Calibration of microscopic traffic-flow models using multiple data sources

BY SERGE HOOGENDOORN* AND RAYMOND HOOGENDOORN

Delft University of Technology, Delft, The Netherlands

Parameter identification of microscopic driving models is a difficult task. This is caused by the fact that parameters—such as reaction time, sensitivity to stimuli, etc.—are generally not directly observable from common traffic data, but also due to the lack of reliable statistical estimation techniques. This contribution puts forward a new approach to identifying parameters of car-following models.

One of the main contributions of this article is that the proposed approach allows for joint estimation of parameters using different data sources, including prior information on parameter values (or the valid range of values). This is achieved by generalizing the maximum-likelihood estimation approach proposed by the authors in previous work.

The approach allows for statistical analysis of the parameter estimates, including the standard error of the parameter estimates and the correlation of the estimates. Using the likelihood-ratio test, models of different complexity (defined by the number of model parameters) can be cross-compared. A nice property of this test is that it takes into account the number of parameters of a model as well as the performance. To illustrate the workings, the approach is applied to two car-following models using vehicle trajectories of a Dutch freeway collected from a helicopter, in combination with data collected with a driving simulator.

Keywords: car-following models; calibration; traffic data

1. Introduction

Microscopic simulation models have become widely applied tools for numerous applications in traffic engineering. Parameter identification remains a difficult task when using these microscopic traffic-flow models. The reasons for this are many: parameters are not directly observable from common (cross-section) traffic data, they are not transferable to other situations (different locations, periods of the day, etc.), real-driving behaviour is variable in time and space, etc. Several approaches to structured model calibration using macroscopic and, in some cases, microscopic traffic data have been proposed, but a good approach based on concepts of statistical theory is still lacking.

*Author for correspondence (s.p.hoogendoorn@tudelft.nl).

One contribution of 10 to a Theme Issue ‘Traffic jams: dynamics and control’.
As examples of automated approaches to model calibration, in the following, the approach developed by Brockfeld et al. (2004), as well as the approach developed by Schultz & Rilett (2005), are considered. Brockfeld et al. (2004) cross-compared different microscopic traffic-flow models using data from car-following experiments. To this end, they established car-following parameter sets minimizing Theil’s U, an error term expressing the difference between prediction and observation. The error rates of the different models in comparison to the data varied between 9 and 24 per cent, for validation between 12 and 30 per cent. No single model appeared to be significantly better than the others (even the more complex ones). It was argued that these errors could probably not be suppressed, irrespective of the model that is used, owing to the different behaviour of each driver, which is in line with the results that will be presented in this contribution.

Schultz & Rilett (2005) proposed a methodology to introduce and calibrate a parameter distribution using measures of central tendency and dispersion (i.e. mean and variance) to generate input parameters for car-following sensitivity factors in microscopic traffic-simulation models. The approach was applied to IH-10 in Houston, TX using the CORSIM model, and subsequently calibrated using an automated genetic algorithm. An overview of calibration approaches as well as parameter values are found in Brackstone & McDonald (1999).

The main contribution presented in this article is the derivation of a statistical parameter estimation approach, enabling statistical analysis of the parameter estimates and cross-comparison of models of different model complexity, respecting the well-known parsimony principle. The approach allows combining multiple sources of data, thereby improving the estimation results when the information in individual trajectories is limited (Ossen & Hoogendoorn 2008). Possible data sources are:

— real-life vehicle trajectories (e.g. Hoogendoorn et al. 2003),
— vehicle-trajectory data collected from a driving simulator (e.g. Hoogendoorn et al. 2010),
— more traditional sources of data, such as cross-section data collected by inductive loops, and
— prior information on the model parameters or on combinations of model parameters.

The proposed approach allows for statistical analysis of the parameter estimates, including the standard error of the parameter estimates and the correlation of the estimates. Also, we can easily test whether a specific model outperforms the other models using the likelihood-ratio test. An important property of this test is that it takes into account the number of parameters of a model as well as the performance. Furthermore, it can deal with the serial correlation in the trajectory data.

In §§2 and 3, we will recall the generic description of car-following models and data-collection approaches from our earlier papers. Section 4 reviews the maximum-likelihood estimation (MLE) first proposed in Hoogendoorn & Ossen (2005). Section 5 presents the generalization of this approach, constituting the main contribution of the presented work. Section 6 shows the application results when the generalized approach is applied to fuse two types of data. Section 7 describes the main conclusions of our work.
2. Modelling car-following behaviour

A microscopic model provides a description of the movements of individual vehicle–driver combinations. These movements are the result of the characteristics of drivers and their vehicles, the interactions between drivers, and between the driver and road characteristics, external conditions and the traffic regulations and control. In general, two types of driver tasks are distinguished:

— longitudinal tasks (acceleration, maintaining speed, distance-keeping relative to leading vehicles), and
— lateral tasks (lane changing, overtaking).

This contribution considers the former, although it would be relatively easy to extend it to lateral driving tasks.

The term car-following model is used here for the general class of dynamic microscopic models describing the longitudinal behaviour of a driver in relation to the lead vehicle(s). Driver \( i \), following driver \( (i - 1) \), may for instance react on (changes in) the spacing \( s_i = x_{i-1} - x_i \) between the vehicles, or his or her relative speed \( \Delta v_i = v_{i-1} - v_i \). Many models have been proposed to describe this longitudinal behaviour. It is beyond the scope of this contribution to provide a comprehensive overview of all models. Rather, in this contribution, only a small selection of models will be considered, which are described in the remainder of the section. The proposed parameter-identification approach is, however, generally applicable, given that sufficient data are available, and that these data contain sufficient information to estimate all parameters in the model (Ossen & Hoogendoorn 2008).

(a) General model of car following

Before presenting two examples of well-known car-following models, let us present the general mathematical model used in the remainder of this article. To this end, we distinguish between the driver state \( z_i \) and the perceived traffic state \( x_i \) on which driver \( i \) reacts. The former generally describes the physical description of the driver (or rather vehicle) state, such as its position and speed. The latter represents some mapping of the actual traffic state (e.g. position and speeds of all vehicles in the system) that explains the behaviour of driver \( i \).

Let \( z_i(t) \) denote the (time-dependent) state of driver \( i \). As it will turn out in the remainder of the text, the specification of the state will depend on the car-following model that is considered. In most cases, the state will consist of the position \( x_i(t) \) and the speed \( v_i(t) \) of the driver. Let the vector \( \xi_i(t) \) denote the time-dependent traffic state on which driver \( i \) may react or anticipate. This vector may, for example, contain (estimates of) positions, speeds and accelerations of driver–vehicle combinations \( j \) in the vicinity of driver \( i \). It always contains the state \( z_i(t) \) of driver \( i \) himself. For most models, the traffic state can be defined completely by

\[
\xi_i(t) = (z_1(t), \ldots, z_i(t), \ldots, z_n(t)),
\]

where \( n \) is the number of vehicles in the system.
Most car-following models can then be described in either of the following generic forms:

\[
\frac{dx_i(t)}{dt} = f(x_i(t), x_i(t - \tau_i) | \theta_i) \tag{2.2}
\]

or

\[
\frac{dx_i(t)}{dt} = f_1(x_i(t)) = v_i(t) \tag{2.3}
\]

and

\[
\frac{dv_i(t)}{dt} = f_2(x_i(t), x_i(t - \tau_i) | \theta_i), \tag{2.4}
\]

where \( \theta_i \) denotes the set of parameters describing the car-following behaviour of driver \( i \), including the (explicit) reaction time \( \tau_i \). Furthermore, \( x_i \) denotes the position of the front bumper of the vehicle \( i \) corresponding to the collected data, as shown in figure 1.

More generally, we can define the state \( z_i(t) \) of driver \( i \) as \( z_i(t) = x_i(t) \) or as \( z_i(t) = (x_i(t), v_i(t)) \) and write

\[
\frac{dz_i(t)}{dt} = f(x_i(t), x_i(t - \tau_i) | \theta_i). \tag{2.5}
\]

Note that we assume that all drivers \( i \) within a population can be described by the same model structure, but may be different in terms of the parameters \( \theta_i \) describing their behaviour. In other words, we will assume that the behaviour of drivers is described using different model parameters that follow some (unknown) distribution.
The estimation approach presented in the contribution entails comparing the predictions of the car-following model at discrete time instants $t_k$, with $k = 1, \ldots, K$, with the available trajectory data. This means that equation (2.5) is integrated given the initial conditions, i.e.

$$z_i(t_{k+1}) = z_i(t_k) + \int_{t_k}^{t_{k+1}} f(\xi_i(s), \xi_i(s - \tau_i) | \theta_i) \, ds$$

$$= F(\xi_i(t_k), \xi_i(t_k - \tau_i)).$$

(b) Example I: model of Gazis, Herman and Rothery

One of the most simple car-following models proposed in the literature is the model of Gazis, Herman and Rothery (GHR; Gazis et al. 1961). This model expresses the acceleration of the follower ($d/dt v_i(t)$) as a delayed response to the relative speed $\Delta v_i(t)$ with respect to the predecessor (stimulus),

$$\frac{d}{dt} v_i(t) = a \Delta v_i(t - \tau),$$

where $\Delta v_i = v_{i-1} - v_i$. The model has only two parameters:

— the sensitivity $a > 0$ describing the reaction strength to the stimulus $\Delta v_i$

and

— the reaction time $\tau \geq 0$ describing the delay in the response.

For the GHR model, the vector $\theta_i = \{a, \tau\}$ will be used to indicate the parameters that will be identified from the data.

Equation (2.8) shows that when the relative speed $\Delta v_i$ is less than zero (follower $i$ driving faster than the leader $i - 1$), the follower will decelerate and vice versa. This will not occur immediately: the response is delayed with the reaction time $\tau$.

Although the model has been criticized by many, the simplicity of the model has resulted in the model’s popularity. For instance, the model can be used to apply stability analysis showing under which conditions—i.e. combinations for parameter values for $a$ and $\tau$—the model is either asymptotically stable (meaning that a disturbance damps out as it travels from the subsequent vehicles in the platoon), locally stable (reaction to change in leader behaviour damps out as time progresses) or unstable (see Gazis et al. 1961). Note that asymptotic (or string) stability is a stronger requirement than local (or controller/exponential) stability: the model is locally stable if $a \tau < \pi/2$, while it is asymptotically stable when $a \tau < 1/2$.

It is beyond the scope of this article to describe all properties of the GHR model.

(c) Example II: intelligent-driver model

The second example that we will consider is the intelligent-driver model (IDM; Treiber et al. 2006), which describes the acceleration of driver $i$ as a function of
the distance $s_i(t)$, speed $v_i(t)$ and the relative speed $\Delta v_i(t)$ using the following expression:

$$\frac{d}{dt} v_i = a \left( 1 - \left( \frac{v_i}{v_0} \right)^4 - \left( \frac{s^*(v_i, \Delta v_i)}{s_i - l} \right)^2 \right), \quad (2.9)$$

where the desired gap $s^*$ is given by

$$s^*(v_i, \Delta v_i) = s_0 + Tv_i - \frac{v_i \Delta v_i}{2\sqrt{ab}}, \quad (2.10)$$

and where $l$ denotes the length of vehicle $i$. The five parameters to be estimated are:

- the maximum acceleration $a$,
- the maximum deceleration $b$,
- the free speed $v_0$,
- the minimum time headway $T$, and
- the net stopping distance $s_0$.

For the IDM, the vector $\theta_i = \{a, b, v_0, s_0, T\}$ is used to indicate the parameters describing the behaviour of driver $i$.

The properties of the IDM have been well researched. We refer to Treiber et al. (2006) for an overview of the main results.

Note that, although in this article, we are focusing on parameter estimation of car-following models, the method that is proposed in the ensuing will be generally applicable to many other microscopic model-identification problems, such as the calibration of microscopic pedestrian-flow models (Hoogendoorn et al. 2006).

### 3. Microscopic traffic data

In this contribution, two types of vehicle-trajectory data will be used to estimate the parameters of car-following models: empirical vehicle-trajectory data and experimental trajectory data collected using a driving simulator.

(a) Remote-sensing vehicle-trajectory data

The empirical vehicle-trajectory data used here were collected using a remote-sensing data-collection approach (Hoogendoorn et al. 2003) using an air-borne observation platform (a helicopter), mounted with a high-frequency digital camera and frame grabber. Using image-processing software, the vehicles are detected and tracked as they move along the roadway. This yields trajectory data covering approximately 500 m of freeway stretch; the spatial resolution is smaller than 40 cm, while the temporal resolution is 0.1 s. Besides the trajectories of all vehicles present, the system also determines the lengths and widths of the vehicles. Vehicles driving in both roadway directions were detected and tracked. Only one driving direction is considered here. The dataset was collected at the 1Note that due to notational differences with the original work of Treiber et al. (2006), the definition of the desired gap was modified.

Phil. Trans. R. Soc. A (2010)
Multiple data-source calibration

multi-lane A2 freeway near the Dutch city of Utrecht, and is characterized by congested flow conditions. Figure 1 shows a sample from the dataset. Note that crossing trajectories indicate lane changes.

(b) Driving-simulator data

The non-moving base Advanced Driving Simulator consists of three screens placed at an angle of 120°, a driver’s seat mockup and also hardware and software interfacing of a central computer system to this mockup. The view from the driver’s seat consists of a projection of 210° horizontally and 45° vertically. The software used was developed by STSoftware. The driving environment was developed using STRoadDesign, while the scenario-determining behaviour of other vehicles was written in STScenario.

Data regarding acceleration, deceleration, speed, distance to the lead vehicle and speed of the lead vehicle were collected through STDataProc at a sampling rate of 10 samples per second and saved to a text file. The driving environment consisted of a long freeway stretch, with three lanes in either direction. The other vehicles consisted of cars as well as trucks, with a maximum speed of 120 km h⁻¹. No adverse weather conditions were present. For details about the experimental set-up, see Hoogendoorn et al. (2010).

Since a driving simulator was used in the estimations, some remarks regarding the validity of driving simulators in general are in order. Recent research performed by Yan et al. (2008), investigating differences between behaviour in the driving simulator and in real life at intersections, shows that driving simulators possess relative validity. This means that drivers react to stimuli in the same direction as in real life. Research regarding the validity of driving simulators with regard to speed was conducted by Godley et al. (2002). They also concluded that, although participants generally drove faster in real life than in the driving simulator, the driving simulator possesses relative validity. Finally, Reimer et al. (2006) conducted a validation study on driving simulators using self-reports. Significant relationships were found with regard to accidents, speeding, velocity, overtaking and behaviour at traffic signs. The researchers concluded that driving simulators are adequate instruments to measure driving behaviour.

4. Maximum-likelihood estimation of car-following models

In this section, we propose the new parameter-estimation approach, for which we recall the MLE approach first presented in Hoogendoorn & Ossen (2005). We assume that for each driver \( i \) on the considered freeway, we have an observed trajectory (observed position \( x_i \) at time instants \( t_k \)) and (consequently) observed speeds \( v_i \) (e.g. collected via the approach described briefly in the previous section). Note that since we assume that the leading vehicles are also observed, we can easily determine the different stimuli present in car-following models (distances, relative speeds, etc.). Let us assume that we can do so sufficiently accurately.

The objective is to estimate the parameters \( \theta_i \) of a car-following model describing the behaviour of driver \( i \) using these data.
Maximum-likelihood estimation approach for a single trajectory

Let us briefly recall and generalize the approach first outlined in Hoogendoorn & Ossen (2005). Let $y_i(t_k)$ denote the observed state of vehicle $i$ of a driver at instants $t_k = kh$, where $h$ is the observation time step. The difference between the observed state and the predicted state is equal to

$$e_i(t_k) = y_i(t_k) - z_i(t_k).$$

(4.1)

Note that since the driver states $z_i(t_k)$ are functions of the parameters $\theta_i$ of the car-following model, so are the residuals: $e_i(t_k) = e_i(t_k|\theta_i)$.

Let $p_K(e_i(t_k), \ldots, e_i(t_k), e_i(t_K))$ denote the joint probability density function of the residuals. The MLE is defined as follows:

$$\hat{\theta} = \text{arg max } L(\theta),$$

(4.2)

with the likelihood $L$ defined by

$$L(\theta) = p_K(e_i(t_k), \ldots, e_i(t_k), e_i(t_K)),$$

(4.3)

where $p_K$ denotes the joint probability density function of the residual vector (for all time periods $k = 1, \ldots, K$).

If we assume that the residuals $e_i(t_k)$ are serially independent—which obviously is not the case here—we write the expression for the likelihood $L$ as follows:

$$L(\theta) = \prod_{k=1}^{K} p(e_i(t_k)),$$

(4.4)

where $p$ denotes the probability density function of the residuals $e_i(t_k)$. Note that the issue of serial correlation will be considered explicitly in the remainder of the article. Furthermore, recall that in practical applications, it is often easier to maximize the logarithm of the likelihood,

$$\tilde{L}(\theta) = \ln L(\theta) = \sum_{k=1}^{K} \ln p(e_i(t_k)).$$

(4.5)

Now, let us furthermore assume that the residuals $e_i$ follow the zero-mean multi-variate normal distribution with covariance matrix $\Sigma$. We then get

$$\ln p(e) = \ln \left( (2\pi)^{-N/2} \det(\Sigma)^{-1/2} \exp \left( -\frac{1}{2} e^T \Sigma^{-1} e \right) \right)$$

(4.6)

and

$$\ln p(e) = -\frac{N}{2} \ln(2\pi) - \frac{1}{2} \ln(\det(\Sigma)) - \frac{1}{2} e^T \Sigma^{-1} e,$$

(4.7)

where $N$ is the number of elements in the state $x$, thereby yielding

$$\tilde{L}(\theta, \Sigma) = -\frac{NK}{2} \ln(2\pi) - \frac{K}{2} \ln(\det(\Sigma)) - \sum_{k=1}^{K} e_i^T(t_k|\theta) \Sigma^{-1} e_i(t_k|\theta).$$

(4.8)
It can be shown easily that the MLE of the covariance matrix $\Sigma$ is equal to
\[
\hat{\Sigma} = \hat{\Sigma}(\boldsymbol{\theta}) = \frac{1}{K} \sum_{k'=1}^{K} e_i(t_{k'}|\boldsymbol{\theta}) e_i^T(t_{k'}|\boldsymbol{\theta}),
\]
which allows us to remove the dependency of $\Sigma$ in equation (4.8). The MLE of the parameters $\boldsymbol{\theta}$ can now be determined by maximizing
\[
L^*(\boldsymbol{\theta}) = \tilde{L}(\boldsymbol{\theta}, \hat{\Sigma}(\boldsymbol{\theta})).
\]

(b) Dealing with serial correlation

The main assumptions which have been made in the derivation of equation (4.10) are the fact that the residuals are uncorrelated and that they follow the zero-mean multi-variate normal distribution.

Serial correlation or autocorrelation entails that the subsequent error terms are not independent. Effectively, this results in reduced information richness of the data. Serial correlation occurs when the covariance between the subsequent errors are larger than zero, i.e.
\[
\rho = \text{cov}\{e_i(t_0), e_i(t_1), \ldots, e_i(t_{K-1})\}, \{e_i(t_1), e_i(t_2), \ldots, e_i(t_K)\}.
\]

To determine the extent in which serial correlation plays a role, we propose the following three-step approach:

— obtain the parameter estimates $\hat{\boldsymbol{\theta}}_i$ by optimization of the log-likelihood equation (4.10),
— determine the residuals $e_i(t_k)$, and
— determine $\rho$ directly by equation (4.11).

The Durbin–Watson test (Durbin 1970) can then be applied to test if the estimate of the autocorrelation coefficient significantly differs from zero. If the correlation coefficient $\rho \neq 0$, we need to transform the model to determine correct values for the parameter variance estimates. Without going into detail, we recall that for linear models, the following transformation eliminates the autocorrelation (see Cochrane & Orcutt 1949):
\[
\tilde{z}_i(t_{k+1}) = z_i(t_{k+1}) - \rho z_i(t_k).
\]

Let us define the observations $\tilde{y}_i$ by
\[
\tilde{y}_i(t_{k+1}) = y_i(t_{k+1}) - \rho y_i(t_k).
\]

Let us now define the residuals
\[
\tilde{e}_i(t_k) = \tilde{y}_i(t_k) - \tilde{z}_i(t_k).
\]

We can again use the MLE to determine the parameter estimates $\hat{\boldsymbol{\theta}}_i$ of the transformed model. In general, the estimates will be the same as in the case of the non-transformed model. The error term will however be different, enabling the correct statistical analysis of the model. More specifically, the autocorrelation between the subsequent time periods is now removed.

Phil. Trans. R. Soc. A (2010)
(c) Statistical properties of parameter estimates

To approximate the covariance matrix of the estimated parameters, we can use the so-called Cramer–Rao lower bound (Cramer 1946), stating that

$$\text{var}(\hat{\theta}) \geq -E(\nabla^2 L^*) \quad (4.15)$$

For the MLE estimators, we can actually show that the asymptotic variance of the parameters is given by the right-hand side of equation (4.15).

(d) Likelihood-ratio test

One of the important applications of the MLE is that models of different complexity (i.e. with a different number of parameters) can be cross-compared, taking into account the fact that a simple model is generally preferred over a complex model. To this end, the so-called likelihood-ratio test can be performed.

Suppose that we have two car-following models and that the number of parameters of the models are equal to $m_1$ and $m_2$. The likelihood-ratio test involves testing the statistic

$$2(L^*_1(\hat{\theta}_1) - L^*_2(\hat{\theta}_2)), \quad (4.16)$$

which follows the $\chi^2$ distribution with $d$ degrees of freedom, where $d = m_1 - m_2$ denotes the number of parameters of the model. The likelihood-ratio test is passed with 95% CI if

$$2(L^*_1(\hat{\theta}_1) - L^*_2(\hat{\theta}_2)) > \chi^2(0.95, d), \quad (4.17)$$

meaning that model 1 outperforms model 2 significantly at 95% CI.

5. Generalizations of the maximum-likelihood estimation approach

In Hoogendoorn & Ossen (2005), the approach is successfully applied using the difference between the predicted speeds and the observed speeds as the residuals. However, it turns out that in some cases, some of the parameters cannot be determined directly from the data. The main reason for this is the fact that the trajectory data considered in Hoogendoorn & Ossen (2005) only contain sufficient ‘information’ to identify a relatively small number of parameters. For example, the estimates for the free speed $v_0$ are impaired by the fact that many drivers are car following during the entire observation period.

For the sake of illustration, the MLE approach was applied to the empirical trajectory data described earlier. Figure 2 shows the estimation results for one specific driver. The asterisks indicate the estimated parameters (i.e. $a$, $b$, $s_0$, $T$ and $v_0$). Based on the values of the log-likelihood, the IDM seems to perform quite well (in comparison to other models; see Hoogendoorn et al. 2006). However, convergence problems occur during estimation, in particular because the log-likelihood is insensitive to changes in some of the parameters (in particular $v_0$ and to a lesser extent to $b$). Also, the parameter estimates are not always realistic (again, in particular, the maximum deceleration value $b$).
Figure 2. (a)–(e) Plots of the log-likelihood values around the optimal estimate $\theta_1^*$ of the parameters of the IDM, for example driver 1. (a) Maximum acceleration, (b) maximum deceleration, (c) stopping distance, (d) minimum time headway and (e) free speed.

(a) Maximum-likelihood estimation approach for multiple trajectories

In the preceding section, we discussed how parameter estimates for a single trajectory could be established using a MLE technique. One way to deal with the aforementioned lack of information is to include prior information, as described in Hoogendoorn et al. (2007). In this subsection, we propose using data from different drivers, either from real-life measurements or from driving simulators.

Let us assume that we have a set of $N$ trajectories for an equal number of drivers $i = 1, \ldots, N$. The driving behaviour of each of these drivers can be described by a driver-specific set of parameters $\theta_i$. Given the observed trajectory data, we can determine the likelihood $L^{(i)}(\theta_i)$ of observing this trajectory given that the car following is described by the set of parameters $\theta_i$ using equation (4.3).

In combining the trajectories, we can easily determine the joint likelihood of observing the trajectories of drivers $i = 1, \ldots, N$,

$$L_{\text{mult}}(\theta_1, \ldots, \theta_N) = \prod_{i=1}^{N} L^{(i)}(\theta_i).$$

(5.1)
Using this expression, we can also determine the likelihood that all drivers in the considered driver population can be described by the same parameter set $\overline{q}$,

$$L_{\text{mult}}(\overline{\theta}) = \prod_{i=1}^{N} L^{(i)}(\overline{\theta}).$$

(5.2)

From equation (5.2), the log-likelihood can be easily determined,

$$\tilde{L}_{\text{mult}}(\overline{\theta}) = \sum_{i=1}^{N} \tilde{L}^{(i)}(\overline{\theta}) = \sum_{i=1}^{N} \ln(L^{(i)}).$$

(5.3)

If we assume that the parameters $\theta_i$ describing the behaviour of individual drivers are drawn from some random distribution with mean $\Theta$ and covariance matrix $\Sigma$, we know that

$$E(\overline{\theta}) = \Theta$$

and

$$\text{var}(\overline{\theta}) = \frac{1}{\sqrt{N}} \Sigma.$$  

(5.4) (5.5)

This expression allows us to construct the mean and covariance of the distribution underlying the individual parameters from the joint-estimation results. This means that although multiple trajectories are considered jointly, we are still able to make inferences regarding population heterogeneity.

(b) Maximum-likelihood estimation approach for multiple data sources

In line with the generalization presented in the previous section, we can extend the estimation approach to include other data sources as well. For instance, combining data collected using the driving simulator with the data collected by remote sensing can be achieved by considering the joint likelihood

$$L_{\text{mult}}^{\text{comb}}(\theta) = L_{\text{emp}}^{\text{mult}}(\theta) L_{\text{sim}}^{\text{mult}}(\theta).$$

(5.6)

Note that the log-likelihood becomes

$$\tilde{L}_{\text{mult}}^{\text{comb}}(\theta) = \tilde{L}_{\text{emp}}^{\text{mult}}(\theta) + \tilde{L}_{\text{sim}}^{\text{mult}}(\theta),$$

(5.7)

showing well how the likelihoods stemming from a specific data source will affect the combined likelihood. Note that several factors will influence the properties of the estimation results.

The difference in the number of observations (the number of observed trajectories and their average length) will determine if the combined likelihood yields a bias to either the empirical trajectory data or the data from the simulator. To reduce this bias, we propose to optimize the following average likelihood:

$$\tilde{L}_{\text{mult}}^{\text{comb}}(\theta) = \frac{1}{K_{\text{emp}}} \tilde{L}_{\text{emp}}^{\text{mult}}(\theta) + \frac{1}{K_{\text{sim}}} \tilde{L}_{\text{sim}}^{\text{mult}}(\theta),$$

(5.8)

where $K_{\text{emp}}$ and $K_{\text{sim}}$ denote the total number of data points in the respective samples. In doing so, the estimation results will not be biased towards the largest sample.
It is well known that validity increases when multiple data sources are used (e.g. Mathiso 1988). Validity of findings are strongly related to the quality of the data used (especially with regard to the extent in which they are an accurate reproduction of the phenomena to be studied), the information contained in the data as well as the adequacy and appropriateness of the methodology used. In this regard, a more triangulated approach\(^2\) may be used.

Triangulation may refer to data triangulation, methodological triangulation as well as theoretical triangulation. In data triangulation, the same phenomenon is studied using multiple data sources in order to determine the extent to which the data used as well as the methodology used to analyse the data are valid. In answering the research question posed in this paper, remote-sensing vehicle-trajectory data were combined with data collected through the Advanced Driving Simulator at Delft University of Technology, Faculty of Civil Engineering and Geosciences, Transport and Planning. However, for a more data-triangulated approach to be appropriate and effective, data derived from the multiple sources have to be comparable. Therefore, use was made of data regarding relatively similar traffic-flow conditions. Nevertheless, care should be taken when combining data from different data sources. The data needs to reflect behaviour for comparable circumstances or rather environments (e.g. similar traffic regulations, road, ambient and weather conditions).

\(^{(c)}\) Including prior information

In addition to using multiple trajectories, prior information can be easily included in the estimation approach (see Hoogendoorn et al. 2007). Prior information means the assumption of the parameter ranges. Let us briefly recall the main results from the aforementioned paper. It is assumed that the parameters \(\theta\) stem from some probability distribution \(G(\theta)\) that describes the inter-driver differences in car-following behaviour. Let the following equation describe the joint probability density function of the parameters:

\[
g(\theta) = g(\theta | \mu, \Theta) . \quad (5.9)
\]

Using this information, we can then extend the log-likelihood function as follows:

\[
L^*_1(\theta) := K \ln g(\theta | \mu, \Theta) + L^*(\theta) , \quad (5.10)
\]

where \(\mu\) is the prior mean and \(\Theta\) is the prior covariance matrix of the random parameters \(\theta\). Let us note that if the standard deviation for a particular parameter is very high, this essentially means that the prior information will effectively not be used in the estimation procedure. To show this, let us assume that \(g\) is a multi-variate normal distribution. We can then easily show that for the gradient of the log-likelihood to the parameter vector, we obtain

\[
\frac{\partial (\ln g(\theta))}{\partial \theta} = \Theta^{-1} (\theta - \mu) . \quad (5.11)
\]

\(^2\)Triangulation is the combination of multiple data-collection methods, often used in qualitative research and based on ‘triangulating’ in geosciences, in order to increase the validity of research results.

Phil. Trans. R. Soc. A (2010)
For the unconstrained optimization problem, we have the necessary conditions

\[
\frac{\partial L^*_1}{\partial \theta} = 0 \quad \Rightarrow \quad \Theta \frac{\partial \tilde{L}_1}{\partial \theta} + K(\theta - \mu) = 0.
\] (5.12)

Equation (5.12) shows well how the prior information drives the estimate towards the prior mean value \(\mu\). The extent to which this happens is determined directly by the values of the covariance matrix \(\Theta\): the larger the elements, the smaller the influence of the corresponding prior estimate.

We emphasize that, in principle, any kind of joint distribution can be used for \(g\) (uniform, normal, log-normal, etc.). These include distributions that describe prior knowledge of the impact of parameter values of the model characteristics. In other words, the definition of \(g\) allows for guaranteeing certain properties of the model (such as stability of the car-following model, or the resulting fundamental diagram; see Hoogendoorn et al. 2007). If we consider, for instance, the model of Gazis et al. (1961) described by equation (2.8), we know that asymptotic stability is guaranteed if the product of the sensitivity \(a\) and the reaction time \(t\) satisfies \(c = at < 1/2\). We can then define \(g\) such that this requirement is met,

\[
g(c) = g(a \cdot \tau) = \begin{cases} 
2, & 0 \leq c \leq \frac{1}{2}, \\
0, & \text{elsewhere}.
\end{cases}
\] (5.13)

For other models, such as the IDM, we can, in the same way, add information to the estimation process to ensure that the resulting model meets certain properties.

Note that prior information can stem from many sources, including alternative data sources. In Hoogendoorn et al. (2007), a method to derive prior information from the individual trajectories is described; furthermore, the authors show application of the approach including prior information.

6. Case study

In this section, we show the results of applying the multiple trajectory estimation approach. Note that the results of using prior information are shown in Hoogendoorn et al. (2007). We aim to address the following research questions:

- How does the number of trajectories affect the parameter estimates?
- How does fusing different kinds of data impact parameter estimates?
- How does the performance of the simple GHR model compare with the more complex IDM?

In the remainder, these questions will be answered.

(a) Number of trajectories and sample size

To assess the functioning of the proposed estimation approach, we have used the dataset described previously. For showing the impact of the size of \(N\), we have considered the scenarios \(N = 1, 10, 25\) and 100. For each scenario, \(N\) trajectories...
were drawn randomly from the set of available trajectories and the log-likelihood (5.3) was optimized to acquire the estimate $\hat{\theta}_N$. This was repeated 25 times to get insight into the distribution of $\hat{\theta}_N$ for different values of $N$.

Let us first consider the results of the MLE for the standard approach using a single vehicle trajectory. Figure 3 shows the results of the estimation procedure. The figure clearly reveals that many estimates are unrealistic in terms of their magnitude. For example, more than 30 per cent of the estimates of the maximum deceleration $b$ are equal to zero, while over 10 per cent are larger than 10 m s$^{-2}$. A large share of the minimum headways $T$ are smaller than zero as well. The main cause for these mediocre estimation results is the fact that only few single trajectories contain sufficient information to allow estimation of all parameters.

Figure 4 shows the results for $N = 10$. Descriptive statistics are presented in table 1. It turns out that the approach is successful in establishing plausible estimation results for all parameters, except perhaps for the free speed for which over 20 per cent of the estimates are larger than 60 m s$^{-1}$. The main cause for this is the fact that the likelihood is very insensitive to the free speed owing to the fact that most drivers in the dataset are car following. Also, one of the estimates for the minimum headway $T$ is negative, which is clearly not a realistic result.
Figure 4. Distribution of parameter estimates for $N = 10$. (a) Maximum acceleration, (b) maximum deceleration, (c) stopping distance, (d) minimum headway and (e) free speed.

Table 1. Overview of estimation results for different $N$.

<table>
<thead>
<tr>
<th>parameter</th>
<th>$N = 1$</th>
<th>$N = 10$</th>
<th>$N = 25$</th>
<th>$N = 100$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$ (m s$^{-2}$)</td>
<td>1.52</td>
<td>1.54</td>
<td>1.40</td>
<td>1.43</td>
</tr>
<tr>
<td>$b$ (m s$^{-2}$)</td>
<td>2.55</td>
<td>0.79</td>
<td>0.95</td>
<td>0.78</td>
</tr>
<tr>
<td>$s_0$ (m)</td>
<td>14.53</td>
<td>13.06</td>
<td>13.57</td>
<td>12.47</td>
</tr>
<tr>
<td>$T$ (s)</td>
<td>1.29</td>
<td>1.11</td>
<td>1.08</td>
<td>1.17</td>
</tr>
<tr>
<td>$v_0$ (m s$^{-1}$)</td>
<td>37.1</td>
<td>37.9</td>
<td>33.0</td>
<td>32.9</td>
</tr>
</tbody>
</table>

Increasing $N$ has a positive effect on the estimation results. Figure 5 shows the results for $N = 25$. In this case, all parameter estimates are very reasonable, including the free speed $v_0$ and the minimum headway $T$. The main cause for this is the fact that the inclusion of more drivers increases the probability that one (or more) driver(s) is driving freely—at least part of the time—or can accelerate towards the free speed, rendering the likelihood more sensitive to changes in the free speed $v_0$.  

*Phil. Trans. R. Soc. A* (2010)
Figure 5. Distribution of parameter estimates for $N = 25$. (a) Maximum acceleration, (b) maximum deceleration, (c) stopping distance, (d) minimum headway and (e) free speed.

To show the differences with using a single parameter, figure 6 shows the value of the maximum log-likelihood function $\tilde{L}_{\text{mult}}(\theta)$ around the estimated parameters for $N = 100$. Clearly, for most parameters, there is a (clear) global optimum that can be easily identified.

Finally, table 2 shows the correlation between the parameters. The table shows that the correlations are generally very high, further explaining the complexity of the estimation problem: when changing the one parameter, the performance of the model compared with the observed data can be maintained by changing another parameter.

(b) Estimation results for the Gazis, Herman and Rothery model and the intelligent driver model

Let us start by considering the estimation results for the GHR model (Gazis et al. 1961). We first consider the estimation results using a randomly selected subsample of $N = 50$. This random selection is repeated 25 times to gain insight into the mean and variability of the estimates. Table 3 shows the estimation results for the empirical data, the data from the driving simulator and the combination of the data sources. The table also shows the log-likelihood values for the respective estimation experiments.
Figure 6. Sensitivity of the log-likelihood function for $N = 100$. (a) Acceleration, (b) deceleration, (c) minimum gross distance, (d) minimum time headway and (e) free speed.

Table 2. Correlation between parameter estimates for $N = 25$.

<table>
<thead>
<tr>
<th></th>
<th>$a$</th>
<th>$b$</th>
<th>$s_0$</th>
<th>$T$</th>
<th>$v_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$</td>
<td>1.0000</td>
<td>0.7025</td>
<td>0.7809</td>
<td>−0.8474</td>
<td>−0.8349</td>
</tr>
<tr>
<td>$b$</td>
<td>1.0000</td>
<td>0.4422</td>
<td>−0.6669</td>
<td>−0.5905</td>
<td></td>
</tr>
<tr>
<td>$s_0$</td>
<td>1.0000</td>
<td>−0.8711</td>
<td>−0.7894</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$T$</td>
<td>1.0000</td>
<td>0.9044</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$v_0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Table 3. Overview of data fusion estimation results for the GHR model.

<table>
<thead>
<tr>
<th>parameter</th>
<th>empirical data</th>
<th>simulator data</th>
<th>mixed data</th>
</tr>
</thead>
<tbody>
<tr>
<td>log-likelihood ($\times 10^4$)</td>
<td>−1.875</td>
<td>−1.232</td>
<td>−3.087</td>
</tr>
<tr>
<td>parameter</td>
<td>mean</td>
<td>s.d.</td>
<td>mean</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.450</td>
<td>0.051</td>
<td>0.579</td>
</tr>
<tr>
<td>$\tau$</td>
<td>0.604</td>
<td>0.262</td>
<td>0.500</td>
</tr>
</tbody>
</table>
Table 4. Overview of data-fusion estimation results for the IDM model.

<table>
<thead>
<tr>
<th>log-likelihood ($\times 10^4$)</th>
<th>empirical data</th>
<th>simulator data</th>
<th>mixed data</th>
</tr>
</thead>
<tbody>
<tr>
<td>parameter</td>
<td>mean</td>
<td>s.d.</td>
<td>mean</td>
</tr>
<tr>
<td>$a$ (m s$^{-2}$)</td>
<td>1.297</td>
<td>0.160</td>
<td>0.995</td>
</tr>
<tr>
<td>$b$ (m s$^{-2}$)</td>
<td>0.768</td>
<td>0.320</td>
<td>1.250</td>
</tr>
<tr>
<td>$s_0$ (m)</td>
<td>11.66</td>
<td>2.21</td>
<td>0.71</td>
</tr>
<tr>
<td>$T$ (s)</td>
<td>1.212</td>
<td>0.194</td>
<td>0.855</td>
</tr>
<tr>
<td>$v_0$ (m s$^{-1}$)</td>
<td>31.37</td>
<td>5.38</td>
<td>31.61</td>
</tr>
</tbody>
</table>

From the table, we can observe the differences in the estimates and their variability—reflecting the heterogeneity in the driving population as well as the estimation variance—when using different data sources. First of all, we conclude that the mean sensitivity is lower when using empirical data than when using synthetic data, while for the reaction time, the opposite can be observed. For the sensitivity $a$ we also see that the standard deviation is much smaller for the empirical data than for the simulator data; for the reaction time $T$, no considerable difference is observed.

When combining the datasets, we see that the average estimates change: the sensitivity alpha of the combined data yields an estimate that is somewhere between the estimates of the respective data sources. The reaction time, however, is lower than the estimates from the empirical data and the simulator data. Given the smaller standard deviation, it appears that the estimation results from the combined data sources are more reliable, which is the likely cause of the unexpected difference in the estimated value for the reaction time.

Finally, note that the sum of the likelihoods is smaller than the likelihood value obtained using the mixed dataset.

Let us now consider the estimation results obtained when we consider the IDM. Table 4 shows the estimation results for the empirical data, the data from the driving simulator and the combination of the data sources, as well as the log-likelihood values for the respective estimation experiments.

With regard to the IDM, we can see from the log-likelihood value that the model performs better than the simple GHR model. The relative improvement in the log-likelihood is not very high, but sufficient to conclude that the model’s performance is statistically better than the GHR model (see also Hoogendoorn & Ossen 2005).

Also for the IDM, differences between the parameter estimates stemming from empirical and experimental data are clearly present. These differences are especially large from the stopping distance $s_0$ and for the minimum time headway $T$. When combining the two data sources, the estimates for $s_0$ and $T$ appear to be somewhere in between the values of the single data source estimates.

This does not hold for the parameters $b$ (maximum deceleration) and $v_0$ (free driving speed). For both parameters, we see a strong reduction in the standard deviation of the parameter estimates, implying that data triangulation here leads
to better estimates for the parameters. This is further supported by the fact that the log-likelihood of the mixed-data estimation is larger than the sum of the log-likelihoods of the individual estimations.

7. Conclusions and future work

This article has presented a new framework to estimate parameters of car-following models. Unlike the many approaches that have been put forward in past decades, the approach presented here allows for statistical analysis of the estimates (standard error, covariances, etc), including testing the statistical significance of the individual parameters and the model as a whole. Furthermore, the approach can be used to effectively cross-compare models of different complexity (in terms of the number of model parameters).

In applying the method, problems occur when the individual trajectory data are not sufficiently rich, i.e. do not contain sufficient information. To resolve this, a generalization of the approach to jointly consider multiple trajectories stemming from different data sources was proposed and applied successfully. Furthermore, generalizations using prior information about the parameters to be estimated are presented. Using prior information, it can be ensured that the parameter estimates are realistic and that the car-following model has specific important characteristics (e.g. local or asymptotic stability, correct fundamental diagram, etc.).

The approach is generic and can be applied to any car-following model. In the paper, the approach has been successfully applied to estimate the parameters of the GHR and IDM. The approach allows for studying heterogeneity in the driving population with respect to the parameters of the car-following model.

The research presented in this article is part of the research programme ‘Traffic and Travel Behaviour in case of Exceptional Events’, sponsored by the Dutch Foundation of Scientific Research MaGW-NWO. The authors are very grateful to the constructive criticism and remarks of the anonymous reviewers.

References


Multiple data-source calibration


