The influence of the road layout on the Network Fundamental Diagram

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

D. de Jong MSc
Delft University of Technology
Transport & Planning
Stevinweg 1
Delft, The Netherlands
daviddejong@outlook.com

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

November 13, 2013

Word count:
nr of words in abstract 183
nr of words (including abstract) 4700
Refs (default) 500
nr of figures & tables 8* 250 = 2000
total 7300

Submitted to the 93th Annual Meeting of the Transportation Research Board
ABSTRACT

The Macroscopic or Network Fundamental Diagram (NFD) describes the relationship between the network flow and the number of travellers within the network. The shape of the NFD is important for control purposes like perimeter control or routing. However, it is unknown how the network structure influences the shape of the NFD. This is studied in this paper. Also, the influence of adding urban arterials, decreasing homogeneity but adding road capacity, is studied. Finally, the effect of the location of these arterials is studied.

For this purpose, a tool is developed which can create realistic networks with preset characteristics (road length and road type). The exact network is created at random. Simulation shows that networks with similar characteristics have different NFDs. Also, networks which have arterials next to urban roads are created with the tool. Simulation shows that spread in the NFD is higher for these more heterogeneous networks and the maximum production is lower. The exact location of a major road crossing the network does not have a large effect. All in all, we conclude that the NFD is network-specific and must be calibrated for a particular network.
1 INTRODUCTION

The Macroscopic Fundamental Diagram or Network Fundamental Diagram (NFD) is the relationship between the number of vehicles in a network and the average flow in that area. This concept has been proposed several decades ago (1), and in recent years it gained attention (e.g., (2, 3, 4, 5, 6)). Recently developed control concepts as for instance perimeter control (7) or routing (5) require the shape of the NFD to be known. This shape can of course be measured in real life. For networks which are not yet implemented in real life, determining this curve empirically is impossible. Therefore it is useful to have techniques to determine this curve based on other principles.

An analytical method has been proposed (8), but this holds only for arterial roads with traffic lights. The network specific effects or influence of road types cannot be captured by this methodology. The question addressed in this paper is there: is the NFD dependent on road types and the specific network layout.

In this paper we will show that the network design can change the shape of NFD, even though the main statistics as roadway length and road types are the same. Mixing road types also has an effect. Methodologically the approach taken in this study is that random networks are being designed with similar properties. This could be networks with and without hierarchical structure (e.g., a ring road and minor roads inside).

The methodology to test different networks with all the same properties raised an issue, namely to create networks with all the same properties. This paper develops a method and a tool to do so, which is described in section 4. This tool is used in the study to create similar networks with and without arterials within the city boundary. The goal of this paper is to test the influence of the road layout. All active traffic management measures are hence not incorporated in this paper and for the traffic lights, a fixed timing is adopted.

The remainder of the paper is set-up as follows. The next section gives an overview of the ways to estimate the NFD. Then, section 3 gives the research set-up and methodology. Section 4 describes the tool which has been developed to create networks. Section 5 presents the resulting NFDs and finally, section 6 presents the conclusions.

2 THE SHAPE OF THE NETWORK FUNDAMENTAL DIAGRAM

The field of research into the NFD is rapidly developing. The number of papers on the NFD which appeared recently is too large to discuss all, so we restrict ourselves to the shape of the NFD. For estimating the NFD, an analytical method has been developed (8). They apply variational theory to traffic operations. Integrating the effect of traffic lights into the variational formulation, they are able to present an analytical approximation of the NFD. This is extended by Leclercq and Geroliminis (6) where the effect of route choice is included. Both papers use routes in one direction, so effects of crossing flows cannot be studied with either of these methods. Effects of signal timing in a regular lattice network are shown by Zhang et al. (9).

The dynamics of traffic play an important role in the shape of the NFD. This has been studied at different levels of complexity, ranging from simple insightful networks (4), via grid networks (10) to complexer real-world network networks (11). They all show that traffic networks tend to get more congested once traffic congestion sets in and that production decreases with decreasing traffic homogeneity. This feature is explained and seems to be independent of the network layout.
The design of a network itself is also of importance. With design we mean what the exact connections of the links are. The link length and the number of connections can be similar, but how these links are connected by intersections or T-junctions. The influence of these effects is – as far as the authors are aware – not been studied in depth yet, and this will be studied further in this paper.

3 METHODOLOGY

The goal of this paper is to study the effect of specific network design on the NFD. We do so by creating different networks which share the same basic properties. This means that the structure of the arterials is the same, as well as the locations of the connections between the main roads and the underlying road network. For the underlying road network, the roadway length is similar. However, the exact layout of this underlying road network is different. For instance, the underlying network may consist of several housing blocks, and a block of sport facilities. These can be arranged in a different order, changing the network connections. We say that the basic properties of these networks are the same, but the exact network is different.

The NFD represents traffic operations at the network level. It can therefore be conceived that the exact layout is an issue. On these changing underlying road networks, there are no traffic signals, so they cannot play a role. The main question addressed in this paper is: “Can the NFD been constructed from the roadway length, speed limit and capacity”.

When adding arterials, one might expect that because the capacity of the arterial is larger, the production might be higher. However, it also breaks the homogeneity of the network, and earlier research, showed that this will lead to lower production. We will test several networks with an additional arterial to test which of the effects is stronger. Finally, we will test different locations of the arterials, in the center or off-center, to see whether that extra inhomogeneity limits production further.

Creating similar networks is challenge on its own, which will be tackled by a tool which will be described in section 4. The tool will be used to create seven different networks. First, we are interested in the effect of random variations in unsignalised roads. We hence create three different networks with only a surrounding arterial and create the inner road structure at random. Then, we will study what the effect of a arterial across the town is, which can be either in one dimension (only north-south) or two, i.e. an north-south and an east-west arterial. Finally, it is also studied what the effect of the location of the arterial (centralised or decentralised) is, by locating the arterial in a different location. Note that these basic networks capture all possibilities: adding further arterials is a repetitive step of adding a single arterial.

The networks are in the end compared on the relationship between production and accumulation. These are calculated as follows. The production is the average flow, here calculated by the distance that all vehicles cover in a aggregation time divided by the aggregation time and the road length in the network (units: veh/h/lane). The accumulation is the average density, calculated here as the total number of vehicles divided by the road length in the network (units: veh/km/lane). The capacity of the network is the highest production.
4 DESIGNING RANDOM NETWORKS IN VISSIM

The networks are created using a tool which takes an input. For a microscopic network simulation program – in this paper, we use Vissim – an exact intersection design and signal timing are required as well. The steps to come to a detailed network design are described here. Due to limitation of space, the paper does not describe all steps of the algorithm in detail – for this, we refer to (12) – but describes the main principles. The working of the tool is described beyond it usage in the remainder of the paper, in particular, as example we consider designing network including freeways, which will not be used for the further analyses.

4.1 Input

In Microsoft Excel, the general layout of the network is created, as is shown in figure 1a. In this layout, the location of freeways and arterials is indicated (including speed limit, see also the meaning of the numbers in table 1), as well as the location and type of the different intersections. Next, the intersections are converted to nodes and are connected to their neighbouring intersections, based on the layout of the roads. For intersections which will connect to subnetworks (urban roads, denoted by 2), temporary nodes are created within the subnetwork (see figure 1b), which are used at a later stage to connect the subnetworks to the general layout. Finally, dangling pieces of the network are removed as well.

4.2 Main network

4.2.1 Intersection infrastructure design

Next, the basic layout of the intersections of the major roads needs to be created, as a microscopic traffic simulation program requires this design to be made in detail. For this we follow the Dutch guidelines CROW (13) and create the intersections. Based on the different turns that have to be made at each intersection (following from the general layout), the location of a number of base points, from which the intersection can be constructed are calculated. As a first assumption in creating the intersections, it is assumed that the rightmost lane of each through going turn aligns with the rightmost lane of the opposite outgoing lane. Using this assumption, the width of the central bank can be calculated and the center point of each turn can be determined, based on the lane width and number of lanes.

When the configuration of the lanes is determined, the most crucial base point, the center of the “nose” of the central bank, can be determined. This point is defined as the point where the inside line of the left turns (with a radius of 25.0 meters) are 3.5 meters apart, as is shown in figure 1c. The next assumption is that the (gantries of the) signals are located at a distance of 2.0 meters from the center of this ‘nose’ and that the stop bar is located at a distance of 12.0 meters from the signals. As last, the start and end point of the right hand turns are calculated, by simply filleting the outer curbs at a radius of 15 meters.

After the start and end point of each turn in the intersection have been calculated, a fictive demand of 100 veh/h is given to each turn. Using an overlay, the conflict areas between the different turns is generated, see figure 4.2.1. The guidelines CROW (13) provide the location of the stop bar, and the resulting clearance times can be calculated. Now the cycle times are calculated. If they are
(a) The general network layout - interpretation of numbers in table 1

(b) The resulting basic network

(c) The intersection design

FIGURE 1 The first designs of the network
**TABLE 1 Meaning of the numbers in the design model**

<table>
<thead>
<tr>
<th>Value</th>
<th>Definition of network component</th>
<th>Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Arterial intersection</td>
<td>This intersection can connect to any other intersection; This intersection can create a connection into any adjacent subnetwork; This intersection can create an additional connection to the outside of the network (create new zone).</td>
</tr>
<tr>
<td>1,x</td>
<td>Arterial link x = speed in km/h divided 10;</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Subnetwork intersection</td>
<td>This intersection can connect to arterial intersections; This intersection can create a connection into any adjacent subnetwork.</td>
</tr>
<tr>
<td>3</td>
<td>Freeway link</td>
<td>Links of the arterial cannot cross, unless specified otherwise.</td>
</tr>
<tr>
<td>3,1</td>
<td>End of freeway</td>
<td>Defines the point where the freeway ends and a zone is to be created.</td>
</tr>
<tr>
<td>3,2</td>
<td>Freeway link</td>
<td>Does not hinder the creation of a connection of a subnetwork connection subnetwork intersection into the subnetwork next to may be made the freeway.</td>
</tr>
<tr>
<td>3,3</td>
<td>Freeway turn</td>
<td>Point where freeway makes a turn.</td>
</tr>
<tr>
<td>3,5</td>
<td>Freeway link, underlying</td>
<td>Link of the arterial may cross, but does not connect to link may cross.</td>
</tr>
<tr>
<td>3,7</td>
<td>Freeway off/onramp the freeway.</td>
<td>Point where freeway connects with main intersections; No connection with adjacent subnetwork intersections is made.</td>
</tr>
<tr>
<td>3,9</td>
<td>Freeway junction</td>
<td>Intersection between two freeway links.</td>
</tr>
<tr>
<td>4</td>
<td>Subnetwork entry blocker</td>
<td>Prevents the creation of a subnetwork exit/entry point.</td>
</tr>
<tr>
<td>5</td>
<td>Network boundary</td>
<td>Boundary of the network; Area in which no subnetwork can be created.</td>
</tr>
<tr>
<td>empty</td>
<td>Subnetwork area</td>
<td>Area in which a subnetwork can be created.</td>
</tr>
</tbody>
</table>

too long, the number of lanes for each direction are updated and a new intersection is designed, and the process starts again.

### 4.2.2 Signalized intersections

Using the different conflict areas and clearance times, the different conflict groups and control structures are determined. Using the fictive demands, the saturation flow and adjustment factors are calculated for each turn. Using the method proposed by (14), a signal scheme for each intersections is calculated. Using a Poisson distribution for the arrival rate, the length of each lane is calculated, in order to avoid blocking back of the auxiliary lanes. Next the intersection is converted to a node-link configuration and is placed in the network layout.
4.2.3 Traffic demand and route generation

Apart from the general network layout, the total demand to be processed in the network within a 2-hour period is input externally as well. From the original layout, as shown in figure 1b, each outer node (little green dot), is converted to an origin and a destination. These origins and destinations will then form the basis for the OD-matrix. Using a uniform distribution, a random number between 1 and the maximum number of trips is drawn for each OD-pair. These numbers are proportionally scaled such that the number of trips over all OD-pairs matches the given demand (accounting for a little deviation, due to rounding). The outer nodes can have their origin in either the network itself, or outside of the network, meaning that both internal and external trips are modeled.

Based on the network layout, link speeds and average delay for each turn of each intersections (based on the signal scheme), free flow times are calculated between each OD-pair. Using these travel times, a k-shortest path algorithm and a DUE-assignment (using the Frank-Wolfe algorithm and MSA) different paths are determined, resulting in traffic flows for each link within the network.

To compare the shape of the NFD there need to be congested and uncongested conditions. Therefore, several traffic simulations are done, all with a different traffic demand. The demand is scaled up and down homogeneously over all of the OD pairs traffic operations need to be present.

4.2.4 Static network update

With the obtained flows, first the number of lanes for each link is determined and the layout of each intersection is updated. Within this update, the flow of each link is assigned to the corresponding turn - as described in section 4.2.1 - and a new intersection layout with corresponding signal scheme is calculated. If the cycle time is too long (> 120 s), an additional lane is added to the turn with the highest flow per lane after which the layout and signal scheme are recalculated. This process is repeated until the cycle time is within the limits.
TABLE 2 Characteristics of subnetwork blocks

<table>
<thead>
<tr>
<th>Type</th>
<th>$\Delta x_{\text{min}}$</th>
<th>$\Delta x_{\text{max}}$</th>
<th>$\Delta y_{\text{min}}$</th>
<th>$\Delta y_{\text{max}}$</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Houses</td>
<td>50</td>
<td>100</td>
<td>80</td>
<td>180</td>
<td>50</td>
</tr>
<tr>
<td>High rise buildings</td>
<td>60</td>
<td>90</td>
<td>120</td>
<td>200</td>
<td>25</td>
</tr>
<tr>
<td>Parks</td>
<td>50</td>
<td>400</td>
<td>50</td>
<td>400</td>
<td>10</td>
</tr>
<tr>
<td>Sport facilities</td>
<td>150</td>
<td>400</td>
<td>200</td>
<td>500</td>
<td>15</td>
</tr>
</tbody>
</table>

(a) The 3km x 3km network separated in blocks

(b) Example of the subnetwork blocks (block 9 in figure 3a)

FIGURE 3 The blocks in the design process

4.3 Random subnetwork creation

After the general network layout is created, the areas between the major roads are converted to subnetworks, as shown in figure 3a. Along the border of each subnetwork a safety zone is added, in order to accommodate the roads of the main network. Also inserts are made, which are used to accommodate the connection to the main network, using the temporary subnetwork nodes created earlier.

For each of these areas a street pattern, consisting of local, bi-directional, single lane roads is created. This is based on the size of the blocks. Using Google maps, the typical sizes of blocks (of a residential area) with different land use are determined, see table 2. Also, it is determined which fraction of the blocks is used for what purpose. Using these values, a block type and its dimensions are drawn at random, after which the block is inserted at the first possible bottomleft position. Then a local road is added around that block. This process is repeated until the block is filled for a certain percentage, or no more blocks can be added. The remaining areas which are too small to fit any other purpose are filled with parks. An example of the way in which a subnetwork is filled, is shown in figure 3b.

Next the subnetwork is divided in smaller sections (see the small black lines in figure 3b), and a feeder, an origin/destination node is added within each of the resulting areas. All trips, assigned to the original temporary subnetwork nodes, are divided among the resulting feeders, based on the inverse of their relative distance to each of the intersections. This avoids demands originating from the eastern intersection to be assigned to a feeder at the west of the subnetwork, resulting in a mismatch of the demand at each intersection. Demand for OD-pairs within the same subnetwork are set to zero, under the assumption that the resulting travel distances are sufficiently
FIGURE 4 Steps in the design process
short to be made by another mode of transport. Again, the number of trips for the remaining OD-pairs is scaled, to match the original given demand. After traffic is assigned, the network is converted to a node-link layout and is connected to the main network.

After the subnetworks have been fully created and the demand pattern is set, the subnetworks are also converted to a link-node structure (see figure 4.3) in which all links are converted into bi-directional roads. For subnetwork intersections, a separate link is created for each turn. The roads connecting the subnetwork to the main network (shown in black in figure 3b), are removed and replaced with two entry-nodes at the point where it connects to the road structure of the subnetwork.

When completely converted, the subnetwork is fitted into the arterial network, using the indents in the subnetwork, to accommodate the intersection branches leading into the subnetwork. Intersection lanes leading into/out of the subnetwork, are then connected to the entry-nodes by adding an additional link. The result of this is shown in figure 4.3.

### 4.4 Configuring the network for use in VISSIM

Before the network can be exported to VISSIM, its link-node structure has to be converted to the link-connector structure used in VISSIM. Within VISSIM, links are connected to each other by connectors, which are roughly similar to links. An important property of both the link and the connector is that they cannot be connected to their same type, e.g. links cannot connect to links and connectors cannot connect to connectors. Another important difference between the link and the connector is that a link is a straight line and that a connector can be curved. So in order to accurately model intersection curves in VISSIM, the link-node structure has to be redesigned. Apart from that, the connection between the different lanes before, on and after the intersection has to be accurately modeled within VISSIM as well. In order to obtain realistic intersections, again the design rules, as shown in figure 1c are applied and the actual curves of the turns on the intersection are constructed.

Next the incoming lanes have to fan out over the different turns of the intersection. For this the following design rules are applied:

- Lanes diverge at an angle of $\pi/10$. As lane width is 3,0 meters, the additional length of each diverging lane is 30 meters;

- If the number of lanes is equal to the number of lanes at the stop bar, a 1-on-1 connection is made;

- If a through going lane is present and the number of through going and left lanes is equal to or larger than the number of incoming lanes, the rightmost incoming lane is connected to the rightmost through going lane. Additional lanes are created on either side of the through going turn;

- If a through going lane is present and the number of through going left and right turn lanes is smaller than the number of incoming lanes, than the rightmost lane of the through going turn should be as close as possible to the rightmost incoming lane. Incoming lanes are then assigned to the turning lanes from left to right;
If only the left and right turns are present, the incoming lanes are divided equally over the turning lanes. In case the number of incoming lanes is unequal, the remaining lane is assigned to the turn with the highest intensity;

In case a through going lane and a single turning lane are present, the incoming lanes are first assigned to the through going turn, from right to left. Any remaining lanes are then assigned to the left or right turn.

The resulting intersection, as present in VISSIM is shown in figure 4.3. When the network is fully converted, it is loaded into VISSIM and a number of initial simulations are run (at 20 percent of the demand), in order to obtain an extensive path set (50) for each OD-pair. Then the simulation is run at full demand and data is obtained for each link and intersection in the network. Using the number of vehicles which have passed each of the links within the network, the number of lanes for each arterial link can be recalculated. Also using the total number of vehicles that have used each turn at an intersection can be determined. Using these intensities, the design and signal timings of each intersection are completed recalculated, using the method as described in section 4.2.4. This update is run only once, but can be repeated multiple times until convergence is reached. Nevertheless, it was found that all networks were capable of processing the assigned demand, without resulting in congestion, making additional updates unnecessary.
5 IMPACT OF NETWORK LAYOUT ON NFD

5.1 Networks used

To study the impact of the network layout we construct different networks. First, to study the impact of different road layouts, one base network is created. This is a 3x3 km square network with an arterial ring road. Internally, this is filled with blocks as described in section 4.3. This random process is repeated three times, each giving a different network, with similar characteristics. The resulting networks are shown in figure 5.

As indicated in section 3, the effect of additional arterials within the network is studied. Several networks are created, with arterials in one (network 4 and 5) or two (network 6 and 7) directions, either centered in the middle (networks 5 and 7) or off-center (networks 4 and 7). The resulting physical layouts are shown in figure 5.

5.2 Resulting traffic operations

Figure 6 shows the NFDs for the different networks. Let us first consider the first question “Can the NFD been constructed from the roadway length, speed limit and capacity”. We do so by considering the NFD of the networks with only a ring road arterial and no arterial inside, networks 1 to 3. We see a clear free flow branch of the fundamental diagram, and a sharp point of maximum production. The NFDs for network 1 differs from the other two, in the sense that the maximum production is lower, and there is more fluctuation near the top. So the exact shape of the NFDs depend on the network.

After the top, the NFD then decreases to the congested branch. This transition is much sharper than the analytical method with cuts for a ring road with traffic lights (8) suggests. Moreover, the congested branch shows a convex part, which cannot be found using the above mentioned method. Possibly network effects with spillbacks to other links cause this shape. The NFDs, created by averaging all traffic operations on the arterial and the inner network, are quite crisp.

Then the second question: does adding a arterial increase capacity? Adding a north-south arterial does not increase the capacity. In fact, the capacity decreases. Note that this is the normalised capacity, in veh/h/ lane, so this includes the length of the arterial stretch as well. An explanation is that the inhomogeneity makes that congested routes are nevertheless attractive to travellers, because the partially congested arterial gives a lower travel time than the completely uncongested local road. In the congested part, there is a larger spread. This is probably due to the fact that different states with this accumulation can exist with different types of spread in the network. Earlier, it has been shown (15), explained (16) and modeled (17) how more variation in densities leads to a lower accumulation. Because the network is inhomogeneous, there might be situations with more congestion and still people will not deviate from the fastest route. This will not always be the case (for instance for east-west movements), so there is more variety in the variation. Hence, the congested branch is wider. Note that the trend is here as well to have a convex end of the NFD. It is also interesting to note that the production at the higher accumulations will not go to zero, but remains slightly above. This gridlock effect is avoided due to the re-routing properties in Vissim.

In network 6 and 7 there are two arterials in the network, north-south and east-west. This also does not increase the normalised capacity, for the same reasons as adding one arterial does.
FIGURE 6 The fundamental diagrams for the seven networks

not increase capacity (it is normalized per road segment). The second arterial seems to partially correct for the effects of the inhomogeneity seen in the network with one arterial north-south. This is because by adding a second arterial, both dimensions, north-south and east-west are equal again. Therefore, the inhomogeneity effects seen in networks 6 and 7 are less considerable. However, the arterial still makes the network inhomogeneous, so some inhomogeneity remains. Especially in network 5 where inhomogeneity might play a larger role, several hysteresis loops can be observed. It is generally known that under recovering the network flows are lower than during the onset of congestion for the same network accumulation (e.g., (16, 18)).

The last question is what the impact is of the location of the inner arterial. The results show that the NFD is practically equal for any location of the arterial. Comparing network 4 (centered north-south arterial) and network 5 (off-center north-south arterial), no difference can be seen. The same holds for the comparison of network 6 (centered north-south and east-west arterials)
and network 7 (off-center north-south and east-west arterials). It seems that the networks with off-center arterials have a slightly higher production, but this is too marginal to quantify here. We could also attribute these changes to the random effects, in the same way that the NFD of network 1 differed from the NFD of network 2 and 3.

6 CONCLUSIONS

This paper studied NFDs of various road networks. It presented a tool to create random networks based on required properties of the network. This tool has been used to analyse the effect of the network layout on the NFD.

It has been known that signal timing plays an important role in the NFD. However, also changes in the underlying road network (in which no traffic signals are present), result in differently shaped NFDs. That implies that the network structure also has an important role. It has been found that the NFD depends on the structure of the network, so even for the inner network without traffic signals the NFD cannot be predetermined. The paper also shows that homogeneous traffic networks, even if they only consist of local roads, give a crisper NFD than if two road types are mixed. The spread in the NFD is larger if one arterial (in one direction) is added than if in both directions a arterial is added. The exact location of the arterial plays no significant role. This way, adding an arterial decreases the maximum production (normalized to the roadway length).

Traffic control schemes based on the NFD are currently being developed. The findings in this paper show that it is required to determine the NFD for each network layout specifically, and it cannot be based on the general characteristics as network length or road type.

Acknowledgement This research was performed in the ITS Edulab.

REFERENCES


