Automatic calibration of the link fundamental diagram for macroscopic traffic simulation models

Sasan Amini[a,*], Gabriel Tilg[a], Ahmet Cagri Tekin[a], Daniele Tiddi[b], Fritz Busch[a]

[a]*Chair of Traffic Engineering and Control, Technical University of Munich, Germany*
[b]*PTV SISTeMA, Rome, Italy*

Keywords: Traffic flow theory, Fundamental diagram, Traffic simulation

---

1. Introduction

The fundamental diagram (FD) of traffic flow theory relates the macroscopic traffic flow parameters i.e. flow (q), density (k) and average speed (v) to each other. If traffic is stationary, by knowing two of the three parameters the third one can be estimated with $q = vk$. Traditionally, the data collected by loop detectors (i.e. average flow and occupancy at fixed locations over a certain period of time) are used to estimate the FD and its shape. It is often argued by several researchers [1, 2] that loop detectors have major limitations: they cannot detect stopped vehicles, the position of the detector influences the calculated density, and the structure of shockwaves cannot be captured at a single location. In addition, averaging values over time causes averaging different traffic conditions that results into non-stationary traffic, for which fundamental relationship is not valid. Therefore, calibrating FDs based on probe vehicle data is becoming more prevalent to address these limitations.

Despite the criticism, loop detectors are still the dominant source of measuring traffic data. The collected data are mainly used to describe the traffic state on a single link. The estimated FD is extensively used for traffic control, traffic state prediction and macroscopic traffic modeling as well. The shape of the FD has been well studied for decades and several deterministic functions have been proposed to fit a curve to empirical data. For this purpose, the parameters of the FD are calibrated for the observed road section, since each road has its own characteristics, leading to a unique FD. In addition, external factors i.e. weather, control measures, work zone etc. affect the shape which should be considered in the calibration process. Surprisingly, in most cases the fitting procedure is done manually. More specifically, based on subjective judgments and depending on the shape of the target function two or three parameters are manually extracted for each detector in order to estimate the others and minimize the error. Only recently, two studies have tried to develop an automated approach to fit functions over the collected data from motorways, which are described in the following.

---

[*] Corresponding author: sasan.amini@tum.de

In an earlier study, Dervisoglu et al. [3] have proposed an approach in which first they identify the free flow speed by performing linear regression on speed-flow data based on observations with speed of higher than 55 mph. This is followed by estimating the capacity as the maximum observed value of flow measurements. Lastly, they use quantile regression to calculate the backward-wave speed. In a more recent study, Knoop and Daamen [4] have developed an automated process to fit Wu's fundamental diagrams [5] to the data from a motorway. They first separate the free flow and congested branch of the FD by fitting a triangular function. While fixing free flow speed and backward-wave speed they estimate the remaining parameters based on least-square fit. In both of these approaches the capacity drop is successfully captured.

Although these two studies have shown promising results on both empirical and simulation data, they have evaluated the methods on individual segments of intercity motorways. This requires to either calibrate the FDs for each location separately or assume similar FDs based on their physical characteristics i.e. speed limit, lane width etc. As Gu et al. [6] discuss, such an approach is not necessarily accurate and the data should be used to cluster detectors which show similar patterns. Moreover, these methods cannot be directly applied on loop detector data from urban roads. Thus, the main goal of this paper, is to propose a methodology that can fit a suitable FD function to loop detector data from both urban and rural roads which will be further used in a macroscopic traffic simulation (e.g. PTV VISUM).

2. Methodology

The proposed methodology consists of three main steps as follows:
First, data cleansing is performed on the empirical data collected from loop detectors to detect outliers and exclude them from data for FD calibration. Outlier detection is commonly performed by employing the box and whiskers [7] approach or median absolute deviation techniques. The former may result in excluding some "true" measurements on urban roads due to higher variance of traffic condition in urban networks and the latter requires manual adjustment of weight factors. Thus, in this paper, the density based spatial clustering of applications with noise (DBSCAN) [8] is used to exclude outliers from data. DBSCAN is a powerful unsupervised clustering algorithm which does not require to specify the number of clusters a priori.

The second step is to cluster similar links to calibrate the FDs only once for all links in each cluster. The clustering is done by comparing only the free flow branch of the FDs. More specifically, first a parabolic function according to Gentile [9] is applied to fit a curve on the hypocongested (free flow) branch of each FD. To achieve this, the values of maximum flow and speed are needed as an input for the procedure. These are assumed to be the 95^{th} percentile of maximum observed flow and speed as suggested in a previous study by Leclercq [10]. To establish the clusters, links with similar fitted curve (using the Frchet distance) are grouped in the same cluster.

In the third step, data inside each cluster are aggregated to fit several form of functions from which the one with the lowest root mean squared error (RMSE) is selected as the best fit. The functions used in this study are: triangular function proposed by Dervisoglu et al. [3], a polynomial function from Gentile [9] and also Wu [5]. The difference between the last two is that in Wu's function data points are weighted based on their reliability to reduce the impact of higher scatter on congested branch on the fitting.

To evaluate the efficiency of the proposed methodology, the data set is split into two parts where 80% is used for calibration of the FDs and the remaining 20% is used for validation over individual detectors. Moreover, the fitted functions are used in a macroscopic traffic simulation to estimate the traffic state on several links of the network and will be compared to the measured data.

3. Preliminary results

3.1. Data set

In total, data from 172 loop detectors that are located either on urban or rural roads are investigated in this paper. The data contains 15-minute aggregated flow and speed for a period of almost 2 years. Density is calculated from speed and flow using the fundamental relationship q=kv.

3.2. First results

The proposed methodology has been partially applied on selective links to compare the performance of the employed approaches. For the first step (outlier detection) DBSCAN shows better results than box and whiskers but it is computationally more expensive. With respect to the third step, fitting two of the three functions has been implemented which shows that the polynomial function of Gentile [9] underestimates the capacity whereas the linear function of Dervisoglu et al. [3] slightly overestimates the maximum flow. With respect to jam density the results look very similar, where Gentile estimates slightly smaller jam density. The results of the fitted curve on the speed-density function are much more satisfying for Gentile Polynomial in comparison to Dervisoglu. Figure 1 shows some of the preliminary results.

It is worth to mention that the current implementation of the algorithm can be applied for static calibration which does not account for within day variation of FD. In working towards this goal, the authors will modify the second step of the approach to account for dynamic clustering of the link FDs. The results of such approach will be provided in the full paper.

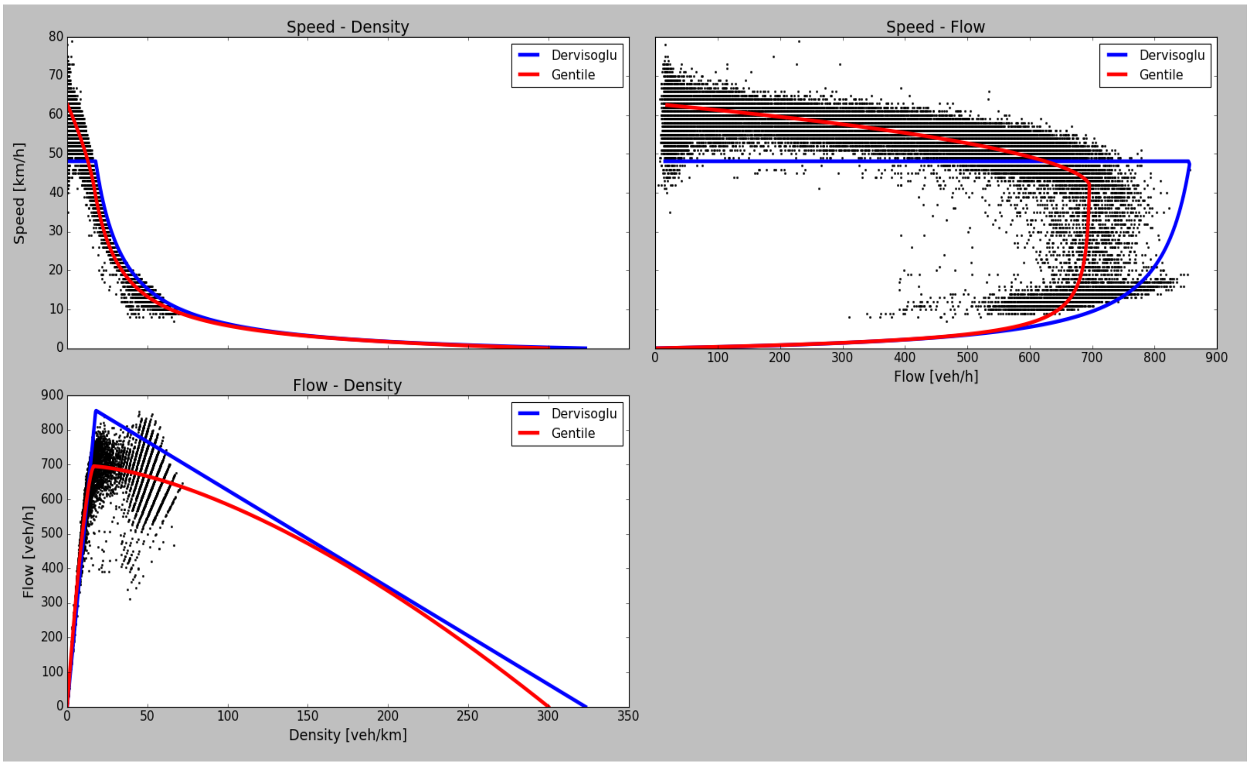

Figure 1: Fitted curves on one example FD using the method of Gentile [9] in red and Dervisoglu et al. [3] in blue.

References

[1] T. Seo, T. Kusakabe, Y. Asakura, Calibration of fundamental diagram using trajectories of probe vehicles: Basic formulation and heuristic algorithm, Transportation Research Procedia 21 (2017) 6–17.

[2] N. Chiabaut, C. Buisson, L. Leclercq, Fundamental Diagram Estimation Through Passing Rate Measurements in Congestion, IEEE Transactions on Intelligent Transportation Systems 10 (2009) 355–359.

[3] G. Dervisoglu, G. Gomes, R. Horowitz, Automatic Calibration of the Fundamental Diagram and Empirical Observations on Capacity, Transportation Research Board 88th Annual Meeting (2009).

[4] V. L. Knoop, W. Daamen, Automatic fitting procedure for the fundamental diagram, Transportmetrica B 5 (2017) 133–148.

[5] N. Wu, A new approach for modeling of Fundamental Diagrams, Transportation Research Part A: Policy and Practice 36 (2002) 867–884.

[6] Z. Gu, M. Saberi, M. Sarvi, Z. Liu, A big data approach for clustering and calibration of link fundamental diagrams for large-scale network simulation applications, Transportation Research Part C: Emerging Technologies (2017).

[7] J. W. Tukey, Exploratory data analysis, volume 2, Reading, Mass., 1977.

[8] M. Ester, H.-P. Kriegel, J. Sander, X. Xu, et al., A density-based algorithm for discovering clusters in large spatial databases with noise., in: Kdd, volume 96, pp. 226–231.

[9] G. Gentile, The general link transmission model for dynamic network loading and a comparison with the due algorithm, New developments in transport planning: advances in Dynamic Traffic Assignment 178 (2010).

[10] L. Leclercq, Calibration of Flow-Density Relationships on Urban Streets, Transportation Research Record: Journal of the Transportation Research Board 1934 (2005) 226–234.