Transportation Research Part C 68 (2016) 311–332

Contents lists available at ScienceDirect

Transportation Research Part C

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

A global optimization algorithm for trajectory data based car-following model calibration

Li Li [a], Xiqun (Micheal) Chen [b,*], Lei Zhang [c]

[a] Department of Automation, Tsinghua University, Beijing 100084, China
[b] College of Civil Engineering and Architecture, Zhejiang University, Hangzhou 310058, China
[c] Department of Civil and Environmental Engineering, University of Maryland, College Park, MD 20742, United States

ARTICLE INFO

Article history:
Received 28 July 2015
Received in revised form 7 April 2016
Accepted 12 April 2016

Keywords:
Microscopic car-following model
Calibration
Lipschitz optimization
Global optimization

ABSTRACT

How to calibrate the parameters of car-following models based on observed traffic data is a vital problem in traffic simulation. Usually, the core of calibration is cast into an optimization problem, in which the decision variables are car-following model parameters and the objective function usually characterizes the difference between empirical vehicle movements and their simulated correspondences. Since the objective function is usually nonlinear and non-convex, various greedy or stochastic algorithms had been proposed during the last two decades. However, the performance of these algorithms remains to be further examined. In this paper, we revisit this important problem with a special focus on the geometric feature of the objective function. First, we prove that, from a global perspective, most existing objective functions are Lipschitz continuous. Second, we show that, from a local perspective, many of these objective functions are relatively flat around the global optimal solution. Based on these two features, we propose a new global optimization algorithm that integrates global direct search and local gradient search to find the optimal solution in an efficient manner. We compare this new algorithm with several existing algorithms, including Nelder–Mead (NM) algorithm, sequential quadratic programming (SQP) algorithm, genetic algorithm (GA), and simultaneous perturbation stochastic approximation (SPSA) algorithm, on NGSIM trajectory datasets. Results demonstrate that the proposed algorithm has a fast convergence speed and a high probability of finding the global optimal solution. Moreover, it has only two major configuration parameters that can be easily determined in practice.

1. Introduction

Microscopic simulation plays an important role in the analysis and design of traffic facilities. It provides a flexible platform on which various traffic scenarios can be studied in a controlled maneuver without disrupting real-world traffic (Chung and Dumont, 2009; Barceló, 2010). Usually, a traffic simulation platform consists of several models which address different aspects of traffic behavior. In this paper, we will focus on longitudinal car-following models (Punzo and Simonelli, 2005; Gunay, 2007; Punzo and Tripodi, 2007; Kesting and Treiber, 2008; Ossen and Hoogendoorn, 2008; Tordeux et al., 2010; Chen et al., 2012, 2015; Ciuffo et al., 2012a,b).

* Corresponding author at: B828 Anzhong Building, College of Civil Engineering and Architecture, Zhejiang University, 866 Yuhangtang Rd, Hangzhou 310058, China.

E-mail address: chenxiqun@zju.edu.cn (X. (Micheal) Chen).

http://dx.doi.org/10.1016/j.trc.2016.04.011

Most microscopic simulation tools adopt certain ordinary difference equations characterized by a few parameters to model drivers' car-following actions. In order to guarantee the selected simulation models well reproduce traffic phenomena observed in practice, different calibration methods have been proposed to appropriately determine the model parameters (Wu et al., 2003; Ahn et al., 2004; Brockfeld et al., 2005; Panwai and Dia, 2005; Punzo and Simonelli, 2005; Kesting and Treiber, 2008; Ossen and Hoogendoorn, 2008; Punzo et al., 2012).

In some microscopic car-following models, a few parameters have physical equivalents in reality, e.g., desired velocity used in Gipps' car-following model (Gipps, 1981; Wilson, 2001; Punzo and Tripodi, 2007; Ciuffo et al., 2012a,b). Other macroscopic features of traffic flow parameters, e.g., free flow velocity, critical density and jam density, are important in many models and can be straightforwardly estimated from loop detector data (Rakha et al., 2007; Rakha and Gao, 2011). However, in lots of car-following models, more parameters cannot be simply derived from macroscopic measurements (Zhang and Kim, 2005; Punzo and Simonelli, 2005; Kesting and Treiber, 2008; Ossen and Hoogendoorn, 2008; Ciuffo and Punzo, 2014).

As a result, the car-following model calibration using empirical vehicle trajectory data gained increasing interests recently, since trajectory data provided much more details of driving actions on the microscopic level (Punzo and Simonelli, 2005; Kesting and Treiber, 2008; Ossen and Hoogendoorn, 2009; Chiabaut et al., 2010; Punzo et al., 2012). There was an approach used for the first time in car-following calibration by Ossen and Hoogendoorn (2008) and successively by Punzo et al. (2012), making use of laboratory experiments to compare algorithms in a fair way, that is using the model itself to generate synthetic data on which performing calibration. This allows the analysist to know what the global optimum is. How to accurately estimate model parameters from trajectory then becomes an issue of importance; since none of existing car-following models can perfectly fit all of the empirical trajectories, due to various influence factors such as time-varying dynamics of driving actions (Wagner, 2012; Koutsopoulos and Farah, 2012), heterogeneity of different drivers (Ossen and Hoogendoorn, 2007, 2011; Wang et al., 2010), measurement noises (Kesting and Treiber, 2008; Ossen and Hoogendoorn, 2008) and asymmetric characteristics in car-following and their impacts on traffic flow evolution (Li et al., 2013a; Wei and Liu, 2013).

Recently, Punzo et al. (2012) made a deep investigation about methodological aspects in the calibration of car-following models, including a comparison of optimization algorithms. Further, different indicators were proposed to measure algorithm performance not only in terms of the objective function value but also in terms of the distance of calibrated parameters from the global ones. In that research, the optimization performance indicator was proposed to provide a measure of the accuracy of the best solution in a calibration experiment in terms of both the parameter values and the score of the objective function. In this paper, we follow the same measures of performance (time series of the follower's speed and spacing between leader and follower) and adopt one of the goodness-of-fit functions (i.e., root mean square error) used in Punzo et al. (2012), since the normalized optimization performance indicator cannot be achieved due to unknown true global optimum. The main methodological contribution of this study is to propose a combined two-stage optimization algorithm for trajectory data based car-following calibration.

In many solutions for the car-following model calibration, parameters were taken as constant and uncertainties were viewed as an additive disturbance to system dynamics. The core of calibration can thus be casted into an optimization problem, in which decision variables are model parameters and the objective function usually characterizes the difference between empirical vehicle movements and their simulated correspondences.

In the simulation, usually two consecutive vehicles are considered at one time. The state of the leading vehicle is updated according to empirical observations, and the state of the following vehicle is updated via the selected car-following model. According to simulation settings, we can further categorize the existing calibration approaches into two types (Wagner et al., 2010; Treiber and Kesting, 2013a).

The first category is called *local-fit* or direct-fit, the endogenous model variables are compared against the data, separately for each data point. Specifically, the empirical position and velocity information of both vehicles are used as input at each simulation time step. Consequently, the outputs of the car-following model at the next time steps can be calculated. The modeled accelerations are compared against the empirical ones, separately for each time step, to obtain the samples of additive disturbance. If the distribution of additive disturbances is further assumed, we can directly fit the car-following model with respect to these sampled disturbances. For example, the maximum likelihood estimation (MLE) was applied to maximize the "agreement" of model parameters with the observed data (Hoogendoorn and Ossen, 2005; Hoogendoorn and Hoogendoorn, 2010; Kim et al., 2013). Since the induced variables (usually the accelerations at different time steps) are separable in the objective function, the optimization problems of this category are usually easy to solve.

The second category is called *global-fit* or indirect-fit, a complete data trajectory is compared with a simulated trajectory. The car-following model is not directly calibrated by independently comparing the endogenous model variables with the observations. Instead, only the empirical position and velocity information of the leading vehicle are used as input at each time step. The initial state of the following vehicle is given, too. Then, the movements of the following vehicle at the rest of time steps are calculated sequentially. Finally, the simulated trajectory is compared with the empirical one. Since the induced variables (usually the velocities or positions at different time steps) are determined in a sequential manner, the objective functions of this category are usually nonlinear, non-convex and difficult to optimize.

It was demonstrated in Punzo and Simonelli (2005), Wagner et al. (2010), and Treiber and Kesting (2013a) that these two kinds of approaches might generate notably different choices of parameter values, although the same trajectory dataset was