Transportation Research Part B 80 (2015) 185–201

Contents lists available at ScienceDirect

Transportation Research Part B

journal homepage: www.elsevier.com/locate/trb

A simple nonparametric car-following model driven by field data

Zhengbing He [a], Liang Zheng [b], Wei Guan [a,*]

[a] MOE Key Laboratory for Urban Transportation Complex Systems Theory and Technology, Beijing Jiaotong University, Beijing, China
[b] School of Traffic and Transportation Engineering, Central South University, Changsha, China

ARTICLE INFO

Article history:
Received 26 November 2014
Received in revised form 14 July 2015
Accepted 15 July 2015
Available online 25 July 2015

Keywords:
Car-following model
Data-driven approach
k-nearest neighbour
Traffic oscillation
Fundamental diagram

ABSTRACT

Car-following models are always of great interest of traffic engineers and researchers. In the age of mass data, this paper proposes a nonparametric car-following model driven by field data. Different from most of the existing car-following models, neither driver's behaviour parameters nor fundamental diagrams are assumed in the data-driven model. The model is proposed based on the simple k-nearest neighbour, which outputs the average of the most similar cases, i.e., the most likely driving behaviour under the current circumstance. The inputs and outputs are selected, and the determination of the only parameter k is introduced. Three simulation scenarios are conducted to test the model. The first scenario is to simulate platoons following real leaders, where traffic waves with constant speed and the detailed trajectories are observed to be consistent with the empirical data. Driver's rubbernecking behaviour and driving errors are simulated in the second and third scenarios, respectively. The time–space diagrams of the simulated trajectories are presented and explicitly analysed. It is demonstrated that the model is able to well replicate periodic traffic oscillations from the precursor stage to the decay stage. Without making any assumption, the fundamental diagrams for the simulated scenario coincide with the empirical fundamental diagrams. These all validate that the model can well reproduce the traffic characteristics contained by the field data. The nonparametric car-following model exhibits traffic dynamics in a simple and parsimonious manner.

1. Introduction

Car-following models are always of great interests of traffic engineers and researchers. As one of the most important traffic analysis tools, the first car-following models (Pipes, 1953; Chandler et al., 1958) were proposed as early as sixty years ago. From then on, a large number of car-following models were addressed. For example, the optimal velocity model (Bando et al., 1995) and the full velocity difference model (Jiang et al., 2001) were proposed based on the intuitive speed changing of vehicles, and the well-known Gipps model (Gipps, 1981) and the intelligent driver model (Treiber et al., 2000) were proposed in the perspective of driver's acceleration and deceleration strategies. More recently, to study the trigger and formation of the stop-and-go traffic oscillations, Laval and Leclercq (2010) and Chen et al. (2012b) incorporated non-equilibrium (timid and aggressive) driver behaviour into Newell's simplified car-following model (Newell, 2002). Laval et al. (2014) captured traffic oscillations by using a desire acceleration model with a white noise term. It showed that small driver errors can result in the

* Corresponding author.
E-mail addresses: he.zb@hotmail.com (Z. He), zhengliang@csu.edu.cn (L. Zheng), weig@bjtu.edu.cn (W. Guan).

http://dx.doi.org/10.1016/j.trb.2015.07.010
0191-2615/

stop-and-go oscillation. Detailed introduction about more car-following models can be found in the review articles (see e.g. Brackstone and Mcdonald, 1999; Hoogendoorn, 2001; Helbing, 2001; Saifuzzaman and Zheng, 2014). Performance comparison amongst important car-following models can be found in the recently published book (Treiber and Kesting, 2013b). In the work, the simulated time–space diagram and the fundamental diagrams were used as core elements to evaluate performance in a figure so-called "fact sheet".

A process prior to analysing or simulating real traffic dynamics is calibration, including the fundamental diagrams, driver's behaviour parameters, etc. Only calibrated models are able to approximate real traffic dynamics and conditions. Calibration, however, is a challenging topic needed to be deeply investigated (see e.g. Brockfeld et al., 2004; Kesting and Treiber, 2008; Punzo et al., 2012; Treiber and Kesting, 2013a). Even the shapes of the fundamental diagrams in assumptions are still controversial (see e.g. Kerner, 1998; Schönhof and Helbing, 2009; Treiber et al., 2010).

Recently arising high-fidelity traffic data and data-driven modelling approaches provide an opportunity to stride over the modelling and calibrating, and to extract traffic dynamics and driver's behaviour directly from mass field data. The "manual" errors produced in modelling and calibrating are not a concern of these approaches. Artificial neural networks (ANN), support vector regression (SVR), nonparametric regression, etc. are all prevailing data-driven modelling approaches. There approaches have been widely applied and investigated in the field of short-term traffic flow forecasting; see e.g. Zhang and Ge (2013), Wu et al. (2004) and Zheng and Su (2014) and a recent review article (Vlahogianni et al., 2014). In recent years more high-fidelity traffic data, such as the trajectory data provided by Next Generation Simulation Project (NGSIM, 2006), are collected and released. This makes it possible to model car-following behaviour directly from a large number of field data.

In the studies utilising the data-driven approaches to model driver's behaviour, Panwai and Dia (2007) proposed a car-following model by using ANN with the back-propagation and fuzzy ARTMAP architectures. The model classified field data into five driving modes. Before moving, each vehicle was classified into a mode first. Space headway and leader's speed were the inputs of ANN and the follower's speed were the output. The dataset for training was limited, which was collected by an instrumented following vehicle for only 300 s. Simulation results showed that the model outperformed the Gipps and psychophysical models. Colombaroni and Fusco (2014) also modelled car-following behaviour by using ANN. Car-following data were collected by using several global-positioning-system-equipped vehicles that followed each other on urban roads. More variables were added into the training inputs in order to incorporate driver's reaction delay. Nevertheless, the analyses and comparisons based on several immediate followers might be insufficient in the above two literature. Forward-propagating traffic waves triggered by leader's deceleration could be observed in Fig. 13 in Colombaroni and Fusco (2014), which might be unrealistic. In Khodayari et al. (2012), four inputs, i.e., reaction delay, relative speed, relative distance and follower's speed, were chosen to train ANN, and the follower's acceleration was output. Although the proposed model was studied and compared, the study was only limited in leader–follower pairs. Zheng et al. (2013) further presented an explicit reaction delay model based on ANN, and tested the model in a nine-vehicle platoon following a given leader. By utilising SVM, Wei and Liu (2013) investigated the asymmetric characteristic in car-following behaviour and its impact on traffic flow evolution. Space headway, follower's speed, and relative speed were taken as the inputs, and follower's speed was output. Similarly, in the model test only an immediate follower was simulated and compared with the real immediate one.

In the existing literature of data-driven modelling, most of the inputs and outputs, such as reaction delay and acceleration, are continuous variables, which are also indirect to computer simulation; only a few followers are simulated to test the model, which is insufficient; ANN and SVR are complex in architecture. These are expected to be improved, in particular when apply them in a simulation scenario. To this end, the paper proposes a simple nonparametric car-following model driven by field data. In this model, all inputs and outputs are based on vehicle positions, which are straightforward to reproduce traffic dynamics; thousands of vehicles are simulated to test and validate the model; k-nearest neighbour, one of the most simplest data-driven approach, is employed, which is simple and parsimonious. It is demonstrated that without making any assumption or calibration, the model is able to well replicate important traffic characteristics contained by the field data, such as the traffic oscillations from the precursor stage to the decay stage and the fundamental diagrams.

The remainder of the paper is organised as follows: The field data driving the nonparametric model are introduced in Section 2; the model is proposed in Section 3; a simulation scenario of platoons following real vehicles is presented in Section 4 in order to primarily validate the model; Section 5 simulates a 1.25 km roadway with rubbernecking, and the results are thoroughly analysed and compared; Section 6 introduces driver errors into the model, and the response of the model to small perturbations is demonstrated; conclusions are made at last.

2. Site description and the data

The nonparametric car-following model proposed in the paper is driven by a large number of field data. The well-known NGSIM providing high-fidelity trajectory dataset satisfies the demand. We propose and analyse the nonparametric car-following model based on the trajectory dataset collected on a 6-lane segment in the vicinity of Lankershim Avenue on southbound US-101 freeway in Los Angeles, California. The time period of data collection ranges from 7:35 a.m. to 8:35 a.m. on June 15, 2005. We adopt the trajectory dataset collected on Lane 1 (median lane), 2, and 3, where lane changes are relatively few. Most of the speeds are less than 60 km h^{-1} (except a few faster vehicles leaving the segment), which indicates that the segment is congested during the study period. Stop-and-go oscillations originate and propagate upstream at a