Transportation Research Part B 97 (2017) 214–239

Contents lists available at ScienceDirect

Transportation Research Part B

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

Efficient calibration techniques for large-scale traffic simulators

Chao Zhang[a,*], Carolina Osorio[a], Gunnar Flötteröd[b]

[a] Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[b] Department of Transport Science, KTH Royal Institute of Technology, 11428 Stockholm, Sweden

A R T I C L E I N F O

Article history:
Received 15 May 2016
Revised 2 December 2016
Accepted 3 December 2016
Available online 25 January 2017

A B S T R A C T

Road transportation simulators are increasingly used by transportation stakeholders around the world for the analysis of intricate transportation systems. Model calibration is a crucial prerequisite for transportation simulators to reliably reproduce and predict traffic conditions. This paper considers the calibration of transportation simulators. The methodology is suitable for a broad family of simulators. Its use is illustrated with stochastic and computationally costly simulators. The calibration problem is formulated as a simulation-based optimization (SO) problem. We propose a metamodel approach. The analytical metamodel combines information from the simulator with information from an analytical differentiable and tractable network model that relates the calibration parameters to the simulation-based objective function. The proposed algorithm is validated by considering synthetic experiments on a toy network. It is then used to address a calibration problem with real data for a large-scale network: the Berlin metropolitan network with over 24300 links and 11300 nodes. The performance of the proposed approach is compared to a traditional benchmark method. The proposed approach significantly improves the computational efficiency of the calibration algorithm with an average reduction in simulation runtime until convergence of more than 80%. The results illustrate the scalability of the approach and its suitability for the calibration of large-scale computationally inefficient network simulators.

1. Introduction

This paper focuses on the calibration (i.e., the estimation of the input parameters) of simulation-based road transportation models. We use the term “traffic simulator” to denote any simulation-based transportation model, whether macroscopic, mesoscopic or microscopic. The problem of model calibration has been extensively studied by the transportation community. A survey of both analytical and simulation-based calibration problems and algorithms is given in Balakrishna (2006). The most extensively studied calibration problem is, arguably, that of the calibration of origin-destination (OD) matrices, with seminal work such as Cascetta and Nguyen (1988); Cascetta et al. (1993) and more recent work such as Zhou and Mahmassani (2007; 2006); Zhou (2004); Ashok and Ben-Akiva (2002) and Ashok (1996), as well as the deployment of path flow estimators (PFEs), with the seminal work of Bell et al. (1997). Recent PFE reviews include Flötteröd (2008) and Buisson et al. (2012).

* Corresponding author.
E-mail address: chaoz@mit.edu (C. Zhang).

http://dx.doi.org/10.1016/j.trb.2016.12.005
0191-2615/

This paper focuses on the calibration of simulators that are computationally costly to evaluate. The high computational cost can be due to: (i) the simulation of high levels of demand along with a high-resolution representation of demand (e.g., disaggregate representation of travelers), (ii) the simulation of a large-scale network, (iii) the use of a stochastic simulator requiring the evaluation of numerous simulation replications, (iv) the desire to evaluate performance under equilibrium conditions, which requires running sequentially multiple simulation-based assignment iterations (Nagel and Flötteröd, 2012).

Given the high computational costs involved in evaluating the simulation models, there is a need for calibration algorithms that can identify solutions with good performance at a low computational cost. That is, algorithms that can identify good solutions within few algorithmic iterations, i.e., with few simulated (near-)equilibrium responses.

We first discuss the main challenges of addressing calibration problems. We then state how the approach proposed in this paper addresses these challenges. The traffic model is often stochastic. It can involve sampling for every traveler from a variety of disaggregate behavioral models (e.g., choice models such as departure time, mode, route, lane-changing, etc.). Thus, a single run of the simulator involves drawing, for each of the thousands or hundreds of thousands of travelers, from a set of behavioral distributions. Given a sample of behavioral choices, a traffic flow model is used to propagate the travelers throughout the network. A review of state-of-the-art simulation models is presented in Barceló (2010). Thus, the mapping between the calibration parameters and the probability distribution of a given performance metric (e.g., the objective function of a calibration problem) is an intricate function.

This mapping is often non-convex and may contain multiple local minima. The stochasticity of the simulator requires the use of optimization algorithms that account for the lack of both: (i) a closed-form expression of the objective function and, (ii) exact function evaluations (since the functions can only be estimated via simulation). The traditional statistical and numerical algorithms are of limited use for calibration problems, since the underlying simulators often lack the strong assumptions required by these methods (e.g., normality, ergodicity) (Buisson et al., 2012). Hence, the traditional approach to calibration has been the use of simulation-based optimization (SO) algorithms. Most SO algorithms are general-purpose algorithms, they are not tailored to the intricacies of transportation simulators. They have been used extensively for calibration, several reviews include Ben-Akiva et al. (2012); Balakrishna (2006) and Antoniou (2004). Algorithms frequently used for calibration include simultaneous perturbation stochastic approximation (SPSA) (Spall, 1992), genetic algorithms (GA) (Holland, 1975), particle filters and Kalman filters.

The generality of these SO algorithms stems from the fact that they treat the simulator as a black-box. The main implication of this is that they are designed to achieve asymptotic (i.e., large-sample size) convergence properties. They are not designed to identify good solutions within few algorithmic iterations, i.e., they are not computationally efficient. Yet they are used by the transportation community under tight computational budgets, i.e., under small-sample size conditions. One approach to derive computationally efficient algorithms, is to exploit the structure of the underlying calibration problem. General-purpose SO algorithms exploit limited problem-structure (e.g., at most they are based on numerical linearizations). One recent work that does exploit problem structure within the calibration algorithm is that of Flötteröd et al. (2011). It formulates and embeds within the algorithm an analytical approximation of the first-order derivative of the simulator's measurement equation. This leads to significant reductions in the computational requirements of the algorithm.

Traditional SO algorithms, although designed to guarantee asymptotic properties, are typically used for calibration under tight computational budgets. This makes them sensitive to the initial points. Given the difficulty of identifying good initial points to the calibration problem, there is a need for algorithms that perform well under tight computational budgets while being robust to the quality of the initial points. Additionally, the lack of computationally efficient calibration algorithms has led recent calibration research to focus on the design of dimensionality reduction methods (e.g., sensitivity analysis methods) (Ge et al., 2014; Ge and Menendez, 2016; 2014; Ciuffo and Azevedo, 2014).

This paper addresses the following challenges.

Computational efficiency We propose an algorithm that can identify points with good performance within few algorithmic iterations. Therefore, it is a computationally efficient algorithm that reflects well the computational conditions under which calibration problems are addressed by both the transportation research and practice communities. This is achieved by designing an algorithm that exploits the transportation-specific structure of the calibration problem. More specifically, the proposed approach solves at every iteration of the calibration algorithm, an analytical (i.e., not simulation-based) approximate calibration problem. This analytical problem is solved by using information from an analytical traffic model. The latter is highly efficient. It is formulated as a system of nonlinear equations. Hence, it can be evaluated with a variety of efficient solvers. This is key for the efficiency of the calibration algorithm.

The proposed approach resorts to the use of a derivative-free algorithm. In other words, it does not rely on estimates of the derivatives of the simulation-based objective function. This further contributes to the efficiency of the algorithm.

Analytical structural information The algorithm embeds an analytical approximation of the simulator. This contributes to a largely unresolved methodological challenge which is the formulation of tractable measurement equations that link available surveillance field data to the simulator's calibration parameters.

Robustness to initial conditions The algorithm can identify good solutions within few iterations regardless of the initial points. It is robust to the quality of the initial points.

Stochasticity The algorithm is a simulation-based optimization algorithm that accounts for the simulator's stochasticity.