Chapter 4

Fundamental diagrams

Contents of this chapter. This chapter introduces the concept of the fundamental diagram. Some models used in practice will be discussed, also one with a so called capacity drop. This capacity drop has implications for using the diagram for capacity estimation. Also the link with a simple model for an indivudual driver is discussed. Finally some aspects of studies about the diagram will be discussed and a practical result of the effect of rain on the diagram.

List of symbols

q	veh/s	flow rate, intensity, volume
k	veh/m	traffic density
u	m/s	instantaneous speed
q_c	veh/s	critical flow / capacity
k_c	veh/m	critical density
k_j	veh/m	jam density
u_c	m/s	speed at critical density

4.1 Introduction

It is reasonable to assume that drivers will on average do the same under the same average conditions: if drivers are driving in a traffic flow that has a certain speed u, they will on average remain the same distance headway s with respect to the preceding vehicle. This implies that *if we would consider a stationary traffic flow, it is reasonable to assume that there exists some relation between the traffic density* $k = 1/s$ *and the instantaneous mean speed* u (and since $q = ku$, also between the density and the flow, or the flow and the speed). This relation – which is sometimes refered to as the equilibrium situation – will depend on the different properties of the road (width of the lanes, grade), the composition of the flow (percentage of trucks, fraction of commuters, experienced drivers, etc.), external conditions (weather and ambient conditions), traffic regulations, etc.

In traffic flow theory the relations between the macroscopic characteristics of a flow are called 'fundamental diagram(s)'. Three are in use, namely:

- intensity - density $q = q(k)$
- speed - density $u = u(k)$
- speed - intensity $u = u(q)$

It is important to understand that these *three relations represent the same information*: from one relation one can deduce the other two; See Fig. 4.1.

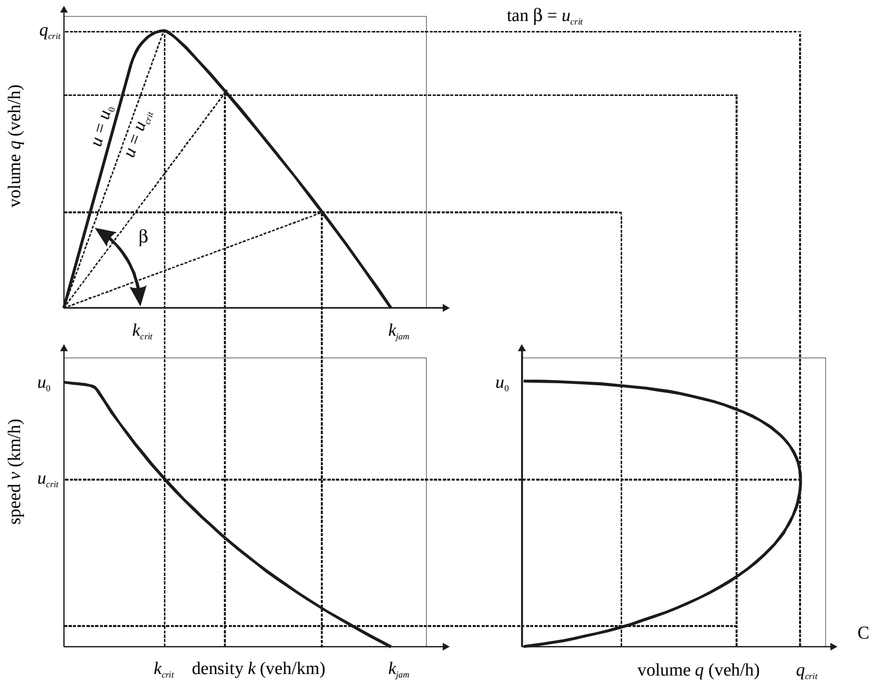

Figure 4.1: Three interrelated forms of the fundamental diagram

Remark 35 *Sometimes density k is replaced by occupancy β but this does not influence the character of the relations substantially.*

It is tempting to assume that the fundamental diagram implicates causality amongst density, flow and speed. One might argue that the average speed u of the drivers is determined by the density k , i.e. $u = u(k)$ (or equivalently, the speed u is determined by the average distance headways $1/k$). However, care should be taken when making such inferences, as is illustrated by the following example.

Consider a situation when traffic conditions are congested, and overtaking opportunities are limited, drivers will be following the vehicle in front and will thus adopt to the speed of the latter to do so. In other words, the follower can only control the distance s with respect to the vehicle in front. It can thus be argued that under these circumstances the distance s is a function of the speed u. In fact, the following relation has been proposed in the literature to express the relation between the distance s and the speed u

$$s = s_0 + T_r u \tag{4.1}$$

where s_0 denotes the minimal gap between two vehicles, and T_r denotes the reaction time. On average, this relation implies the following relation between the speed and the density

$$k = k\left(u\right) = \frac{1}{s_0 + T_r u} \tag{4.2}$$

or

$$q = ku = \frac{u}{s_0 + T_r u} \tag{4.3}$$

Since by definition the average speed will always be smaller than the free speed, i.e. $u \leq u_0$, from Eq. (4.3) we can conclude that the flow is bounded from below, i.e. $q \geq \frac{u_0}{s_0+T_r u_0}$ and can thus not be valid for dillute traffic conditions, where the flow is very small.

On the contrary, assuming free flow conditions, the speed u is generally not determined by the vehicle in front, but rather by the free speed u_0 and an occasional interaction between a slower vehicle (assuming sufficient passing conditions). Then, the density is a function of the flow q and the free speed u_0, i.e.

$$k = k(q) = \frac{q}{u_0} \tag{4.4}$$

Note that derivation of the fundamental relation from a microscopic model is discussed later in this chapter.

4.1.1 Special points of the fundamental diagram

Special points of the diagram are:

- (mean) Free speed u_0; this is the mean speed if $q = 0$ and $k = 0$; it equals the slope of the function $q(k)$ in the origin;
- Capacity q_c; this is the maximal intensity, sometimes called critical intensity;
- Capacity density or critical density k_c ; i.e. the density if $q = q_c$;
- Capacity speed u_c ; i.e. the mean speed if $q = q_c$;
- Jam density k_j ; i.e. the density if $u = 0$ and $q = 0$.

Note that $u(q)$ is a two-valued function, i.e. at one value of q there are two possible values for u, and k. Note also that capacity is not a special point of the function $u(k)$.

Notation 36 *The part of $q(k)$ with a constant speed is called the 'stable region' of the diagram*[1]*. As soon as speed decreases with increasing density, one enters the 'unstable region'.*

Notation 37 *The region in which densities are greater than the capacity density, is called 'congestion region' or 'congestion branch'. The entire region with $k < k_c$ is sometimes called 'free operation' and the congestion region 'forced operation'.*

4.1.2 Importance of the fundamental diagram

- Definition of Level of Service (LOS) - the LOS - a yard stick for the quality of the traffic operation - is related to the fundamental diagram. The LOS is discussed further in chapter 6.
- The diagram contains also information about the relation between intensity and travel time; this is used in traffic assignment. Fig. 4.2 shows the relations between density and intensity on the one hand and travel time on the other.
- Determination of the capacity q_c, which is an important parameter in traffic planning and control, is sometimes based on a specific form of the diagram, as will be shown in section 4.2.3.
- Often the fundamental diagram is a basic element of more comprehensive models that describe the traffic operation on a network, e.g. METANET, or describe assignment and traffic operation in combination, e.g. 3DAS [15].
- On the other hand microscopic models, e.g. the model FOSIM, can be used to derive fundamental diagrams.

[1] In chapter 4 on shockwave analysis, this namegiving will become clear.

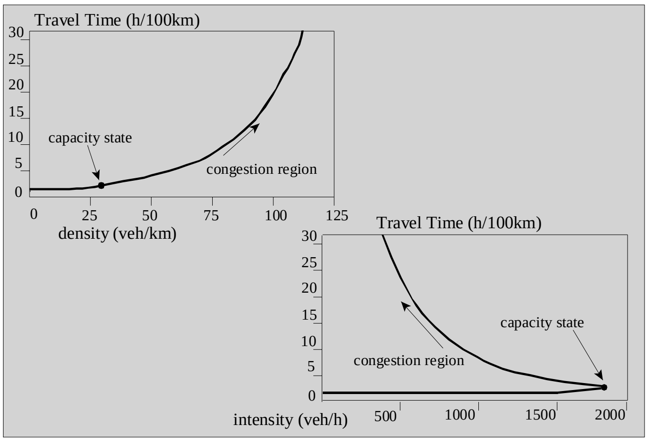

Figure 4.2: Travel time per distance as a function of density and intensity

4.1.3 Factors influencing the diagram

The fundamental diagram is not a physical law but is dependent on characteristics of:

1. the road;
2. the drivers and their vehicles;
3. conditions such as lighting and weather.

In practical terms the most important factors are:

- Type of road and, for a given type, characteristics such as lane width, curvature, grades, and quality of surface.
- Vehicle composition; this is normally made operational by a division of the flow between cars and trucks (in the USA recreational vehicles are often taken into account as a separate vehicle category).
- Type of traffic with respect to travel purpose and familiarity of the drivers with the road. E.g., home-work travel over relatively short distances shows a different behaviour than long distance holiday traffic.
- Measures, such as speed limits, either permanent or dynamic.
- Lighting conditions, e.g. darkness (with or without road lighting) and daylight.
- Weather conditions: dry, rain, fog.

To give some indication of the influence of these different aspects, Tab. 4.1 shows an overview of the capacity reductions for different weather and ambient conditions.

Over the years vehicles have changed (better road-holding, better brakes, more comfort) and so have drivers, especially their experience in handling dense traffic at relatively high speeds. These factors have led to an increase in capacity of motorways during the last few decades; estimates range from 0.4 to 1.0 % per year[4].

[4] American Highway Capacity Manual: under ideal conditions (road, traffic, external factors) the capacity of one lane of the motorway has increased from 2000 veh/h in 1950 to 2400 veh/h in 2000, corresponding to a growths of 0.37 % per year.

Circumstances	Capacity
Ideal conditions	100%
Darkness (no illumination)	95%
Darkness (with illumination)	97%
Dense Asphalt Concrete (DAC) [2]with rain	91%
Open Asphalt Concrete (OAC) with rain[3]	95%
DAC / rain / darkness	88%
DAC / rain / darkness / illumination	90%
OAC / rain / darkness	91%
OAC / rain / darkness / illumination	92%

Table 4.1: Capacity reduction effects

4.2 Models of the fundamental diagram

Looking through the literature one finds a lot of models for the diagram. They are based on:

- *Driving behaviour theories*; e.g. car-following theory. By aggregating the individual microscopic behaviour, one can obtain the macroscopic fundamental diagram. The aggregation can be done analytically or by means of a microscopic simulation model. It's obvious that keeping distance is especially relevant for the congestion region. At free operation a mixture of keeping distance and free driving determines the diagram. Analogies with a queueing system exist; the queueing is the driving behind a slower leader that cannot be overtaken immediately. Wu's diagram in the next section uses this analogy.

- *Empirical observations.* Using a 'curve fitting' technique a function is determined that fits the data well and preferably does not contradict elementary conditions. The importance of this type of model has increased with the abundant availability of data. The models of Smulders and De Romph in the next section are Dutch examples of this development.

- *Analogy with flow phenomena from different fields of science*, e.g. the mechanics of a stream of gas of fluid. The importance of these type of models has decreased because the knowledge about the behaviour of drivers has increased and it is really different from gas molecules and fluid particles. This type of model will not be included in these notes; interested readers are referred to [18].

Often an analysis starts with the function $u(k)$ because that function is monotonic. It is always advisable to determine the other relations $q(k)$ and $u(q)$, too, and inspect them for consistency and plausibility. Because the derivative of $q(k)$ is the speed of the kinematic wave, this function is also useful for checking the model.

4.2.1 Model specifications

- Fundamental diagram of Greenshields [22].

This is certainly one of the oldest models of traffic flow theory. It is based on the simple assumption that mean speed decreases linearly with density. The model was validated using 7 (seven!) data points obtained by aerial photography.

$$u(k) = u_0 \left(1 - \frac{k}{k_j}\right) \tag{4.5}$$

with u_0 = free speed and k_j = jam density. The function $q(k)$ follows simply:

$$q = ku \Rightarrow q(k) = ku_0 \left(1 - \frac{k}{k_j}\right) = ku_0 - k^2 u_0/k_j \text{ (parabola)} \tag{4.6}$$

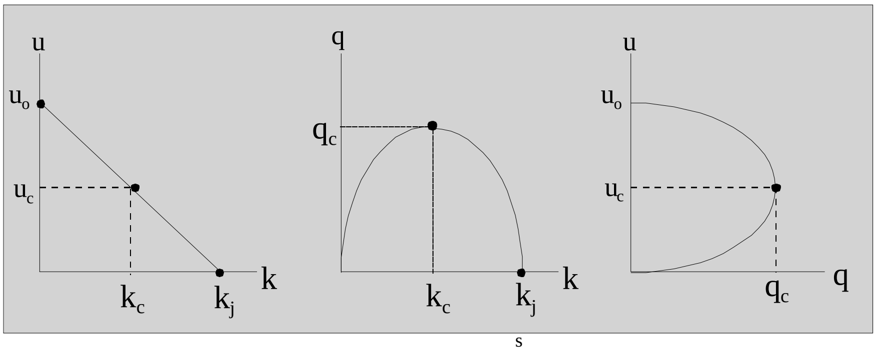

Figure 4.3: Greenshields' fundamental diagram

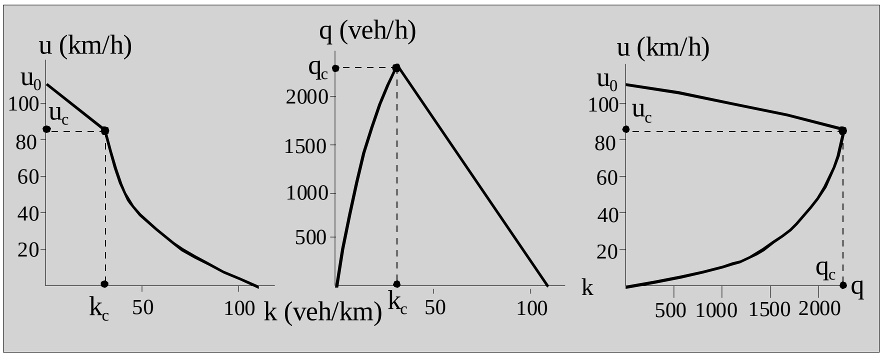

Figure 4.4: Smulders' fundamental diagram

The critical values for q, k, and u follow from the condition that at the capacity point:

$$\frac{dq}{dk} = 0 \Rightarrow \frac{dq}{dk} = u_0 - 2ku_0/k_j = 0 \Rightarrow k_c = k_j/2 \tag{4.7}$$

and thus, by substition into eqn. (4.5)

$$u_c = u_0\left(1 - k_c/k_j\right) = u_0/2 \tag{4.8}$$

For the capacity, we find

$$q_c = k_c u_c = \frac{1}{4} u_0 k_j \tag{4.9}$$

The function $q(u)$ is also simple to derive: substitute in (4.5) $k = q/u$

$$u = u_0\left(1 - (q/u)/k_j\right) \Rightarrow 1 - u/u_0 = (q/u)/k_j \tag{4.10}$$

or

$$q(u) = uk_j - (k_j/u_0)\,u^2 \tag{4.11}$$

Greenshields' model is still used very much because of its simplicity; see Fig. 4.3. However, for motorways it deviates from present reality since u does not decrease that strongly for small densities, and the fact that capacity density is much smaller than half of the jam density (0.2 would be a better number).

- Fundamental diagram of [49]

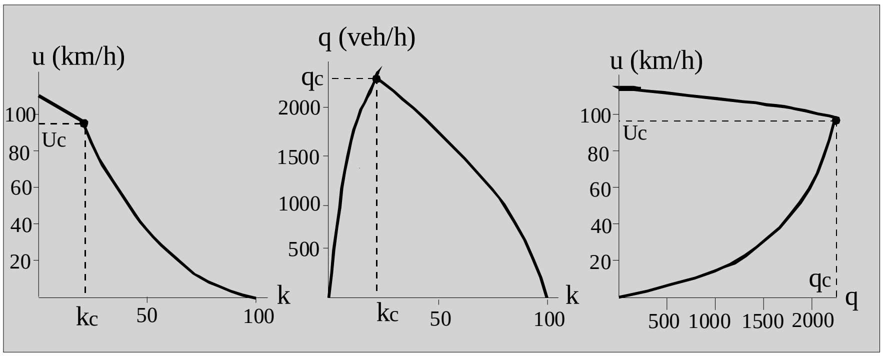

Figure 4.5: De Romphs' fundamental diagram for a 3-lane roadway

In developing a dynamic macroscopic model for traffic flow on motorways, Smulders introduced the diagram: (see Fig 4.4):

$$u(k) = \begin{cases} u_0\,(1 - k/k_j) & \text{for } k < k_c \\ \gamma\,(1/k - 1/k_j) & \text{for } k > k_c \end{cases} \tag{4.12}$$

Parameters are: u_0, k_c, k_j, γ. The parameter γ is not free but follows from the requirement that $u(k)$ is continuous at point $k = k_c$. It is left to the reader to derive that $\gamma = u_0 k_c$.

Smulders chose for a Dutch motorway: $u_0 = 110km/h$, $k_c = 27veh/km$, and $k_j = 110veh/km$ (all variables per lane). Capacity then becomes: $q_c = 2241veh/h$.

Remark 38 *For this diagram the requirement $dq/dk = 0$ in the capacity point does not necessarily hold.*

- Fundamental diagram of De Romph

De Romph (1994) has generalised Smulders' diagram to:

$$u(k) = \begin{cases} u_0\,(1 - \alpha k) & \text{for } k < k_c \\ \gamma\,(1/k - 1/k_j)^\beta & \text{for } k > k_c \end{cases} \tag{4.13}$$

At first glance there are 6 parameters: u_0, k_c, k_j, α, β, γ. The required continuity at $k = k_c$ reduces the number of independent parameters to 5.

The reader is invited to derive that: $\gamma = u_0\,(1 - \alpha k_c) \,/\, (1/k_c - 1/k_j)\,\beta$. Using extended measurements on the ring road of Amsterdam, the parameters of this model have been estimated for different elements of the motorway; e.g. 2-lane roadways, 3-lane roadways; single lane onramps; roadways in a tunnel; etc.

Example 39 *For a 3-lane roadway the results (average values for one lane) are:*

- **Example 40** – $u_0 = 110km/$, $k_c = 23veh/km$, $\alpha = 5.710^{-3}$, *and* $\beta = 0.84$.
  - *Assumed is:* $k_j = 100veh/km$. *Then it can be calculated:*
  - $\gamma = 1672$, *and capacity* $q_c = 2215veh/h$; *see Fig. 4.5.*

Remark 41 *The speed of the kinematic wave at jam density for this model is minus infinity, which is far from reality. Derivation:*

$$q(k) = \gamma k\,(1/k - 1/k_j)^\beta \Rightarrow \frac{dq}{dk} = \gamma\,(1/k - 1/k_j)^\beta - \gamma k^{-1}\beta\,(1/k - 1/k_j)^{\beta-1} \tag{4.14}$$

If $k \to k_j$ then the first term of $dq/dk \to 0$, but the second term $\to \infty$ because $\beta < 1$. β has to be < 1 because otherwise the congested branch of q(k) is convex, which is not feasible either. Moreover, De Romph has found $\beta < 1$ based on data.

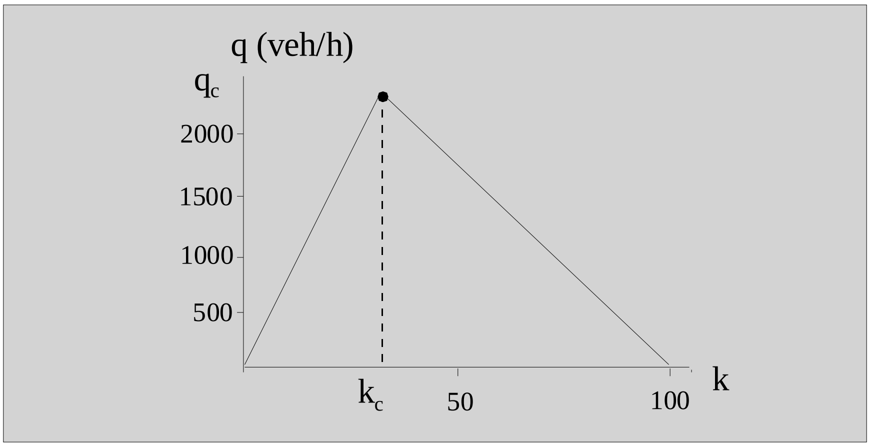

Figure 4.6: Daganzo's fundamental diagram

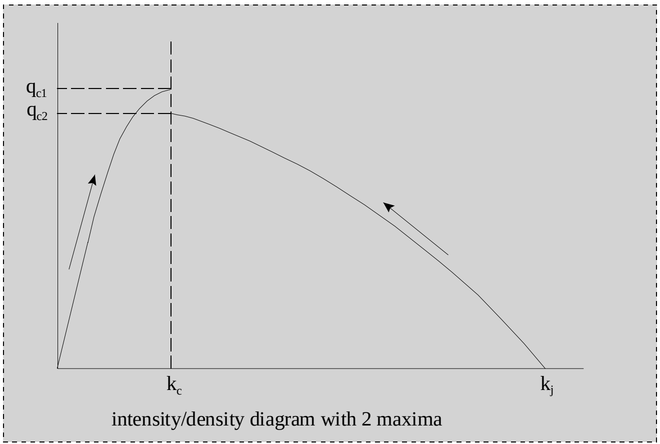

Figure 4.7: Example of a fundamental diagram with a discontinuity

- Schematised fundamental diagram of Daganzo [14]

When developing new models for traffic operation one needs a fundamental diagram that is simple but represents the essential properties of the traffic flow correctly. The simple parabola model of Greenshields is an example fulfilling these requirements. Daganzo has introduced an alternative in which the function $Q(k)$ is represented by two straight lines; see Fig. 4.6. It is left to the reader to determine the relations $u(k)$ and $u(q)$.This model has three parameters: u_0, q_c (or k_c) and k_j. It is remarkable that the diagrams of De Romph and Daganzo are rather similar if one chooses the parameter values correctly.

4.2.2 Concept of discontinuous diagram

Edie [19] was the first researcher to indicate the possibility of a discontinuity in the diagram around the capacity point. This idea is based on the observation that a traffic stream with increasing density (starting from stable or free flow) reaches a higher capacity value ('free flow capacity', q_{c1}) than a traffic stream starting from a congested state (in the extreme case from a standing queue) that ends in the so called 'queue discharge capacity', q_{c2}. This idea is illustrated in Fig. 4.7 in which a so called 'capacity drop' is present.

4.2.3 Wu's fundamental diagram with capacity drop

Recently Wu [58] has developed a model for the diagram with a capacity drop, based on assumptions about microscopic behaviour. In this model two regimes are distinguished: free flow and congested flow. Free flow has densities from $k = 0$ up to $k = k_1$. Congested flow has a density range from $k = k_2$ up to the jam density k_j. Both regimes are overlapping in terms of density range, i.e. $k_1 > k_2$.

Main assumption for free flow state

In this state it is assumed the traffic flow is a mixture of free driving vehicles with mean speed u_0 and platoons with speed u_p. If the fraction of free driving vehicles (in terms of density) is p_{free}, then the fraction of vehicles in platoon is $(1 - p_{free})$, and the overall space mean speed equals:

$$u = p_{free}u_0 + (1 - p_{free})u_p \tag{4.15}$$

Parameters: $p_{free} = 1$ for $k = 0$ and $p_{free} = 0$ for $k = k_1$.

It can be argued, based on queueing theory, that for a two lane roadway (for one-directional traffic) p_{free} decreases linearly with density, i.e.

$$p_{free} = 1 - k/k_1 \tag{4.16}$$

For a three-lane roadway $p_{free} = 1-(k/k_1)^2$ and in general the exponent equals the number of lanes minus 1, i.e. for n lanes:

$$u(k) = (1 - (k/k_1)^{n-1})u_0 + (k/k_1)^{n-1}u_p \tag{4.17}$$

It is left to the reader to derive that (4.17) implies $u(k)$ starts flatter if the roadway has more lanes.

At density k_1 every vehicle drives in a platoon with speed u_p. This implies a relation between the nett time headway, h^f_{nett}, in this platoon, the effective vehicle length $(1/k_j)$ and speed. It is left to the reader to derive:

$$k_1 = \left(u_p h^f_{nett} + \frac{1}{k_j}\right)^{-1} \tag{4.18}$$

Given parameters u_p and k_j, either h^f_{nett} or k_1 is a free parameter of the model. Wu has chosen h^f_{nett} because it can be observed in practice more easily than k_1, especially if it is assumed that h^f_{nett} is a constant parameter for all free flow states.

Assumption for congested flow state

In congested flow every vehicle is (more or less) in a car-following state and maintains a constant nett time headway, h^c_{nett}, over the density range $k_2 < k < k_j$. This assumption implies a straight line for the congested part of the function $q(k)$. Hence in this aspect the model is the same as Daganzo's and Smulders'.

Derivation:

$$s_{nett} = uh^c_{nett} \quad \text{with } u = \text{mean speed} \tag{4.19}$$

$$s_{gross} = 1/k_j + uh^c_{nett} \tag{4.20}$$

$$k = \frac{1}{s_{gross}} = \frac{1}{1/k_j + uh^c_{nett}} \tag{4.21}$$

Solving for u leads to

$$u = \frac{1}{h^c_{nett}} \left(\frac{1}{k} - \frac{1}{k_j} \right) \quad \text{and apply } q = ku \tag{4.22}$$

yielding

$$q = \frac{1}{h^c_{nett}} \left(1 - \frac{k}{k_j} \right) \tag{4.23}$$

Parameter k_2: This parameter can be determined by assuming that the maximum speed of the congested state is at most equal to the speed that corresponds to 100 percent platooning at free flow state. Consequently using eq. (4.21):

$$k_2 = \frac{1}{1/k_j + u_p h^c_{nett}} = \left(u_p h^c_{nett} + \frac{1}{k_j} \right)^{-1} \tag{4.24}$$

Note the similarity of eq. (4.24) and eq. (4.20).

Example 42 *The model has 5 parameters. Typical values (standardized to lane values) for a two-lane roadway with 100 % cars are:* u_0 *= free flow speed = 110 km/h;* u_p *= speed of free flow platoon = 80 km/h;* k_j *= jam density = 150 veh/km;* h^f_{nett} *= nett time headway in free flow platoon = 1.2 s; and* h^c_{nett} *= nett time headway at congested flow = 1.6 s. Because the nett time headway at free flow is smaller than at congestion the two branches of the fundamental diagram in the u-k plane do not coincide. In the example the free flow capacity is 2400 veh/h and the discharge capacity 1895 veh/h (21 % less); see Fig. 2.8. Capacity drops found in practice are usually smaller.*

Remark 43 *The capacity drop offers a substantial possible benefit of ramp metering. If one can control the input flows such that the intensity on the main carriageway does stay below, say, 2200 veh/h per lane, than most of the time the smaller discharge capacity is not relevant. However, in practice the benefits of ramp metering are less than according to this reasoning.*

4.2.4 Diagram for roadway and lane

The interpretation of diagrams for separate lanes of a roadway and for the total roadway is not straightforward. The diagram for the roadway is not just the sum of the diagrams of the lanes because the distribution of the traffic over the lanes plays a role.

The mean speed as a function of intensity on a roadway of a motorway, the upper branch of $u(q)$, is often remarkably flat until rather high intensities. This can be partly explained by the changing distribution of the roadway intensity over the lanes. We explain this mechanism by means of a schematised situation.

Consider a two-lane roadway of a motorway with two types of vehicles; with free speed v_1 and v_2. The fast ones will use the left lane more and more with increasing roadway intensity. Denote

- q_t = roadway intensity
- q_r= intensity on right lane and q_l = intensity on left lane
- v_1 = speed of fast vehicles and v_2 = speed of slow vehicles
- c = slope of $u(q)$ for right lane
- α = fraction of q_t on right lane

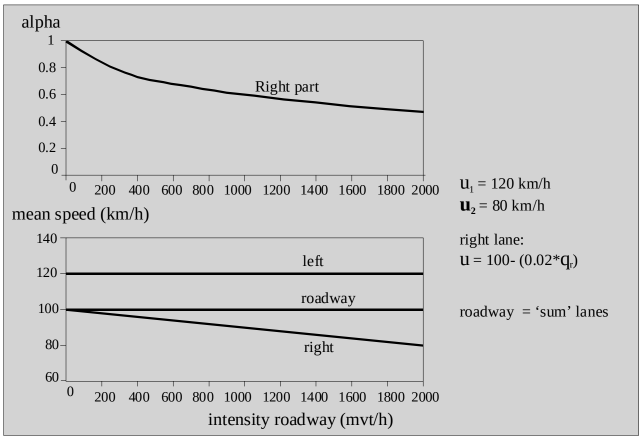

Figure 4.8: Distribution of roadway densities over the lanes and part of the fundamental diagram for roadway and lanes

The distribution of vehicles over v_1 and v_2 is 1:1. When q_t is near 0, all vehicles use the right lane and with increasing q_t more and more fast ones will use the left lane. We can divide the roadway intensity as follows:

$$q_t : \begin{cases} \text{Right} \quad \alpha q_t : \begin{cases} 0.5q_t \text{ (slow)} \\ \alpha q_t - 0.5q_t \text{ (fast)} \end{cases} \\ \text{Left} \qquad (1-\alpha)\, q_t \text{ (fast)} \end{cases} \tag{4.25}$$

yielding

$$u_r = \{0.5v_2 + (\alpha - 0.5)v_1\}/\alpha \tag{4.26}$$

We require that for the right lane the mean speed decreases linearly with its intensity:

$$u_r = 0.5\,(v_1 + v_2 - cq_r) \tag{4.27}$$

The question is which function $\alpha\,(q_t)$ brings about correspondence between (4.26) and (4.27). Set (4.26) = (4.27) and substitute $q_r = \alpha q_t$

$$\{(\alpha - 0.5)\,v_1 + 0.5v_2\}\,/\alpha = 0.5\,(v_1 + v_2) - c\alpha q_t \tag{4.28}$$

This is a second degree equation for α with the solution:

$$\alpha = \frac{-\delta v + \sqrt{\delta v^2 + 8cq_t\delta v}}{4cq_t} \quad \text{with} \quad \delta v = v_1 - v_2 \tag{4.29}$$

It turns out that $\alpha\,(q_t)$ corresponds qualitatively with reality.

Example 44 $v_1 = 120km/h$, $v_2 = 80km/h$, $c = 0.02$. *At* $q_t = 2000veh/h$ *all fast vehicles drive on the left lane and all slow ones drive on the right lane; Note:* u_r *is linear as function of* q_r *but not as function of* q_t.

4.2.5 Fundamental diagram based on a car-following model

Consider a situation in which vehicle i drives behind vehicle $i-1$. Vehicle i considers a gross distance headway of s_i. Both vehicles have the same speed v. Driver i includes the following in determination of the gross distance headway

1. Vehicle $i-1$ may suddenly brake and come to a complete stop
2. Driver i has a reaction time of T_r seconds
3. Braking is possible with an deceleration of a (with $a>0$)
4. When coming to a full stop behind the preceding vehicle, the net distance headway between the vehicles $i-1$ and i is at least d_0
5. The decelaration of the vehicle $i-1$ is α times the deceleration of vehicle i

It should be clear that the parameter α is somehow a measure for the aggressiveness of the driver: larger values of α imply that the follower assumes more abrupt deceleration of the follower, which will result in larger distance headways. It is left to the reader to determine that the minimal distance s_i to the leader of vehicle i

$$s_i + v_i^2 - \frac{1}{2\alpha a_i} = L + d_0 + v_i T_r + \frac{v_i^2}{2a_i} \tag{4.30}$$

or in words: the gross distance headway + braking distance of the leader = length of the follower + margin + reaction distance + braking distance follower.

If we assume that $s_0 = L + d_0$ (gross stopping distance headway = vehicle length + safety distance margin) then we have

$$s_i = s_0 + v_i T_r + \frac{v_i^2}{2a_i}\left(1 - \frac{1}{\alpha}\right) \tag{4.31}$$

To transform this microscopic relation into a macroscopic description of traffic flow, we need to make the following substitutions: replace the gross distance headway s_i by the inverse of the density $1/k$, replace the individual speed v_i by the flow speed u and replace the gross stopping distance headway s_0 by the inverse of the jam density $1/k_j$. We get

$$\frac{1}{k} = \frac{1}{k_j} + uT_r + \frac{u^2}{2a}\left(1 - \frac{1}{\alpha}\right) \tag{4.32}$$

Since $q = ku$, we get the following expression for the relation between flow and speed

$$q(u) = \frac{u}{\frac{1}{k}} = \frac{u}{\frac{1}{k_j} + uT_r + \frac{u^2}{2a}\left(1 - \frac{1}{\alpha}\right)} \tag{4.33}$$

Fig. 4.9 shows the resulting relation between the speed and the flow. Note that the speed is not restricted in this model.

Let us now show how we can determine the capacity from the car-following model. Fig. 4.9 shows how the flow rate depends on the speed for a large range of speeds for a certain set of parameters. We can now distinguish two situations

1. The flow speed u is – within certain boundaries – free and the drivers can choose it arbitrarily. This is the case for motorway traffic and it appears from practise that the maximum flow can be attained in this situations
2. The maximum flow speed u_m is prescribed and is lower that the optimal speed u_c yielding the capacity. In that case, the capacity is given by $q(u_m)$ (Eq. (4.33)).

If we first consider situation 1, then we can determine the speed u_c for which capacity results by taking the derivative of q with respect to u and setting it to zero. After checking that this indeed yields the maximum (and not the minimum), we can determine that the optimal (critical) speed, capacity and critical density are given by

$$u_c = \sqrt{\frac{2}{k_j}\frac{a}{(1-1/\alpha)}} \qquad q_c = \frac{u_c k_j}{2 + T_r u_c k_j} \qquad k_c = \frac{k_j}{2 + T_r u_c k_j} \tag{4.34}$$

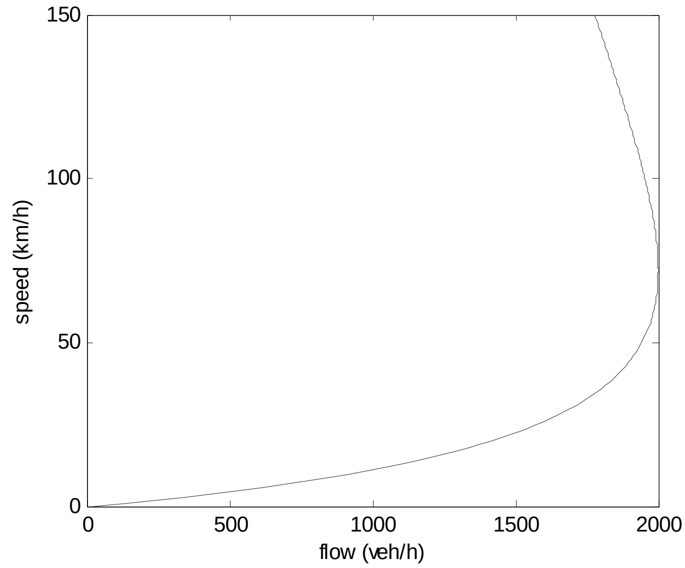

Figure 4.9: Flow - speed curve $u(q)$ derived from simple car-following strategy

Sensitivity of capacity

We can also express q_c as a function of α, T_r, k_j and a to determine the effect of changes in the parameters on the capacity

$$q_c = \frac{1}{T_r + \sqrt{\frac{2}{k_j}\frac{1-1/\alpha}{a}}} < \frac{1}{T_r} \tag{4.35}$$

From Eq. (4.35), we observe that the capacity increases when

- the reaction time T_r decreases;
- the jam-density k_j increases (shorter cars);
- the braking deceleration a is larger (improved braking system, e.g. ABS);
- the parameter α is closer to 1 (drivers are less cautious).

Furhtermore, it is remarkable that the critical speed u_c is not a function of the reaction time T_r. How can that be explained?

From Eq. (4.34) we can see that when $\alpha \to 1$ that $u_c \to \infty$; this implies that we need to restrict our choice for α by assuming that $\alpha > 1$. Clearly, this does not hold when there is a speed limit.

The model provides us an explanation for the following observations: it turns out that as time goes by, capacity increases steadily. For example, in the USA the capacity has increases from 2000 veh/h (in 1950) to 2400 veh/h (in 2000) *under ideal circumstances*. Using the derived model, we can explain this 20% increase. It is left to the reader to verify that the parameter α has the biggest influence on the capacity. We can thus explain the capacity increase by a decrease in α (drivers are more daring and are more experienced), a shorter reaction time T_r and higher deceleration a.

Brick wall following strategy

A special case of the model described here results when the following driver assumes that the leading vehicle is able to abruptly come to a full stop. This implies that the leader may have

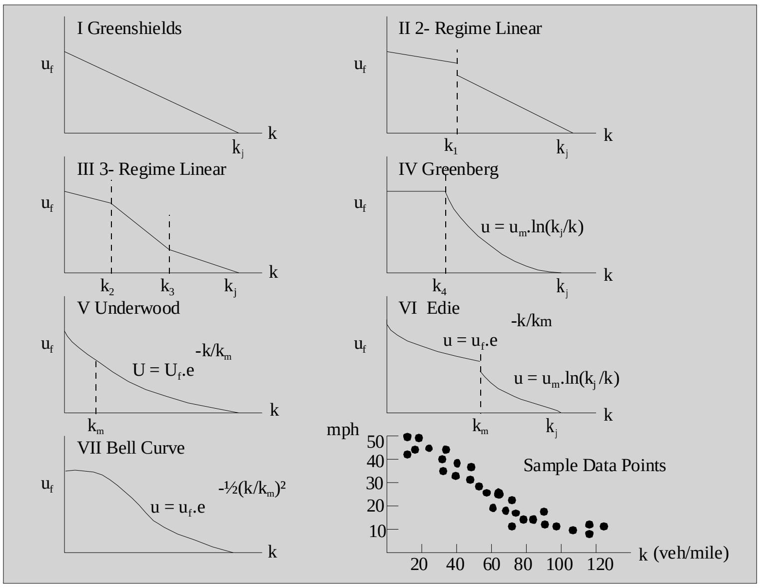

Figure 4.10: Models for $u(k)$ and a set of data points

an infinite deceleration – in practise, a very large deceleration. Consider for instance a car that crashes into a stopped truck. This pessimistic scenario is especially important for systems in which not the individual drivers have responsibility, but the operator (e.g. Combi-road or people movers).

4.3 Studies of the fundamental diagram

4.3.1 General points

If one wants to determine the fundamental diagram for a road section the following points are relevant:

- Does one need the complete diagram or only a part of it; e.g. only the free operation part ($k < k_c$) or only the congestion branch. A more fundamental point is whether it is possible to determine the complete diagram at one cross-section; see next section.
- Is the road section homogeneous? If this is the case, one can do with observations at a single cross-section. Otherwise road characteristics are variable over the section and a method such as the moving observer might be suitable.
- Period of analysis: If this is chosen too short, random fluctuations will have too much influence; if it is too long then it is questionable whether the state of the flow is stationary over the period. In practice the balance between randomness and stationarity has led to periods of 5 to 15 minutes.
- Finally, one has to estimate the parameters of the model chosen. This is mostly done by using a regression technique.

Fig. 4.10 (from [17]) shows that a given set of data points can be used to fit quite a few different models. In general models without too many parameters are preferred.

4.3.2 Influence and importance of data collection location

In order to get quantitative data about a fundamental diagram one has to carry out measurements at selected sites and periods. This will be illustrated with the data one can get when

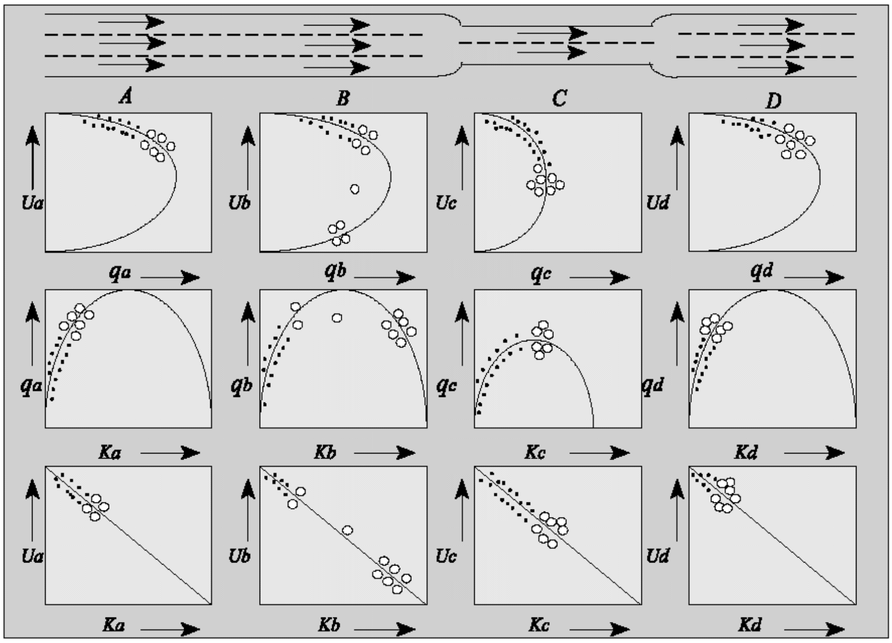

Figure 4.11: Traffic flow conditions at different cross-sections for a under and oversaturated bottle-neck

taking measurement around an overloaded bottle-neck. Fig. 4.11 depicts a roadway of a motorway of 3 lanes with a bottle-neck (b-n) section of 2 lanes wide. Measurements will be carried out at 4 cross-sections:

A This cross section is so far upstream the bottle-neck that congestion due to an overloading of the b-n will not reach it.

B This cross-section is closer to the b-n than A and congestion will reach it.

C A cross-section inside the b-n.

D A cross-section downstream of the b-n.

The fundamental diagrams of Greenshields have been assumed to hold for all cross-sections. They are the same for cross-sections A, B and D. For cross-section C the form is similar but the capacity and jam density are 2/3 of the values at A. We assume that the intensity increases gradually from a low value to a value that is just a little smaller than the capacity of the b-n; this capacity is $2C_0$ with C_0 = capacity of one lane. The data points resulting from such a demand pattern are depicted as * in the diagrams. At cross-section A, B and D intensity is not higher than 2/3 of the capacity. This means free operations with high speeds. At cross-section C the capacity is nearly reached. In a realistic case this would mean somewhat lower speeds but here that is not the case because Greenshield's diagram is used.

Now we assume demand increases further until a value that equals 2.5 times C_0 and discuss which data points we will get on the 4 cross-sections.

A It has been assumed that congestion will not reach this cross-section, so the state of the flow remains free. The data points, depicted by open circles, are in a range of 4000 to 5000 vehicles per hour and speeds remain high.

B When the higher demand reaches the beginning of the b-n, congestion will start, move upstream and reach cross-section B after some time. Before that moment data points are still on the free flow part of the diagram and afterwards on the congestion part. The intensity then equals (on average) the capacity of the b-n and the mean speed equals the speed corresponding to this intensity according to the congested part of the diagram (in this case around 30 km/h).

C In the b-n intensity is limited to the capacity value of the b-n and the state does not change much relative to the state at the end of the demand increase discussed before.

D Here the intensity is not larger than $2C_0$ because the b-n does not let through more vehicles and traffic operation remains free.

When demand is reduced to low values the process will develop in reverse order. Looking to the total results it appears that at cross-section A and D only free traffic operation can be observed. In the b-n one can observe the total free part of the diagram. Most information about the diagram is seen at cross-section B. But it should be realised that also here the information about the congested part of the diagram is rather limited. To collect data about the complete congested part of the diagram, one requires a b-n with a capacity varying from zero to, in this case, $3C_0$.

Concluding:

- The site of data collection determines which traffic flow states one can observe.
- Only in a bottle-neck one can observe a long lasting capacity state. If one wants to estimate capacity and has observations carried out at sites not being a bottle-neck, some form of extrapolation is always required.

An important consequence of the discussion above is that the capacity state can only occur in a bottle-neck. If one wants to determine the capacity from observations, not carried out in a bottle-neck, some form of extrapolation is needed.

4.3.3 Estimation of capacity using the fundamental diagram

In the light of the capacity drop two capacities do exist; free flow capacity and queue discharge capacity. Both values are relevant and require a different estimation method.

Estimating free flow capacity

Capacity can be estimated by fitting a model of the function $q(k)$ to data. The calculated maximum of $q(k)$ is an estimate of capacity. It turns out that this model is not suitable for nowadays motorway traffic, where it does not hold that the model for the free flow part of $q(k)$ has a maximum with the derivative $dq/dk = 0$ at capacity.

An alternative method is to assume a value of the capacity density k_c, fit a curve to a model for the free flow part of $q(k)$ and estimate the capacity by $q(k_c)$. It is obvious that the outcome of the estimation depends on the value of k_c and this makes the methods more suitable for comparisons than for absolute values. The procedure has been applied in determining the effect on capacity of road lighting; see [6].

"To estimate a capacity shift due to road lighting a before-and-after study set-up has been chosen. In its most simple form this implies estimating the average capacity of a roadway before and after installing of lighting. However, the disadvantages of such a 'naive' before-and -after study are great because there are so many factors that could have an influence on capacity and be different in both periods. Therefore the capacities have also been estimated under daylight conditions in before and after periods.

This set-up also enables determining the effect on capacity of darkness, without and with lighting, with daylight capacity as the reference situation. Such a comparison can give additional information on the relative contribution of lighting to a capacity improvement.

	Daylight		Darkness	
	Before	After	Before	After
Site I NB Treatment	100	99.1	90.3	92.9
Site I SB Treatment	100	96.8	93.5	95.4
Site II Treatment	100	100.4	95.5	97.1
Site III Comparison	100	98.8	95.1	94.4

Table 4.2: Set-up of study. Cells with double lining indicate cases with road lighting

In total, data from 3 motorway sites has been used. At Site I (motorway A50) data was available from both directions (North Bound=NB and SB), which was analysed separately. At the second motorway, A12, only data from one roadway was available. However, along the same roadway (approximately 14 km downstream) data was available from a road section with lighting in before and after period, which has been used as an extra comparison group. The total set-up is shown in Table 4.2, in which double lined cells represent cases with road lighting." Results in the table are presented as an index. The estimated capacities of the cells 'Daylight-Before' have been set at 100. Overall the positive effect of road lighting on capacity could be convincingly shown but it is rather small (2.5 % at the 2-lane roadway and 1.7 % at the three lane roadway).

Estimating queue discharge capacity

This capacity can only be estimated in an overloaded bottle-neck. Ideally three measuring stations are required:

- Upstream of the bottle-neck; data from this cross-section can be used to determine if overloading occurs (mostly mean speed is used as a criterion).
- Downstream of the bottle-neck; the state of the traffic flow should be free at this cross-section. If this is not the case, then a more severe bottle neck is present downstream.
- At the bottle-neck; intensities at this spot are capacity measurements if the traffic state upstream is congested and the state downstream is free.

Remark 45 *At an overloaded bottle-neck intensity at all three measuring stations is the same. It is an advantage to use the intensity at the downstream section to estimate capacity because traffic flow is more smooth than at the other stations and this reduces measuring errors.*

Remark 46 *The fact that large parts of motorway networks all over the world exhibit daily overloading is beneficial for the knowledge of capacity values.*

4.4 Capacity state on a motorway and the effect of rain

Finally an illustration of the fundamental diagram by means of data from a motorway (roadway with three lanes) near Rotterdam under two conditions: dry and fair weather, and rainy conditions; see Fig. 4.12 and 4.13.

The road section considered is more or less a bottle-neck with a length of 4 km. At a speed of 70 to 80 km/h drivers stay 3 to 3.5 minute in this traffic flow state. Apparently they are capable to deliver a high performance over this relatively short period, leading to high capacities. It is questionable if drivers are able to maintain this over a period of, say, one hour.

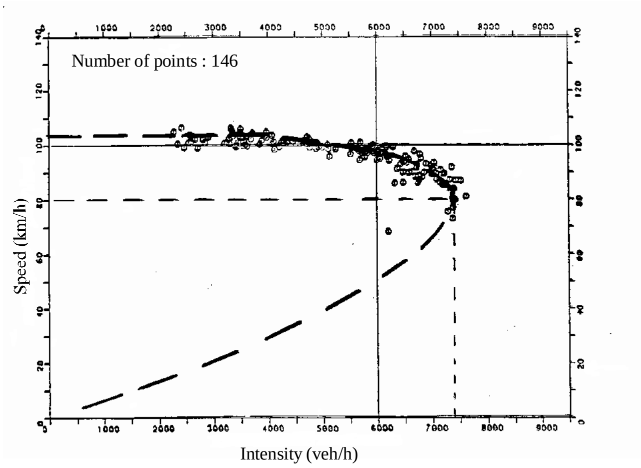

Figure 4.12: Relation speed-intensity in dry weather

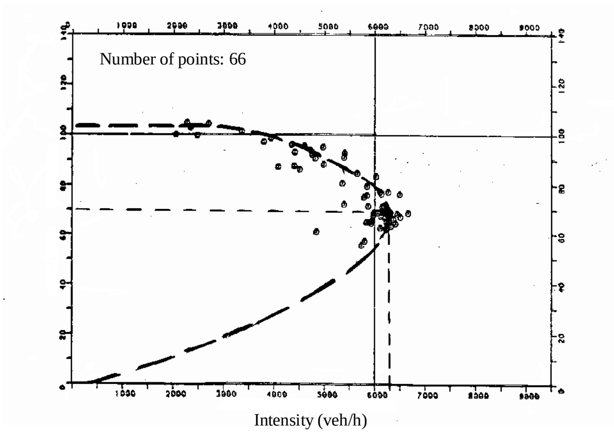

Figure 4.13: Relation speed-intensity in case of rain

MO-team	Roadway (3 lanes)			Average per lane			
	q_c	u_c	k_c	q_c	k_c	$\bar{s}$	$\bar{h}$
	veh/h	km/h	veh/km	veh/h	veh/km	m	s
Dry	7400	80	92.5	2460	30.8	32.4	1.46
Rain	6300	70	92.7	2100	30.9	30.9	1.71

Table 4.3: Characteristics capacity state at dry and wet weather