

This paper is also published in a journal. Please cite instead:  
Knoop, V.L. and S.P. Hoogendoorn (2013) Empirics of a Generalised Macroscopic  
Fundamental Diagram for Urban Freeways,  
Transportation Research Records No. 2391, p. 133-141

# Empirics of a Generalized Macroscopic Fundamental Diagram for Urban Freeways

V. L. Knoop PhD  
Delft University of Technology  
Transport & Planning  
Stevinweg 1  
Delft, The Netherlands  
+31 15 278 8413  
v.l.knoop@tudelft.nl

S.P. Hoogendoorn PhD  
Delft University of Technology  
Transport & Planning  
Stevinweg 1  
Delft, The Netherlands  
s.p.hoogendoorn@tudelft.nl

October 25, 2012

Word count:

nr of words in abstract	240
nr of words (including abstract)	4392
nr of figures& tables	12* 250 = 3000
total	7392

Submitted to the 92th Annual Meeting of the Transportation Research Board

## ABSTRACT

Due to increasing exchange of data between measurement sites, the area over which traffic control is applied increases. This leads to three new challenges: (1) working with the large quantities of data (transmit, store), (2) estimating the traffic state, (3) controlling a large area with many controllers (and hence large solution space). This paper introduces a new way of describing the traffic state for a large area, which requires much less data and nevertheless gives an accurate representation of the state.

The Macroscopic Fundamental Diagram (MFD) which links the production, i.e. the average flow, to the accumulation, i.e. the average number of vehicles in an area. This paper shows that this can be generalized to a *generalized macroscopic fundamental diagram* (GMFD) for urban freeways, which relates the production to the accumulation and the spatial spread of density. Analysis of 10 months of data of the Amsterdam ring road freeway shows that this is a continuous function, which increases and decreases with accumulation like a fundamental diagram, and decreases with the spatial spread of density. The predictive performance is tested by using a non-parametrized fit and by fitting a functional form, which perform equally well. Predicting the production is important especially near the maximum production. The GMFD explains much more of the spread in the production than the MFD, especially near this maximum production. Thus, this lean traffic state description can be used in a setting a target for traffic control.

# 1 INTRODUCTION

Traffic congestion is a serious problem in urbanised areas because of delays and air pollution. Traffic control measures are being developed to intervene with the traffic operations and to improve the traffic situations. However, with increasing spatial scope of traffic control and increasing number of control variables, the joint optimisation of all traffic flow measures is difficult. This is for instance due to large data sets which need to be processed, but also the large solution space. For more information on traffic measures and the requirements, we refer to (1).

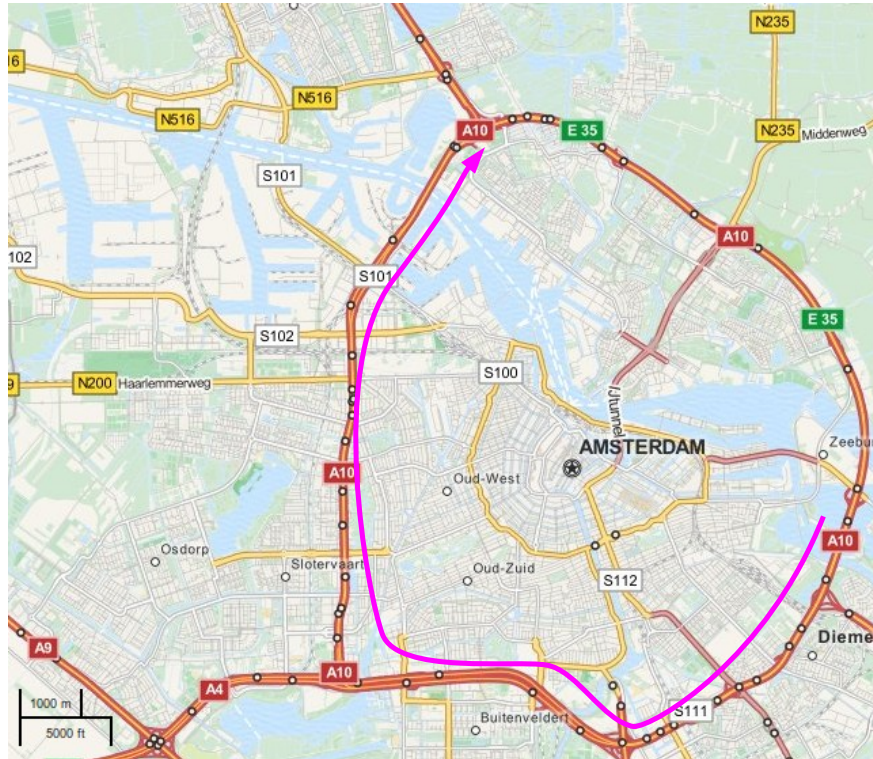
After earlier publications (e.g., (2, 3)), Daganzo (4) re-introduced the concept of the Macroscopic Fundamental Diagram (MFD) in modern traffic engineering, which relates the number of vehicles in an area to the average flow in the area. This potentially simplifies traffic control enormously. Instead of a detailed description of the traffic state, one variable, the number of vehicles in the area, is sufficient to describe the traffic state. Geroliminis and Daganzo (5) show that the concept works in practice for traffic in the center of a major city. It is claimed, though, that the concept of the MFD only works for homogeneous traffic conditions.

Ji et al. (6) studied the combination of an urban arterial and an urban freeway. During congestion build-up and congestion resolution, the MFD found for equilibrium conditions did not hold. Buisson and Ladier (7) focussed specifically at non-equilibrium conditions, analysing the traffic state of an urban freeway and the main arterials during a strike. They conclude that in case of inhomogeneous traffic conditions, the MFD does not longer hold. Similar findings are done for urban networks, for instance by Mazlounian et al. (8). They propose quantifying the spatial spread of density by the standard deviation of the densities, and show that for various values of this spatial spread of density, the MFD had a different shape. They found so in a simulation of a urban grid network, and they related the impact to the turning movements of the individual vehicles.

Cassidy et al. (9) studied the empirical MFDs for freeway networks. In this paper, we consider only urban freeways in the network. Cassidy et al. (9) explain that due to averaging states from different branches of an idealised, triangular fundamental diagram leads to a lower production than the production on a triangular fundamental diagram corresponding to the average traffic state. He also showed using data from a short stretch of freeway that indeed the top of the triangular fundamental diagram is not reached. Recently, Saberi and Mahmassani (10) investigated the network dynamics of the variation in density, checking data and phenomena. They conclude that hysteresis loops in the network are due to the dynamics of the traffic.

This paper describes the effect of inhomogeneity on the production for an (urban) *freeway* network. In fact, we aim to come with a simple quantitative description of the traffic production as function of the accumulation and the spatial spread of density. This we will call a *generalized macroscopic fundamental diagram* (GMFD). This relationship can be used in traffic control systems which aim for a high production, so therefore it is critical that the maximum production is well represented. We will find this relationship using 10 months of data of the A10 freeway ring road around Amsterdam.

The next section describes the data used for the analysis. Then, section 3 describes the data processing. Section 4 presents the macroscopic fundamental diagram, which links the production to the accumulation; for the data set, this shows some scatter. Section 5 presents the generalized macroscopic fundamental diagram for urban freeways, describing the production as continuous function of the accumulation and the spatial spread of density. In section 6 it is determined how good the predictions are if the traffic state is predicted on only aggregated data, both using the MFD and using the GMFD. Both non-parametrized and functional forms are studied. Section 7 compares the results and 8, finally, will present the conclusions.



**FIGURE 1 The studied ring road**

## **2 STUDY SITE: AMSTERDAM RING ROAD**

This section describes the data used for this study. First, the study site is described, and then the collected data in section 2.2.

### **2.1 A10 ring road**

In this paper we use data from the A10 freeway, the freeway around the capital city Amsterdam. Figure 1 shows a map of the area. We will consider a 22 km long section, from the Zeeburger tunnel at km 10 to intersection Coenplein at km 32 in the counter-clockwise direction. This part is indicated with an arrow in the figure. There are 4 freeway intersection and 13 on and off ramps in the section, all indicated with white dots. The road has different sections with a different number of lanes. Table 1 summarizes these characteristics.

A large part of the stretch has a speed limit of 100 km/h. From km 22 to km 28, the speed limit is 80 km/h which is checked by an average speed check, i.e. the average speed on that stretch is enforced, and drivers will be fined automatically if they exceed the speed limit. Due to this average speed check, the fraction of drivers speeding is very low (11). The difference in speed is not expected to play an important role, since the fraction of traffic in the stretch with the lower speed limit (i.e., the number of cars in that stretch relative to all cars in the network) is approximately the same. We will therefore consider the whole stretch as one area.

**TABLE 1 Key characteristics of the data analysis**

Length	22km
Nr of intersections	13
Nr of cross-sections	56(+3 excluded)
Nr of lane specific detectors	179(+9 excluded)
Nr of cross-sections with 2 lanes	18
Nr of cross-sections with 3 lanes	17(+3 excluded)
Nr of cross-sections with 4 lanes	13
Nr of cross-sections with 8 lanes	
Time period	1 January 2011 - 20 October 2011
Time of day	6 am – 8 pm
Nr of days included	271
Aggregation time	1 minute
Nr of aggregation intervals	227,911
Nr of intervals used for calibration	151,941
Nr of intervals used for validation	75,970

## 2.2 Data availability

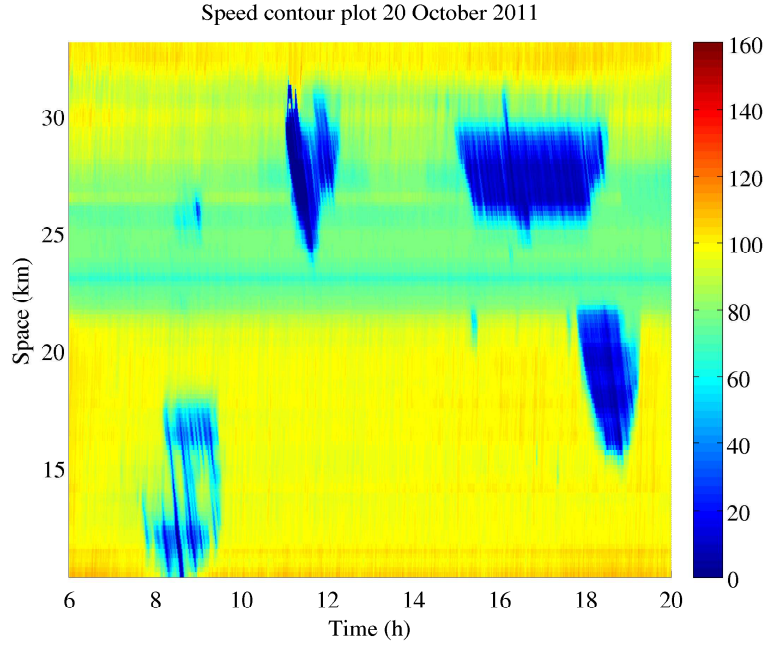
The area is equipped with double loop detectors approximately every 500 meters, and more in some parts. In the section, there are 188 lane-specific detectors distributed over 59 cross-sections. These produce the flow and time mean speed for intervals of one minute.

The period for which data is available is 1 January 2011 to 21 October 2011. From these data, the weekend days are removed. Furthermore, we only consider the period from 6.00 am to 8.00 pm. In this section, at two cross-sections the detector data is not available due to technical errors, and at one cross-section, there is a consistent over-estimation of the speed. Because the data from these detectors are unavailable or incorrect, they are removed from the data set. Apart from that, no data has been removed, so incidents and all weather conditions are all in the data set. Figure 2 shows a typical pattern of the traffic operations on the road in a speed contour plot. Note the effect of the different speed limits in the figure: from km 22 to 28, the speed is lower due to the lower speed limit and the average speed check. Furthermore, we see typical queuing patterns upstream of localised bottlenecks, which mainly are the freeway intersections.

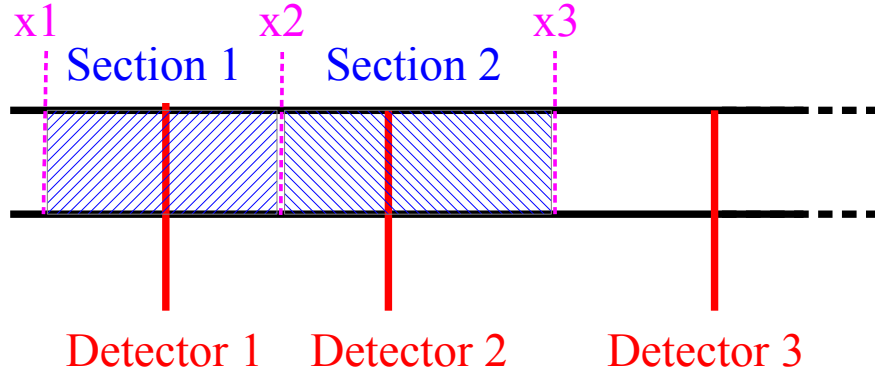
## 3 DATA PROCESSING

The output of the detectors is flow,  $q_l$ , and time mean speed,  $v_l$ , per lane  $l$ . To calculate the density, it would be preferable to have the space mean speed ( $u$ ) or the harmonically averaged speed. However, this is not available for the data. The difference between time mean speed and space mean speed is due to the spread in speed of the individual vehicles. In the Netherlands it is forbidden to overtake on the right hand side, so the left lanes are the faster lanes, and in fact, drivers choose their lane for a large part based on their speed. Therefore, most of the differences in driving speed between vehicles can be expressed in terms of differences in average speeds across the lanes, rather than within one lane. Therefore, we average the time mean speeds per lane harmonically over the roadway, and remove the largest part of the bias usually found in time mean speed (14). In this average we weight according to the flow  $q$  in each lane. So use the following equation to approximate the space mean speed:

$$u \approx \left( \frac{\sum_{\text{lanes } l} q_l \frac{1}{v_l}}{\sum_{\text{lanes } l} q_l} \right)^{-1} \quad (1)$$



**FIGURE 2** A speed contour plot of 20 October 2011 for the selected site; traffic drives from bottom to the top. For illustration purposes in this figure, the traffic data has been smoothed using the methods proposed by Yuan et al. (12) and Schreiter et al. (13).



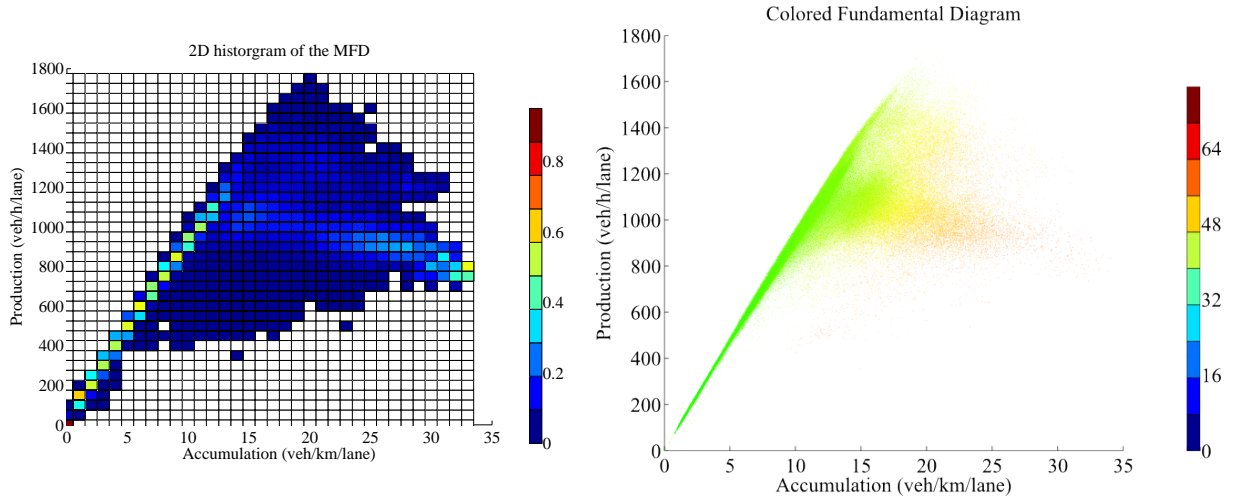
**FIGURE 3** The sections and the construction thereof based on the detector locations

With this speed, we calculate the density as

$$k = \frac{q}{u} \quad (2)$$

This means that for all detectors, we have a flow and density. We assume this traffic state for a section from halfway the detector and the upstream detector up to halfway the detector and the downstream detector. This is illustrated in figure 3, where  $x_2$  is exactly halfway detector 1 and detector 2. For the first and last detector, the downstream boundary is assumed at an equal distance from the detector as the upstream boundary. For the example in figure 3, this means that the distance from the  $x_1$  to the detector is equal to the distance from  $x_2$  to the detector.

Now we need to calculate the production  $P$ , accumulation  $A$  and the spatial spread of density  $\gamma$  from the flow and density of sections  $s$ . Basically, the production is the weighted averaged flow in all cells,



(a) Histogram of the measurements in the accumulation production plane. The color shows which fraction of the measurements at a certain accumulation has the corresponding production.

(b) The MFD in the accumulation production plane. The color shows the spatial spread of density (vehicles/km/lane). It shows that the higher the accumulation, the higher this spread. Furthermore, *for a certain accumulation* the lower the production the higher the spread.

**FIGURE 4 Traffic operations in the accumulation-production plane.**

and the accumulation is the weighted averaged density Geroliminis and Daganzo (5). The weight for all measurements is the amount of lane length in section  $s$ , indicated by  $L_s$ . This is the length of the section multiplied by the number of lanes. Similarly, we define the spatial spread of density as the square root of the weighted variance of the densities in all sections, in which we choose the same weights as for the production and accumulation. In equations, we thus define:

$$\text{Production } P = \frac{\sum_{\text{Sections } s} q_s * L_s}{\sum_{\text{Sections } s} L_s} \quad (3)$$

$$\text{Accumulation } A = \frac{\sum_{\text{Sections } s} k_s * L_s}{\sum_{\text{Sections } s} L_s} \quad (4)$$

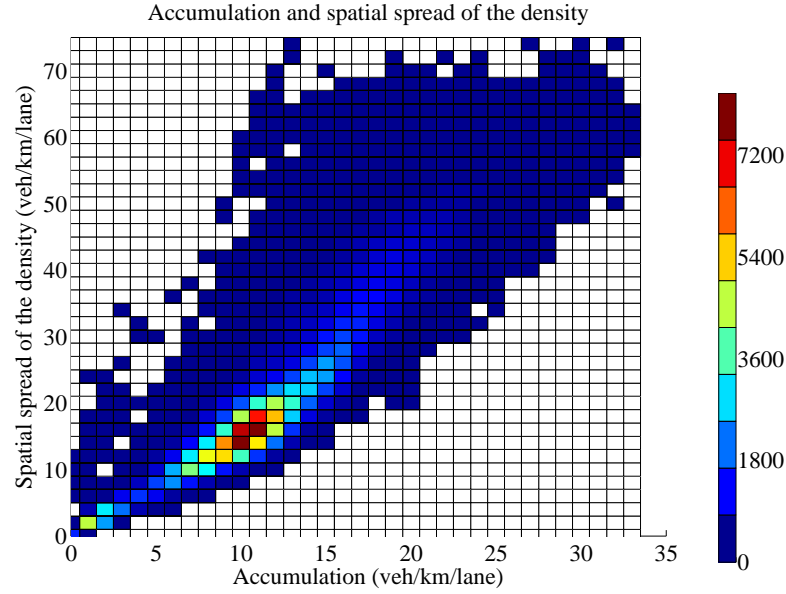
$$\text{Spatial spread of density } \gamma = \sqrt{\frac{\sum_{\text{Sections } s} \left( L_s (k_s - \bar{k})^2 \right)}{\sum_{\text{Sections } s} L_s}} \quad (5)$$

In the next sections we will compute these for all aggregation periods for the network, and we will relate the production to the accumulation and the spatial spread of density.

#### 4 MACROSCOPIC FUNDAMENTAL DIAGRAM FOR URBAN FREEWAYS

Figure 4 shows the accumulation and the production. This relationship is not as crisp as for an homogeneous urban network, as shown by Geroliminis and Daganzo (5). In situations with low density, the speed across the detectors is approximately the same, so the average speed is approximately the same as well for all points with a low accumulation. The most interesting part near the maximum production, and especially near that part, there is much noise. This is due to the fact that especially there, traffic states from different regimes (congested, uncongested) are mixed. These mixed states lead to a lower production, as is shown by Cassidy et al. (9). The two dimensional histogram of the traffic states, figure 4a, show per bin of equal density,





**FIGURE 5 Relationship between the accumulation and the spatial spread of density in a 2 dimensional colorbar. The color indicates the number of observations with a certain accumulation and a certain spatial spread of density.**

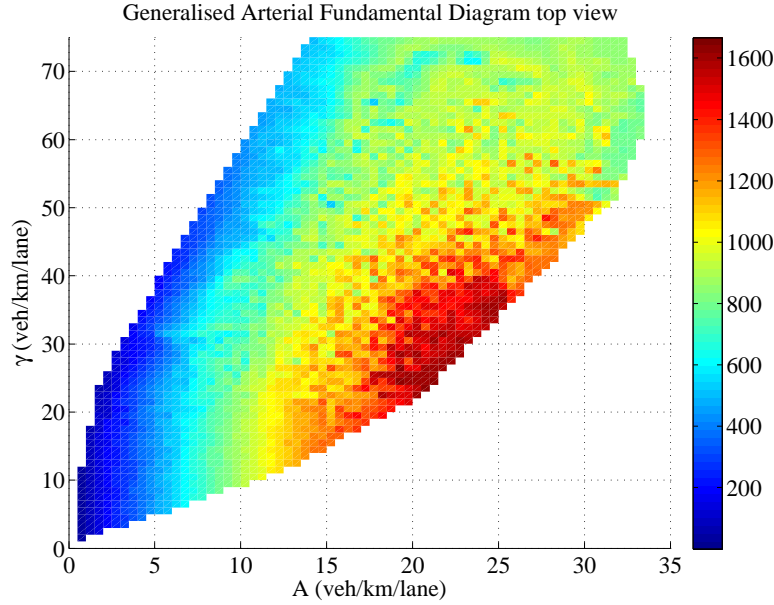
what fraction of the measured traffic states has which production. It shows that at densities near the critical density, it is indeed relatively spread. Quite certainly, the maximum production is not measured often.

Figure 4b gives an MFD in which the points are colored according to the spatial spread of density. The spatial spread of density changes from left to right, indicating that with a larger accumulation the spatial spread of density is also larger. This internal correlation is checked, and shown in figure 5, which indeed shows correlation. More interestingly, for the same accumulation, the points with a higher spatial spread of density have a lower production. The figure thus shows that predicting the production with only the accumulation gives scatter. Moreover, it also gives reason to believe that including the spatial spread of density as explanatory variable improves the prediction.

## 5 GENERALIZED MACROSCOPIC FUNDAMENTAL DIAGRAM FOR URBAN FREEWAYS

The Generalized Macroscopic Fundamental Diagram (GMFD) for urban freeways describes the production on a lengthy road stretch without intersections as function of the accumulation and the spatial spread of the density. This production, shown in figure 6, is constructed by averaging the measurements in bins of similar accumulation (interval steps of 0.5 veh/km) and spatial spread of density (interval steps of 1 veh/km/lane) simultaneously.

This estimated production for the combination of a specific combination of accumulation and spatial spread of density is thus an average of the surrounding points. No further smoothing has been applied to the function. Nevertheless, the production seems to be a rather smooth, and continuous function of accumulation and spatial spread of density. As expected, the production decreases with the increase of the spatial spread in density.



**FIGURE 6 Generalized Macroscopic Fundamental Diagram for urban freeways: the color indicates the production in veh/h/lane.**

## 6 PREDICTING PRODUCTION

Both the MFD and the GMFD can be used to predict the production. This is for instance needed to apply it in a control scheme. This can be done by either a non-parametrized form (section 6.1) or a fitting a functional form (section 6.2). This section describes the process of constructing the relationships, and how these relationships are used to predict the production based on the accumulation (MFD) or the accumulation and the spatial spread of density (GMFD). Section 7 later compares the predictions.

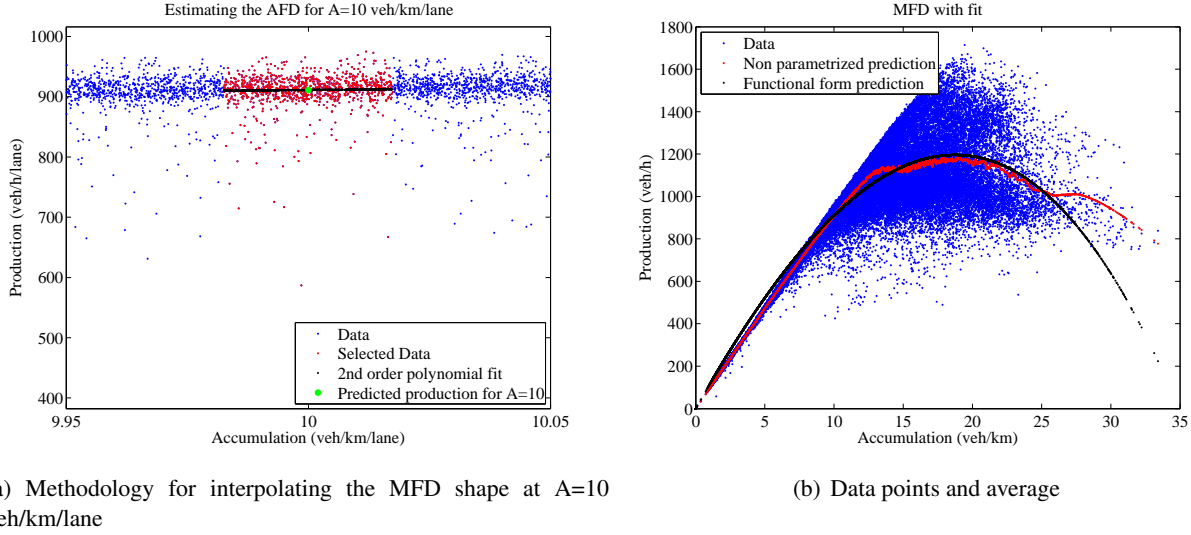
For all methods, we first select 2/3 of our initial dataset of 227,911 measurements (171 days times 841 aggregation intervals of 1 minute between 6am and 8 pm per day), resulting in 151,941 observations, to describe the process or to fit a function (calibration). The remaining 1/3 (75,970 observations) will be used later on to test the prediction power of this method.

### 6.1 non-parametrized form

This section will introduce the two estimation methods, the first based on the MFD (described in section 6.1.1), the second on the GMFD (section 6.1.2), both not using any parametrized form. We will also call these methods interpolation methods. The disadvantage of fitting a predefined shape to the (G)MFD, is that functional form is a priori unknown. We thus first choose for a data driven interpolation method to get the shape of the (G)MFD without further assumptions of shapes. For both the MFD and the GMFD, we then predict the production at a certain accumulation with the average production of the traffic states in that accumulation range.

#### 6.1.1 Macroscopic Fundamental Diagram

This method finds a production for each accumulation. Mathematically, this method implements a polynomial fit of the order of two of the 1000 closest points to a series of regular spaced grid points. The intersection with the y-axis is taken as estimation for the production at that grid point. These 100 points



**FIGURE 7 Predictive power of accumulation for the production**

should be related to the total range of accumulation, spanned by all 151,894 observations. The accumulation range of these 1000 points is on average  $(1000/151,894) \times 35 \text{ veh/km/lane} = 0.23 \text{ veh/km/lane}$ . The influence of the quadratical term of the fit ( $m_3$ ) is then (using the value we will find later after fitting, see table 2,  $-0.044 * (0.23)^2 = 0.0023 \text{ veh/h/lane}$ . This is two orders of magnitude than the stochastic error, as figure 7a illustrates. Therefore, we conclude that the bins sizes are small enough.

The result is plotted in figure 7b. As expected from the data analysis which shows many points under the top, the production line does not increase to the top of the production. As a consequence, the figure comparing the predicted and realised production (figure 8a) shows that the predicted production is stable for a whole range of realised productions. Since traffic control is most likely to aim for the high production, this is not a very good predictor since (1) no traffic state is predicted to have a high prediction and (2) there is no variation in the prediction above a certain threshold level.

Predicting the production is done by first determining which are the nearest neighbors in accumulation for which an average production is calculated using the method described above. The predicted production then is an interpolation of these estimated productions.

### 6.1.2 GMFD

Also for the method implementing an estimate based on the GMFD, we choose a free form as predictor. In fact, we use the same method as explained in section 5 to create the surface: averaging the production in a certain bin.

The surface is described by points  $P = P(A, \gamma)$ . For the traffic prediction of a traffic state outside the calibration set, we first determine the accumulation and the standard deviation of the densities. This combination will not lie exactly at the (grid) values for which the surface is calculated. The predicted production is now the interpolation of the productions from surrounding points of the surface.

## 6.2 Fitting a functional form

The other option is fitting a functional form. The advantage of this method is that there are a few parameters that need to be estimated, so the production function can be described by a few numbers. Similar to the interpolation methods of the previous section, we split the data set into a calibration part (2/3) and a validation

**TABLE 2 Interpretation of the variables**

Description	symbol	equation	Value fit		unit
			MFD	GMFD	
1st order term	$m_1$		116	129	km/h
2nd order term	$m_2$		-2.2	-1.25	km <sup>2</sup> lane/h
3rd order term	$m_3$		-0.033	-0.044	km <sup>3</sup> lane <sup>2</sup> /h
Influence of spread	$r$			13	km/h
Average speed in free flow conditions	$v_{\text{free}}$	$m_1$	116	129	
Gridlock accumulation	$A_g$	$\frac{-m_2 + \sqrt{4m_1m_3}}{2m_3}$	34.9	41.6	veh/km/lane
Critical accumulation	$A_c$	$\frac{-2m_2 + \sqrt{(2m_2)^2 + 4(3m_3)m_1}}{6 * m_3}$	19	23	veh/km/lane
Maximum production	$P_c$	$P(A_c)$	1.2E3	1.8E3	veh/h/lane

part (1/3).

There are many functional forms to describe the shape of the fundamental diagram in the flow-density plane. We choose a polynomial form for reasons of simplicity; in this form, we choose the lowest order possible for reasons of parsimony. If we would fit a polynomial function of the order of two, the critical accumulation, i.e. the accumulation with the maximum production, is half the accumulation for which the area is gridlocked ( $A_g$ ). Since this is not the case, the function  $P(A)$  should have an order of more than two. Choosing the lowest order then gives a third order polynomial function.

We furthermore know that a accumulation of zero should give a production of zero, so the constant must be zero. Also for the impact of the spatial spread of density we choose the simplest reasonable form for reasons of parsimony. Since there are no other restrictions, we choose a linear form for the fit with the GMFD. This results in the following function:

$$P(A) = m_1 A + m_2 A^2 + m_3 A^3 + r\gamma \quad (6)$$

We choose to present the equation in this simple algebraic form, and not a more complex form which shows the physical meaning of the parameters more clearly. The three degrees of freedom of  $M = [m_1, m_2, m_3]$  have a clear interpretation from a traffic flow point of view. For the derivation of these parameters, it is useful to get the derivative of the production to the accumulation.

$$\frac{dP}{dA} = m_1 + 2m_2 A + 3m_3 A^2 \quad (7)$$

The free speed is the derivative of  $P$  at zero accumulation. The accumulation at which grid lock occurs is the accumulation larger than zero for which the production is zero. The critical accumulation can be found by solving  $\frac{dP}{dA} = 0$  for the accumulation, which is a second order polynomial equation. The maximum production can then be found by using this accumulation and calculating the production (equation 6, in which we assume zero spread of density).

We fit the parameters of the function such that the deviation (RMSE) between the measurement and the fit is minimum. Table 2 shows the values of the parameters, as well as the interpretation of these variables for traffic characteristics; figure 7b shows the fit of the MFD in the flow-density plane. Both for the MFD and the GMFD the free flow speed is overestimated (116 and 129 km/h respectively, compared to

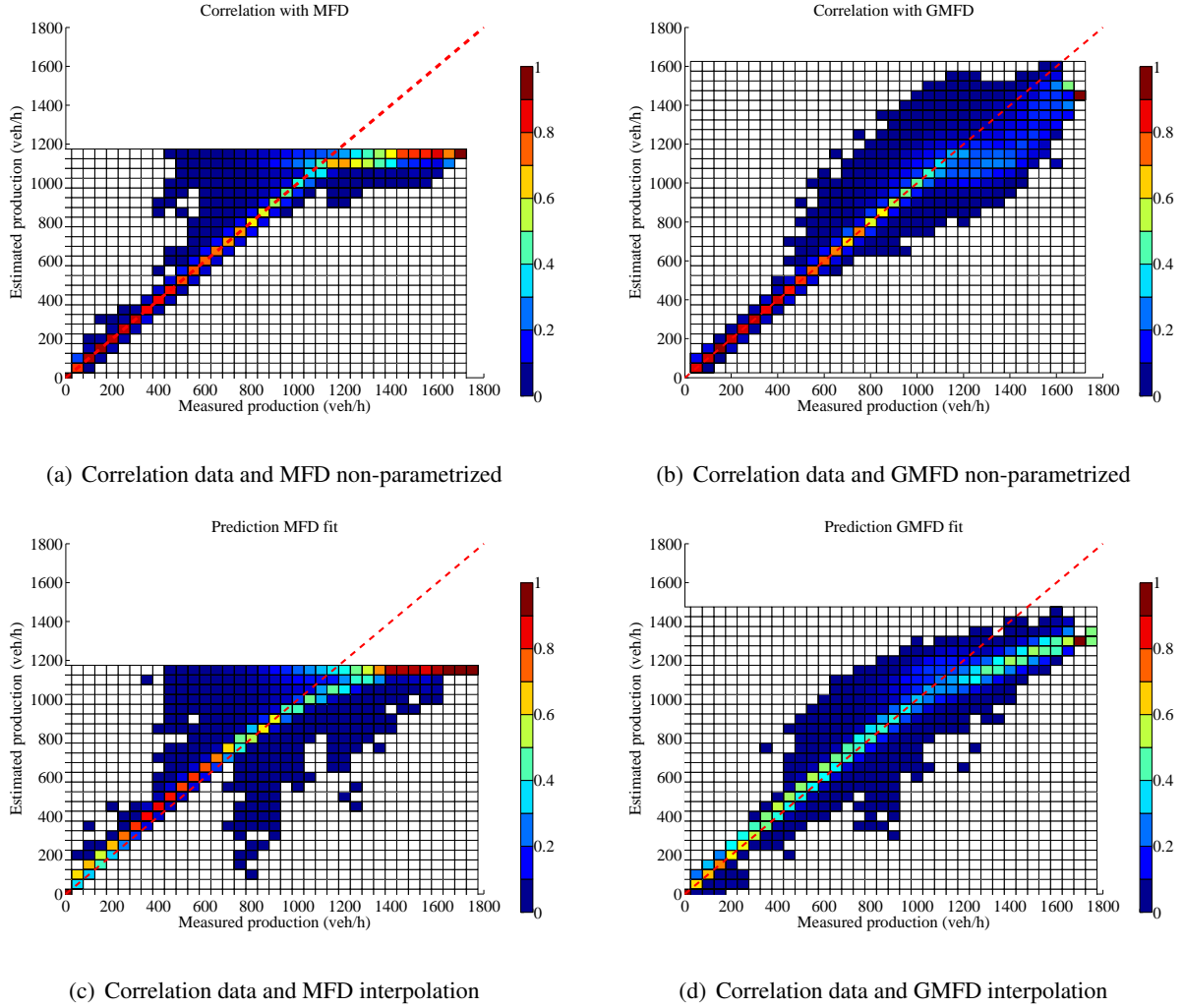
**TABLE 3 Sensitivity of results for the part of data used for calibration, varying from 10% to 100%. For each calibration part, the parameter set is determined 20 times. The mean values for the parameters are given, as well as the standard deviation. Also the predictive power for the validation part (i.e., all data not used for calibration) is indicated by the correlation coefficient  $R^2$ .**

Calibr. part	$m_1$ (km/h)		$m_2$ (km <sup>2</sup> lane/h)		$m_3$ (km <sup>3</sup> lane <sup>2</sup> /h)		$m_4$ (km/h)		$R^2$ of remaining data	
	mean	stdev	mean	stdev	mean	stdev	mean	stdev	mean	stdev
10%	128	3	-1.1	0.4	-0.05	0.01	13.3	0.2	0.8978	0.0009
20%	129	2	-1.2	0.3	-0.047	0.009	13.3	0.2	0.8980	0.0007
30%	129.1	0.49	-1.24	0.05	-0.045	0.001	13.24	0.08	0.8982	0.0003
40%	129.1	0.2	-1.24	0.02	-0.0449	0.0007	13.26	0.06	0.8980	0.0004
50%	129.1	0.2	-1.24	0.03	-0.0451	0.0008	13.26	0.06	0.8981	0.0006
60%	129.1	0.2	-1.23	0.02	-0.0452	0.0006	13.25	0.06	0.8982	0.0006
70%	129.1	0.1	-1.23	0.02	-0.0451	0.0005	13.27	0.04	0.8980	0.0009
80%	129.1	0.1	-1.24	0.01	-0.0449	0.0004	13.26	0.02	0.898	0.001
90%	129.11	0.05	-1.237	0.007	-0.0451	0.0002	13.26	0.02	0.898	0.002
100%	129.14	N/A	-1.240	N/A	-0.0450	N/A	13.26	N/A	N/A	N/A

traffic speeds of approximately 110 km/h at maximum). Figure 7b shows why: the speed is overestimated at lower accumulations, and then it can be lowered to have a flat part near the top. The curvature is higher than what would be expected based on fundamental diagrams. This is due to a lack of data in highly congested situations (no complete gridlock situations occurred). Therefore, the quality of the fit is not reduced by the rapid decrease of production after the critical density, resulting in fitted jam densities of 35 and 42 veh/km/lane for the MFD and the GMFD respectively. Note that these values are outside the range of accumulations for which calibration data were available, so the interpretation of that value being the accumulation at which grid lock occurs, is not very accurate.

The critical accumulation is estimated more in line with values of literature for the fundamental diagram: 19 and 23 veh/km/lane for the MFD and the GMFD respectively. The maximum production for the MFD fit is 1200 vehicles/h/lane respectively. This is in agreement with the maximum average production found in by interpolation in section 6.1. If the spatial spread of density is taken as explanatory variable separately, the maximum production is higher because the state with maximum production there is no spatial spread in density. The interpolation will average this with traffic states with a spatial spread in density, and thus lowers the production. By separating the two variables, we can reconstruct a theoretical maximum production for a homogeneous road stretch (no spatial spread of density). The fit of the GMFD functional form predicts 1800 veh/h/lane as maximum production. It also shows that per veh/km/lane spatial variation of density the production reduces with 13 veh/h/lane.

To analyze the sensitivity of the method for the amount of data included in the calibration, we repeated the fitting procedure for the functional form of the GMFD. We varied the calibration part from 0.1 with steps of 0.1 to 1, and for each of these parts, we estimated the model parameters 20 times. Table 3 shows the results for the parameter estimates and their spreads. For the situation where the whole set is used for calibration, there is no spread since this there is no random inclusion of points (all points need to be included). The table shows that the results (both in found parameter values as in the resulting predictive power, indicated by the correlation coefficient  $R^2$ ) hardly depend on the amount of data included in the calibration, provided it is enough. With the current data set, even 10% is enough.



**FIGURE 8 Predictive power of both interpolation methods. The figure show a two-dimensional probability distribution function: the color indicates which fraction of the data points within a range of production falls in the corresponding bin of predicted production. For the MFD interpolation, the predictions never exceed 1200 veh/h.**

## 7 COMPARISON

This section discusses the advantages of the methodology of a Generalized Macroscopic Fundamental Diagram compared to a regular Macroscopic Fundamental Diagram for both the interpolation (section 6.1) and the fitting of the functional form (section 6.2). For this comparison, we will compare the predictive power of the methods. The end of the section compares the non-parametric form with the functional form.

To quantify the predictive power, we analyse which part of the variation in the production is predicted as such by the estimate. To this end, we compute the squared sum of the differences with the mean,  $S$ , and the squared sum of the differences between the measurement and the prediction,  $Q$

$$S = \sum_i (P_i - \bar{P})^2 \quad (8)$$

**TABLE 4 The R-squares for the fits**

	Data included	MFD	GMFD
Interpolation	All	0.86	0.89
	$A > 18$ veh/km/lane	0.067	0.37
Functional fit	All	0.85	0.90
	$A > 18$ veh/km/lane	0.017	0.39

$$Q = \sum_i (P_i - \tilde{P}_i)^2 \quad (9)$$

In these equations, the bar indicates the mean value and the tilde indicates the estimate by either the MFD or the GMFD by using the interpolation or the functional form. We now calculate the R-square by

$$R^2 = 1 - \frac{Q}{S} \quad (10)$$

This value shows which part of the variation in the production is explained in the prediction.

Figure 8 graphically shows the correlations of the predictions of production with the measured production. Generally, both the interpolation methods and the functional forms show the same: in the free flow part, the prediction is quite good. Once the accumulation approaches the critical accumulation, the methods which only include accumulation as variable estimate a production of 1000-1200 vehicles/h/lane. The spatial spread of density clearly adds useful information to the prediction. Including this as explanatory variable improves the prediction, and the predicted flows reach higher values.

Table 4 shows the R-squares for all methodologies. The R-square is slightly higher for the prediction with the GMFD than the prediction with the MFD (0.89 and 0.86 respectively for the interpolation methods and 0.85 and 0.90 for the functional fit methods). This is mainly due to the fact that the general shape can be predicted: low production with low accumulation, and then increasing with increasing accumulation. However, predicting that low accumulations produce low flow is not interesting, nor difficult.

The interesting part is the top, which is better described by the GMFD. In fact, the MFD will never predict high productions, due to averaging with states with lower production. Let's first focus on the interpolation methods. When calculating the R-squares (equation 10) for the traffic production for traffic states with an accumulation of 18 veh/km/lane and higher, the interpolation of the GMFD performs considerably better than the interpolation of the MFD. The method using the MFD (so only accumulation) returns a R-square of 0.067. This means that in the interesting part, near the highest production, the MFD can only explain 6.7%. The results of the predictions based on the functional form are similar: the addition of the spatial spread of density as explanatory variable improves the fraction of the variation in production which can be explained by the model. For the same reasoning as with the MFD, this effect is much stronger for higher productions. Here we find that the MFD predicts 1.7%. Comparing the results of the parametrized description and the description with a functional form, we find that their performance in terms of predicting the production are similar. For both methods, non-parametrized and a functional form, there are advantages. A non-parametrized form is not bounded to a shape, so not restricted to that shape which is an advantage. However, no smooth conditions are imposed, and many points are needed to describe the function.

## 8 CONCLUSIONS

In this paper we analysed almost ten months data from the A10 urban freeway. We computed the production and related that to the accumulation and the spatial spread of the density. The production is related to the accumulation, but especially near the maximum production, the spatial spread of density plays an important role. Including the spatial spread of density explains a much larger part of the variation in the production.

The resulting surface, production as function of the accumulation and the spatial spread of density, is a smooth function which we call the Generalized Macroscopic Fundamental Diagram (GMFD) for urban freeways. First, it increases and then it decreases with increasing accumulation, as can be expected based on the fundamental diagram which has a similar concave shape. It decreases as function of the spatial spread of density. A third order polynomial function was tested and found to describe the surface approximately as good as a non-parametrized average.

In practice, there is a correlation between the accumulation and the spatial spread of the density: with more vehicles on the road the spatial spread of densities is higher. Nonetheless, the production can not be predicted by only the accumulation.

This GMFD can be used for traffic control on the highway in a similar way as proposed earlier by Landman et al. (15) or Knoop et al. (16), especially coordinated traffic control. For instance, if dynamic speed limits can be operated together with ramp metering installations, the control variables are quite complex, spanning a large solution space. If the objective function of the traffic control needs to be determined by a traffic simulation, each iteration will cost a lot of time, and with a large solution space, many iterations are needed. If the control parameters can be linked to the accumulation and the spatial spread of density, which is possible with ramp metering at different locations, the GMFD can produce the production easily.

**Acknowledgement** This research was sponsored by a IP-CC subsidy from ICTregie/NWO in the project *SI4MS*, Sensor Intelligence for Mobility Systems, and by the foundation Next Generation Infrastructures.

## REFERENCES

- [1] Knoop, V. L., J. W. C. van Lint and J. Vries, L. Kester, and I. Passchier, Relationship between application scale and maximum time latency in Intelligent Transport Solutions. In *Proceedings of the 92nd Annual Meeting of the Transportation Research Board*, Washington D.C, 2013.
- [2] Godfrey, J., The mechanism of a road network. *Traffic Engineering and Control*, Vol. 11, No. 7, 1969, pp. 323–327.
- [3] Mahmassani, H., J. Williams, and R. Herman, . Performance of urban traffic networks. In *Proceedings of the 10th International on Transportation and Traffic Theory Symposium* (N. Gartner and W. N.H.M., eds.), Elsevier, Cambridge, Massachusetts, 1987.
- [4] Daganzo, C., Urban gridlock: Macroscopic modeling and mitigation approaches. *Transportation Research Part B: Methodological*, Vol. 41, No. 1, 2007, pp. 49–62.
- [5] Geroliminis, N. and C. F. Daganzo, Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. *Transportation Research Part B: Methodological*, Vol. 42, No. 9, 2008, pp. 759–770.
- [6] Ji, Y., 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, 2010, pp. 42–48.
- [7] Buisson, C. 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.
- [8] Mazloumian, A., N. Geroliminis, and D. Helbing, The spatial variability of vehicle densities as determinant of urban network capacity. *Philosophical Transactions of the Royal Society A*, Vol. 368, 2010, pp. 4627–4647.
- [9] Cassidy, M., K. Jang, and C. Daganzo, Macroscopic Fundamental Diagram for Freeway Networks: Theory and Observation. In *Proceedings of the 90th Annual Meeting of the Transportation Research Board*, 2011.



- [10] Saberi, M. and H. S. Mahmassani, Exploring the Properties of Network-wide Flow-Density Relations in a Freeway Network. *Transportation Research Records, Journal of the Transportation Research Board*, Vol. in print, 2012.
- [11] Goemans, J., *Evaluatie Dynamax proef A20*. Witteveen + Bos, 2012.
- [12] Yuan, Y., J. W. C. Van Lint, T. Schreiter, S. P. Hoogendoorn, and J. L. M. Vrancken, Automatic speed-bias correction with flow-density relationships. In *Proceedings of the 2010 International Conference on Networking, Sensing and Control (ICNSC)*, Chicago, 2010, pp. 1–7.
- [13] Schreiter, T., Y. Yuan, J. W. C. Van Lint, and S. P. Hoogendoorn, The Adaptive Smoothing Method with Spatially Varying Kernels: ASM-svK. In *Proceedings of the Seventh Triennial Symposium on Transportation Analysis*, Tromsø, Norway, 2010, pp. 1–4.
- [14] Knoop, V. L., S. P. Hoogendoorn, and H. J. Van Zuylen, Empirical Differences between Time Mean Speed and Space Mean Speed. In *Proceedings of Traffic and Granular Flow 07*, Springer, Paris, France, 2007.
- [15] Landman, R., S. Hoogendoorn, M. Westerman, J. Van Kooten, and S. Hoogendoorn-Lanser, Design and Implementation of Integrated Network Management in the Netherlands. In *Proceedings of the 89th TRB Annual meeting*, 2010.
- [16] Knoop, V., J. Van Lint, and S. Hoogendoorn, Route Advice and its Effect on the Macroscopic Fundamental Diagram. *Transportation Research Records*, in print.