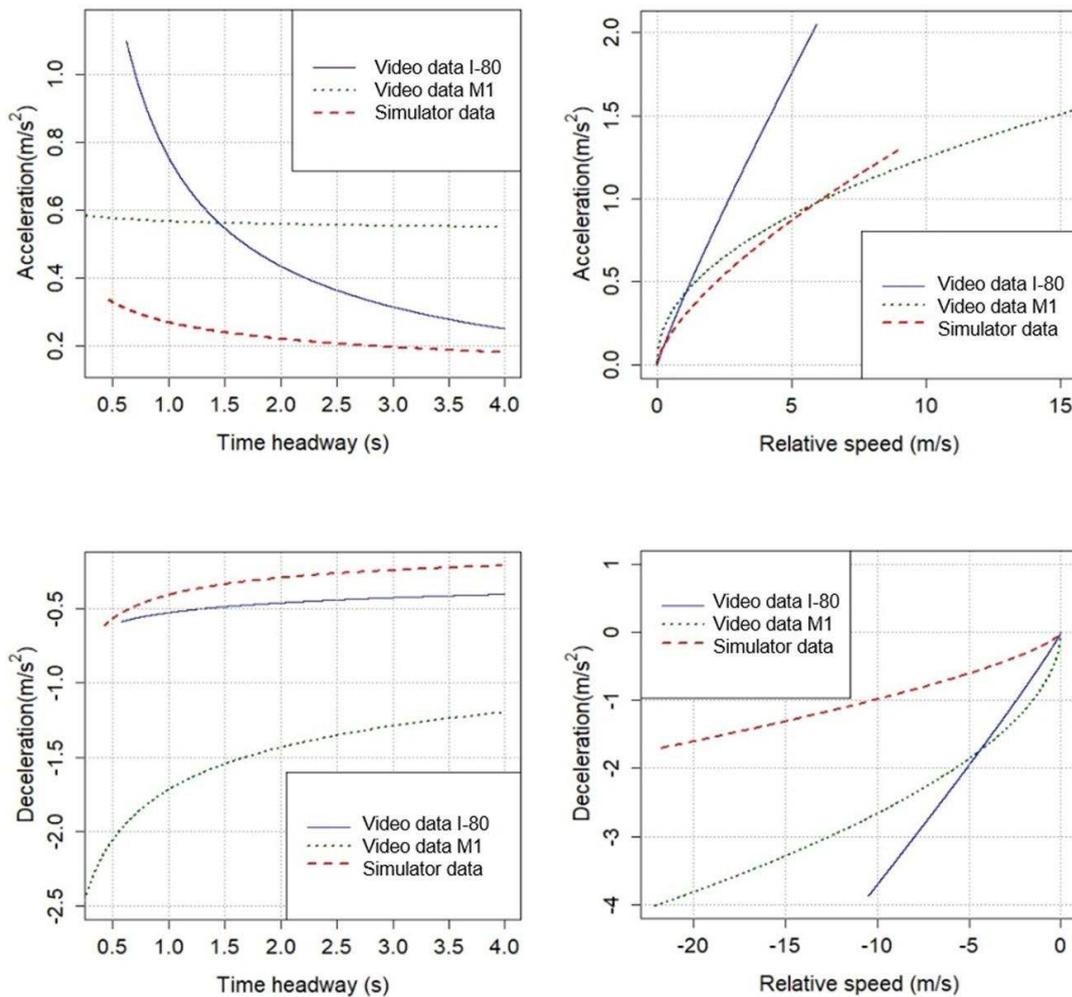


Figure 4: Sensitivity analysis of the car-following models

The results from the transferability analysis show that a car-following model developed by driving simulator data cannot be directly used for real-driving applications (e.g. in microsimulation). As later discussed in the Conclusion section, the differences between the models may have also occurred also for other reasons, other than the different nature of the two datasets. In the next section a series of approaches is investigated in order to address this issue and reduce the gap.

## 5. Model updating and joint estimation

The analysis described in the previous section highlights the lack of transferability from driving simulator models to the field. The current section investigates two different

updating approaches that aim to reduce the potential behavioural bias of driving simulator data and identify the most suitable of them in order to develop a context for its application in a real driving framework. Moreover, the results of the car-following models are compared with the results of joint model estimated combining the simulator data with each of the video datasets.

### 5.1.1 Model updating

The parameters of the driving simulator model have been updated using the Bayesian updating (Dey & Fricker, 1994) and Combined Transfer Estimation (Ben-Akiva & Bolduc, 1987) approaches. The updated parameters and the results of the TTS after the application of model updating are presented in Table 4 referring to the I-80 data and Table 5 referring to the M1 data. The TTS value after applying Bayesian updating indicates that the null hypothesis of model equivalence is in both cases rejected at the 0.05 level ( $\chi^2_{\text{critical}} = 18.31$ ). However, the TTS value of the Combined Transfer Estimation shows that after updating, the null hypothesis is not rejected and thus, driving simulator data can be considered transferable for both the I-80 and M1 datasets. This finding is consistent with existing literature, where Combine Transfer Estimation outperforms Bayesian updating.

Table 4: Parameters and TTS results after model updating – I-80 data

Variable	Bayesian updating	Combined Transfer Estimation
	Parameter estimate	Parameter estimate
Reaction time distribution		
$\mu_t$	-0.162	-0.398
$\sigma_t$	0.326	0.326
Car-following acceleration		
Constant	0.548	0.838
time headway (s)	0.667	0.809
relative speed (m/s)	0.837	0.908
$\sigma^{\text{acc}}$	0.598	0.732
Car-following deceleration		
Constant	-0.430	-0.517
time headway (s)	0.243	0.149
relative speed (m/s)	0.887	0.935
$\sigma^{\text{dec}}$	0.794	0.802
Transferability Test Statistic (TTS)		
Summary statistics	Bayesian updating	Combined Transfer Estimation
Degrees of freedom (Dof)	10	10
LLapplic( $\beta_{\text{transf}}$ )	-17884.1	-17245.46
LLapplic( $\beta_{\text{applic}}$ )	-17240.88	-17240.88
$-2[\text{LLapplic}(\beta_{\text{transf}}) - \text{LLapplic}(\beta_{\text{applic}})]$	1286.44	9.16

The effects of the updating techniques and the amendments of each, on the parameters of Model 1, can be demonstrated more rigorously with the application of the sensitivity analysis described in previous section. The results of both approaches are illustrated in Figures 6 and 7 for Model 2 and Figures 8 and 9 for Model 3. The parameters after applying Bayesian updating on Model 1 with respect to the parameters of Model 2 (Figure 6), result in the occurrence of lines which are in higher proximity – compared to the initial sensitivity analysis as presented in Figure 5 – but still there is a distinct difference between the two cases. On the other hand, the new set of parameters, after

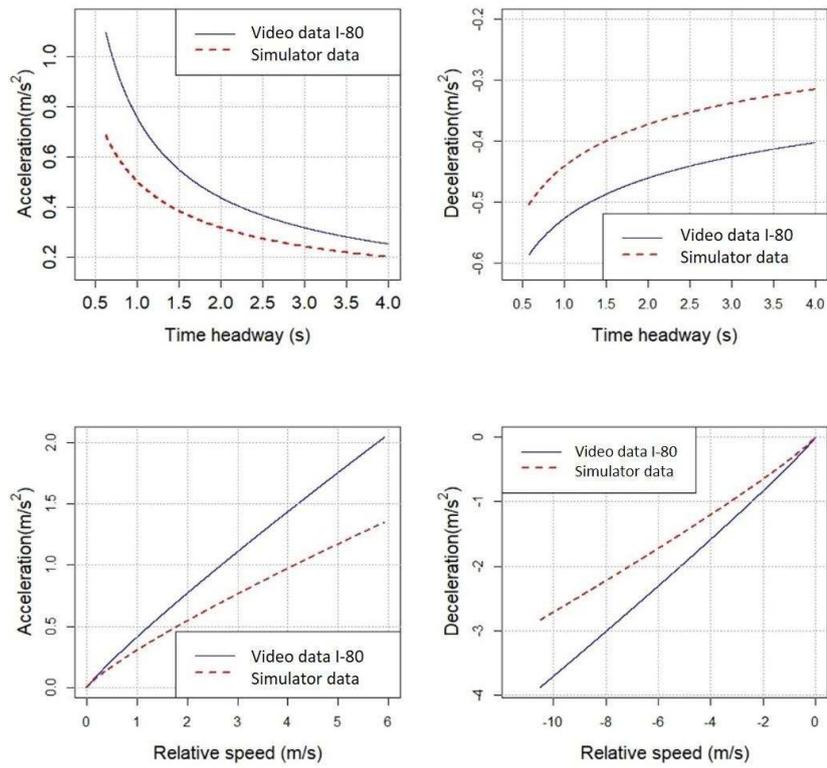


Figure 5: Sensitivity analysis after Bayesian updating – I-80 data

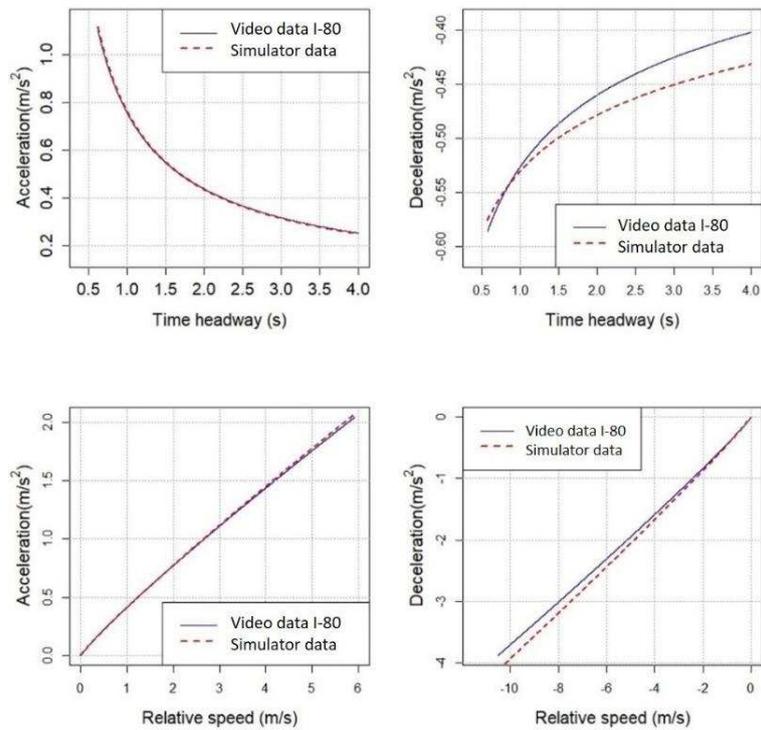


Figure 6: Sensitivity analysis after Combined Transfer Estimation – I-80 data

Combined Transfer Estimation (Figure 7), produces very similar outcomes. In particular, the acceleration regime of all models is almost identical, while some differences can be noticed in the deceleration regime. The results of the sensitivity analysis are consistent with the outcomes presented in Table 4 and provide a more detailed investigation regarding the effects of each updating technique on each of the elements of the car-following model based on driving simulator data.

The impact of the two examined updating techniques is similar, with respect to the M1 data. In particular, the application of Bayesian updating (Figure 8) results in some non-significant improvement however, after Combined Transfer Estimation (Figure 9), the results from the updated Model 1 and Model 3 are significantly closer. It may be noted in Figure 9, the acceleration regime of Model 3 is not approximated entirely accurately, even after the application of CTE. This could have been potentially caused by the ‘noise’ in the M1 data caused by the increased number of lane-changes occurring in the section. The deceleration regime of Model 3 was however approximated more accurately, compared to the respective one of Model 2.

Table 5: Parameters and TTS results after model updating – M1 data

	Bayesian updating	Combined Transfer Estimation
Variable	Parameter estimate	Parameter estimate
<hr/>		
Reaction time distribution		
$\mu_t$	0.658	0.663
$\sigma_t$	0.492	0.766
Car-following acceleration		
Constant	0.368	0.299
time headway (s)	0.113	-0.079
relative speed (m/s)	0.670	0.781
$\sigma^{acc}$	0.347	0.715
Car-following deceleration		
Constant	-0.668	-0.923
time headway (s)	0.267	0.247
relative speed (m/s)	0.541	0.514
$\sigma^{dec}$	0.744	0.764
<hr/>		
Transferability Test Statistic (TTS)		
<hr/>		
Summary statistics	Bayesian updating	Combined Transfer Estimation
<hr/>		
Degrees of freedom (Dof)	10	10
LLapplic( $\beta_{transf}$ )	-4556.65	-3861.82
LLapplic( $\beta_{applic}$ )	-3857.24	-3857.24
$-2[LLapplic(\beta_{transf}) - LLapplic(\beta_{applic})]$	1398.81	9.16

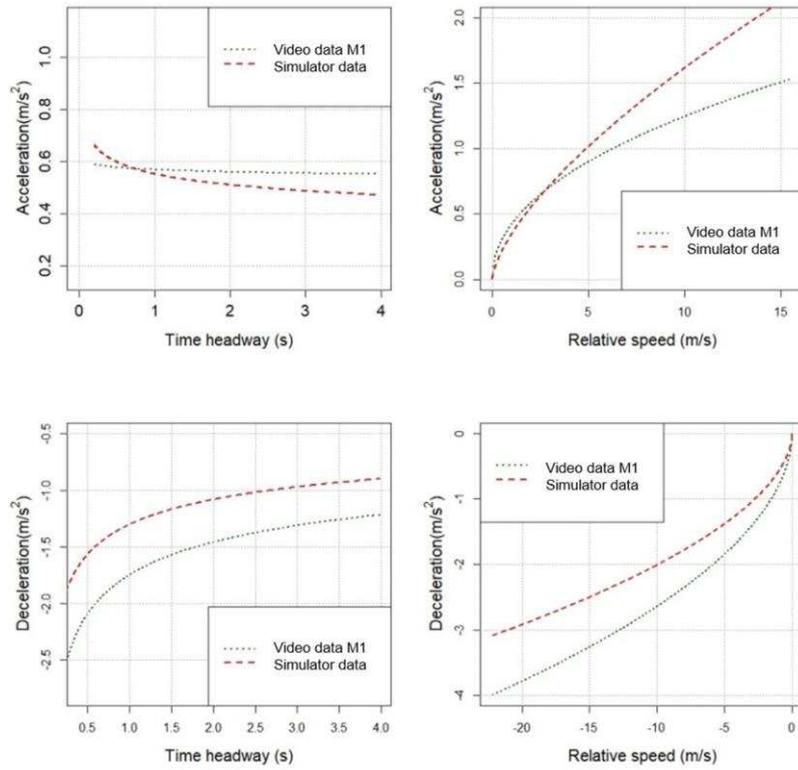


Figure 7: Sensitivity analysis after Bayesian updating – M1 data

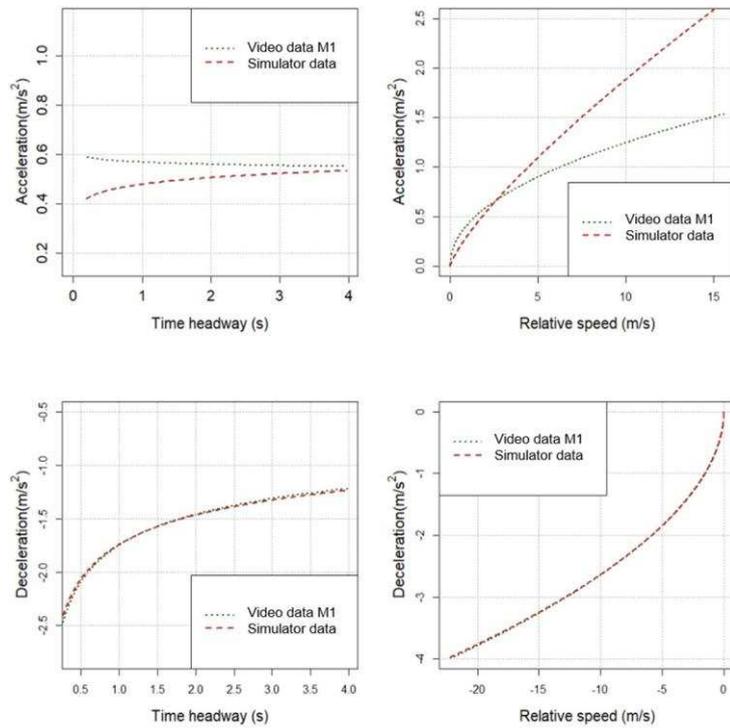


Figure 8: Sensitivity analysis after Combined Transfer Estimation – M1 data

### 5.1.2 Joint estimation results

The differences between the driving simulator and field trajectory datasets are further investigated in the context of joint model estimation. In this approach, the car-following model is estimated combining simultaneously both data sources, using the driving simulator data and each of the video datasets separately. Initially, the datasets have been considered as the same source and a single set of parameters have been estimated. The results of this model are not presented in the context of the present analysis, but its final log-likelihood value has been used for comparison purposes in Table 6. As a next step, a series of scale parameters is introduced to account for the differences between trajectory and simulator data. The scale parameters are applied to the sensitivity $\times$ stimulus terms, the standard deviation parameters and the reaction time parameters with the formulation  $(\delta^{\text{trajectory}} + \delta^{\text{simulator}} \times \text{scale})$ , where  $\delta^{\text{trajectory}}$  is a dummy variable equal to 1 if the observation belongs to the trajectory dataset and  $\delta^{\text{simulator}}$  is a dummy variable equal to 1 if the observation belongs to the driving simulator dataset. Six scale parameters are used in total. In essence, given that for every density function involved in the model specification (acceleration, deceleration and reaction time) a mean and a standard deviation component is estimated, each of the scale parameters is used to approximate the difference of the driving simulator data estimates with respect to the video trajectory data estimates. After incorporating the aforementioned scale specification in Equation 7, the acceleration/deceleration density function is given from Equation 26

$$f(a_n^{\text{cf,g}}(t)|\tau_n) = \frac{1}{\sigma_{\text{e,cf,g}}} \varphi \left( \frac{a_n^{\text{cf,g}}(t) - s[X_n^{\text{cf,g}}(t - \tau_n)] \times f[\Delta V_n(t - \tau_n)](\delta^{\text{traj}} + \theta^\mu \delta^{\text{sim}})}{\sigma_{\text{e,cf,g}}(\delta^{\text{traj}} + \theta^\sigma \delta^{\text{sim}})} \right) \quad (26)$$

where  $\theta^\mu$  and  $\theta^\sigma$  represent the scale parameters of mean and standard deviation respectively.

The parameter estimates of the model are presented in Table 6. Owing to the model specification, the t-ratio values of the scale parameters refer to the comparison with 1 rather than 0. Regarding Model 2 (I-80 data), all scale parameters, apart from reaction time standard deviation, are significantly different from 1. This result consists an additional indication to the tests applied in the previous section, that for joint estimation, the differences between the various data sources should be considered and captured. With respect to Model 3, less of the scale parameters are significant. Some of the results are consistent with a-priori expectations. For instance, the initial t-test of parameter equivalence between Models 1 and 3 (Table 3) showed that parameters related to acceleration were transferable while the standard deviation of was not. This has been reflected in the scale results with a non-significant and a significant scale parameter respectively. Moreover, most of the deceleration parameters of the same models were not equivalent while the differences in deceleration variance were not significant. This finding is also consistent with the scale parameters.

To further asses each joint model, they have been compared, using the likelihood ratio test, with (a) a joint model estimated without any scale parameters and (b) Models 1 & 2 (I-80 data) and Models 1 & 3 (M1 data) respectively. Regarding case (b), the log-likelihood of the joint model is compared with the sum of log-likelihood values of the relevant pair of models, with degrees of freedom equal to the sum of the parameters of the initial individual models minus the estimated parameters of the joint model. The

null hypothesis is rejected for both cases (a) and (b) and for both field datasets, indicating two different types of outcomes, with respect to the use of scale parameters. At first, the joint model without accounting for scale differences, does not perform as well as the model including the scale parameters. This result is expected and consistent with all findings presented in the current analysis. The differences in driving behaviour between simulated and real road traffic driving affect model fit and need to be considered. On the other hand, the use of scale parameters, for both video datasets, does not manage to improve the model sufficiently, since the results of the likelihood ratio test with the individual models show that the joint model does not perform equally to the two separate ones. As a brief conclusion, it should be mentioned that the use of scale parameters, in the suggested model specification, improves model fit but does not fully capture the differences between the different data sources and further approaches should be considered.

Table 6: Parameter estimates of the joint models

Variable	Video data I-80		Video data M1	
	Parameter estimate	Robust t-statistic	Parameter estimate	Robust t-statistic
Reaction time distribution (Video trajectory data)				
$\mu_t$	-0.3964	-16.45	0.6874	5.46
$\sigma_t$	0.3264	65.2	0.7136	5.03
Car-following acceleration				
Constant	0.7209	14.88	0.3902	9.67
time headway (s)	0.5562	6.73	0.1975	1.62
relative speed (m/s)	0.7801	15.1	0.6577	10.21
$\sigma^{acc}$	0.7337	77.11	0.6916	9.43
Car-following deceleration				
Constant	-0.5644	-12.01	-0.8607	-23.87
time headway (s)	0.2584	2.77	0.2932	9.52
relative speed (m/s)	0.8539	19	0.5721	21.77
$\sigma^{dec}$	0.8008	74.95	0.7525	42.5
Scale parameters				
Car-following acceleration mean	0.5435	-9.67 (1)	0.8558	-1.04 (1)
Car-following acceleration std. dev	0.461	-27.90 (1)	0.3372	-31.89 (1)
Car-following deceleration mean	0.3052	-29.92 (1)	0.4869	-9.28 (1)
Car-following deceleration std. dev	0.8629	-2.61 (1)	0.9201	-1.34 (1)
Reaction time mean	-1.65	-16.72 (1)	0.9672	-0.18 (1)
Reaction time std. dev	1.0908	0.23 (1)	0.5004	-2.29 (1)
LL	-22892.67		-9496.16	
$\rho^2$ :	0.19		0.14	
Adj. $\rho^2$ :	0.19		0.14	
LR (compared to a joint model without scale parameters):	3024.48 (dof = 6)		2617.28 (dof = 6)	
LR (compared to Models 1 and 2 combined):	81.89 (dof = 4)		56.15 (dof = 4)	

## 6. Conclusion

The current paper presents a detailed investigation of the applicability of car-following models estimated using driving simulator data to study real road traffic scenarios. This relates primarily to the transferability of model parameters estimated on data from simulator experiments and their suitability for representing driving behaviour in field traffic. While previous studies have conducted such transferability tests with simple models (that ignore heterogeneity in reaction time), a more advanced modelling approach has been adopted which indicated that the differences become more pronounced when the model specifications are more complex.

The analysis is based on the comparison of car-following models estimated using driving simulator data collected at the University of Leeds Driving Simulator and field traffic data from the widely used I-80 trajectory dataset (NGSIM project) and data collected at M1 (a UK weaving section). Transferability between the two contexts has been primarily examined with basic approaches such as the t-test of individual parameter equivalence and the TTS. The results of the initial transferability tests suggest that driving simulator data should be used with caution. For instance, the t-tests for individual parameter equivalence show that almost all parameters are not directly transferable, with respect to the I-80 data, while significant differences have been also found in the deceleration regime of the driving simulator and M1 models. Moreover, the sensitivity analysis shows that in real life, drivers are more sensitive to changes in traffic conditions compared to simulated environments – this can have crucial safety implications. As an example, the results indicate that drivers apply smaller deceleration rates in the simulated environment. Discrepancies like this may misguidedly lead to false interpretation of drivers' behaviour not only in terms of road safety and crash investigation but also in microscopic modelling applications. The findings indicate that the parameters estimated from driving simulator data are not suitable for direct application in such models, which prompted us to investigate methods for improving transferability.

We have applied a series of techniques to improve the transferability of the simulator data, based on a) parameter updating and b) joint estimation that accounts for differences in scale. While Bayesian updating did not validate model transferability, the results of Combined Transfer Estimation have indicated that driving simulator data can be made transferable to a real driving context, opening up new prospects for further research. The joint model estimation consists of several steps where specifications without and with parameters that would account for the differences in scale have been tested. The results of the joint model estimation reveal that there is a statistically significant difference in the scale of both acceleration and deceleration behaviour with respect to the I-80 data and the deceleration regarding the M1 data. Moreover, this model specification performs significantly better compared to a model where a single vector of parameters is used for both datasets (driving simulator and field), without accounting for differences in scales. While the two separate models are always expected to outperform the joint model, our results indicate that the gap in performance was so large as to suggest that joint estimation did not adequately capture the differences between the two cases.

It may be noted that given the secondary nature of the video trajectory data, the traffic variables (speed, acceleration, space headway etc.) used for model estimation are significantly different between the two data sources. These differences are likely to

have arisen not only because of the different nature of the data (field traffic and simulator) but a series of other factors could have also contributed to these discrepancies. In particular, the road environments differ in terms of the surroundings (i.e. landscape) while the motorways have different numbers of lanes, as the I-80 data refers to a 5-lane infrastructure with an additional HOV, the M1 data is a weaving location of three main plus two auxiliary lanes while the simulated motorway has 3 lanes. Although the study focuses on car-following behaviour and thus the number of lanes should not have a major impact, it is still possible that the aforementioned differences can have an effect on the results of the analysis, as drivers' behaviour might have been affected overall. Moreover, The I-80 dataset refers to more congested USA traffic while this is not the case for the M1 data, as reflected by the speed distributions observed in the data. Moreover, in the simulator environment, traffic has been variable with a sense of congestion in one of the segments of the simulator data. Another limitation with respect to the data used the current work regards the differences and variations in driving behaviour owing to the cultural/regional differences of the drivers between the USA and UK for one of our two video datasets. Although these differences can have a significant impact on the model parameters, our results show that the same transferability issues occur also with the UK M1 data. Thus, although regional differences in behaviour may affect the parameter estimates of the models, it may not be the most significant reason as the same patterns in all steps of the analysis have been observed for both the I-80 and UK M1 data, in terms of transferability. Another issue relevant to the previously mentioned, could be related to the direction of traffic (right-hand vs left-hand drive system) however, given that only car-following behaviour has been considered, this effect is not expected to have a major impact.

Some other differences refer to vehicle capabilities in terms of dynamics i.e. the field data includes trajectories from multiple vehicles while the data driving simulator data contains behavioural data from different drivers using the same vehicle. Although this can have a significant impact on the parameters of car-following models, similar studies usually omit the impact of vehicle type, which is averaged in the model parameters, or the categorise vehicles in cars and heavy vehicles (Durrani et al., 2016). This is the case for the majority of driving simulator studies which consider the same vehicle for all the participants. Another source of bias on the estimation results could be related to the temporal differences in data collections, as the I-80, M1 and the simulator data have been collected in 2005, 2013 and 2017 respectively. Since the acceleration and deceleration capabilities and other vehicular features have become more advanced over the years, there is some possibility that the transferability between the I-80 and the simulator data may be affected by the time difference. However, the M1 and the simulator data are temporally closer and less likely to be affected by such differences. Also, although this is a matter that could affect results, the I-80 data is still being used in recent studies related to traffic flow and driving behaviour modelling (I. Ahmed et al., 2019; Mercat et al., 2019; M. Zhao et al., 2019), while given the observed acceleration/deceleration values in the data, it has not been expected to be a major source of bias in transferability.

The analysis in this paper has been undertaken acknowledging all the aforementioned limitations. In order to reduce the potential impacts because of all the issues raised in the previous paragraphs, the datasets have been processed to reduce some of the potential effects. Given that the additional differences between the two data sources pose extra challenges in model transferability or joint estimation, our results are

potentially on the conservative side and present an upper bound on issues of transferability, as these differences may have had a negative impact on the efficiency of the tested approaches, since they have been possibly captured along with the differences in the nature of the data sources. Similar issues are likely to be the case also in future efforts focused on transferability. Field observations cannot be controlled thus difficulties in reproducing the variability of observed real-life traffic flow patterns in simulator will always be present.

In terms of results, Combined Transfer Estimation is found to be the most efficient approach for improving the transferability of a car-following model estimated using driving simulator data to a field traffic context as the joint model estimation does not significantly capture the differences between the two contexts. Though the superior performance of the Combined Transfer Estimation approach, compared to Bayesian updating, has been previously demonstrated for other models (e.g. car ownership, mode choice), to the best of our knowledge, this is the first comparison of the performance of the updating methods in the context of driving behaviour and especially car-following models. Based on the current findings, there is scope to extend the current study to other forms of car-following models e.g. the latent class GM (Koutsopoulos & Farah, 2012), IDM model (Treiber et al., 2000), Optimal Velocity model (Bando et al., 1995), as well as other driving behaviour models e.g. lane-change, passing etc. Moreover, although the approach of joint estimation still requires further improvements, it allows for flexible model specifications that incorporate human factors in driving behaviour models. For instance, Paschalidis et al. (2019) presented a car-following model, estimated with simulator data, that incorporates the impact of acute stress and sociodemographic characteristics. An efficient joint estimation approach would allow for the specification of a similar model combining multiple data sources and accounting for the differences in scale, that would minimise the negative impact on the behavioural realism of the field data.

The transferability of the models estimated using the driving simulator data to the real-world can be improved by combining it with real-world data. To the best of our knowledge, this is the first research that demonstrates this in a systematic approach. While this indeed demonstrates the need of real-world data, given that driving simulator data has unique advantages (e.g. testing the effects of new layouts/design factors that can lead to potentially unsafe conditions, incorporation of socio-demographic attributes of the driver) but also disadvantages (potential behavioural validity issues, smaller sample of participants etc.), the datasets are complementary as opposed to stand-alone sources. The applicability of the techniques presented in the present paper can potentially be extended to any type of driving behaviour modelling. As stated previously, driving simulators allow for the implementation and evaluation of behaviour during extreme or risky scenarios. Although real-world data on extreme scenarios is rare and it may be difficult to empirically test if the exact conclusions will hold, theoretically, the methods presented in the present paper could be implemented in any scenario where driver behaviour is approximated or investigated via econometric modelling approaches and the standard errors of the estimated parameters can be determined. Regarding the data requirements for this case, the closer the gap between the real-world and the driving simulator scenario, the better the transferability should be. To that end, rich naturalistic datasets like SHRP2 (which includes crash and near-crash driving behaviour data) could be used to test if comparable driving simulator scenarios can be created.

Based on the aforementioned discussion, in terms of practical use of the model updating, although the proposed approach still requires actual road data, the framework makes it possible to combine human factors in the driving behaviour models and develop models for emerging technologies and/or risky scenarios (Paschalidis et al., 2019) without compromising the behavioural realism of the real-world data. Thus, bridging the gap between simulated and real driving context enables researchers to utilise the best of both sources of data.

### **Acknowledgement**

The core part of this research is supported by the Next Generation Driving Behaviour Model (NGDBM) project funded by FP7 Marie Curie Career Integration Grant of the European Union (PCIG14-GA-2013-631782) and the Economic and Social Research Council, UK. Prof Stephane Hess' contribution is supported by the European Research Council through the consolidator grant 615596-DECISIONS. We would like to thank Michael Daly of UoLDS team for creating the driving simulator scenarios and Dr Daryl Hibberd for his feedback on the design of the experiments.

### **Data Availability Statement**

Some or all data, models, or code generated or used during the study are available in a repository online in accordance with funder data retention policies (Field data can be found on the NGSIM website:

<https://ops.fhwa.dot.gov/trafficanalysistools/ngsim.htm>).

Some or all data, models, or code generated or used during the study are available from the corresponding author by request (Anonymised driving simulator data can be made available based on request).

### **References**

- Aghabayk, K., Sarvi, M., & Young, W. (2012). Understanding the dynamics of heavy vehicle interactions in car-following. *Journal of Transportation Engineering*, 138(12), 1468-1475.
- Ahmed, I., Xu, D., Roupail, N., & Karr, A. (2019). Lane change rates at freeway weaving sites: trends in HCM6 and from NGSIM trajectories. *Transportation Research Record*, 2673(5), 627-636.
- Ahmed, K. I. (1999). Modeling drivers' acceleration and lane changing behavior. Massachusetts Institute of Technology,
- Atherton, T. J., & Ben-Akiva, M. (1976). Transferability and updating of disaggregate travel demand models.
- Bando, M., Hasebe, K., Nakayama, A., Shibata, A., & Sugiyama, Y. (1995). Dynamical model of traffic congestion and numerical simulation. *Physical review E*, 51(2), 1035.
- Bella, F., Garcia, A., Solves, F., & Romero, M. A. (2007). Driving simulator validation for deceleration lane design. Retrieved from
- Ben-Akiva, M., & Bolduc, D. (1987). Approaches to model transferability and updating: the combined transfer estimator: Département d'économie, Université Laval.
- Ben-Akiva, M., Bradley, M., Morikawa, T., Benjamin, J., Novak, T., Oppewal, H., & Rao, V. (1994). Combining revealed and stated preferences data. *Marketing Letters*, 5(4), 335-349.

- Brackstone, M., & McDonald, M. (2003). Driver behaviour and traffic modelling. Are we looking at the right issues? Paper presented at the IEEE IV2003 Intelligent Vehicles Symposium. Proceedings (Cat. No. 03TH8683).
- Branzi, V., Domenichini, L., & La Torre, F. (2017). Drivers' speed behaviour in real and simulated urban roads—A validation study. *Transportation Research Part F: Traffic Psychology and Behaviour*, 49, 1-17.
- Buse, D. S., Schettler, M., Kothe, N., Reinold, P., Sommer, C., & Dressler, F. (2018). Bridging worlds: Integrating hardware-in-the-loop testing with large-scale VANET simulation. Paper presented at the 2018 14th Annual Conference on Wireless On-demand Network Systems and Services (WONS).
- Chang, D. (2018). Transferability Of Merging Behaviour Models Based On Gap Acceptance Theory. University of Leeds.
- Chen, S., Chen, Y., Zhang, S., & Zheng, N. (2019). A novel integrated simulation and testing platform for self-driving cars with hardware in the loop. *IEEE Transactions on Intelligent Vehicles*, 4(3), 425-436.
- Dey, S. S., & Fricker, J. D. (1994). Bayesian updating of trip generation data: combining national trip generation rates with local data. *Transportation*, 21(4), 393-403.
- Durrani, U., Lee, C., & Maoh, H. (2016). Calibrating the Wiedemann's vehicle-following model using mixed vehicle-pair interactions. *Transportation Research Part C: Emerging Technologies*, 67, 227-242. doi:10.1016/j.trc.2016.02.012
- Farah, H., Bekhor, S., Polus, A., & Toledo, T. (2009). A passing gap acceptance model for two-lane rural highways. *Transportmetrica*, 5(3), 159-172.
- Fouladinejad, N., Fouladinejad, N., Jalil, M. A., & Taib, J. M. (2011). Modeling virtual driving environment for a driving simulator. Paper presented at the 2011 IEEE International Conference on Control System, Computing and Engineering.
- Galbraith, R. A., & Hensher, D. A. (1982). Intra-metropolitan transferability of mode choice models. *Journal of Transport Economics and Policy*, 7-29.
- Gazis, D. C., Herman, R., & Rothery, R. W. (1961). Nonlinear follow-the-leader models of traffic flow. *Operations research*, 9(4), 545-567.
- Godley, S. T., Triggs, T. J., & Fildes, B. N. (2002). Driving simulator validation for speed research. *Accident Analysis & Prevention*, 34(5), 589-600.
- Hadayeghi, A., Shalaby, A. S., Persaud, B. N., & Cheung, C. (2006). Temporal transferability and updating of zonal level accident prediction models. *Accident Analysis & Prevention*, 38(3), 579-589.
- Halkias, J., & Colyar, J. (2006). NGSIM interstate 80 freeway dataset. US Federal Highway Administration, FHWA-HRT-06-137, Washington, DC, USA.
- Hancock, P. (1999). Is car following the real question—are equations the answer? *Transportation Research Part F: Traffic Psychology and Behaviour*, 2(4), 197-199.
- Herman, R., & Rothery, R. W. (1965). Car Following and Steady-State Flow. Paper presented at the 2nd International Symposium on the Theory of Traffic Flow., Paris.
- Hoogendoorn, S., & Hoogendoorn, R. (2010). Calibration of microscopic traffic-flow models using multiple data sources. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 368(1928), 4497-4517.
- Hoogendoorn, S. P. (2005). Unified approach to estimating free speed distributions. *Transportation Research Part B: Methodological*, 39(8), 709-727.

- Hou, Y., Zhao, Y., Hulme, K. F., Huang, S., Yang, Y., Sadek, A. W., & Qiao, C. (2014). An integrated traffic-driving simulation framework: Design, implementation, and validation. *Transportation Research Part C: Emerging Technologies*, 45, 138-153.
- Hussain, Q., Alhajyaseen, W. K., Pirdavani, A., Reinolsmann, N., Brijs, K., & Brijs, T. (2019). Speed perception and actual speed in a driving simulator and real-world: A validation study. *Transportation Research Part F: Traffic Psychology and Behaviour*, 62, 637-650.
- Karasmaa, N. (2007). Evaluation of transfer methods for spatial travel demand models. *Transportation Research Part A: Policy and Practice*, 41(5), 411-427.
- Karimi, A., Bassani, M., Boroujerdian, A., & Catani, L. (2020). Investigation into passing behavior at passing zones to validate and extend the use of driving simulators in two-lane roads safety analysis. *Accident Analysis & Prevention*, 139, 105487.
- Koppelman, F. S., & Wilmot, C. G. (1982). Transferability analysis of disaggregate choice models. *Transportation Research Record*, 895, 18-24.
- Koutsopoulos, H. N., & Farah, H. (2012). Latent class model for car following behavior. *Transportation Research Part B: Methodological*, 46(5), 563-578.
- Kusuma, A. (2015). Modelling driving behaviour at motorway weaving sections. University of Leeds,
- Laagland, J. (2005). How to model aggressive behavior in traffic simulation. Paper presented at the Twente Student Conference on IT, Enschede.
- Lee, H. C. (2003). The validity of driving simulator to measure on-road driving performance of older drivers. *Transport engineering in Australia*, 8, 89-100.
- McGehee, D. V., Mazzae, E. N., & Baldwin, G. S. (2000). Driver reaction time in crash avoidance research: validation of a driving simulator study on a test track. Paper presented at the Proceedings of the human factors and ergonomics society annual meeting.
- Mercat, J., Zoghby, N. E., Sandou, G., Beauvois, D., & Gil, G. P. (2019). Inertial Single Vehicle Trajectory Prediction Baselines and Applications with the NGSIM Dataset. arXiv preprint arXiv:1908.11472.
- Montanino, M., & Punzo, V. (2013). Making NGSIM data usable for studies on traffic flow theory: Multistep method for vehicle trajectory reconstruction. *Transportation Research Record*, 2390(1), 99-111.
- Papadimitriou, S., & Choudhury, C. F. (2017). Transferability of car-following models between driving simulator and field traffic. *Transportation Research Record*, 2623(1), 60-72.
- Paschalidis, E., Choudhury, C. F., & Hess, S. (2019). Combining driving simulator and physiological sensor data in a latent variable model to incorporate the effect of stress in car-following behaviour. *Analytic methods in accident research*, 22, 100089.
- Punzo, V., Borzacchiello, M. T., & Ciuffo, B. (2011). On the assessment of vehicle trajectory data accuracy and application to the Next Generation Simulation (NGSIM) program data. *Transportation Research Part C: Emerging Technologies*, 19(6), 1243-1262.
- Risto, M., & Martens, M. H. (2014). Driver headway choice: A comparison between driving simulator and real-road driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 25, 1-9.
- Rossi, R., Meneguzzer, C., & Gastaldi, M. (2013). Transfer and updating of Logit models of gap-acceptance and their operational implications. *Transportation*

- Research Part C: Emerging Technologies, 28, 142-154.  
doi:10.1016/j.trc.2011.05.019
- Saifuzzaman, M., & 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.
- Santoso, D. S., & Tsunokawa, K. (2005). TRANSFERABILITY AND UPDATING ANALYSIS OF MODE CHOICE MODEL BETWEEN TWO DEVELOPING COUNTRIES—PREDICTIVE PERFORMANCE POINT OF VIEW. *Journal of the Eastern Asia Society for Transportation Studies*, 6, 173-185.
- Santoso, D. S., & Tsunokawa, K. (2010). Comparison of updating techniques in transferability analysis of work trip mode choice models in developing countries. *Journal of Advanced Transportation*, 44(2), 89-102.
- SHRP2. (Strategic Highway Research Program). Retrieved March 13, 2016, from <http://www.shrp2nds.us>
- Sikder, S., Pinjari, A. R., Srinivasan, S., & Nowrouzian, R. (2013). Spatial transferability of travel forecasting models: a review and synthesis. *International Journal of Advances in Engineering Sciences and Applied Mathematics*, 5(2-3), 104-128.
- Team, R. C. (2013). R: A language and environment for statistical computing.
- Toledo, T. (2007). Driving behaviour: models and challenges. *Transport Reviews*, 27(1), 65-84.
- Treiber, M., Hennecke, A., & Helbing, D. (2000). Congested traffic states in empirical observations and microscopic simulations. *Physical review E*, 62(2), 1805.
- UDRIVE. (European Naturalistic Driving Study). Retrieved March 28, 2016, from <http://www.udrive.eu>
- University of Leeds Driving Simulator (UoLDS) - <https://uolds.leeds.ac.uk/>
- Van Hinsbergen, C., Schakel, W., Knoop, V., van Lint, J., & Hoogendoorn, S. (2015). A general framework for calibrating and comparing car-following models. *Transportmetrica A: Transport Science*, 11(5), 420-440.
- Van Winsum, W. (1999). The human element in car following models. *Transportation Research Part F: Traffic Psychology and Behaviour*, 2(4), 207-211.
- Wilmot, C. G. (1995). Evidence on transferability of trip-generation models. *Journal of Transportation Engineering*, 121(5), 405-410.
- Yan, X., Abdel-Aty, M., Radwan, E., Wang, X., & Chilakapati, P. (2008). Validating a driving simulator using surrogate safety measures. *Accident Analysis & Prevention*, 40(1), 274-288.
- Yun, M., Zhao, J., Zhao, J., Weng, X., & Yang, X. (2017). Impact of in-vehicle navigation information on lane-change behavior in urban expressway diverge segments. *Accident Analysis & Prevention*, 106, 53-66.
- Zhao, M., Wang, S.-H., Sun, D., & Wang, X.-J. (2019). A Car-Following Model Considering Preceding Vehicle's Lane-Changing Process. *IEEE Access*, 7, 89913-89923.
- Zhao, Y., Wagh, A., Hou, Y., Hulme, K., Qiao, C., & Sadek, A. W. (2016). Integrated traffic-driving-networking simulator for the design of connected vehicle applications: eco-signal case study. *Journal of Intelligent Transportation Systems*, 20(1), 75-87.
- Zhao, Y., Wagh, A., Hulme, K., Qiao, C., Sadek, A. W., Xu, H., & Huang, L. (2012). Integrated traffic-driving-networking simulator: A unique R&D tool for connected vehicles. Paper presented at the 2012 International Conference on Connected Vehicles and Expo (ICCVE).

Zheng, Z. (2014). Recent developments and research needs in modeling lane changing. *Transportation Research Part B: Methodological*, 60, 16-32.

# Appendix

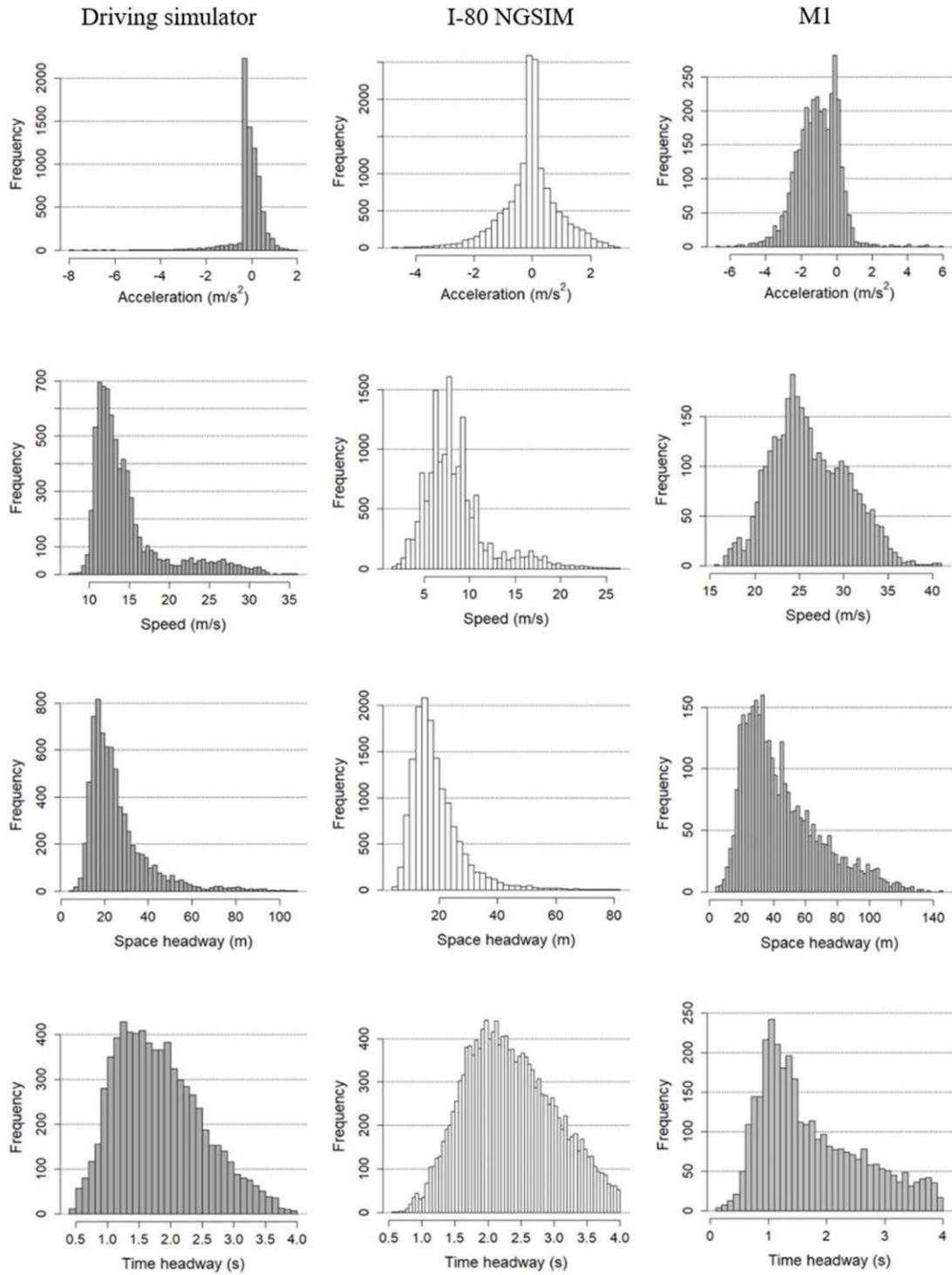


Figure A.1: Histograms of key variables in the video and driving simulator datasets