IMPROVING THE TRANSFERABILTY OF CAR-FOLLOWING MODELS BETWEEN DRIVING SIMULATOR AND FIELD TRAFFIC

Evangelos Paschalidis, Corresponding Author
Choice Modelling Centre
Institute for Transport Studies
University of Leeds, UK
36-40 University Rd, Leeds LS2 9JT
Email: tsep@leeds.ac.uk

Charisma F. Choudhury
Choice Modelling Centre
Institute for Transport Studies
University of Leeds, UK
36-40 University Rd, Leeds LS2 9JT
Email: C.F.Choudhury@leeds.ac.uk

Stephane Hess
Choice Modelling Centre
Institute for Transport Studies
University of Leeds, UK
36-40 University Rd, Leeds LS2 9JT
Email: s.hess@its.leeds.ac.uk

Word count: 5705 words text + 7 tables/figures x 250 words (each) = 7455 words

Submission Date: August 1st 2017

ABSTRACT

Over the last few decades, there have been two different 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 further details and compare alternative methods to overcome the problem. Two main approaches are tested: 1) econometric approaches for increasing model transferability (Bayesian Updating and Combined Transfer Estimation) and 2) joint estimation using both data sources simultaneously. The stimulus-response based car-following models developed using experimental data collected from the University of Leeds Driving Simulator (UoLDS) and detailed trajectory data collected from Interstate 80 (I-80), CA, USA have been used in this regard. T-tests for individual parameter equivalence and Transferability Test Statistic for model transferability (TTS) are used for evaluating the performance of each proposed approach. The results indicate that the transferability can be improved after parameter updating and combined transfer estimation outperforms the other approaches. The findings of this study can be useful in more effectively using driving simulator data for development of mainstream mathematical models of driving behaviour.

Keywords: car-following model, driving simulator, video data, transferability, joint estimation, Bayesian updating, combined transfer estimations

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), have been widely studied in the past few decades *(1, 2)*. Microscopic driving behaviour models are typically developed using two types of data, (a) driving simulator (where drivers drive an instrumented vehicle in a simulated roadway) and (b) real traffic data. Driving simulator data are collected following standardized procedures and are more controllable and reproducible. Furthermore, driving simulators allow researchers to manipulate the surrounding conditions (e.g. geometric layout of the road, number and type of vehicles etc.) as well as driver specific conditions (e.g. level of distraction and fatigue) and run various hypothetical scenarios. However, there is scepticism regarding simulator fidelity (physical and behavioural) and how well drivers' behaviour in a simulator matches with their behaviour in real roads *(3)*. On the other hand, real traffic data best represents true driving behaviour, but have several limitations: measurement errors, complex confounding of influencing factors, less control on the external factors, absence of driver characteristics etc. Given the difference of the two data sources, it is important to investigate the transferability of the model parameters between driving simulator and real traffic. It may be noted that besides these two sources, naturalistic driving data collected using instrumented vehicles (e.g UDRIVE *(4)*, SHRP2 *(5)* etc.) have also been used in research, but given the high costs involved, the availability of these data is still limited. Moreover, similar to driving simulator data, naturalistic data are likely to be prone to behavioural incongruence; and similar to field 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, concrening 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. *(6)* 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 held. Towards the same direction, Yan et al. *(7)* 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 only relative validity was found, regarding safety. Bella et al. *(8)* reproduced a real two-lane road section composed by 11 parts and tested validity in speed. This study confirmed relative but also absolute validity for most of the examined cases. Risto and Martens (*9*) compared the differences in headway choice between an instrumented vehicle and driving simulator without finding significant deviations. Finally, McGehee et al. (*10*) compared drivers' reaction times in real and simulated environment and found statistical equivalence between the two cases.

The development of driving behaviour models based on simulator data has already been reported in literature *(11, 12)*. However, since only relative validity has been established, it remains questionable whether this type of data is suitable for real world applicationos. 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 *(13)*. In this research, we investigate the differences in further details and compare alternative methods to overcome the problem and improve transferability. Morover, we consider drivers' reaction time as a random variable, in order to address this limitation in the previous approach *(13)* and investigate transferability more rigorously.

The present analysis focuses on improving the transferability of car-following models[1] developed using experimental data collected from the University of Leeds Driving Simulator (UoLDS) and detailed trajectory data collected from Interstate 80 (I-80), CA, USA. Based on a review of the literature, two main approaches are tested:

1. Econometric approaches for increasing model transferability
2. Joint Estimation using both data sources simultaneously

The concept of transferability refers to the transfer of a model estimated in a context to a different one, and has been applied in several fields of transportation research. The lion's share is dedicated to the investigation of transferability with the application of discrete choice modelling e.g. *(14, 15, 16, 17)*, however, other modelling approaches can also be found *(18, 19)*.

The joint estimation of models using various data sources was introduced in the discrete choice modelling field *(20)* and mostly refers to the combination of stated-preference and revealed-reference data. The motivation for data combination is the development of enhanced models that exploit the advantages of the various data sources while at the same time minimise their shortcomings, by allowing differences in their scale vary.

The rest of the paper is organised as follows: Section 2 describes the methodological background. This section is followed by the case study description. In section 4 are presented the results of the model estimation and in section 5 the transferability and joint estimation results. The paper concludes with a discussion section.

METHODOLOGICAL BACKGROUND

Car-following model

Basic structure

The model structure is based on the stimulus-response GM car-following model *(21)*. 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 (Equation 1):

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

where: ΔX_n is the space headway at time t, V_n is the following vehicle speed, ΔV_n is the relative speed between the following and the lead vehicle and τ_n is the reaction time. Finally, α, β and γ are constants.

Based on the GM model, several extensions have been suggested. Herman and Rothery *(22)* were the first to highlight that passenger cars have different acceleration and deceleration capacity. In order to address this shortcoming in the GM model, Ahmed *(2)* introduced acceleration-deceleration asymmetry within a stimulus-response framework (Equation 2):

[1] The concept of car-following refers to the applied acceleration of a driver while closely following a leader.

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

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 error term while g represents the car-following regime (acceleration or deceleration). In the present study, the sensitivity and stimulus parts are analysed in (Equations 3 and 4):

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

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

where: ΔX_n is the time headway, ΔV_n is the relative speed between the subject and the lead vehicle and τ_n is the reaction time. Finally, α^g, γ^g and λ^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 was modified in order to consider only time headway, as in *(13)*.

The reaction time distribution

The current model specification also allows for the incorporation of reaction time. Following examples in literature *(2, 23)*, the reaction time is assumed to follow a log-normal truncated distribution (Equation 5):

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

where: $\varphi(.)$ is the standard normal distribution density function, $\Phi(.)$ is the cumulative normal distribution, τ_n is the reaction time of driver n, μ_τ is the mean of the distribution of $\ln(\tau_n)$, σ_τ is the standard deviation and τ^{max}, τ^{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 model parameters. The bounds of reaction time were set between 0 and 4 seconds *(2, 23)*.

Likelihood Function

The assumption of the car-following model is that a driver accelerates if the relative speed is positive and decelerates if negative. Given this, the distribution of acceleration decisions is given, conditionally on reaction time τ, as (Equation 6):

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

where:

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

Assuming that the error terms are normally distributed, the acceleration decisions can be expressed as (Equation 7):

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

where, g ∈ {acc,dec}.

In the current specification, the acceleration observations of each driver n are assumed to be independent while the 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 (Equation 8):

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

The unconditional form of the distribution above is (Equation 9):

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

At the final step, the model is estimated by maximizing the log-likelihood function of the acceleration observations (Equation 10):

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

Evaluating Models Performance and Transferability

A review of the literature revealed several formal statistical tests of transferability *(24)* among which the *t-tests of individual parameter equality* and *Transferability Test Statistic* (TTS) have been found to be most widely used and were thus selected for this study.

The t-tests of individual parameter equality compares individual pairs of coefficients by testing the t-stat difference between the parameter estimates of equivalent variables between the two models as e.g. in *(15)*. The t-stat differences can be expressed as (Equation 11):

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

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

The TTS *(14)* assesses whether the null hypothesis of statistical equivalence between the transferred and the application context model, is rejected or not (Equation 12):

$$TTS_{appl} = -2\left[LL_{appl}(\beta_{est}) - LL_{appl}\left(\beta_{appl}\right)\right] \quad (12)$$

where, $LL_{appl}(\beta_{est})$ is log-likelihood on the application context data using transferred context parameters and $LL_{appl}(\beta_{appl})$ is the log-likelihood on the application context data using application context parameters. The TTS value follows a chi-squared (χ^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 *(17)*. At 95% level of confidence, the models are classified statistically different (i.e. non-transferable) if $\chi^2 > \chi^2_{critical}$.

Evaluating Methods to Improve Transferability

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. *(25)*. Two main methods for model updating are explained below:

Bayesian Updating

The Bayesian process follows the Bayes theorem in which prior information about the model is combined with a random sample from the application context to get updated information that is important in reducing doubt during prediction *(26)*. The parameters estimated with the video data can be used as the prior information in this case and the following formula can be used (Equation 13):

$$\beta_{upt} = \left(\frac{\beta_{est}}{\sigma_{est}^2} + \frac{\beta_{appl}}{\sigma_{appl}^2}\right)\left(\frac{1}{\sigma_{est}^2} + \frac{1}{\sigma_{appl}^2}\right)^{-1}, \tag{13}$$

where β_{est} is the parameter of the estimation (driving simulator) context model, σ_{est} is its standard deviation, β_{appl} is the parameter of the application (real driving) context model and σ_{appl} is its standard deviation.

Combined Transfer Estimation

The combined transfer estimation method *(16)* acknowledges the variations between parameters due to long time gaps and other differences between the estimation and application contexts such that the updated parameters are estimated as (Equation 14):

$$\beta_{upt} = \left(\frac{\beta_{est}}{\sigma_{est}^2 + aa\prime} + \frac{\beta_{appl}}{\sigma_{appl}^2}\right)\left(\frac{1}{\sigma_{est}^2 + aa\prime} + \frac{1}{\sigma_{appl}^2}\right)^{-1} \tag{14}$$

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

CASE STUDY

Data

Video trajectory data

The vehicle trajectories data, used in the analysis, was collected at the Interstate 80 (I-80), CA, USA, within the framework of the Next Generation SIMulation (NGSIM) project *(27)*. The observations took place on the 13th April, 2005. The length of the road segment is approximately 500 meters (1650 feet) and composed by five lanes plus a high occupancy vehicle (HOV) lane (Figure 1a). The vehicles' trajectories referring to the observations from 4.00 p.m. to 4.15 p.m. were further processed by *(28, 29)*. The final dataset contained information regarding the position, speed, acceleration, lane, size and type of each vehicle.

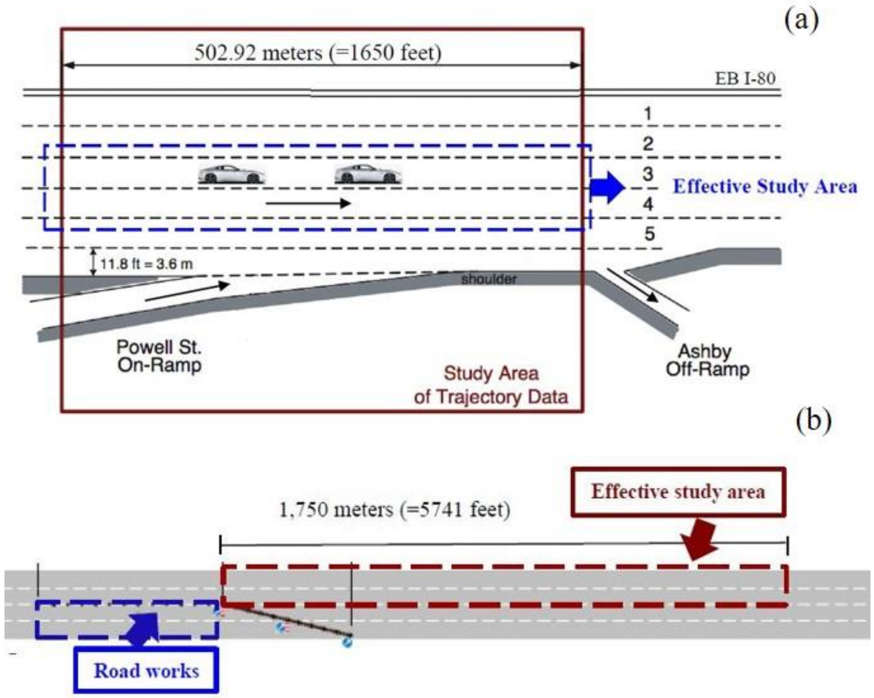

FIGURE 1 (a) I-80 motorway data collection site *(30)*, (b) Schematic of road section used in UoLDS experimental study *(31)*

Driving simulator data

The driving simulator data was collected at the University of Leeds Driving Simulator (UoLDS). The data collection took place in the context of the "Smart Motorway-All Lanes Running" project *(32)*, funded by Highway Agency (UK) *(31)*. The main aim of the project was to investigate drivers' behaviour during motorway driving, under the presence of roadworks (Figure 1b). In particular, participants drove four different scenarios (1:light, low density; 2:light, high density; 3:dark, low density; 4:dark, high density) for approximately 40 minutes each. The road consisted of four lanes and traffic signals along the leftmost lane were warning participants about the roadworks ahead (e.g. lane blockage). In total, 40 drivers (20 females, 20 males) aged from 19 to 83 years old participated in the study. For the present analysis, only the observations from the second scenario (light, high density) were considered.

Data description

The raw datasets were further processed to better meet the requirements for the development of a car-following model. As a first step, relationships regarding the surrounding traffic such as relative speed, acceleration of lead vehicle etc. were extracted in both datasets.

Regarding the I-80 video dataset, only cars that did not attempt lane-changing during the observation period were included in the analysis. A similar approach was also applied for the driving simulator data. However, given that the observation period for that case was longer, each road segment was split in smaller parts approximately 500m long. Only the parts where no lane-changing was detected considered for the model development. Moreover, the data was further processed to account only for the road segments without roadworks.

For both datasets, the considered observation frequency was 1 observation/sec. Also, in order to avoid free-flow observations and following the findings in *(33)*, an upper bound of 4s was applied in the observed time headway; all the values above that threshold, were treated as free-flow and excluded. Moreover, only the observations where the leader was the same for the whole range of reaction time were considered. For the final estimation, the video dataset was composed of 447 individuals and 16314 observations while the driving simulator dataset 40 individuals and 3895 observations. Table 1 summarises the descriptive statistics for some variables of the two datasets.

TABLE 1 Descriptive statistics of the two datasets variables

Variable	I-80 Video data				Driving simulator data				Levene's test for equality of variances	t-test for equality of means
	Min	Mean	Max	sd	Min	Mean	Max	sd	p-value	p-value
Speed (m/s)	1.340	8.420	27.100	3.668	16.800	29.300	37.500	3.227	0.100	0.000
Acceleration (m/s^2)	-3.990	-0.030	4.320	0.960	-3.650	-0.019	1.430	0.297	0.000	0.000
Time headway (s)	0.493	2.430	4.000	0.728	0.022	2.410	4.000	0.858	0.000	0.157
Space headway (m)	4.230	19.600	85.100	9.314	5.090	74.903	146.065	26.566	0.000	0.000
Front speed (m/s)	0.000	8.250	26.200	3.720	14.600	29.100	40.900	4.263	0.000	0.000

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 (Table 1). Regarding the test's results, the p-value for the Leven's test is significant for all variables which indicates that the variances of all the variables are different between the video and the driving simulator datasets. Additionally, the results of the t-test for the equality of means show that, besides time headway, the means of the subject speed, acceleration, leader speed and space headway with the lead vehicle are significantly different. These findings show that there are some differences in the variables (and thus the traffic conditions) between the two datasets which may be influential for the models' results. 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 and fixed variations in traffic whereas the field traffic will have larger variability).

ESTIMATION RESULTS

The estimation results are summarized in Table 2 and explained below.

Individual Models

Model1: Car-following model based on driving simulator data

The signs of acceleration and deceleration constants are both expected, however, the first parameter is not statistically significant at 0.05 level. A similar pattern is also observed regarding the parameters of time headway. More specifically, the positive sign of the time headway parameter for the acceleration regime implies that drivers tend to follow their leader's speed less as time headway increases *(2)*. This parameter is however not statistically significant in the present estimation. Regarding the deceleration regime, the positive sign of the time headway parameter indicates that drivers adopt smaller decelerations at larger headways. Finally, the relative speed parameter is significant for both acceleration and deceleration regimes. It is worth mentioning that although the relative speed parameter is apriori expected to be smaller than 1 *(2)*, because of the limited acceleration or deceleration a driver can apply, larger values are also allowed. The impact of each parameter in acceleration is better illustrated in the next section of model comparison. Figure 2 depicts the reaction time distribution as expressed by the estimated mean and standard deviation. The distribution is centered approximately around 3 seconds while it extends from approximately 2 to 4 seconds.

Model 2: Car-following model based on video trajectory data

The results of the car-following model estimation based on the video data are presented in Table 2. The parameters have all expected signs and are significant at 0.05 level. Moreover, the values of relative speed parameters are below 1, as apriori expected. The reaction time distribution is presented in Figure 2. The estimated mean is lower, compared to Model 1 and the distribution approximately extends between 0.8-1.3s.

Model comparison and sensitivity analysis

In the current section is investigated the effect of models' variables in the car-following acceleration (deceleration). Figure 3 depicts the sensitivity analysis for the various parameters. Focusing on the driving simulator data, the results indicate that the absolute value of acceleration remains at constant levels as time headway increases, while, on the other hand absolute deceleration decreases with the increase of time headway. Moreover, acceleration and deceleration reach their maximum absolute values when relative speed is maximum and minimum, respectively. The observed deceleration patterns are expected, since they indicate drivers' safety concerns; as time headway or relative speed decrease drivers decelerate to avoid collision.

With respect to acceleration and time headway, a different pattern is observed for the video data, compared to the driving simulator. More specifically, acceleration decreases as time headway increases. The potential interpretation of this finding is that as time headway increases, drivers reach their desired speed in a free-flow regime and thus adopt constant speed. The deceleration trend is similar to driving simulator observations. Regarding the effect of relative speed in acceleration, the type of slope is different, compared to the simulator data, however the general pattern is the same, concerning the observed min-max acceleration absolute values occurrence. A similar finding is also noticed for the deceleration case. Deceleration patterns for Model 2 (Video data) imply the same safety concerns as in Model 1 but in this case drivers are more sensitive in the traffic conditions changes and thus higher absolute values are observed. It is worth mentioning that acceleration-time headway and deceleration-relative speed plots of Model 1 produce acceleration almost as a straight line close to zero. Both models behave in a similar way in the deceleration-time headway case where however, the minimum values differ.

The results of the t-test of individual parameter equivalence indicate that for most of the parameter pairs, the null hypothesis of equivalence is rejected. The t-stat of the difference is insignificant for the standard deviation of reaction time distribution, the time headway parameter of acceleration regime and the relative speed parameter of deceleration regime, thus only these parameters can be transferred. The results of the TTS regarding transferability from driving simulator to real driving context show that the null hypothesis of equivalence between the two models is rejected, therefore, transferability cannot be validated.

TABLE 2 Models parameter estimates, t-test of individual parameter equivalence and Transferability Test Statistic (TTS) results

	Driving simulator data		Video data		T-tests of individual parameter equivalence
Variable	Parameter estimate	Robust t-statistic	Parameter estimate	Robust t-statistic	Difference t-stat
Reaction time distribution					
μ_t	1.0898	19.96	0.0445	4.06	18.7705
σ_t	0.1392	0.73	0.059	7.92	0.4203
Car-following acceleration					
constant	0.0005	0.25	0.7434	13.13	-13.1130
time headway (s)	0.1627	0.08	0.5446	6.33	-0.1876
relative speed (m/s)	3.3199	3.03	0.8672	20.98	2.2369
σ^{acc}	0.2481	19.63	0.779	80.51	-33.3536
Car-following deceleration					
constant	-0.1397	-2.74	-0.6066	-15.71	7.3003
time headway (s)	1.5727	4.51	0.471	5.96	3.0812
relative speed (m/s)	0.7614	4.87	0.8103	25.7	-0.3066
σ^{dec}	0.3258	7.82	0.813	75.34	-11.3204

Transferability Test Statistic (TTS)

Summary statistics	Simulator to real driving transferability
Degrees of freedom (Dof)	12
LLapplic (β_{transf})	-73206.12482
LLapplic(β_{appl})	-19461.7394
-2[LLapplic(β_{transf}) -LLapplic(β_{appl})]	107488.7709

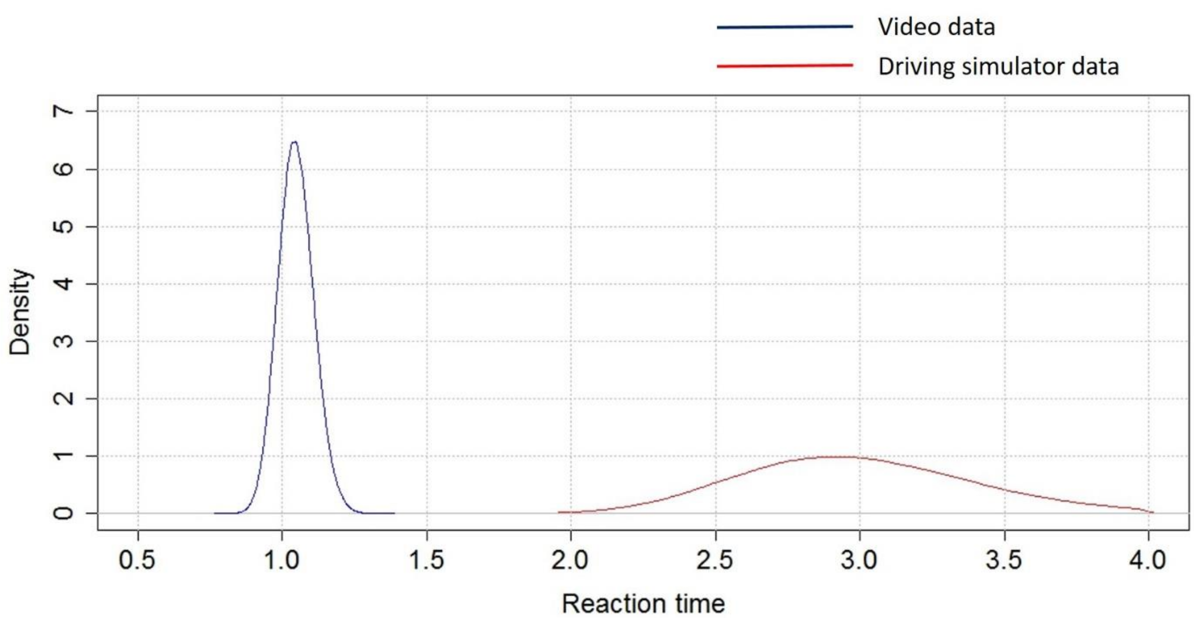

FIGURE 2 Reaction time distributions of the car-following models

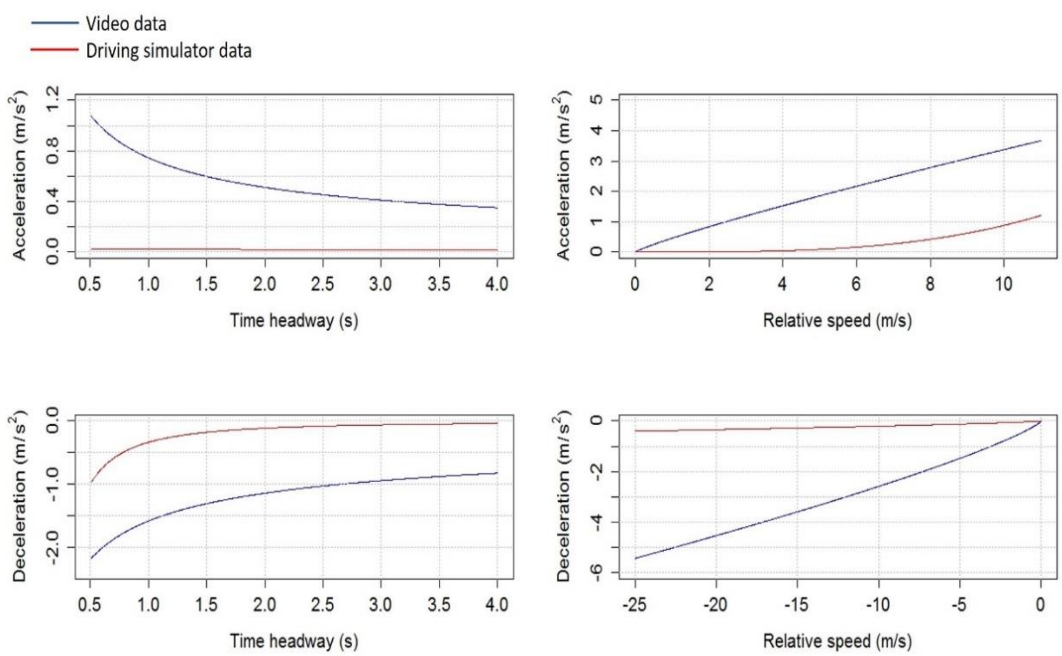

FIGURE 3 Sensitivity analysis of the car-following models

MODEL UPDATING AND JOINT ESTIMATION

The work in the previous section highlighted a 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. The results are compared with the results of a joint model estimated using both datasets.

Model updating

The parameters of the driving simulator model were updated using the Bayesian Updating *(26)* and Combined transfer estimation *(16)* approaches. The results of the TTS after the application of model updating are presented in Table 3. The TTS value after applying Bayesian updating indicates that the null hypothesis of model equivalence is rejected. However, the TTS value of the combined transfer estimation shows that after updating, the null hypothesis cannot be rejected and thus, driving simulator data can be transferable.

TABLE 3 The Transferability Test Statistic results after model updating

Transferability Test Statistic (TTS)		
Summary statistics	Bayesian updating	Combined transfer estimation
Degrees of freedom (Dof)	10	10
LLapplic (βtransf)	-23079.9493	-19449.7262
LLapplic(βapplic)	-19446.7394	-19446.7394
-2[LLapplic(βtransf) -LLapplic(βapplic)]	7267.3290	6.8830

Joint estimation results

For this approach, the car-following models was re-estimated combining simultaneously both data sources. Initially, the datasets were considered as a single source and unique parameters were estimate; more parameters and scales were gradually added. The new models were assessed with the likelihood ratio test; the log-likelihood value of each model was compared with the sum of log-likelihood values of Models 1 and 2 with degrees of freedom equal to the sum of the parameters of the initial models minus the estimated parameters of the joint model. The scale factors were applied individually on specific parameters or the sensitivity×stimulus term in total with the following form: $\delta^{video}+\delta^{simulator}\times$scale, where δ^{video} is a dummy variable equal to 1 if the observation belongs to the video dataset and $\delta^{simulator}$ is a dummy variable equal to 1 if the observation belongs to the driving simulator dataset.

TABLE 4 Parameter estimates of the joint model

Variable	Parameter estimate	Robust t-statistic
Reaction time distribution (Video data)		
μ_t	0.0334	3.11
σ_t	0.0389	4.84
Reaction time distribution (Driving simulator data)		
μ_t	1.1386	28.95
σ_t	0.1121	1.82
Car-following acceleration		
constant	0.7523	12.45
time headway (s)	0.5686	6.28
relative speed (m/s)	0.8765	20.75
σ^{acc} (Video data)	0.7788	80.8
σ^{acc} (Driving simulator data)	0.2494	20.04
Car-following deceleration		
constant	-0.6589	-13.31
time headway (s)	0.5951	5.69
relative speed (m/s)	0.8174	24.66
σ^{acc} (Video data)	0.8129	76.34
σ^{acc} (Driving simulator data)	0.3306	7.86
Scale parameters		
Car-following acceleration mean	0.0545	4.27
Car-following deceleration mean	0.1123	4.57

LL: -19446.284
ρ^2: 0.22
-2LL difference (compared to Models 1 and 2): 40.84

Among the various joint models are presented only the parameter estimates of the model with the best log-likelihood score (Table 4). The selected specification consists of 16 estimated parameters. Different parameters were estimated for the two datasets, regarding the reaction time distribution and the standard deviation of the acceleration (deceleration) density function. Moreover, two scale parameters (for acceleration and deceleration means) were considered. In brief, the best log-likelihood was achieved when different parameters were estimated for each dataset and the sensitivity×stimulus term between the two cases was scaled. It is worth mentioning that the case where both datasets were treated as the same source, produced the worst log-likelihood score (LL = -23434.4), which indicates that for estimation from multiple data sources, it is crucial to account for the differences in scale.

The parameter estimates of the joint model are all significant at the 95% level, including the two scale parameters. This result shows that there is a significant difference in the applied acceleration (deceleration) in the two contexts, as it is expressed through the explanatory variables, that should be considered in simultaneous estimation. The log-likelihood value of the joint model was compared with the sum of the log-likelihood values of the two separate models. The difference (40.84) is larger than the χ^2 critical value for 4 degrees of freedom (9.488) and the null hypothesis is thus rejected. This finding implies that the joint model does not better capture the acceleration decisions, compared to the separate models, although the estimated differences in scale are significant.

DISCUSSION

The current study investigated the development of a car-following model from multiple data sources, focusing on the adequacy of driving simulator data. One of the main motivations of this approach is the potential incorporation of variables captured only by driving simulators within a modelling framework that may strengthen existing specifications. However, it is acknowledged that driving behaviour might be biased in a simulated environment, therefore, this weakness must be minimized to ensure the reliability of the models.

Two main approaches were applied in the present paper to account for the potential bias of driving simulator data, namely, model transferability and joint model estimation. As a first step, two separate models were developed, using driving simulator (Model 1) and video (Model 2) datasets, respectively. The first model was considered as the transferred model while the latter as the application case. Regarding transferability, three different techniques were applied. As a first step, the equivalence of individual parameters and models' equivalence were tested without however identifying transferability. Following literature indications, the parameters of Model 1 were updated, with two different techniques (Bayesian updating and combined transfer estimation). Bayesian updating did not validate model transferability however, the results of combined transfer estimation indicated that driving simulator data can be transferable to real driving context. The second approach of joint model estimation revealed that there is a statistically significant difference in the scale of both acceleration and deceleration values. However, despite the identification of this difference, the joint model did not perform better, compared to Models 1 and 2 separately. This finding should be further investigated, in order to examine the scale differences in more datasets.

The results of the transferability tests and joint estimation suggest that driving simulator data should be used with caution. For instance, the t-tests for individual parameter equivalence showed that not all the parameters are directly transferable, while also the mean reaction time is different in the simulated environment. Moreover, the sensitivity analysis, showed that in real life, drivers are more sensitive in the changes of traffic conditions.

As a limitation of the present study, might be considered the differences of the datasets in two ways, (a) they refer to different countries, (b) the existence of the roadworks parts in the simulator data. Although in the latter case the roadworks parts were removed, they might still have affected drivers' behaviour e.g. drivers might choose to drive more cautiously. Given that these differences pose extra challenges in model transferability or joint estimation, our findings are more on the conservative side. The simulator dataset presented an extreme case (i.e. roadworks) but results might be improved if two more similar settings are compared. This is a case should be further exmined in the future.

This study consists a first step towards the more efficient use of driving simulator data in a driving behaviour modelling framework. The development of an approach that would accommodate for the deficiencies of driving simulator data would also allow in specifications that could benefit from their enriched information regarding drivers' characteristics or the wider variety of scenarios. A simple further application could be e.g. a general acceleration model (2), where reaction time or critical headway between car-following and free-flow state are expressed as a series of socio-demographic characteristics. Moreover, the approach of the current study can be also extended to other driving behaviour models e.g. lane-change. Developing models that better capture driving behaviour, may lead in improved representation of traffic phenomena in microsimulation and at the same time in better predictions regarding e.g. the implications of specific safety measures that are tested on a driving simulator basis, in real driving conditions.

ACKNOWLEDGMENT

The research is supported by the Next Generation Driving Behaviour Model (NG-DBM) project funded by FP7Marie Curie Career Integration Grant of the European Union (PCIG14-GA-2013-631782), the European Research Council through the consolidator grant 615596-DECISIONS and the Economic and Social Research Council, UK.

REFERENCES

1. Toledo, T. Driving behaviour: models and challenges. *Transport Reviews.* Vol. 27, No. 1, 2007, pp. 65-84.
2. Ahmed, K. I. *Modeling drivers' acceleration and lane changing behavior (Doctoral dissertation, Massachusetts Institute of Technology).* 1999.
3. Lee, H.C. The validity of driving simulator to measure on-road driving performance of older drivers. *Transport Engineering in Australia.* Vol. 8, No. 2, 2003, p. 89.
4. UDRIVE. European Naturalistic Driving Study. Available on-line at: http://www.udrive.eu, Assessed Feb. 28, 2016.
5. SHRP2. Strategic Highway Research Program. Available on-line at: http://www.shrp2nds.us, Assessed Mar. 13, 2016.
6. Godley, S. T., T. J. Trigg, and B. N. Fildes. Driving Simulator Validation for Speed Research. *Journal of Accident Analysis & Prevention.* Vol. 34, No. 5, 2002, pp. 589-600.
7. Yan, X., M. Abdel-Aty, E. Radwan, X. Wang, and P. Chilakapati. Validating a driving simulator using surrogate safety measures. *Journal of Accident Analysis & Prevention.* Vol. 40, No. 1, 2008, pp. 274-288.
8. Bella, F., A. Garcia, F. Solves, and M. A. Romero. Driving simulator validation for deceleration lane design. *Proceedings of the 86th Transportation Research Board Annual Meeting.* No. 07-0894, Washington, D.C., 2007.
9. Risto, M., and M.H. Martens. Driver Headway Choice: A Comparison between Driving Simulator and Real-Road Driving. *Transportation Research , Part F: Traffic Psychology and Behavior.* Vol. 25, 2014, pp. 1-9.
10. McGehee, D. V., E. N. Mazzae, and G. S. Baldwin. Driver Reaction Time in Crash Avoidance Research: Validation of a Driving Simulator Study on a Test Track. *In proceedings of the Human Factors and Ergonomics Society Annual Meeting.* Vol. 44, No. 20, 2000, pp. 3-320.

11. Hou, Y., Zhao, Y., Hulme, K. F., Huang, S., Yang, Y., Sadek, A. W., & C., Qiao. An integrated traffic-driving simulation framework: Design, implementation, and validation. *Transportation Research Part C: Emerging Technologies.* Vol. 45, 2014, pp. 138-153.
12. Farah, H., S. Bekhor, A. Polus and T. Toledo. A passing gap acceptance model for two-lane rural highways. *Transportmetrica.* Vol.5, No.3, 2009, pp. 159-172.
13. Papadimitriou, S. and C.F. Choudhury. Transferability of Car-Following Models between Driving Simulator and Field Traffic. *Transportation Research Record (in press).* 2017.
14. Atherton, T.J., and M. E. Ben-Akiva. *Transferability and Updating of Disaggregate Travel Demand Models.* No. 610, 1976.
15. Galbraith, R.A., and D. A. Hensher. Intra-Metropolitan Transferability of Mode Choice Models. *Journal of Transport Economics and Policy.* 1982, pp. 7-29.
16. Ben-Akiva, M., and D. Bolduc. Approaches to model transferability and updating: the combined transfer estimator. *Transportation Research Record.* No. 1139, 1987, pp. 1-7.
17. Koppelman, F. S., and C. G. Wilmot. Transferability analysis of disaggregate choice models. *Transportation Research Record.* No. 895, 1982, pp. 18-24.
18. Wilmot, C. G. Evidence on transferability of trip-generation models. *Journal of Transportation Engineering.* Vol. 121, No. 5, 1995, pp. 405-410.
19. Hadayeghi, A., A. S. Shalaby, B. N. Persaud, and C. Cheung. Temporal transferability and updating of zonal level accident prediction models. *Accident Analysis & Prevention.* Vol. 38, No. 3, 2006, pp. 579-589.
20. Ben-Akiva, M., M. Bradley, T. Morikawa, J. Benjamin, T. Novak, H. Oppewal, and V. Rao,. Combining revealed and stated preferences data. *Marketing Letters.* Vol. 5, No. 4, 1994, pp. 335-349.
21. Gazis, D. C., R. Herman, and R. W. Rothery. Nonlinear follow-the-leader models of traffic flow. *Operations research.* Vol. 9, No. 5, 1961, pp. 545-567.
22. Herman, R., and R. W. Rothery. Car-following and steady state flow. *In Proceedings of the 2nd International Symposium of the Theory of Traffic Flow.* 1965, OECD, Paris, pp. 1-13.
23. Kusuma, A. *Modelling driving behaviour at motorway weaving sections (Doctoral dissertation, University of Leeds).* 2015.
24. Sikder, S., A. R. Pinjari, S. Srinivasan, and R. Nowrouzian. Spatial Transferability of Travel Forecasting Models: A Review and Synthesis. *International Journal of Advances in Engineering Sciences and Applied Mathematics.* Vol. 5, No. 2-3, 2013.
25. Santoso, D. S., and K. Tsunokawa. Transferabiliy 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.* Vol. 6, 2006, pp. 173-185.
26. Dey, S. S., and J. D. Fricker. Bayesian updating of trip generation data: combining national trip generation rates with local data. *Transportation.* Vol. 21, No.4, 1994, pp. 393-403.
27. Halkias, J., and J. Colyar. *Interstate 80 Freeway Dataset. Fact Sheet.* s.l. : U.S. Department of Transportation, NGSIM Program., Federal Highway Administration, 2006.
28. Montanino, M., and V. Punzo. Making NGSIM data usable for studies on traffic flow theory: Multistep method for vehicle trajectory reconstruction. *Transportation Research Record: Journal of the Transportation Research Board.* No. 2390, 2013, pp. 99-111.
29. Punzo, V., M. T. Borzacchiello, and B. Ciuffo. On the assessment of vehicle trajectory data accuracy and application to the Next Generation SIMulation (NGSIM) program data. *Transportation Research Part C: Emerging Technologies.* Vol. 19, No. 6, 2011, pp. 1243-1262.

30. Choudhury, C., V. Ramanujam, and M. E. Ben-Akiva. Modeling Acceleration Decisions for Freeway Merges. *Transportation Research Record: Journal of the Transportation Research Board, Transportation Research Board of the National Academies.* No. 2124, 2009.
31. Agency, Highways. *Smart Motorways.* UK Government. Available on-line at: https://www.gov.uk/government/collections/smart-motorways#history, Accessed Aug. 9, 2015.
32. University of Leeds Driving Simulator, UoLDS. Smart Motorways: All Lanes Running. Faculty of Environment. Institute for Transport Studies. Official Data. Leeds, United Kingdom, 2014.
33. Hoogendoorn, S. P. Unified approach to estimating free speed distributions. *Transportation Research Part B: Methodological.* Vol. 39, No. 8, 2005, pp. 709-727.