



19th International Symposium on Transportation and Traffic Theory

Capacity Drops at Merges: an endogenous model

Ludovic Leclercq^{a,*}, Jorge A. Laval^b, Nicolas Chiabaut^a

^aLICIT, IFSTTAR / ENTPE, Université de Lyon, Rue Maurice Audin, 69518 Vaulx-en-Velin Cedex, France

^bSchool of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA, USA

Abstract

The Newell-Daganzo merge model is not only very simple but also accurately reproduces experimental findings. However, the capacity downstream of the merge is an exogenous variable in the model. This is a serious limitation for merges that behave as active bottlenecks because their downstream capacity is a direct consequence of the merging behavior. This paper proposes an analytical model that extends the Newell-Daganzo model by incorporating, endogenously, the capacity drop related to the merging process. Two cases are investigated depending on the traffic states on the on-ramp. The model properties are analyzed and a sensitivity analysis is performed to quantify the relative contribution of the each parameter in the capacity drop. Finally, the extended Newell-Daganzo model is validated with experimental data coming from an active merge bottleneck on the M6 freeway in UK.

© 2011 Published by Elsevier Ltd.

Keywords: Capacity drop, merge, merge ratio, kinematic wave model, active bottleneck

1. Introduction

Traffic flow behavior at freeway merges has been extensively investigated in the literature. One of the simplest models for merges was originally proposed in (Newell, 1982) and later formalized in (Daganzo, 1995). This model, referred to as ND model in this paper, postulates that flows on both incoming roads share the available downstream supply following a specific allocation scheme, see Fig. 1. Notably, when the main road is congested (high demand), the on-ramp may be either (i) in free-flow (low demand) or (ii) congested (high demand). In case (i), vehicles on the on-ramp force their way and all the demand succeeds in merging. In case (ii) the merging behavior is mostly deterministic and dictated by a fixed merge ratio. This simple model has been verified repeatedly to accord surprisingly well with experimental findings (Troutbeck, 2002; Cassidy and Ahn, 2005; Bar-Gera and Ahn, 2010).

The main limitation of the ND model is that the downstream capacity is exogenous. This prevents any meaningful modeling of the capacity drop phenomenon or the effects of the merging section length. This is particularly troublesome when the merge behaves as an active bottleneck. Indeed, capacity drops between 10 to 30 % are commonly observed in that case (Elefteriadou et al., 1995; Persaud et al., 1998; Kerner, 2002; Cassidy and

* Corresponding author. Tel.: +33-472047716; fax: +33-472047712.

E-mail address: leclercq@entpe.fr.

Bertini, 1999; Sarvi et al., 2007; Chung et al., 2007; Yi and Mulinazzi, 2007). Lane-changing, low-speed inserting vehicles and heterogeneous lane behavior have been suspected to induce the capacity drop (Cassidy and Rudjanakanoknad, 2005; Cassidy and Ahn, 2005; Laval et al., 2005, Treiber et al., 2006, Laval and Daganzo, 2006). Indeed, these phenomena generate variations between over and mid-saturated traffic states at merges (Mauch and Cassidy, 2002; Ahn and Cassidy, 2007; Laval et al., 2009; Li et al., 2010) that prevent to reach the full freeway capacity.

Several models have been proposed to account for capacity drops, but with the exception of (Laval and Daganzo, 2006), it is treated exogenously. For example, (Koshi et al. 1983; Hall and Hall, 1990) propose models that are mainly descriptive and based on fundamental diagrams with reverse lambda shape; (Evans et al., 2001; Kerner, 2000 ; 2004) postulate stochastic approaches; and (Siegel et al., 2009 ; 2009b) assume second order models derived from (Aw and Rascle, 2000) and (Greenberg, 2001) works. However, none of these models propose an explicit formulation of the relationship between local traffic interactions related to lane-changing and their global impact on the capacity. Furthermore, most of these models are hard to implement and to calibrate in practice.

The aim of this paper is to overcome these issues by introducing, in the ND model, an endogenous analytical description of the merge bottleneck activation and the ensuing capacity drop. Basic idea is to incorporate in this model the physical mechanism unveiled in (Laval and Daganzo, 2006) pertaining to the effects of the bounded acceleration of merging vehicles. Toward this end, section 2 of the paper presents the analytical model that estimates the capacity drop with respect to the demand on the on-ramp and the different model parameters. For mathematical tractability, this model is not fully realistic because both incoming roads are assumed single-lane. However, it will significantly improve the global understanding of how the merging mechanism contributes to the capacity drop. A sensitivity analysis is performed in section 3 to precisely quantify the contribution of the different parameters. This will provide some interesting insights on how road design can improve merge capacity. Finally, the extended ND model will be validated with experimental data coming from a merge on the M6 freeway in UK in section 4. To this end, an estimation method will be presented to partly overcome the one-lane limitation of the model. Finally, section 5 presents a discussion.

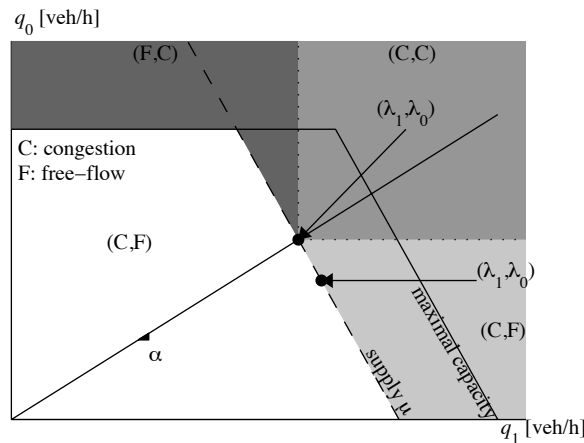


Fig. 1: Merge diagram for the Newell-Daganzo (ND) model

2. An analytical model for capacity drop

In this section, a simple analytical model that accounts for the bounded acceleration of inserting maneuvers is presented. Two cases will be studied separately depending on the state of the on-ramp, which can either be congested (right top quadrant of ND model) or not (right bottom quadrant), see Fig. 1. The demand on main road is always high enough to encounter congestion.

2.1. Congested on-ramp

Consider a merge with two one-lane incoming roads as in Fig. 2a. The main road is labeled 1 and the on-ramp is labeled 0. The length of the insertion section is L . Traffic on each road is described by the kinematic wave model (Lighthill and Whitham 1955; Richards 1956) and a triangular fundamental diagram with free-flow speed u , wave speed w , and jam density κ . The capacity on one-lane is equal to $Q = wu\kappa/(w+u)$. The upstream demands are denoted λ_i ($i=0,1$) and the mean flows just before the beginning of the insertion section ($x=0$) are noted q_i . The merge always behaves as an active bottleneck, i.e. the downstream flow q_0+q_1 is never constrained by the downstream supply μ .

Vehicles from the on-ramp are assumed to insert themselves in the main road with speed v_0 and then accelerate at a constant rate a until they reach the free-flow speed, see Fig. 2b. Both v_0 and a are assumed identical for all vehicles. This seems reasonable as the on-ramp is congested and thus the traffic states are quite homogeneous. Vehicle i is inserting at time t_i . The time headway between two consecutive insertions is denoted $h_i = t_{i+1} - t_i$. Such headways follow an unknown distribution $H(h_0, s)$ with mean h_0 and standard deviation s . When s is equal to 0, vehicles regularly insert every h_0 , as illustrated in Fig. 2b. After their insertion, vehicles from the on-ramp behave as moving bottlenecks (Newell, 1998; Lebacque et al., 1998; Munoz and Daganzo, 2002; Leclercq et al., 2004). Thus, these moving obstructions constrain the flow on the main road and create voids in front of them. It was shown in (Laval and Daganzo, 2006) that these moving obstructions are responsible for the capacity drop. Two cases will be distinguished in the present work: $L=0$ and $L>0$.

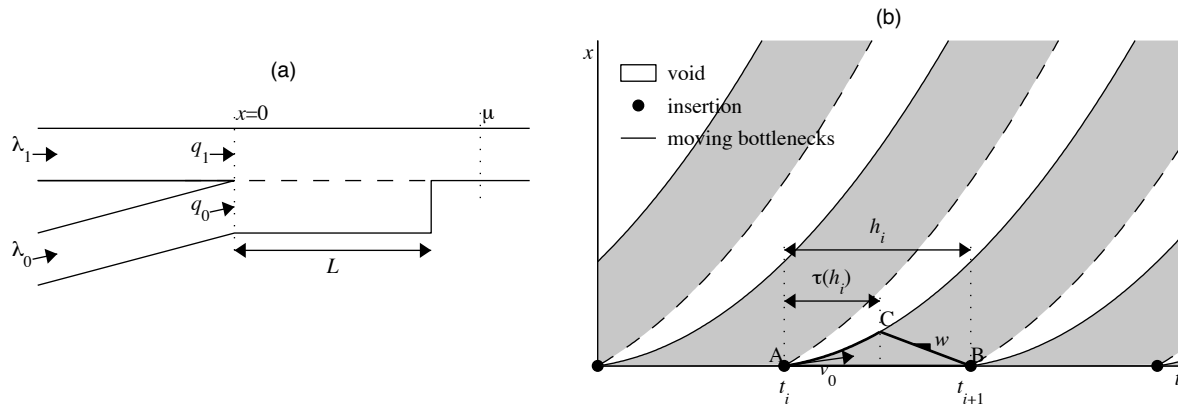


Fig. 2: (a) sketch of the merge (b) Inserting process when $L=0$

When $L=0$, all insertions take place at $x=0$, see Fig. 2a. Let N_i be the cumulative number of vehicles that have crossed this location by time t_i . At a large time scale, the mean total flow downstream of the merge, i.e. the effective capacity C , is the same whatever the location is. Thus, C can be simply expressed at $x=0$ with respect to N_i :

$$C = \sum_{i=1}^n (N_{i+1} - N_i) / \sum_{i=1}^n h_i \quad \text{with } n \rightarrow +\infty \tag{1}$$

Variational theory (Daganzo, 2005) provides a convenient way to calculate the increase in N between t_i and t_{i+1} , i.e. along the horizontal path from A to B, see Fig. 2b. Indeed, this increase is the same as on the alternative path $A \rightarrow C \rightarrow B$ that first follows the moving bottleneck trajectory and then reaches B along a characteristic with slope w . From A to C, the passing rate is equal to 0 because no vehicle can overpass the moving obstruction. From C to B, the passing rate is equal to $w\kappa$ because the path slope is w . Thus, the increase in N between t_i and t_{i+1} is equal to $w\kappa(h_i - \tau(h_i))$, where $\tau(h_i)$ represents the time duration between points A and C, see Fig. 2b and (Leclercq, 2005) for more details. It can be shown that the effective capacity C can be obtained with:

$$C = w\kappa \left(1 - \sum_{i=1}^n \tau(h_i) / \sum_{i=1}^n h_i \right) \quad \text{and} \quad \tau(h_i) = -\frac{w + v_0}{a} + \frac{1}{a} \sqrt{(w + v_0)^2 + 2awh_i} \tag{2}$$

The law of large numbers tells us that:

$$(1/n) \sum_{i=1}^n h_i \rightarrow E(h_i) \text{ and } (1/n) \sum_{i=1}^n \tau(h_i) \rightarrow E(\tau(h_i)) \tag{3}$$

where $E(x)$ is the mathematical expectation of x . $E(h_i)$ is equal to h_0 . As the distribution H of h_i is unknown, $E(\tau(h_i))$ cannot be analytically derived. However, one can estimate its second-order approximation (Oehlert, 1992):

$$E(\tau(h_i)) \approx \tau(E(h_i)) + \frac{1}{2} s^2 \frac{\partial^2 \tau}{\partial h_i^2}(E(h_i)) = \tau(h_0) - \frac{as^2w^2}{2((w+v_0)^2 + 2awh_0)^{3/2}} \tag{4}$$

Thus, the effective capacity C is equal to:

$$C \approx w\kappa + \frac{w\kappa}{h_0} \left(\frac{w+v_0}{a} - \frac{1}{a} \sqrt{(w+v_0)^2 + 2awh_0} + \frac{as^2w^2}{2((w+v_0)^2 + 2awh_0)^{3/2}} \right) \tag{5}$$

This effective capacity is shared by the two incoming flows q_0 and q_1 . As both upstream roads are congested, the merge ratio α holds and $q_0 = \alpha q_1$ and the effective capacity can also be expressed as:

$$C = q_0 + q_1 = (1 + 1/\alpha)q_0 \tag{6}$$

Note that q_0 and v_0 are related by the fundamental diagram, i.e. $v_0 = wq_0 / (w\kappa - q_0)$, and that $h_0 = 1/q_0$. Thus, by combining (5) and (6), we obtain an equation in q_0 with respect to the following parameters: a, w, κ, α and s :

$$w\kappa + q_0 \left(\frac{w^3\kappa^2}{a(w\kappa - q_0)} - \frac{w\kappa}{a} \sqrt{\frac{w^4\kappa^2}{(w\kappa - q_0)^2} + \frac{2aw}{q_0}} + \frac{as^2w^3\kappa}{2(w^4\kappa^2/(w\kappa - q_0)^2 + 2aw/q_0)^{3/2}} \right) - \left(1 + \frac{1}{\alpha} \right) q_0 = 0 \tag{7}$$

Unfortunately it is impossible to derive the explicit formulation of q_0 with respect to these parameters. However, we can easily compute the q_0 values for any given sets of parameters. We propose the indicator c to quantify the relative capacity drop, i.e. the complement of the ratio between the effective capacity C and the capacity Q given by the fundamental diagram:

$$c = 1 - \frac{C}{Q} = 1 - \left(1 + \frac{1}{\alpha} \right) \frac{q_0}{Q} \tag{8}$$

Note that capacity drop here cannot directly be compared with experimental values found in the literature. Indeed, we have chosen a fixed reference, Q , to calculate the capacity drop. In reality, the capacity drop is often defined in reference to the maximal flow observed just before the capacity drop, which is always lower than Q . Further in the paper we will evaluate how a, w, κ, α and s influence c by performing a sensitivity analysis.

Now we consider the case of a merge where the insertion section has a spatial extension ($L > 0$). In this case, vehicle i 's insertion takes place at a random location $x_i \in [0, L]$. The distribution X of x_i is assumed to be uniform, as suggested by the empirical evidence in (Daamen et al., 2010). Indeed, these authors have computed merge location histograms at a merge in Netherlands. This distribution appears to be piecewise uniform in congestion, see figure 6 in that reference. We have performed extensive tests that demonstrate that a piecewise uniform distribution leads to the same results in terms of capacity drop as a uniform distribution.

To simplify the exposition, vehicles are first supposed to enter the mainline every h_0 , i.e. $s=0$, see Fig. 3a. A supplementary assumption is made here: we neglect the effect that the void just downstream of a moving bottleneck can have on its own trajectory by making it react to downstream congestion waves carrying lower speeds. This implies that a moving bottleneck disappears when it crosses such waves; see square dots in Fig. 3a. We have confirmed that this assumption does not introduce significant error when L is not too high (roughly below 300 m), and therefore appears to be a valid assumption in our context. Note that this moving bottlenecks behavior is close to the relaxation process described in (Laval and Leclercq, 2008) for lane-changers.

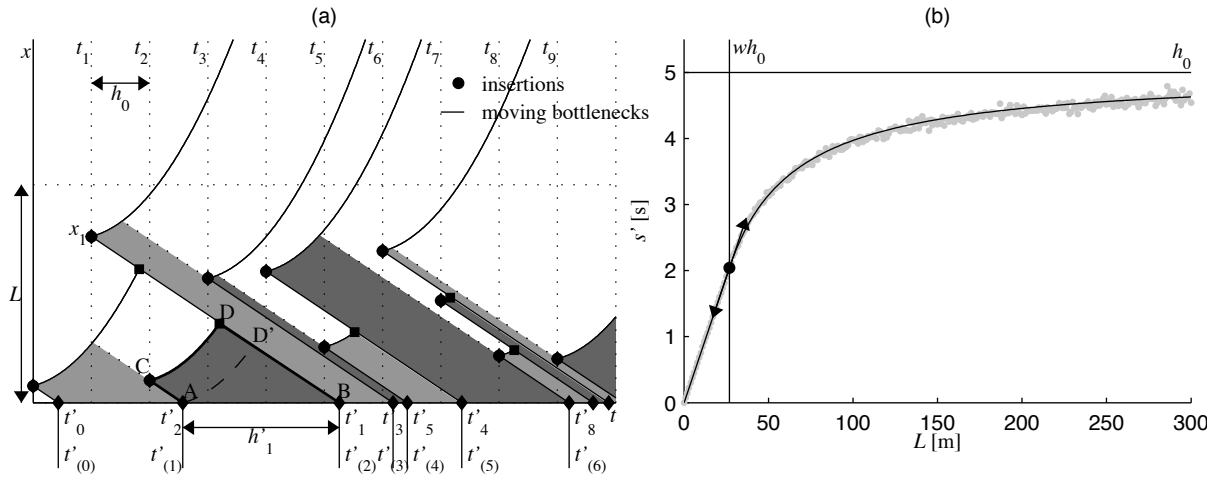


Fig. 3: (a) Inserting process when $L > 0$ (b) Study of the standard deviation of H

Let t'_i be the time when the wave coming from (t_i, x_i) reaches $x=0$ and $t'_{(i)}$ the ordered series built from the realizations of t'_i , see Fig. 3a. The effective capacity C can still be calculated at $x=0$ with (1) by focusing on the increase in N between $t'_{(i)}$ and $t'_{(i+1)}$, i.e. between points A and B in Fig. 3a. This increase can be calculated along the alternative path $A \rightarrow C \rightarrow D \rightarrow B$ in figure. Note here a crucial property: if the moving boundary condition from C to D is translated by the vector CA , we can see another alternative path $A \rightarrow D' \rightarrow B$ where the increase in N is equivalent, see Fig. 3a. Thus, the problem appears to be equivalent to the previous one with $L=0$ by substituting H by H' , the latter being defined as the distribution of $h'_i = t'_{(i+1)} - t'_{(i)}$. The distributions H and H' have the same mean h_0 . We only need to determine s' the standard deviation of H' and the effective capacity C will be obtained by substituting s by s' in (5).

The computation of s' is a quite challenging task because there is no obvious relation in most cases between t_i and $t'_{(i)}$ due to the ordering process. A particular case can first be studied. When $L < wh_0$, the order of t'_i is not modified as every wave emitted at time t_i reaches $x=0$ before t_{i+1} . Thus, $h'_i = t'_{(i+1)} - t'_{(i)} = t_{i+1} - t_i = h_0 + (x_{i+1} - x_i)/w$ in this case. Inserting positions of two consecutive vehicles may reasonably be considered as independent. The standard deviation s' of H' is then given by:

$$s'^2 = V(h'_i) = (1/w^2)V(x_{i+1} - x_i) = L^2/6w^2 \tag{9}$$

where $V(x)$ is the variance of x . When L is very large (formally tends to infinity), it can be demonstrated that H' follows an exponential distribution. Thus, its standard deviation becomes equal to its mean h_0 . When L is higher than wh_0 but not too large. We have performed extensive numerical analysis of s' with respect to L for different values of h_0 and w . It turns out that in all cases the best fit is obtained with the following formula:

$$s'(L) = (bL + c)/(L + d) \tag{10}$$

Fig. 3b presents the fitting result for $h_0=5$ s and $w=5.38$ m/s. The r -square is higher than 0.99. Similar fits are observed for a wide range of values for h_0 and w . Interestingly, it is possible to determine the analytical formulations of b , c and d . First, as s' tends to h_0 when L tends to infinity, b should be equal to h_0 . Second, we can reasonably impose that s' and its first derivative be continuous at $L=wh_0$, see Fig. 3b. Thus, the final expression of s' is:

$$s'(L) = \begin{cases} L/\sqrt{6}w & \text{if } L < wh_0 \\ h_0(L - wh_0/\sqrt{6}) / (L + (\sqrt{6} - 2)wh_0) & \text{if } L \geq wh_0 \end{cases} \tag{11}$$

We can now relax the assumption $s=0$ when $L > 0$. Let H'' be the distribution of the time gaps h''_i between the start of two consecutive moving obstructions at $x=0$. H'' has the same mean h_0 as H . Its standard deviation s'' is driven by two processes, one related to the distribution of the inserting positions and the other related to the distribution of the inserting times. If we suppose that these two processes are independent, s'' can be estimated by:

$$s'' = \sqrt{s^2 + s'^2} \tag{12}$$

To summarize, it appears that when both incoming roads are congested, (7) and (8) can be applied to estimate the capacity drop in the below mentioned four cases. One only has to substitute s in (7) by the right expressions:

1. $s \leftarrow 0$, point merge with constant time gaps between two successive insertions ($L=0, s=0$);
2. $s \leftarrow s$, point merge with distributed time gaps between two successive insertions ($L=0, s>0$);
3. $s \leftarrow s'$ given by (11), extended merge with constant time gaps between two successive insertions ($L>0, s=0$);
4. $s \leftarrow s''$ given by (12), extended merge with distributed time gaps between two successive insertions ($L>0, s>0$).

2.2. Uncongested on-ramp

In this section, we focus on the right bottom part of the ND merge diagram in Fig 1. Here, λ_0 is low enough to prevent congestion on the on-ramp and therefore $q_0 = \lambda_0$. Thus, speeds on the on-ramp are no longer constrained. We assume then that inserting vehicles will adopt the mean main road speed v_1 before accelerating. The inserting process remains the same as in the previous section. The effective capacity C is given by substituting v_0 by v_1 in (5). Note that h_0 is now equal to $1/\lambda_0$ and becomes an additional parameter. Note too that v_1 and q_1 are related by the fundamental diagram, i.e. $v_1 = wq_1 / (w\kappa - q_1)$. Since C is also equal to $q_1 + \lambda_0$, we obtain the following equation in q_1 :

$$\frac{1}{w\kappa} \left(1 + \frac{q_1}{\lambda_0} \right) = \frac{1}{\lambda_0} + \frac{w^2\kappa}{a(w\kappa - q_1)} - \frac{1}{a} \sqrt{\frac{w^4\kappa^2}{(w\kappa - q_1)^2} + \frac{2aw}{\lambda_0}} + \frac{aw^2s^2}{\left(w^4\kappa^2 / (w\kappa - q_1)^2 + 2aw / \lambda_0 \right)^{3/2}} \tag{13}$$

Interestingly, the above expression can be greatly simplified when $s=0$ by defining $q'_1 = w\kappa - q_1$ and $\beta = 2w^3\kappa^2/a$:

$$\lambda_0 = q_1^{(3/2)} \left(\sqrt{\beta} - \sqrt{q'_1} \right) / (\beta - q'_1) \tag{14}$$

This expression shows that when β increases, q'_1 increases and q_1 decreases for a given value of λ_0 . We fail to determine a convincing physical interpretation of β but interestingly, β increases when a decreases. This corroborates an intuitive result: the capacity drop is higher when the acceleration rate is lower. To complete this study, we still have to derive the expression of s' when $L>0$ because (11) is no longer applicable. Indeed, figure 6 in (Daamen et al., 2010) shows that the merge location histogram in free-flow fits a normal distribution rather than a uniform one. Furthermore, this figure suggests that the mean inserting position L_m is close to $L/4$. Thus, we assume that X follows a normal distribution with mean L_m and standard deviation $L_m/2.57$. The standard deviation has been adjusted so that 99.5% of the insertions take place between 0 and L . The expression of s' can then be obtained using a similar method as in previous sections:

$$s'(L) = \begin{cases} L/2.57\sqrt{8}w & \text{if } L < 2.7wh_0 \\ h_0 \left(L - 2.7^2wh_0 / 2.57\sqrt{8} \right) / \left(L + (2.57\sqrt{8} - 2.7)wh_0 \right) & \text{if } L \geq 2.7wh_0 \end{cases} \tag{15}$$

Note that the threshold from when the ordering process of t'_i influences the standard deviation of h'_i is now higher than wh_0 . This is because most insertions are generated at the beginning of the insertion section, i.e. roughly between 0 and $L/2$. By combining (15)-(13) or (12)-(15)-(13) it is possible to determine the value of q_1 with respect to λ_0 in all four case studies. The relative capacity drop c is then given by:

$$c = 1 - (q_1 + \lambda_0) / Q \tag{16}$$

3. Analysis of the model properties

3.1. Results for a reference scenario

Both analytical models introduced previously are applied for the following reference scenario: $w=19.4$ km/h, $u=115$ km/h, $\kappa=145$ veh/km, $a=2$ m/s², $\alpha=0.76$. The theoretical capacity Q given by the fundamental diagram is then equal to 2400 veh/h. The length L is set to 0 m (point merge) or to 160 m (extended merge). The standard deviation s is either equal to 0 s (constant time gaps) or 2 s (distributed time gaps). Fig. 4a presents the evolution of c with respect to λ_0 . For λ_0 values lower than 100 veh/h, no capacity drop occurs. The relative capacity drop then gradually increases until it reaches its maximum value when the on-ramp is fully congested. The threshold value for λ_0 is about 600 veh/h. When λ_0 exceeds this threshold, c remains constant. Note that the maximal value for c is very high (up to 46%). This is because the relative capacity drop is defined with respect to the theoretical reachable capacity Q and because there is only one lane downstream.

For lower values of λ_0 , neither the spatial extension of the merge nor the standard deviation s have a noticeable influence on the capacity drop. When λ_0 increases, the four curves begin to diverge. At the end when the on-ramp is fully congested, it appears that the relative capacity drop is reduced for 46% to 40% when L changes from 0 to 160 m and $s=0$. This corresponds to an increase of the effective capacity from 1310 to 1450 veh/h. Thus, the spatial extension of the merge significantly improves the insertion process.

Fig. 4b presents the same results but plotted in the merge diagram. Note that when the on-ramp is uncongested, the inserting flow q_0 is always equal to λ_0 . The diagonal line in this figure corresponds to the theoretical capacity Q derived from the fundamental diagram. It helps to better appreciate the absolute value of the capacity drop with respect to the flow on the on-ramp.

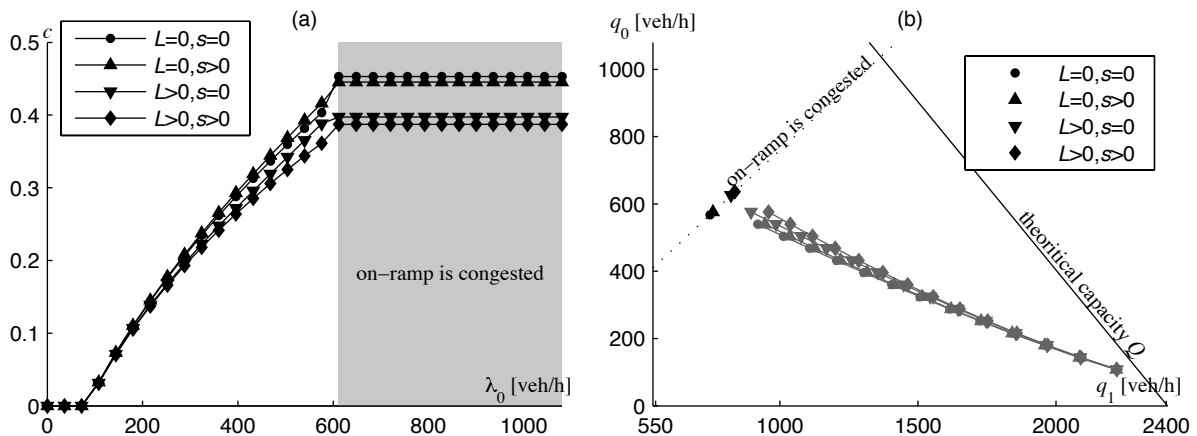


Fig. 4: (a) Relative capacity drop with respect to on-ramp demand (b) Traffic states in ND framework

3.2. Sensitivity analysis

The sensitivity of the relative capacity drop c to the different parameters a , w , κ , α , L and s can be evaluated with (7) and (13). We first focus on the acceleration rate a for the reference scenario when $L>0$ and $s=0$. Fig. 5a presents the evolution of c with respect to λ_0 for increasing values of a from 0.5 to 3 m/s². This parameter appears to have a high influence. For example, when the on-ramp is congested, c is reduced from 0.49 to 0.34 when the on-ramp is congested and a varies from 1 to 3 m/s². This corresponds to an increase of the effective capacity from 1220 to 1580 veh/h (+23%). When the on-ramp is uncongested and $\lambda_0=300$ veh/h, c decreases from 0.32 to 0.16 for the same variation in a . The effective capacity thus increases from 1630 to 2020 veh/h (+22%). It is clear that merges should preferentially be implemented where vehicles could optimally accelerate, i.e. no uphill, adequate sight distance, ... Fig. 5a also highlights that both capacity drop beginnings and full congestion appearance on the on-ramp are

observed for higher λ_0 values when a increases. Fig. 5b shows the evolution of the effective capacity C in the merge diagram for a between 0.5 and 3 m/s^2 . This confirms that the capacity drop is highly influenced by the acceleration. Indeed, the lower the acceleration is the farther the effective capacity is from the theoretical capacity line.

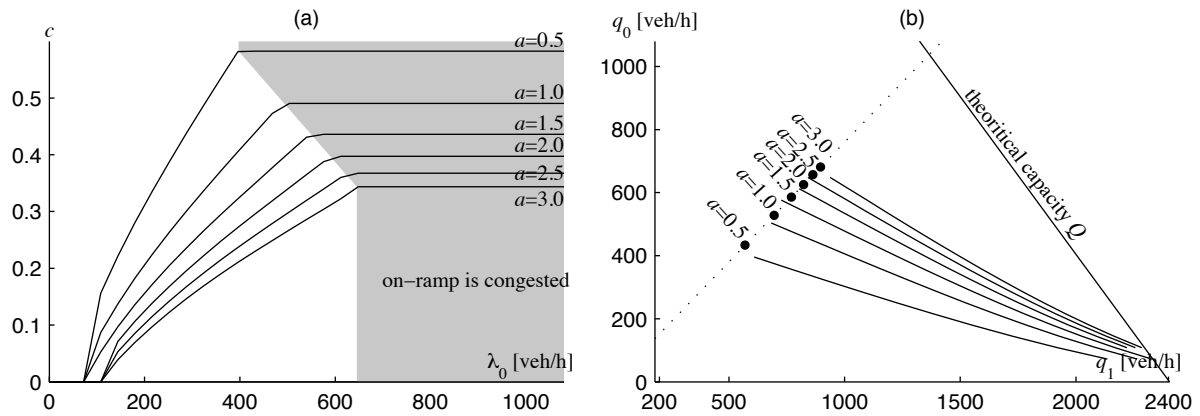


Fig. 5: Influence of the acceleration on the capacity drop (a) merge diagram representation (b) relative capacity drop with respect to λ_0

The sensitivity analysis is now extended to all parameters when the on-ramp is congested. Free-flow situations are disregarded because they mostly lead to similar conclusions. Each parameter is first separately tested for a typical range of values, see Fig. 6. The four identified cases for the insertion process are considered and the reference value for each parameter is pointed out by a dotted line in each subplot. Fig. 6a first reinforces that the acceleration rate is the most influent parameter. The wave speed and the jam density have also a significant impact, see Fig. 6b-c. However, note that a wide range of values have been tested for these parameters. In reality, w and κ are known to vary little for a given freeway. Furthermore, these parameters also modify the capacity Q of the main road. Thus, when w increases (respectively κ), Q also increases and c then increases even if the effective capacity C is constant.

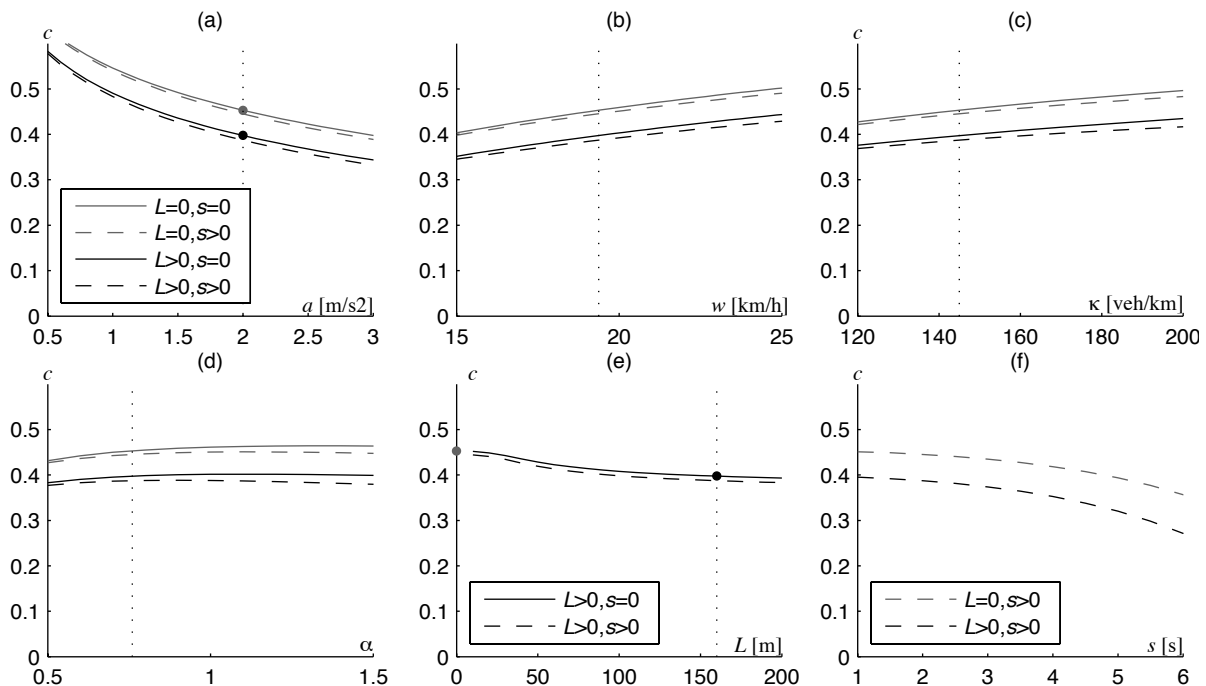


Fig. 6: Sensitivity analysis when on-ramp is congested; influence of (a) the acceleration (b) the wave speed (c) the jam density (d) the merge priority ratio (e) the length of the insertion section (f) the standard deviation of H .

The merge priority ratio seems to have almost no influence on capacity drop, see Fig. 6d. This is quite surprising because when α increases the inserting flow q_0 increases and so do the number of moving bottlenecks. This would have to result in a drop of the effective capacity. However, the analytical model sheds some light on the global mechanism. Indeed, increasing q_0 increases the number of inserting vehicle but also their speed v_0 because traffic conditions improve on the on-ramp. Both phenomena compensate each other and the capacity drop is almost steady regardless of the merge ratio. Thus, all traffic states fall on the same capacity line in the merge diagram.

Fig. 6e focuses on the influence of the length of the insertion section. The figure shows that increasing L reduces the relative capacity drop and then improves the effective capacity. It is noticeable that most of the benefit is obtained before L reaches 100 m. If this result could be generalized with more than one-lane on the main road, it would be an interesting insight for road design. Subplots a to d also highlight the benefits of an extended merge by comparing plots when $L=0$ and $L=160$ m ($L>0$). Note in Fig. 6e that when $s=0$ and L tends to 0, the relative capacity drop converges to the value obtained when $L=0$ (gray dot). Thus, the transition between $L=0$ and $L>0$ is smooth.

Fig. 6f shows the influence of the standard deviation of H when $L=0$ and $L>0$. It appears that increasing s reduces c . Indeed, randomness in inserting times for on-ramp vehicles sometimes generates larger time headways where the flow on main road is able to reach higher values. Thus, the effective capacity increases. This also explains why increasing L reduces the capacity drop. Indeed, we have previously demonstrated that a random distribution of inserting positions is equivalent to introducing randomness in H .

The last part of the sensitivity analysis consists in studying the joint influence of the different parameters when the on-ramp is congested. We disregard α because it has no influence on the capacity drop and fixe s to 0. Fig. 7a shows the joint influence of a and L on the relative capacity drop c . Dots corresponds to the values provided by the analytical model when $L>0$ and $s=0$. It appears that a polynomial regression with degrees not higher than 3 perfectly fits dots (r -square equal to 0.9994). Furthermore, the joint terms of the regression are negligible because r is not significantly reduced when they are omitted (r turns to 0.9993). We finally obtain the following expression:

$$c = 0.774 - 0.331a + 0.122a^2 - 1.85 \cdot 10^{-2} a^3 - 6.76 \cdot 10^{-4} L + 1.73 \cdot 10^{-6} L^2 \tag{17}$$

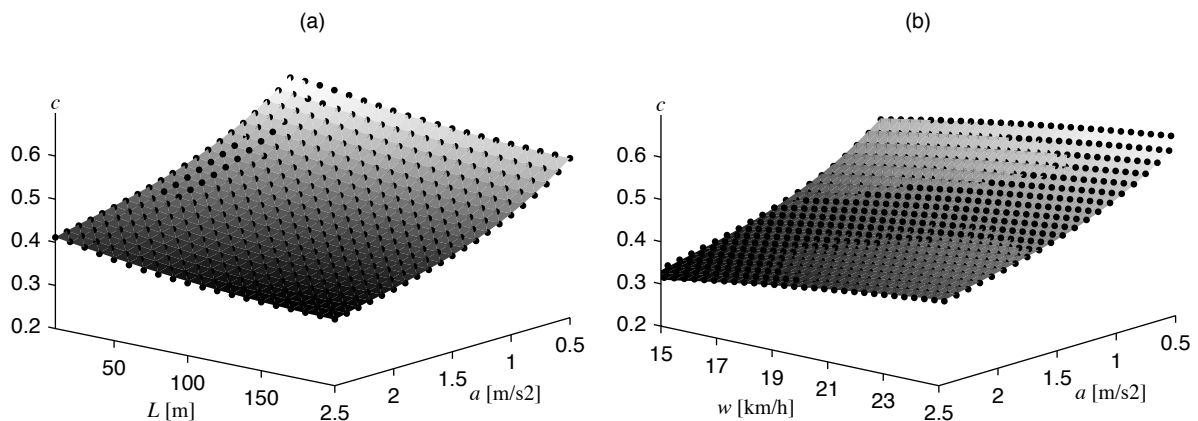


Fig. 7: Multiple regression (a) c with respect to a and L (b) c with respect to a and w .

Thus, we can conclude that a and L have independent contributions to the relative capacity drop. Similar studies are performed for a and w , see Fig. 7b, and for a and κ . Polynomial regressions with no joint terms provide accurate fits with r equal respectively to 0.9993 and 0.9994. Thus, all parameters look independent. It is then possible to combine all the obtained regressions into one:

$$c = 0.402 - 0.332a + 0.122a^2 - 1.85 \cdot 10^{-2} a^3 - 6.76 \cdot 10^{-4} L + 1.73 \cdot 10^{-6} L^2 + 6.84 \cdot 10^{-2} w - 3.12 \cdot 10^{-3} w^2 + 0.724\kappa \tag{18}$$

Where a is expressed in m/s^2 , L in m, w in m/s and κ in veh/m. This global regression is surprisingly accurate: when (18) is tested against model dots for the three planes (a,L) , (a,w) and (a,κ) , the associated RMSE are respectively $3.4 \cdot 10^{-3}$, $4.8 \cdot 10^{-3}$ and $3.2 \cdot 10^{-3}$. Note that (18) predicts the upper limit for the capacity drop because $s=0$, and provides a ready to use extended ND model with an endogenous calculation of the capacity drop. Indeed, when the on-ramp is congested, one only has to compare the downstream supply with the effective capacity derived from (18) to

determine the effective supply and then apply the merge ratio. Unfortunately, this formula is only valid for one-lane on main road. We will see in the next section how we can overcome this limitation.

4. Model validation

4.1. Experimental data

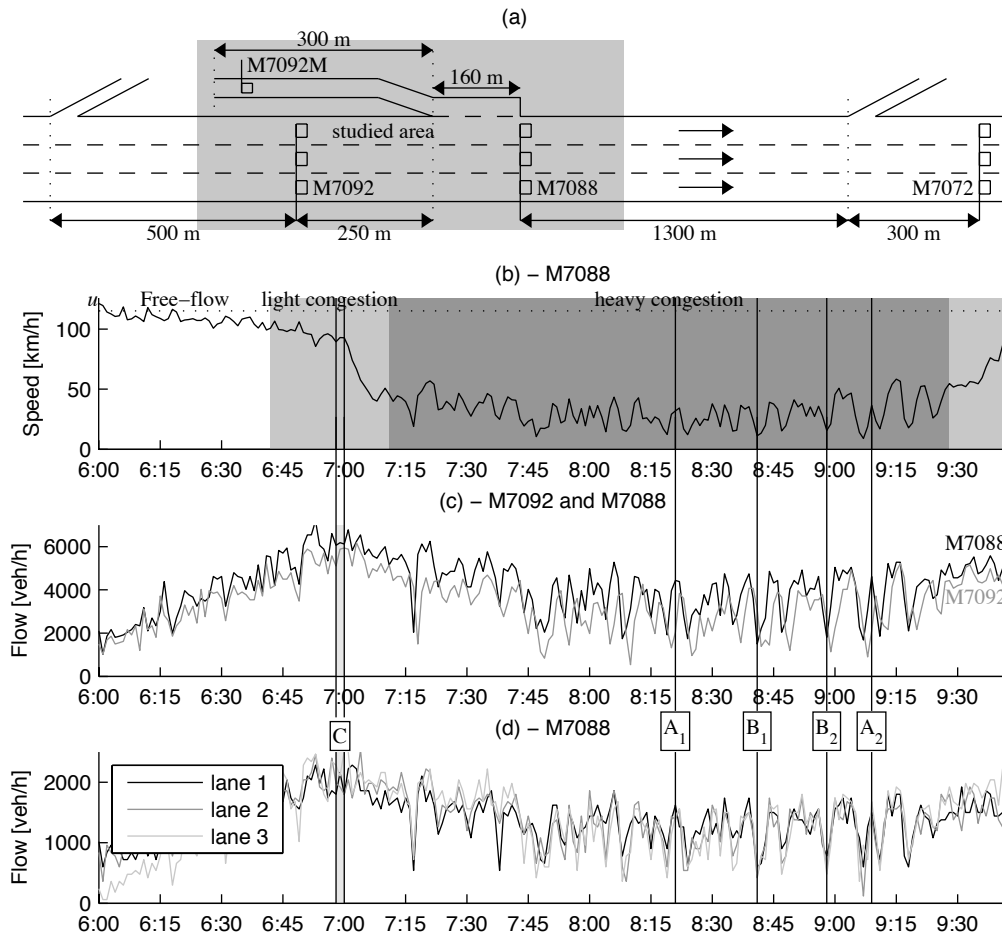


Fig. 8: (a) sketch of the experimental site (b) speed evolution at detector M7088 (c) flow evolution at detectors M7092 and M7088 (d) flows by lane at detector M7088.

To perform the model validation, we focus on a merge located on a southbound three-lane segment of the M6 highway near Manchester, UK. A sketch of the site is depicted in Fig. 8a. An aerial photograph can be consulted on Google Earth® by looking for the following geographical coordinates: 53°25'14.85"N, 2°34'42.18"O. Regularly staggered loop detectors can be found on this highway section. These detectors provide average flow, speed and occupancy rate per lane every minute. Detector M7072 is located 1600 m downstream of the selected merge and never encounters congestion. This confirms that the selected merge is really an active bottleneck. In this study, we mainly focus on data from detectors M7092 and M7088 that are respectively located 250 m upstream and just downstream of the merge, see Fig. 8a. Another detector M7092M is also useful because a drop in speed here indicates that the on-ramp is fully congested. Unfortunately, this loop is far upstream of the beginning of the merging section. Thus, it is not possible to deduce the inserting flow from these observations in a synchronized manner with observations at M7088 and M7092. Inserting flows are then estimated by the difference between

observations at M7088 and M7092 with a one-minute lag. This lag roughly represents the time that a vehicle needs to travel from M7092 to M7088 in congestion.

All data from May 2006 is available. However, congestion is not observed every day nearby the merge. We have then selected the morning peak hours for four days, i.e May, 2, 3, 9 and 16. Those days suffer the highest levels of congestion. Fig. 8b-d present the observations for May, 3rd between 6:00 and 10:00. Fig. 8b shows the evolution of the speed at M7088, where one can identify the time when the freeway is congested, i.e between 6:45 and 9:30. Combining this with information from M7092M, it is possible to determine when the on-ramp is in free-flow (light congestion on the freeway) or in congestion (heavy congestion on the freeway). In that latter case, Fig. 8c clearly shows stop-and-go waves that appear nearby the merge and that form oscillations.

When the ramp is uncongested, we select time periods when the flows observed at M7092 and M7088 are nearly stable during at least 3 min, e.g. point C in Fig. 8c. When the ramp is congested, we focus on situations where flows on each lane at M7088 are nearly equal (homogeneous lane distribution), e.g. points A₁, A₂, B₁ and B₂ in Fig. 8d. Such a choice will be justified in the next section. At the end of this data filtering process, we obtain couples of insertions versus upstream flows that can be plotted in the merge diagram.

4.2. Comparison with the analytical model predictions

Fig. 9 presents the results of the experimental observations (void circles) for the four selected days. Note that the other marks and the meaning of the grey shaded areas will be defined later in this section. Note also that the fundamental diagram has already been calibrated on this site during congestion (Chiabaut et al, 2008): $w=19.4$ km/h and $\kappa=145$ /km/lane. The free-flow speed u can easily be estimated as 115 km/h by looking at Fig. 8b. This is very close to the speed limitation on UK motorways. The maximum capacity per lane given by the fundamental diagram is thus $Q=2400$ veh/h.

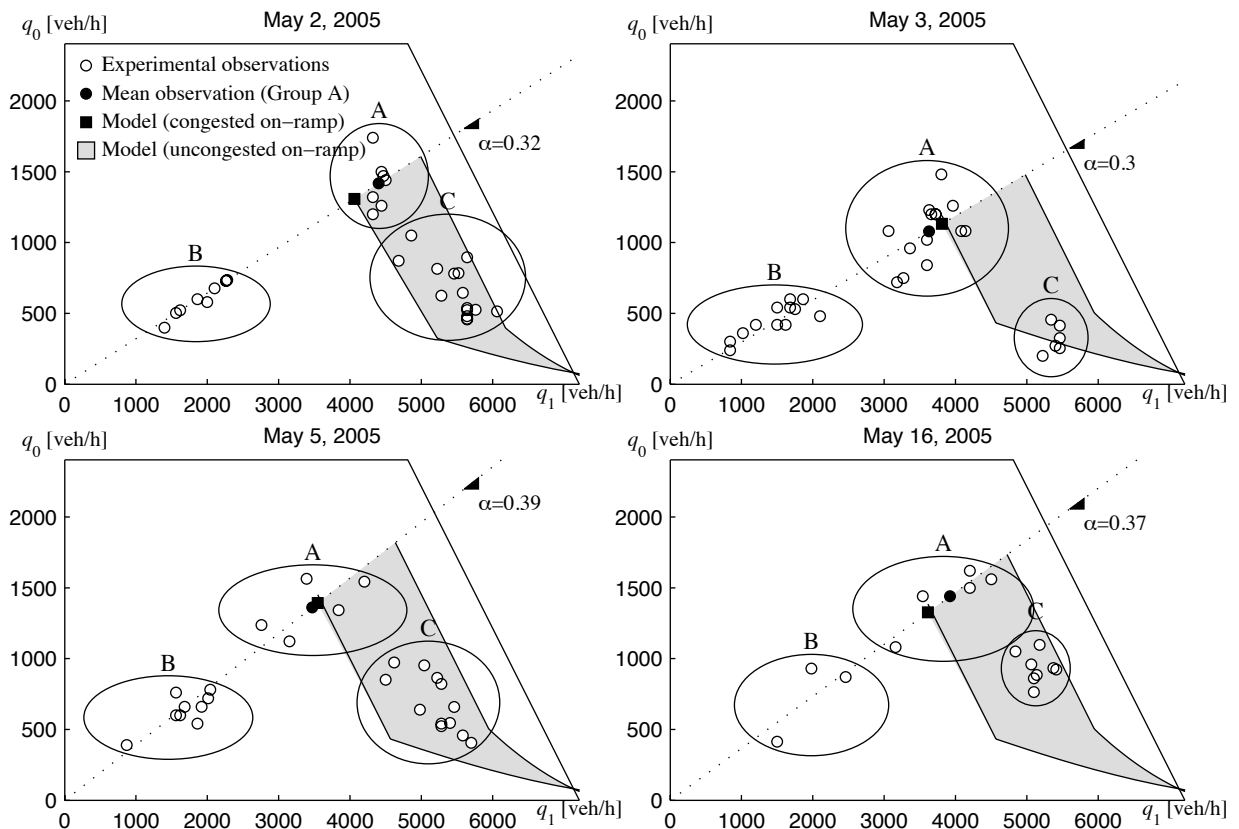


Fig. 9: Comparison between experimental data and model estimations

Fig. 9 demonstrates one more time the relevance of the ND model. Indeed, when the on-ramp is congested, the merging behavior respects a fixed priority ratio; see groups A and B on Fig. 9. When the on-ramp is uncongested, traffic states fall into the right bottom quadrant of ND model; see group C. Notice here that a capacity drop is observed in both cases as predicted by the analytical model. Note also that when the on-ramp is congested, traffic states can be split in two groups. Group B corresponds to heavy congested situations correlated with appearances of stop-and-go waves. Group A corresponds to regular congested situations.

To compare experimental data with the analytical model predictions, the one-lane limitation should be overcome. To this end, we assume that the analytical model properly deals with the merging behavior between the on-ramp and the shoulder lane. The challenge is to account for the lane flow distribution and discretionary lane-changing. We first focus on situations when the on-ramp is congested. The merge ratio α represents the ratio between the inserting flow q_0 and the total upstream flow q_1 . The local merge ratio α' refers to the ratio between q_0 and the upstream flow on the shoulder lane $q_{1,1}$. The sensitivity analysis of the analytical model brings to light that α' does not influence the relative capacity drop c . Thus, even if α' is unknown and potentially variable in time, it is possible to estimate the downstream capacity on the shoulder lane $C=(1-c)Q$. With $a=2 \text{ m/s}^2$, $s=2 \text{ s}$ and $L=160 \text{ m}$, the analytical model leads to $c=0.314$ and $C=1647 \text{ veh/h}$. When the lane flow distribution is homogeneous downstream of the merge, it is easy to determine the total downstream capacity $C'=3C$. This explains why we have only selected, during congestion on the on-ramp, experimental data with homogeneous lane flow distribution at M7080. These observations can adequately be compared with the analytical model.

The model estimation is represented by a solid square in the four subplots of Fig. 9. Note that q_0 and q_1 are derived from C' using the observed merge ratio α . Fig. 9 highlights that the analytical result fit the average observed traffic states of group A, i.e the relative error is always below 7%. This average corresponds to the solid black dots. We can then conclude that the analytical model accurately estimates the regular congested situations for a congested on-ramp when downstream lane distribution is homogeneous. However, Fig. 9 also shows that the analytical result is far from group B observations. This is explained – because the proposed model does not take into account stop-and-go waves.

Let us now focus on situations when the on-ramp is uncongested. It is difficult to find downstream homogeneous lane distributions in that latter case because the global freeway is not heavily congested. Most of the time, flows are highly different on each lane, see Fig. 8d. The median lane can even be in free-flow while the shoulder lane is congested. To compare experimental data with analytical results, we first determine the evolution of the capacity C of the shoulder lane with respect to the on-ramp demand λ_0 using (13). Note that the analytical model in free-flow is only valid when $\alpha' < 1$ because it implicitly assumes that the inserting flow is lower than the main road upstream flow. This is most always the case when the main road has only one lane but this is not necessarily the case here. Indeed, when the congestion is not severe on the freeways, a great proportion of vehicles on the shoulder lane may switch to the center and the median lane. This increases the inserting flow. This limitation can be easily overcome by considering that $C(\lambda_0)$ has a lower-bound equal to the value calculated when the on-ramp is congested.

To derive the total capacity $C'(\lambda_0)$ of the freeway from $C(\lambda_0)$ we consider the two following borderline cases: (a) all downstream lanes experience the same capacity restriction (high level of congestion) and (b) only the shoulder lane experiences a capacity drop while center and median lanes are in free-flow (low level of congestion). The total capacity is then either equal to (a) $C'(\lambda_0)=3C(\lambda_0)$ or (b) $C'(\lambda_0)=C(\lambda_0)+2Q$. The region between these two curves $C'(\lambda_0)$ describes the capacity drop for all possible downstream lane flow distributions. This region is filled with light gray in Fig. 9. It appears that almost all the experimental points from group C fall into the region predicted by the analytical model. This model is therefore able to accurately define the range of possible observed capacity drop when the on-ramp is in free-flow. Note that we have only selected stable traffic flow during at least 3 min in that latter case because the analytical model only predicts mean capacity drop in time, see Fig. 8a.

5. Discussion

We have introduced an analytical model that predicts the capacity drop at an active merge bottleneck with respect to the demand on the on-ramp. This model is quite simple while accounting for most of the key elements of the merging mechanism. It predicts that the relative capacity drop c almost linearly increases with λ_0 when the on-ramp is in free-flow, then reaches its maximal value and remains constant when the on-ramp is congested, see Fig. 4a and Fig. 5a. This analytical model sheds some light on the relative influence of the different parameters involved in the

merging process. The acceleration rate has the highest impact because low rate significantly reduces the available capacity. This is not surprising but the insight here is that the capacity drop value can be directly derived from the mean acceleration rate. The analytical model also predicts that the relative capacity drop is reduced when the length of the insertion section increases. Note here that most of the reduction is observed before the length reaches 100 m. Thus, longer insertion sections do not seem to significantly improve the capacity on freeways. This result has to be confirmed experimentally but it is potentially valuable for road design. The last insight provided by the analytical model is that the merge ratio does not influence the capacity drop when the main road has only one lane. This is not obvious because higher merge ratios lead to higher inserting flows. But this increase is offset by a speed increase on the on-ramp that reduces the perturbations caused by inserting vehicles. This property is convenient when estimating capacity drops for freeways that have more than one-lane, see section 4. All the results synthesized here are certainly useful to improve ramp-metering algorithms.

The first experimental results presented in this paper show that the analytical model is promising. The main current limitation is that it needs to be completed with a lane flow distribution model that accounts for the mean effect of discretionary lane changing when the main road has several lanes. However, it already provides an efficient and elegant way to estimate the upper (when the freeway is uniformly congested) and lower (when congestion only occurs on shoulder lane) bounds of the capacity drop. The authors currently investigate lane flow distribution at merges. The second limitation is that the merge ratio has been fully considered as an exogenous parameter. This hypothesis is reasonable as a first approximation as existing works show that this ratio mostly depends on the merge design (Bar-Gera and Ahn, 2010). However, this should be more deeply investigated in the future because the inserting process may also influence such a ratio when both upstream roads are congested. The last limitation of the analytical model is that it does not reproduce stop-and-go waves when the on-ramp is congested but only the regular capacity drop due the merging process. However, we found that oscillations do not appear due to the random insertion process when vehicles are homogeneous and have the same acceleration rate. We can then conjecture that drivers and vehicles heterogeneities play a major role in the oscillation process. For example, stop-and-go waves may appear when several trucks with low acceleration rate are in the vicinity of the merge at the same time or try to overtake each other. Investigations are made in this direction. They will take benefit of the general framework proposed in (Laval and Leclercq, 2010b) to account for drivers' aggressiveness.

Another great interest of the proposed analytical model is that it can improve any merge model based on the demand / supply framework. We have shown in this paper how to implement it in the ND model by modifying the supply function. This also applies to other classical macroscopic merge models, e.g. (Jin and Zhang, 2003; Lebacque and Koshyaran, 2005) or to microscopic merge models that explicitly accounts for downstream supply (Chevallier and Leclercq, 2009). Indeed, the analytical model directly assesses the capacity drop that reduces the available downstream supply when the merge is active. Note that it can account for the length of the insertion section even if the merge is represented as a point. This model can also be used to test any microscopic behavioral merge models at a macroscopic scale. Indeed, it provides an estimation of the mean capacity drop that should be observed depending on the different parameters. The authors are currently investigating the microscopic merge model proposed in (Laval and Leclercq, 2008) to test its consistency at the macroscopic scale with the findings of this paper and its ability to properly reproduce lane flow distributions and to account for vehicles heterogeneity. Preliminary results are encouraging but research in this realm has to continue.

Acknowledgements

The authors are grateful to the Highways Agency of England for providing data on M6 freeways. The authors would also greatly thank Cécile Becarie for her daily assistance and Estelle Chevallier for her careful reading of this paper.

References

- Ahn, S., Cassidy, M.J., 2007. Freeway Traffic Oscillations and Vehicle Lane-Change Maneuvers, In: Allsop, R., Bell, M.G.H., Heydecker, B.G. (Eds.). *Proceedings of the 17th International Symposium on Transportation and Traffic Theory*, Londres: Elsevier, 691-710.

- Aw, A. and Rascle, M., 2000. Resurrection of « second order » models of traffic flow ? *SIAM Applied Mathematics*, 60, 916-938.
- Bar-Gera, H., Ahn, S., 2010. Empirical macroscopic evaluation of freeway merge-ratios, *Transportation Research C*, 18(4), 457-470.
- Cassidy, M. J., Bertini, R. L., 1999. Some traffic features at freeway bottlenecks, *Transportation Research B*, 33(1), 25-42.
- Cassidy, M.J., Ahn, S., 2005. Driver turn-taking behavior in congested freeway merges. *Transportation Research Record - Journal of the Transportation Research Board*, 1934, 140-147.
- Cassidy, M.J., Rudjanakanoknad, J., 2005. Increasing capacity of an isolated merge by metering its on-ramp. *Transportation Research B*, 39(10), 896-913.
- Chevallier, E., Leclercq, L., 2009. Do microscopic merging models reproduce the observed priority sharing ratio in congestion? *Transportation research part C*, 17(3), 328-336.
- Chiabaut, N., Leclercq, L., Bretin, T., Buisson, C., 2008. Calibrating the Fundamental Diagram in Congestion: Methods Based on Observations at Consecutive Loop-Detectors. *Proceedings of the IWTDSC Conference*, 8-9 September, Barcelona, (Spain) [CDROM]. Tokyo: University of Tokyo, 9 p.
- Chung, K., Rudjanakanoknad, J., Cassidy, M., 2007. Relation between traffic density and capacity drop at three freeway bottlenecks, *Transportation Research B*, 41(1), 82-95.
- Daamen, W., Loo, L., Hoogendoorn, S., 2010. Empirical analysis of merging behavior at a freeway on-ramp. *Proceedings of the Transportation Research Board 89th Annual Meeting*. Washington: TRB (DVD-Rom), 15 p.
- Daganzo, C.F., 2005. A variational formulation of kinematic waves: basic theory and complex boundary conditions. *Transportation Research B*, 39(2), 187-196.
- Daganzo, C., 1995. The cell transmission model, Part II: network traffic. *Transportation Research B*, 29(2), 79-93.
- Elefteriadou, L., Roess, R.P., McShane, W.R., 1995. Probabilistic nature of breakdown at freeway merge junctions, *Transportation Research Record - Journal of the Transportation Research Board*, 1484, 80-89.
- Evans, J., Elefteriadou, L., Natarajan, G., 2001. Determination of the probability of breakdown on a freeway based on zonal merging probabilities. *Transportation Research B*, 35(3), 237–254.
- Greenberg, J., 2001. Extensions and amplifications of a traffic model of Aw and Rascle, *SIAM Journal on Applied Mathematics*, 62, 729–745.
- Kerner, B.S., 2000. Theory of breakdown phenomenon at highway bottlenecks. *Transportation Research Record - Journal of the Transportation Research Board*, 1710, 136–144.
- Kerner, B.S., 2002. Empirical macroscopic features of spatial-temporal traffic patterns at highway bottlenecks, *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 65(4), 1-30.
- Kerner, B.S., 2004. *The Physics of Traffic*. Heidelberg: Springer, 296 p.
- Koshi, M., Iwasaki, M., Okhura, I., 1983. Some findings and an overview on vehicular flow characteristics, In: Hurdle, V.F., Hauer, E., Steuart, G.F. (Eds.) *Proceedings of the 8th International Symposium on Transportation and Traffic Theory*, Toronto: University of Toronto Press, 403-426.
- Hall, F.L., Hall, L.M., 1990. Capacity and speed-flow analysis at the queen Elisabeth way in Ontario, *Transportation Research Record - Journal of the Transportation Research Board*, 1287, 108-118.
- Jin, W.L., Zhang, H.M., 2003. On the distribution schemes for determining flows through a merge. *Transportation Research B*, 37(6), 521-540.
- Laval, J.A., Cassidy, M.J., Daganzo, C.F., 2005. Impacts of lane changes at on-ramp bottlenecks: a theory and strategies to maximize capacity. In: Kühne, R., Poeschl, T., Schadschneider, A., Schreckenberg, M., Wolf, D. (Eds.), *Traffic and Granular Flow '05'*. Berlin: Springer, 577-586.
- Laval, J.A., Daganzo C.F., 2006. Lane-changing in traffic streams. *Transportation Research B*, 40(3), 251-264.
- Laval, J.A., Chen, D., Ben Amer, K., Guin, A., Ahn, S., 2009. Evolution of Oscillations in Congested Traffic: Improved Estimation Method and Additional Empirical Evidence. *Proceedings of the Transportation Research Board 88th Annual Meeting*. Washington: TRB (DVD-Rom), 18 p.
- Laval, J.A., Leclercq, L., 2008. Microscopic modeling of the relaxation phenomenon using a macroscopic lane-changing model. *Transportation Research part B*, 42(6), 511-522.
- Laval, J.A., Leclercq, L., 2010. Continuum Approximation for Congestion dynamics along freeway Corridors. *Transportation Science*, 44, 2010, 87-97.
- Laval, J.A., Leclercq, L., 2010b. A mechanism to describe the formation and propagation of stop-and-go waves in congested freeway traffic. *Philosophical Transactions of Royal Society A*, 368(1928), 4519-4541.
- Lebacque, J.P., Lesort, J.B., Giorgi, F., 1998. Introducing buses into first-order macroscopic traffic flow models. *Transportation Research Record*, 1644, 70-79.
- Lebacque, J.P., Koshiyaran, M.M., 2005. First order macroscopic traffic flow models: intersections modeling, network modeling. In: Mahmassani, H.S. (Ed.), *16th ISTTT*, Pergamon, London, 365-386.
- Leclercq, L., Chanut, S., Lesort, J.B., 2004. Moving bottlenecks in the LWR model: a unified theory. *Transportation Research Record*, 1883, 3-13.
- Leclercq, L., 2005. A new numerical scheme for bounding acceleration in the LWR model. In: Heydecker, B. (Ed.) *4th IMA Conference on Mathematics in Transport studies*, 7-9 September, London, 279-292.
- Li, X., Peng, F., Ouyang, Y., 2010. Measurement and estimation of traffic oscillation properties, *Transportation Research B*, 44(1), 1-14.
- Lighthill, M.J and Whitham, J.B., 1955. On kinematic waves II: A theory of traffic flow in long crowded roads. *Proceedings of the Royal Society*, A229, 317-345.
- Mauch, M., Cassidy, M. J., 2002. Freeway traffic oscillations: observations and predictions. In: Taylor, M.A.P (Ed.) *Proceedings of the 15th International Symposium on Transportation and Traffic Theory*, Amsterdam : Pergamon, 653-674.

- Munoz, J.C., Daganzo, C.F., 2002. Moving Bottlenecks: A Theory Grounded on Experimental Observation. In: Taylor, M. (Ed.) 15th ISTTT, Elsevier, London, 441-462.
- Newell, G.F., 1982. Applications of queueing theory (2nd edition). New York: Chapman & Hall, 303 p.
- Newell, G.F., 1998. A moving bottleneck. *Transportation Research B*. 32(8), 531-537.
- Oehlert, G. W., 1992. A note on the Delta method. *The American Statistician* 46(1), 27-29.
- Persaud, B., Yagar, S., Brownlee, R., 1998. Exploration of the Breakdown Phenomenon in Freeway Traffic. *Transportation Research Record - Journal of the Transportation Research Board*, 1634, 64-69.
- Richards, P.I., 1956. Shockwaves on the highway. *Operations Research*, 4, pp. 42-51.
- Sarvi, M., Kuwahara, M., Ceder, A., 2007. Observing freeway ramp merging phenomena in congested traffic, *Journal of Advanced Transportation*, 41(2), 145-170.
- Siebel, F., Mauser, W., Moutari, S., Rascle, M., 2009. Balanced vehicular traffic at bottleneck, *Mathematical and Computer Modelling*, 49(3-4), 689-702.
- Siebel, F., Mauser, W., Moutari, S., Rascle, M., 2009b. Modeling synchronized flow at highway bottlenecks, In: Appert-Rolland, C., Chevoir, F., Gondret, P., Lassarre, S., Lebacque, J.P., Schreckenberg, M. (Eds), *Traffic and Granular Flow '07'*. Berlin: Springer, 201-210.
- Treiber, M., Kesting, A., Helbing, D., 2006. Understanding widely scattered traffic flows, the capacity drop, and platoons as effects of variance-driven time gaps. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 74(2), 1-10.
- Troutbeck, R.J., 2002. The performance of uncontrolled merges using a limited priority process. In: Taylor, M.A.P., (Ed.), *Proceedings of the 15th International Symposium on Transportation and Traffic Theory*, Amsterdam : Pergamon, 463-482.
- Yi, H., Mulinazzi, T., 2007. Observed distribution patterns of on-ramp merge lengths on urban freeways, *Transportation Research Record - Journal of the Transportation Research Board*, 2023, 120-129.