# Network features and MFD parameters 

Allister Loder<br>Lukas Ambühl<br>Monica Menendez<br>Kay W. Axhausen

# Network features and MFD parameters 

Allister Loder
IVT
ETH Zürich
CH-8093 Zürich
phone: +41-44-633 6258
fax: +41-44-633 1057
allister.loder@ivt.baug.ethz.ch
Monica Menendez
IVT
ETH Zürich
CH-8093 Zürich
phone: +41-44-633 6695
fax: +41-44-633 1057
monica.menendez@ivt.baug.ethz.ch
May 2017

Lukas Ambühl<br>IVT<br>ETH Zürich

CH-8093 Zürich
phone: +41-44-633 3251
fax: +41-44-633 1057
lukas.ambuehl@ivt.baug.ethz.ch
Kay W. Axhausen
IVT
ETH Zürich
CH-8093 Zürich
phone: +41-44-633 3943
fax: +41-44-633 1057
axhausen@ivt.baug.ethz.ch


#### Abstract

While there is a well-studied body of literature on the relationship between the built environment and travel behavior, the relationship between the built environment and the performance of urban road networks is less understood. Macroscopic traffic models, such as the macroscopic fundamental diagram (MFD), model the response of the transport system to travel demand. Many macroscopic models try to establish a functional relationship between the design of urban road network, traffic control schemes and traffic flows or speeds. However, empirical work on linking network design to the macroscopic relationship is scarce and usually limited to a single city. We use traffic data from inductive loop detectors or similar from more than 20 cities worldwide to address this gap from empirical perspective.


In this paper, we first present the available datasets from Bern, London, Madrid and Zurich. Second, we discuss the existing macroscopic traffic flow models and approaches to describe network features. Finally, we present preliminary findings from our ongoing research.

## Keywords

MFD

## 1 Introduction

An extensive body in literature studied the relationship between travel behavior and the built environment (Ewing and Cervero, 2010, Cervero and Kockelman, 1997), whereas it is much less studied between traffic and the built environment. Seminal - but almost the only - work in this context is by Smeed (1961, 1968), Herman and Prigogine (1979), and was re-imitated after thirty years by Daganzo and Geroliminis (2008). Understanding the relationship between traffic and the built environment has key implications for the way we build cities, but, so far, a lack of data, methodologies and tools precluded such analyses at large scale.

Early work by Smeed (1961, 1968) linked available road space to the capacity and trip completion, while - twenty years later - network features such as link length and intersection density were linked to the parameters of the two fluid-theory using empirical data (Ardekani and Herman, 1985, Ardekani et al., 1992, Ayadh, 1986). This theory considers only speed, but neither flow nor density, which are essential variables in traffic control. In this paper, we use empirical traffic data from more than thirty cities around the globe and the recently introduced theory of the macroscopic fundamental diagram (MFD) to address the gap in the relationship between traffic and the built environment. The MFD's consistency with the physics of traffic and large urban scale applicability (Daganzo and Geroliminis, 2008, Geroliminis and Daganzo, 2008) provide a profound framework for studying the relationship between traffic and the built environment.

## 2 Background

In his seminal work, Smeed (1968) was interested in modeling the relationship between city design and traffic performance measured by the number of vehicles that can reach their destinations per hour in a town center as a function of speed, area of the town center and fraction of area dedicated to road infrastructure. He collected empirical data from many British cities from which he concluded that the relationship between network's capacity and area dedicated to traffic is bound between a lower limit and upper limit as shown in Figure 1 from his work. To the best of our knowledge, this is the first representation of macroscopic traffic showing that cities perform differently. Smeed introduced $\tilde{f}$ as a measure how effectively road space is used to account for the differences between cities. However, we found no further work on explaining why differences between cities exist.

The influence of network features on traffic performance was then addressed with the Two-fluid

Figure 1: Smeed's fundamental relationship between speed and flow in urban areas.


Source: Smeed (1968)
theory. In the Two-fluid theory, Herman and Prigogine (1979) divided vehicles in an urban network into two groups: moving and stopped vehicles. The two key assumptions are

- Average journey speed is proportional to the fraction of moving vehicles
- The fraction of time an observer stops in the network is equal to the fraction of stopped vehicles.

The following equation describes the two fluid theory. The two key parameters are $n$ (measures how the network responds to increasing traffic) and $T_{m}$ (corresponds to the free-flow travel time). The two measured variables $T$, the travel time and $T$ the running time.
$\log T_{r}=\frac{1}{n+1} \log T_{m}+\frac{n}{n+1} \log T$
For $n$ and $T_{m}$, Williams (2001) reported on two field studies that linked network features to the two parameters. Ayadh (1986) suggested that the minimum travel time is a function of the average number of lanes, while the measure of traffic performance $n$ depends on the fraction of one-way streets and average block length to block width ratio. Ardekani et al. (1992) suggested that the minimal travel time per unit distance depends on the traffic signal density, average speed limit and the fraction of approaches with good progression, while the traffic performance parameter depends on the fraction of one-way streets, average number of lanes per street, traffic signal density and fraction of actuated signals. Note, however, that the sample size was small
with four cities in Ayadh (1986), whereas nineteen cities / networks have been considered in Ardekani et al. (1992).

Building on the fundamentals of the two-fluid theory, Mahmassani et al. (1987) proposed an exponential relationship between three fundamental variables of traffic: flow $Q$, speed $v$ and density $K$ in the network.

$$
\begin{align*}
& v=v_{f} \exp \left(-\alpha\left(\frac{K}{k_{\text {crit }}}\right)^{d}\right)  \tag{2}\\
& Q=K v_{f} \exp \left(-\alpha\left(\frac{K}{k_{\text {crit }}}\right)^{d}\right)
\end{align*}
$$

$v_{f}, \alpha, k_{c r i t}$ and $d$ are parameters to be estimated. The authors tested these relationship with simulated data, while the empirical evidence and theory were pending and provided twenty years later by Daganzo and Geroliminis (2008) and Geroliminis and Daganzo (2008). In their theory on the Macroscopic fundamental diagram (MFD), a tight upper bound on averages of density and flow for an arterial can be explained by network features and traffic control schemes alone, invariant of demand. The authors concluded that the tight bound for the network MFD equals the bound for a link if, first, demand is slowly-varying and distributed, second, redundant network with many desirable routes, third, homogeneous network with similar links, and, last, the fundamental diagram of links is not significantly affected by turning movements. Arguably, some of these assumptions are not met in reality, although Daganzo and Geroliminis (2008) yield a convincing match between the analytical and simulation and loop detector solution for San Francisco and Yokohama, respectively. Leclercq et al. (2014) showed with simulation for a grid network how the analytical method for different kind of route considerations diverge from the true underlying MFD in the simulation. Daganzo et al. (2011) and Gayah and Daganzo (2011) showed the importance of availability of routes and route choice in the shape of the MFD. Laval and Castrillon (2015) concluded that the MFD's shape is influenced by the ratio of green-time to block length and the mean red to green ratio.

Aside macroscopic flow models, the link between built environment and traffic prediction has also attracted interest. See Sarlas and Axhausen (2016), Hackney et al. (2007) or for some examples Gao et al. (2013).

## 3 Methodology

The objective of this work is to explain and subsequently predict MFD parameters with network features using statistical methods. In a functional modeling approach as Equation 3, MFDs are characterized by a network capacity $Q_{c a p}$ at the critical density $k_{c r i t}$, the free flow speed $v_{f}$, and at least two parameters describing the rate of degeneration of network performance. Network features span a wide range and can basically be divided into two groups: metric and topological measures. In this section, we provide a brief overview on network features, but refer to comprehensive reviews by for example Levinson (2012) and Barthélemy (2011) for more information. At this place, we have to acknowledge the contributions by OpenStreetMap (OSM) and Geoff Boeing; without this data availability and Boeing's 2017 software tool, we could not have made progress in a way we did so far.

In Table 1 we provide an overview of network feature measures we consider at the moment as important. As this is ongoing research, the list is not complete yet. We queried for each city in our sample the road network via the OpenStreetMap API. As we believe that residential and service roads should not be considered for MFD estimation, we estimated the network features first including residential streets and second without residential and service roads. As an example we show in Figure 2 the road hierarchy levels.

| Network feature | Formula | Comment |
| :---: | :---: | :---: |
| Urban area | A | Total area of experimental site, waterays and forests subtracted. |
| Total street network length | $L=\sum_{i}^{N} l_{i} n_{\text {lanes }, i}$ | Using OSM attribut lanes; where no lane information was given, we replaced by the mean of the respective functional road class. |
| Road density | $d_{\text {road }}=\frac{L}{A}$ |  |
| Road network area | $R=\sum_{i=1}^{N} l_{i} n_{\text {lanes }, ~} w_{i}$ | $w_{i}$ is approximated and differentiated by functional road class. |
| Average street length | $\bar{l}=\frac{\sum_{i=1}^{N} l_{i}}{N}$ |  |
| Average number of lanes per street | $\bar{n}_{\text {lanes }}=\frac{\sum l_{i} n_{\text {lanes }, i}}{\sum l_{i}}$ | Using OSM attribut lanes; where no lane information was given, we replaced by the mean of the respective functional road class. |
| Intersection density | $d_{\text {intersection }}=\frac{N_{\text {non-dead-end nodes }}}{A}$ |  |
| Signalized intersection density | $d_{L S A}=\frac{N_{\text {signalized }} \text { nodes }}{}$ |  |
| Average node degree |  | Mean number of edges incident to the nodes. |
| Treeness | $\phi_{\text {tree }}=\frac{L_{\text {tree }}}{L}$ | A tree is defined as a set of connected lines that do not form a complete circuit. |
| Average closeness centrality | $1 / N \sum_{i=1}^{n} \frac{1}{\sum_{y \neq i} d(y, i)}$ | Closeness centrality is for each nodei the reciprocal of the sum of the distance $d(y, i)$ from one node to all other nodes in the graph, weighted by length. |
| Characteristic path length | $l_{p a t h}=\frac{\sum_{i} \sum_{j} d_{i, j}}{N(N-1)}$ | Average shortest path length between two nodes, averaged over all pairs of nodes. |
| Average betweenness centrality | $g=1 / n \sum_{i=1}^{N} \sum_{s \neq t} \frac{\sigma_{s t}(i)}{\sigma s t}$ | For a node $i$, betweenness is the fraction of all shortest paths $\sigma_{s t}$ fro0m $s$ to $t$ that go through this node. |
| Circuity | $C=L_{e} / L_{g c d}$ | Total edge length $L_{e}$ divided by the sum of the great circle $L_{g c d}$ distances between the nodes incident to each edge. |
| Average node connectivity | $\operatorname{con}_{n}$ | Expected number of nodes need to be removed to disconnect a randomly selected pair of non-adjacent nodes |
| Edge connectivity | $\mathrm{con}_{e}$ | Minimum number of edges need to be removed to disconnect the graph |
| Average cycle length of signals | $\bar{c}$ | Information from local experts (experience value) |

Figure 2: Road network of Basel, Switzerland. Colors of the streets follow their functional class. Motorways are given in blue, primary, secondary and tertiary roads are colored in red, all other roads are given in black.


Source: OpenStreetMap

## 4 Data

In this analysis, we use data from stationary traffic detectors: inductive loop (single and double loops), ultrasonic, passive infrared and camera detectors. We believe that this data source is most reliable in terms of a cross-comparison of cities for several reasons. First, using a very similar method to collect traffic data reduces the measurement bias. Second, correction methods can be applied to overcome some spatial biases and yield results close to Edie's method (Leclercq et al., 2014, Ambühl et al., 2017). Third, for the probe method, we believe that for a conclusive cross-comparison similar data sources should be used, e.g. only taxis data, navigation devices, or automated vehicle location devices of busses, but neither is this data available to us, nor is it guaranteed that we can measure and control for all important (unobserved) factors such as
bus lanes, taxis serving mainly certain routes (between the airport and the CBD), and most importantly probe penetration rates are difficult to estimate accurately and their levels differ between cities. However, we acknowledge that loop detectors also face a sample bias if only a subset of the network is sampled. Nevertheless, research has shown that the MFD can be estimated with reasonable accuracy with only a subset of links in the sample Ortigosa et al. (2014). It is clear that additionally, there exists a measurement bias since we deal with empirical data affected by noise and potential errors.

We received the data in many different ways. For few cities, the traffic data could be queried via an application programming interface and stored on our computers, whereas for most cities the data was exported from the traffic management computer and provided by employees of the local transport authority. The traffic management software imposed in some cases restrictions on the volume of exportable data, because it was not designed for mass export of traffic data. For example, some software only allows to export the raw data of a single detector for a single day at a time. With limited time, we had to restrict the exported days to just a few. Most cities provided us with already aggregated data, while Frankfurt and Dresden provided measurements for each vehicle passing a loop.

Each city provided us with information on the localization of each detector. Most cities provided us with construction plans of intersections and roads where the position of each detector was indicated. We digitized each detector in a geographic information system. For computing network-wide average flows and densities weighted by the link length $l$, where the length is defined as the distance on the link from major intersection to major intersection for each detector. We define major intersection when two streets intersect with traffic lights, roundabouts, motorway exits, and if a pedestrian traffic light that have an impact on traffic flow.

Since some difficulties occurred when we tried to obtain the desired information (link length, traffic lights, etc) from maps, e.g. open street map, with an automated routine, we decided to draw each link manually in the geographic information system. In addition, the different data formats made it often impossible to automate the identification of driving directions. To accurately determine the link, we used aerial photography and panoramic scenes from the roads. As we draw unique lines for each link (lane), we were able to automatically identify whether multiple loops cover a link. In the final data set, we attribute each loop detector to a single link or lane. The geographic location of the loops is stored in a .shp or .kml file, which includes:

- The identification number of its associated link.
- The length $l$ of its associated link.
- The position pos in meters from the downstream major intersection.
- The road name (from OSM).
- The functional road class as a measure of the road level hierarchy (from OSM); see next paragraph.
- In case we did not receive the data per lane, but per road, we also attributed the number of lanes to each detector. If possible we used the number of lanes given by the data provider, otherwise we inspected the detector location in a panoramic scene from the road.
- In cities, where we identified multiple detectors per lane and link, we flag the detector, that has the largest distance pos to the downstream major intersection as a part of the sample and disregarded the others.

As literature suggests some of our loops tend to show faulty behavior. Moreover, the implemented automatic fault detection routine from some transport authorities did not always yield robust results. Thus, we inspected each detector's scatter and time series plots in order to remove faulty detectors. Figure 4(c) shows scatter plots of two detectors we considered as well operating; they show similarities to a fundamental diagram; while we consider the loop detectors in Figure $4(\mathrm{~d})$ as incorrect because they exhibit a random scatter. Other faulty detectors for example showed a constant flow or occupancy measurement.

Figure 3: Available data.
(b) Loop covered network in Stuttgart, Germany. All links covered by a loop are highlighted in black, detector locations are given by black dots. All other roads drawn in gray.

(c) Scatter plots of a loop detectors in Bordeaux, France, considered as well operating.

(d) Scatter plots of a loop detectors in Graz, Austria, with some observations considered as false and, subsequently removed from the sample.


Table 2: Sample overview
 For Brisbane, London, Madrid and Singapore we reduced the sample in space to a more manageble size

## 5 Results

In this ongoing research we are still in the stage of preparing the sensor data and road network; at the moment we have not prepared both data sets in a way that allows further comparison. Therefore, we present here the MFDs from the data we already prepared from table 2, Bern, London, Madrid and Zurich. We have created the MFDs for these four cities according to the loops method defined by Leclercq et al. (2014). For each time interval we create the link length weighted mean flow and occupancy. We then apply a local smoothing using a moving mean consisting of the values of two intervals before the actual interval, the value of the actual interval, and the values of two intervals after the actual interval. This allows to remove some noise from the data. Figures $5-8$ show the MFDs with the corresponding sample region besides. Differences are apparent. Bern, the smallest city in the sample does not show any congested branch, whereas London and Zurich, both, have a well defined congested branch of the MFD. Madrid shows only slight signs of congestion. At the same time we record that the average speeds in these cities do decrease significantly. In London, Madrid and Zurich to around $1 / 3$ of the free flow speeds (these ratios were calculated using $\tilde{v}=q / o$, where $\tilde{v}$ is a proxy for speed, $o$ is the occupancy and serves as proxy for density, $q$ is the flow). It is interesting that the decrease in speeds have different effects on the MFD. Note, these are preliminary results and more research is under-going to verify these results and interpretations.

## 6 Outlook

At the moment, we are still preparing the sensor data by adding the location information, identifying the affected link and cleaning the data from errors. If we have this data ready for many cities, we can work on identifying regions within each city suitable for estimating MFD (size, shape, coverage with loop detectors). The importance of this data cleaning is emphasized by the scatter plot in Figure 11 of a loop detector in Madrid. The detector seems to be wrongly connected to the computer. In addition, we are in negotiations for some more cities to participate in our study, thus the sample will most probably increase as well. After identifying the regions of interest for each city, we will use the software tool by Boeing (2017) to obtain the networks from OpenStreetMap and compute the indicators. Subsequently, we can start with our statistical analysis.

Figure 5: MFD for the city of Bern, Switzerland. Region used is shown on the right.

$\qquad$

Figure 6: MFD for the city of London, United Kingdom. Region used is shown on the right.

Days: 3, N=184


Figure 7: MFD for the city of Madrid, Spain. Region used is shown on the right.


Figure 8: MFD for the city of Zurich, Switzerland. Region used is shown on the right.


Figure 9: Speed and occupancy relationship in the urban networks of Bern, London, Madrid and Zurich.


Figure 11: Scatter plot of detector 61062 in Madrid.


## 7 References

Ambühl, L., A. Loder, M. Menendez and K. W. Axhausen (2017) Empirical macroscopic fundamental diagrams: New insights from loop detector and floating car data, Paper presented at the 96th Annual Meeting of the Transportation Research Board, Washington, D.C.

Ardekani, S. A. and R. Herman (1985) A comparison of the quality of traffic service in downtown networks of various cities around the world., Traffic Engineering and Control, 574-581.

Ardekani, S. A., J. C. Williams and S. Bhat (1992) Influence of urban network features on quality of traffic service, Transportation Research Record: Journal of the Transportation Research Board, 1358, 6-12.

Ayadh, M. T. (1986) Influence of the city geometric features on the two fluid model parameters, Master thesis, Virginia Polytechnic Institute and State University.

Barthélemy, M. (2011) Spatial networks, Physics Reports, 499 (1) 1-101.
Boeing, G. (2017) Osmnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks, Manuscript under review.

Cervero, R. and K. Kockelman (1997) Travel demand and the 3ds: Density, diversity, and design, Transportation Research Part D-Transport and Environment, 2 (3) 199-219.

Daganzo, C. F., V. V. Gayah and E. J. Gonzales (2011) Macroscopic relations of urban traffic variables: Bifurcations, multivaluedness and instability, Transportation Research Part B: Methodological, 45 (1) 278 - 288.

Daganzo, C. F. and N. Geroliminis (2008) An analytical approximation for the macroscopic fundamental diagram of urban traffic, Transportation Research Part B: Methodological, 42 (9) 771-781.

Ewing, R. and R. Cervero (2010) Travel and the built environment, Journal of the American Planning Association, 76, 265-294.

Gao, S., Y. Wang, Y. Gao and Y. Liu (2013) Understanding urban traffic-flow characteristics: a rethinking of betweenness centrality, Environment and Planning B: Planning and Design, 40 (1) 135-153.

Gayah, V. V. and C. F. Daganzo (2011) Clockwise hysteresis loops in the macroscopic fundamental diagram: An effect of network instability, Transportation Research Part B: Methodological, 45, 643-655, ISSN 0191-2615.

Geroliminis, N. and C. F. Daganzo (2008) Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings, Transportation Research Part B: Methodological, 42 (9) 759-770.

Hackney, J. K., M. Bernard, S. Bindra and K. W. Axhausen (2007) Predicting road system speeds using spatial structure variables and network characteristics, Journal of Geographical Systems, 9 (4) 397-417.

Herman, R. and I. Prigogine (1979) A two-fluid approach to town traffic, Science, 204, 148-151.
Laval, J. A. and F. Castrillón (2015) Stochastic approximations for the macroscopic fundamental diagram of urban networks, Transportation Research Procedia, 7, 615-630.

Leclercq, L., N. Chiabaut and B. Trinquier (2014) Macroscopic fundamental diagrams: A cross-comparison of estimation methods, Transportation Research Part B: Methodological, 62, 1-12.

Levinson, D. (2012) Network structure and city size, PLoS ONE, 7 (1).
Mahmassani, H., J. C. Williams and R. Herman (1987) Performance of urban traffic networks, in N. Gartner and N. H. M. Wilson (eds.) Proceedings of the 10th International Symposium on Transportation and Traffic Theory, 1-20.

Ortigosa, J., M. Menendez and H. Tapia (2014) Study on the number and location of measurement points for an MFD perimeter control scheme: a case study of Zurich, EURO Journal on Transportation and Logistics, 3 (3-4) 245-266.

Sarlas, G. and K. W. Axhausen (2016) Prediction of aadt on a nationwide network based on an accessibility-weighted centrality measure, Paper presented at the 5th Symposium of the European Association for Research in Transportation (hEART 2016).

Smeed, R. J. (1961) The traffic problem in towns, Statistical Society, Manchester.
Smeed, R. J. (1968) Traffic studies and urban congestion, Journal of Transport Economics and Policy, 2 (1) 33-70.

Williams, J. C. (2001) Macroscopic flow models, in Revised Monograph on Traffic Flow Theory, $1-31$.

