Driver reaction time estimation from real car following data and application in GM-type model evaluation

Xiaoliang Ma* and Ingmar Andréasson

Center for Traffic Simulation Research (CTR)
Royal Institute of Technology (KTH), Stockholm 10044, Sweden

Abstract

Driver behavior plays an important role in modeling vehicle dynamics in a traffic simulation environment. To study one element of the general driver behavior, that of car following, an advanced instrumented vehicle has been applied in dynamic data collection in real traffic flow on Swedish roads. This paper briefly introduces our car following data collection and smoothing methods. Moreover, we introduce spectrum analysis methods based on Fourier analysis of car following data to estimate driver reaction times, a crucial parameter of driver behavior. As an example, we calibrate a generalized GM-type model, an extension of the classical nonlinear GM model, in stable following regime based on the estimated driver reaction times. The calibrated model is then evaluated by closed-loop simulations.

keyword: Car following data, reaction time estimation, spectrum analysis methods, GM-type model calibration and closed-loop simulation.

1 Introduction

Driver behavior is an important factor for the evaluation of modern transportation systems. The new technologies in the form of Intelligent Transportation System (ITS) make understanding of behavioral response of drivers even more necessary than ever. On the other hand, ITS technologies offer modern researchers a good opportunity to observe and analyze human behavior in different real traffic situations. Recently, the use of microscopic simulation to assess the effects of new ITS systems both on-road and in-vehicle is becoming increasingly popular and credibility of the simulation models is crucial to their successful applications. Hence calibration of the model based on measurements from real traffic environment attracts increased attention in transportation and traffic planning.

In microscopic traffic simulation, driver behavior is one of the most important part of the models to reflect general traffic characteristics. Moreover, the car following model is an essential component of driver behavior models. Therefore, calibration or adjustment of parameters of the car following model directly affects the applicability of the simulation models.

1.1 Car following behavior

Car following models describe the longitudinal behavior of drivers when they are following a leading vehicle and trying to maintain a driver specific safe headway distance to that vehicle. Normally,
the driver in the car following stage is assumed to move forward without the intention of changing lane or overtaking. Research of driving behavior in traffic science can be traced back to the 1950s when traffic flow theory was initially developed and car following models were extensively studied. The well-known mathematical car following model introduced by Gazis et al. in 1961 [1] was both an extension of the early models, e.g. Chandler et al. [2] and a summary of the early idea of stimulus-response type car following models. The model takes the form of a differential equation

\[ a_n(t + \tau_n) = \alpha \frac{v_n(t + \tau_n)^3}{[x_{n-1}(t) - x_n(t)]\gamma} [v_{n-1}(t) - v_n(t)] \]  

where \( x(t) \), \( v(t) \) and \( a(t) \) are the position, speed and acceleration of the vehicles. \( \tau_n \) is called reaction time of the driver and assumed to be a fixed value for a certain driver \( n \); \( \alpha \), \( \beta \) and \( \gamma \) are constant parameters. This model is often called the nonlinear General Motor (GM) model and has the intuitive hypothesis that the follower’s acceleration is proportional to the speed difference term, \( \delta v(t) = v_{n-1}(t) - v_n(t) \), and the exponent of its own speed but being inversely proportional to that of distance headway, \( D(t) = x_{n-1}(t) - x_n(t) \). The speed difference part at the right hand side is always translated as the ‘stimulus’ term while the fraction between the exponent of following speed and that of distance headway is called the ‘sensitivity’ term. The nonlinear GM model has been calibrated on different data sets by many researchers, though a variety of estimated parameter sets was reported [3]. Besides the GM-type models, there are many other model forms such as fuzzy inference models [4] and action point models [5], and recent research on this basic topic of traffic science has fulfilled a promising progress, which promotes both the elementary research of human behavior science and its applications in traffic engineering.

1.2 Driver reaction delay

Driver reaction delay is a common characteristic of humans in operation and control, such as driving a car. The operational coefficients and delay characteristic of humans can vary rapidly due to changes of factors such as task demands, motivation, workload and fatigue [6]. However, estimation of these variations is almost impossible in the classical paradigm. Therefore, an assumption of a fixed reaction delay in a certain regime still can not be completely circumvented.

Driver reaction time was defined as the summation of perception time and foot movement time by earlier car-following research [7]. In psychological studies, the driver reaction process is further represented in four states: perception, recognition, decision and physical response. In microscopic traffic simulation, the driver and vehicle are normally modeled as an integrated unit and the delay within the mechanical system of the vehicle is often neglected. Although research on car following models has been historically focused on exploration of different modeling frameworks and variables that affect this behavior, it has been recognized that the reaction delay \( \tau_n \) of each driver \( n \) is an indispensable factor for the identification of car following models. It affects the traffic dynamics not only in a microscopic way but also macroscopically [8].

Many studies have estimated the reaction time based on indoor experiments and driving simulators. For example, in the study by Johansson and Rummer [9] more than 300 subjects were instructed to brake a pedal as soon as they heard a sound. The estimated reaction time varied from 0.4 second to 2.7 seconds with a mean value of 1.0 second. A recent study using both a real driving environment and a simulator [10] shows that the reaction time of drivers to an anticipated danger in a real environment has a mean value of 0.42 seconds and a standard deviation of 0.14 seconds whereas the mean value of the reaction time distribution to an unanticipated danger by extreme braking is about 1.1 seconds and that in a simulator is about 0.9 seconds. In real traffic, the driver reaction to expected and unexpected stimuli are also different [11], and Fambro reported that the mean of reaction times for unexpected and expected stimuli are 1.3 seconds and 0.7 seconds respectively.

To estimate driver reaction delays from real data, several approaches have been proposed. Ozaki [12] developed a graphical method to identify the individual driver reaction time based on speed difference and acceleration profiles. Ranjitkar et al. [13] applied the graphical method in stability analysis of car-following behaviors, and based on car-following data collected on a test track, they estimated that the average driver reaction time for individual drivers ranged from 1.27 to 1.55 seconds. Ahmed computed the reaction time jointly with other parameters of the car-following model, and the estimated mean value of reaction times was 1.34 seconds [14]. The reaction time
distribution, assumed log-normal, was estimated from empirical data using the maximum likelihood method under assumptions of a predetermined GM-type model form.

1.3 Objective and structure

To study general driver behavior and identify parameters of car following models for micro-simulation, an advanced instrumented vehicle has been applied to collect data using a high measurement frequency in real traffic on Swedish roads. The collected data is smoothed using the Kalman smoothing algorithm and then analyzed in the frequency domain based on Fourier Transform and spectrum analysis, through which the reaction delay time can be estimated for each driver observed. Based on the estimated delay and smoothed data, a car following database is created for model calibration and evaluation purposes.

This article is organized as follows: in the second section the data collection method for this study is briefly introduced; the third section illustrates the mathematical foundation of spectrum analysis and its applications in the driver reaction delay estimation; in the fourth section, we calibrate a generalized GM-type model based on reaction times identified by spectrum analysis methods and then evaluate the model using data from our car following database; finally, we summarize this paper with conclusions.

2 Data collection and car following regimes

An instrumented vehicle developed by Volvo Technology was employed in our car following data collection. The car was equipped with a GPS-based navigation system and an on-board trip computer for recording travel time, distance, speed, fuel economy etc. Two lidar sensors and corresponding video cameras were installed to observe objects in both the front and rear directions. Each lidar sensor can capture at most four objects at the same time. The distance between the instrumented car and the observed object can be continuously measured by sensors at a maximal frequency of 50 times per second (Hz), that is the measuring interval is \( \Delta t = 0.02 \) sec. Meanwhile, the Volvo ERS software installed on a portable computer connected to the equipment assembles all infused information, synchronizes them with video signals and then writes them as a binary file. A real-time adaptive filtering algorithm is adopted in the software to smooth the measured distance data, from which relative speed and acceleration are numerically approximated. Hence, the data and traffic situations can be compared and analyzed at the same time using the software. Figure 1 shows the interface of the Volvo ERS logger. The collected car following data can be written as

\[
x(t) = [s_i(t) \, v_i(t) \, a_i(t) \, D(t) \, dv(t) \, da(t)]^T
\]

where \( s_i(t), v_i(t) \) and \( a_i(t) \) are position, speed and acceleration of the instrumented vehicle \( i \), and \( D(t), dv(t) \) and \( da(t) \) are relative position, speed and acceleration between the instrumented car and observed one.

Our car following experiments were at this moment mostly conducted on a motorway section of Road E18 near Stockholm with speed limits of 70 kph and 90 kph. The follow-the-leader behavior of random vehicles behind our equipped car were observed. The motorway section includes the ’2+1’ type road where ‘pure’ car-following behavior can be mostly captured at the single-lane parts. The unique features of our equipped car and the high measuring frequency provide an opportunity to obtain car following data of high quality from real traffic. However, relative speed and acceleration profiles are only derived based on real-time adaptive filtering of the distance measurement, by which certain delays might be introduced and the filtering results are not smooth enough for our car following study. In an earlier publication [15], we formulate a state-space model for the tracked vehicles based on existing physical relations in vehicle states and on a random walk model of acceleration time series. The model can be represented as follows

\[
X(t + 1) = F \cdot X(t) + V(t)
\]

\[
Y(t) = H \cdot X(t) + W(t).
\]

where \( X(t) = [s_n(t) \, v_n(t) \, a_n(t)]^T \) is the physical state vector of position, speed and acceleration; \( V(t) = [0 \, 0 \, v(t)]^T \) is the noise vector in the state transition; \( Y(t) = [\hat{s}_n(t) \, \hat{v}_n(t) \, \hat{a}_n(t)]^T \) is the
As far as is known, the Kalman filter can be applied to estimate vehicle states with optimal results in the MMS sense. In our former studies [15], the Kalman smoothing algorithm, also called the Rauch-Tung-Striebel smoother [16], is adopted to get a more accurate estimate of the states of the following vehicle only based on the distance measurement, that is, $H = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$. Figure 2 shows very clear speed and acceleration patterns in the car following mode estimated from the distance measurement and the comparison with real-time estimates. Therefore, based on our unique instrumented car and data estimation method, random car-following patterns can be obtained with high-accuracy estimation results comparable to those from other measurement methods.

It has been revealed that the longitudinal behavior of drivers changes under different conditions, and hence it is necessary to separate behavioral patterns of drivers in different situations [17]. In a former study by Bengtsson [18], four common stages were proposed: following, cut-in, braking and approaching. Each of them adopted a unique parameter set. It was reported that the resulting model is more robust than trying to estimate a parameter set for all stages. In our study [19], car-following data is also classified into four regimes: approaching, stable following, braking and accelerating. Approaching describes the process that a vehicle starts to react to a leading vehicle and ends when the following vehicle reaches the area of his desired distance. In the stable following regime drivers can both accelerate or decelerate mildly to keep in the area of a desired distance and this process may result in a temporary headway closing or opening. The braking process includes more intensive deceleration behavior when the following vehicle comes too close to the leading one or the following speed is much faster than that of the leading vehicle. The acceleration process describes how the following vehicle accelerates intensively to a speed level when a leading vehicle exists. This process may result in either the closing or the opening of two cars. Figure 3 illustrates all car following regimes.

### 3 Spectrum analysis and driver reaction time estimation

In this section, we introduce spectrum analysis methods to analyze the car following data in the frequency domain. Before applying the method to our driver reaction delay estimation, the mathematical foundations of spectrum analysis and delay estimation methods are reviewed first.

#### 3.1 Mathematical review

Spectrum analysis is a common methodology in system identification [20], signal processing [21] and time series analysis [22]. It is based on the analysis of cross and auto spectrum of two stationary stochastic processes, which are defined as the Fourier Transform (FT) of the auto-covariance and cross-covariance functions of the process as follows

$$C_{xy}(\omega) = \sum_{\tau=-\infty}^{\infty} \gamma_{xy}(\tau)e^{-i\omega\tau} \quad (4)$$

$$S_{xx}(\omega) = \sum_{\tau=-\infty}^{\infty} \gamma_{xx}(\tau)e^{-i\omega\tau} \quad (5)$$

where $\gamma_{xy}(\tau)$ and $\gamma_{xx}(\tau)$ are the cross covariance function between the input $x(t)$ and the output $y(t)$ and the auto-covariance function of the input $x(t)$ respectively. The coherency spectrum is defined as the normalized cross spectrum of $C_{xy}(\omega)$

$$\text{Coh}_{xy}(\omega) = \frac{|C_{xy}(\omega)|}{\sqrt{S_{xx}(\omega)S_{yy}(\omega)}} \quad (6)$$

and it is a bounded measure for the linear association between the time series. The gain spectrum is defined as

$$G_{xy}(\omega) = \frac{|C_{xy}(\omega)|}{S_{xx}(\omega)} \quad (7)$$
which describes the amplitude of the transfer function of the linear system between the input process \( x(t) \) and the output process \( y(t) \). At last, the phase spectrum \( \Phi_{xy}(\omega) \) is defined by

\[
\Phi_{xy}(\omega) = \arg[C_{xy}(\omega)]
\]  

or

\[
C_{xy}(\omega) = |C_{xy}(\omega)|e^{i\Phi_{xy}(\omega)}.
\]

It is worth mentioning that the phase spectrum of a pure delay system, \( H(\omega) = e^{-i\omega \tau} \), is a straight line and this is because the cross spectrum can be written as

\[
C_{xy}(\omega) = H^*(\omega)S_{xx}(\omega)
\]

where \( H^*(\omega) \) is the conjugate of the system transfer function. Moreover, since \( S_{xx}(\omega) \) is real,

\[
\Phi_{xy}(\omega) = \arg[H^*(\omega)] = \omega \tau.
\]

### 3.2 Delay estimation methods

There are a number of methods for estimation of delay between a pair of time series or signals, which can be classified as parametric estimations and non-parametric estimations. The former one is based on the mathematical modeling of time series such as ARMA models while the latter tries to explore the delay by correlation and spectral analysis of data. Björklund [23] reviewed the delay estimation methods in general from a control and system identification point of view. In this section we will mainly focus on two non-parametric delay estimation methods using spectrum analysis.

A conventional way to estimate the delay between two time series in the frequency domain is to fit a straight line to the phase spectrum due to the relation in equation (11). Hannan [24] proposed to fit a straight line in the least square sense by finding the delay \( \tau \) that maximizes the following objective function

\[
L(\tau) = \sum_{\omega_i \in B} \frac{\text{Coh}^2(\omega_i)}{1 - \text{Coh}^2(\omega_i)} \cos(\arg[H(\omega_i)] - \omega_i \tau)
\]

where \( B \) represents available frequencies with significant coherence. The cosine function in equation (12) is used to take care of the periodicity problem of the phase spectrum. In principle, this method gives exact delay estimation when the linear system is a pure delay system. In reality, many physical and biological systems can be modeled as a minimum phase system with an all-pass system like in figure 4a (note that the \( Z \)-transform is used in the system representation) , although the all-pass part may be more than a pure delay. Muller et al. [25] suggest to correct the phase by eliminating the minimum phase part, which can be estimated using the Hilbert Transform method [26], and then to estimate the delay by fitting a straight line to the all pass part. The idea is to find the delay \( \tau \) maximizing the updated objective function

\[
L_2(\tau) = \sum_{\omega_i \in B} \frac{\text{Coh}^2(\omega_i)}{1 - \text{Coh}^2(\omega_i)} \cos(\arg[H(\omega_i)] - \log |G_{xy}(\omega_i)| - \omega_i \tau).
\]

where \( \log |G_{xy}(\omega_i)| \) is the Hilbert Transform of the logarithm of the gain or amplitude of the transform function of the linear system. However, this method is valid only when a broad band of coherence exists but is limited to the delay estimation between signals correlated within a very narrow band [25].

Another extensively used method is based on coherence analysis, as a delay \( \tau \) between time series in principle will cause a reduction in the estimated coherence between them. To compensate the reduction in coherence due to the delay, artificial shifts are applied to realign one of the time series by a lag \( \tau \) while keeping the other one constant. Then the delay shift giving the maximum coherence at a significantly coherent frequency may be found. Mathematically, the idea can be represented as

\[
\tau_m = \arg \max_{\tau} \text{Coh}_{xy}(\tau, \omega_0)
\]

where \( \omega_0 \) is a frequency or frequency band of significant coherence and spectral power density. This method is suitable for the delay estimation between time series of very narrow-banded coherence for which the phase spectrum method can not be reliably applied.
3.3 Validation

Both delay estimation methods of (13) and (14) have been validated in a number of publications such as [25]. To evaluate these methods scientifically before applying to our own time series, we first validate them by filtering an input signal through a system like figure 4b where the all-pass part is a pure delay i.e. \( Q(Z) = 1 \) and the minimum phase is set to be first to third order all pole models or AR(N) models, which can be represented by

\[
y(t) = \sum_{n=1}^{N} a_n y(t - n) + x(t - \tau)
\]

(15)

where \( N = 1...3 \). First, we test the method when the input is white noise. Both the phase spectrum and coherence analysis methods give a satisfying estimation of the delay time. Then we change the input to a narrow band signal such as our measured car following data in figure 5. The phase spectrum method does not lead to a satisfying result while the coherence method gives a rather good estimation of the delay. Furthermore, we examine the methods by delay estimation in a more complex system in figure 4b where \( H_1(Z) \) is an ARMA(I,3) system and \( H_2(Z) \) is a third order AR system. The signal to noise ratio (SNR) in the experiment ranges from 0 to 10 dB. When the input \( x(t) \) is white noise, both methods can be successfully applied to estimate the delay, e.g. figure 6-1. When input is a narrow band signal such as the speed difference time series of the car following data, the coherence method can be more successfully applied. It seems that coherence analysis performs better in the validation procedures for narrow band signals. However, all these tests are conducted for true linear systems, because cross-spectral analysis assumes linear and stationary properties of the input and output relation. But this assumption is always violated to some degree due to the nonlinearity and noise involvement in real life systems. Nevertheless, the study by Barton and Cohn [27] shows that the response of the visual system of humans to oncoming objects in a driving task exhibits a band-pass characteristic, which is well described by a linear, minimum phase, third order transfer function. This serves as a concrete support to our application of the spectrum analysis methodology.

3.4 Application

After evaluation of both delay estimation methods above, we apply them in our driver reaction time estimation from real car following data. It is not an easy task in practice to apply spectrum estimation since discrete FT of the finite data sequence results in the spectral leakage problem [21]. In many cases, some window types \( W_s \) with smoothed edges may apply to the data in order to reduce the leakage when computing the periodogram, a straightforward estimator of the power spectrum. However, the periodogram is not a consistent estimator for the spectrum since the variance of it does not reduce when the length of data increases. In literature such as [21], many methods have been proposed to estimate the spectrum but a trade-off always exists between the variance and bias of the estimation. In general, the approximate variance can be derived for the gain and phase spectrum [22]

\[
\var[\hat{G}_{xy}(\omega)] = \frac{1}{\nu} \left( \frac{1}{\text{Coh}_{xy}(\omega)^2} + 3 \right) |G_{xy}(\omega)|
\]

(16)

\[
\var[\hat{\Phi}_{xy}(\omega)] = \frac{1}{\nu} \left( \frac{1}{\text{Coh}_{xy}(\omega)^2} - 1 \right)
\]

(17)

where \( \nu \) is defined as the effective number of degrees of freedom and depends on the smoothing window \( W_s \). Moreover, the phase estimate has its 95% confidence interval given by

\[
\hat{\Phi}_{xy}(\omega) \pm 1.96 \sqrt{\frac{1}{2M} \left( \frac{1}{\text{Coh}_{xy}(\omega)^2} - 1 \right)}
\]

(18)

where \( M \) is the number of disjoint segments that the data sequence is divided into when calculating the power spectra and cross spectrum. Therefore, both the phase and gain spectrum can not be reliably estimated in the case of small coherency.

In the reaction delay estimation, we try to apply both methods with a variation of window types, frequency resolutions and so on [21]. Certain criteria have been used to determine a satisfying
reaction delay. First of all, it should be within a reasonable range, e.g. 0 to 3 second; secondly, since car following data are all narrow band signals, the delay estimated from a frequency of narrow-banded coherence by the coherence analysis method is more reliable than the result from phase estimation in this case; finally, if the estimations from both methods are consistent or close, the result is more accepted. Figure 6-2 shows an example in which we get consistent estimations of driver reaction delay using both coherence and phase analysis methods. Based on the methodology, we have estimated the delays between the speed difference inputs and acceleration outputs for ten random drivers in the whole car following process and the results range from 0.52 seconds to 1.24 seconds. However, when applying both methods to estimate delays between inputs of speed of the following car and distance headway and the acceleration, the methods may give inconsistent estimations and some of the estimated delays are zero or even positive values. This might be explained by several facts. First, the linear relation between the speed of the following car (or the distance headway) and the acceleration is often not significant, and therefore the cross-spectrum method can not be reliably applied. Second, unlike speed difference between two vehicles which is directly related to the changing rate of the visual range detected by driver’s eyes [27], human drivers have difficulty to tell, even roughly, the distance and speed values when driving; In addition, human drivers also show certain anticipation ability and this may give the estimated system some non-causal characteristics in their sensitivity. Consequently, it is practically more appropriate to adopt the time delay between the stimulus, the speed difference and the acceleration reactions. Since the sample is not large enough in our study, collecting more data will be necessary to summarize the estimated reaction times as a distribution. However, the estimated results in this section can be applied for model calibration and evaluation in the next section.

4 Evaluation of the GM-type model

In this section, a generalized GM-type car following model is calibrated based on the empirical data and the reaction time estimated by spectrum analysis. Moreover, the calibrated model is evaluated by closed-loop simulations.

4.1 A generalized GM-type model

First of all, we will introduce a more generalized form of the classical GM-type model, an extension of equation (1), and the corresponding assumptions on this model form. At first, this generalized GM model assumes asymmetry between the acceleration and deceleration processes; moreover, it is assumed that positive speed differences will result in positive acceleration of the following car while negative speed differences lead to the deceleration of the following vehicle. Secondly, we propose that a reaction delay exists only between the speed difference stimulus and the acceleration output while the sensitivity term is determined by the real-time information of space headway and following speed. At last, we assume that the model parameters are different for different car following regimes including stable following, approaching, accelerating and braking. This means that we intend to estimate different parameter sets for the corresponding car following stages in the micro-simulation model. Mathematically, the model is expressed by

\[ a_{n}^{s,g}(t + \tau_{n}) = \alpha_{n}^{s,g} \frac{v_{n}(t + \tau_{n})^{\beta_{n}^{s,g}}}{[x_{n-1}(t + \tau_{n}) - x_{n}(t + \tau_{n})]^{\gamma_{n}^{s,g}}} |\Delta v_{n}(t)| + \epsilon_{n}^{s,g}(t + \tau_{n}). \]  

(19)

where

\[ s \in \{\text{stable following, approaching, accelerating, braking}\}, \]

\[ g \in \{\text{acceleration, deceleration}\}, \]

\[ \tau_{n} = \text{reaction time to the stimulus} \]

\[ \Delta v_{n}(t) = v_{n-1}(t) - v_{n}(t), \]

\[ \epsilon_{n}^{s,g}(t + \tau_{n}) = \text{stochastic random term at time } t + \tau_{n}. \]
4.2 Model calibration

The current car following data from ten randomly observed drivers are filtered, realigned according to the reaction delay estimation and stored in a database using a computer program developed for this purpose. With more data collection by the instrumented vehicle, the size of the database will increase accordingly and the calibration will be repeated to include the new datasets. It is worth mentioning that we mainly report the results from calibration in the stable following regime in this article since the principle and procedure are similar for other regimes and the stable following pattern is most abundant in our datasets but is traditionally difficult for regression analyses. In practice, the calibration is often reformulated as a nonlinear optimization problem i.e. to minimize the summation of squared errors between the model output and real output data. Mathematically, this can be represented by

$$\min_{\alpha, \beta, \gamma} \sum_i \sum_t \epsilon_i^n(t)^2 = \min_{\alpha, \beta, \gamma} \sum_i \sum_t (\tilde{a}_i^n(t) - \alpha \frac{v_i^n(t)^\beta}{|x_{i-1}(t) - x_i^n(t)|} |\Delta v_i^n(t - \tau_n)|)^2$$

(20)

where $\tilde{a}_i^n$ is the real acceleration of the following vehicle $n$ for the data sequence $i$ and the parameter sets of $\alpha$, $\beta$ and $\gamma$, are different for acceleration and deceleration processes. This problem can be solved numerically either by the gradient-based methods, e.g. the iterative gradient search and conjugate gradient method, or by the derivative-free methods, e.g. the grid search and genetic algorithm [28]. Our calibration is conducted using MATLAB and table 1 shows the calibrated parameters for the generalized GM-type model using data from stable following regimes. Figure 7 illustrates the distribution of the stochastic residual term $\epsilon(t)$ after the calibration process.

<table>
<thead>
<tr>
<th>Action</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\alpha$</td>
</tr>
<tr>
<td>acceleration</td>
<td>0.0025</td>
</tr>
<tr>
<td>deceleration</td>
<td>-5.20</td>
</tr>
</tbody>
</table>

Table 1: The calibrated parameters of the generalized GM-type model in the stable following regime

4.3 Model evaluation

After calibration of the model based on real data, evaluation of the model is an indispensable procedure for real applications. Since the model is calibrated based on the acceleration output, it is necessary to study whether the model can replicate the speed and position profiles of the same data sequence in a closed-loop simulation in which the states of the leading car and the initial condition of the following vehicle are given and then the car following process is simulated. Figure 8-1 shows an example of the comparison between the original data and simulated states using the generalized GM-type model calibrated on the same car following data sequence. The speed profile of the following vehicle can be replicated more accurately than its positions since it is only the first order integral of the acceleration process. The bias in certain acceleration points and therefore the speed values may result in an unredeemable error in positions of the following vehicle. Recently, Punzo and Simonelli [29] have proposed to formulate a dynamic calibration process as follow

$$\min_{\theta} (x_{obs} - x_{sim})^T C^{-1} (x_{obs} - x_{sim})$$

(21)

where $C$ is the covariance matrix, $\theta$ is the parameters of the acceleration model ($\theta = [\alpha \beta \gamma]^T$ in the GM model) and $x_{obs}$ is the observed measure of performances (MOPs) vector while $x_{sim}$ is the MOPs vector from a closed-loop simulation based on the acceleration model and the initial condition. This formulation generalizes the situations where either acceleration or speed or position can be used as MOPs. It has been reported that using vehicle trajectories as MOPs gives better calibration results than using acceleration profiles. Hence, our calibration strategy can be further improved, and it is also worth investigating whether the strategy will enhance the prediction ability of the model.

Robustness of a model means the ability of it to predict the data not being used in the calibration process. In our car following study, this can be examined by a combination of cross-validation and
closed-loop simulation in which several data sequences are adopted in the calibration of the model and then a new car following dataset is validated by the outputs of a closed-loop simulation. Figure 8-2 shows an example of the comparison between the original data and simulated outputs using the generalized GM-type model calibrated on several other car following datasets. From the figure, the model can basically replicate the states of the following vehicle in reality without local divergence [18], though certain bias still exists. From the cross-validations on our datasets in stable following regime, the prediction errors are between 11% and 28% using the average gap measure [30]

\[
e = \frac{\sum_{t=0}^{T} |s^{\text{sim}}(t) - s^{\text{obs}}(t)|}{\sum_{t=0}^{T} |s^{\text{obs}}(t)|}.
\]  

(22)

This means that using acceleration to calibrate the model sometimes gives a high bias on the trajectory prediction of the following car. In addition, a more justifiable criteria on the validation process might study the prediction error in the whole states of the following vehicle, that is

\[
e = \frac{\sum_{t=0}^{T} \sqrt{\|Y^{\text{sim}}(t) - Y^{\text{obs}}(t)\|^2 W Y^{\text{sim}}(t) - Y^{\text{obs}}(t))}}{\sum_{t=0}^{T} \sqrt{Y^{\text{obs}}(t)^T W Y^{\text{obs}}(t)}}.
\]  

(23)

where \(Y(t) = [a(t) v(t) s(t)]^T\) is the state vector at time \(t\) and \(W\) is the weight matrix.

5 Summary and conclusion

Studying driver behavior based on data from real traffic has recently become a promising direction for understanding behavior response of drivers, for modeling driver behavior in traffic simulation and for designing driver assistant systems such as Adaptive Cruise Control systems. In this paper, we have briefly presented our car following data collection and reduction method based on an advanced instrumented vehicle. Moreover, the coherence and phase spectrum analyses are introduced as two methods to estimate an important parameter of car following models, the reaction time or reaction delay. The advantage of spectrum analysis is that we may directly estimate the reaction delay from the empirical data without assumptions of the model form, and the estimated value can be understood as an independent parameter of driver properties without any relation with adopted car following models. However, the methods are limited by the linear and stationary presumptions on the system and the difficulties of spectral estimations in practice. Hence the validation of the estimated reaction times is still not sufficient and further exploration on this topic both in the time series analysis domain and in psychology aspects is still necessary. In addition, the fixed delay assumption also puts a restriction on the ability of model to represent reality; estimating the delay as a statistical or uncertainty term base on a quantitative analysis of real data may improve this deficiency.

Based on the fixed reaction time estimation, a database of car following has been initiated using datasets from ten randomly observed drivers. We have modified the classical GM-type model and calibrated it using datasets in the stable following regime. To test the properties of the model, closed-loop simulations are applied to evaluate the replicability and robustness of the model. The results show that the calibrated model can basically describe the stable car-following regime, though certain bias exists. Therefore, more effort should be put on exploring better strategies on dynamic calibration of the model; meanwhile, similar studies are necessary to extend to other car-following regimes and the global stability of the model should also be evaluated either analytically or in a microscopic simulation environment in our future research.

Acknowledgments

The authors would like to thank professor Karl-Lennart Bång for his consistent supports on the data collection based on the advanced instrumented vehicle. This research work is undertaken as part of the SIMLAB project, which is funded by the Swedish National Road Administration (Vägverket) and Center for Traffic Simulation Research (CTR) at Royal Institute of Technology (KTH).
References


Figure 1: The interface of Volvo ERS software
Figure 2: Comparison of measured real-time vehicle states by the ERS software and offline estimates of vehicle states by the Kalman smoothing algorithm (the measuring interval is $\Delta t = 0.04$ seconds).
Braking regime: the follower tends to brake hard when it enters a dangerous area or its speed is too high.

Approaching regime: the following vehicle intends to adapt its speed before entering stable following area $D_{stab}$.

Stable following regime: the following vehicle is within its desired area, $D_{stab}$, and its acceleration $a_n$ fluctuates mildly according to changes in speed difference.

Braking regime: the follower tends to brake hard when it enters a dangerous area or its speed is too high.

Accelerating regime: the follower intends to accelerate to a desired speed when it follows another vehicle.

Figure 3: Classification of interactive car following regimes (note: the distance $D$ is often represented in time headway in seconds to generalize different speed levels)
General linear system: \( H(Z) = H_a(Z)H_m(Z) \)

(a)

input \( x(t) \)

All pass filter

\[ H_a(Z) = Z^{-\tau}Q(Z) \]

output \( y(t) \)

(b)

input \( x(t) \)

Minimum phase filter

\[ H_m(Z) = \frac{A_1(Z)}{B_1(Z)} \]

output \( y(t) \)

\[ H(Z) = H_a(Z)H_m(Z) \]

\[ H(Z) = Z^{-\tau}Q(Z) \]

\[ H(Z) = \frac{A_1(Z)}{B_1(Z)} \]

\[ H(Z) = Z^{-\tau} \]

Figure 4: a) Representation of a general linear system by two filters: a minimum phase filter \( H_m(Z) \) and an all-pass filter \( H_a(Z) \); b) The system for the validation of the delay estimation methods.
Figure 5: Power spectra (dB) of the car following time series estimated via periodogram and the estimated transfer function (amplitude and phase) between speed difference \( \Delta v(t) \) and acceleration \( a_2(t) \).
Figure 6: Validation results of delay estimation methods based on a pure linear system with noise (the real delay is −10)(upper); Estimation of the driver reaction time between the speed difference and acceleration (the estimated lag is about 13 and the interval $\Delta t = 0.04$)(lower).
Figure 7: Residual distribution of the calibration on the stable following regime of the generalized GM-type model
Figure 8: Comparison of the real vehicle states with results from the replication test using a closed-loop simulation (upper); comparison of the real vehicle states with results from the model prediction test by cross-validation (lower).