

Empirical MFDs using Google Traffic Data

Knoop, Victor; van Erp, Paul; Leclercq, Ludovic; Hoogendoorn, Serge

DOI

[10.1109/ITSC.2018.8570005](https://doi.org/10.1109/ITSC.2018.8570005)

Publication date

2019

Document Version

Final published version

Published in

2018 IEEE Intelligent Transportation Systems Conference, ITSC 2018

Citation (APA)

Knoop, V., van Erp, P., Leclercq, L., & Hoogendoorn, S. (2019). Empirical MFDs using Google Traffic Data. In *2018 IEEE Intelligent Transportation Systems Conference, ITSC 2018: November 4-7, 2018, Maui, Hawaii, USA* (Vol. 2018-November, pp. 3832-3839). Article 8570005 IEEE. <https://doi.org/10.1109/ITSC.2018.8570005>

Important note

To cite this publication, please use the final published version (if applicable). Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.

Green Open Access added to TU Delft Institutional Repository

'You share, we take care!' – Taverne project

<https://www.openaccess.nl/en/you-share-we-take-care>

Otherwise as indicated in the copyright section: the publisher is the copyright holder of this work and the author uses the Dutch legislation to make this work public.

Empirical MFDs using Google Traffic Data

1 st Victor L. Knoop <i>Transport & Planning</i> <i>TU Delft</i> Delft, the Netherlands v.l.knoop@tudelft.nl	2 nd Paul B.C. van Erp <i>Transport & Planning</i> <i>TU Delft</i> Delft, the Netherlands p.b.c.vanerp@tudelft.nl	3 rd Ludovic Leclercq <i>LICIT, ENTPE, IFSTTAR</i> <i>Univ. of Lyon</i> Lyon, France ludovic.leclercq@ifsttar.fr	4 th Serge P. Hoogendoorn <i>Transport & Planning</i> <i>TU Delft</i> Delft, the Netherlands s.p.hoogendoorn@tudelft.nl
---	--	---	--

Abstract—The **Macroscopic Fundamental Diagram (MFD)** describes the relation between accumulation and speed in a zone. While theoretically expected, empirical validations have been done with limited numbers of floating car (e.g., taxi) data, and loop detectors. This paper will verify existence, shape and crispness of the MFD using Floating Car Data (FCD) from Google. Due to the large amount of users (i.e., high penetration rate), this unique data-set contains traffic information with a high spatial resolution, a high spatial scope and high reliability. We use the data for 3 purposes. First, an MFD for the city of Amsterdam is constructed, revealing a strong relationship between the average density and the average flow. It also shows that the urban road network never reaches its capacity. Secondly, inhomogeneity is analysed. Traffic is well spread over the network, hence inhomogeneity is low. Moreover, if present, the inhomogeneity has only a minor effect on the flow. Also traffic in different directions is homogeneous. Thirdly, for the first time, an MFD is created for a whole country, which turned out to be very crisp as well. This suggests that small areas or a directional split are not needed to create crisp MFDs. That, in turn, implies that a crisp MFD is not a sufficient condition to apply control without considering internal dynamics.

Index Terms—Traffic flow, macroscopic fundamental diagram, urban traffic, traffic congestion, traffic dynamics

I. INTRODUCTION

An area-wide relation between number of vehicles and their speed was already presented in the sixties [1]. In the last decade the concept of zonal traffic relationships has been revitalised, by the work of [2]. The Macroscopic Fundamental Diagram (MFD) relates the average flow for an area to the average density in the area. In this paper, we will consider average density, being the accumulation divided by the road length. Moreover, we will consider average flow, being the production divided by the roadway length [2]. Scaling

Co-sponsored by the Amsterdam Institute of Advanced Metropolitan Solutions, and Google. This study has received funding from the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation program (grant agreement No 646592 MAGnUM project)

production and accumulation with the roadway length makes it easier to compare the MFDs for different areas. Following the generalized definitions of Edie [3], density is the quotient of flow and speed. Measuring speed and average flow hence suffices to construct an MFD (see also [4]). This principle will be used here.

Extensive literature has been developed on traffic control using the MFD: perimeter control has been already suggested by [2], and discussed extensively since, e.g. [5], [6]. Also other control measures like routing have been proposed, e.g. [7]. Apart from a control measure, the MFD can also be used in a large-scale traffic description, e.g. [8], [9]. For both means, it is required that the MFD can be found, and is crisp (i.e., there is not too much scatter).

It was ground-breaking when [10] first showed the MFD for the city of Yokohama. Their data was based on loop detectors and the GPS traces of taxis. The MFDs of several cities have been presented or used afterwards, usually obtained by taxi data or by simulation data, e.g., [5], [11]–[14]. It is unclear for now whether an MFD of the same crispness also holds for city if measured in real life for a percentage of passenger cars. Loop detector data have been used to construct the MFDs, e.g., for a French cities for three time periods [15], and for the Amsterdam ring road [16]. Loop detector data, however, are only installed on the major roads. Moreover, because loop detectors only measure traffic that pass the detector, the loops do not provide an accurate representation of traffic in stopped or stop-and-go traffic. Moreover, the location of the detectors influence the found shape of the MFD [17]. One solution to overcome this is to fuse loop detector data and floating car data [18]. A completely different approach is presented in this paper, exploiting big data in traffic using Google data, and creating an MFD from floating car data collected at large scale.

This paper aims to give insight into traffic relationships on a zonal level. In particular, it shows overall and road-type specific MFDs for the area of Amsterdam, and even the whole Netherlands. In this study we check the

existence of an MFD, and investigate its shape using the new type of data. Scatter might be due to random effects, but also traffic dynamics can play a role [19], [20]. We will explicitly discuss the scatter of the MFD for Amsterdam. We will also discuss the suitable size of a zone, showing MFDs for parts of Amsterdam, but also showing that MFD seems to exist for a area (much) larger than a trip length. This indicates that the basis for an MFD can be a common cause (people leaving at the same time), and is not caused by causality (no traffic influence).

The data used for this research is provided by Google in their Better Cities program [21]. These data have for instance also been used for traffic management purposes for San Francisco [22]. The data themselves will be described in section II-A.

The further organisation is straightforward: first, the research approach will be discussed, followed by the results in section III. Section IV presents the conclusions.

II. RESEARCH APPROACH

In this section, we describe the research approach. We start by the description of the raw data (section II-A), then section II-B describes how the data are manipulated to get the desired information. Finally, section II-C describes the analyses we will carry out.

A. Data description

For this paper, we use floating car data, which in this case are collected by Google. The location data are recorded from any mobile device of which the user opted to enable location services, and has agreed to share his location data. Data come from a range of sensor types: besides GPS, cell phone towers and Wi-Fi detectors are used. There are many mobile devices for which Google collects the location, most notable devices which have Google Maps activated, and Android phones. Their location data are map-matched and their speed is recorded. The data show the transport mode (i.e., bicycle, walking, car, transit). For this research, we will be using only vehicle data. It should be stressed that Google has shared data only in an aggregate way, and applied a differential privacy filter, so individual data cannot be extracted.

For aggregation intervals of 5 minutes, speed and relative flow on a road segment are available. The flow is given as ratio of the flow in the particular aggregation interval and the maximum flow on that segment in the period June 2015 to January 2016. The speed is given as absolute value. The distorting effect of the differential privacy filter is especially strong for low flows. Penetration rates are not available, nor are absolute flow values,

so the penetration rate cannot be computed. Given the estimated amount of users (Android phones, Google Map users), and the availability of data at moments the roads are not used much, we believe this to be high, well over 10%.

The city of Amsterdam has approximately 1 million inhabitants; we consider roads in the metropolitan area between 4.839°E and 4.98°E and between 52.32°N and 52.43°N. The considered roads are freeways or arterial roads. These arterial roads typically have two lanes per direction, traffic lights and a speed limit of 50 km/h. The road segments are classified as freeway or non-freeway segment, based on the mean speed over a full day (including congestion and traffic lights). If this mean speed exceeds 17.5 m/s (63 km/h), the road is considered a freeway, otherwise it is considered a non-freeway. For Amsterdam, this yields 811 freeway segments and 7620 non-freeway segments. Figure 1(b) shows the area and the road segments, colored by type (orange is freeway, green is urban). A manual check of the road types shows that the separation works very well.

Similar data have been obtained for the whole Netherlands. Manual check for other cities have shown that other cities are similar to Amsterdam in type of sampling and representation: similar road types are included (freeways and arterials), and spatial aggregation of the data is similar. The Netherlands have approximately 17 million inhabitants. The Google data provides 25,787 freeway segments and 340,439 nonfreeway segments.

The total data-set has a size of 2.5 TB. Due to export limitation, we will use data for all days of September 2015. Due to the noise for low flows, we select only aggregation periods during daytime, between 7am and 10pm, which include the morning and evening peak. Speeds of 0 or 45 m/s, the extremes of the scale, occur relatively frequent in the data set. This is due to a large noise term, which then is bounded to the scale; these values are unrealistic. Since these extreme speeds affect the averages (and very low speeds highly affect the density), we only include speeds between 0.1 and 40 m/s. Note that we expect to exclude wrong measurements, and still expect to keep congested measurements. This is because in those measurement intervals, there are many vehicles, hence the accuracy is better (better average, and less intentional distortion from the privacy filter). The actual 5-minute average speed is non-zero (not all vehicles stand still for the full 5 minutes), hence we expect a non-zero speed in those intervals.

B. Extracted data

1) *Measuring flow and density: assumptions and scaling:* For each road segment i and each aggregation

period t , the data provides the speed $v_{i,t}$ and the relative flow $\phi_{i,t}$. Moreover, the location of the road segment and the heading (in degrees) is known. To get to the average flow Q and average density K , we should compute:

$$Q_t = \frac{\sum_i l_i q_{i,t}}{\sum_i l_i} \quad (1)$$

$$K_t = \frac{\sum_i l_i q_{i,t} / v_{i,t}}{\sum_i l_i} \quad (2)$$

In this equation, q denotes flow for a segment, and l the relative lane length of a the segment. Note we express the MFD as relationship between average density and average flow. From the flow and density we can derive the average speed in the area:

$$V_t = \frac{Q_t}{K_t} \quad (3)$$

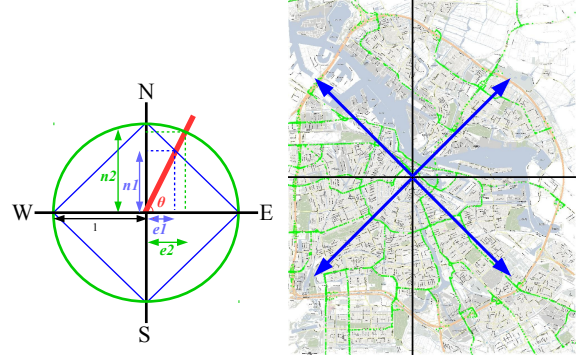
Note that throughout the paper, capital symbols relate to zonal properties and lower case symbols to link properties.

Unfortunately, link lengths are not provided, so to compute the above, we need an estimate. available. We assume that all road segments have the same length. The same assumption has worked in the successful creation of the MFD of Yokohama [10]. Moreover, a manual check of the location of the measurement points showed indeed similar distances for all segments. Furthermore, we assume a maximum flow of 1500 veh/h for the non-freeway sections and a maximum flow of 6000 veh/h for the freeway sections. (Note that these are estimates for a the maximum flow in a 5 minute interval during 6 month on a segment-level. Even if the total network will not reach capacity, on individual links capacity can still be reached.) We now estimate the flow on a road segment by multiplying the relative flow of the segment by the maximum flow of the segment. Then, we compute the average flow (equation 1) and density K (equation 2).

2) *Inhomogeneity*: Literature shows that if the fixed number of vehicles are homogeneously spread within the network, the flow is higher than if they are inhomogeneously spread. A ‘‘Generalised MFD’’ has been created, which relates the average flow as function of the average density and inhomogeneity [20]. We use the standard deviation of density as measure of inhomogeneity (as also in [20], [23]):

$$\Gamma_t = \sqrt{\frac{\left(\sum_i \left[(q_{i,t}/v_{i,t}) - K_t\right]\right)^2}{N}}, \quad (4)$$

in which N denotes the number of links.



(a) Contribution of a road (red) to each of the directions (b) The four quadrants of the city, and the outside direction

Fig. 1. Contributions of directionality

3) *Directionality*: To analyse the effect of directions, the average density and flow per direction are computed. The density (respectively flow) will be split into a contribution per direction.

We will be using four directions of interest, denoted Ω (capital letter because it is a property of the system), i.e. North, East, South and West. Ω is measured from the East, so the $\Omega = 0$ represents the East. We want a link to contribute to at most 2 of the 4 directions, and we want the sum of its contribution to be 1. This way, the sum of the directional contributions is equal to 1, and thus the sum of the directional density (or flow) equals the sum of the density (or flow).

Consider the construction of the directional contribution in figure 1(a). Note that by design (sum of the contributions is 1), $e1 + n1 = 1$. For all points on the blue square it holds that the sum of the directional contribution is 1. The intersection of the square with the road (under angle θ) with the square give the desired contributions to the East and North ($e1$ and $n1$), indicated by the dashed blue line.

Let us consider the contribution to the East (i.e., consider $\Omega=0$) of a direction, indicated $e1$ in figure 1(a). We can compute so by rescaling $e2$, i.e., the East component of the intersection of the road with a circle with radius 1 (green solid line). Since the shape of the dotted rectangles are the same, we know $\frac{e1}{e2} = \frac{n1}{n2}$. The reader can verify that combined with $e1 + n1 = 1$, this yields

$$e1 = \frac{e2}{n2 + e2} \quad (5)$$

Moreover, $e2$ and $n2$ are known: $e2 = \cos(\theta)$ and $n2 = \sin(\theta)$.

For a general direction of interest, we find the directional contribution of a link under an angle θ to direction

Ω :

$$d_{\Omega,i} = \begin{cases} \frac{\cos(\Omega - \theta_i)}{\sin(\Omega - \theta_i) + \cos(\Omega - \theta_i)} & \text{if } |\Omega - \theta_i| < \pi/2 \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

We use these fractions for each link to compute the directional average density and directional average flow, by multiplying the density respectively the flow with the directionality component d .

C. Analyses

This section describes the analyses which will be performed. The first part describes constructing an MFD. Then, it is described how the (effect of) inhomogeneity of traffic is analysed. Section II-C3 describes how the effect of directionality can be analysed. Finally, section II-C4 describes how we check the effect of area size on the existence and crispness of the MFD.

1) *Macroscopic fundamental diagram*: First, the MFDs will be presented, which show the relationship between the average density and average flow. The first analyses are done for the network of Amsterdam, which is considered to be one network, with trips crossing all through the network. These analyses show the existence and scatter of the MFD. It will also show to which extent the average flow reduces for higher average densities. These reductions are due to the fact that due to other congestion, some road segments cannot be loaded to capacity. The extent to which the average flow reduces is hence related to amount of internal spillbacks. This indicates the quality of the urban traffic control in Amsterdam, and the potential for perimeter control.

Besides, it is tested whether an MFD holds for a network much larger than the typical trip length, in our case the country. Especially the local roads do not form a proper infrastructure for these trips. However, the motorway network is relatively dense, and longer trips are made on this network. Therefore, as exploratory research, a MFD is constructed for all roads in the Netherlands, and for the freeway network. If a MFD holds, that means that we can use the description for a network being (much) larger than the trip length. The fact that there is an MFD does not imply causality between the crowdedness at different parts of the network. It could be caused by the fact that the network is crowded *at the same times*.

2) *Inhomogeneity of traffic*: The GMFD (generalised macroscopic fundamental diagram, [20]) shows the effect of inhomogeneity in the traffic. Homogeneity means that traffic is equally spread over the city, and inhomogeneity means traffic is concentrated at several bottleneck

locations. In this paper we study to which extent inhomogeneity causes a reduction in average flow.

To quantify, a two-variable fit of an MFD is made: the average flow as function of average density and inhomogeneity Γ . For the base MFD, the Drake fundamental diagram [24] is considered to estimate the average flow (\tilde{Q}), supported by results in [25]. However, to include skewness, average density is first transformed non-linearly, using a power-transformation. The power (parameter Z) indicates the skewness. The inhomogeneity is included as linear decrease with the standard deviation of density, following [20]; see [23] for more background. All together, the fitted MFD is described by the following equations:

$$K' = K^Z \quad (7)$$

$$K'_{\text{crit}} = K_{\text{crit}}^Z \quad (8)$$

$$V = V_0 * \exp(-1/2 * K'/K'_{\text{crit}}) \quad (9)$$

$$\tilde{Q} = K'V - B\Gamma \quad (10)$$

The GMFD is hence described by 4 parameters: a skewness power parameter Z , a density scaling parameter K_{crit} , a speed scaling parameter V_0 and the influence of the inhomogeneity B .

Beside the impact of the inhomogeneity, we quantify the inhomogeneity itself, and relate it to average density. This will show to which extent the potential problem of nucleation (i.e., a part of the network is congested and attracting more congestion, whereas the other part is in free flow conditions) is a problem in practice.

3) *Directionality*: We want to analyse the effect of directions. When the MFD was introduced [10], it was claimed to work in homogeneous conditions. However, typically the roads have a difference in load per direction, so this assumption does not hold. This section will analyse to which extent directionality has an effect. For all analyses in this section and the following section, we use data limited to urban roads. This is because we do not want the analysis to be largely determined by through traffic on the freeways, which in an analysis of mixed roads have a high impact due to the high capacity of these roads.

The first analysis on directions investigates to which extent the loads for the various directions (North, East, South, West) are correlated. This can give information whether a zonal description incorporating all directions is reasonable, or that a directional description per zone would better suit traffic dynamics. To this end, we will analyse the correlations in average density per direction. If the density per direction is not strongly correlated, an MFD as function of the average density in 4 directions (one for each direction) might be better to use

as explanatory variables. Similarly, if flow for different directions is not strongly correlated, a separate MFD might be fitted for each direction.

For correlations, we use a Pearson's correlation. We will be reporting the Pearson correlation coefficient, ranging from -1 if there is a perfect negative linear correlation (y decreases linearly with x) to +1 if there is a perfect positive linear correlation (y increases linearly with x). Note that correlation indicates to which extent x and y are linearly dependent; it cannot reveal causality.

The second analysis, directions are chosen relative to going towards the center. Ideally, one would like to assess the directionality of a link and split this into inflow towards the center and outflow from the center (perhaps combined with clockwise and counter-clockwise flows, which here are mainly the ring road freeway, and therefore excluded). Unfortunately, this could technically not be achieved. Instead, the city was divided into four zones (see figure 1(b)), and an outflow direction was set at 45 degrees, and the midpoint being the city center. The outflow direction is traffic into the direction of the arrow. Links are assessed using directional contributions (as described in section II-B3 (equation 6) and shown in figure 1(a)), under a direction of interest (Ω) offset by 45 degrees. Similarly, the outbound contribution can be determined.

As example, an eastbound road in the North-East quadrant contributes 50% to the outbound traffic and 50% to the clockwise traffic. An eastbound road in the South-East quadrant contributes 50% to the outbound traffic, and 50% to counter-clockwise traffic. This method works well for links close to the arrow, and gets less accurate the closer the links are to the edges of the quadrant (i.e., closer to the black lines in figure 1(b)).

The average flow towards the city center ("inflow") is the average of the flow towards the city center in all four quadrants. The average flow outside of the city center ("outflow") is the average of the flow away from the city center in all four quadrants. We will check the correlation between the density for the inbound traffic and the density for the outbound traffic at a particular time. For a city center with an inflow peak in the morning and an outflow peak in the afternoon, one would expect a low (or even negative) correlation. The correlation coefficient can be compared with the correlations for average flow in each zone. This way, it becomes clear whether a separation by flow direction (aggregated over the whole town) will yield a more homogeneous zone, or a separation by geographical area (aggregated over all directions).

Both analyses are done for flow as well as density.

4) *Area size*: This analysis consists of checking the correlations of all (urban) roads within a area. It can be thought that areas should not be too large since inhomogeneities occur. We will therefore check the correlation for smaller and larger geographical areas as well.

We will therefore compute the correlation between average density and average flow between the four quadrants in Amsterdam as introduced in the section above. This will give 6 correlation coefficients (NE-SE, NE-SW, NE-NW, SE-SW, SE-NW, SW-NW), of which the average will be taken as representative for the correlation between two quadrants.

The second analysis is the correlations for a larger area. We check the correlation levels between the traffic states the Amsterdam area and the Netherlands. Indeed, the Netherlands are too large to consider it one traffic system. Correlations will at most reveal similarities in traffic state, and no causality. A high correlation coefficient shows that at the same time, the traffic state fluctuates in the same way.

Here again, both correlations are computed for average flow and average density.

III. RESULTS

This section presents the results, organised along the same way as the analyses. Hence, section III-A presents the MFDs for Amsterdam and the Netherlands, section III-B presents the (effect of) inhomogeneity, section III-C presents the results of the directionality and section III-D presents the effects of the area size on the MFD.

A. *Macroscopic fundamental diagrams*

Figure 2 shows the MFDs (a-c). The lines show the medians as well as the 17.5 and 82.5 percentiles, meaning that 65% of the measurements lies within the points. For a normal distribution, these percentiles match the the mean plus or minus a standard deviation. The lines are very crisp, showing that the average density is well explaining the average flow. This is in line with earlier research. Remarkable is that the flow does not decrease at higher densities. Using their traffic management strategies, the Amsterdam municipality largely succeeds in avoiding spillbacks and related loss in average flow. The figure also shows fits according to equations 8-10, which are well in line with the mean. Note that although the flow is increasing with the density, and hence the zone is not congested in the traditional interpretation of the word, speeds are reducing considerably – see figure 2 (d-f). The low speeds at low densities are most likely due to added noise for low densities.

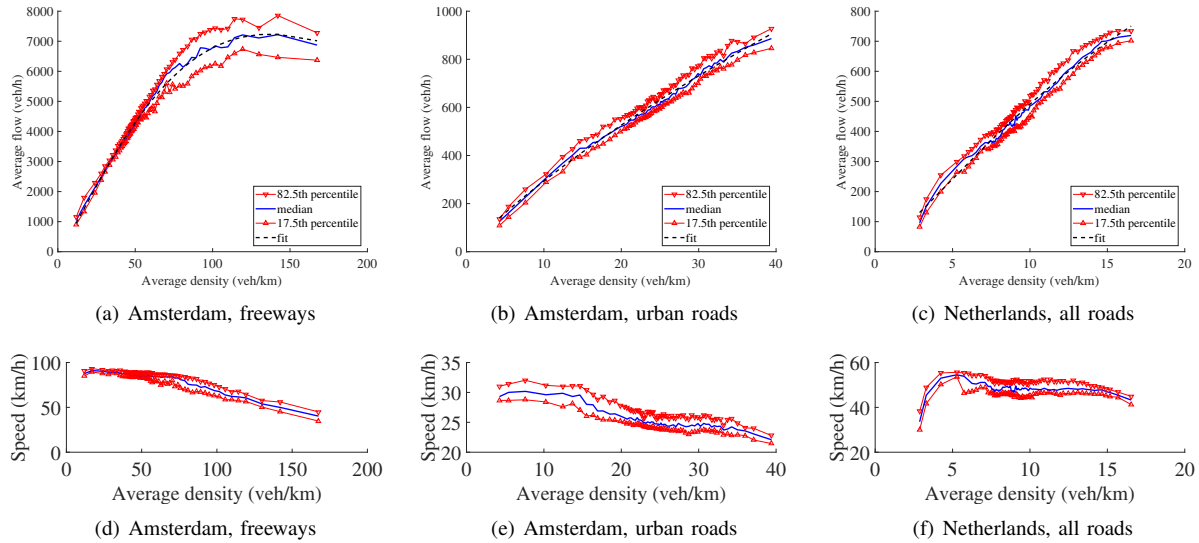


Fig. 2. MFDs for different roads. The fits are made using of equations 8-10

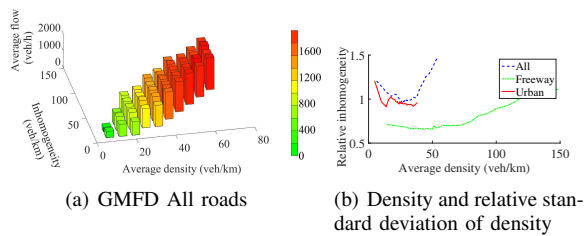


Fig. 3. Effect of inhomogeneity for the network of Amsterdam

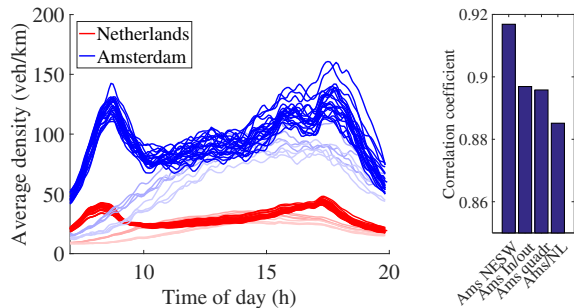
Figure 2(c) show the MFD for the whole Netherlands. Surprisingly, the relationship between the average density and average flow is also very crisp in this case. This does not imply there is causal relation between the queues in different parts of the network, or drivers need to share the same physical infrastructure. A likely reason is that there is temporal correlation between the loads at the different parts of the network, and this is not necessarily caused by the same drivers crossing the whole network. For example, the speed in the north of the country might be correlated with the density if the peak hour is at the same time. That does not mean that the drivers in both parts of the country physically influence each other. Also here, the functional form is well able to describe the pattern.

B. Relation between average density and spread in density

The inhomogeneity has an influence on the average flow: for the same average density, each unit of standard

deviation reduces the average flow. Figure 3 shows the Generalised MFD (GMFD). The main influencing factor is the average density; the standard deviation of density plays a minor role. The effect of the standard deviation differs. The slope of GMFD, in terms of increasing flow per increasing inhomogeneity (i.e., parameter B in equation 10), varies from -1 to -20 (veh/h)/(veh/km). Given that flow values are much higher than (spread in) density, this means the MFD is relatively flat as function of inhomogeneity. This holds also for the more congested freeway-only network.

Inhomogeneity itself is closely related to average density; there is an almost linear relationship. To visualise this better than the proportional increase of the inhomogeneity as function of average density, we show the inhomogeneity divided by the average density, i.e. the relative inhomogeneity. Figure 3(b) shows that the inhomogeneity varies between 0.7 and 1.5 times the average density. As function of average density, the relative inhomogeneity first decreases and then increases. The increase for high average densities means that the network becomes more and more heterogeneous when loads increase. The fact that inhomogeneity depends on density means that the MFD is not necessarily expressed a function of both variables, but one can be derived from the other. Both the strong dependency (i.e., the standard deviation of density depends on the average density) as well as the limited effect (low absolute value for B) suggest that a fitting a two-variable MFD is not essential.



(a) Average density over time; Saturdays are shaded a bit lighter, Sundays are shaded much lighter
 (b) Correlation coefficients

Fig. 4. Correlations

C. Directionality

Figure 4(a) shows the average density over time, showing clear peaks at various days. This is found to be largely the same for various directions: the correlation coefficients between the average densities in the various directions are generally very high, 92% over all roads, 96% over the urban roads and 81% over the freeways. It is remarkable that the correlation for the urban roads is higher than of all roads. That implies that the urban network can be considered as one entity, more so than the freeway network, even though the travel time over the urban roads is higher.

Since the correlations are this high, the average densities for 4 different directions cannot be considered independent. Hence, describing the average flow as function of the average density in each of the directions is not useful. The fact that the correlation coefficient is lowest for the freeways, suggests that in peak hours, buffer capacity remains in some directions, whereas the urban network gets congested in all directions.

The most traditional situation is a flow into the city center in the morning and out of the city center in the evening peak. The coefficients for correlation between average density in the inflow and outflow directions is 0.90. That means that there is not a strong tidal effect. Instead of the traditional situation, people go in all directions. This could be due to people living in the city center and moving out, but this is unlikely given low car ownership and good public transport possibilities in the city center. More likely is that there are travelers that live in the north of Amsterdam, or north of Amsterdam, that work in the south travel across town, and vice versa for other directions.

D. Area size

The correlations in the different quadrants of Amsterdam are on average (over all combinations between quadrants) 0.90. The area of the quadrants is smaller than Amsterdam as a whole, however, the correlations are similar in value to the correlations for Amsterdam as a whole (at 0.92). This indicates that the smaller area size does not show higher correlations, or a more homogeneous traffic state.

Next, we analyse the correlation between the traffic state in Amsterdam and the traffic state in the Netherlands. Interestingly, these correlations are also high, as shown in the fourth bar. Although the whole country, and especially the urban roads considered here, cannot be seen as one system (no vehicles will traverse a considerable part of the country on urban roads), the correlations are high. This means that busier periods occur simultaneously in various cities. That the congestion starts at the same times at various locations in the country is also shown by the function of average density over time, see figure 4(a). Note thus that correlations should not be seen as causality, but in this case clearly only indicates simultaneous occurrence of events.

IV. DISCUSSION AND CONCLUSIONS

In this paper we have constructed MFDs using large amounts of floating car data of Google. The Google dataset allows us to analyse data for different types of roads. The data shows that traffic shows a crisp fundamental diagram based on a considerable fraction of all cars. Compared to earlier approaches, where MFDs have been shown using a limited number of floating car data and/or loop detector data, the empirical validation of an MFD using floating car data of private vehicles is scientifically a major step.

With regard to the traffic conditions, it has been found that the traffic in Amsterdam is operating at a reduced speed during some times of the day, but not yet to the level that it reduces the average flow. In fact, in urban roads the capacity of the network is not yet reached: a further increase of traffic average density (a higher traffic load) would still cause for a higher average flow. The freeway network, however, is operating at the top of its capacity.

The MFD is a sensible way to describe the traffic state in a network. The inhomogeneity in the network is directly related to the average density. A two-variable MFD as proposed by [20] is not needed for recurrent conditions in normal operating conditions, since one variable (the inhomogeneity) can be directly derived from the other (the average density). A fit of the average flow as function of both variables shows that the

inhomogeneity has a minor effect on the average flow. If inhomogeneity could be independently controlled, the increase of average flow would be minor.

Directionality did not play a major role either: different directions (North, East, South and West) or inbound and outbound traffic were well balanced.

Finally, we found traffic conditions are relatively homogeneous for the Amsterdam network, and, remarkably, for a large networks as well, including the network of the whole country. We found that the MFD for the country of the Netherlands is crisp. This is due to the fact that the loading and unloading pattern of the network is similar in time. So a crisp MFD is not necessarily related to a small area of interacting users. However, since there are no interacting users, this MFD should not be used for instance perimeter control. This combination shows that an area with a crisp MFD is not necessarily suited to apply network-wide control without considering the internal traffic operations.

Further analyses by the municipality could investigate the effects at dedicated days (e.g., Christmas) on particular parts or directions in the network. Furthermore, a real-time use of this type of data (including information to road users) is an subject to consider.

REFERENCES

- [1] J. Godfrey, "The mechanism of a road network," *Traffic Engineering and Control*, vol. 11, no. 7, pp. 323–327, 1969.
- [2] C. Daganzo, "Urban gridlock: Macroscopic modeling and mitigation approaches," *Transportation Research Part B: Methodological*, vol. 41, no. 1, pp. 49–62, 2007.
- [3] L. Edie, "Discussion of traffic stream measurements and definitions," in *Proceedings of the Second International Symposium on the Theory of Traffic Flow*, OECD, Paris, France, 1965.
- [4] J. Du, H. Rakha, and V. V. Gayah, "Deriving macroscopic fundamental diagrams from probe data: Issues and proposed solutions," *Transportation Research Part C: Emerging Technologies*, vol. 66, pp. 136–149, 2016.
- [5] N. Geroliminis, J. Haddad, and M. Ramezani, "Optimal perimeter control for two urban regions with macroscopic fundamental diagrams: A model predictive approach," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 14, no. 1, pp. 348–359, 2012.
- [6] M. Keyvan-Ekbatani, M. Papageorgiou, , and V. Knoop, "Controller design for gating traffic control in presence of time-delay in urban road networks," *Transportation Research part C*, vol. 59, pp. 308–322, 2015.
- [7] M. Hajiahmadi, V. L. Knoop, B. De Schutter, and H. Hellendoorn, "Optimal dynamic route guidance: A model predictive approach using the macroscopic fundamental diagram," in *Intelligent Transportation Systems-(ITSC), 2013 16th International IEEE Conference on*. IEEE, 2013, pp. 1022–1028.
- [8] M. Yildirimoglu and N. Geroliminis, "Approximating dynamic equilibrium conditions with macroscopic fundamental diagrams," *Transportation Research Part B: Methodological*, vol. 70, pp. 186–200, 2014.
- [9] V. L. Knoop and S. P. Hoogendoorn, "An area-aggregated dynamic traffic simulation model," *EJTIR*, vol. 15, no. 2, pp. 226–242, 2015.
- [10] N. Geroliminis and C. F. Daganzo, "Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings," *Transportation Research Part B: Methodological*, vol. 42, no. 9, pp. 759–770, 2008.
- [11] Y. Ji, W. Daamen, S. Hoogendoorn, S. Hoogendoorn-Lanser, and X. Qian, "Macroscopic fundamental diagram: investigating its shape using simulation data," *Transportation Research Record, Journal of the Transportation Research Board*, vol. 2161, pp. 42–48., 2010.
- [12] J. Ortigosa, M. Menendez, and H. Tapia, "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*, vol. 3, no. 3-4, pp. 245–266, 2014.
- [13] M. Saberi, H. Mahmassani, T. Hou, and A. Zockaie, "Estimating network fundamental diagram using three-dimensional vehicle trajectories: Extending edie's definitions of traffic flow variables to networks," *Transportation Research Record: Journal of the Transportation Research Board*, no. 2422, pp. 12–20, 2014.
- [14] J. A. Laval and F. Castrillón, "Stochastic approximations for the macroscopic fundamental diagram of urban networks," *Transportation Research Part B: Methodological*, vol. 81, pp. 904–916, 2015.
- [15] C. Buisson and C. Ladier, "Exploring the impact of the homogeneity of traffic measurements on the existence of the macroscopic fundamental diagram," *Transportation Research Records, Journal of the Transportation Research Board*, 2009.
- [16] V. L. Knoop and S. P. Hoogendoorn, "Empirics of a generalized macroscopic fundamental diagram for urban freeways," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2391, no. 1, pp. 133–141, 2013.
- [17] T. Courbon and L. Leclercq, "Cross-comparison of macroscopic fundamental diagram estimation methods," *Procedia - Social and Behavioral Sciences*, vol. 20, no. 0, pp. 417 – 426, 2011.
- [18] L. Ambühl and M. Menendez, "Data fusion algorithm for macroscopic fundamental diagram estimation," *Transportation Research Part C: Emerging Technologies*, vol. 71, pp. 184–197, 2016.
- [19] V. V. Gayah and C. F. Daganzo, "Clockwise hysteresis loops in the macroscopic fundamental diagram: An effect of network instability," *Transportation Research Part B: Methodological*, vol. 45, no. 4, pp. 643 – 655, 2011.
- [20] V. L. Knoop, H. van Lint, and S. P. Hoogendoorn, "Traffic dynamics: Its impact on the macroscopic fundamental diagram," *Physica A: Statistical Mechanics and its Applications*, vol. 438, pp. 236–250, 2015.
- [21] A. Erland, Tech. Rep., 2015, <https://europe.googleblog.com/2015/11/tackling-urban-mobility-with-technology.html>, accessed 21 July 2017.
- [22] B. Sana, J. Castiglione, D. Cooper, and D. Tischler, "Using google's aggregate and anonymized trip data to support freeway corridor management planning in san francisco," in *Proceedings of the Annual Meeting of the Transportation Research Board*, 2017.
- [23] S. P. Hoogendoorn, W. Daamen, V. L. Knoop, J. Steenbakkers, and M. Sarvi, "Macroscopic fundamental diagram for pedestrian networks: theory and applications," *Transportation Research Procedia*, vol. 23, pp. 480–496, 2017.
- [24] J. S. Drake, J. L. Schöfer, and A. May, "A statistical analysis of speed density hypotheses," in *Proceedings of the Third International Symposium on the Theory of Traffic Flow*, L. C. Edie, R. Herman, and R. Rothery, Eds. Elsevier North-Holland, New York, 1967.
- [25] V. L. Knoop, S. P. Hoogendoorn, and J. W. C. Van Lint, "Routing strategies based on macroscopic fundamental diagram," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2315, no. 1, pp. 1–10, 2012.