A functional form for the macroscopic fundamental diagram with a physical meaning

Working Paper

Author(s):
Ambühl, Lukas; Loder, Allister; Bliemer, Michiel; Menendez, Monica; Axhausen, Kay W.

Publication date:
2017-11

Permanent link:
https://doi.org/10.3929/ethz-b-000207171

Rights / license:
In Copyright - Non-Commercial Use Permitted

Originally published in:
Arbeitsberichte Verkehrs- und Raumplanung 1306
A functional form for the macroscopic fundamental diagram with a physical meaning

Lukas Ambühl\textsuperscript{a}, Allister Loder\textsuperscript{a,*}, Michiel Bliemer\textsuperscript{b}, Monica Menendez\textsuperscript{a}, Kay W. Axhausen\textsuperscript{a}

\textsuperscript{a}Institute for Transport Planning and Systems, ETH Zurich, Switzerland
\textsuperscript{b}Institute of Transport and Logistics Studies, University of Sydney, Australia

Abstract

The macroscopic fundamental diagram (MFD) relates the vehicle accumulation and the production of travel in an urban network with a well-defined and reproducible curve. Thanks to its elegance, the MFD offers a wide range of applications, most notably for traffic control. Recently, more and more empirical MFDs have been documented, providing further insights and facilitating their application in real urban networks. So far, however, no generally accepted functional form has been identified.

This paper proposes a functional form for the MFD that is based on the smooth approximation of an upper bound of technologically feasible traffic states. We show the applicability of the new functional form in an empirical context with data from Marseille (France), London (UK), Lucerne, and Zurich (both Switzerland). Without loss of generality, we define the upper bound as a trapezoidal diagram which depends on four physical parameters: free-flow speed, capacity, jam density, and wave speed. The new functional form allows a physically meaningful estimation of the MFD and thus enhances the applicability thereof. Moreover, the smooth approximation uniquely determines the critical density and quantifies the network’s infrastructure potential given a technology frontier, measured as the difference between an upper bound for the MFD and the traffic states observed in real life. Based on the empirical findings, we conclude that the values of the infrastructure potential are similar across cities. Therefore, even if the empirical MFD is unknown, the MFD can be approximated for any city with a defined technology frontier, using a value of the infrastructure potential within the range found in this study.

Keywords: MFD, functional form; infrastructure potential

1. Introduction

The macroscopic fundamental diagram (MFD) relates the vehicle accumulation (average network density) and the production of travel (average network flow) in an urban network \cite{GeroliminisDaganzo2008}. The MFD can be seen as the network counterpart of a fundamental diagram which describes the steady state relationship between flow, density, and speed at the level of a road segment. The resulting well-defined curve of the MFD depends on the characteristics of the roads (i.e. the fundamental diagrams of the individual roads), the network design and signal settings, and the route choice. Given slowly varying demand, the MFD is a reproducible characteristic of an urban transportation network \cite{DaganzoGeroliminis2008, GeroliminisBoyaci2012, LeclercqGeroliminis2013}. Thanks to its elegance, the MFD offers a

*Corresponding author
Phone: +41/(0)44-633-6258
E-mail address: allister.loder@ivt.baug.ethz.ch

Preprint submitted to Transportation Research Part B: Methodological November 9, 2017
wide range of applications, most notably for traffic control (Haddad and Geroliminis, 2012; Zheng et al., 2012). Recently, more and more empirical MFDs have been documented, providing further insights and facilitating the applications of the MFD in cities. So far, however, no generally accepted functional form has been identified.

Figure 1 shows an MFD estimated using empirical data from the city of Zurich, Switzerland. Note that the density and flow are proportional to the accumulation and the travel production (assuming a constant trip length), respectively. From Figure 1 we identify two regimes: the uncongested and congested branches of the MFD. The traffic states separating both regimes are found at the maximum flow (i.e. capacity $Q$) with a critical density $k_{\text{crit}}$. With the density $k$ and the flow $q$ known, the speed follows from the fundamental relationship $v = q/k$, where the free flow speed is $u_f$. Figure 1 does not show a complete congested branch, as complete traffic gridlocks, where densities approach the jam density $\kappa$, are rarely observed in practice.

Since the formal introduction of the MFD, many potential applications have been explored: perimeter control (Haddad and Geroliminis, 2012; Yang et al., 2017), road pricing (Zheng et al., 2012), the discussion of urban design (Zheng and Geroliminis, 2013; Tsikeris and Geroliminis, 2013; Ortigosa et al., 2015), parking (Cao et al., 2017; Cao and Menendez, 2015) and traffic safety (Alsalhi and Dixit, 2015). In order to capture the multi-modality of urban transport networks, the MFD has also been extended into the 3D-MFD (Geroliminis et al., 2014; Loder et al., 2017).

Similar to the MFD in Figure 1, other empirically estimated MFDs, e.g. for Yokohama (Geroliminis and Daganzo, 2008) and Toulouse (Buisson and Ladier, 2009), exhibit a range of flows for any given density. There are at least three relevant explanations for this observed scatter. First, the upper bound corresponds to the steady state conditions, which are rarely observed in dynamic networks as second and third order effects take place (Mariotte et al., 2017; Gao and Gayah, 2017; Gayah and Daganzo, 2011). Second, inhomogeneity in the spatial distribution of vehicles always leads to a flow smaller than in the case of homogeneously distributed traffic (Daganzo et al., 2011; Ji and Geroliminis, 2012; Doig et al., 2013; Muhlich et al., 2015). Third, public transport operations can interfere with vehicle flows, e.g. public transport priority or rigid timetables (Arnet et al., 2015; Nikias et al., 2016). Nevertheless, a distinct curve, which describes the equilibrium or steady state in the network given the current characteristics of the roads, the network, routes, and vehicle
Defining a meaningful relationship between the traffic variables flow, density and speed using a functional form has attracted attention ever since the automobile became a popular means of transportation (Smeed 1968). At the road level, Schaar (1925) derived a speed-flow relationship based on the reaction time of drivers and vehicle length. Greenshields (1935) suggested a linear relationship between speed and density characterized by the parameters of free flow speed, capacity and jam density. Later, network wide relations were investigated by Herman and Prigogine (1979) and Mahmassani et al. (1987). At the network level, Smeed (1968) observed that the speed-flow relationships across different cities exhibs remarkable differences. He argued that these differences resulted from how efficiently road space was used in a city. From a policy perspective, such an investigation is very appealing, i.e., how well is the infrastructure used given the current characteristics of the roads, the network, the routes and the vehicle technology - especially in light of the automation of vehicles and the expected changes of capacities and speeds (Fagnant and Kockelman 2015). The MFD seems a promising tool, however, no methodology exists so far to illustrate and quantify how well the infrastructure is used.

This paper proposes a functional form for the MFD that is based on the smooth approximation of a trapezoidal diagram. This specific diagram is based on four physical parameters, namely free-flow speed, capacity, jam density, and wave speed. We prove that the functional form satisfies the common mathematical properties as defined for fundamental diagrams by Del Castillo and Benitez (1995). Additionally, we show the applicability of the new functional form in an empirical context with data from Marseille (France), London (UK), Lucerne, and Zurich (both Switzerland). The new functional form allows a physically meaningful estimation of the MFD and thus enhances its applicability. Moreover, the smooth approximation determines a unique critical density and quantifies the network’s infrastructure potential given a technology frontier. More importantly, we find that the values of the infrastructure potential are similar to each other across cities. Therefore we conclude that even if the empirical MFD is unknown, the MFD can be approximated for any city with a defined technology frontier, using a value of the infrastructure potential within the range reported in this paper.

The remainder of this paper is organized as follows. We introduce the new method in Section 2, including a discussion of the mathematical properties. In Section 3 we show the results of the empirical estimation. We then conclude with a discussion of the results and the implications of the new functional form for MFD research.

2. Methodology

2.1. A new functional form for the MFD

As a first order estimate of the MFD, we use a fundamental diagram, which is characterized by a free-flow speed $u_f$, a wave speed $w$, and a jam density $\kappa$. In contrast to a single road segment with infinite length, the capacity of a network is further limited by traffic signals at intersections, which constrain the network at an average flow of $Q$ (Daganzo 2007, Daganzo and Geroliminis 2008). Thus, Eqn. 1 gives the first order approximation of the MFD where flows are bounded by the lower envelope defined by the three arguments $u_f k$, $Q$ and $(\kappa - k)w$ (Daganzo 1994, 2007).

$$q (k) = \min (u_f k; Q; (\kappa - k)w)$$

This lower envelope results in a simplified trapezoidal fundamental diagram and can be seen as a theoretical upper bound of feasible traffic states, see Figure 2.
We propose to model the MFD as a smooth approximation of this trapezoidal fundamental diagram. This approximation, also referred to as a quasi-min operator or a soft-min (see Bliemer et al. 2017; Cook 2011), transforms the minimum operator into a ln-sum-exp function. Eqn. 2 provides an example thereof. Here, $\lambda$ is the smoothing parameter, with values closer to zero indicating that the function matches closely the true minimum.

$$\min \left( x_1; x_2; x_3; \ldots; x_n \right) \approx -\lambda \ln \left( \exp \left( -\frac{x_1}{\lambda} \right) + \exp \left( -\frac{x_2}{\lambda} \right) + \exp \left( -\frac{x_3}{\lambda} \right) + \ldots + \exp \left( -\frac{x_n}{\lambda} \right) \right)$$  \hspace{1cm} (2)

When we apply this smooth approximation of the minimum operator to the trapezoidal fundamental diagram, we obtain Eqn. 3 where $\lambda$ determines how far the curve lies beneath the trapezoidal diagram.

$$q(k) = \min \left( u f k; Q; (\kappa - k)w \right) = \lim_{\lambda \to 0} -\lambda \ln \left( \exp \left( -\frac{ufk}{\lambda} \right) + \exp \left( -\frac{Q}{\lambda} \right) + \exp \left( -\frac{(\kappa - k)w}{\lambda} \right) \right)$$  \hspace{1cm} (3)

Figure 3 illustrates the behavior of Eqn. 3 for two different values of $\lambda$. We observe that for $\lambda = 0.06$ the curve resembles the shape of a typical MFD. An important note, $q(k)$ is equal to zero at $k = 0$ and $k = \kappa$ only when $\lambda \to 0$, but negative otherwise. Thus, we propose to introduce a numerical boundary condition as given by Eqn. 4, where $\varepsilon$ is the maximum deviation allowed from 0. Note that this value has to be defined in an empirical context, but must be relatively small. Some discussion on the values of $\varepsilon$ in an empirical context is given in Section 3.

$$|q(0)| \leq \varepsilon, \quad |q(\kappa)| \leq \varepsilon$$  \hspace{1cm} (4)

Based on the shape defined in Eqns. 1, 3 and 4 we discuss in the next section the mathematical properties of this function. The condition of Eqn. 4 implies that the shape function must be chosen carefully and small values of $\lambda$ are expected.

### 2.2. Mathematical properties

Del Castillo and Benitez (1995) summarized the mathematical properties which should be satisfied by a functional form for the speed (flow) - density relationships on a road segment. Note that most functional forms for fundamental diagrams do not satisfy all properties at once. In this section, we test these properties on our proposed functional form for the MFD, since such properties must be satisfied at the network level.
as well. We will later apply a trapezoidal shape function on our empirical data, thus, we will evaluate the mathematical properties for such shape. In mathematical terms

(i) \( v \in [0; u_f] \), ranges from zero to the free flow speed,
(ii) \( k \in [0; \kappa] \), is bounded between zero and the jam density \( \kappa \),

for the functional relationship \( v(k) \),

(iii) free flow speed occurs when density approaches zero, thus \( v(0) = u_f \),
(iv) at \( \kappa \) no vehicle is moving, thus \( v(\kappa) = 0 \),
(v) speed is strictly monotonically decreasing with density, therefore \( v'(k) < 0 \) for \( 0 < k \leq \kappa \),
(vi) the interaction of vehicles diminishes in light traffic, i.e. \( \lim_{k \to 0} v'(k) \to 0 \),
(vii) \( q(k) \) is strictly concave or \( q''(k) < 0 \), which results from the equation of continuity of traffic flow \( \frac{\partial k}{\partial t} + \frac{\partial q}{\partial x} = 0 \).

Properties (i) and (ii) are satisfied by the very design of Eqn. 3. Hence, we only prove properties (iii) to (vii) in the following and we assume that Eqn. 4 holds.
Condition (iii): Boundary condition for free flow speed

Following the fundamental equation of traffic flow, speed \( v \) is obtained as \( v = q/k \). Using Eqn. 3 and \( k \) and \( \lambda \) approaching zero, \( u_f k \) is the smallest value for the minimum operator. Then, Eqn. 5 solves to \( u_f \).

\[
v(0) = \lim_{k \to 0} \frac{q(k)}{k} = \lim_{k \to 0} \frac{\min (u_f k; Q; (\kappa - k)w)}{k} = \lim_{k \to 0} \frac{u_f k}{k} = u_f
\]

Condition (iv): Boundary condition for jam density

With traffic density approaching jam density, the speed must drop to zero. Using Eqn. 3 and \( k \) approaching \( \kappa \) and \( \lambda \) approaching zero, \( (\kappa - k)w \) is the smallest value for the minimum operator. Then, we find in Eqn. 6 that at the jam density the speed is zero.

\[
v(\kappa) = \lim_{k \to \kappa} \frac{q(k)}{k} = \lim_{k \to \kappa} \frac{\min (u_f k; Q; (\kappa - k)w)}{k} = \lim_{k \to \kappa} \frac{(\kappa - k)w}{k} = 0
\]

Condition (v): Monotonically decreasing \( v(k) \)

The proof of \( v'(k) < 0 \) for \( 0 < k \leq \kappa \) is trivial as it is based on the shape of the MFD alone. Note that for the following conditions, \( \lambda \) is small but not zero. Consider the first derivative of speed \( v \) with respect to \( k \) in its most general form in Eqn. 7 derived from \( v = q/k \). Since \( k^2 \) in the denominator is always positive, we only discuss the difference in the numerator.

\[
v'(k) = \frac{kq'(k) - q(k)}{k^2}
\]

Reordering the numerator results in

\[
kq'(k) < q(k)
\]

And substituting \( q'(k) \) and \( q(k) \) based on Eqn. 3 leads to the following expression

\[
k \left( \frac{u_f \exp \left( \frac{-u_f k}{\lambda} \right) - w \exp \left( \frac{-w}{\lambda} \right)}{\exp \left( \frac{-u_f k}{\lambda} \right) + \exp \left( \frac{-Q}{\lambda} \right) + \exp \left( \frac{-w}{\lambda} \right)} \right) < -\lambda \ln \left( \exp \left( \frac{-u_f k}{\lambda} \right) + \exp \left( \frac{-Q}{\lambda} \right) + \exp \left( \frac{-w}{\lambda} \right) \right)
\]

Here, one can easily identify that for \( k = 0 \) the inequality does not hold, but [Del Castillo and Benitez 1995] excluded zero in their condition. Let us then consider three cases for the discussion of the monotonic decrease.
of speed. In the first case, with \( k \in \left] 0; k_{\text{crit}} \right[ \) the term \( \exp \left( - (\kappa - k) w / \lambda \right) \) on both sides of Eqn. 9 can be omitted as it is approximately zero. Then, Eqn. 9 simplifies to Eqn. 10.

\[
k \left( \frac{u_f \exp \left( - \frac{u_f k}{\lambda} \right)}{\exp \left( - \frac{u_f k}{\lambda} \right) + \exp \left( - \frac{Q}{\lambda} \right)} \right) < - \lambda \ln \left( \exp \left( - \frac{u_f k}{\lambda} \right) + \exp \left( - \frac{Q}{\lambda} \right) \right) \tag{10}
\]

In Eqn. 10, the left hand side is positive in the considered interval as \( k \) as well as \( q'(k) \) are positive. However, the right hand side is negative at very small values of \( k \) (see also Section 2.1). Thus, the inequality holds for large parts of the interval. In the second case, \( k = k_{\text{crit}} \) which leads to a simplification of Eqn. 9 and results in \( 0 < Q \). In the third case, \( k \in \left] k_{\text{crit}}; \kappa \right[ \) makes the term \( w \exp \left( - (\kappa - k) w / \lambda \right) \) dominant on the left hand side of Eqn. 9 and makes the left hand side negative for all values within the considered interval while the right hand side is always positive. We conclude then that \( v(k) \) is monotonically decreasing for \( 0 < k \leq \kappa \) for most parts of the interval when the shape function is well selected so that \( \lambda \) does not become too large.

**Condition (vi): Vanishing interactions between cars**

Intuitively, with only a few cars in the network, the interactions between cars vanishes for very low densities. Therefore, we consider Eqn. 5 where the speed is evaluated at \( k = 0 \). The derivative of \( v(0) \) is 0, as \( u_f = \text{constant} \).

**Condition (vii): Strict concavity of \( q(k) \)**

We prove the property of strict concavity of \( q(k) \) by showing that \( q''(k) < 0, \forall k \in \left] 0; \kappa \right[ \). We find the first derivative of \( q(k) \) by applying the chain rule. After reordering the terms of the second derivative in Eqn. 11 one can observe that the term in parentheses is always positive and that the entire expression is always negative with \( \lambda > 0 \). Therefore, \( q(k) \) is strictly concave.

\[
q''(k) = - \frac{1}{\lambda} \left( \exp \left( \frac{Q + \kappa w + (u + w) k}{\lambda} \right) \left( u_f^2 \exp \left( \frac{(\kappa - k) w}{\lambda} \right) + w^2 \exp \left( \frac{u_f k}{\lambda} \right) + (u_f + w)^2 \exp \left( \frac{Q}{\lambda} \right) \right) \right) < 0 \tag{11}
\]

2.2.1. Conclusions

We have shown that Eqn. 3 satisfies all mathematical properties of a flow-density relationship as proposed by Del Castillo and Benitez (1995) when the shape function is appropriately selected and \( \lambda \) is small. In addition, Eqn. 3 allows for unique solutions in MFD applications and due to its continuous form it is easy to use in optimization.

2.3. A physical meaning of the \( \lambda \) parameter

In the smooth approximation of the trapezoidal function introduced in Eqn. 3 all parameters have a physical meaning, and are linked to traffic, except for \( \lambda \). Here, we consider two interpretations of \( \lambda \) in a traffic related context.

First, recall that the magnitude of \( \lambda \) indicates how far the observed curve deviates from the shape function (i.e. upper bound). If the shape function is chosen to follow a theoretical technology frontier, \( \lambda \) can be considered as the *infrastructure potential*. In other words, it measures the gap between the observed and
the physically possible flows in the network for any given density. Smaller values of $\lambda$ indicate that the infrastructure is used more efficiently. Especially, when one bears in mind the dawning of autonomous vehicles, this interpretation becomes more policy relevant. However, autonomous vehicles might not only change how efficiently the network is used, but could change also other aspects (e.g. capacity). It is now possible to capture the effects on the network efficiency given certain changes in technology.

Second, although not considered yet, the variation of $\lambda$ in space and time can be used as a measure of homogeneity in the network given that the respective shape function remains constant. When we observe only little variation of $\lambda$ between different neighborhoods in a network, arguably, the network can be considered more homogeneous than in the case of a larger variation. Thus, $\lambda$ could potentially be used in network partitioning algorithms, as the objective of such partitions is usually to achieve more homogenous regions.

2.4. Using other shapes of the diagram

The methodology presented above is not restricted to a simplified trapezoidal fundamental diagram, but can use different shapes with combinations of linear and nonlinear congested and uncongested branches. In this section, we briefly discuss three families of functional forms that can serve as shape functions as well. Note that all three families are inspired by the fundamental diagram.

The first family are linear models. Seminal work here was done by Greenshields (1935) using a linear relationship between traffic density and speed. In turn, the relationship between density and traffic flow is quadratic. This relationship has two degrees of freedom, i.e. the free flow speed $u_f$ and the jam density $\kappa$. With free flow speed corresponding to a large extent to the speed limit and jam density corresponding to the inverse of car size, arguably, this function is less capable of reflecting the influence of the design of a city, e.g. intersections limiting the capacity.

The second family are non-linear models (Greenberg, 1959; Underwood, 1961; Drake et al., 1967). This family uses either the logarithm or an exponential form to account for a non-constant interaction between cars with increasing density. Mahmassani et al. (1987) proposed to model the macroscopic relationship between flow and density with such a relationship. In general, for such models it is necessary to estimate certain smoothing parameters. As investigated by Mahmassani et al. (1987), this functional form is capable of capturing the effects of the networks’ geometric features. Unfortunately, this family of functional forms is not capable of describing gridlock at $\kappa$ and the smoothing parameters do not have strong physical meaning.

The third family of functional forms are multi-regime functions. The idea is to use rather simple functional forms, e.g. linear or quadratic functions, for intervals of density. The simplified trapezoidal fundamental diagram used in Section 2 also belongs to this family. Bliemer and Raadsen (2017) introduce a dual-quadratic approach for the fundamental diagram. Here, the uncongested branch of the flow density relationship is modeled by a different functional relationship than the congested branch. A dual regime function ensures that both curves intersect exactly at the point defined by the capacity and the critical density. Such functional relationship accommodates five degrees of freedom, all physically meaningful parameters: free flow speed $u_f$, capacity $Q$, critical density $k_{crit}$, wave speed $w$, as well as jam density $\kappa$. However, this formulation can only accommodate the influence of the network and signal settings in the choice of $Q$ and $k_{crit}$. The method of cuts presented by Daganzo and Geroliminis (2008) can also be considered as a multi-regime stepwise linear function. In this method, the effects of the network design and traffic signal control parameters are considered in the intercepts and slopes of each linear function. However, in cities with adaptive signal control, it is difficult to identify such parameters (Leclercq and Geroliminis, 2013).

We summarize the three families of functional forms in Figure 4. We have chosen the parameters in each functional form such that free flow speed and capacity are identical. We observe that in the uncongested branch all families exhibit a similar trend, but for the congested branch the behavior differs remarkably.

As stated before, the smooth approximation method presented in this paper can be used with the shapes
mentioned above. However, it is clear that the proofs discussed in Section 2.2 should be re-evaluated for every shape function separately. In the next section, we will apply the proposed method with a simplified trapezoidal fundamental diagram as the shape function. Empirical data from four cities is used to illustrate the performance of our functional form in a real world environment and to show the resulting values of $\lambda$.

3. An empirical application

This section demonstrates the proposed methodology on a set of empirical data from four different European cities: Marseille, London, Lucerne, and Zurich. For every city, we collected data from at least one week of measurements from loop detectors, which record occupancy and flows in aggregation intervals between 3 and 5 minutes. Most cities operate a system that detects relatively simple systematic errors automatically. Nonetheless, we further scrutinized the data by identifying outliers (Chen and Liu, 1993) and by reducing noise with a moving average technique (Gómez and Maravall, 1996). The occupancy is converted to density by a linear transformation according to the findings of Geroliminis and Daganzo, 2008; Hall et al., 1992. Furthermore, densities from loop detector data are known to be sensitive to the location of the loop detector within the length of the link (Geroliminis and Daganzo, 2008; Buisson and Ladier, 2009; Leclercq et al., 2014), thus, we apply a recently developed correction method that reduces the bias in the estimation of traffic density (Ambühl et al., 2017).

Here, we investigate the proposed method. First, we study the magnitude of $\lambda$ given a physically meaningful envelope function. Second, we analyze the sensitivity of $\lambda$ and the functional form to variations in the parameters of the envelope function. Third, we econometrically estimate all relevant parameters ($\lambda$ and the rest of the parameters of the shape function, see Eq. 3). This distinction is motivated by the potential applications of the proposed methodology: While the first two aspects focus on measuring the efficiency or performance of a network, the latter one might be more useful for advanced modeling purposes.
For the first two aspects, we choose a simplified trapezoidal fundamental diagram for the shape function as introduced in Section 2. We refrain from using the methodology by Daganzo and Geroliminis (2008) for the shape function, because all four cities have an extensive adaptive traffic control system. However, we acknowledge the findings by Daganzo and Geroliminis (2008) that the free flow and wave speed, and the capacity is reduced in a network due to the delay encountered at intersections. Therefore, we approximate the physical parameters of the macroscopic fundamental diagram by the equations of Daganzo and Geroliminis (2008) based on representative values for cycle length, green time and offset. In the next section we address the issue of uncertainty of these parameters by evaluating the robustness of the estimation with a sensitivity analysis. We estimate $\lambda$ with non-linear least squares in Stata 15.

3.1. Single parameter estimation

All parameters for the simplified trapezoidal macroscopic fundamental diagram for this application are listed in Table 1. We refer the interested reader to Table A.1 in the Appendix with a more detailed description on the calculation of these parameters, which are obtained with equations by Daganzo and Geroliminis (2008). Table 1 also provides the value for the estimated $\lambda$. All estimates are significant at the 1% level of significance.

Table 1: In the upper part of this table the shape parameters for the trapezoidal diagram are given for each city while in the lower part we present the estimates of $\lambda$ for Section 3.1. All estimates are statistically significant at the 1% level of significance.

<table>
<thead>
<tr>
<th>City</th>
<th>Marseille</th>
<th>London</th>
<th>Lucerne</th>
<th>Zurich</th>
</tr>
</thead>
<tbody>
<tr>
<td>$u_f$ [m/s]</td>
<td>9.85</td>
<td>7.71</td>
<td>9.42</td>
<td>7.45</td>
</tr>
<tr>
<td>$w$ [m/s]</td>
<td>1.55</td>
<td>1.58</td>
<td>1.88</td>
<td>1.61</td>
</tr>
<tr>
<td>$\kappa$ [veh/m]</td>
<td>0.150</td>
<td>0.150</td>
<td>0.145</td>
<td>0.145</td>
</tr>
<tr>
<td>$Q$ [veh/s]</td>
<td>0.145</td>
<td>0.168</td>
<td>0.177</td>
<td>0.149</td>
</tr>
</tbody>
</table>

| Estimation results | $\lambda$ | 0.065  | 0.053  | 0.068  | 0.038  |

Figure 5 shows the empirical MFDs for the 4 cities, as well as the proposed functional form. Let us first discuss the empirical MFDs. All cities exhibit a congested branch and thus indicate network-wide congestion. When analyzing the curves of the MFD approximations, we observe that they fit the data well. This shows that the introduced methodology is appropriate as a functional form for the MFD. Regarding the estimated values of $\lambda$ we find that they are less than 0.1 in all cases. It can be seen from Figure 5 that this magnitude of $\lambda$ results in a functional form that satisfies the mathematical properties discussed in Section 2. Now that $\lambda$ is estimated for all four cities using the same underlying trapezoidal shape of the fundamental diagram function, a comparison of the network performance between cities is possible based on a single parameter. Given the parameters in Table 1 we find that Zurich has the smallest value of $\lambda$. This can be seen as evidence of a well operating adaptive control scheme that utilizes much of the available infrastructure. Note, however, that especially in the case of Lucerne it is apparent that the estimated function does not fully capture the empirically observed curve, which emphasizes that other diagram shapes (or parameters) could further improve the fit. As discussed in Section 2 and Equation 4 $\varepsilon$ describes how much the functional form deviates from 0 at $q(0)$ and $q(\kappa)$. The smaller $\varepsilon$, the better this boundary condition is satisfied, which coincides with a low value of $\lambda$. We find that $\varepsilon$ ranges from 0.0001 for Zurich to 0.008 for Marseille. We explain the difference by the difference in the amount of scatter, which is significantly lower in Zurich than Marseille and in turn makes the fit through the mean of the points better. Nevertheless, these values are very small, and can be decreased with less scatter in the MFD or a more accurate choice for the points of the estimation (e.g. removing all observations that exhibit higher levels of inhomogeneity).
An important conclusion is that even in the absence of empirical data, we can assume a range of $\lambda$ between 0.04 and 0.07 for any city, which allows us to estimate the MFD given a technology frontier (shape function/upper bound). In other words, with the proposed methodology, MFDs can be reasonably approximated for any city with only little information.

3.2. Sensitivity analysis

We now account for the uncertainty of the parameters listed in Table 1 by first adjusting their value by ±5%. Using this variance in the data, we perform a sensitivity analysis and shed light on the results in the case where the parameters for the shape function are subject to some errors.

Figure 6 shows the functional form for the four cities. For readability we have excluded the empirical points and we have marked the estimated range of values in dark gray. It is apparent that for the region where data is available the resulting curve changes only slightly, leading to a robust estimation of the functional form of the MFD. The value of capacity as well as critical density, both relevant for traffic control applications, do not change. Nevertheless, the estimates of $\lambda$ change as shown in Figure 7 where the parameters listed
in Table 1 are adjusted by ±5 – 20%. We find that the rank of cities based on λ does not change when we consider the introduced uncertainty. Furthermore, the size of the range for the same uncertainty level remains similar for the four cities (e.g. for the 5% uncertainty the values of λ span a range of 0.03, i.e. ± 0.015 and for the 20% uncertainty, the values of λ span a range of 0.1).

3.3. Full econometric estimation

The fitting of the curve could be further improved, if the parameters of the shape function were estimated instead of a-priori defined. This approach is appealing for advanced modeling applications as used in e.g. Mariotte et al. (2017); Batista et al. (2017). That being said, it is clear that the estimated parameters of the shape must not necessarily follow the physically meaningful boundary of infeasible traffic states anymore (see also Figure 2).

In this empirical context, we estimate λ, u_f, Q using non-linear least squares, while fixing jam density and wave speed to the values of Table 1, because the congested branch is not fully observed in any of the investigated cities. Table 2 shows the estimates for the four cities. For this econometric estimation we use

Figure 6: Estimation results for a sensitivity analysis based on ±5% change in the shape parameters in Table 1. Figure 7 summarized the range of λ for each city with varying uncertainty.
Table 2: Estimation results for the full econometric analysis with jam density and wavespeed fixed because of a lack of data. All estimates are statistically significant at the 1\% level of significance.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Marseille</th>
<th>London</th>
<th>Lucerne</th>
<th>Zurich</th>
</tr>
</thead>
<tbody>
<tr>
<td>$u_f$ [m/s]</td>
<td>9.93</td>
<td>8.67</td>
<td>6.58</td>
<td>7.44</td>
</tr>
<tr>
<td>$Q$ [veh/s]</td>
<td>0.121</td>
<td>0.208</td>
<td>0.165</td>
<td>0.147</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.048</td>
<td>0.073</td>
<td>0.045</td>
<td>0.037</td>
</tr>
</tbody>
</table>

all observed data points, resulting in a regression through the mean of the data. First thing to notice is that all the parameters, including $\lambda$, are within the same order of magnitude as in Section 3.1. However, the larger degree of freedom leads to a better fit through the mean of the observations.

Note that if we were interested in modeling the uppermost observed points in the flow-density relationship, we would choose only these points for the regression (i.e. using a non-linear quantile regression for example).

4. Conclusions

Del Castillo and Benitez (1995) concluded that “the search of a general analytical expression for the speed-density relationship is [...] an open problem” and there are numerous approaches for a functional form between the speed and the density for fundamental diagrams. The MFD, on the other hand, has received only little attention in this context and no generally accepted functional form for the MFD exists.

In this paper, we introduced a new functional form for the MFD based on the smooth approximation of the trapezoidal diagram. Combined with a physically meaningful and appropriate shape function, our proposed functional form results in a concave, continuous and single regime function described by one free parameter. We have shown that the proposed functional form satisfies the mathematical properties for fundamental diagrams, which are equally valid for MFDs. From these properties, we conclude that the proposed functional form is appropriate and simplifies the use of the MFD in many applications. We applied the proposed functional form in an empirical analysis of traffic in four European cities for two different
purposes: (i) measuring the efficiency of the network with respect to the boundary of infeasible traffic states and (ii) fitting the data allowing for a higher degree of freedom and thus offering a better fit to the data for more advanced modeling. We find that the functional form fits the data well for both approaches, validating the proposed methodology. Furthermore, based on the empirical findings, we conclude that the values of the infrastructure potentials $\lambda$ are close to each other across cities. For any city with a defined technology frontier (shape function) and an imputed value of the infrastructure potential $\lambda$ in the range of 0.04-0.07, the MFD can be reasonably estimated without any empirical data and only little information.

Nevertheless, there are some limitations to this study. First, the estimation of $\lambda$ from empirical data is sensitive to the aggregation interval. Larger intervals tend to smoothen traffic states which in turn leads to an increase in $\lambda$. It is clear that the shape function should be adapted to the interval (e.g. average saturation flows tend to be lower when measured over a longer interval). Second, the mathematical properties of the functional form are only guaranteed when $\lambda$ is small. For the discussed shape function and parameters this is not an issue. Last, we have shown that uncertainty in the parameters leads to a shift in the critical density and the congested branch, but not in the uncongested branch. Therefore, it is important to identify the parameters of the shape function carefully. However, it is clear that the value of $\lambda$ should not be compared across different shape functions.

Future research will concentrate on scrutinizing other shape functions and their robustness. Especially the stochastic approximation of the MFD provided by Laval and Castrillón (2015) seems to be a promising framework for the shape function. As autonomous vehicles are expected to increase capacity and speeds (Fagnant and Kockelman 2015), future research can also investigate the relationship between technology changes and $\lambda$. Last, we estimated $\lambda$ using non-linear least squares that fit the curve through the mean of points, but we could also use the concept of quantile regression to fit the curve through the 95th quantile of data (Bates and Watts 2007).

From a policy perspective, the value of $\lambda$ can give interesting insights in how efficiently the infrastructure is used given current characteristics of the roads, the network, routes and vehicle technology, as it measures the gap between the shape resulting from the aforementioned characteristics and the observed MFD. Future research can explore the relationships between urban design and the value of $\lambda$ with a cross-sectional analysis of different cities as it has been done before for the two-fluid theory (Ardekani et al. 1992; Ardekani and Herman 1985). This kind of analysis is relevant for planners as it allows to identify ways to improve the performance of urban road networks.

Acknowledgement

This work was supported by ETH Research Grants ETH-04 15-1 and ETH-27 16-1. We wish to acknowledge the support by Franck Lavergne from Marseille, Ashley Turner and Andy Emmonds from London, Thomas Karrer and Milena Scherrer from Lucerne and Gian Dönier, Urs Birchmeier and Karl Weberruss from Zurich. Finally, we thank Mireia Roca-Riu and Ambra Toletti for their support in this research.

References


Cook, J., 2011. Basic properties of the soft maximum. Working document, Department of Biostatistics, University of Texas, Houston.


Ortigosa, J., Menendez, M., Gayah, V. V., 2015. Analysis of network exit functions for various urban grid network configurations. Transportation Research Record: Journal of the Transportation Research Board 2491, 12–21.


Appendix A. Data information

In this appendix, we summarize the approach to obtain the parameters used for the definition of the envelope, namely the free flow speed, the capacity, the wave speed and the jam density. Table A.1 summarizes the auxiliary parameters needed for this calculation with the equations by Daganzo and Geroliminis (2008). Note that most of these values are provided directly by local experts and traffic signal design guidelines.

We cross-validated our estimates for the free flow speed by considering the average speed limit obtained from OpenStreetMap. We furthermore validated the occupancy-density conversion by comparing the highest observed speed (99th quantile) to the free flow speeds queried from the GoogleMaps API.

Table A.1: Input parameters for the computation of the diagram shape function as defined by Eqn. 3.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Marseille</th>
<th>London</th>
<th>Lucerne</th>
<th>Zurich</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma$ [\text{-}]</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>$u_0$ [m/s]</td>
<td>12.5</td>
<td>12.5</td>
<td>11.1</td>
<td>12.5</td>
</tr>
<tr>
<td>$w_0$ [m/s]</td>
<td>6.5</td>
<td>6.25</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>$s$ [veh/s]</td>
<td>0.51</td>
<td>0.51</td>
<td>0.51</td>
<td>0.51</td>
</tr>
<tr>
<td>$C$ [s]</td>
<td>60</td>
<td>70</td>
<td>90</td>
<td>55</td>
</tr>
<tr>
<td>$G$ [s]</td>
<td>17</td>
<td>23</td>
<td>31</td>
<td>16</td>
</tr>
<tr>
<td>$l$ [m]</td>
<td>280</td>
<td>254</td>
<td>560</td>
<td>284</td>
</tr>
</tbody>
</table>