NEILA BHOURI, Ph.D.[1]
(Corresponding author)
E-mail: neila.bhouri@ifsttar.fr
MAURICE ARON, Statistician[1]
E-mail: mauraron@gmail.com
HABIB HAJSALEM, Ph.D.[1]
E-mail: habib.haj-salem@ifsttar.fr
[1] University Paris-Est, IFSTTAR
  COSYS Department, GRETTIA Laboratory
  14-20, Boulevard Newton, 77447 Champs sur Marne,
  Cedex 2, France

A DATA-DRIVEN APPROACH FOR ESTIMATING THE FUNDAMENTAL DIAGRAM

ABSTRACT

The fundamental diagram links average speed to density or traffic flow. An analytic form of this diagram, with its comprehensive and predictive power, is required in a number of problems. This paper argues, however, that, in some assessment studies, such a form is an unnecessary constraint resulting in a loss of accuracy. A non-analytical fundamental diagram which best fits the empirical data and respects the relationships between traffic variables is developed in this paper. In order to obtain an unbiased fundamental diagram, separating congested and non-congested observations is necessary. When defining congestion in parallel with a safety constraint, the density separating congestion and non-congestion appears as a decreasing function of the flow and not as a single critical density value. This function is here identified and used. Two calibration techniques – a shortest path algorithm and a quadratic optimization with linear constraints – are presented, tested, compared and validated.

KEY WORDS

non-analytical; calibration; empirical data; shortest-path algorithm; convex quadratic program; safety constraint; critical density function;

1. INTRODUCTION

Traffic flow theory is the basis for understanding, controlling and predicting the movements of vehicles. It deals with variables at different levels of aggregation in time or space and determines relations between them [1].

The most popular, simplest and oldest relation in the traffic flow theory is the fundamental speed-density relationship. Its history and developments are described in [2–3]. The graphical representation of the relationship between any two of these variables is called the fundamental diagram (FD). It describes how speed decreases with density – this decrease is due to safety reasons.

An analytical FD allows an easier representation of traffic phenomena and is often necessary in at least three cases: when relations between car-following models and the FD are investigated [4]; when considering a stochastic traffic flow model, which leads to specifying a stochastic FD and making analytical assumptions about its form and the form of its random fluctuations [5]; when constructing explicit solutions to the Lighthill-Whitham-Richards (LWR) traffic flow model [6].

In other cases, however, such as with macroscopic traffic simulation models or assessment studies [7], an analytic FD is not necessary. Furthermore, recent sensor approaches, such as floating vehicle data (FCD) or similar means, have a wide spatial coverage regarding their measurements but do not provide information about traffic state variables such as flow or density. An already calibrated FD is used for the estimation of these variables ([8] and [9]). The use of an analytical FD does not allow the estimation of state variables for high speeds since the FD curve often has a very flat branch near free-flow speeds. Using an empirical FD allows to take into account all data and to have more accurate values for different speeds.

In this contribution, the proposed speed-density relationship is neither analytic nor a set of analytic functions, but just the function which best fits the data.

To respect the traffic flow theory, it is mandatory to have a decreasing speed as a function of density. When density increases, the average space headway between two consecutive vehicles decreases, thereby decreasing the space gap. A decreasing speed must accompany this decreasing gap; otherwise, the stopping distance (which increases with the speed) becomes unsafe. Appendix A shows that, when defining congestion in parallel with a safety constraint, the density separating congestion and non-congestion is, most often, a decreasing function of the flow. This function generalizes the critical density value; it

impacts the fundamental diagram. Additionally, it should be noted that, for high flows, the safety constraint might be not complied with.

The only constraint used to build the non-analytical FD in this paper is to respect the assumption that speed decreases when density increases.

In this paper, the use of the FD and the state of the art are presented in Section 2. The type of data and the calibration methodology are presented in Section 3. Two approaches are explored for determining the FD. In the first one, the shortest path approach (SPA, Section 4) gives the congested and non-congested speeds related to the flow; the non-congested speed is constrained to decrease when the flow increases, whereas, in congestion, speed is constrained to increase with the flow. In the second one, a decreasing speed related to density is obtained due to an algorithm – the linear quadratic optimization with linear constraints (QOLC, Section 5).

Section 6 is dedicated to validation and transferability. Then, some conclusions and perspectives are outlined.

Two appendices are dedicated to the density threshold separating congestion and non-congestion – the first one in relation with the safety constraint and the second one in relation with a particular FD (Underwood, [10]). These appendices show that this density threshold is a decreasing function of the flow.

2. THE USE OF THE FUNDAMENTAL DIAGRAM

According to Coifman [11], "much of traffic flow theory depends on the existence of a fundamental relation between flow, Q, density, K, and space mean speed, V." In first-order traffic models, an FD is used in conjunction with a conservative equation, initial conditions and conditions on demand [12] [13]. In classical second-order models [14], the FD is included in a speed equation that takes into account dynamic space and time effects. The new family of second-order models, the generic second-order modeling (GSOM), combines a first-order model with the dynamics of driver-specific attributes [15].

A number of analytical speed-flow or speed-density relationships have been proposed and calibrated, including by Greenshields [16], Greenberg [17], Edie [18], Underwood [10] and May [19]. These models, however, do not take into account certain phenomena such as spontaneous congestion, random variations, capacity drop and hysteresis, that are described in [20], [21]. A functional form for the FD, based on generative functions applied on an inverse of a generalized space interval, is proposed in [22]. This functional form is the solution of a system of constraints (speed decreases with density; a concave flow density relationship...).

With a global form being somewhat restrictive, a set of piecewise linear regressions that is able to reproduce the capacity drop is successfully proposed in [23].

Kerner [24] argued that congestion includes two different traffic phases: the synchronized traffic (when the downstream front is fixed at a bottleneck) and the wide moving jam. This reduces the scope of use of the FD, which should be used only in cases where the congestion structure is near the one prevailing when the FD has been calibrated.

Forerun by Daganzo & Geroliminis [25], who provided empirical evidence for the existence of an urban-scale FD, many promising approaches and applications explore network-scaled relations between vehicle density and space mean flow – called macroscopic fundamental diagram [15, 26].

In this paper, an empirical fundamental relationship is reworked to ensure that speed decreases with density. Until recently, a low computational cost was required, but nowadays computers perform fast calculations, even with a great number of parameters. When a large number of parameters is justified and does not decrease the robustness of the model, there is no justification for calibrating a simplified model using only a few parameters, which implies a loss of information. Building an FD which best fits the data leads inevitably to increased accuracy.

3. TYPE OF DATA AND THE CALIBRATION PROCESS

3.1 Type of data used for calibrating an FD

Data for calibrating an FD consists either of trajectory data [27] or of loop data by vehicle at some points, or of loop data aggregated for a time period (flow, occupancy and possibly average speed).

The first category of data enables comprehension of phenomena and accurate preprocessing in order to remove the noise. The second category allows specifying FDs by length of vehicle [28] and acutely studying the variability. Data of the third category are more commonly available and used.

The fundamental equation of traffic flow establishes the relationship between the three main macroscopic variables – flow, speed and density. If a relationship is established between any two of the variables, the relationship of the third one can be controlled by the following equation:

$$Q = K \cdot V_s \qquad (1)$$

where K is the density, Q is the flow and V_s is the space mean speed.

The space mean speed is computed when the speed v_j is available for every vehicle j on the whole section – the space mean speed is their harmonic speed average.

When only flow and density are available, the space mean speed is defined as their ratio.

When data come from a dual loop detector, the spot speed of vehicle j when it reaches the detector is marked as v_j; although these spot speeds are different from the previous definition, their harmonic average is used for the space mean speed:

$$\frac{1}{V_s} = \frac{1}{Q} \sum_{j=1}^{Q} \frac{1}{v_j} \qquad (2)$$

When data come from a loop detector giving only flow and occupancy τ, and if an estimation of the average length of vehicles is available, a relation analogous to *Equation 1* arises. Indeed, let l be the length of the loop detector and L_j be the length of vehicle j; each vehicle j passes the sensor during a time equal to $\frac{L_j + l}{v_j}$, v_j is the spot speed of vehicle j at the detector; occupancy (for one time unit) is the sum of the passage times of vehicles passing the detector – their number, in a time unit, is the traffic flow Q:

$$\tau = \sum_{j=1}^{Q} \frac{L_j + l}{v_j} \qquad (3)$$

Let L_{eff} be the effective average of lengths L_j, weighted by $\frac{1}{v_j}$.

$$L_{eff} = \frac{V_s}{Q} \cdot \sum_{j=1}^{Q} \frac{L_j}{v_j} \qquad (4)$$

implies

$$L_{eff} + l = \frac{V_s}{Q} \cdot \sum_{j=1}^{Q} \frac{L_j + l}{v_j} = \frac{V_s}{Q} \cdot \tau \qquad (5)$$

thus

$$Q = \frac{V_s}{L_{eff} + l} \cdot \tau \qquad (6)$$

Density K in *Equation 1* is replaced in *Equation 6* by the ratio $\frac{\tau}{Leff + l}$. It is not really the spatial density required by the traffic flow theory, but it is suitable for a lot of traffic models which only simulate what happens at the beginning or at the end of sections.

3.2 Calibration of the fundamental diagram

Data used in the following sections consist of six-minute records of average speed and flow (Q_i, V_i), with i as the time index; density K_i is derived as $\frac{Q_i}{V_i}$. Data were collected in 2009–2010 on the A1 motorway linking Paris to Lille, in the north of France. The speed-density or speed-flow relationships are calibrated using a dual loop detector on a section situated at 4 km from the city of Lille. The speed there was limited to 110 km/h as on other urban motorways in France; since traffic operators wanted to reduce this speed limit to 90 km/h, an ex-ante assessment of this measure was made [7], where fundamental diagrams, as close as possible to the data, were required to facilitate the comparison between the simulated new scheme and the empirical data. This was the motivation for this paper. On this road section, the motorway has five lanes in each direction. The relationships presented here are calibrated for the third lane, from Paris towards Lille. We used data from the year 2010. After the elimination of empty periods (six-minute periods without vehicles) and irrelevant data (speed < 2 km/h or speed > 200 km/h, or traffic flow ≥ 320 vehicles/6 minutes), 58,000 six-minute data were used out of the 87,600 of total data.

The curve corresponds to the generalized exponential model $Q = V_f \cdot K \cdot e^{-(K/K_0)^\alpha}$ with V_f=111.9 km/h, K_0=36.8, α=1.48.

Drivers might adapt their speed with respect to their time headway, it being linked to the flow, which is in a certain unit and is the inverse of the average time headway. The data used to calibrate a speed-flow
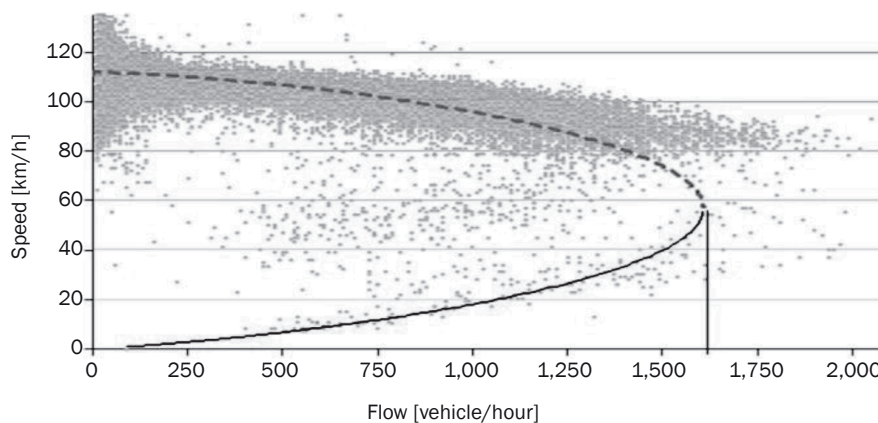

Figure 1 – Speed-flow FD for motorway A1: plotted empirical data corresponds to the third lane, January-June 2010

relationship often consists of flows (Q_i) and average speeds (V_i) at a traffic detector for a set of periods i; density is estimated by the ratio $\frac{Q_i}{V_i}$.

The FD generally does not synthesize the data properly because these data are, for different reasons, very scattered (see *Figure 1*). An important cause of variability arises from inhomogeneous conditions due to accelerations, transient states, platoons, "synchronized" vehicles. In an inhomogeneous period, the observed flow Q is partly obtained during non-congested sub-periods with low occupancy (i.e., low density) and partly during congested sub-periods with high occupancy (i.e., high density).

To calibrate a speed-flow diagram, it is necessary to use a threshold for separating periods of non-congestion and periods of congestion in the empirical data. Non-congested periods contribute to the calibration of the non-congested branch of the FD, where the speed $V^{free}(q)$ is a decreasing function of the flow, and the ratio $K^{free}(q) = \frac{q}{V^{free}(q)}$ is increasing. Congested periods are used to calibrate the congested branch of the FD, where the speed $V^{congested}(q)$ increases with the flow – the ratio $K^{congested}(q) = \frac{q}{V^{congested}(q)}$ not being constrained to decrease.

The two branches of the FD meet at the point where the flow is maximum – the speed is then the critical speed, and the density (the capacity divided by the critical speed) is the critical density. The capacity might be taken as the highest traffic count observed; the critical speed is the average of harmonic mean speeds observed at periods of the highest traffic count.

When browsing the non-congested branch of the FD from zero flow to capacity, the ratio $K^{free}(q) = \frac{q}{V^{free}(q)}$ increases from zero to the critical density. Also, the critical density is generally used to separate congested and non-congested points. However, the lower flows are little concerned by the critical density. Nothing prevents defining congestion (then separating congested periods from non-congested ones) not with a single critical density threshold, but with a threshold function of the flow $k(q)$. This function is assumed to be monotonous – swinging is not explainable and should result from over-identification of the FD from the dataset.

When it comes to a density threshold function, there are several arguments for decreasing rather than increasing with the flow:
- It results in a better numerical adjustment (see Remark 2).
- A parallel between the density required by a safety constraint, and the density threshold function shows that this function is decreasing, at least until the flow reaches a certain value (see Appendix A).

– Due to individual or contextual (meteorology,...) reasons, some drivers avoid high speeds (at free-flow) but have a common behavior at capacity. This speed reduction implies an increase of the flow/speed ratio (thus of density) which vanishes when the flow increases to capacity.

At a given flow q, the empirical dataset (q_i, v_i) leads to a dataset in the (flow x density) space $\left(q_i = q,\ k_i = \frac{q_i}{v_i}\right)$. For a given flow q, the splitting of the $\left(k_i = \frac{q}{v_i}\right)$ in two parts – round $K^{free}(q)$ and $K^{congested}(q)$ – is optimal (i.e., the deviation is minimum) when the threshold on k_i is equal to $\frac{K^{free}(q) + K^{congested}(q)}{2}$. This function decreases when q increases for the well-known Underwood FD [10] (see Appendix B). This is a clue for a decreasing density threshold.

The equations and the process giving the FD (within the above constraints) are below. For readability, the width of each class is one unit, so there are q_{max} classes for a flow from 1 to q_{max}. The periods are grouped in a class according to their flow (subscript q).

For every flow class q, let n_q be the number of periods of the class, let Q_q^j and V_q^j, $j=1...n_q$ be the empirical flows and average speeds, respectively.

Without any loss of generality, for every flow class q, the $\left(Q_q^j, V_q^j\right)$ pairs are assumed to be sorted when j varies from 1 to n_q, according to the ratio $\frac{Q_q^j}{V_q^j}$ (the first j corresponds to the smallest ratio).

Let $W_{q,k}^{free}$ and $W_{q,k}^{congested}$ be the average of the empirical speeds V_k^j when density is smaller or greater than k, respectively.

Let $n_{q,k}$ be the number of class q periods with density smaller or equal to k.

If considered that, for the flow class q, k is the density threshold separating congested and non-congested periods, the empirical FD will consist of two values at the traffic flow q.

The average of non-congested empirical speeds is:

$$W_{q,k}^{free} = \sum_{j=1}^{n_{q,k}} \frac{V_q^j}{n_{q,k}} \tag{7}$$

The average of congested empirical speeds is:

$$W_{q,k}^{congested} = \sum_{j=n_{q,k}+1}^{n_q} \frac{V_q^j}{n_q - n_{q,k}} \tag{8}$$

The measure of the scatter at the flow level q is the residual between the observed and modeled speeds; if the threshold separating congested and non-congested regimes were the density $k(q)$ (typically, whatever is q, the critical density is k_{crit}) the measure of the scatter would be the following:

$$E^2_{q,k(q)} =$$

$$= \left(\frac{1}{n_q}\right)\left[\sum_{j=1}^{n_{q,k(q)}}\left(V^j_q - W^{free}_{q,k(q)}\right)^2 + \sum_{j=n_{q,k(q)}+1}^{n_q}\left(V^j_q - W^{congested}_{q,k(q)}\right)^2\right] \quad (9)$$

The clustering of the n_q pairs in two classes – from 1 to $n_{q,kcrit}$ and from $n_{q,kcrit}$ +1 to n_q is not optimal for decreasing $E^2_{q,k(q)}$. Additionally, the paper proposes to identify the solution of function $k(q)$ (Equation 10) for the whole set of traffic flows:

$$E^2_{SPA} = \left(\frac{1}{n}\right)\min_{k(q)}\sum_{q=1}^{q_{max}} n_q E^2_{q,k(q)} \quad (10)$$

with two constraints:

$$k_{max} \geq k(q\text{ - }1) \geq k(q) \geq k_{crit} \quad (11)$$

$$W^{congested}_{q\text{-}1,k(q\text{-}1)} \leq W^{congested}_{q,k(q)} \leq W^{free}_{q,k(q)} \leq W^{free}_{q\text{-}1,k(q\text{-}1)} \quad (12)$$

The minimization provides $k(q)$, decreasing with q.

Remark 1. At a given flow, the level of $k(q)$ impacts both $W^{congested}_q$ and W^{free}_q. For instance, $k(q)$ higher than the critical density implies that some periods (those the density of which is comprised between k_{crit} and $k(q)$) are no longer considered to be congested. Their speed is lower than the average non-congested speed and higher than the average congested speed. As these periods pass from congestion to non-congestion, this makes both $W^{congested}_{q,k(q)}$ and $W^{free}_{q,k(q)}$ lower than $W^{congested}_{q,kcrit}$ and $W^{free}_{q,kcrit}$, respectively. In turn, if for all flows q the density thresholds $k(q)$ are higher than the critical density, both branches of the empirical FD will be lower. The higher the density thresholds, the lower the empirical speeds for both branches.

The following are the observations made on the main features of this approach: the graph, the constraint of a decreasing speed related to density, the algorithm.

Nodes and links

The SPA leads to an approximation of the solution in q_{max} steps. The nodes are quadrupled (q, W^{free}, $W^{congested}$, k), where q is a flow class; k is a density $k_{min} \leq k \leq k_{max}$; W^{free} and $W^{congested}$ are speed values, included in specified discrete intervals, typically: W^{free}, $W^{congested}$, and k are integers, in the units specified by the user. This constraint avoids the creation of an infinite number of nodes. Then, it is recommended (not mandatory) to reduce the number of nodes created by adding the constraints:

$$V_{min}(q) \leq W^{congested} \leq V_{max,congested}(q)$$
$$V_{min,free}(q) \leq W^{free} \leq V_{max}(q) \quad (13)$$
$$W^{congested} < W^{free}$$

where $V_{min}(q)$, $V_{min,free}(q)$, $V_{max,congested}(q)$, $V_{max}(q)$, k_{min}, k_{max} are specified by the user.

Speed must decrease with density

Equation 12 does not always imply a decreasing speed with density. Both parts of the FD, where the functions are monotonous, are examined below:

1) The non-congested part

In this condition, the two last terms of *Equation 12* imply that, when q varies, the speeds $W^{free}_{q,k(q)}$ can be extended in a decreasing function related to the flow $V(Q)$. It can be assumed that this function is continuous and differentiable and $V'(Q)$ is derivative related to Q. With $V(Q)$ decreasing, $V'(Q)$ is negative (or null). The fundamental diagram says that Q is a function $Q(K)$ of the density. Then the composition $V(Q(K))$ (i.e., $V \circ Q$) can be formed, giving the speed related to the density.

We show below that vice versa also applies, that speed also decreases with density (i.e., the constraint is satisfied). Indeed, as $K(Q) = \dfrac{Q}{V(Q)}$ is the density related to the flow, its derivative related to Q is:

$$K'(Q) = \frac{1}{V(Q)} \text{-} Q \cdot \frac{V'(Q)}{V^2(Q)} \quad (14)$$

As $V'(Q)$ is negative, *Equation 14* implies that $K'(Q)$ is positive and $Q'(K)$, the derivative of Q related to K, is also positive. Indeed, at every point $K=k$, the value of $Q'(K)$ is the inverse of the value of $K'(Q)$ at $Q=q$.

The derivative related to K of the compound function $V \circ Q$ is:

$$\frac{dV}{dK} = V'(Q(K)) \cdot Q'(K) \quad (15)$$

It is negative because $V'(Q(K))$ is negative, whereas $Q'(K)$ is positive.

2) The congested part

In this part, the two first terms, $W^{congested}_{q\text{-}1,k(q\text{-}1)}$ and $W^{congested}_{q,k(q)}$, of *Equation 12* are extended in an increasing function (also marked as $V(Q)$, assumed to be continuous and differentiable with $V'(Q) \geq 0$ as well). This does not imply that the speed also decreases with density. Indeed, as $V'(Q(K)) \geq 0$, $K'(Q)$ given in *Equation 14* is the sum of positive and negative terms, it is, therefore, not always negative. This is verified only if the following applies:

$$Q \cdot V'(Q) \geq V(Q) <=> V'(Q) \geq \frac{1}{K} \quad (16)$$

This constraint is introduced in the SPA by removing the links between the nodes that do not respect this constraint.

The steps of the algorithm

1) Initial step, for q=1

For the flow class q=1, nodes (q=1, W^{free}, $W^{congested}$, k) are reached at a cost depending only on k, equal to:

$$C(q, W^{congested}, W^{free}, k) = n_{q,k} \cdot \left(W^{free} - W_{q,k}^{free} \right)^2 + \\ + \left(n_q - n_{q,k} \right) \cdot \left(W^{congested} - W_{q,k}^{congested} \right)^2 \qquad (17)$$

When $W_{q,k}^{free}$ or $W_{q,k}^{congested}$ are not defined, the node is removed.

It is easy to see that the optimization of *Equation 10* is equivalent to the optimization of the deviation between the empirical speed averages (by flow class) and the FD speeds. This deviation is equal to the square root of the sum of *Equation 17* for all flows *q*, divided by the total number of periods:

$$D_{SPA} = \sqrt{\frac{1}{n} \sum_q C(q, W^{congested}, W^{free}, k)} \qquad (18)$$

Note that the speed averages might be replaced by their median in *Equations 7–9* and *17*.

2) Steps 2 to q_{max}

For $q>1$, links are created towards every node $(q, W^{free}, W^{congested}, k)$, from existing nodes of the preceding flow level *q*-1 $(q\text{-}1, W^f, W^c, k')$, for any W^f greater or equal to W^{free}, any W^c smaller or equal to $W^{congested}$ and any k' greater or equal to *k*. These links have the same cost, depending only on *q* and *k*, equal to $C(q, W^{congested}, W^{free}, k)$, as defined in *Equation 17*.

When $W_{q,k}^{free}$ or $W_{q,k}^{congested}$ are not defined, the cost of the link is assigned to a very high value. When the flow class *q*-1 is empty, *q*-1 is replaced by *q*-2, etc.

3) Final step

The algorithm ends when the highest flow class is treated. The speed-flow relationship is provided by the path(s) leading to the nodes of minimal cost of the last step $(q=q_{max})$.

The number of nodes, then the memory and the time required for applying the algorithm, depend on the width of flow classes and on the size of the discrete sets used for speed and density.

3.3 Results

Figure 2 shows the relationship obtained with the dataset used in this Section 3 for the year 2010 (width of flow classes: 50 vehicles/hour, speed unit of 1 km/hour, density unit 1 vehicle/km).

The optimization provides a deviation D_{SPA} (between the empirical average speeds and the FD speeds) equal to 0.35 km/h.

The critical speed is 83 km/h; the critical density is 26 vehicles/km. For very low flows (less than 450 vehicles/hour) there were no data for estimating the speed in congestion. The capacity (2,190 vehicles/hour) comes from the maximum traffic count observed (219 vehicles) in a six-minute period. This points to the main drawback of the approach: the algorithm aims to reach
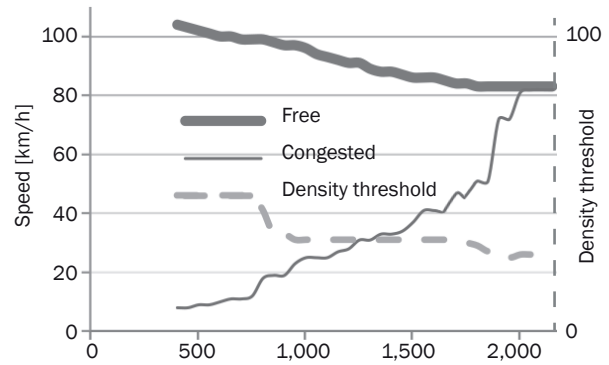

Figure 2 – Speed-flow relationships with the SPA

at least one node in the highest traffic flow class, even if the number of periods which constitutes this class is low. This can be avoided either by:

– eliminating the highest flow class(es),
– grouping the highest flow classes into a single class,
– simplifying *Equations 7, 8, 9, 12, 17* and *18* and, in the final step of the algorithm, not distinguishing the congested and non-congested speeds at $q=q_{max}$.

Remark 2. Solving the problem with an increasing (instead of decreasing) density threshold would lead to lessening the adjustment: the deviation D_{SPA} would increase from 0.35 to 0.4 km/h.

Remark 3. The same approach applies to estimating a density-flow relationship, which provides, in turn, a flow-density relationship. The roles of speed and density are inverted, and so are the roles of congestion and non-congestion.

Remark 4. The simplified form of this approach would apply to calibrate a speed-density relationship: instead of flow classes, density classes *k* are considered; for every density $k(k>0)$ and admissible speed *W*, a node (W, k) is created; for every speed *Z*, $0 \leq Z \leq W$, a link is created between (W, k) and $(Z, k+1)$; the cost of every link towards this node is $n_k \left(W - \overline{V}_k \right)^2$, where n_k is the number of observations of class *k*, and $\overline{V}_k = \sum_{j=1}^{n_k} \frac{V_k^j}{n_k}$.

4. SPEED-DENSITY RELATIONSHIP CALIBRATED BY QOLC

The estimated density $K_i = \frac{Q_i}{V_i}$ is generally lower than 80 vehicles/km. In this section, we built density classes of subscript $k=1...m$ – here $m=160$, and the width of a density class is 0.5 vehicle/km. To each period *i*, a density class *k* is assigned according to the value of K_i. Data become (Q_k^j, V_k^j), the flow and average speed of the *j*-th period, the density of which belongs to class *k*.
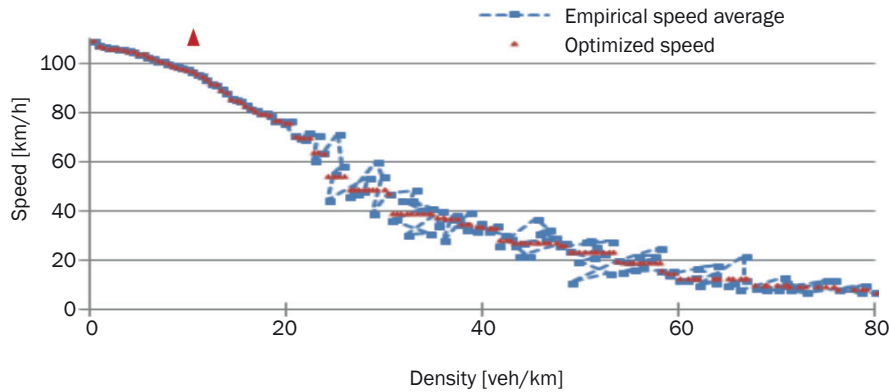
Figure 3 – Speed-density relationship with the QOLC approach

The objective is to find the speed vector (W_k), decreasing with density k, the closest to the data. It is the solution for the quadratic programming problem with linear constraints:

$$E^2 = \min \sum_{k=1}^{m} \sum_{j=1}^{n_k} \left(W_k - V_k^j \right)^2 \qquad (19)$$

with the following constraints:

$$W_k \geq W_{k+1} \quad \text{for} \quad 1 \leq k < m \quad \text{and} \quad W_m \geq 0 \qquad (20)$$

where n_k is the number of periods for the density class k.

This problem is solved using the software R® and SCICOS® packages [29]. If the empirical mean speed for class k is marked as $\overline{V}_k$, it is easy to see that replacing V_k^j with $\overline{V}_k$ does not change W_k. Therefore *Equation 19* can be replaced with:

$$D_{QOLC}^2 = \frac{1}{n} \min \sum_{k=1}^{m} n_k \left(W_k - \overline{V}_k \right)^2 \qquad (21)$$

The empirical speed average and the optimized speed are plotted by density class in *Figure 3*. There are very few differences between both curves.

The capacity of the road is derived as the maximum of products $k.W_k$ whatever the density class k, $1 \leq k \leq m$. This maximum is reached when the density class k is 25 veh/km; in this class, the average density, average speed and flow were K=25.2 vehicles/km; V=71.5 km/h, thus, Q=1,802 vehicles/h. This critical density is in accordance with the critical density of the first approach (26 vehicles/km). However, the capacity is lower than the one of the first approach (Q=2,179 vehicles/h). Indeed, W_k is linked to the average and not to the maximum speed recorded in class k, it is the same for the product $k.W_k$. So, this technique underestimates the capacity.

The sensitivity of the results to the number of classes and to their width was tested with m=80 density classes of width 1 (instead of 0.5) vehicle/km; the optimized speeds remained very close (less than 0.7 km/h) to the previous in most cases, except for density 41–43 (the speed difference was 1.2 km/h), and for density 23 (the speed difference was 2.2 km/h).
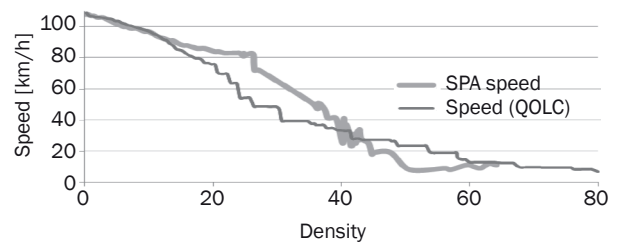

Figure 4 – Speed-density relationships obtained by both approaches

Remark 5. Reversing the roles of speed and density, a density-speed relation can be calibrated. Then, speed classes are built, which should be in accordance with [23], where constant-speed fluctuations are highlighted. The average of the empirical densities for a speed class v, marked as $\overline{K}_v$, replaces $\overline{V}_k$ used for density class k. To follow [23], this average may be replaced in *Equation 21* by the median of the (K_v^j, j) being such a period that its average speed is in class v.

In *Figure 4*, where the X-axis is density and the Y-axis is speed, the QOLC speed-density relationship and the points derived from the SPA relationships were plotted concomitantly.

The following can be seen:
– Small oscillations of the SPA speed for high densities (greater than 43 vehicles/km). These appear because the additional constraint of decreasing speed with density was not implemented.
– A probable underestimation of capacity in the QOLC approach.
– A maximum density higher for the QOLC than for the SPA. Indeed, the SPA uses flow classes; each flow class including two sub-classes, the first one for the congested values (high densities), the second one for the non-congested values (low densities); the mean density of the first sub-class, since it is a mean, is always lower than the maximum density. This type of aggregation is not made in the QOLC.
– Some differences between the SPA and QOLC speeds for medium and high densities. The SPA speeds are higher than the QOLC ones for

densities between 20 and 40 vehicles/km, then lower for densities greater than 40 vehicles/km (congestion).

Are these differences explainable?

1) There is a possible explanation for lower SPA speeds at high densities (beyond 40 vehicles/km): when flows are low, high-density thresholds have been used for specifying the congestion. A low SPA FD speed is in accordance with Remark 1 of Section 4 – the higher the density thresholds, the lower the empirical speeds for both FD branches.

2) There is no direct explanation for high SPA FD speeds at higher flows (densities between 20 and 40 vehicles/km): the density thresholds used are no longer high, which makes the SPA FD speeds not lowered, but this does not explain why they are higher than the QOLC FD speeds. Much lower density thresholds should have been used to make the SPA and QOLC FD speeds equal. This highlights the sensitivity of the SPA FD speeds to these density thresholds. It is both a benefit and a danger for the calibration of a speed-flow FD. This can contribute to feeding the debate on the still sensitive subject of the fundamental diagram.

5. VALIDATION AND TRANFERABILITY

The assessment of the FD is based on the deviations between the FD and empirical data. For a speed-density relationship (and a density class pattern), the mean squared error (MSE) is given according to the value $\frac{E^2}{n}$, where E^2 is given in *Equation 19*. E cannot be lower than the value obtained when $W_k = \overline{V}_k$, the empirical speed average. In this case, $\sqrt{\frac{E^2}{n}}$, computed on the calibration dataset is equal to 7.6 and 8.5 km/h for the QOLC and the SPA approaches, respectively.

The validity of the non-analytic approach of this paper is discussed below. This discussion is based on the deviation between the obtained FD and the empirical speed average, by class of density in the QOLC method or by flow class in the SPA. These deviations are equal to D_{SPA} or D_{QOLC}, the quantities minimized in *Equation 18* for the SPA approach and in *Equation 21* for the QOLC. The penultimate and ultimate columns of *Table 1* give D for the QOLC and the SPA methods.

D must be low not only on the calibration dataset but also on other datasets. The results were validated on other datasets of the same section of the same motorway. *Table 1* gives the results of the calibration, validation, and transferability of the FD. The calibration is based on the first six months of the year 2010 (*Table 1*, line 1), the validation on the last six months of the same year (*Table 1*, line 2) or on the lane of the same section with the same lane number in the opposite direction (line 3).

The transferability of the method is assessed by the values of D obtained by applying the same calibrated FD on the faster or slower lanes of the same motorway section (*Table 1*, lines 4 and 5) or on middle lanes of other close motorways (A22, A25, lines 6 and 7).

Deviations are very small on the calibration set as well as on the validation datasets; this validates the methods. The calibration of the FD is rarely transferable to the other lanes of the motorway, or to the same lane of other motorways. This is not surprising. Since the objective is to be the closest to the data, a new calibration is needed for each lane.

6. CONCLUSION

We developed a methodology in this paper to establish a fundamental diagram which best fits the data, free of any analytical form, considering just the assumption that drivers adapt their speed to their

Table 1 – Deviation (in km/h) between the average empirical speed and the FD speed

Line	Section	Lane[1]	Year	No of 6' Periods	Deviation [km/h]	
					QOLC[4]	SPA
1*	A1.D1[2]	3rd/5	January–June 2010	23,406	0.7	0.3
2	A1.D1	3rd/5	July–December 2010	35,306	1.4	1.1
3	A1.D2	3rd/5	2010	58,216	1.8	1.9
4	A1.D1	2nd/5	2010	58,417	7.9	2.8
5	A1.D1	4th/5	2010	58,888	12.2	4.7
6	A1.D1	3rd/5[3]	2010	52,405	7.4	4.3
7	A22	3rd/5	2010	61,038	7.2	4.8
8	A25	3rd/5	2009	42,919	5.3	4.7

[1] Lane 1 is the fastest and lane 5 is the slowest.
[2] Mark A1.D1 indicates the motorway number A1 in direction 1 (toward Borth). D2 means the opposite direction.
[3] The sensor is located at 2 km upstream.
[4] The width of density classes is 1 vehicle/km.
* Calibration set.

distance headway (i.e., to the inverse of density) or to their time headway (i.e., to flow). The objective of this research is to minimize the loss of information due to the use of an analytic form when it is not required. Two methodologies have been tested: a quadratic optimization with linear constraints (QOLC) and a shortest path algorithm (SPA).

Analytic FDs have a comprehensive and a predictive power, which is required in a number of problems: regarding car-following [4], with stochastic models [5], construction of explicit solutions of traffic flow models [6]. However, as only few parameters are calibrated for determining the whole relation between speed and density, some loss of information appears between the empirical data and the analytic FD. This loss can be drastic for certain problems, such as estimating traffic state variables (the flow) from empirical FCD speeds. This issue occurs when FCD or Bluetooth sensors replace the common traffic loops. The relationship between speed and flow that an analytic FD provides at high speeds is too flat and not so accurate to be inverted, which is necessary to obtain the flow from the speed. For this problem, or for an ex-ante assessment of a traffic management strategy, the power of the analytic FD is not needed; it is better to use an FD close to data.

For the calibration of the FD, sampling is an important task with regard to the clustering variable as well as the number and size of classes. Depending on whether we consider the speed-density or the speed-flow relationship of the FD, density or flow classes are built. Using density classes does not enable us to have the exact value of the road capacity. Indeed, in the QOLC approach, the class corresponding to the critical density contains not only periods at capacity, but also periods with simultaneously a lower flow and a lower speed; this makes the average flow of the periods composing this class lower than the capacity. On the other hand, in the SPA, using flow classes results in an easy capacity identification – the capacity corresponds to the highest empirical flow class.

Furthermore, the optimization method can be extended to more constraints, such as the concavity of a flow-density FD, which is required in the first-order LWR traffic flow model and in the GSOM models.

The QOLC or SPA results consist of as many parameters as the number of classes. It is well known that the fewer the parameters, the more significant they are. But the quantity of data (in this case, 58,000 six-minute periods) allows the calibration of a high number of parameters, even if the traffic variables are not fully independent. A calibrated FD can be used on other periods of the same section, or on a symmetric section; elsewhere, other calibrations are necessary.

Within the SPA, congested and non-congested periods might be separated either according to a constant density threshold (the critical density) or to a threshold function of the flow. Such a function rightly reduces the importance of the critical density and impacts the FD speeds related to the flow. In the case of a density threshold decreasing from a high value (at very low flows) to the critical density (at capacity), both FD congested and non-congested speeds are increased. Besides, at low flows, periods whose empirical density is between the critical density and the threshold function are no longer considered to be congested. This is consistent with a definition of congestion based on an extended safety constraint. In every case, it would be sensible to first carefully check the results, to derive the critical density, the free-flow speed and the capacity, and to check these values with the same parameters obtained by other methods.

A correct free-flow speed is obtained with both methods, the QOLC and the SPA. Both methods could be used, achieving the main objective of this research, which is to give a fundamental diagram relationship, free of any analytical form but with respect to the traffic flow phenomena.

ACKNOWLEDGMENTS

NEILA BHOURI, Ph.D, Hdr.[1]
(Auteure correspondant)
Courriel : neila.bhouri@ifsttar.fr
MAURICE ARON, Statisticien-économiste.[1]
Courriel : mauraron@gmail.com
HABIB HAJSALEM, Ph.D, Hdr.[1]
Courriel : habib.haj-salem@ifsttar.fr
[1] Université Paris-Est, IFSTTAR
   Département COSYS, Laboratoire GRETTIA
   14-20, Boulevard Newton, 77447 Champs sur Marne,
   Cedex 2, France

UNE APPROCHE GUIDEE PAR LES DONNEES POUR ESTIMER LE DIAGRAMME FONDAMENTAL

RESUME

Le diagramme fondamental relie la vitesse moyenne, la concentration et le débit du trafic. Nombre de problèmes requièrent une forme analytique pour ce diagramme, parce qu'une telle forme est susceptible d'apporter un potentiel de compréhension et de prédiction. Nous pensons cependant que dans certaines études d'évaluation, une forme analytique n'est pas nécessaire et peut même occasionner une perte de précision. Nous développons ici un diagramme fondamental non analytique, qui transcrit au plus près les données empiriques, en respectant les relations qui existent entre les variables du trafic. Pour éviter un biais, nous séparons les données en et hors congestion. Avec une définition de la congestion fondée sur un seuil proportionnel à celui utilisé dans la contrainte de sécurité, la concentration qui sépare congestion et non-congestion apparaît comme une fonction décroissante du débit – et non comme une simple valeur de concentration critique. Nous identifions cette

fonction, puis l'utilisons. Deux techniques de calibrage - un algorithme de plus court chemin, et une optimisation quadratique avec contraintes linéaires - sont ici présentées, testées, comparées et validées.

MOTS-CLEFS

non-analytique; calibrage; données empiriques; algorithme du plus court chemin; programmation quadratique convexe; contrainte de sécurité; fonction de concentration critique;

REFERENCES

[1] Gerlough DL, Huber MJ. *Traffic Flow Theory: a Monograph.* Transportation Research Board. Special Report 165, 1975.

[2] Transportation Research Circular E-C 149. *75 Years of the Fundamental Diagram for traffic flow theory, Greenshields Symposium.* TRB Traffic Flow, Theory and Characteristics Committee; 2011.

[3] Dhingra SL, Gull I. Traffic Flow Theory Historical Research Perspectives. In: *Transportation Research Circular E-C 149. 75 Years of the Fundamental Diagram for traffic flow theory. Greenshields Symposium.* TRB Traffic flow, Theory and Characteristics Committee; 2011. p. 45-62.

[4] Pipes LA. Car-Following models and the Fundamental Diagram of road traffic. *Transportation Research. Part B: Methodological.* 1967;1(1): 21-29.

[5] Li J, Chen QY, Ni D, Wang H. Analysis of LWR model with Fundamental Diagram subject to uncertainty. In: Transport. Research Circular E-C 149. *75 Years of the Fundamental Diagram for Traffic Flow Theory. Greenshields Symposium.* TRB Traffic flow Theory and Characteristics Committee; 2011. p. 74-83.

[6] Lu Y, Wong SC, Zhang M, Shu CW, Chen W. Explicit construction of entropy solutions for the Lighthill-Whitham-Richards traffic flow model with a piecewise quadratic flow-density relationship. *Transportation Research Part B: Methodological.* 2008;42(4): 355-372.

[7] Aron M, Seidowsky R, Cohen S. Ex-ante assessment of a speed limit reducing operation – A data-driven approach. In: Yannis G, Cohen S. (Eds.) *Traffic Safety.* Wiley; 2016. p. 177-197.

[8] Cohen S, Christoforou Z. Travel time estimation between loop detectors and FCD: A compatibility study on the Lille network, France. *Transportation. Research Procedia.* 2015;10: 245-255. Available from: https://ac.els-cdn.com/S2352146515002616/1-s2.0-S2352146515002616-main.pdf?_tid=cb-da321a-d945-11e7-9781-00000aab0f02&acdnat=1512428125_be15da67407e1b228fff263a8ed88619 [Accessed 5th Dec 2017].

[9] Neumann T, Böhnke PL, Touko Tcheumadjeu LC. Dynamic representation of the fundamental diagram via Bayesian networks for estimating traffic flows from probe vehicle data. In: *16th IEEE ITSC, October 6-9, 2013, The Hague, The Netherlands.* IEEE; 2013. p. 1870-1875.

[10] Underwood RT. *Speed, volume and density relationships: Quality and theory of traffic flow.* New Haven, Connecticut: Yale Bureau of Highway Traffic; 1961. p. 141-188.

[11] Coifman, B. Revisiting the empirical fundamental relationship. *Transportation Research Part B: Methodological.* 2014;68: 173-184.

[12] Lighthill MJ, Whitham GB. On kinematic waves. II. A theory of traffic flow on long crowded roads. In: *Proceedings of the Royal Society of London,UK. Mathematical and Physical Sciences, series A.* 1955;229(1178): 317-345.

[13] Richards P. Shock waves on the highway. *Operations Research.* 1956;4(1): 42-51.

[14] Payne HJ. Models of freeway traffic and control. In: Bekey GA. (ed.) *Simulation Councils Proceedings. Mathematical Models Public Systems.* 1971;1(1): 51-61.

[15] Lebacque JP, Khoshyaran MM. A variational formulation for higher order macroscopic traffic flow models of the GSOM family. *Transportation Research Part B: Methodological.* 2013;57: 245-265.

[16] Greenshields BD. A study in highway capacity. In: *Proceedings of the Highway Research Board. Washington DC.USA.* Vol. 14; 1935. p. 448-477.

[17] Greenberg H. An analysis of traffic flow. *Operations Research.* 1959;7(1): 79-85.

[18] Edie LC. *Car-Following and steady state theory for non congested traffic. Tunnel traffic capacity study.* Port of New York Authority, New York, USA. Report number: 6, 1960.

[19] May AD. *Traffic flow fundamental.* Englewood Cliffs: Prentice Hall; 1990.

[20] Treiterer J, Myers JA. The hysteresis phenomenon in traffic flow. In: *Proceedings of the 6th Int. Symp. on Transportation and Traffic Theory, 26-28 August 1974, Sydney, Australia.* Vol. 6; 1974. p.13-38.

[21] Zhang HM. A mathematical theory of traffic hysteresis. *Transportation Research Part B: Methodological.* 1999;33(1): 1-23.

[22] Castillo J, Benítez F. On the functional form of the speed-density relationship I: General theory. *Transportation Research Part B: Methodological.* 1995;29(5): 373-389.

[23] Li J, Zhang HM. Fundamental Diagram of traffic flow. New identification scheme and further evidence from empirical data. *Transportation Research Records, Journal of the Transportation Research Board.* 2011;2260: 50-59.

[24] Kerner BS. Three-phase theory of city traffic: Moving synchronized flow patterns in under-saturated city traffic at signals. *Physica A: Statistical Mechanics and its Applications.* 2014;397: 76-110.

[25] Daganzo CF, Geroliminis N. An analytical approximation of the macroscopic fundamental diagram. *Transportation Research Part B: Methodological.* 2008;42(9): 771-781.

[26] Ji Y, Xu M, Li J, van Zuylen HJ. Determining the Macroscopic Fundamental Diagram from Mixed and Partial Traffic Data. *Promet – Traffic & Transportation.* 2018;30(3): 267-279. Available from: https://traffic.fpz.hr/index.php/PROMTT/article/view/2406/ [Accessed 29 August 2018].

[27] Lu XY, Varaiya P, Horowitz R, Skabardonis A. Fundamental Diagram modeling and analysis based NGSIM data. In: *12th IFAC Symposium on Control in Transportation Systems, 2009, Redondo Beach, California, USA.*

[28] Coifman B. Jam occupancy and other lingering problems with empirical fundamental relationships. *Transportation Research Records, Journal of the Transportation Research Board.* 2014;2422: 104-112.

[29] Goldfarb D, Idnani A. A numerically stable dual method for solving strictly convex quadratic programs. *Mathematical Programming.* 1983;27: 1-33.

Appendix A. Density related to flow with regard to a safety constraint

Let V be the velocity of a vehicle, Γ its maximum deceleration, T its reaction time, H its distance headway (between its front bumper and the front bumper of the prior vehicle).

Assuming that an object falls from the prior vehicle, the following driver is able to avoid it if his stopping distance $\frac{V^2}{2 \cdot \Gamma} + V \cdot T$ is lower than the distance headway minus the length L of the prior vehicle. The constraint is even heightened or generalized when the stopping distance is lower than the headway, multiplied by a coefficient α lower than 1, and when L includes a safety distance:

$$\frac{V^2}{2 \cdot \Gamma} + V \cdot T + L \leq \alpha \cdot H \tag{22}$$

"Non-congestion" can be defined as periods when, for a certain α, *Equation 22* is valid.

Taking the meter as the space unit (instead of the kilometer) makes $H = \frac{1}{K}$. Then, using the relation $Q = K.V$:

$$\frac{Q^2}{2 \cdot \Gamma \cdot K^2} + \frac{Q \cdot T}{K} + L \leq \frac{\alpha}{K} \Leftrightarrow$$
$$\frac{Q^2}{2 \cdot \Gamma} + Q \cdot T \cdot K + L \cdot K^2 \leq \alpha \cdot K \Leftrightarrow \tag{23}$$
$$\frac{Q^2}{2 \cdot \Gamma} + (Q \cdot T - \alpha) \cdot K + L \cdot K^2 \leq 0$$

According to $\Delta = (Q \cdot T - \alpha)^2 - 2 \cdot L \cdot \frac{Q^2}{\Gamma}$, the quadratic equation associated to *Equation 23* has either no root, when $\Delta<0$, then drivers cannot respect the constraint, or, when $\Delta \geq 0$, has two roots, $K_1(Q)$ and $K_2(Q)$, relating density to flow:

$$K_1(Q) = \frac{(a - Q \cdot T) + \sqrt{\Delta}}{2 \cdot L} \tag{24}$$

$$K_2(Q) = \frac{(a - Q \cdot T) - \sqrt{\Delta}}{2 \cdot L} \tag{25}$$

For these roots, the stopping distance just complies with the constraint. When Q is near zero, $K_1(0) \approx \frac{\alpha}{L}$ and $K_2(0) \approx 0$. When $Q=0$, the root $K_1(Q)$ convenes for the congested branch of the FD. By continuity, $K_1(Q)$ convenes any Q. We assume that when $K_2(Q) \leq K \leq K_1(Q)$, there is no congestion, i.e., the generalized safety constraint is complied with. This assumption says that the function $K_1(Q)$ replaces the critical density. Its variations are studied below.
Let u be such that

$$u^2 \cdot T^2 = \frac{2L}{\Gamma} \tag{26}$$

Then

$$\Delta = (Q \cdot T - \alpha)^2 - u^2 \cdot T^2 \cdot Q^2 \tag{27}$$

The derivative of $K_1(Q)$ related to Q is

$$K_1'(Q) = -\frac{T}{2 \cdot L}\left(1 + \frac{\alpha + T \cdot Q(u^2 - 1)}{\sqrt{\Delta}}\right) \tag{28}$$

Equations 24, 25 and *28* are defined when:

$$\Delta = (Q \cdot T(1+u) - \alpha) \cdot (Q \cdot T(1-u) - \alpha) \geq 0 \tag{29}$$

The roots of the equation $\Delta = 0$ are

$$Q_1 = \frac{\alpha}{(T \cdot (1+u))} \text{ and } Q_2 \frac{\alpha}{(T \cdot (1-u))} \tag{30}$$

When considering $\alpha = 0.6$, there are three possibilities, according to u.

a) case u>1
This is the common case: when $T=1$ second, $\Gamma=6$m/s^2, $L=4$m, then $u=1.15$. In that case, $Q_2<0$ and $Q_1>0$. *Equation 29* – and, in turn, the safety constraint – are satisfied when

$$0 \leq Q \leq Q_1 \tag{31}$$

The safety constraint is not satisfied when traffic flows are higher.

When $Q<Q_1$, *Equations 24, 25* and *28* are defined and the derivative *28* is negative. This means that $K_1(Q)$ is decreasing from $\frac{a}{L}$ (in this case, 150 vehicles/km) at $Q=0$ until a positive value of $\frac{\alpha \cdot u}{2 \cdot L(1+u)}$, reached when $Q=Q_1$ (thus, with the given parameters, until 0.04 vehicle/m or 40 vehicles/km for $Q_1=0.279$ vehicle/second, i.e., 1,004 vehicles/hour).

b) case u=1
$K'(Q)$, given by *Equation 28* is still negative, thus $K_1(Q)$ is still decreasing from $\frac{\alpha}{L}$ (in this case, 150 vehicles/km) at $Q=0$ until the value $\frac{\alpha}{4.L}$ (in this case, 37.5 vehicles/km) at $Q = Q1 = \frac{\alpha}{2.T}$ (this corresponds to 0.3 vehicle/second, i.e., 1,080 vehicles/hour). The safety constraint is not satisfied when traffic flows are higher.

c) case u < 1
Equation 29 is satisfied when

$$Q \leq Q_1 \text{ or when } Q \geq Q_2 \tag{32}$$

From *Equation 28*, $K_1'(Q)$, is negative or null when:

$$\frac{\alpha + Q \cdot T \cdot (u^2 - 1)}{\sqrt{(Q \cdot T(1+u) - \alpha) \cdot (Q \cdot T \cdot (1-u) - \alpha)}} \geq -1 \tag{33}$$

The root of the numerator of the first term of *Equation 33* is:

$$Q_n = \frac{\alpha}{T \cdot (1 - u^2)} \tag{34}$$

When $Q > Q_N$, the numerator is negative; in this case, *33* is satisfied only when this numerator is, in absolute value, lower than the denominator, which is equivalent to:

$$(\alpha + Q \cdot T \cdot (u^2 - 1))^2 \leq (Q \cdot T \cdot (1+u) - \alpha) \cdot (Q \cdot T \cdot (1-u) - \alpha) \tag{35}$$

Equality 35 is reordered in

$$Q \cdot T^2 \cdot u^2 (1 - u^2) \left(Q - \frac{2 \cdot \alpha}{T \cdot (1 - u^2)} \right) \geq 0 \qquad (36)$$

The roots of the quadratic equation associated with *Equation 36* are $Q_0 = 0$ and $Q_4 = \frac{2 \cdot \alpha}{T \cdot (1 - u^2)}$. As

$$\frac{Q_4}{Q_2} = \frac{2}{1+u} \succ 1; \frac{Q_2}{Q_N} = 1 + u \succ 1; \text{ and}$$

$$\frac{Q_N}{Q_1} = \frac{1}{1-u} \succ 1 \Rightarrow 0 = Q_0 < Q_1 < Q_N < Q_2 < Q_4 \qquad (37)$$

When $Q \geq Q_4$, *Equations 29* and *32–36* are satisfied, as well as the safety constraint *23*. Then, from *Equation 28*, $K'_1(Q)$ is negative and $K_1(Q)$ decreases with Q.

When $Q_2 \leq Q \leq Q_4$, *Equations 29* and *32* and the safety constraint *23* are satisfied. *Equation 34* remains satisfied, *Equations 33*, *35* and *36* are reversed, $K'_1(Q)$ is positive and $K_1(Q)$ increases with Q.

When $Q_1 \leq Q \leq Q_2$, *32* is not verified; the safety constraint *23* is not complied with.

When $Q \leq Q_1$, *32*, *29* and the safety constraint *23* are satisfied; as *Equation 34* is reversed, *Equation 33* remains satisfied, $K'_1(Q)$ is negative; $K_1(Q)$ decreases with Q.

Appendix B. A decreasing density threshold, in the Underwood fundamental diagram

Whatever the FD, the value of the function $\frac{K^{free}(q) + K^{congested}(q)}{2}$ is greater at q=0 than at q=qmax. Indeed, for very low flows, the density on the congested branch of the FD is much higher than twice the critical density. This does not mean that this function always decreases between q=0 and q=qmax. But it is the case at least for a well-known fundamental diagram, the Underwood speed-density relationship [10], which is $V = V_f \cdot e^{-K/K_{crit}}$, where V_f is the free speed (valid at q=0) and K_{crit} the critical density. The resulting flow is

$$Q = K \cdot V = K \cdot V_f \cdot e^{-K/K_{crit}} \qquad (38)$$

Inverting this relationship, two density-flow relationships $K^{free}(q)$ and $K^{congested}(q)$ are obtained; although no analytic form is available, there is a (tedious) mathematical proof that the function $\frac{K^{free}(q) + K^{congested}(q)}{2}$ is decreasing. A graphical approach, with a sufficiently small resolution step, is good enough as well.

This relationship has only two parameters; it reduces, after two changes of scale (in distance and in time), to the unique negative exponential curve $V = \exp(-K)$:

– The unit of length, instead of 1km is set to $\frac{1}{K_{crit}}$ km; in this new unit the critical density is equal to 1. In this unit, speeds must be multiplied by K_{crit}. In particular, the free speed V_f is replaced by $W_f = V_f \cdot K_{crit}$

– The unit of time, instead of 1 hour, is set to $\frac{1}{W_f}$ hour; in this new unit the speeds are divided by W_f; in particular the free speed becomes $\frac{W_f}{W_f} = 1$
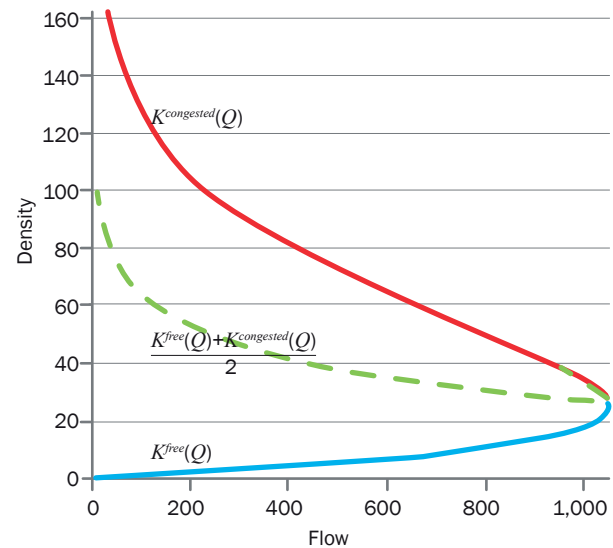

Figure 5 – Graphical decrease of the threshold in the Underwood exponential model