

Jean-Patrick Lebacque
Maurice Aron
Vincent Aguiléra
Coordination scientifique

MODÉLISATION DU TRAFIC



Actes du groupe de travail 2007



LES COLLECTIONS DE L'INRETS

Conformément à la note du 04/07/2014 de la direction générale de l'Ifsttar précisant la politique de diffusion des ouvrages parus dans les collections éditées par l'Institut, la reproduction de cet ouvrage est autorisée selon les termes de la licence CC BY-NC-ND. Cette licence autorise la redistribution non commerciale de copies identiques à l'original. Dans ce cadre, cet ouvrage peut être copié, distribué et communiqué par tous moyens et sous tous formats.



Attribution — Vous devez créditer l'Oeuvre et intégrer un lien vers la licence. Vous devez indiquer ces informations par tous les moyens possibles mais vous ne pouvez pas suggérer que l'Ifsttar vous soutient ou soutient la façon dont vous avez utilisé son Oeuvre.



Pas d'Utilisation Commerciale — Vous n'êtes pas autorisé à faire un usage commercial de cette Oeuvre, tout ou partie du matériel la composant.



Pas de modifications — Dans le cas où vous effectuez une adaptation, que vous transformez, ou créez à partir du matériel composant l'Oeuvre originale (par exemple, une traduction, etc.), vous n'êtes pas autorisé à distribuer ou mettre à disposition l'Oeuvre modifiée.

Le patrimoine scientifique de l'Ifsttar

Le libre accès à l'information scientifique est aujourd'hui devenu essentiel pour favoriser la circulation du savoir et pour contribuer à l'innovation et au développement socio-économique. Pour que les résultats des recherches soient plus largement diffusés, lus et utilisés pour de nouveaux travaux, l'Ifsttar a entrepris la numérisation et la mise en ligne de son fonds documentaire. Ainsi, en complément des ouvrages disponibles à la vente, certaines références des collections de l'INRETS et du LCPC sont dès à présent mises à disposition en téléchargement gratuit selon les termes de la licence Creative Commons CC BY-NC-ND.

Le service Politique éditoriale scientifique et technique de l'Ifsttar diffuse différentes collections qui sont le reflet des recherches menées par l'institut :

- Les collections de l'INRETS, Actes
- Les collections de l'INRETS, Outils et Méthodes
- Les collections de l'INRETS, Recherches
- Les collections de l'INRETS, Synthèses
- Les collections du LCPC, Actes
- Les collections du LCPC, Etudes et recherches des laboratoires des ponts et chaussées
- Les collections du LCPC, Rapport de recherche des laboratoires des ponts et chaussées
- Les collections du LCPC, Techniques et méthodes des laboratoires des ponts et chaussées, Guide technique
- Les collections du LCPC, Techniques et méthodes des laboratoires des ponts et chaussées, Méthode d'essai



Institut Français des Sciences et Techniques des Réseaux,
de l'Aménagement et des Transports
14-20 Boulevard Newton, Cité Descartes, Champs sur Marne
F-77447 Marne la Vallée Cedex 2

Contact : diffusion-publications@ifsttar.fr

www.ifsttar.fr



Jean-Patrick Lebacque
Maurice Aron
Vincent Aguiléra
Coordination scientifique

Modélisation du trafic

Actes du groupe de travail 2007

Actes INRETS n° 123
Décembre 2009

Coordination scientifique

Jean-Patrick Lebacque INRETS–GRETIA
jean-patrick.lebacque@inrets.fr

Maurice Aron INRETS–GRETIA
maurice.aron@inrets.fr

Vincent Aguiléra LVMT–ENPC
vincent.aguilera@enpc.fr

Unités de recherche

GRETIA, Laboratoire génie des réseaux de transport et informatique avancée,
INRETS, université Paris-Est
2, rue de la Butte Verte, F-93166 Noisy-le-Grand cedex, France.

LVMT, Laboratoire ville, mobilité, transports
INRETS–ENPC–université de Marne-la-Vallée, université Paris-Est
19, rue Alfred-Nobel, cité Descartes, Champs-sur-Marne
F-77455 Marne-la-Vallée cedex 02, France

Institut national de recherche sur les transports et leur sécurité – INRETS

Direction scientifique / politique éditoriale – Aude Lauby
25 avenue François-Mitterrand, case 24, F-69675 Bron cedex, France
Tél. : +33 (0)4 72 14 23 00 – Fax : +33 (0)4 72 37 68 37 – www.inrets.fr

© Les collections de l'INRETS
N° ISBN 978-2-85782-676-7 – N° ISSN 0769-0266

En application du code de la propriété intellectuelle, l'INRETS interdit toute reproduction intégrale ou partielle du présent ouvrage par quelque procédé que ce soit, sous réserve des exceptions légales

Table des matières

Modèle d'écoulement du second ordre : modèle ARZ (Aw-Rascle-Zhang)	5
Salim Mammam, Jean-Patrick Lebacque, Habib Haj-Salem	
Analyse du trafic et du routage dans un réseau ad hoc de véhicules	37
Anthony Busson, Muriel Mabilia, Véronique Vèque	
Ajustement et optimisation des paramètres de poursuite de Newell.....	47
Aurélien Duret, Christine Buisson, Nicolas Chiabaut	
Contrôle d'accès par platitude différentielle	63
Hassane Abouaïssa, Violina Iordanova	
Effects of variable speed limits on motorway traffic flow	81
Markos Papageorgiou, Elias Kosmatopoulos, Ioannis Papamichail	
Real-time merging traffic control with applications to toll plaza and work zone management.....	101
Markos Papageorgiou, Ioannis Papamichail, Anastasia Dimitriou Spiliopoulou, Antonis Fanouriou Lentzakis	
Fiche bibliographique / Publication data form	128

Modèle d'écoulement du second ordre : modèle ARZ (Aw-Rascle-Zhang)

Salim Mammari¹, Jean-Patrick Lebacque², Habib Haj-Salem²

¹ Service d'Étude sur les Transports, les Routes et leurs Aménagements
salim.mammari@developpement-durable.gouv.fr

² Institut National de Recherche sur les transports et leur Sécurité (INRETS)
Laboratoire Génie des Réseaux de Transport et Informatique Avancée
lebacque@inrets.fr
haj-salem@inrets.fr

Résumé

Les modèles de trafic constituent un outil de connaissance en proposant d'importants concepts tels que l'offre et la demande locales de trafic, l'équilibre statique ou dynamique des réseaux, les lois de comportement des usagers et la représentation des écoulements de trafic. Ces modèles constituent un outil de dialogue et d'échange d'informations particulièrement utile pour la prise de décision. Par ailleurs, les modèles se traduisent informatiquement par des outils de simulation, de gestion des réseaux, de contrôle et de planification qui répondent aux attentes des exploitants, des communautés territoriales et des autorités organisatrices.

Dans cet article, on s'intéresse plus particulièrement à un modèle de trafic du second ordre, notamment aux propriétés physiques du modèle et sa résolution du point de vue analytique et numérique.

Introduction

La croissance de la population au niveau mondial ainsi que l'intensification des échanges économiques ont induit une augmentation importante des besoins en matière de transport.

Le domaine des transports, par sa complexité et l'importance du facteur humain, se prête mal à la construction de théories structurées universelles que ce

soit de type mathématique ou physique. Cependant, il existe un besoin important de concepts opératoires pour l'analyse des problèmes et leur quantification. Les modèles de trafic répondent à ce besoin par l'application d'une démarche scientifique aux problèmes posés par le transport.

Afin d'aider les exploitants des réseaux autoroutiers dans leur tâche de gestion du trafic, des modèles de simulation ont été mis au point (Lighthill and Whitham, 1955 ; Richards, 1956 ; Payne, 1971 ; Del Castillo *et al.*, 1993 ; Ross, 1988 ; Aw and Rascle, 2000 ; Zhang, 1998 ; 2002 et Jiang *et al.*, 2002 ; Colombo, 2002 ; Lebacque *et al.*, 2007b)). Le champ d'application de ces modèles de simulation est très diversifié. À titre d'exemples, on peut citer les modèles dédiés à la planification, à la gestion et au contrôle du trafic, les modèles d'aide à la mise en œuvre de stratégies de régulation et l'aide à la prise de décision.

Dans cet article, on s'intéresse plus particulièrement à la résolution analytique du modèle de trafic de second ordre proposé par Aw-Rascle et Zhang (modèle ARZ). Comme on le verra par la suite, la mise en évidence des solutions analytiques exactes relatives à un modèle de second ordre demeure une tâche assez difficile. Par conséquent, nous utiliserons un schéma de Godunov (Kröner, 1997) pour la résolution numérique du modèle.

La validation du schéma numérique de Godunov est réalisée par l'étude de la convergence des solutions numériques obtenues par discrétisation spatio temporelle vers les solutions analytiques.

Par ailleurs, une étude en simulation est fournie à la fin du papier afin de montrer les points forts d'un modèle de second ordre à reproduire les situations de trafic avec des données réelles. Une comparaison à un modèle de simulation du premier ordre LWR (Lighthill and Whitham, 1955) est également établie.

1. Modèle ARZ (Aw-Rascle-Zhang)

Le modèle ARZ proposé par Aw-Rascle et Zhang est basé sur le système d'équations suivant :

Équation de conservation

$$\partial_t \rho(x,t) + \partial_x (\rho v)(x,t) = 0 \quad (1)$$

Équation de vitesse

$$\partial_t v(x,t) + \partial_x \left(v + \rho V_e'(\rho) \right)(x,t) \partial_t v(x,t) = 0 \quad (2)$$

Pour résoudre analytiquement le modèle ARZ (Aw-Rascle-Zhang), il est nécessaire de réécrire le modèle sous la forme conservative. L'un des avantages de cette mise en forme est que les solutions de ce type d'équations sont fonction de la matrice d'états $A(U)$ définie comme la matrice jacobienne de la matrice du flux

$f(U)$. La cinétique des modèles du second ordre est contrôlée par la matrice $A(U)$. La résolution du problème de Riemann conduit à l'apparition d'ondes de choc, d'éventails ou bien de discontinuités de contact. Ces phénomènes sont fonction de la nature des valeurs propres de la matrice $A(U)$. Ces dernières correspondent à la vitesse de propagation des caractéristiques, qui propagent les petites perturbations du trafic à l'intérieur d'un tronçon autoroutier.

Le modèle ARZ mis sous forme conservative se réduit à une équation de la forme :

$$\partial_t U + \partial_x f(U) = 0 \quad (3)$$

avec :

$$U = \begin{pmatrix} \rho(x,t) \\ y(x,t) \end{pmatrix} \quad f(U) = \begin{pmatrix} \rho(x,t)v(x,t) \\ p(x,t) \end{pmatrix}$$

où :

- $l(x,t) \stackrel{def}{=} v(x,t) - V_e(\rho(x,t), x)$: la vitesse relative, i.e. le différentiel entre la vitesse actuelle et la vitesse en régime d'équilibre,
- $y(x,t) \stackrel{def}{=} \rho(x,t)l(x,t) = q(x,t) - Q_e(\rho(x,t), x)$: le débit relatif, i.e. le différentiel entre le débit actuel et le débit en régime d'équilibre comme le montre la Figure 1,
- $Q_e(\rho, x)$: le diagramme fondamental,
- $q(x,t) \stackrel{def}{=} \rho(x,t)v(x,t)$: le débit à l'instant t au point x ,
- $p(x,t) = q(x,t)(v(x,t) - V_e(\rho(x,t), x)) = v(x,t)y(x,t)$: la pression relative.

Comme on peut le remarquer, dans le cas où la vitesse relative du trafic est nulle ($l = 0 \Leftrightarrow v = V_e(\rho)$), le modèle ARZ se réduit à une seule équation qui n'est rien d'autre que le modèle LWR du premier ordre.

Le système (3) est strictement hyperbolique car la matrice $A(U) = \nabla_U f(U)$ admet deux valeurs propres distinctes :

$$\begin{aligned} \lambda_1 &= \frac{y}{\rho} + Q_e'(\rho) = v + \rho V_e'(\rho) \\ \lambda_2 &= \frac{y}{\rho} + V_e(\rho) = v \end{aligned} \quad (4)$$

Les vecteurs propres à droite associés respectivement aux valeurs propres $\lambda_1(U)$ et $\lambda_2(U)$ et que l'on note $R_1(U)$ et $R_2(U)$ sont donnés par :

$$R_1(U) = \begin{pmatrix} -\rho \\ -y \end{pmatrix} \quad R_2(U) = \begin{pmatrix} 1 \\ \frac{y}{\rho} - \rho V_e'(\rho) \end{pmatrix} \quad (5)$$

Les ondes associées à la valeur propre $\lambda_1(U)$ correspondent à des ondes de raréfaction du fait que la valeur propre $\lambda_1(U)$ est vraiment non linéaire (VNL) :

$$\nabla \lambda_1(U) \cdot R_1(U) = Q_e''(\rho) < 0$$

De même, les ondes associées à la seconde valeur propre correspondent à des singularités de contact du fait que $\lambda_2(U)$ est linéairement dégénérée (LD) :

$$\nabla \lambda_2(U) \cdot R_2(U) = 0$$

Par conséquent, les ondes associées à la seconde valeur propre correspondent à des singularités de contact.

2. Résolution analytique du modèle ARZ

Dans cette section, nous allons discuter les différentes solutions du problème de Riemann dans le cas homogène, c'est-à-dire, dans le cas où les deux diagrammes fondamentaux des états amont et aval sont égaux (pas de changement de nombre de voies).

2.1. Mise en évidence des ondes élémentaires

Considérons deux états de trafic amont U_0 et aval U . Comme les ondes de raréfactions sont des fonctions auto-similaires, les ondes de raréfaction associées à la 1-onde sont fonction du paramètre $\xi = \frac{x}{t}$. Par conséquent, si une onde de raréfaction connecte l'état de trafic à gauche $U_0(x \leq 0, t = 0)$ à l'état de trafic à droite $U(x \geq 0, t = 0)$, on aura donc :

$$\begin{aligned} \lambda_1(U(\xi)) &= \frac{x}{t} = \xi \\ \det\left(\frac{d}{d\xi} U(\xi), R_1(U(\xi))\right) &= 0 \end{aligned} \quad (6)$$

D'après l'équation (6), les deux états de trafic U_0 et U sont connectés par une onde de raréfaction si et seulement si :

$$v(\xi) - V_e(\rho(\xi)) = v_0 - V_e(\rho_0) \quad (7)$$

De plus, la valeur propre λ_1 doit satisfaire la condition suivante :

$$\lambda_1 = v(\xi) + \rho(\xi) V_e'(\rho(\xi)) = \xi \quad (8)$$

En combinant les équations (7) et (8), l'onde de raréfaction est donnée par :

$$\begin{cases} \xi = \frac{x}{t} \\ \rho(\xi) = Q_e^{-1}(\xi - (v_0 - V_e(\rho_0))) \\ v(\xi) = V_e(\rho(\xi)) + v_0 + V_e(\rho_0) \end{cases} \quad (9)$$

De la même manière, deux états de trafic U_g et U_d localisés respectivement à gauche et à droite d'une discontinuité de contact au point d'abscisse x_0 sont connectés par une 1-onde ou une 2-onde si et seulement si (condition de Rankine-Hugoniot (Kröner, 1997)) :

$$v_d - v_g = 0 \quad (2 - \text{discontinuité de contact}) \quad (10)$$

$$v_d - V_e(\rho_d) - (v_g - V_e(\rho_g)) \quad (1 - \text{onde de choc}) \quad (11)$$

Deux états du trafic U_g et U_d connectés par une 1-onde et vérifiant l'équation (11) sont reliés par une 1-onde de raréfaction (resp. une 1-onde de choc) si et seulement si $\rho_g > \rho_d$ (resp. $\rho_g < \rho_d$).

De plus d'après l'équation (11) et sachant que $V_e(\rho)$ est une fonction décroissante, on déduit que les états U_g et U_d sont connectés par une onde de raréfaction (resp. 1-onde de choc) si $v_g < v_d$ (resp. $v_g > v_d$).

Les discontinuités de contact (2-onde) propagent les discontinuités de la vitesse relative ou du débit relatif. Cependant, les vitesses des deux côtés de la discontinuité sont égales comme indiqué par l'équation (10).

2.2. Problème de Riemann

En considérant comme conditions initiales deux états de trafic U_g et U_d définis par :

$$\begin{aligned} U(x,0) &= U_g & \text{si } x < 0 \\ U(x,0) &= U_d & \text{si } x > 0 \end{aligned} \tag{12}$$

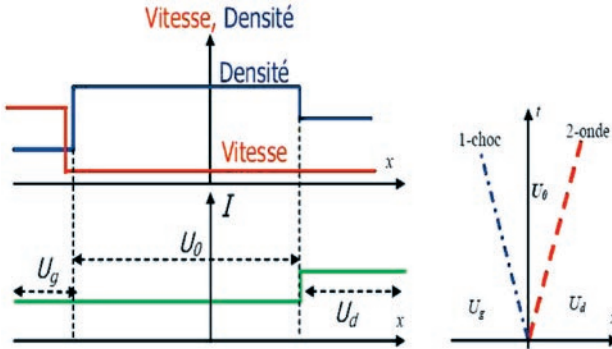
La solution générale du problème de Riemann (eq. 3 et 12) est composée de :

- une 1-onde connectant l'état du trafic U_g à un état du trafic intermédiaire noté U_0 à déterminer,
- une 2-onde connectant l'état du trafic intermédiaire U_0 à l'état du trafic U_d .

Deux types de solutions sont à distinguer en fonction de la nature des 1-ondes qui peuvent être soit des ondes de chocs ou de raréfaction :

- Type 1 : 1-choc connectant l'état amont U_g à un état intermédiaire U_0 suivi par une 2-discontinuité de contact qui relie l'état U_0 à l'état aval U_d (Voir Figure 1)

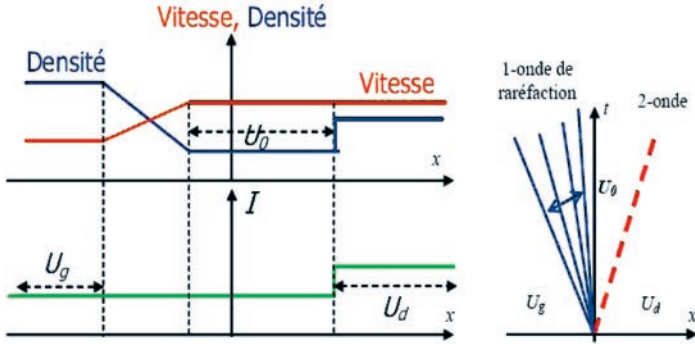
Figure 1 : solutions au problème de Riemann : type 1



- Type 2 : 1-onde de raréfaction connectant l'état amont U_g à un état intermédiaire U_0 suivi par une 2-discontinuité de contact qui relie l'état U_0 à l'état aval U_d (Voir Figure 2).

D'après Lebacque (Lebacque *et al.*, 2007a), le modèle ARZ tel que formulé et sous certaines conditions, le problème de Riemann n'admet pas toujours de solutions. Par conséquent :

Figure 2 : solutions au problème de Riemann : type 2

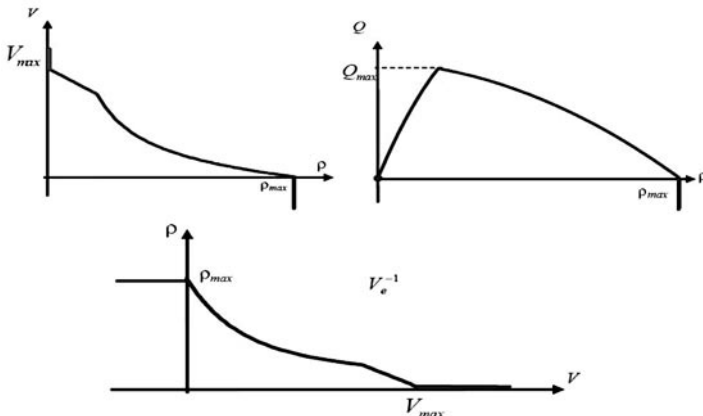


- Le modèle n'admet aucune solution pour certaines conditions initiales qui sont néanmoins sensées représenter des situations réalistes du trafic,
- Le modèle ne peut être résolu numériquement en utilisant par exemple le schéma de Godunov.

Afin de garantir l'existence et l'unicité de solutions physiques au problème de Riemann pour tous les états initiaux possibles du trafic, la solution proposée par Lebacque (Lebacque *et al.*, 2007a) consiste à prolonger le diagramme fondamental et plus particulièrement la réciproque de la relation d'équilibre comme suit :

$$\begin{aligned} V_e^{-1}(v) &= 0 & \text{si } v \geq V_{\max} \\ V_e^{-1}(v) &= \rho_{\max} & \text{si } v \leq 0 \end{aligned} \quad (13)$$

Figure 3 : extension du diagramme fondamental



L'état intermédiaire U_0 entre l'état amont U_g et l'état aval U_d se déduit par combinaison des équations (10) et (11). On obtient ainsi :

$$\begin{aligned} \rho_0 &= V_e^{-1}(v_d - v_g + V_e(\rho_g)) \\ v_0 &= v_d \end{aligned} \quad (14)$$

Comme on peut le remarquer, la réciproque de la fonction d'équilibre $V_e^{-1}(\cdot)$ joue un rôle central pour la résolution du problème de Riemann.

2.3. Solution au problème de Riemann

Dans cette section, nous allons étudier plus en détails les différents cas possibles du problème de Riemann.

En tenant compte de l'extension de la relation d'équilibre vitesse-densité, plusieurs cas sont à distinguer. Ces cas dépendent de l'argument de la fonction $V_e^{-1}(\cdot)$ qui est égal à $v_d - v_g + V_e(\rho_g)$. Sachant que cet argument varie entre 0 et V_{\max} , trois principaux cas sont à distinguer :

- Cas 1 : $v_d - v_g + V_e(\rho_g) > V_{\max}$
- Cas 2 : $0 \leq v_d - v_g + V_e(\rho_g) \leq V_{\max}$
- Cas 3 : $v_d - v_g + V_e(\rho_g) < 0$

Dans cette section, nous allons nous intéresser plus particulièrement au cas 2. Pour les deux autres cas, les solutions analytiques se construisent de manière tout à fait similaire (Mammar *et al.*, 2008a).

- Cas 2 : $0 \leq v_d - v_g + V_e(\rho_g) \leq V_{\max}$

Dans ce cas, nous supposons que l'argument de la réciproque de la fonction d'équilibre V_e^{-1} est compris entre la vitesse nulle et la vitesse maximale autorisée V_{\max} sur un tronçon d'autoroute. On distingue deux cas de figure qui dépendent de la vitesse de l'état amont v_g et celle de l'état aval v_d .

- a) Cas 2.1 : $v_d \leq v_g$

Lorsque la vitesse du trafic de l'état amont v_g est plus grande que la vitesse du trafic de l'état aval v_d , la solution du problème de Riemann est composée de :

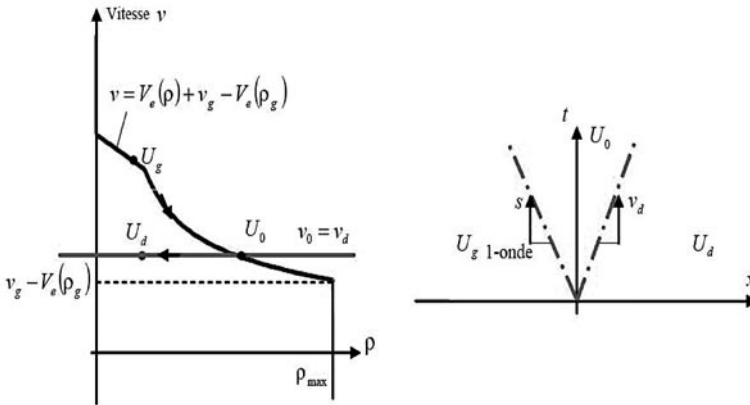
- 1-onde de choc (1-OC) qui connecte l'état U_g à l'état intermédiaire U_0 ,
- 2-discontinuité de contact (2-DC) qui connecte l'état intermédiaire U_0 à l'état amont U_d .

L'état intermédiaire U_0 est donné par l'équation (14). La vitesse de propagation (s) de l'onde de choc peut être évaluée par la relation de Rankine-Hugoniot suivante :

$$s = \frac{q_0 - q_g}{\rho_0 - \rho_g} \Rightarrow v_d - s = \rho_g \frac{v_g - v_d}{\rho_0 - \rho_g} > 0 \quad (15)$$

car $\rho_0 > \rho_g$, $v_g > v_d$

Figure 4 : caractéristiques de l'onde de choc et la discontinuité de contact dans les plans (ρ, v) et (x, t)



Afin de calculer les caractéristiques de l'état à l'origine défini par v_w et ρ_w , on distingue deux sous-cas qui sont fonction du signe de la vitesse de propagation de l'onde de choc (s).

- Sous-cas 2.1.1 : $q_0 - q_l \leq 0$ ($s \leq 0$)

Lorsque la vitesse de propagation de l'onde de choc est négative, la solution du problème de Riemann est donnée par :

$$\begin{cases} v_w = v_0 = v_d \\ \rho_w = \rho_0 = V_e^{-1}(v_d - v_g + V_e(\rho_g)) \\ q_w = q_0 = \rho_0 v_d \\ p_w = q_w (v_g - V_e(\rho_g)) \end{cases}$$

- Sous-cas 2.1.2 : $q_0 - q_l \geq 0$ ($s \geq 0$)

Lorsque la vitesse de propagation de l'onde de choc est positive, la solution du problème de Riemann est donnée par :

$$\begin{cases} v_w = v_g \\ \rho_w = \rho_g \\ q_w = q_g = \rho_g v_g \\ \rho_w = \rho_g = q_g (v_g - V_e(\rho_g)) \end{cases}$$

- b) Cas 2.2 : $v_d \geq v_g$

Dans ce cas, lorsque la vitesse du trafic de l'état aval v_d est supérieure à la vitesse du trafic de l'état amont v_g , la solution du problème de Riemann est composée de :

- une onde de raréfaction qui connecte l'état amont U_g à l'état intermédiaire U_0 ,
- une discontinuité de contact qui connecte l'état intermédiaire U_0 à l'état aval U_d .

La vitesse de la caractéristique qui porte l'état intermédiaire U_0 est donnée par :

$$\lambda_1 = v_0 + \rho_0 V_e'(\rho_0)$$

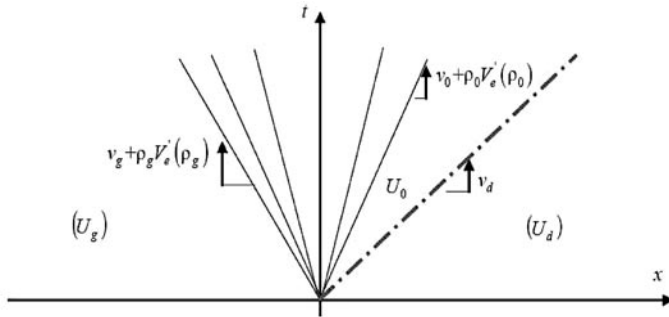
avec

$$\begin{cases} v_0 = v_d \\ \rho_0 = V_e^{-1}(v_d - v_g + V_e(\rho_g)) \end{cases}$$

Dans ce cas de figure, trois sous-cas sont à distinguer :

- $\lambda_1(U_g) \geq 0$ et $(U_w = U_g)$,
- $\lambda_1(U_0) \leq 0$ et $(U_w = U_0)$,
- $\lambda_1(U_g) \leq 0 \leq \lambda_1(U_0)$.

Figure 5 : caractéristiques de l'onde de raréfaction et la discontinuité de contact dans le plan (x, t)



- Cas 2.2.1 : $\lambda_1(U_g) = v_g + \rho_g V_e'(\rho_g) \geq 0$: dans ce cas, la solution au problème de Riemann est donnée par :

$$\begin{cases} v_w = v_l \\ \rho_w = \rho_l \\ q_w = q_l = \rho_l v_l \\ p_w = p_l = q_l (v_l - V_e(\rho_l)) \end{cases}$$

- Cas 2.2.2 : $\lambda_1(U_0) = v_0 + \rho_0 V_e'(\rho_0) \leq 0$: dans ce cas, la solution au problème de Riemann est donnée par :

$$\begin{cases} v_w = v_d \\ \rho_w = \rho_0 = V_e^{-1}(v_d - v_g + V_e(\rho_g)) \\ q_w = \rho_w v_w \\ p_w = q_w (v_g - V_e(\rho_g)) \end{cases}$$

- Cas 2.2.3 : $\lambda_1(U_g) = v_g + \rho_g V_e'(\rho_g) \leq 0 \leq \lambda_1(U_0) = v_0 + \rho_0 V_e'(\rho_0)$: dans ce dernier sous-cas, la solution au problème de Riemann est donnée par :

$$\begin{cases} v_w = V_e(\rho_w) + v_g - V_e(\rho_g) \\ \rho_w = Q_e^{-1}(-v_g + V_e(\rho_g)) \\ q_w = \rho_w v_w \\ p_w = q_w (v_g - V_e(\rho_g)) \end{cases}$$

3. Calcul de certaines solutions analytiques complexes

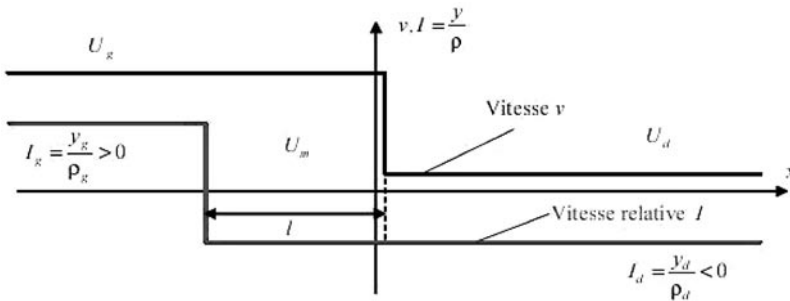
Comme on vient de le voir, le calcul des solutions analytiques du modèle ARZ reste une tâche très difficile même dans le cas de scénarios assez classiques. Pour illustrer cette complexité, nous allons nous intéresser à la résolution du problème de Riemann dans le cas d'une interaction d'une onde de choc (1-onde) avec une onde de débit relatif (2-onde) (Mammar *et al.*, 2008b ; Lebacque *et al.*, 2008a). Comme les 1-onde peuvent être de choc ou de raréfaction deux cas de figure sont à considérer :

- Interaction d'une 1-onde de choc avec une 2-onde,
- Interaction d'une 1-onde de raréfaction avec une 2-onde.

3.1. Interaction d'une 1-onde de choc avec une 2-onde

Considérons une situation de trafic avec un état initial défini par la figure 6. Ce cas correspond à une discontinuité du débit relatif y (ou vitesse relative $l = \frac{y}{\rho}$) suivie d'une onde de choc (discontinuité de vitesse).

Figure 6 : interaction d'une 2-onde et d'une 1-onde, conditions initiales



La discontinuité du débit relatif va générer une 2-onde qui se propage à la vitesse du trafic. L'onde de choc (1-onde) se propage à une vitesse beaucoup plus petite que la vitesse du trafic. Par conséquent, il y a une interaction entre ces deux types d'ondes. Ceci se produit lorsque la 2-onde rattrape la 1-onde. Les deux états du trafic amont U_g et aval U_d sont séparés par un état que l'on note U_m (Voir Figure 6). Cet état a pour caractéristiques :

$$\begin{aligned}
 v_m &= v_g \\
 \rho_m &= V_e^{-1}(v_g - v_d + V_e(\rho_d))
 \end{aligned}
 \tag{16}$$

Ces équations expriment l'existence d'une 2-onde entre l'état U_g et l'état U_m et une 1-onde de choc entre l'état U_m et l'état U_d . Les deux ondes s'intersectent à l'instant t_0 donné par :

$$t_0 = \frac{l}{v_d - v_{md}}$$

avec v_{md} la vitesse de propagation de la 2-onde et l la distance initiale qui sépare les deux états du trafic U_g et U_d . À l'instant t_0 , les deux états de trafic U_g et U_d sont côte à côte, dans ce cas on obtient un problème de Riemann simple dont la solution est composée d'une 1-onde de choc qui connecte l'état U_g à un état intermédiaire U_p , et d'une 2-onde qui connecte cet état intermédiaire à l'état U_d . L'état intermédiaire U_p est donné par les équations suivantes :

$$v_p = v_d$$

$$\rho_p = V_e^{-1}(v_g - v_g + V_e(\rho_g))$$

La solution analytique complète dans le plan (ρ, v) et (x, t) est décrite par les Figures 7 et 8.

Figure 7 : solution analytique dans le plan (ρ, v) : interaction d'une 1-onde et une 2-onde

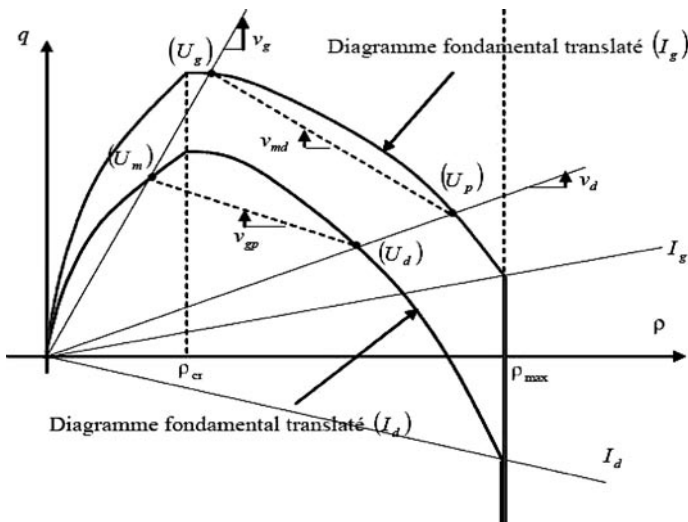
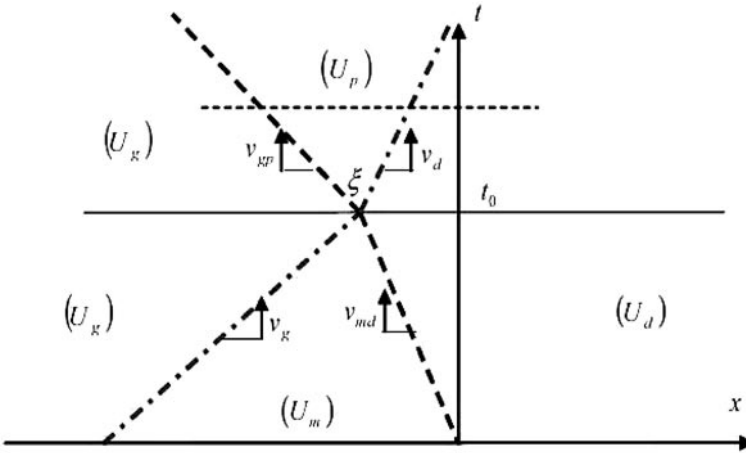


Figure 8 : solution analytique dans le plan (x, t) : interaction d'une 1-onde et une 2-onde



Les coordonnées du point d'intersection des deux ondes sont données par :

$$\xi = \frac{lv_{md}}{v_g - v_{md}} = v_{md} t_0$$

La solution est donnée par ce qui suit :

– Si $t \leq t_0$

$$U = \begin{cases} U_u & \text{si } x \leq -l + v_g l \\ U_m & \text{si } -l + v_g l \leq x \leq v_{md} l \\ U_d & \text{si } x \geq v_{md} l \end{cases}$$

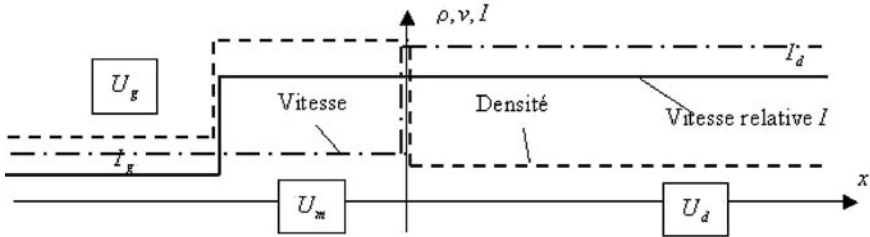
– Si $t \geq t_0$

$$U = \begin{cases} U_u & \text{si } x \leq \xi + v_{lp}(l - l_0) \\ U_m & \text{si } \xi + v_{lp}(l - l_0) \leq x \leq \xi + v_d(l - l_0) \\ U_d & \text{si } x \geq \xi + v_d(l - l_0) \end{cases}$$

3.2. Interaction d'une 1-onde de raréfaction avec une 2-onde

Considérons maintenant le problème de Riemann dans le cas d'une interaction d'une 1-onde de raréfaction avec une 2-onde avec des conditions initiales définies par la Figure 9.

Figure 9 : interaction d'une 1-onde de raréfaction avec une 2-onde, conditions initiales



Comme on peut le remarquer, il y a deux valeurs de la vitesse relative I_g et I_d qui conduisent à une translation du diagramme fondamental. Par ailleurs, la discontinuité initiale $I_g \rightarrow I_d$ qui sépare les états de trafic U_g et U_m ont des vitesses identiques égales à v_g du fait qu'à l'instant initial, l'onde de raréfaction se produit à l'origine et la 2-onde, qui apparaît, propage la discontinuité de la vitesse relative.

Considérons dans un premier temps le cas de la 2-onde. Cette dernière se propage dans l'état de trafic U_g à la vitesse v_g . Lorsque la 2-onde rencontre l'onde de raréfaction, la 2-onde se propage dans l'état de trafic U_d à la vitesse v_d . Ce phénomène conduit à un changement de l'état U_d (commutation de son état de trafic à un état U_p qui est caractérisé par $I = I_g$ et $v = v_d$) :

$$v_p = v_d$$

$$\rho_p = V_e^{-1}(v_d - I_g)$$

Considérons à présent le cas d'une 1-onde de raréfaction. Dans le plan (x, t) , toutes les droites, à l'origine et contenues dans l'onde, sont des lignes de caractéristiques de pente ξ égale à la première valeur propre donnée par l'équation (4), c'est-à-dire, $\lambda_1 = \xi = I_d + Q_e'(\rho_\xi^d)$. Ces caractéristiques se distinguent par un état de trafic $U_\xi^d = (\rho_\xi^d, v_\xi^d)$ uniforme. Il est ainsi possible d'exprimer l'état du trafic en fonction de la pente ξ :

$$\rho_\xi^d = Q_e^{-1}(\xi - I_d)$$

La détermination de la trajectoire de la 2-onde dans la 1-onde de raréfaction se fait par intégration du champ de vitesse :

$$\dot{x} = V_e(\rho(x, t)) + I_d = V_e(Q_e^{-1}(\xi - I_d)) + I_d$$

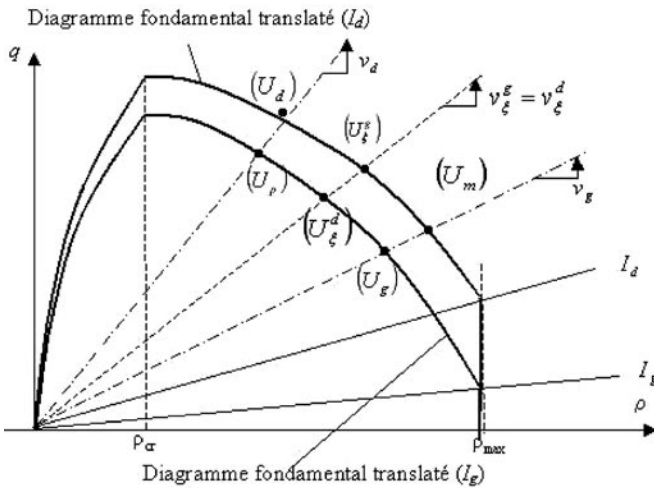
avec $\xi = x/t$.

Le passage de la 2-onde conduit à un changement de la 1-onde de l'état $U_{\xi}^d = (\rho_{\xi}^d, v_{\xi}^d)$ à l'état $U_{\xi}^g = (\rho_{\xi}^g, v_{\xi}^g)$. Les deux états de trafic sont caractérisés par des vitesses identiques et des vitesses relatives différentes. Ce futur état $U_{\xi}^g = (\rho_{\xi}^g, v_{\xi}^g)$ est déterminé par :

$$v_{\xi}^d = l_d + V_e(\rho_{\xi}^d) = v_{\xi}^g = l_d + V_e(\rho_{\xi}^g) \Rightarrow \rho_{\xi}^g = V_e^{-1}(V_e(\rho_{\xi}^d) + l_d - l_g)$$

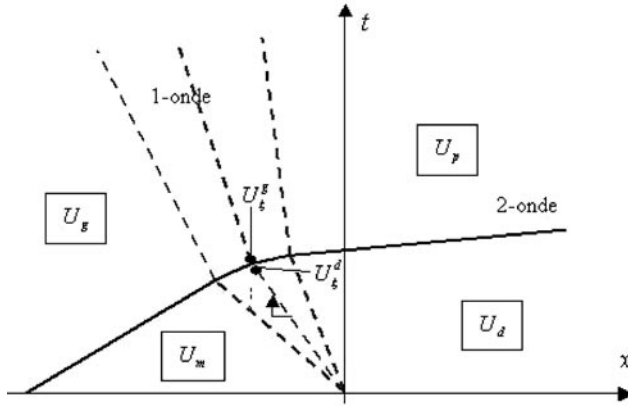
En conclusion, les deux états de trafic U_g et U_p sont reliés par une 1-onde du côté gauche de la 2-onde. Comme le montre la figure ci-dessous, cette onde contient les états situés entre U_g et U_p du diagramme fondamental translaté d'une quantité l_g .

Figure 10 : interaction d'une 1-onde de raréfaction avec une 2-onde : solution analytique dans le plan (ρ, q)



Dans le plan (x, t) les états de trafic U_g et U_p sont reliés par un éventail de raréfaction (1-onde) qui commence depuis l'état U_{ξ}^g situé sur la 2-onde (voir Figure 11).

Figure 11 : interaction d'une 1-onde de raréfaction avec une 2-onde : solution analytique dans le plan (x, t)

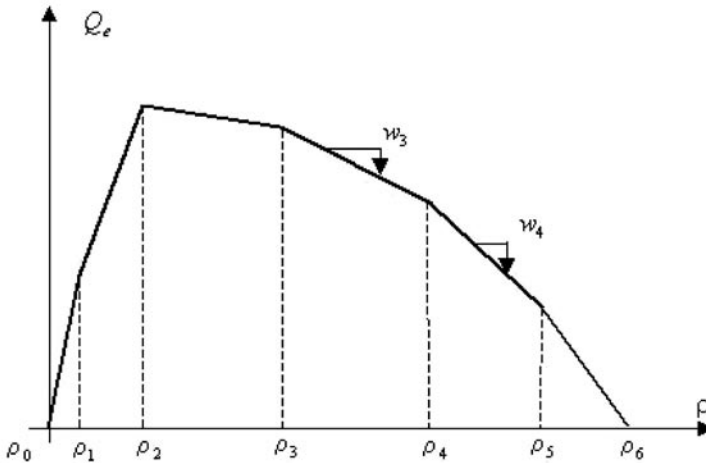


3.3. Cas d'un diagramme fondamental linéaire par morceau

À présent, considérons le cas d'un digramme fondamental linéaire par morceaux, c'est-à-dire, avec un nombre fini d'ondes de vitesses, telles que :

$$Q_e'(\rho) = w_j \quad \text{si} \quad \rho_j \leq \rho \leq \rho_{j+1}$$

Figure 12 : diagramme fondamental linéaire par morceaux



De même que dans les cas précédents, lorsque le diagramme fondamental linéaire par morceaux est translaté d'une quantité l , il reste toujours continu par morceau à la seule différence que les vitesses des ondes sont translatées d'une quantité l :

$$(Q_e + l)'(\rho) = w_j + l \quad \text{if} \quad \rho_j \leq \rho \leq \rho_{j+1}$$

Les 1-ondes de raréfaction basées sur de tels diagrammes ont une structure très spéciale. L'espace des ondes de raréfaction dans le plan (x, t) reliant les états de trafic en amont U_g et aval U_d est caractérisé par une densité constante $\rho = \rho_i$ quel que soit la valeur de la vitesse relative l avec :

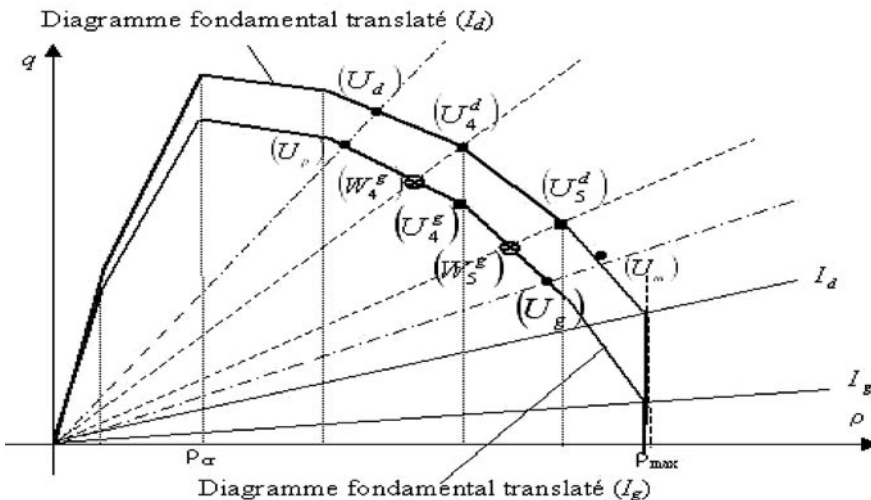
$$\rho_d \leq \dots \leq \rho \leq \dots \leq \rho_g$$

L'espace dans lequel $\rho = \rho_i$ est séparé du secteur dans lequel la densité $\rho = \rho_{i+1}$ par une discontinuité qui se propage à la vitesse que l'on note par w_i (pour plus de précision le lecteur se référer aux travaux de Velan et à Florian (Velan and Florian, 2002)). Comme les ondes de tous les états, tels que $\rho_i \leq \rho \leq \rho_{i+1}$, se propagent à vitesse constante, l'extension physique de tous ces états est constante dans le temps.

Dans le cas d'une prolongation initiale nulle (les états provenant d'un seul point comme les ondes de raréfaction), cette prolongation reste nulle à chaque instant. Seuls les états avec deux vitesses d'onde peuvent avoir une prolongation différente de zéro pour les ondes de raréfaction. Le point crucial est que les discontinuités entre les états ont des propriétés analogues aux ondes de chocs.

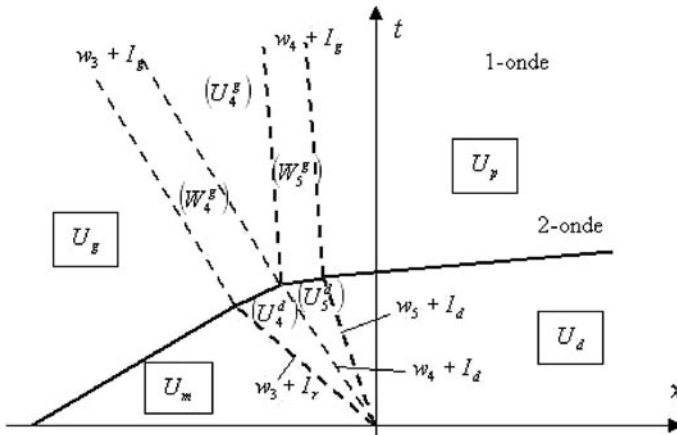
Afin d'analyser le phénomène d'interaction entre une 2-onde et une 1-onde de raréfaction dans les cas d'un diagramme fondamental continu par morceaux, nous allons nous appuyer sur l'analyse déjà faite dans la sous-section précédente.

Figure 13 : interaction d'une 1-onde de raréfaction avec une 2-onde : solution analytique dans le plan (ρ, q) avec un diagramme fondamental linéaire par morceaux



Pour cela, considérons les conditions initiales données par la Figure 9 et un diagramme fondamental illustré par la Figure 12. Dans un premier temps, il faut identifier les états de trafic impliqués en construisant des courbes linéaires par morceau par analogie à la Figure 10. En considérant l'exemple de la Figure 13, la 1-onde $U_m \rightarrow U_d$ contient deux secteurs associés aux états de trafic U_5^d et U_4^d . Après le passage de la 2-onde, il se produit une commutation des états de trafics U_5^d et U_4^d aux états W_5^g et W_4^g (vitesse identique mais avec des vitesses relatives différentes). Par ailleurs, il y a une 1-onde de raréfaction $W_5^g \rightarrow W_4^g$ qui contient un secteur dans lequel le trafic est à l'état U_4^g . La solution résultante dans le plan (x, t) est décrite par la Figure 14.

Figure 14 : interaction d'une 1-onde de raréfaction avec une 2-onde : solution analytique dans le plan (x, t) avec un diagramme fondamental linéaire par morceaux



4. Résolution numérique du modèle ARZ

Les solutions analytiques du modèle ARZ restent assez complexes à déterminer. Ceci nous mène à résoudre par des schémas numériques le modèle ARZ. Dans la littérature plusieurs schémas numériques ont été proposés. On peut citer les schémas WENO (Zhang *et al.*, 2003) (*weighted essentially non-oscillatory*), les schémas de suivi d'ondes (*wave-tracking*) (Henn, 2005) qui reposent sur des solutions analytiques obtenues en approximant le diagramme fondamental par une fonction linéaire par morceaux et le schéma de Godunov (Kröner, 1997).

Dans cette section nous allons considérer le schéma de suivi d'ondes ainsi que le schéma de Godunov.

4.1. Schéma de suivi d'onde (wave-tracking)

L'analyse présentée dans la sous-section 3.3 montre qu'il est assez simple de prolonger la technique de suivi d'ondes au modèle ARZ. L'idée de base du schéma de suivi d'ondes est de calculer la solution exacte du système 3 en considérant :

- Des conditions initiales constantes par morceaux,
- Un diagramme fondamental linéaire et constant par morceaux,
- Une densité et une vitesse relative en entrée d'un segment d'autoroute constant par morceaux,
- Une condition limite en sortie d'un segment d'autoroute constante par morceaux.

Le calcul de la solution du système (3) est entièrement décrite par la dynamique de la 1-onde et de la 2-onde qui interagissent entre elles comme nous venons de le voir dans les sous-sections 3.1 et 3.3.

La solution analytique exige le calcul des temps d'intersection des ondes (comme dans la sous-section 3.1), des vitesses des ondes et le suivi des densités.

Les seules densités à propager sont :

- les densités initiales et en entrée du segment (conditions limites),
- les densités ρ_j et les densités qui se produisent par l'interaction avec la 2-onde (W_5^g, W_4^g comme dans l'exemple de la sous-section 3.3).

Les seules valeurs des vitesses relatives l à propager sont les valeurs initiales et en entrée d'un segment.

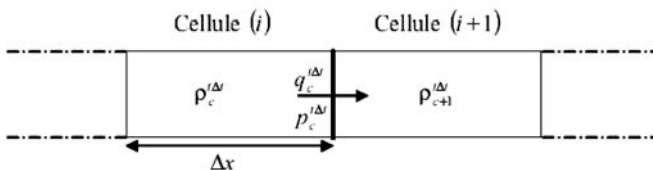
Pour plus de détails, le lecteur peut se référer aux travaux de Henn (Henn, 2005) pour l'application de cette méthode au modèle de trafic du premier ordre LWR ainsi qu'aux travaux de Wong (Wong, 1998).

4.2. Schéma de Godunov

La résolution numérique par le schéma de Godunov pour une section d'autoroute homogène est basée sur les idées suivantes :

- On décompose une section d'autoroute en cellules $(c) = [x_{c-1}, x_c]$. De même, on décompose le temps avec un pas égal à $(t) = [t\Delta t, (t+1)\Delta t]$ comme le montre la Figure 15.

Figure 15 : principe de discrétisation d'un réseau autoroutier



- La densité ρ et le débit relatif y sont supposés homogènes sur chaque cellule (c) au début de chaque pas de temps (t), ils sont égaux à ρ_c^t et y_c^t respectivement. Ainsi, la densité $\rho(x, t)$ et le débit relatif $y(x, t)$ sont approxi-
més au début de chaque pas par des fonctions constantes par morceaux $\tilde{\rho}(x, t)$ et $\tilde{y}(x, t)$:

$$\begin{aligned}\tilde{\rho}(x, t\Delta t) &= \rho_c^t & \text{si } x \in (c) &= [x_{c-1}, x_c] \\ \tilde{y}(x, t\Delta t) &= y_c^t & \text{si } x \in (c) &= [x_{c-1}, x_c]\end{aligned}\quad (17)$$

- Le flux entre deux cellules pendant un pas de temps est calculé en résolvant un problème de Riemann en chaque point x_c de la frontière entre la cellule (c) et la cellule ($c + 1$). Le respect de la condition de Courant-Friedrichs-Lewy (condition CFL) assure la non-interaction entre les solutions du problème de Riemann aux points x_{c-1} et x_c . Par conséquent, le flux est donné par :

$$\begin{pmatrix} q_c^t \\ p_c^t \end{pmatrix} = F \left(\begin{pmatrix} \rho_c^t \\ y_c^t \end{pmatrix}, \begin{pmatrix} \rho_{c+1}^t \\ y_{c+1}^t \end{pmatrix} \right)\quad (18)$$

- La solution au problème de Riemann en chaque point x_c d'une cellule est approximée par deux solutions que l'on note $\tilde{\rho}(x, \tau)$ et $\tilde{y}(x, \tau)$, et ce quel que soit $\tau \in (t) = [t\Delta t, (t+1)\Delta t]$.
- Les moyennes des densités $\tilde{\rho}(x, (t+1)\Delta t)$ et des débits relatifs $\tilde{y}(x, (t+1)\Delta t)$ à l'instant $(t+1)$ représentent respectivement la densité ρ_c^{t+1} et le débit relatif y_c^{t+1} .
- Les moyennes des cellules à l'instant $(t+1)\Delta t$ peuvent être calculées en appliquant la propriété de conservation de la densité ρ (conservation du nombre de véhicules) et du débit relatif y pendant un pas de temps $(t+1)$ pour chaque cellule (c). Cette moyenne est donnée par :

$$\begin{aligned}\rho_c^{t+1} &= \rho_c^t + \frac{\Delta t}{x_c - x_{c-1}} \left(-q_c^t + q_{c-1}^t \right) \\ y_c^{t+1} &= y_c^t + \frac{\Delta t}{x_c - x_{c-1}} \left(-p_c^t + p_{c-1}^t \right)\end{aligned}\quad (19)$$

5. Étude en simulation

Dans cette section, nous allons nous intéresser à l'étude en simulation de la convergence des solutions numériques vers les solutions analytiques ainsi qu'à la mise en évidence de quelques phénomènes de trafic telles que les ondes d'arrêts et de redémarrages. Par ailleurs, une confrontation du modèle ARZ sur des données réelles est fournie à la fin de la section. Pour toutes ces études en simulation, on considère la relation d'équilibre suivante :

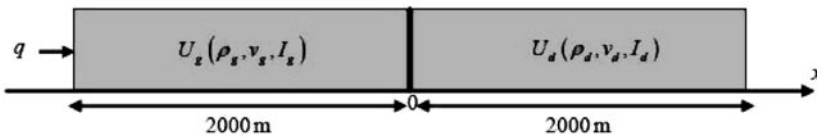
$$\begin{aligned}
 Q_e(\rho) &= \rho V_{\max} - \frac{\rho^2}{\rho_{cr}}(V_{\max} - V_{cr}) & \text{si } 0 \leq \rho \leq \rho_{cr} \\
 Q_e(\rho) &= Q_{\max} \left[1 - \left(\frac{\rho - \rho_{cr}}{\rho_{\max} - \rho_{cr}} \right)^2 \right] & \text{si } \rho_{cr} \leq \rho \leq \rho_{\max}
 \end{aligned} \tag{20}$$

5.1. Étude comparative des solutions analytiques et numériques

Afin de montrer la validité du schéma numérique, on se propose de faire une étude comparative entre les solutions analytiques exactes et les solutions numériques obtenues par discrétisation spatiotemporelle.

Pour cela, on considère un tronçon d'une longueur de 4000 mètres comme le montre la Figure 16.

Figure 16 : conditions initiales pour le problème de Riemann



Les conditions de trafic amont et aval sont :

$$U_g \left(\rho_g = \frac{\rho_{cr}}{2}, v_g < V_e(\rho_g) \right) \text{ et } U_d \left(\rho_d = \frac{\rho_{\max}}{2}, v_d < V_e(\rho_d) \right)$$

La Figure 17 montre que la solution numérique est très proche de la solution analytique. De plus, une réduction des pas de discrétisations spatiale et temporelle a pour effet une meilleure convergence de la solution numérique vers la solution analytique (Voir Figure 18).

Figure 17 : étude comparative entre les solutions analytique et numérique.
Cas 2.1.1

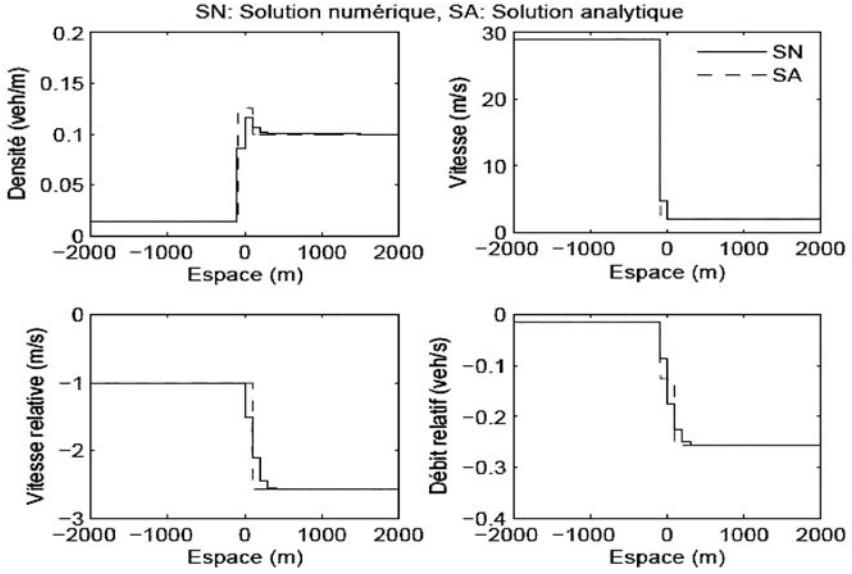
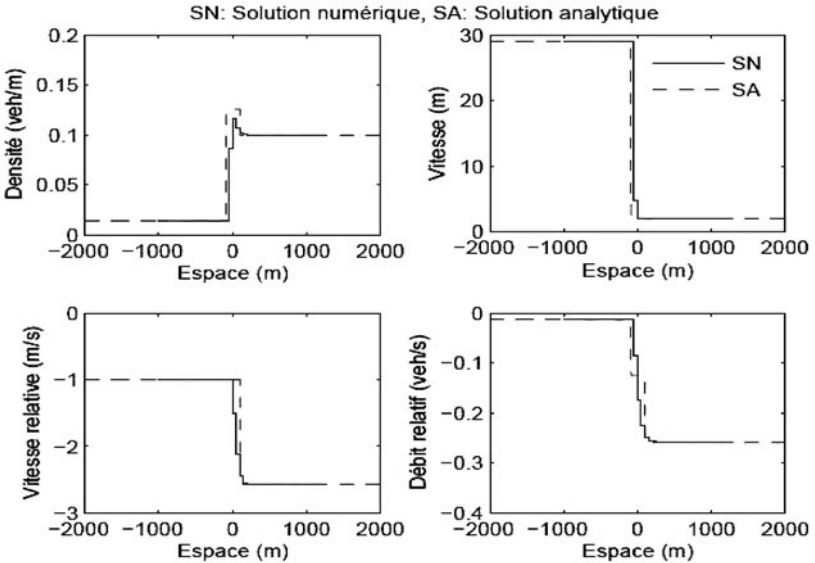


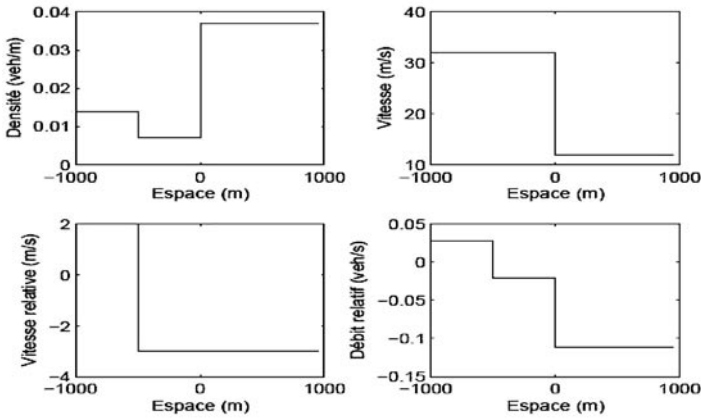
Figure 18 : influence de la réduction de l'échelle spatio-temporelle.
Cas 2.1.1



De même, les Figures 19 (conditions initiales), 20, 21, 22 et 23 représentent une étude comparative entre la solution analytique et numérique dans le cas d'une interaction entre une 2-onde et d'une 1-onde. Chaque figure représente la solution analytique avec la solution numérique et ce pour les quatre variables du modèle : la densité, la vitesse, la vitesse relative et le débit relatif. La résolution espace-temps de la discrétisation est de 50 mètres et 1 seconde respectivement. Les simulations montrent un excellent accord entre les solutions numérique et analytique même si la situation est compliquée par l'interaction de deux ondes.

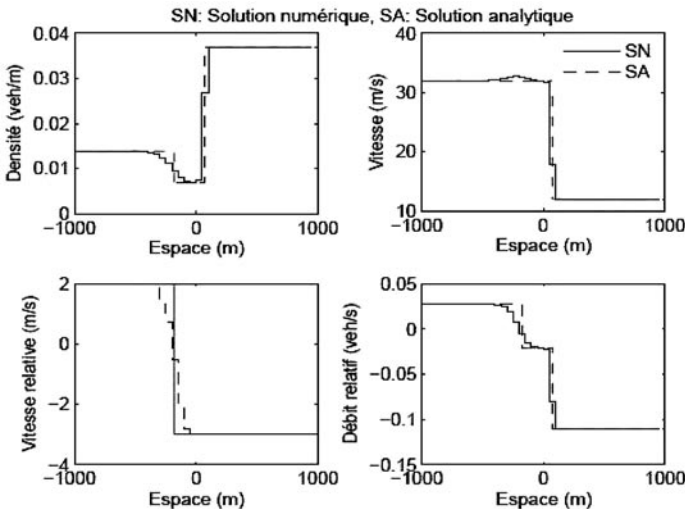
- $t = 0$

Figure 19 : interaction d'une 2-onde et d'une 1-onde, conditions initiales



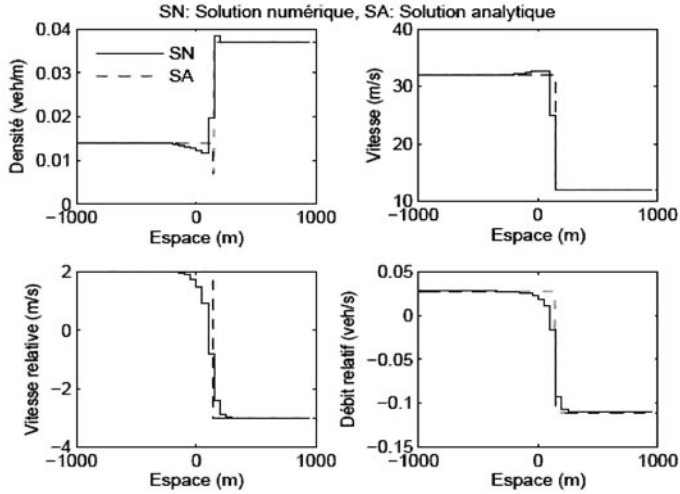
- $t = 10\text{s} < t_0 = 20.25\text{s}$

Figure 20 : interaction d'une 2-onde et d'une 1-onde, $t = 10\text{ s}$



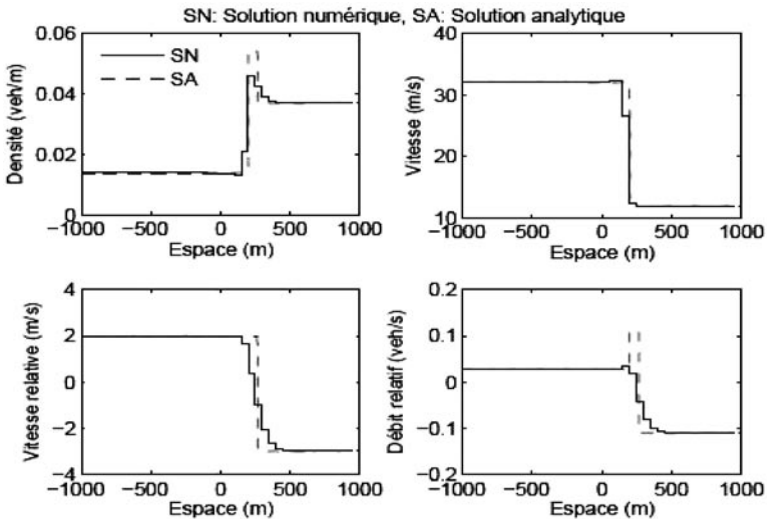
– $t = 20\text{ s} < t_0 = 20.25\text{ s}$

Figure 21 : interaction d'une 2-onde et d'une 1-onde, $t = 20\text{ s}$



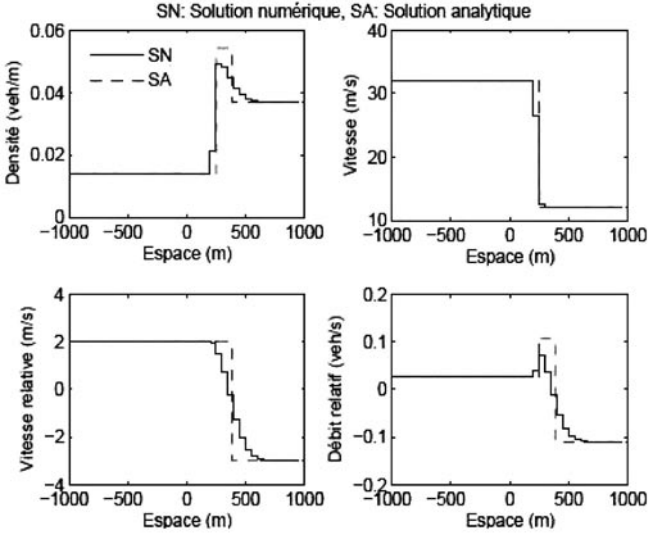
– $t = 30\text{ s} > t_0 = 20.25\text{ s}$

Figure 22 : interaction d'une 2-onde et d'une 1-onde, $t = 30\text{ s}$



- $t = 40\text{ s} > t_0 = 20.25\text{ s}$

Figure 23 : interaction d'une 2-onde et d'une 1-onde, $t = 40\text{ s}$



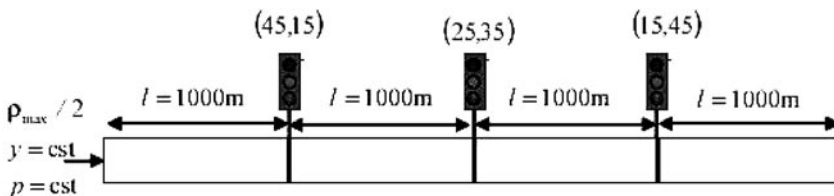
5.2. Mise en évidence de quelques phénomènes de trafic

Afin de reproduire les phases de trafic qui évoluent hors régime d'équilibre, il est nécessaire que les conditions initiales du trafic soient différentes de l'équilibre. Pour mettre en évidence ces types de phénomènes, on étudie l'évolution du trafic sur un réseau homogène (pas de changement du nombre de voies) en présence de feux tricolores avec des conditions initiales différentes du régime d'équilibre.

On considère un réseau autoroutier d'une longueur $L = 4$ kilomètres en présence de trois feux tricolores espacés d'une longueur $l = 1$ kilomètre (Figure 24). Les durées de cycle de chaque feu sont les suivantes :

- Feu 1 : durée du vert 45 secondes et durée du rouge 15 secondes.
- Feu 2 : durée du vert 25 secondes et durée du rouge 35 secondes.
- Feu 3 : durée du vert 15 secondes et durée du rouge 45 secondes.

Figure 24 : configuration géométrique du réseau



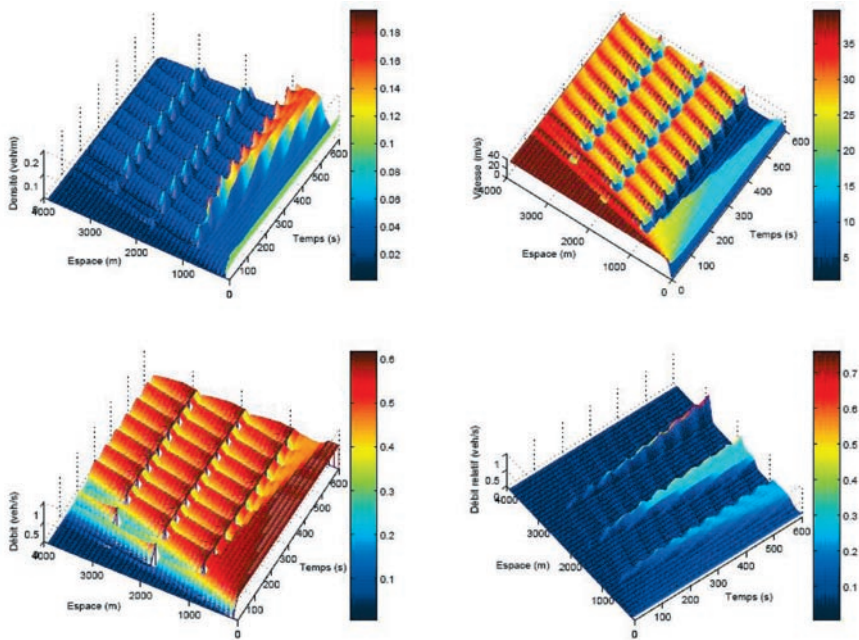
L'intérêt de ces feux est qu'ils permettent d'observer la création d'ondes de choc (ondes de décélération) quand les feux passent au rouge et des ondes de raréfaction lorsque les feux passent au vert (ondes d'accélération).

Les pas de discrétisation spatio-temporelle sont respectivement égaux à $\Delta x = 100$ m et $\Delta t = 1$ s. Dans un premier temps, on considère que le réseau est alimenté avec une demande fixe égale à $2160 \text{ véh}\cdot\text{h}^{-1}$ ($\rho = \frac{\rho_{\max}}{2}, y = \text{cst}, p = \text{cst}$).

La Figure 25 représente l'évolution de la densité, de la vitesse, du débit et du débit relatif à travers le réseau autoroutier en présence de feux tricolores.

Comme on peut le remarquer, l'évolution des variables fondamentales du modèle ARZ en présence de feux tricolores est sous la forme d'un accordéon. Cette forme est due à la présence d'ondes d'arrêt lorsque le feu tricolore est rouge, et d'ondes de redémarrage lorsque le feu tricolore passe au vert.

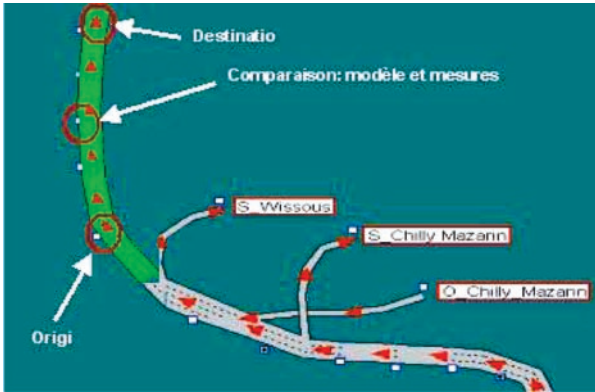
Figure 25 : évolution des variables fondamentales du modèle ARZ en présence de feux tricolores



5.3. Étude en simulation avec des données réelles

Nous allons maintenant nous intéresser à la capacité du modèle ARZ à reproduire les phénomènes de trafic en situations réelles. Pour cela, on considère une portion d'autoroute d'une longueur de 2500 mètres et comprenant cinq stations de mesures de trafic avec une période d'échantillonnage d'une minute (Figure 26). Cette présente partie d'autoroute est en particulier très fortement encombrée au

Figure 26 : configuration géométrique du site

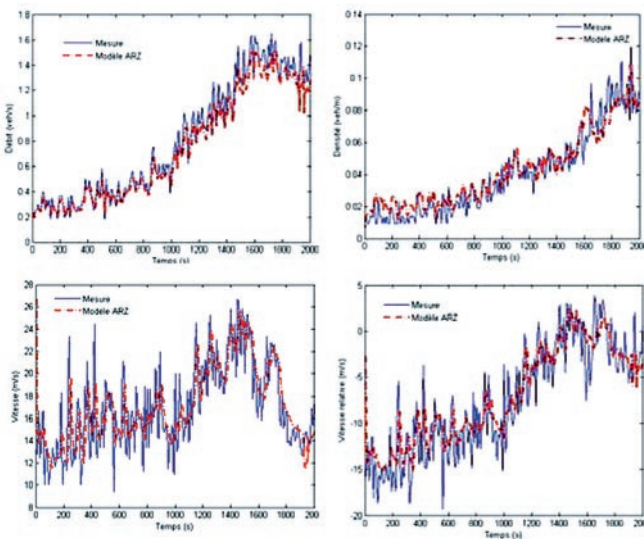


cours de la période de pointe du matin. Cette section est discrétisée en 5 segments : trois segments, un à l’origine et un à la destination.

Par ailleurs, on considère la même relation d’équilibre qui est donnée par l’équation (20) avec comme paramètre $V_{max} = 126 \text{ km h}^{-1}$, $\rho_{cr} = 35 \text{ véh/km/voie}$, $v_{cr} = 65 \text{ km h}^{-1}$ et $\tau_{max} = 58 \%$ (ces paramètres sont approximatifs et sont obtenus sans calibrage fin).

La Figure 27 représente l’évolution en fonction du temps des quatre variables de trafic : débit, densité, vitesse et vitesse relative. En dépit de la simplicité du procédé de calibrage de la relation d’équilibre, l’évolution des quatre variables du

Figure 27 : étude comparative de l’évolution des variables de trafic entre le modèle ARZ et la mesure



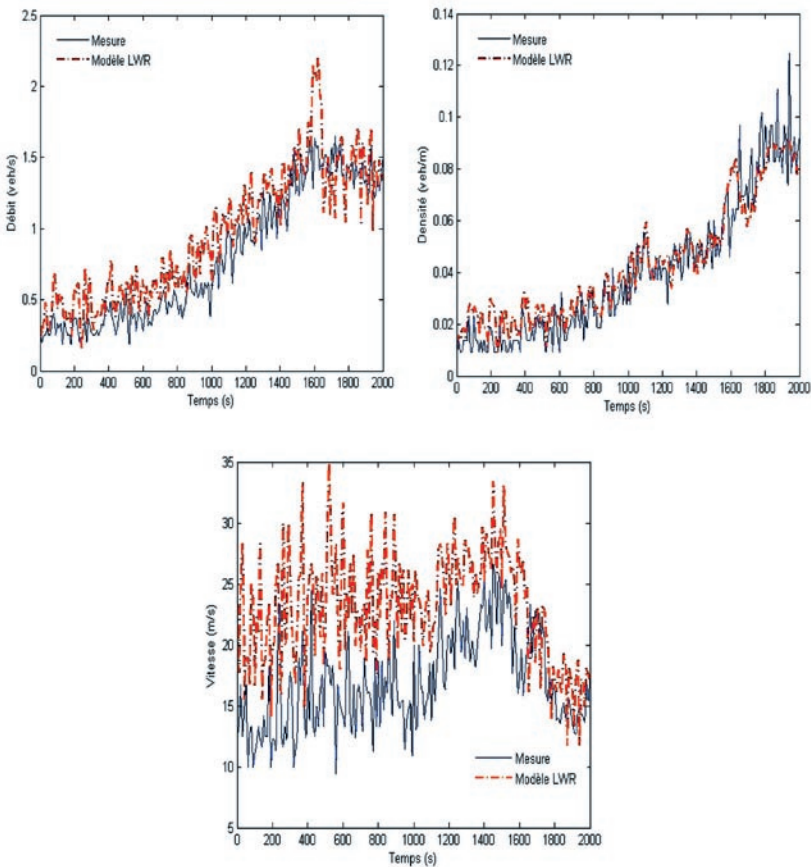
modèle est très proche de la mesure. Par ailleurs, le modèle ARZ reproduit parfaitement en espace et en temps la congestion comme c'est le cas pour la mesure.

Considérons maintenant le cas où la vitesse relative du trafic est nulle ($I = 0 \Leftrightarrow v = V_e(\rho)$). Dans ce cas présent, le modèle ARZ se réduit à une seule équation qui n'est rien d'autre que le modèle LWR du premier ordre :

$$\partial_t \rho + \partial_x (Q_e(\rho)) = 0 \quad \text{avec} \quad Q_e(\rho) = \rho V_e(\rho)$$

Comme on peut le remarquer sur la Figure 28, l'erreur entre la mesure et la simulation dans le cas du modèle LWR est très grande contrairement au modèle ARZ. Ce constat est dû au fait que le modèle LWR ne permet de reproduire que les états de trafic qui évoluent hors régime d'équilibre.

Figure 28 : étude comparative de l'évolution des variables de trafic entre le modèle LWR et la mesure



Conclusions et perspectives de recherche

Dans cet article, nous avons présenté la résolution analytique du modèle ARZ dans le cas d'une section d'autoroute homogène ainsi que la complexité qu'il pouvait y'avoir lorsque on s'intéresse au calcul de certaines solutions au problème de Riemann. Nos motivations pour une résolution exacte du modèle ARZ sont qu'elle permet une discrétisation correcte du modèle, la définition de condition aux limites pour la modélisation des intersections (Lebacque *et al.*, 2008) et le développement de modèle hybride (Mammar *et al.*, 2006).

Dans (Lebacque *et al.*, 2006), Lebacque a montré que les notions d'offre et de demande souvent utilisées dans le cas de la résolution analytique d'un modèle de premier ordre LWR peuvent être étendues à la résolution analytique du modèle ARZ dans le cas hétérogène (diagrammes fondamentaux à gauche et à droite différents).

Par ailleurs, des travaux récents menés par Lebacque (Lebacque *et al.*, 2007b) ont permis de généraliser le modèle ARZ. Cette généralisation est fondée sur la définition d'un invariant I comme une fonction de la densité et de la vitesse : $I = f(\rho, v)$.

Ce modèle est en cours d'intégration dans la plateforme multi-modèles MAGISTER qui permet des simulations suivant différents modèles de représentation du trafic.

Actuellement, plusieurs pistes sont à l'étude : extension du modèle ARZ aux multivoies et la prise en compte des différentes classes de véhicules (PL, VL, ...).

Références

- Aw, A., Rascle, M., (2000). Resurrection of Second Order Models of Traffic flow. *SIAM J. Appl. Math.*, 60(3): 916-938.
- Colombo, R., (2002). Hyperbolic phase transitions in traffic flow. *SIAM Journal of applied Math.* 63, 2, 708-721.
- Del Castillo, J-M., Pintado, P., Benitz, F.G., (1993). A Formulation of Reaction Time of Traffic Flow Models. In Daganzo, C.F. (Ed), *Proceeding of the 12th International Symposium on Transportation and Traffic Flow Theory*, Berkeley. Amsterdam: Pergamon, pp. 387-405.
- Jiang, R., Wu, Q.S., Zhu, Z.J., (2002). A new continuum model for traffic flow and numerical tests. *Transportation Research Part B*, Vol 36, 5, pp. 405-419.
- Kröner, D., (1997). Numerical schemes for conservation laws. Wiley Teubner.

- Lebacque, J.P., (1996). The Godunov scheme and what it means for first order traffic flow models. In: *Transportation and Traffic Theory, proceeding of the 13th ISTTT* (J.B. Lesort ed.). 647-677.
- Lebacque, J.P., Mammar, S., Haj-Salem, H., (2005). Second order traffic flow modelling: the Riemann problem resolution using supply/demand based approach. *Proceedings of the 10th Euro Working Group on Transportation (EWGT)*, 13-16 September, Poznan-Poland.
- Lebacque, J.P., S., Mammar, S., Haj-Salem, H., (2007a). The Aw-Rascle-Zhang model: vacuum problems, existence and regularity of the solutions of the Riemann problem. *Transportation Research Part B*, Vol 41, 7, pp. 710-721.
- Lebacque, J.P., Mammar, S., Haj-Salem, H., (2007b). Generic second order traffic flow modelling. Accepted for publication in the *Proceedings of the 17th ISTTT* (International Symposium on Transportation and Traffic Theory). London.
- Lebacque, S. Mammar, H. Haj-Salem., (2008a). Second order traffic flow modelling: the Riemann problem resolution using supply/demand based approach. Accepté pour publication dans *Special Issue of Archives of Transport*, Vol. 20, No. 1-2.
- Lebacque, S. Mammar, H. Haj-Salem., (2008b). An intersection model based on the GSOM model. 7th IFAC World Congress, July 6-11, Seoul, Korea.
- Lighthill, M.H., and Whitham, G.B., (1955). On kinematic waves II: A theory of traffic flow on long crowded roads. *Proc. Royal Soc. (London) A* 229: pp. 317-345.
- Mammar, S., Lebacque, J.P., Haj-Salem, H., (2005). Second order traffic flow modelling: the Riemann problem resolution in homogeneous case without relaxation term. *Proceedings of the 10th Euro Working Group on Transportation (EWGT)*, 13-16 September, Poznan-Poland.
- Mammar, S., Lebacque, J.P., Haj-Salem, H., (2006). A hybrid Model Based on a Second Order Traffic Model. *Transportation Research Board*.
- Mammar, S., Lebacque, J.P., Haj-Salem, H., (2008a). Riemann problem resolution and Godunov scheme for the Aw-Rascle-Zhang model. Accepté pour publication dans *Transportation Science*.
- Mammar, S., Lebacque, J.P., Haj-Salem, H., (2008). Second order traffic flow modelling: the Riemann problem resolution in homogeneous case without relaxation term. Accepté pour publication dans *Special Issue of Archives of Transport*, Vol. 20, No. 1-2.
- Payne, H.J., (1971). Models of Freeway Traffic and Control. *Simulation Councils Proc. Ser. Math. Models Public Syst.* 28(1), pp. 51-61.
- Ross, P., (1988). Traffic Dynamics. *Transportation Research Part B*, Vol 22B, n° 6, pp. 421-435.

- Richards, P.I., (1956). Shock-waves on the highway. *Opns. Res.* 4: pp. 42-51.
- Velan, S., Florian M., (2002). A note on the entropy solution of the hydrodynamic model of traffic flow. *Transportation Science* 36, 4, pp. 435-446.
- Wong, S., (1998). Multi-commodity traffic assignment by continuum approximation of network flow with variable demand, *Transportation Research Part B*, vol. 32, n° 8, pp. 567-581.
- Zhang, H.M., (1998). A Theory of Nonequilibrium Traffic Flow. *Transportation Research Part B*, Vol 32, pp. 485-498.
- Zhang, H.M., (2002). A Non-Equilibrium Traffic Model Devoid of Gas-Like Behavior. *Transportation Research Part B*, Vol 36, pp. 275-290.
- Zhang, H.M., Shu, C., Wong, G., Wong, S., (2003). A weighted essentially non oscillatory numerical scheme for a multi-class Lighthill-Whitham-Richards traffic flow model, *Journal of computational physics*, vol. 191, n° 2, pp. 639-659.

Analyse du trafic et du routage dans un réseau ad hoc de véhicules

Anthony Busson, Muriel Mabiala, Véronique Vèque

Université Paris 11

Institut d'électronique fondamentale

Bat 220, Université Paris-Sud, Centre Scientifique d'Orsay

91405 cedex

anthony.busson@u-psud.fr

Résumé

Dans cet article, nous présentons des résultats théoriques et pratiques de modélisation de réseaux de véhicules où les nœuds communicants sont des voitures équipés de carte réseau sans fil. Les propriétés du trafic routier ont un impact très important sur les performances du réseau. Nous donnons un certain nombre de résultats analytiques (taille des graphes connexes, probabilité qu'il existe un chemin entre deux véhicules) et des résultats obtenus par simulation (durée de vie d'une connexion). À partir d'un modèle réaliste du trafic routier, nous avons développé un modèle analytique des réseaux ad hoc de véhicules, dont nous dérivons les caractéristiques suivantes : taille des graphes connexes, probabilité qu'il existe un chemin entre deux véhicules ; ainsi que la durée de vie d'une connexion obtenue par simulation. Enfin, nous comparons deux protocoles de routage (DSR et GPSR) afin d'étudier leur comportement dans ce réseau fortement dynamique.

Introduction

Nous considérons dans cet article, les réseaux de véhicules. On suppose que les véhicules sont équipés d'interfaces réseaux sans fil type Wi-Fi. Deux véhicules sont alors capables de communiquer entre eux directement dès qu'ils sont à portée radio l'un de l'autre. Pour cela, ils peuvent par exemple utiliser le mode ad hoc de la technologie Wi-Fi. Ces communications peuvent être étendus dans la mesure où les véhicules peuvent servir de relais pour d'autres véhicules. Tant que

le trafic est suffisamment dense, des communications entre véhicules éloignés sont possibles.

Ces réseaux appelés VANET (*Vehicles Ad hoc NETWORK*), peuvent servir à deux familles d'applications distinctes. Des applications de sécurité routière, qui consiste par exemples à remonter des alarmes pour prévenir d'un accident, de la présence d'une plaque de verglas, du déclenchement d'un airbag ou de feux de détresses, etc. Le second type d'applications, dites orientés utilisateurs, peuvent être par exemple des logiciels de chat entre véhicules, d'échanges de fichiers, de voix sur IP, de collectes et de diffusion d'information sur le trafic ou encore d'accès à Internet au travers de point d'accès disséminer de manière éparse le long de l'autoroute.

Figure 1 : graphe représentant le réseau de véhicules

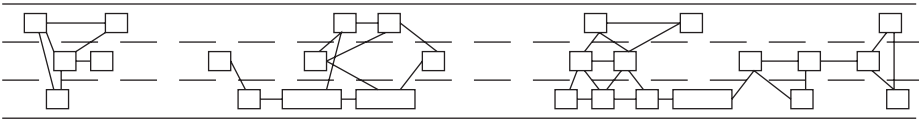
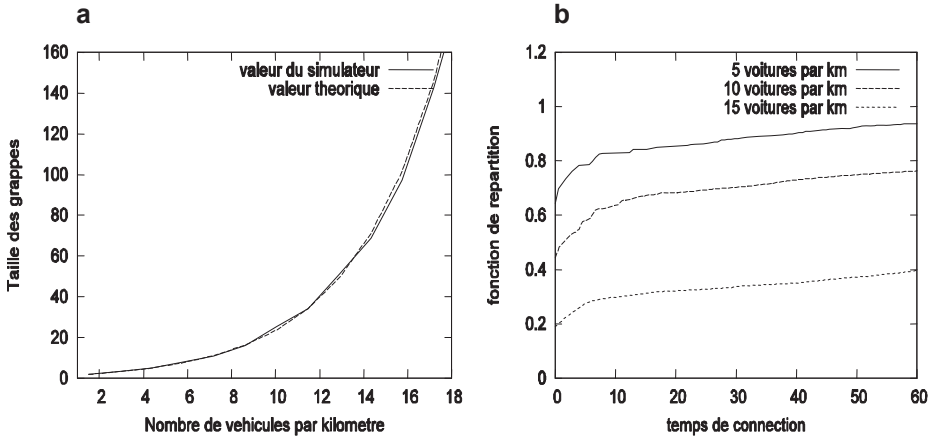


Figure 2 : a) Nombre moyen de voiture par grappe. b) Fonction de répartition durée de vie d'une connexion



Dans cet article, nous nous intéressons aux propriétés structurelles de ces réseaux et de leurs impacts sur les performances réseaux. Dans une première partie nous proposons un modèle analytique, déjà bien connu de la théorie des files d'attente, qui représente la position des véhicules à un instant donné. Cela a pour but d'évaluer le nombre de véhicules capables de communiquer entre eux, la longueur maximum des chemins (nombre de relais séparant une source et une

destination), ou la probabilité qu'il existe un chemin entre deux véhicules sachant la distance entre celles-ci. Dans la deuxième partie, nous étudions les performances des protocoles de routage dans ce type de réseau. Il s'agit de mécanisme permettant aux voitures de savoir à quel nœud envoyé les données pour atteindre une certaine destination. Afin de les comparer, nous utilisons un simulateur routier, effectuant une micro-simulation, qui émule le comportement de chaque véhicule. Ce simulateur nous permet d'obtenir des statistiques que l'on ne pouvait pas évaluer avec le modèle analytiques, comme par exemple la durée de vie des connexions, et il nous permet donc également de générer des trajectoires réalistes de véhicules pour évaluer les performances des protocoles de routage.

1. Caractéristique du réseau de véhicules

Dans cette partie, nous nous intéressons aux caractéristiques topologiques du réseau ad hoc de véhicules. Si on considère le réseau comme un graphe où les nœuds sont les véhicules et les arêtes sont les liens radio comme représenté sur la figure 1, l'une des caractéristiques de ce graphe est qu'il est très fragmenté. En effet, le graphe est constitué de « grappes » de nœuds connexes mais il n'existe pas de chemin entre ces grappes. Les communications sont alors uniquement possibles entre les voitures d'une même grappe. Des statistiques sur la répartition géographique des voitures ont montré que les positions des nœuds suivaient un processus de Poisson [3]. Ceci nous permet de déduire certaines propriétés du réseau ad hoc. Si on suppose que la portée radio des véhicules est nettement plus grande que la largeur de l'autoroute alors on peut modéliser l'ensemble des véhicules par un processus de Poisson sur la ligne. L'intensité du processus est notée λ et peut être interprétée comme le nombre moyen de véhicules par kilomètre sur l'ensemble des voies. On suppose que la portée radio des véhicules est constante et égale à $R > 0$ pour tous les véhicules. Ce modèle a été largement étudié dans [4]. On sait par exemple que la probabilité d'avoir k nœuds (où k est un entier positif) dans une grappe est égale à

$$P(C = k) = k \left(1 - e^{-\lambda R}\right)^{k-1} e^{-2\lambda R}$$

La taille moyenne d'une grappe est alors

$$E[C] = 2e^{\lambda R} - 1$$

Sur la figure 2.a, nous comparons le nombre moyen de véhicules par grappe obtenu analytiquement et avec notre simulateur de trafic (voir paragraphe 2.2). Nous obtenons des résultats très proches. Avec ce modèle, il est également possible de connaître la loi de la longueur d'un cluster ([4] théorème 2.1 page 88). Cette longueur est défini comme la distance entre les deux véhicules les plus éloignés de la grappe plus deux portées correspondant aux portées radio aux extré-

mités de cette grappe. Il est facile d'en déduire le nombre moyen de grappe de véhicules par kilomètre. On peut aussi en déduire la probabilité p que deux véhicules à distance d soient connectés, autrement dit qu'il existe un chemin entre ces deux nœuds, donc qu'ils appartiennent à la même grappe :

$$p = 1 \text{ si } d < R \text{ et } p = \int_R^d f(x) dx \text{ pour } d > R \text{ avec}$$

$$f(x) = \frac{\lambda}{e^{R\lambda} - 1} \left[1 + \sum_{j=1}^{\left\lfloor \frac{x}{R} - 1 \right\rfloor} \frac{(-1)^j}{j!} (\lambda(x - (j+1)R))^{j-1} e^{-jR\lambda} (\lambda(x - (j+1)R) + j) \right]$$

Cette distribution peut nous permettre de dimensionner un réseau hybride. Un réseau hybride est un réseau où une partie du réseau est un réseau d'infrastructure, un certain nombre de point d'accès est placé le long des routes mais en nombre insuffisant pour assurer la couverture totale de la route. L'autre partie du réseau (les véhicules) est en mode ad hoc. Pour qu'un véhicule puisse accéder à un des points d'accès, il faut qu'un certain nombre de véhicules relais les données jusqu'au point d'accès le plus proche. La formule ci-dessus peut permettre de dimensionner le réseau hybride en termes de nombre de bornes et d'écart entre elles. Les distances entre ces bornes sont calculées de manière à ce que la probabilité qu'un véhicule ne puisse accéder à une borne par le réseau ad hoc soit inférieure à une certaine valeur.

Le modèle que nous venons de présenter et d'étudier ne nous permet pas de prendre en compte la dynamique du réseau. Par exemple, il est intéressant d'évaluer la durée pendant laquelle deux voitures appartiennent à la même grappe, afin de savoir si des applications nécessitant des connexions de longue durée peuvent être envisagées dans ce type de réseau. Pour obtenir ces statistiques, nous utilisons le simulateur décrit au paragraphe 2.2. Sur la figure 2.b., nous traçons la fonction de répartition $P(\text{durée de vie} < t)$ de la durée de vie d'une connexion c'est à dire du délai durant lequel deux voitures appartiennent à la même grappe. Ces statistiques sont obtenues pour une autoroute à 3 files avec 5, 10 et 15 voitures par file et par kilomètre (en moyenne). Pour la sélection des deux voitures, nous choisissons une première voiture aléatoirement puis nous prenons la voiture qui est juste devant sur une file voisine. Bien sûr, on note que plus le trafic est dense et plus grand sera le temps durant lequel il existe un chemin entre les deux véhicules. Moins intuitivement, on voit que la distribution de cette durée a des propriétés qui semblent proches des distributions à queues lourdes. Par exemple, pour la densité la plus faible, 70 % des connexions durent moins d'une seconde. Par contre, il existe une proportion non négligeable (7 %) de connexions de plus de 1 minute et même 0,2 % de connexions qui durent plus de 15 minutes.

2. Routage dans les réseaux ad hoc de véhicules

Dans cette deuxième partie, nous nous intéressons au comportement des protocoles de routage dans ces réseaux de véhicules. Les protocoles de routage sont des mécanismes basés sur l'échange de messages entre les nœuds permettant aux nœuds de connaître le chemin que les données doivent emprunter pour atteindre leur destination. Nous comparons les deux protocoles les plus à même de fonctionner : DSR (Dynamic Source Routing protocol) et GPSR (Greedy Perimeter Stateless Routing protocol). Ils sont plus efficaces car plus réactifs, ils prennent en compte plus rapidement les changements topologiques, fréquents dans les réseaux de véhicules. Des études du même type ont déjà été effectuées dans [2] mais peu de résultats ont été publiés et un certain nombre de paramètres ne sont pas donnés ou nous semblent peu réalistes.

2.1. DSR et GPSR

Développé dans le projet Monarch à l'Université Carnegie Mellon [5], le protocole DSR (Dynamic Source Routing protocol) est un protocole simple et efficace dédié aux réseaux ad-hoc sans fil comportant des nœuds mobiles. DSR introduit les propriétés d'auto-configuration et d'auto-organisation qui permettent ni infrastructure ni administration additionnelles. Le protocole DSR est un protocole de routage pour les réseaux ad hoc. Il est de type réactif et utilise un routage à la source. Une route est construite lorsque le chemin vers une destination avec qui l'on souhaite communiquer n'est connu par l'émetteur. Il utilise alors deux mécanismes principaux : *Route Discovery* et *Route Maintenance*. La découverte de route est utilisée lorsqu'on ne possède pas de route vers le destinataire et consiste à construire une nouvelle route. La source diffuse dans le réseau une requête de recherche de route (RREQ). Lorsque le destinataire reçoit la requête, il envoie une Route Reply qui contient la liste des nœuds traversés par la requête. L'initiateur mettra sa table à jour à réception de RREPLY. Afin d'améliorer les performances, les paquets intermédiaires peuvent recopier dans leur cache la liste d'adresses contenue dans un message Route Reply ainsi que celle transportée par les paquets de données.

Le protocole de routage GPSR (*Greedy Perimeter Stateless Routing*), défini dans [6], est un protocole de routage géographique proactif qui utilise la position des routeurs en plus de la destination pour décider du prochain nœud auquel retransmettre le paquet. Par défaut, un routeur GPSR utilise le mode glouton (d'où son nom, *greedy*) et retransmet le paquet à son voisin le plus proche de la destination. La connaissance de ses voisins se fait par diffusion périodique de messages *Hello* contenant l'adresse du nœud et sa position(x,y). Dans les réseaux fortement dynamiques, la fréquence de ces *Hello*s a un impact important sur la pertinence de la liste des voisins. Dans les simulations, nous comparerons donc deux versions de GPSR avec des *Hello*s échangés toutes les 0,5 et 1,5 seconde. Lorsque le routeur n'a pas de voisins immédiats dans la direction du destinataire, il utilise le mode *perimeter* décrit dans [6].

2.2. Simulateur de trafic et simulateur réseau

Afin d'obtenir des trajectoires proches de celles des automobiles, nous utilisons un simulateur de trafic que nous avons codé en C. Ce simulateur utilise la micro simulation qui consiste, pour chaque nœud, à reproduire le comportement type d'un automobiliste. Un automobiliste aura un comportement qui dépendra de son environnement (vitesse, distance avec la voiture de devant, etc.). À chaque véhicule est associée une vitesse cible qui est la vitesse à laquelle le conducteur roulerait s'il était tout seul sur sa voie. S'il est effectivement tout seul, il adapte son accélération pour atteindre cette vitesse cible. S'il n'est pas seul, il va adapter sa vitesse à son environnement, et en particulier à la vitesse de la voiture juste devant. Enfin, si l'automobiliste n'est pas content de sa file, et si une file voisine semble lui offrir un meilleur environnement, alors il a la possibilité de changer de file. Toutes ces décisions sont des fonctions de l'environnement auxquelles on introduit une composante aléatoire permettant de modéliser le fait que deux conducteurs dans deux situations égales n'ont pas forcément les mêmes réactions. Le modèle que l'on utilise pour notre simulateur est présenté de manière très précise dans [1].

Les simulations réseaux, c'est-à-dire simulation de l'échange de données, des liens Wi-Fi, des protocoles de routages, ont été réalisées avec l'outil NS-2 (Network Simulator) dans sa version 6, qui fournit une implémentation de DSR et GPSR. Chaque véhicule transmet grâce à un lien radio 802.11 de débit égal à 2 Mbit/s avec une portée radio de 250 mètres. NS2 est un simulateur à événements discrets permettant l'étude des protocoles réseaux (filaire et mobile). Il est gratuit et son code source est libre. La couche physique utilisée est le 802.11b. Chaque nœud a une portée radio de 250 mètres. Nous avons choisi comme modèle de propagation le « *Two-Ray ground reflection model* », qui prend en compte, le chemin direct et la réflexion du sol. Nous utilisons le simulateur de trafic décrit plus haut, pour générer le mouvement des nœuds. La position des nœuds et leur vitesse sont alors introduits dans NS. L'environnement simulé est une autoroute de 14 km comportant 3 voies (un seul sens est pris en compte). Lorsque l'intensité ne varie pas, le nombre moyen de véhicules par voie est de 6. Pour établir les connexions entre les voitures, nous choisissons des voitures au hasard dans les deux premiers kilomètres de l'autoroute. Dans les figures, chaque point est obtenu comme la moyenne de 10 échantillons. Chaque échantillon correspond à une simulation d'une durée de 100 secondes. Sur l'ensemble des graphiques, les barres verticales correspondent à un intervalle de confiance à 95 %.

2.3. Comparaison de DSR et GPSR

Sur la figure 3.a., nous donnons le PDR (*Packet Delivery Ratio*) qui est défini comme le rapport du nombre de paquets reçus sur le nombre de paquets envoyés. Afin de dissocier les pertes dues à la topologie (il n'existe pas de route) et celles dues au protocole de routage, nous avons tracé sur cette même figure la probabilité (obtenue à partir du simulateur de trafic) qu'il existe une route entre deux véhicules (choisis dans les mêmes conditions que pour les simulations NS). Cette courbe représente le maximum que le PDR peut atteindre tous protocoles de

routage confondus. Il est noté PDRmax sur la figure. Il apparaît que DSR offre des performances assez proches de ce maximum et qu'il est meilleur par rapport à GPSR.

Sur la figure 3.b, nous avons tracé le PDR (*Packet Delivery Ratio*) en fonction du débit des connexions. Le PDR diminue avec le débit pour tous les protocoles. Pourtant, la somme des débits pour l'ensemble des dix connexions ne dépasse pas 160 kbit/s. Ce résultat s'explique par deux phénomènes. Nous utilisons ici des paquets très courts (de 64 octets) ce qui génère un grand nombre de paquets et diminue les performances au niveau 2 (backoff, collisions, etc.). L'étroitesse

Figure 3 : a) PDR lorsque le nombre de véhicules par voie varie ;
b) PDR lorsque le débit par voie augmente

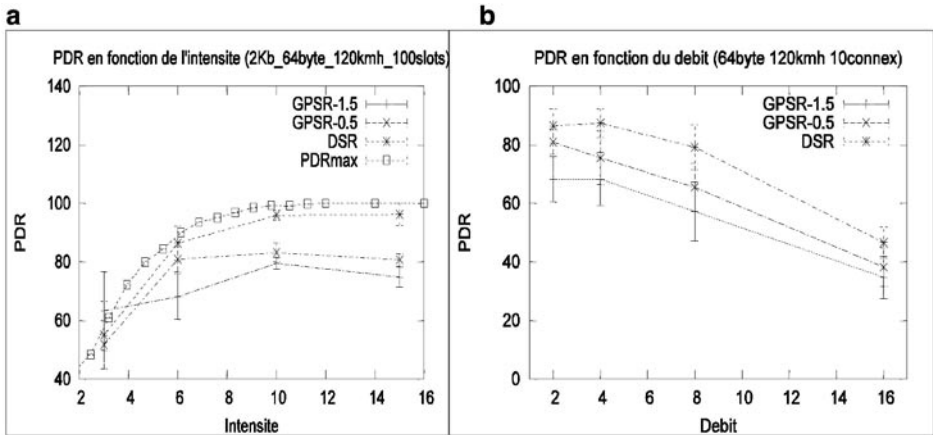
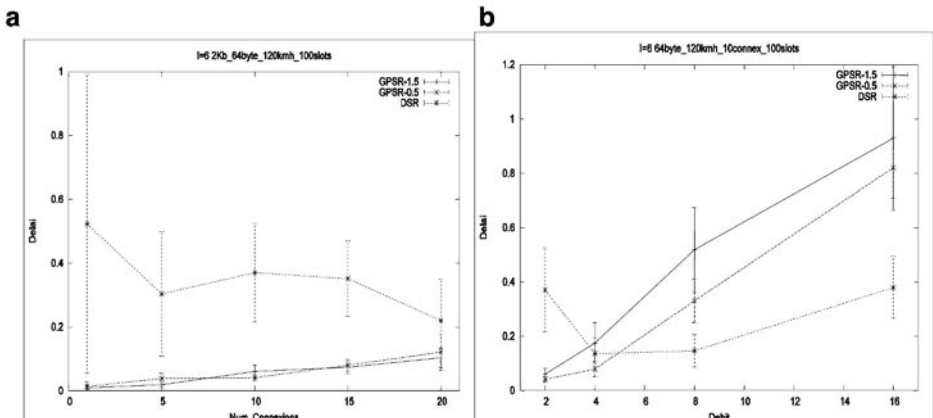


Figure 4 : a) Délai lorsque le débit pour chaque connexion varie ;
b) Délai lorsque le nombre de connexions varie



du réseau (la largeur de l'autoroute) fait que tous les noeuds entre une source et une destination reçoivent les paquets ce qui génère des goulots d'étranglements importants.

Sur les figures 4.a. et 4.b., on observe le délai moyen défini comme la différence entre le moment où le paquet est donné à la couche réseau pour émission et le moment où il est reçu par la destination. On observe qu'avec DSR le délai est plus important. Ceci est dû au délai introduit par les requêtes de découverte de chemin. Mais quand le débit augmente (figure 4.a.), la route mise en cache permet de diminuer le délai moyen. Mais ceci vaut pour une route particulière, aussi, lorsque le nombre de connexions augmente (figure 4.b.), les routes sont différentes et on ne peut bénéficier des routes mises en caches.

Conclusion

Dans cet article, nous avons présentés quelques résultats sur nos travaux sur la modélisation des réseaux de véhicules. Les réseaux de véhicules ayant une mobilité et une topologie très particulière, il est important de les modéliser ou de les simuler avec des modèles les plus réalistes possibles. Nous avons modélisé un réseau ad hoc de véhicules en appliquant des résultats existants sur les processus de couverture. D'autres statistiques, prenant en compte une composante temporelle ont été obtenues par simulation. Enfin, une comparaison de deux protocoles de routage a été effectuée. Nous avons observé un PDR (*Packet Delivery Ratio*) assez faible pour ce réseau en raison, entre autres, de la fragmentation du réseau. Mais les changements topologiques fréquents peuvent permettre malgré tout d'atteindre tous les noeuds du réseau. Nous pensons que la particularité du trafic spatial des véhicules permet d'influencer le protocole de routage et pourrait permettre de le simplifier. Nous envisageons donc de prendre en compte les spécificités de ce type de réseau afin d'élaborer des mécanismes de routage plus adaptés.

Références

- [1] Kazi I. Ahmed, *Modeling drivers'acceleration and lane change behaviour*, Doctor of science in transportation systems, Massachussetts Institute of Technology, Boston (MA), 1999.
- [2] H. Füßler, M. Mauve, H. Hartenstein, M. Käsemann, D. Vollner, *A Comparison of Routing Strategies for Vehicular Ad Hoc Networks*, Tech. Report TR-02-003, Department of Computer Science, University of Mannheim, July 2002.
- [3] D.C. Gazis, *Traffic theory*, Kluwer Academic Publishers, 2002.

- [4] P. Hall, Introduction to the theory of coverage processes, John Wiley and Sons, 1998.
- [5] D.B. Johnson, D.A. Maltz, Yih-Chun Hu, *The Dynamic Source Routing protocol for mobile ad hoc networks (DSR)*, July 2004, Internet-Draft, adraft-ietf-manet-dsr-10.txt.
- [6] B. Karp, H.T. Kung, *GPSR : Greedy Perimeter Stateless Routing for wireless networks*, Proc of 6th ACM/IEEE Int. Conf. On Mobile Computing and Networking (MobiCom 2000), 2000, pp. 243-254.

Ajustement et optimisation des paramètres de poursuite de Newell

Aurélien DURET, Christine BUISSON, Nicolas CHIABAUT

*Laboratoire d'Ingénierie Circulation Transports LICIT – ENTPE / INRETS –
Université de Lyon, Vaulx-en-Velin, France*

La version anglophone de cet article a été publiée dans *Transportation Research Record* 2008, journal du 87^e Transportation Research Board (01/2008)

Résumé

L'objet de cette présentation est d'introduire deux méthodes de calibrage individuel du modèle de poursuite simplifié de Newell en régime congestionné à partir de données de trajectoires. En plus de calibrer le modèle de poursuite, ces deux méthodes permettent par la même occasion de quantifier les distributions de ces paramètres au sein du flux. Ces quantifications seront commentées à la lumière du sens physique de chacun de ces paramètres.

Les données de trajectoires utilisées dans cette étude sont issues de NGSIM (ngsim.fhwa.dot.gov). Elles fournissent tous les dixièmes de seconde la position de l'ensemble des véhicules présents sur les 500 mètres d'une section autoroutière californienne à six voies.

La première méthode de calibrage quantifie le diagramme fondamental à l'échelle individuelle. Elle estime les deux paramètres d'entrée du modèle de Newell (vitesse de remontée de congestion et distance à l'arrêt) sous les hypothèses d'un diagramme fondamental individuel triangulaire.

La seconde méthode consiste à trouver le couple de paramètres de poursuite qui minimise la différence entre la simulation et l'observation.

Introduction

Le diagramme fondamental est une description macroscopique des états d'équilibre du trafic. Son allure a longtemps été débattue mais de récentes études expérimentales [1-4] ont montré que le diagramme fondamental est triangulaire

sur voies rapides (inter urbaines ou autoroutières). Ainsi, la relation débit-densité est linéaire en congestion et ne dépend que de deux paramètres : la vitesse maximale de remontée de congestion w et la densité maximale k_x .

Par ailleurs, un flot de véhicules est composé de différents conducteurs présentant potentiellement des caractéristiques de conduite différentes. En guise d'illustration, se référer à (5) qui présente une étude empirique montrant que les conducteurs ont leur personnalité propre. Par conséquent, le trafic ne peut pas être considéré comme strictement homogène, et pour cette raison, cette étude propose une quantification individuelle des paramètres de poursuite.

Grâce aux données particulièrement détaillées diffusées librement par NGSIM (6) et qui contiennent les trajectoires de véhicule circulant sur l'I80 durant l'heure de pointe le soir d'une semaine d'avril 2005, nous avons analysé la relation individuelle vitesse-espacement (nommée *ssr* pour *speed-spacing relationship* dans la suite de l'étude). L'objectif est d'observer et de quantifier l'hétérogénéité du comportement de trafic en congestion.

Pour cela, deux méthodes sont employées. Elles assument toutes les deux le fait que la *ssr* est une représentation individuelle du diagramme fondamental. Ainsi, la *ssr* est également linéaire en congestion et ne dépendra que de deux paramètres individuels : w_i and $k_{x,i}$, équivalents respectifs de w et k_x .

La première méthode estime ces paramètres individuels par le biais d'une régression linéaire par la méthode des moindres carrés sur le nuage de points (vitesse, espacement) calculés le long des trajectoires de véhicules en poursuite. La seconde méthode consiste à optimiser le modèle de poursuite en trouvant le couple de paramètres en minimisant la distance euclidienne entre la trajectoire simulée et la trajectoire observée d'un véhicule en poursuite.

Après avoir introduit la notion de diagramme fondamental, les sections 2 et 3 présentent respectivement les méthodes d'estimation et d'optimisation. Enfin, la dernière partie discute des différences entre ces deux méthodes aux fonctions-objectif différentes.

1. Background

1.1. Diagramme fondamental

Le diagramme fondamental est une relation d'équilibre qui lie le flux de trafic q à sa densité k . De manière générale, elle est étudiée sur des données relevées en un point de la section d'étude et agrégées sur une certaine période (en général 6 minutes). Deux problèmes se posent alors : (i) contrairement au débit et à la vitesse, la concentration ne peut pas se mesurer ponctuellement (ii) l'agrégation temporelle n'est pas une mesure de l'équilibre puisqu'elle agrège elle-même plusieurs états d'équilibre. (7) présente différents résultats de données expérimentales agrégées.

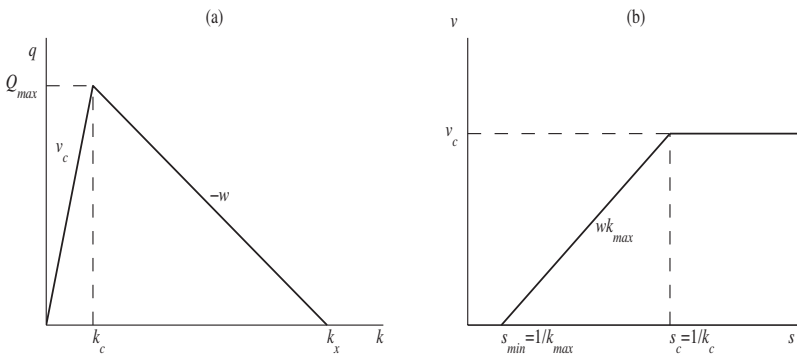
En situation fluide, le diagramme fondamental est linéaire et le débit q est proportionnel à la concentration k ($q = v_f k$, où v_f est la vitesse libre du flux). En

régime congestionné, le diagramme est également linéaire mais décroissant (1-4) et la pente négative w correspond alors à la vitesse de remontée de congestion. La forme triangulaire du diagramme fondamental est illustrée sur la figure 1.a.

Un diagramme équivalent peut être défini en utilisant la relation entre le débit q , la concentration k et la vitesse v ($q = k \cdot v$) ainsi que la définition de l'espacement ($s = 1/k$). La forme triangulaire de la relation (q, k) est alors équivalente à la forme bi-linéaire de la relation (v, s) décrite sur la figure 1.b. Sur cette dernière, la vitesse croît linéairement avec l'espacement entre $s_c = 1/k_c$ et $s_x = 1/k_x$. s_x correspond à l'espacement minimum (à l'arrêt) et s_c est appelé espacement critique, au-delà duquel la vitesse du véhicule est indépendante de l'espacement (situation fluide).

Nous ne nous intéresserons qu'à la partie congestionnée du diagramme fondamental, partie pour laquelle vitesse et espacement sont liés.

**Figure 1 : diagramme fondamental triangulaire ;
a) relation débit-concentration ; b) relation vitesse espacement**



1.2. Filtre global : sélection des véhicules en poursuite stable

Les données individuelles utilisées ici sont issues du programme de recherche NGSIM (6). Nous les utilisons ici pour le calibrage du modèle de poursuite de Newell. Ces données contiennent l'ensemble des positions des véhicules toutes les 0,1 s sur une section d'étude de 500 mètres.

La première étape consiste à identifier les véhicules en poursuite stable. En effet, certains phénomènes tels que le changement de voie ou le changement de leader peuvent amener un véhicule hors équilibre. Pour cela, l'étude se concentrera uniquement sur les véhicules ne changeant pas de voie et conservant le même leader sur les 500 mètres de la section d'étude.

Nous nous affranchissons alors des effets de changement de voie et nous permettons pour plus tard une analyse des résultats voie par voie. Par ailleurs, la voie 1 est une voie HOV, i.e. sur laquelle ne sont autorisés à circuler que les véhicules transportant au moins deux personnes. La vitesse moyenne de circulation sur cette voie est de 46 km/h, contre 20 km/h sur les autres voies. Elle ne

présente pas de congestion assez dure pour garantir que les paires de véhicules retenues soient en poursuite et sera donc exclue de notre étude. À cette étape, nous disposons alors d'un échantillon de 1116 paires de véhicules.

Même si les vitesses individuelles instantanées sont disponibles dans les jeux de données, des erreurs de mesures ont été observées (dues au manque de précision des données). Pour cette raison, les vitesses instantanées ont été recalculées à chaque pas de temps, par une moyenne glissante dont la largeur de la fenêtre est de 1,2 s. Cette période d'agrégation semble être un bon compromis puisqu'elle permet d'atténuer le bruit initialement du aux erreurs de mesure tout en conservant la représentativité de la variation de la vitesse le long de la trajectoire des véhicules étudiés. La figure 2 illustre le type de nuage de point (v, s) obtenu par des données traitées.

2. Méthode d'estimation

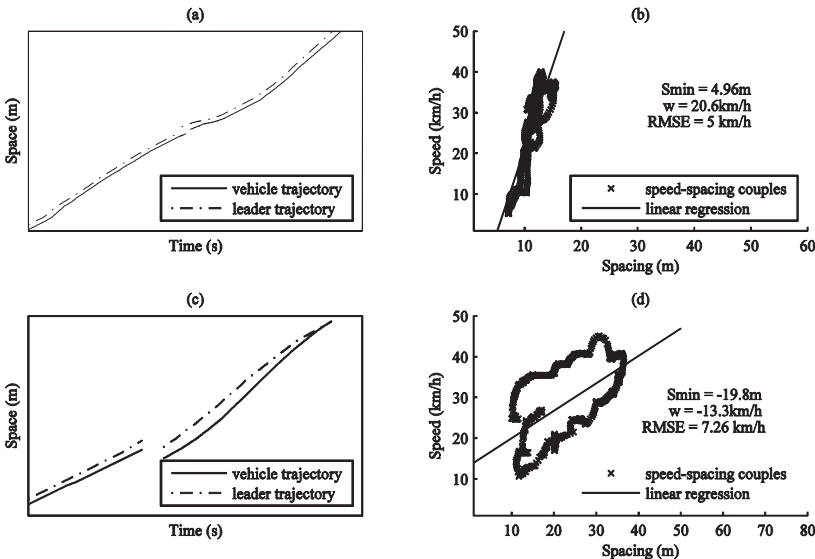
2.1. Présentation de la méthode

La méthode consiste à effectuer une régression linéaire au sens des moindres carrés sur le nuage de point (v, s) d'un couple de véhicule donc le suiveur est en poursuite stable. Ainsi, la fonction-objectif est la suivante :

$$RMSE_i = \sqrt{\frac{1}{\#s_{obs,i}} \sum_{s \in S_{obs,i}} (V_{est,i}(s) - V_{obs,i}(s))^2}$$

où $\#s_{obs,i}$ est le nombre d'espacement mesuré pour le véhicule i .

Figure 2 : a) et c) trajectoires d'une paire de véhicule ; b) et d) sss respective

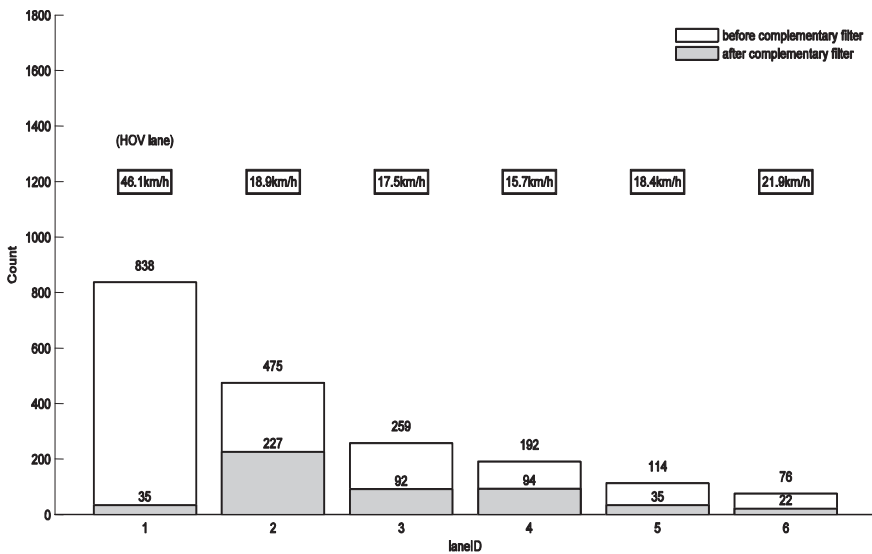


La figure 2 donne les résultats obtenus pour deux véhicules : (a) pour un véhicule ordinaire en situation de congestion, (c) pour un véhicule pouvant accepter des espacements très différents pour une même vitesse, et (b) et (d) leur ssr respective.

2.2. Filtre complémentaire pour la méthode d'estimation

Les véhicules présentant des valeurs d'estimation inconsistantes pour \hat{s}_x et \hat{w} sont ensuite exclus de l'étude. En effet, le sens physique de ces deux paramètres leur interdit les valeurs négatives d'une part, et de prendre des valeurs trop élevées d'autre part. C'est ainsi que les véhicules dont la valeur de \hat{s}_x est supérieure à 12 m sont exclus, ainsi que ceux présentant un \hat{w} supérieur à 36 km/h. Par conséquent, l'échantillon est réduit à 470 véhicules, soit 42 % de l'échantillon initial. La figure 3 présente les échantillons initiaux et finaux par voie (la voie 1 est uniquement donnée à titre d'information).

Figure 3 : échantillonnage avant et après le filtrage complémentaire

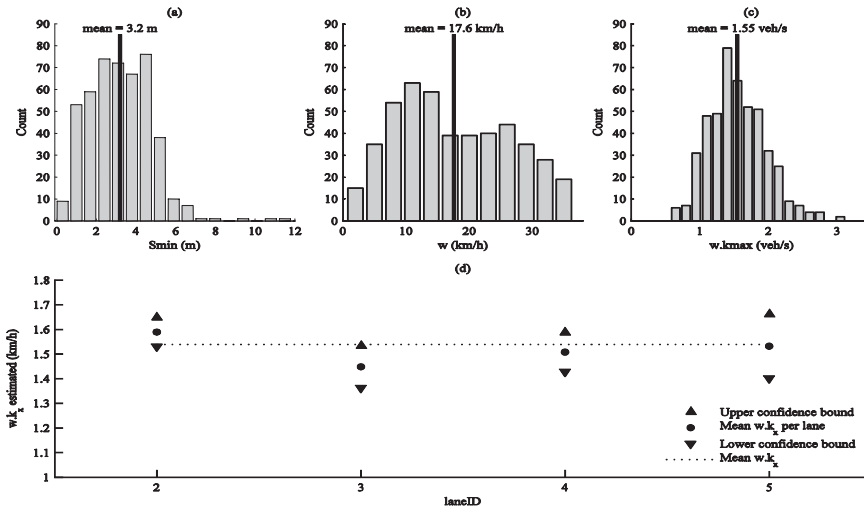


2.3. Résultats de l'estimation

Les distributions de \hat{s}_x et \hat{w} sont respectivement présentées sur les figures 4.a et 4.b.

La distribution de \hat{s}_x est relativement concentrée autour d'une valeur proche de 3,2 m, alors qu'au contraire, la distribution de \hat{w} est plus éparse autour d'une valeur moyenne de 17,6 km/h, correspondant à l'intervalle couramment cité de [15 km/h ; 20 km/h] (4, 5).

Figure 4 : histogramme des estimations de a) s_x ; b) w ; c) r ; d) intervalle de confiance sur r



Le test de Kolmogorov-Smirnov ne rejette pas l'hypothèse de distributions différentes d'une voie à une autre (à un niveau de confiance de 5 %). En revanche, le test de Jarque-Bera rejette le fait que les distributions soient issues d'une distribution normale.

Un autre paramètre couramment utilisé en théorie du trafic est le débit de dépassement. Il correspond au débit croisant un observateur fictif se déplaçant de

l'aval vers l'amont à une vitesse w (3). Il s'exprime sous la forme : $r = w.k_x = \frac{w}{s_x}$.

La figure 4.c présente la distribution des débits de dépassement estimés. Le test de Jarque-Bera ne rejette pas l'hypothèse stipulant que les distributions de r sont issues d'une distribution normale de moyenne $\mu = 1,5$ véh/s et d'écart-type $\sigma = 0,42$ véh/s. Notons que cette méthode tend à sur-estimer le débit de dépassement puisque sa valeur courante est plutôt comprise entre 0,6 et 0,9 véh/s. Notons également que l'écart-type de la distribution du débit de dépassement est relativement important ici, supportant l'hypothèse d'une certaine dispersion des comportements de poursuite au sein d'une flotte de véhicules. Cette hypothèse avait déjà été supportée par (5).

Les intervalles de confiance à 95 % sur les valeurs moyennes des estimations de r pour les voies 2 à 5. La voie 6 a été exclue de l'analyse parce qu'elle ne présente pas un effectif statistiquement suffisamment grand (22 véhicules). Pour les 4 voies restantes : d'une part les voies 2 et 3 présentent respectivement les moyennes les plus haute (1,59 véh/s) et plus basse (1,45 véh/s) et leurs intervalles de confiance sont strictement distincts, ce qui suggère de ne pas agréger les

résultats issus des voies 2 et 3 à moins de faire une étude à un degré de précision supérieur à 5 % ; d'autre part, nous pouvons remarquer ici que les valeurs moyennes (ainsi que les écart-type) obtenues sur les voies 4 et 5 sont relativement proches et peuvent éventuellement être agrégées.

2.4. Limites de la méthode d'estimation

Cette méthode, même si elle présente l'avantage de la simplicité et de la vitesse de calcul, a deux défauts majeurs :

- D'une part \hat{s}_x et \hat{w} présentent des valeurs inconsistantes pour une proportion importante de l'échantillon initial (plus de 57 %) ;
- D'autre part, la valeur moyenne de \hat{s}_x est bien trop faible (3,2 m) puisque inférieur à la longueur moyenne des véhicules étudiés.

Dans le cadre de la validation du modèle de Newell, la partie suivante présente une seconde méthode basée sur la simulation.

3. Méthode d'optimisation des paramètres du modèle de Newell

3.1. Le modèle de poursuite

Le modèle de poursuite de Newell, en congestion, consiste à calculer la position du véhicule suiveur en fonction de sa position actuelle, de la position actuelle de son leader ainsi que la vitesse actuelle de son prédécesseur. Le modèle s'écrit alors :

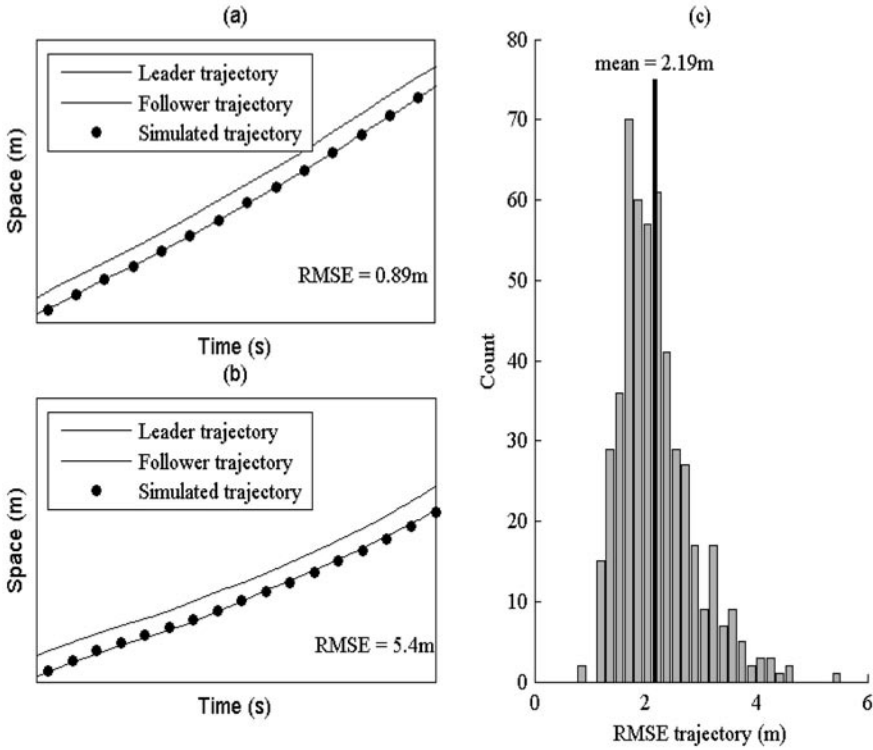
$$x_f(t + \Delta t) = x_f(t) + \Delta t \cdot V_f^{eq} (x_l(t) - x_f(t))$$

où x_f est la position du véhicule en poursuite, x_l la position du véhicule leader, Δt le pas de temps et V_f^{eq} la SSR du véhicule suiveur. (9) est l'unique papier présentant une validation macroscopique du modèle de Newell (à la connaissance des auteurs au moment de cette étude).

Ainsi, la méthode d'optimisation consiste à appliquer le modèle de Newell et trouver le couple (s_{min}^*, w^*) minimisant la distance euclidienne entre la trajectoire simulée et la trajectoire observée. Cette distance se mesure suivante la formule :

$$RMSE_i = \sqrt{\frac{1}{\#(t_{obs})} \sum_{t \in} (x_{obs}(i, t) - x_{sim}(i, t))^2}$$

Figure 5 : a) Meilleur résultat de simulation ; b) pire résultat de simulation ; c) histogramme du RMSE



Pour atteindre le couple optimal, L'ensemble des couples de valeurs sont testés sur le pavé numérique définit sur $0 < s_{\min} < 12$ m et $0 < w < 36$ km/h. La solution optimale est celle qui minimise la fonction de coût présentée ci-dessus. Ce processus a été appliqué à l'ensemble des 1116 véhicules issus du filtrage initial.

Tableau 1 : s^* , w^* et r^* par voie

Voie	s_{\min}^* (m)	w^* (km/h)	$w \cdot k_x^*$ (véh/s)	Nombre de véhicules
2	6,48	19,1	0,82	472
3	6,86	17,5	0,71	271
4	6,65	19,9	0,83	189
5	6,32	19,5	0,86	111
6	4,83	13,9	0,80	73
2 à 6	6,57	18,7	0,79	1116

Sur la voie 2, la moyenne de s_x^* est de 6,5 m et la plupart des valeurs sont incluses dans l'intervalle [5,5 m ; 8 m]. Sur la voie 3, la moyenne est un peu plus élevée (6,9 m) et la distribution légèrement plus éparse. Notons également que la distribution de w^* est largement étalée autour de sa valeur moyenne égale à 18,7 km/h. Le coût optimal moyen obtenu sur l'ensemble des 1116 véhicules est égal à 2,6 m. En excluant les 5 % donnant les plus mauvais résultats, ce coût minimal descend à 2,3 m (voir tableau 2).

Tableau 2 : résumé des résultats d'optimisation

Total	% kept	S_{\min}^* (m)	W^* (km/h)	$W.K_X^*$ (véh/s)	Mean RMSE (m)
1116	100 %	6,5	18,6	0,79	2,7
1061	95 %	6,5	19	0,81	2,3

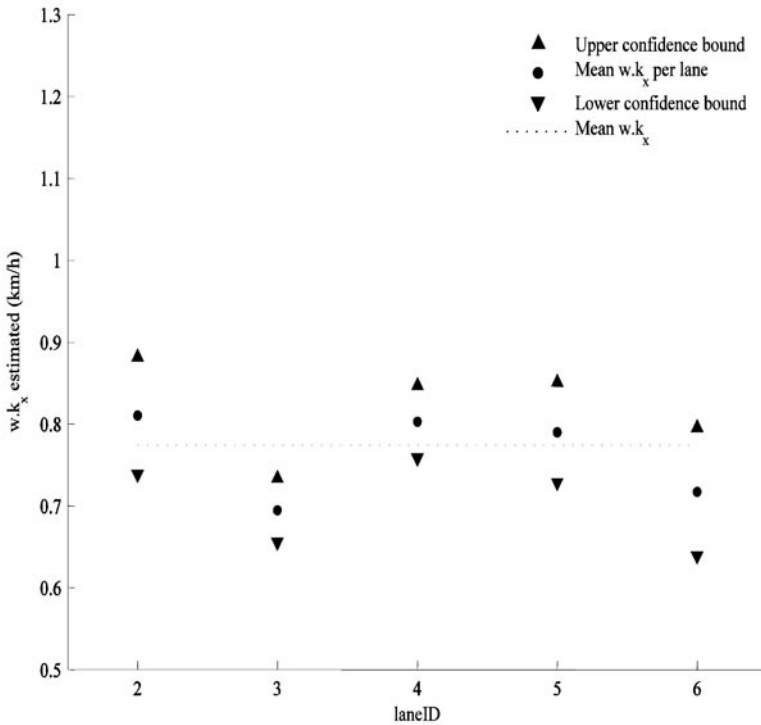
Finalement, même si la méthode d'estimation proposée précédemment ne mène pas au coût minimal du modèle de Newell, il en reste assez proche du minimum possible. Par ailleurs, les différences entre paramètres estimés et paramètres optimaux seront discutées dans la dernière partie du papier.

Une analyse voie par voie a été faite sur les résultats optimaux et la figure 6 présente les intervalles de confiance de r^* à 95 % pour les voies 2 à 6 (cette fois-ci, l'effectif est suffisant sur la voie 6 pour l'inclure dans l'analyse statistique). La valeur moyenne pour les cinq voies est représentée par la ligne horizontale pointillée.

Tout d'abord, contrairement à la méthode d'estimation, la moyenne du débit de dépassement optimal est proche de 0,8 véh/s et est donc en accord avec les valeurs couramment rencontrées dans la littérature. De plus, les deux méthodes (estimation et optimisation) s'accorde sur le fait que les intervalles de confiance du débit de dépassement sur les voies 2, 4, 5 et 6 sur distinct de celui de la voie 3. Cela confirme le fait qu'il est nécessaire de différencier les voies les unes des autres lors de l'attribution des paramètres de poursuite.

Le débit de dépassement optimal de la voie 2 (0,81 véh/s) est égal à la valeur calculée en utilisant la méthode d'évaluation du diagramme fondamental en congestion par voie proposée dans (4). Le débit de dépassement de la voie 3 est légèrement plus faible et égale à 0,69 véh/s. Les voies 4 et 5 montrent que le moyenne et écart-type sont très proches pour ces deux voies. La voie 6, pour laquelle le flot est issu de la voie principale et de la voie d'accélération, montre que le débit de dépassement optimal est égal à 0,71 véh/s.

Figure 6 : intervalle de confiance du débit de dépassement optimal (à 95 %), par voie



3.2. Résultats pour les paramètres moyens

Comme précisé précédemment, une autre étude a été menée dans le but d'évaluer s_x^* et w^* à l'échelle de la voie pour les voies 2 et 3 des mêmes échantillons étudiés ici. Leurs investigations mènent au couple ($s_x = 7,7$ m, $w = 18,9$ km/h) pour la voie 2 et ($s_x = 7,2$ m, $w = 14,5$ km/h) pour la voie 3. Ces quantifications sont proches des moyennes obtenues par la méthode d'optimisation présentée ici. Ces deux couples de valeurs ont ensuite été implémentés dans le modèle de Newell pour calculer la différence euclidienne moyenne entre les trajectoires simulées avec ces deux couples et les trajectoires observées. Le coût moyen obtenu est proche de 4 m. Comparé au coût moyen obtenu en implémentant individuellement les paramètres (2,7 m), nous sommes confortés dans l'idée qu'il est important de tenir compte de la dispersion des paramètres au sein du flot de véhicules.

4. Comparaison entre les deux méthodes

4.1. Implémentation des paramètres d'estimation dans le modèle de poursuite

En implémentant les couples de paramètres individuels estimés par la première méthode (470 véhicules) dans le modèle de Newell, nous pouvons comparer le coût des résultats issus de la première méthode au coût de la méthode d'optimisation. La figure 5 montre les résultats empiriques des meilleur et pire résultats obtenus.

Le RMSE moyen obtenu pour les 470 véhicules est égal à 2,2 m. Ces bons résultats obtenus par le modèle de Newell avec les paramètres estimés sont principalement dus à la robustesse du modèle vis-à-vis des deux paramètres d'entrée : en effet, le modèle semble avoir une faible sensibilité autour de la valeur estimée (nous approfondirons ce point lors de la dernière partie de l'étude).

4.2. Différences quantitatives entre paramètres estimés et paramètres optimaux

Il est intéressant de se pencher désormais sur les différences quantitatives entre les paramètres estimés et les paramètres optimaux.

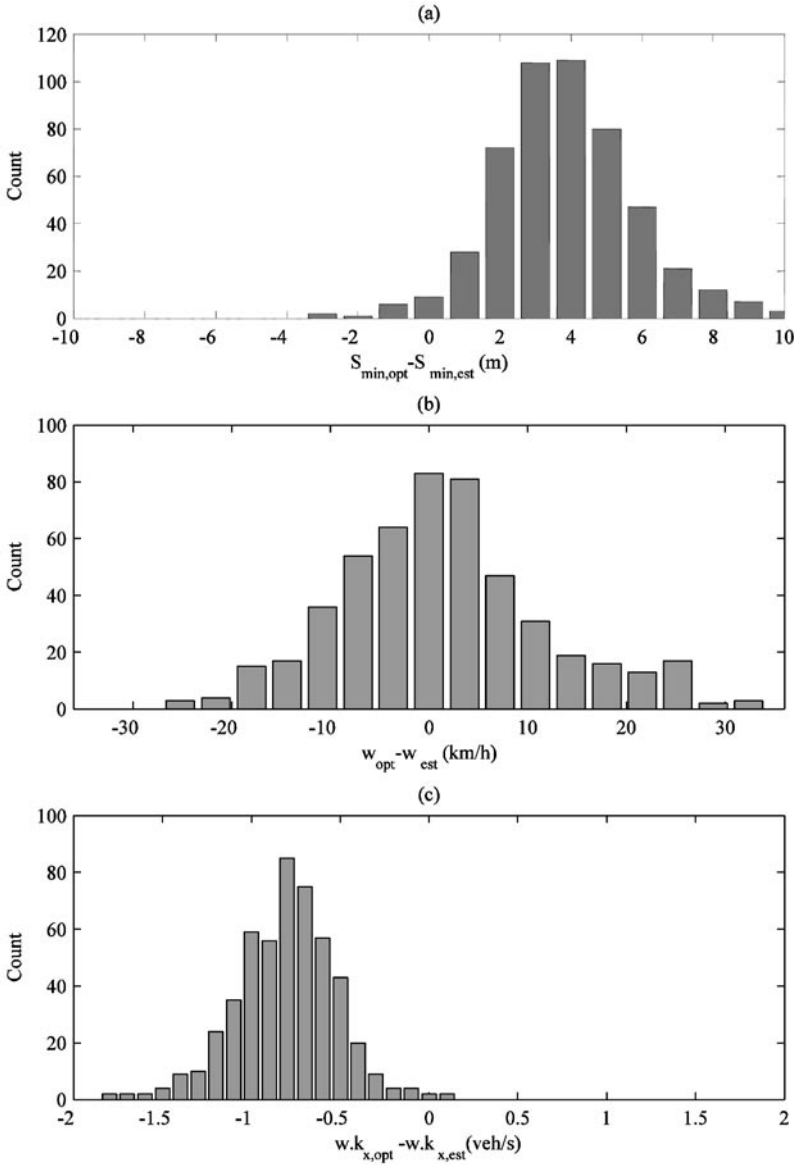
La figure 7 représente les histogrammes de ces différences pour les paramètres s_x , w et r . Sur la figure 4.a, nous avons vu que la méthode d'estimation tend à sous-estimer la valeur de s_x : la figure 7.a confirme cette tendance puisque la différence moyenne entre l'estimation et l'optimisation de ce paramètre est d'environ trois mètres.

Le débit de dépassement est surestimé par la méthode d'estimation de part la sous-estimation de s_x ($r = w/s_x$). Concernant w , la distribution de la différence est centrée sur 0 et aucune tendance particulière ne ressort ici.

Pour autant, la distribution des différences pour entre \hat{w} et w^* montre que la distribution des écarts semble assez éparse (voir figure 7.c). On trouve l'explication de cet étalement de la distribution des différences en observant la forme de la nappe représentant la fonction de coût pour chaque véhicule (voir figure 8). En effet, cette nappe fait apparaître une vallée qui indique qu'il existe un certain nombre de couples de paramètres (s_x , w) menant à un coût proche du coût minimal. Cela explique sans doute la différence quantitative importante parfois trouvée entre l'estimation et l'optimisation. Cela peut également expliquer l'effet de sous-estimation de s_x par la méthode d'estimation.

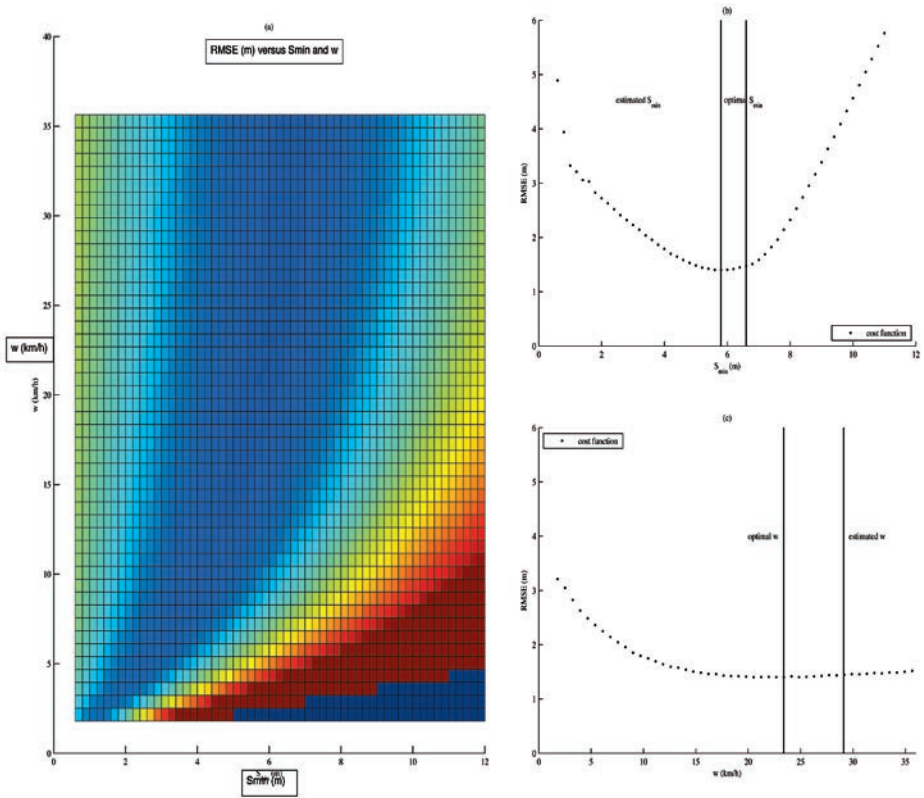
Sur l'exemple de la figure 8, nous pouvons constater que, malgré le fait que l'estimation de s_x soit du même ordre de grandeur que le s_x optimal, ils ne sont pas égaux. Mais pour autant, au regard de l'allure du fond de vallée, nous constatons que ces deux valeurs mènent à un RMSE sur la trajectoire relativement proche de l'optimal (1,6 m contre 1,45 m pour l'optimal).

Figure 7 : histogramme des différences entre paramètres estimés et optimaux



Ceci peut être vu comme un défaut de nos indicateurs qui semblent peu sensibles autour du couple optimal de paramètre. Cela signifie donc que ces méthodes sont bonnes pour calibrer un modèle et obtenir de bon résultats en simulation, mais que leur efficacité est plus relative pour quantifier des paramètres caractéris-

Figure 8 : a) illustration d'une fonction de coût ; b) projection du fond de vallée sur l'axe s_x ; c) projection du fond de vallée sur l'axe w



tiques du trafic. Pour cela, d'autres méthodes de quantification plus sensibles autour de la valeur optimale ont été mises à jour (3).

Discussion

Cet article étudie deux méthodes simples pouvant facilement être appliquées pour quantifier les paramètres de congestion à une échelle microscopique. La première méthode, qui évalue les deux paramètres à partir de l'évolution spatio-temporelle de la ssr , a l'avantage de la simplicité et de la rapidité, et semble donner des valeurs raisonnables de w . Pour autant, elle semble largement sous-estimer s_x pour un nombre important de véhicules identifiés en poursuite dans les jeux de données I80 issus du programme NGSIM. Les résultats de cette méthode permettent malgré tout d'obtenir une bonne reproduction des trajectoires individuelles lorsque ces paramètres estimés sont implémentés en entrée du modèle de poursuite de Newell : ils mènent à un coût de reproduction de trajectoire moyen

de 2,6 m. La seconde méthode de quantification optimise la reproduction de trajectoire par le modèle de Newell. Cette méthode est plus lourde en implémentation et en temps de calcul, mais elle permet une reproduction de trajectoire avec une erreur moyenne de 2,3 m pour 95 % de l'échantillon d'étude. Enfin, ce travail met en exergue la présence de forte variabilité intra voie, ainsi que des comportements de poursuite distincts d'une voie à une autre.

En dépit des différences quantitatives observées entre les estimations et les optimisations, la perte de précision sur la reproduction de trajectoire est raisonnable. En effet, au regard de la variation du RMSE avec s_x , il apparaît que beaucoup de couple de paramètres mènent à un coût proche du coût minimal. Ainsi, tant que le couple de paramètre est localisé dans la vallée de la fonction de coût, la trajectoire peut être reproduite relativement fidèlement (à 2,5 m près).

Pour la suite, la principale direction restant à explorer concerne le processus de validation de ce modèle. En effet, même si les résultats obtenus sur la quantification individuelle des paramètres sont satisfaisants et que la reproduction de trajectoire est fidèle à l'observation, le papier met en lumière le fait qu'un simple calibrage ne mène pas toujours à des résultats corrects de part la fait que la section d'étude n'est parfois pas suffisamment longue pour que la trajectoire du véhicule traverse une onde de choc (condition nécessaire et suffisante à l'estimation des paramètres de poursuite).

Le travail présenté s'inscrit dans un cadre de travail plus global. Il est défini par l'hypothèse de l'influence de la variabilité et de l'inhomogénéité sur la qualité de l'écoulement de trafic. Il requiert donc l'identification d'autre variabilité caractéristique du trafic ainsi que des méthodes de quantification telles de celles présentées ici. Ces méthodes devront, en outre, être applicables sur des données issues de relevés de données classiques (tels que des boucles).

Références

- M. J. Cassidy, M. Mauch, An observed traffic pattern in long freeway queues, *Transportation Research Part A: Policy and Practice*, 35 (2) 2001, pp. 143-156.
- J. R. Windover, M. J. Cassidy, Some observed details of freeway traffic evolution, *Transportation Research Part A: Policy and Practice*, 35 (10) 2001, pp. 881-894.
- N. Chiabaut, Ch. Buisson, L. Leclercq, Fundamental Diagram Estimation through Passing Rate Measurements in Congestion, submitted to IEEE. *Transactions on Intelligent Transportation Systems*.
- B. A. Coifman, Y. Wang, *Average velocity of waves propagating through congested freeway traffic*, in 16th International Symposium on Transportation and Traffic Theory, H. S. Mahmassani, Ed. Maryland, USA: Elsevier, 2005, pp. 165-179.

- M. J. Cassidy, J. R. Windover, Driver memory: Motorist selection and retention of individualized headways in highway traffic, *Transportation Research Part A: Policy and Practice*, 32 (2) 1998, pp. 121-137.
- NGSIM. Next generation simulation. URL: <http://ngsim.fhwa.dot.gov/>, 2006.
- F. L. Hall, *Traffic flow theory: A state of the art report – revised monograph on traffic flow theory*. Washington, USA: US Federal Highway Administration, 1996, chapter “Traffic stream characteristics”.
- C. Thiemann, M. Treiber, A. Kesting, Estimation acceleration and lane-changing dyn S. Ahn, M.J. Cassidy, and J. Laval, Verification of a simplified car-following theory. *Transportation Research Part B: Methodological*, 38(5), 2004, pp. 431-440.

Contrôle d'accès par platitude différentielle

Hassane ABOUAÏSSA¹, Violina IORDANOVA²

¹ *Laboratoire de Génie Informatique et d'Automatique de l'Artois-LGI2A*

EA.3968

Université d'Artois, IUT de Béthune – Rue de l'Université

62408 Béthune Cedex, France

hassane.abouaissa@univ-artois.fr

² *Direction Interdépartementale d'Île de France (DIRIF/DI/PCET)*

DIRIF – 2 à 6 rue Olof Palme

94046, Créteil Cedex, France

violina.iordanova@developpement-durable.gouv.fr

Résumé

Cet article présente une nouvelle méthodologie pour le contrôle d'accès autoroutier basée sur la platitude différentielle. La propriété principale des systèmes dits « différentiellement plats » est que l'ensemble des variables (d'entrée et d'état) peut être exprimé directement, sans intégration d'aucune équation différentielle, en terme d'un ensemble de sorties plates et d'un nombre de ses dérivées dans le temps. La pertinence de cette méthodologie est démontrée à l'aide d'un ensemble de simulations numériques pour le contrôle d'accès d'un axe autoroutier.

Introduction

Le contrôle d'accès représente la seule solution viable pour faire face aux problèmes récurrents des congestions. De plus, l'implémentation de ce type de régulation ne nécessite pas de gros investissements en comparaison avec la construction de nouvelles infrastructures. Par ailleurs, les études menées dans ce domaine ont montré des résultats très satisfaisants dans des situations réelles (Middelham, 2001 ; Cambridge Systematics, 2001 ; Papageorgiou, 2002).

Le problème de la commande du trafic est étudié depuis plus de quarante ans et continu à ce jour d'être un sujet d'actualité (Wattleworth, 1964 ; Drew, 1965), (Haj-Salem, 1988 ; Papageorgiou, 2002). La majeure partie de ces travaux est

consacrée au contrôle d'accès, au guidage automatique, à la limitation dynamique de la vitesse, etc.

La mise en œuvre des lois de commande peut s'effectuer à l'aide des techniques issues de l'automatique classique. Dans ce contexte, plusieurs stratégies ont été proposées allant du simple contrôleur PI, à des approches plus robustes basées sur la commande LQR, la commande non linéaire, la commande optimale, la commande prédictive, la commande par mode glissant, etc. (Papageorgiou, 2003 ; Kachroo, 2000 ; Kostialos, 2002 ; Hegyi, 2005 ; Mammam, 2006). D'autres techniques reposent sur la programmation linéaire. Enfin, la littérature mentionne aussi plusieurs approches basées sur les systèmes experts ou à base des réseaux de neurones et/ou la commande floue (Ho, 1996).

La plupart de ces techniques sont validées par simulation et donnent des résultats très intéressants. Néanmoins, leur application et leur implémentation sur le terrain posent encore beaucoup de problèmes de calibrage.

Une autre technique récemment développée dans le domaine industriel est celle basée sur la platitude différentielle, (Fliess *et al.*, 1992 ; Fliess *et al.*, 1995), (Rudolph, 2003 ; Ramirez, 2004). L'avantage d'une telle approche est sa simplicité et sa puissance pour la planification et le suivi de trajectoires. Néanmoins, bien qu'elle soit de plus en plus appliquée dans le domaine de l'industrie, aucune méthode, à notre connaissance, basée sur ce concept n'a été employée dans le domaine du contrôle d'accès.

La première section rappelle quelques définitions de base de la platitude différentielle. La seconde section, met l'accent sur l'application de ce concept dans le cadre du contrôle d'accès isolé. L'efficacité de l'approche proposée est illustrée dans la troisième section, à travers un ensemble de simulations numériques avec des données réelles. Enfin, la dernière section discute des principales conclusions ainsi que de quelques problèmes ouverts.

1. Platitude différentielle : définitions et propriétés

La propriété de platitude, partagée par de nombreux systèmes pratiques, a été développée en 1992 par M. Fliess et ses collaborateurs (Fliess *et al.*, 1992). Ce concept est basé sur le formalisme de l'algèbre différentielle (ce qui justifie le terme plat). En effet, l'algèbre différentielle est une discipline mathématique dont l'objectif est de généraliser les méthodes de l'algèbre commutative à l'étude des systèmes d'équations différentielles non linéaires polynomiales. Basée principalement sur les travaux de Ritt et Kolchin (Ritt, 1950 ; Kolchin, 1973), elle a été introduite dans le domaine de l'automatique en France en 1985 par M. Fliess. La motivation première était de traiter les problèmes d'inversion entrées-sorties des systèmes non linéaires (Fliess, 1986). Cependant, elle s'est révélée très utile pour l'étude des propriétés structurelles telles que : l'observabilité, la linéarisation, le découplage, le rejet de perturbation, etc. C'est dans cette optique qu'a pu être dégagée la notion de la platitude et de système différentiellement plat. Dans ce qui suit, on se focalise principalement sur la définition de la platitude ainsi qu'à ses

propriétés. Cependant, pour les outils mathématiques de l'algèbre différentielle, tels que les notions de groupe, de corps, d'extension de corps ainsi que leurs développements, le lecteur est invité à consulter les références suivantes : (Cohn, 1985 ; Lam, 2001 ; Kolchin, 1973 ; Ritt, 1950).

1.1 Définition de la platitude

La notion de platitude est fermement reliée à la capacité générale de linéariser un système non linéaire par le choix approprié d'un bouclage dynamique endogène (Fliess *et al.*, 1995). Plus précisément, le système non linéaire décrit par la représentation dans l'espace d'état :

$$\begin{cases} \dot{x} = f(x, u); & x \in R^n, u \in R^m \\ z = g(x, u); & z \in R^p \end{cases} \quad (1)$$

où $n \geq m$, et f est un champ de vecteur analytique de dimension appropriée avec $\text{rang} \frac{\partial f}{\partial u} = m$ et g une fonction analytique, est dit différentiellement plat (ou simplement « plat »), s'il existe une sortie $y = (y_1, y_2, \dots, y_m)$ de la forme :

$$y = \psi \left(x, u, \dot{u}, \ddot{u}, \dots, u^{(l)} \right) \quad (2)$$

telle que les variables d'état, de sortie et de commande, x , z et u respectivement peuvent être exprimées à l'aide de cette sortie y et d'un nombre fini de ses dérivées dans le temps :

$$\begin{cases} x = \Phi_1 \left(y, \dot{y}, \ddot{y}, \dots, y^{(r)} \right) \\ u = \Phi_2 \left(y, \dot{y}, \ddot{y}, \dots, y^{(r+1)} \right) \\ z = \Phi_3 \left(y, \dot{y}, \ddot{y}, \dots, y^{(r)} \right) \end{cases} \quad (3)$$

- où r est un entier positif.

Dans ce cas, à toute trajectoire de référence arbitraire $t \rightarrow (y(t))$ suffisamment dérivable, correspond une trajectoire :

$$t \mapsto \begin{pmatrix} x(t) \\ u(t) \end{pmatrix} = \begin{pmatrix} x = \Phi_1 \left(y, \dot{y}, \ddot{y}, \dots, y^{(r)} \right) \\ u = \Phi_2 \left(y, \dot{y}, \ddot{y}, \dots, y^{(r+1)} \right) \end{pmatrix}$$

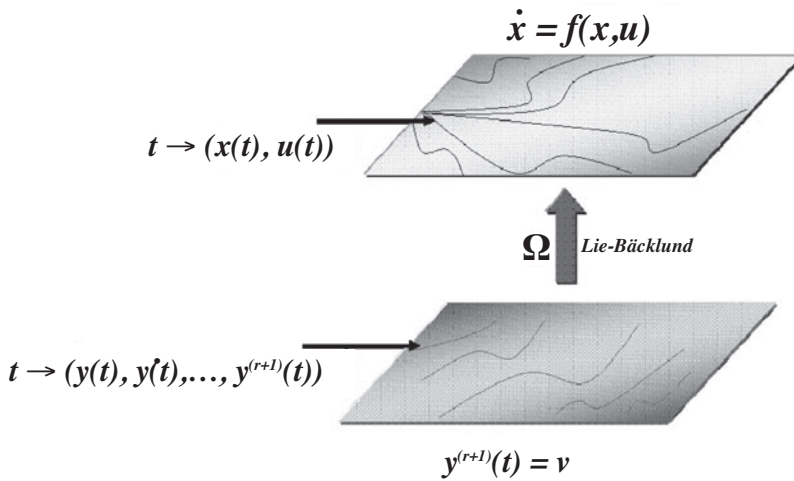
qui vérifie identiquement les équations du système.

Inversement, à toute trajectoire $t \rightarrow (x(t), u(t))$, suffisamment dérivable et vérifiant les équations du système, correspond une trajectoire

$$t \mapsto \gamma(t) = \psi(x, u, \dot{u}, \ddot{u}, \dots, u^{(l)})$$

La correspondance de trajectoires (Figure 1) (Laroche, 2004) est réalisée sans intégration d'aucune équation différentielle du système.

Figure 1 : correspondance de trajectoires



Dans cette figure, Ω est une application appelée « isomorphisme de Lie-Bäcklund ».

La propriété de la platitude permet d'apporter une solution simple et élégante aux problèmes de planification de trajectoires (Fliess *et al.*, 1995). En effet, les systèmes différentiellement plats sont très utiles dans les situations où la génération de trajectoires explicite est nécessaire. Puisque le comportement des systèmes plats est déterminé par les sorties plates, on peut planifier les trajectoires dans l'espace de sortie et les relier ensuite aux entrées appropriées. Plus précisément, à partir des trajectoires de y on peut en déduire immédiatement les trajectoires de x et de u par projections dans l'espace d'état avec une relation algébrique (Fliess, 1995 ; Ramirez, 2004 ; Rudolph 2003).

Dans le domaine industriel, plusieurs systèmes de différents domaines : mécanique, électrique, hydraulique, chimique, etc., sont plats. Leurs sorties plates ont souvent une interprétation physique.

Remarque : Les sorties plates ne sont pas toujours définies pour toutes valeurs de l'état. Un système peut être plat sur un intervalle ouvert. Inversement, l'expression de l'état en fonction des sorties plates et de leurs dérivées peut n'être que locale.

1.2. Propriétés de la platitude

- a. La définition présentée ci-dessus n'est pas limitée aux systèmes représentés par des variables d'état explicites, mais à toute représentation par variables d'état de la forme :

$$\Psi\left(x, \dot{x}, \dots, x^{(n)}, u, \dot{u}, \dots, u^{(m)}\right) = 0 \quad (4)$$

- b. Le nombre de sorties plates est égal au nombre de variables de commande indépendantes.
- c. Les sorties plates ne sont pas uniques. Si y est une sortie plate, alors toute fonction $\xi = \psi(y)$ est aussi une sortie plate, avec ψ inversible.
- d. *Mutatis mutandis* ; la notion de platitude peut être définie dans le cadre des systèmes discrets en remplaçant alors la dérivation par une avance.
- e. On parle de δ -platitude lorsque la sortie plate $y(t)$ dépend de $u, \dot{u}, \dots, u^{(\delta)}$.

2. Application de la platitude au contrôle d'accès

2.1. Principe du contrôle d'accès

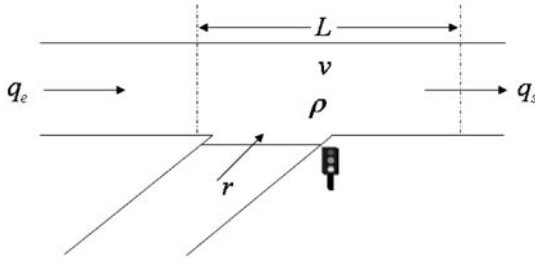
Le contrôle d'accès autoroutier consiste à réguler le débit des rampes d'entrées aux voies principales à l'aide de feux de signalisation, afin de gérer les conflits d'insertion au niveau de l'accès, ce qui permet de maintenir la densité sensiblement au-dessous d'une valeur critique (densité critique)¹ assurant aussi une meilleure fluidité de la circulation (Haj-Salem *et al.*, 2007).

La figure 2 illustre le principe de cette action de régulation.

Le contrôle d'accès peut-être isolé, à proximité de la jonction entre la rampe d'accès et la voie principale. Elle peut-être coordonnée, mettant en œuvre plusieurs rampes simultanément (Haj-Salem, 1998). L'objectif dans ce cas est de répartir la demande dans l'espace et dans le temps afin de limiter les congestions récurrentes (typiques aux heures de pointe) sur les axes principaux.

¹ Notons que la densité critique correspond généralement à la capacité du segment étudié.

Figure 2 : principe du contrôle d'accès



2.2. Algorithme du contrôle d'accès isolé

Le caractère fortement non linéaire du trafic impose d'appliquer des techniques de commande plus sophistiquées. À cet effet, la platitude différentielle permet d'apporter une solution simple et élégante notamment au problème du contrôle d'accès (Iordanova, 2006 ; Abouaïssa, *et al.*, 2007).

Dans ce paragraphe, nous allons montrer la faisabilité d'application de la platitude différentielle pour le contrôle d'accès isolé.

Considérons la section de la figure 2. Pour la description de l'évolution du flux de trafic, trois variables principales, supposées fonctions du temps t et de l'espace x sont utilisées :

- la densité du trafic ρ exprimée en nombre de véhicules par kilomètre,
- le débit q en nombre de véhicules par heures,
- la vitesse moyenne v , en kilomètres par heure.

À partir de ces trois variables, la loi de conservation du flux de véhicules s'écrit (Lighthill, Whitman et Richard ; Lighthill & Whitman, 1955 ; Richard, 1956) :

$$\frac{\partial \rho}{\partial t} + \frac{\partial q}{\partial x} = 0 \tag{5}$$

Cette équation est complétée par deux autres équations :

$$q(t) = \rho(t)v(t) \tag{6}$$

$$q(t) = f(\rho(t)) \tag{7}$$

La discrétisation de l'équation (5) dans l'espace s'écrit :

$$\frac{d\rho(t)}{dt} = \frac{1}{L}(q_e(t) - q_s(t) + r(t)) \tag{8}$$

En substituant l'équation (6) dans (8) on obtient :

$$\frac{d\rho(t)}{dt} = \frac{1}{L}(q_e(t) - \rho(t)v(t) + r(t)) \quad (9)$$

La section décrite par l'équation (9) est caractérisée par une variable d'état (la densité²) et une entrée (r).

Proposition

La section étudiée est plate avec la variable d'état ($y = \rho$) sa sortie plate.

La première dérivée de la sortie plate s'écrit :

$$\dot{y}(t) = \frac{1}{L}(q_e - y(t)v(t) + r) \quad (10)$$

On remarque que la variable de commande apparaît dès la première dérivée. De cette façon toutes les variables du système (d'état et de commande) peuvent être exprimées en fonction de la sortie plate et de sa première dérivée dans le temps.

$$\begin{cases} \rho(t) = y(t) \\ r(t) = L\dot{y}(t) + y(t)v(t) - q_e(t) \end{cases} \quad (11)$$

L'équation de la variable d'état permet de choisir la trajectoire souhaitable pour la densité. L'expression de la variable de commande permet d'ajouter des contraintes additionnelles de la densité à cette trajectoire de sorte qu'elle ne dépasse pas une valeur critique.

2.1.1. Planification de trajectoire : commande en boucle ouverte

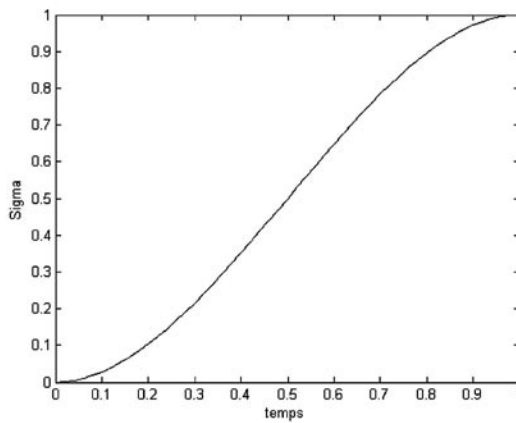
Afin d'élaborer la planification des trajectoires, il est nécessaire de déterminer une loi de commande en boucle ouverte. Pour cela, on définit une trajectoire désirée $y^*(t)$. D'après l'expression de la commande dans le système d'équation (11), cette trajectoire doit avoir des dérivées lisses au moins d'ordre deux. À partir des conditions initiales et finales de la densité $y(t_i) = y_i ; \dot{y}(t_i) = 0$ et $y(t_f = T) = y_f ; \dot{y}(t_f = T) = 0$, on peut construire cette trajectoire de référence pour la densité (sortie plate) en utilisant une interpolation polynomiale dans le

² En pratique, la variable mesurable est le taux d'occupation que l'on traduit ensuite en densité. Dans cet article, par soucis de simplicité on utilise directement la densité.

temps qui est la mieux adaptée puisqu'elle ne nécessite pas un grand effort de calcul dans un environnement temps réel³. On peut choisir la fonction suivante (Figure 3) :

$$\sigma(t) = \begin{cases} 0 & \text{si } t < 0 \\ 3t^2 - 2t^3 & \text{si } 0 \leq t \leq 1 \\ 1 & \text{si } t > 1 \end{cases} \quad (12)$$

Figure 3 : fonction $\sigma(t)$



De cette manière, la trajectoire de référence est définie par :

$$\begin{cases} y^*(t) = y_i + (y_f - y_i) \left(3 \left(\frac{t - t_i}{T - t_i} \right)^2 - 2 \left(\frac{t - t_i}{T - t_i} \right)^3 \right) \\ y^*(t) = y_i + (y_f - y_i) \sigma \left(\frac{t - t_i}{T - t_i} \right) \end{cases} \quad (13)$$

À l'aide de l'expression (11) et la relation (13), on en déduit les trajectoires de densité et le taux de débit que la rampe d'entrée est capable de générer en l'absence de perturbation (Figure 4).

³ Voir (Rudolph, 2003, Ramirez, 2004) pour plus d'information sur la construction des trajectoires de références

Figure 4 : structure de la commande en boucle ouverte


$$\begin{cases} \rho^* = y^* \\ r^* = L\dot{y}^* + y^* v^* - q_e \end{cases} \quad (14)$$

Puisque l'expression de la commande (14) est une relation explicite, les contraintes d'entrée $r_{min} < r^* < r_{max}$ peuvent être considérées hors ligne, par simple évaluation de l'équation fondamentale (13), sans avoir recours à des intégrations d'équations différentielles. Le temps de transfert T doit être choisi afin d'éviter la violation de ces contraintes. Dans cette optique, une procédure itérative permet facilement d'obtenir la trajectoire et la commande en boucle ouverte respectant les contraintes avec des marges appropriées.

2.1.2. Suivi de trajectoire : commande en boucle fermée

La commande en boucle ouverte suppose entre autres que le modèle utilisé soit parfait et que les conditions initiales soient connues avec exactitude (Brault *et al.*, 2000). Dans la pratique, ces connaissances ne sont jamais exactes. De plus, du fait des incertitudes paramétriques des modèles du flux de trafic, de l'apparition de perturbations et des changements dans les conditions de circulation, la commande en boucle ouverte n'est généralement pas suffisante pour obtenir une régulation du flux de trafic efficace. Pour assurer un état d'équilibre et réduire l'influence des variations des paramètres, la commande doit opérer en boucle fermée. En utilisant l'algorithme de la commande plate en boucle ouverte (13), un bouclage additionnel peut être déterminé afin d'obtenir le comportement dynamique désiré compensant les perturbations externes.

Afin d'assurer le suivi de trajectoire, on utilise un bouclage endogène dont le principe est le suivant. On pose $w = \dot{y}$ dans l'expression de la première dérivée de la sortie plate. En remplaçant cette variable dans l'expression de la commande le système prend la forme suivante :

$$w = \dot{y}^* - K_p (y - y^*) \quad (15)$$

où K_p est un paramètre qui doit être choisi de telle sorte qu'il puisse satisfaire les performances désirées du système en boucle fermée et assurer la stabilité asymptotique. Cependant, ce type de bouclage est un bouclage par retour d'état qui nécessite donc que les variables d'état du système puissent être mesurées ou reconstruites.

En remplaçant (15) dans l'équation (13), on obtient l'expression de la commande par retour endogène sous la forme :

$$r = L \left(\dot{y}^* - K_p (y - y^*) + \frac{1}{L} y^v - \frac{1}{L} q \right) \quad (16)$$

2.3. Études préliminaires en simulation

L'objectif de ce paragraphe est d'illustrer, à l'aide d'un exemple simple, la pertinence de l'algorithme basée sur la platitude. On considère pour cela le cas de la section (Figure 2). Pour la simulation, nous avons utilisé le modèle amélioré LWR prenant en compte les accélérations bornées (Lebacque 2002) et introduisant les fonctions d'offre et de la demande, ainsi que un modèle utilisant la loi des nœuds⁴. Le tableau 1 résume l'ensemble des paramètres utilisés :

Tableau 1 : données de simulation

Paramètre	Valeur
Densité critique ρ_c	35 véh/km
Densité maximale ρ_m	180 véh/km
Vitesse libre V_f	110 km/h
Longueur de la section L	1500 m
a	1,66
Nombre de voies	3

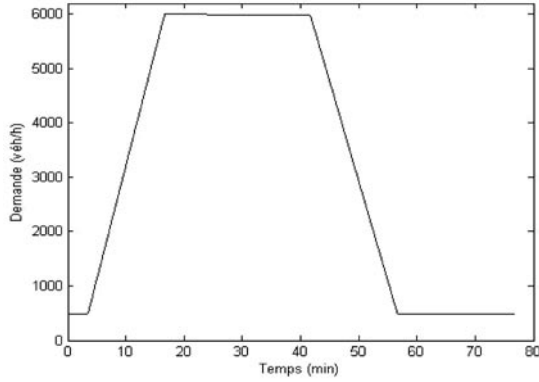
a est un paramètre utilisé dans le diagramme fondamental de May (May, 2002) :

$$q = \rho V_f e^{\left(-\frac{1}{a} \left(\frac{\rho}{\rho_c} \right)^a \right)} \quad (17)$$

On considère aussi un profil de la demande sur la section principale schématisé dans la Figure 5 :

⁴ Nous sommes conscients que les algorithmes du contrôle d'accès sont généralement appliqués avec les modèles macroscopiques du second ordre (Voir Haj-Salem *et al.* 2006).

Figure 5 : profil des demandes sur la section principale et sur la rampe d'accès

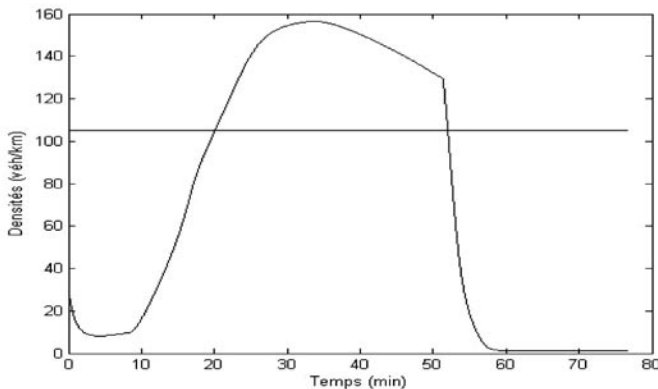


La densité initiale est fixée à $\rho_0 = 10$ véh/km, ce qui correspond à une situation où le trafic est plutôt en régime fluide. L'objectif de la commande est de maintenir une désirée fixée à $\rho_c = 35$ véh/km.

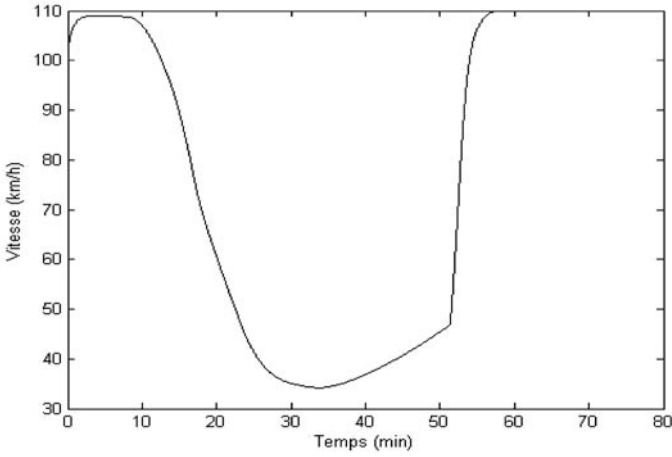
Les figures (6a-b) montrent l'évolution de la densité et de la vitesse sans contrôle d'accès. On remarque que l'augmentation de la demande aux origines a pour conséquence l'apparition d'une congestion avec une chute de la vitesse du flux qui atteint à peu près 32 km/h.

Figure 6 : état du trafic dans le cas sans contrôle d'accès

a. Évolution de la densité



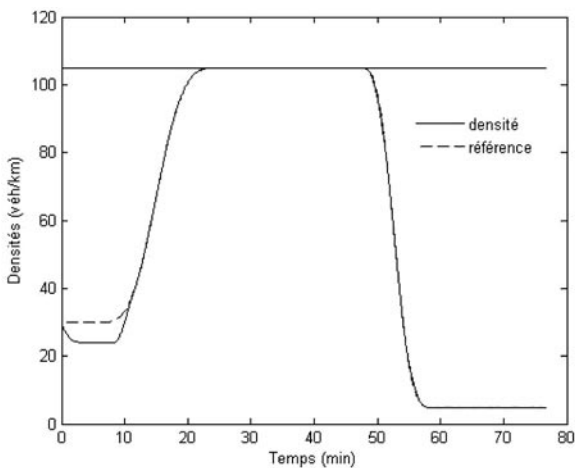
b. Évolution de la vitesse



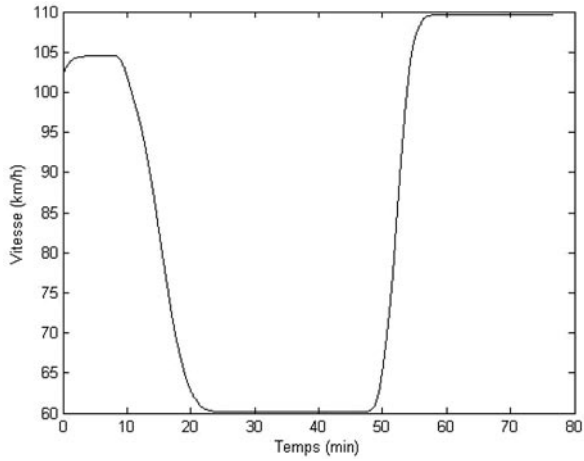
Les figures (7a-c) montrent l'évolution de la densité, de la vitesse, ainsi que le débit de la rampe d'accès. À noter que ce débit de la rampe est borné entre une valeur minimale $r_{\min} = 200$ véh/h et une valeur maximale $r_{\max} = 2000$ véh/h. La mise en œuvre du contrôleur d'accès permet de constater une nette amélioration de la circulation.

Figure 7 : contrôleur d'accès basé sur la platitude différentielle

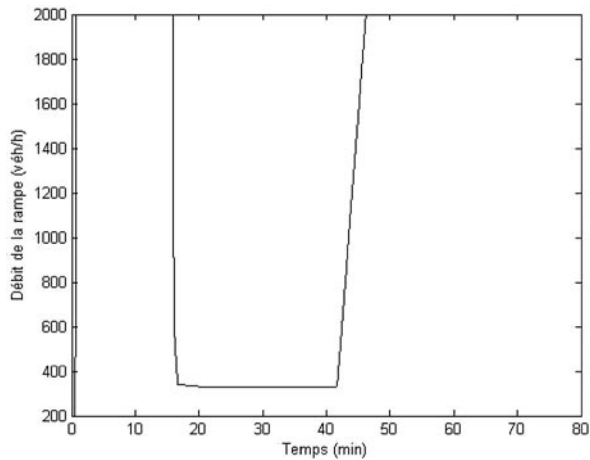
a. Évolution de la densité



b. Évolution de la vitesse



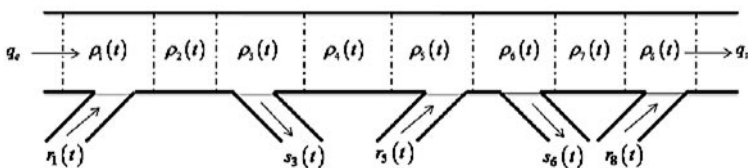
c. Débit de la rampe d'accès



3. Simulations numériques : cas d'un réseau à trois rampes d'accès

Considérons maintenant une section autoroutière composée de trois rampes d'accès (Figure 8).

Figure 8 : portion d'autoroute pour l'évaluation du contrôle d'accès

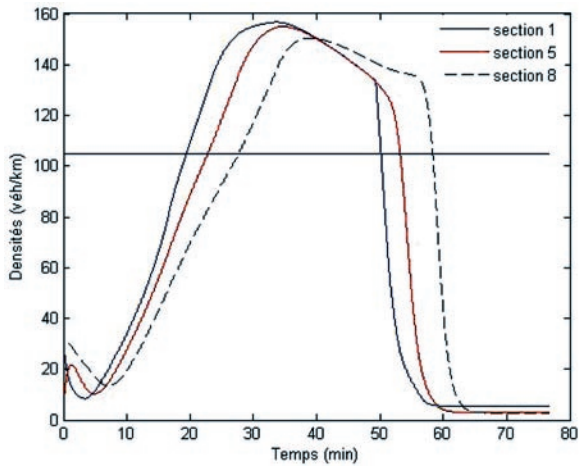


On considère le même profil de la demande sur la voie principale que celui illustré dans la figure 4 ainsi que les mêmes paramètres du modèle de simulation. De plus, on suppose que le débit au niveau des deux rampes de sorties est égal à 400 véh/h.

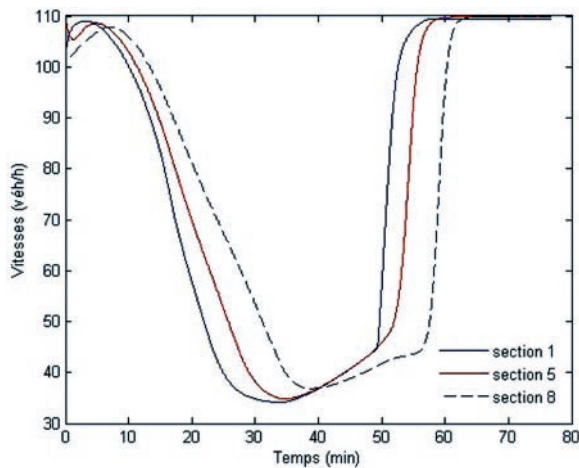
On remarque (Figures 9 a-b), que lorsque aucun contrôle n'est appliqué, l'accroissement des demandes aux origines conduit à la formation de congestions.

Figure 9 : évolution des variables du trafic dans le cas sans contrôle d'accès

a. Évolution des densités



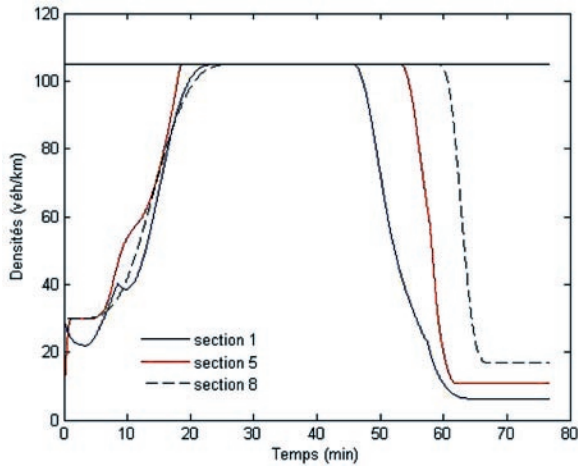
b. Évolution des vitesses



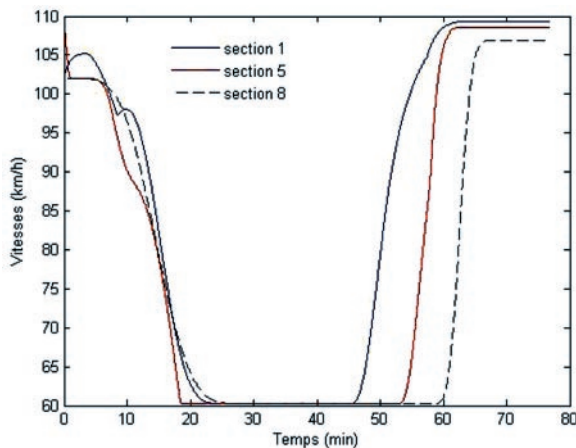
Appliquons l'algorithme basé sur la platitude différentielle pour chacune des trois rampes d'accès. Les figures (10) montrent que lorsque le trafic est fluide, le signal de commande est caractérisé par une valeur élevée, ce qui signifie que tous les véhicules présents au niveau des trois rampes d'accès sont autorisés à emprunter la voie principale. Durant la congestion, les débits des trois rampes diminuent sous l'effet de la commande.

Figure 10 : contrôle d'accès par platitude de l'axe autoroutier

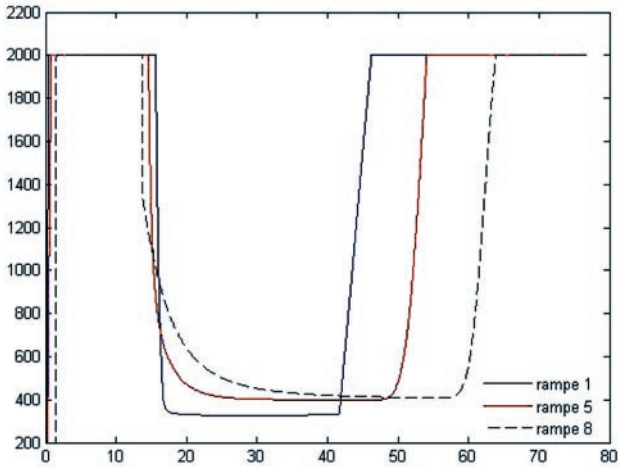
a. Contrôle des densités dans les trois sections



b. Évolution des vitesses dans les trois sections



c. Débits des trois rampes



Conclusion

Les travaux présentés dans cet article nous ont permis d'introduire une nouvelle alternative aux méthodes existantes du contrôle d'accès. L'algorithme proposé est basé sur l'inversion du système réalisé grâce à la platitude différentielle. Un avantage certain dans l'utilisation de cette méthode est sa simplicité ainsi que le fait qu'elle ne nécessite l'intégration d'aucune équation différentielle. La nature du système étudié montre qu'un contrôleur basé sur une commande plate en boucle ouverte s'avère insuffisant pour obtenir une régulation du trafic satisfaisante dans la section étudiée. Ainsi, afin de prendre en compte les différentes fluctuations dans le schéma de contrôle et de traiter les incertitudes de la modélisation ainsi que le rejet de perturbation, nous avons proposé de mettre en œuvre une commande basée sur un bouclage endogène.

La qualité de commande dépend principalement de la qualité des mesures obtenues via les capteurs. Cependant, ces mesures sont généralement entachées de bruit. Un des objectifs de nos travaux futurs est d'exploiter les dernières avancées dans le domaine d'identification afin d'élaborer une commande basée sur l'estimation d'état du trafic (principalement, l'estimation de la densité).

Références

- Abouaïssa, H., Iordanova, V., Jolly, D., 2007. "Integrated Control of Highway Systems Using Flatness-Based Concept", SCS-IFAC, International Conference on Advances in Vehicle Control and Safety, Buenos-Aires, Argentina.

- Brault, P., Mounier, H., Petit, N., Rouchon, P. 2000. "Flatness based tracking control of a manoeuvrable vehicle: the π -Car", MTNSOO Mathematical Theory of Networks and Systems.
- Cambridge Systematics, Inc., Oakland, California, 2001. Twin Cities Ramp Meter Evaluation – Final Report, Prepared for the Minnesota Department of Transportation.
- Cohn, P. 1985. "*Free Rings and their Relations*". 2nd Ed. Academic Press. London.
- Drew, D. R. 1968. "*Traffic Flow Theory and Control*". Chapter 12, McGraw Hill Book Company, New York.
- Fliess, M. 1986, "A note on Invertibility of Nonlinear Input-Output Differential Systems". *Systems and Control Letters*, 8, pp. 147-151.
- Fliess, M., J. Lévine, P. Martin, P. Rouchon. 1992. "Sur les systèmes non linéaires différentiellement plats". *Compte Rendu de l'Académie des Sciences de Paris*, I-315, pp. 619-624.
- Fliess, M., J. Lévine, P. Martin, P. Rouchon. 1995. "Flatness and defect of nonlinear systems : Introductory theory and examples". *International Journal of Control*, 61(6), pp. 1327-1361.
- Haj-Salem, H. Davée M., J.M. Blosseville, Papageorgiou M. 1988. « ALINEA : Un nouvel outil de contrôle d'accès isolé Étude comparative sur site réel », Rapport de recherche INRETS n° 80.
- Haj-Salem, H., J-P. Lebacque, S. Mammar, 2006, "Theoretical Analysis of the Efficiency of Ramp Metering, Speed Management Based on Braess-Like Paradox", *Advanced OR and AI Methods in Transportation*, pp. 450-455.
- Haj-Salem H., S. Mammar, S. Mammar, 2007. "Le contrôle des rampes d'accès", dans *Systèmes de transport intelligents*, Ed. Hermes, Lavoisier, Ch. 8, pp. 227-260.
- Hegyi, A., B. De Schutter, H. Hellendoorn, 2005. "Model Predictive Control for Optimal Coordination of Ramp Metering and Variable Speed Limits", *Transportation Research Part C*, pp. 185-209.
- Ho, F.-S., P. Ioannou, 1996. "Traffic flow modeling and control using artificial neural networks", *IEEE Control Systems*, pages 16–26.
- Iordanova, V. 2006. "Contribution à la modélisation et la commande du trafic : Approches par Bond-Graph et commande par platitude", Thèse de Doctorat de l'Université d'Artois,
- Kachroo, P., K. Krishen, 2000, "System Dynamics and Feedback Control Design Problem Formulations for Real Time Ramp Metering", *Transactions of Society for Design and Process Science*.

- Kolchin, E. R. 1973. "*Differential Algebra and Algebraic Group*". Academic Press, New York, USA.
- Kotsialos, A., M. Papageorgiou, M. Mangeas, H. Haj-Salem, 2002. "Coordinated and integrated control of motorway networks via non-linear optimal control". *Transportation Research C*, 10(1), pp. 65–84.
- Lam, T. Y. 2001. "A first Course in Noncommutative Rings". 2nd Ed. Springer Verlag, New York, USA.
- Laroche B., Ph. Martin, N. Petit. 2004. "Commande par platitude, Équations différentielles ordinaires et aux dérivées partielles". Notes de cours, Cours B7-4 du Module Automatique Avancée, <http://cas.ensmp.fr/~petit/ensta/index.html>.
- Lebacque. J.P. 2002. "A Two Phase Extension of the LWR Model Based on the Boundedness of Traffic Acceleration". *Transportation and Traffic Theory in the 21st Century*, ISTTT.
- Mammar, S., S. Mammar, M. Netto, 2006. "Coordinated Ramp Metering via Second Order Sliding Mode Control", *Proceedings of the IEEE ITSC 2006 IEEE Intelligent Transportation Systems Conference Toronto, Canada*, pp. 261-266, September 17-20.
- May, A. D. 1990, "Traffic flow fundamentals", Englewood Cliffs, NJ: Prentice-Hall.
- Middelham F., 2001. "A synthetic study of the network effects of ramp metering". In *Proceedings of the Seminar Advanced Traffic Control, Delft, The Netherlands*, TRAIL Research School.
- Papageorgiou M. and A. Kotsialos, 2002., "Freeway ramp metering: An overview", *IEEE Transactions on Intelligent Transportation Systems*, vol. 3, no. 4, pp. 271-280,
- Papageorgiou, M., C. Diakaki, V. Dinopoulou, A. Kostialos, Y. Wang, 2003. "Review of Road Traffic Control Strategies". *Proceedings of the IEEE*, Vol. 91, N° 12.
- Ramirez, S., S. Agrawal, 2004. "Differentially Flat Systems", Marcel Dekker, New-York.
- Ritt, J. F. 1950. "Differential Algebra". American Mathematical Society, New York.
- Rudolph, J., 2003. "Flatness Based Control of Distributed Parameter Systems". Shaker Verlag.
- Wattleworth, J. A., 1965. "Peak-period analysis and control of a freeway system", *Highway Res. Board*, Washington, DC, *Highway Res. Record* 157.

Effects of variable speed limits on motorway traffic flow

Markos Papageorgiou, Elias Kosmatopoulos, Ioannis Papamichail

Dynamic Systems and Simulation Laboratory

Department of Production Engineering and Management

Technical University of Crete

73100 Chania, Greece

markos@dssl.tuc.gr

kosmatop@dssl.tuc.gr

ipapa@dssl.tuc.gr

From Transportation Research Record: Journal of the Transportation Research Board N° 2047, Transportation Research Board of the National Academies, Washington D.C., 2008, pp. 37-48. Reprinted with permission of TRB

Abstract

Variable Speed Limits (VSL) displayed on road-side Variable Message Signs (VMS) have emerged as a quite widespread traffic control measure on motorways in many countries leading to substantial traffic safety benefits; on the other hand, there is no clear evidence of improved traffic flow efficiency in operational VSL systems. The paper summarizes available information on the VSL impact on aggregate traffic flow behavior (flow-occupancy diagram) and investigates this issue in more detail by use of real traffic data from a European motorway. It is found that VSL decrease the slope of the flow-occupancy diagram at undercritical conditions; shift the critical occupancy to higher values; and enable higher flows at the same occupancy values in overcritical conditions. Implications of the derived findings for more efficient VSL control strategies are discussed.

Introduction

A particular control measure for the traffic flow on motorways is the display of variable speed limits (VSL) on appropriate variable message signs (VMS) in response to the prevailing traffic conditions. In most cases, VSL are mandatory,

i.e. legally equivalent to fixed speed limits, and may even be enforced to increase driver compliance and hence impact. VSL installations were first introduced in Germany more than three decades ago; today, numerous VSL installations are encountered in many European countries, in North America and elsewhere (e.g., a total of more than 800 km of VSL-equipped motorway stretches are currently in operation in Germany).

The main impact of VSL on traffic flow is deemed to be [1]:

- the reduction of the mean speed at undercritical densities;
- the homogenisation of speeds, i.e. reduction of speed differences among vehicles and of mean speed differences among lanes.

A main targeted impact of VSL is enhanced traffic safety and, indeed, the selection of motorway stretches for VSL installation in several countries is guided by the frequency of registered accidents. The positive impact of VSL on traffic safety is due to speed reduction and speed homogenisation which are correlated with a reduction of accident probability. Multi-year evaluations of the VSL impact on traffic safety indicate a reduction in accident numbers by as much as 20-30% after VSL installation. VSL are also envisaged by some authorities as a means to reduce vehicle emissions and road noise. On the other hand, to the best of our knowledge, there is no evaluation of the VSL impact of available installations that would demonstrate a measurable and consistent improvement of traffic flow efficiency, e.g. in the sense of reduced travel times.

The ideal exploitation of the opportunities offered by VSL would be to preserve the safety and environmental benefits offered by the current systems along with an increase of traffic flow efficiency. The fact that an efficiency increase could not be demonstrated in the conducted field-assessments, does not necessarily mean that VSL per se is not an appropriate measure for the enhancement of traffic flow efficiency. As a matter of fact:

- The impact of VSL on aggregate traffic behavior, e.g. on the fundamental diagram (i.e. the flow-occupancy or speed-flow diagrams), has not been sufficiently investigated with real data; as a consequence, our understanding of even qualitative (let alone quantitative) impacts of VSL is limited to conjectures and assumptions; this lack of reliable understanding, by its turn, hinders the insightful development of VSL control strategies that would target an increase of traffic flow efficiency.
- Current VSL installations employ simple rule-based control strategies for VSL switching, which base their real-time decisions on pre-selected thresholds of traffic flow or occupancy or mean speed. The utilized thresholds are usually selected in an ad-hoc way which does not necessarily exploit the (anyhow unknown) potential impact of VSL on traffic flow efficiency.

The main scope of the investigations reported in this paper is to analyze available data from a VSL-equipped European motorway so as to come up with a better understanding of the impact of VSL on aggregate traffic behavior. The enhanced understanding of the VSL impact may eventually be exploited in order to assess

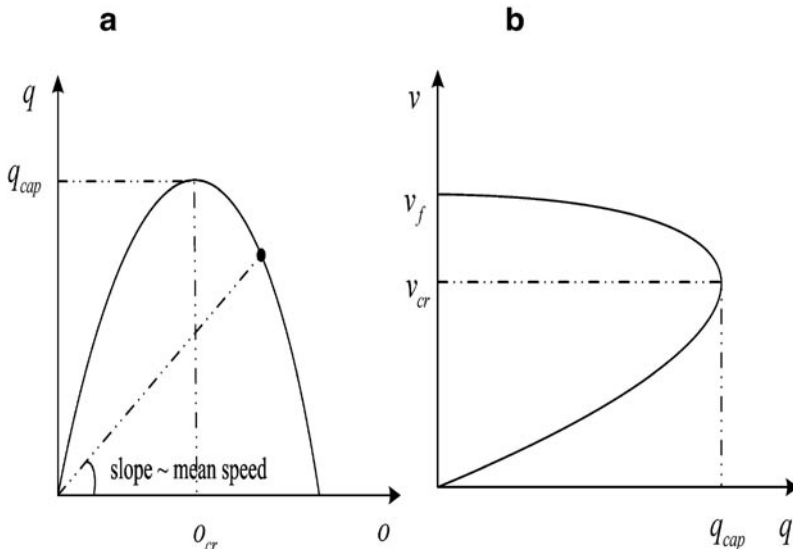
and enhance the current VSL control strategies or to develop new control strategies that would target traffic flow efficiency whilst delivering similar benefits for traffic safety and environmental improvements as achieved by current systems.

1. Background issues

1.1. The fundamental diagram

Under the assumption that traffic conditions do not change substantially in space (i.e., along a motorway stretch) and time (e.g. due to arrival of shock waves from downstream), the traffic flow states may be approximated by the so-called fundamental diagram, which may be a flow-occupancy diagram (inverse-U shape) or a speed-flow diagram (left-turned-U shape), see Figure 1. Recall that the mean speed of a particular traffic state on the flow-occupancy diagram (Figure 1a) is proportional to the slope of the line that connects the particular traffic state point with the origin. A fundamental diagram may be (partially) obtained by collecting measurements of the related traffic variables (flow, occupancy, mean speed) at a specific motorway location and fitting an appropriate mathematical function. This procedure, however, may lead to flawed results if the underlying spatio-temporal traffic flow phenomena are not appropriately considered. In particular, the area around the critical occupancy (capacity flow) is properly visible in real data only at active bottleneck locations (see [2], for more details).

Figure 1: flow-occupancy (a) and speed-flow (b) diagrams where:
 q is flow (veh/h); o is occupancy (%); v is mean speed (km/h);
 q_{cap} is capacity flow; o_{cr} is critical occupancy; v_f is free speed;
 v_{cr} is critical mean speed



The critical parameters (q_{cap} , o_{cr} , v_{cr}) of the fundamental diagram are not necessarily constant. Various data-driven investigations demonstrated that the flow capacity q_{cap} may vary considerably from day to day without any obvious reason. These variations were found to be more pronounced under different weather and lighting conditions. In contrast, the critical occupancy o_{cr} (at which capacity flow occurs) was found to be less sensitive with respect to different weather conditions while no related results are known for the critical speed v_{cr} . However, due to the relationship $q = \alpha \cdot o \cdot v$, where α is a constant, we also have $q_{cap} = \alpha \cdot o_{cr} \cdot v_{cr}$; then, if o_{cr} is relatively stable while q_{cap} is not, it may be deduced that v_{cr} varies as q_{cap} varies, i.e. the sensitivity of this parameter to stochastic or weather influences may be similar as the sensitivity of q_{cap} . The stochastic or weather-induced change of parameters implies that the fundamental diagram may change accordingly in the short term. Moreover, various changes in infrastructure or in vehicle technology may lead to medium or long-term changes in the aggregate traffic flow behavior.

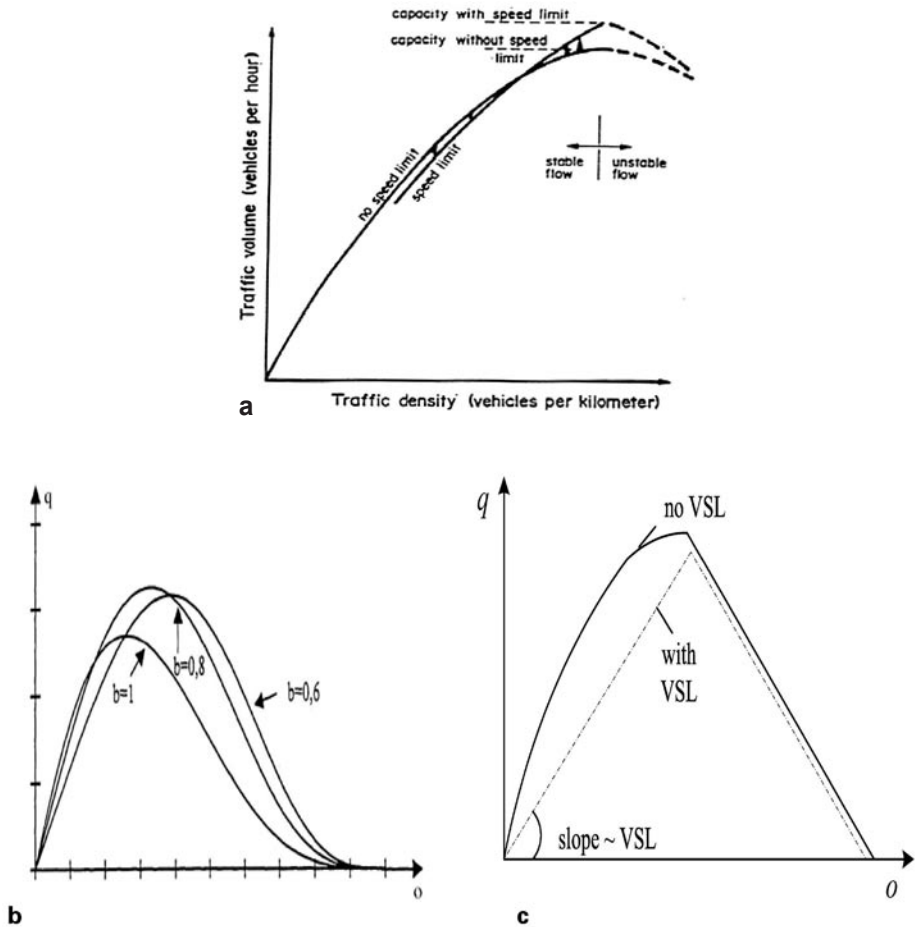
1.2. The impact of VSL

As mentioned earlier, there were very few investigations in the past addressing the precise impact of VSL on aggregate traffic flow behavior, e.g. on the fundamental diagram. Some early investigations (from the 1970s) are summarized by Zackor [1] based on traffic data with and without VSL on a two-lane German motorway. The results indicate a speed homogenisation effect (less speed differences) for individual vehicles as well as for motorway lanes under the impact of VSL. These results are useful for a better understanding of the VSL impact on the individual vehicle speed distribution, but they do not reveal the impact of VSL on the aggregate traffic flow behavior. The latter was also addressed in [1], see Figure 2a, but in a rather qualitative way. Figure 2a illustrates that *“at lower or mean traffic volumes, the mean speed is lower due to the reduction effect whereas, at higher volumes, an increase is detected due to the stabilizing effect. Thus, both capacity and speed rise by about 5% to 10% at the same time”* [1]. Zackor [1] did not comment on the possible increase of the critical occupancy (or critical density) under the influence of VSL.

The results reported in [3] (summarized in [1]) were the basis for Cremer [4] to propose a quantitative model for the VSL-induced fundamental diagram change as displayed in Figure 2b where b is the ratio of the applied VSL divided by the free speed without VSL, and, by convention, $b = 1$ corresponds to the no-VSL case. It is quite likely that the displayed increase of flow capacity is rather exaggerated. In fact, later Dutch investigations could not identify any capacity increase that could be attributed to VSL [5], albeit under advisory (not mandatory) VSL.

In other, more recent research work regarding VSL control, the assumed VSL impact was to merely replace the left part of the flow-occupancy curve by a straight line with slope corresponding to the displayed VSL, see Figure 2c [6].

Figure 2: (a) change of fundamental diagram due to speed limits (1) ; (b) Cremer (4) model for VSL impact, where : $b = 1, 0.8, 0.6$ correspond to no speed limit, $VSL = 0.8 v_f$, $VSL = 0.6 v_f$ respectively ; (c) Hegyi (6) model for VSL impact



Microscopic simulation was also used for assessing the impact of VSL and related strategies, see, e.g., [7].

In conclusion, there seems to be very limited empirical evidence and, indeed, no factual consensus on the potential impact of VSL on aggregate traffic behavior, let alone quantitatively reliable results that could be used for efficient control strategy development. Before proceeding to the analysis of available traffic data in the next sections, it is useful to summarize at this point our expectations of the VSL impact along with their implications for potentially more efficient traffic flow.

1.3. Reducing mean speed at undercritical occupancies

It seems quite reasonable to assume that a VSL displayed at undercritical occupancies will reduce (with reasonable driver compliance) the (otherwise higher) mean speed (Figure 3a). The magnitude of this effect is likely to depend on the displayed VSL as well as on driver compliance. The new, VSL-affected states serve the same flow at lower speed and higher occupancy than the original states, which implies that the travel time increases accordingly. Thus, applying VSL to undercritical traffic states is likely to increase travel times and hence deteriorate traffic flow efficiency.

The described state transition when applying VSL at undercritical occupancies could, however, be exploited in a different context. The application of VSL **upstream** of a bottleneck that is close to becoming active, will **temporarily** (for the duration of the traffic state transition triggered by the VSL) decrease the main-stream flow arriving in the bottleneck area, thus delaying the bottleneck activation and the resulting congestion. Note that the temporary flow decrease during the VSL-triggered traffic state transition is due to the fact that occupancy (and density) in the VSL state is higher than in the original non-VSL state; thus, during the transition, the flow is temporarily reduced so as to “create” the higher traffic density of the VSL-state. It should be noted that this is the only VSL impact exploited in [6].

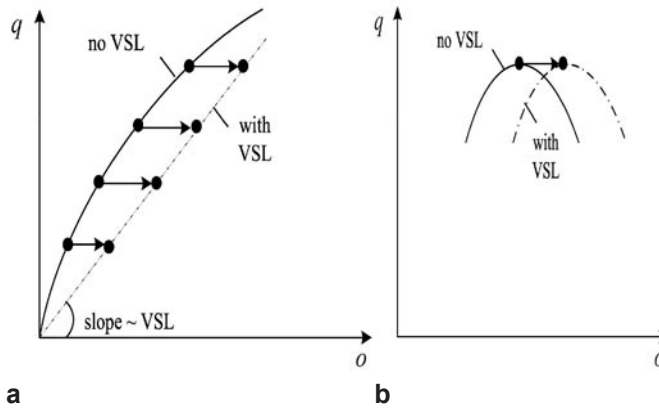
1.4. Increasing throughput and delaying congestion at (over) critical occupancies

According to Hegyi [6] modeling (Figure 2c) both curves (for VSL and non-VSL) meet, but do not actually cross; while Zackor [1] suggests that there is actually a genuine cross-point of both curves somewhere near the critical occupancy (Figure 2a). The cross-points (if any) are likely to lie at increasing occupancy values for decreasing VSL due to the accordingly decreasing slope of the undercritical VSL-affected curves, and in fact there may be no cross-point for very low VSL. While evaluating the VSL impact at occupancies near or higher than the cross point, the following (partly overlapping) questions are of interest (see also Figure 3b):

- Where is the cross-point (if any) located with respect to the non-VSL critical occupancy?
- Are VSL-induced critical occupancies higher than their non-VSL counterpart?
- Are VSL-induced flows higher at overcritical occupancies than their non-VSL counterparts?
- Is there any flow capacity increase for some VSL?

Answering these questions is crucial towards the development of appropriate control strategies for VSL capable of retarding (or avoiding) the onset of congestion.

**Figure 3: (a) potential VSL impact on undercritical mean speeds;
(b) cross-point of diagrams with and without VSL**



1.5. Implications for VSL control strategies

The conclusions derived in the previous sections converge to the following global rules for efficient switching of VSL:

(a) VSL-activation at occupancies lower than the crosspoint of the VSL and non-VSL flow-occupancy diagrams decreases the traffic flow efficiency (increases travel times), unless it is used to address a downstream bottleneck.

(b) VSL-activation at the crosspoint occupancy or (the latest) at the non-VSL critical occupancy is likely to improve the traffic flow efficiency due to avoidance or delay of congestion as well as improved traffic flow stability which might allow for higher flows under overcritical occupancies.

These statements consider essentially **one** single VSL-induced flow-occupancy diagram whereas, in reality, there is a series of such diagrams, one for each available VSL. However, the rules (a) and (b) above could be followed for each subsequent switching, e.g. from no-VSL to 60 mph then to 50 mph and then to 40 mph.

In conclusion, a possible VSL control strategy (practiced in many VSL installations) would call for prior determination of some thresholds reflecting the above rules, e.g. three thresholds for three available VSL, and would switch the VSL in real time accordingly. Some comments are necessary with regard to this procedure:

- To avoid switching oscillations, each activation threshold should be accompanied by a different de-activation threshold so as to create a hysteresis for the switching process.
- The thresholds may be determined in terms of measured flows or occupancies or speeds or a combination thereof. Given the sensitivity of flow and speed behavior with regard to stochastic and weather conditions, it

seems that decision thresholds should mostly rely on occupancies for a more robust behavior that would avoid (or reduce) VSL activation “too early” or “too late”.

- The appropriate threshold values may be different from location to location which may call for a tedious calibration procedure. Notice that the calibration procedure cannot be conducted before VSL installation and operation, because of the need to determine the location-specific cross-points of the various flow-occupancy curves.

These statements, related to a potential rule-based VSL control strategy, will be revisited later, following the presentation of the derived VSL-impact results.

2. The motorway site and data available

2.1. The motorway site

The data used for the analysis of the effects of Variable Speed Limits (VSL) on motorway traffic conditions stem from both directions (3 lanes each) of a motorway stretch equipped with loop detectors placed every 500 meters and providing with occupancy, flow, and speed measurements. A VSL control system is used that determines in real time the speed limits of this motorway stretch by providing appropriate displays to the drivers via variable message signs. This control system uses a rule-based algorithm which relies on flow as well as speed thresholds. The flow and speed threshold values initially applied were revised as a result of a major fine-tuning process in January 2006. Among others, this fine-tuning had the effect of more frequent activation of the speed limits.

2.2. Data treatment

Data from two different periods were available for the present study:

- The period May 2005 – October 2005, when the VSL control algorithm with the initial threshold values was operating and setting advisory speed limits. The utilized threshold values during this period were rather “loose”, i.e. VSL activations were relatively sparse. This is beneficial for the present analysis as the 2005 data contain non-VSL measurements also at relatively high occupancy levels.
- The period January 2006 – May 2006, after the major fine-tuning; during this period, mandatory speed limits were set, and threshold values for the control algorithm were rather “tight”.

Both periods will be referred to as “2005-period” and “2006-period” for brevity. Data were available for both motorway directions, referred to as directions A and B in the following. More details on the available data may be found in [2].

A first step of data treatment was to validate the available data and neglect data that are faulty or irrelevant for the present analysis. In all cases where occu-

pancy/flow/speed data were removed as faulty, the v values were replaced by zero. Moreover, daily data that provided little or no information regarding the effect of speed limits on traffic conditions were not taken into account in our analysis (e.g. weekends, incidents).

Subsequently, the available data were classified according to the VSL that was active at the minute of data collection. More specifically, data were classified according to the VSL displayed on the first gantry upstream of the corresponding detector location, the underlying assumption being that the distance of 500-2000 m between the gantry and the detector are sufficient for the drivers to adjust their speed according to the displayed VSL. The classification comprised four cases: NSL $\hat{=}$ no speed limit, 60 mph, 50 mph, and 40 mph.

2.3. Data analysis methods

Since the main goal of the investigation is to examine whether the speed limits modify the shape of the flow-occupancy diagram, curve-fitting methods were used to analyze the effect of speed limits as follows: The occupancy axis was divided into intervals of equal length Δo and 2nd-order polynomial curve-fitting was performed for each of these intervals $[0, \Delta o]$, $[\Delta o, 2\Delta o]$, $[2\Delta o, 3\Delta o]$,... At each interval, four different 2nd-order curves were produced, each for data corresponding to different speed limits (NSL, 60, 50, 40). After experimenting with different values for Δo , it was concluded that a value of $\Delta o = 4\%$ was producing the most meaningful results. The reason for producing separate fitting curves for different Δo -regions was in order to avoid unbalanced fitting due to much higher numbers of data in some occupancy regions compared to other regions. In order to avoid the situation where curve fitting is using a very small number of available data, curve fitting was actually performed in an interval Δo , only if the number of available data exceeded 9 data points.

2.4. Data selection

A total of 27 days (6 days from the 2005-period and 21 days from the 2006-period) were selected randomly for further treatment and were classified into three categories according to the weather conditions of the particular day: "clear" (18 days), "rainy" (4 days) and "?" (5 days). The last category comprises days where it was raining on and off during the pm peak.

Before proceeding to the evaluation of the impact of VSL control on the aggregate motorway traffic conditions, some preliminary observations regarding the particular conditions on both motorway directions may be helpful:

- The longest part of both directions of motorway seems to resemble a motorway "pipeline" without pronounced bottlenecks and very few on/off ramps.
- Before the major fine-tuning, the VSL activation was rather scarce; this enables for availability of no-VSL data at and beyond the critical occupancy. After the major fine-tuning, the VSL activation was more frequent than before.

- For direction A, the main active bottleneck was always at or close to the specific data location 441A. This did not happen in direction B where different locations were active bottlenecks on different days.

3. Evaluation of VSL impact on motorway traffic

3.1. Average analysis

The investigation of the VSL impact on the aggregate traffic flow behavior is hardly possible on the basis of individual one-day data, because the critical occupancy areas, which are of most interest, are frequently not covered with no-VSL data. In addition, various stochastic effects and weather conditions may distort the VSL impact, particularly with respect to the possible existence and location of cross points between no-VSL and VSL curves. In order to investigate the effect of weather conditions on the shape of the flow-occupancy diagram (with and without VSL), we performed three kinds of global curve-fitting by location, using:

- All traffic data from all 27 representative days.
- The traffic data from the 18 representative days with “clear” weather conditions.
- The traffic data from the 4 representative days with “rainy” weather conditions.

Figures 4-7 present the curve-fitting results for different locations, each time for the case where all 27 representative days’ traffic data are used (upper plots of each figure), for the case where “clear” weather traffic data is used (middle plots) and for the case where “rainy” traffic data is used (lower plots). Figures 4 and 5 refer to A-direction locations, whereby the location 441A of Figure 4 is near the typical bottleneck of this direction; while the location 402A of Figure 5 is upstream of the typical bottleneck. On the other hand, Figures 6, 7 refer to B-direction locations, for which no typical active bottleneck location exists, i.e. active bottlenecks may appear at different locations on different days.

No-VSL data

By inspecting the no-VSL data on Figures 4-7 we may derive the following observations:

- The average flow capacity for both directions is slightly above 100 veh/min for “clear” days but reduces to around 10% lower values on “rainy” days.
- The average critical occupancy for both directions is around 15% for both “clear” and “rainy” days which indicates that the critical occupancy is little sensitive to changing weather conditions.
- Since the critical occupancy is quite stable under different weather conditions while the flow capacity is not, it may be concluded that the critical speed is also sensitive to weather conditions. This may also be derived from the provided diagrams by connecting the origin with the corresponding (q_{cap}, o_{cr}) -point and checking the slope of the corresponding connecting lines.

Figure 4: curve-fitting results for all 27 representative days (upper plot), “clear” representative days (middle plot) and “rainy” representative days (lower plot). Location 441A

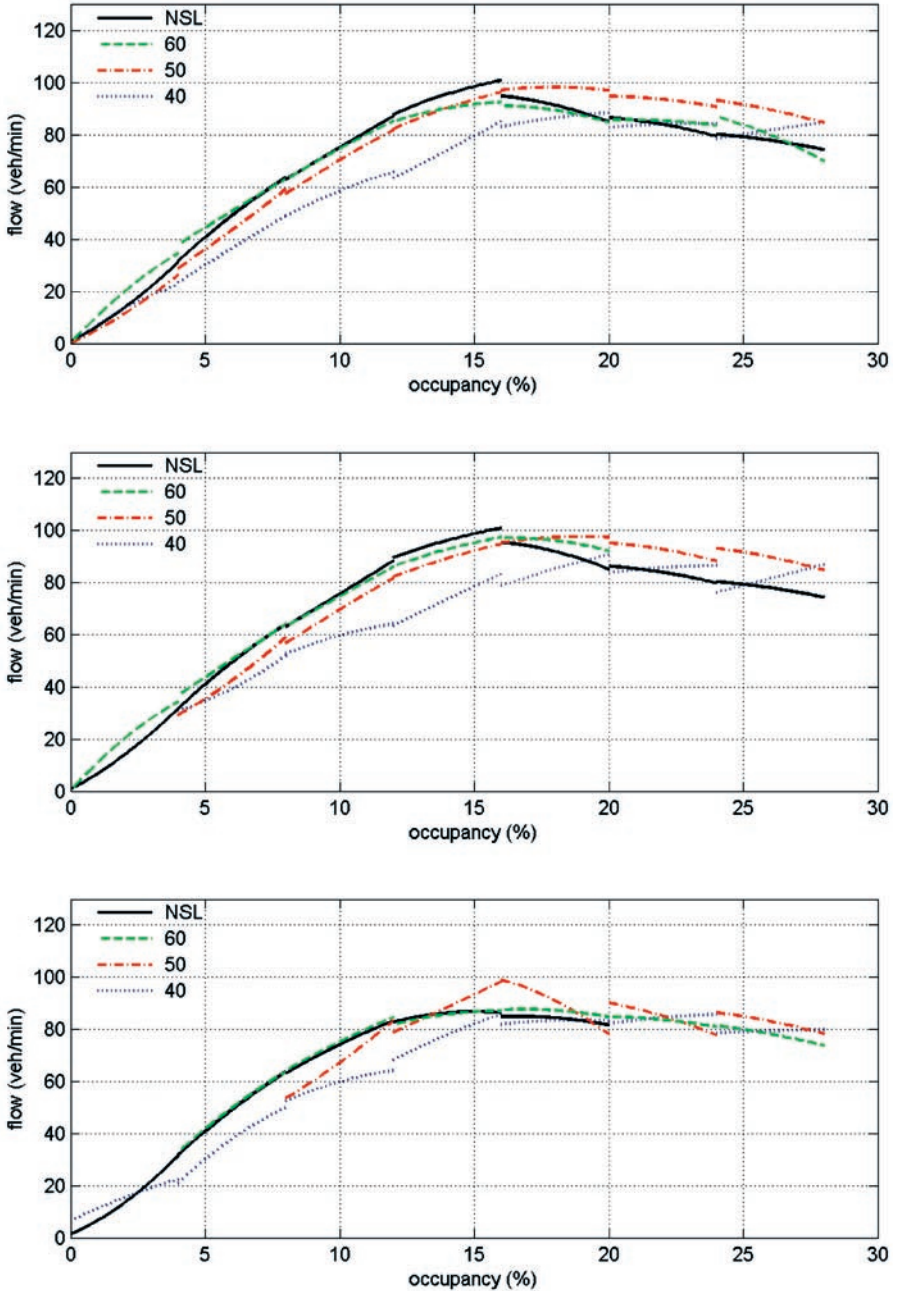


Figure 5: curve-fitting results for all 27 representative days (upper plot), “clear” representative days (middle plot) and “rainy” representative days (lower plot). Location 402A

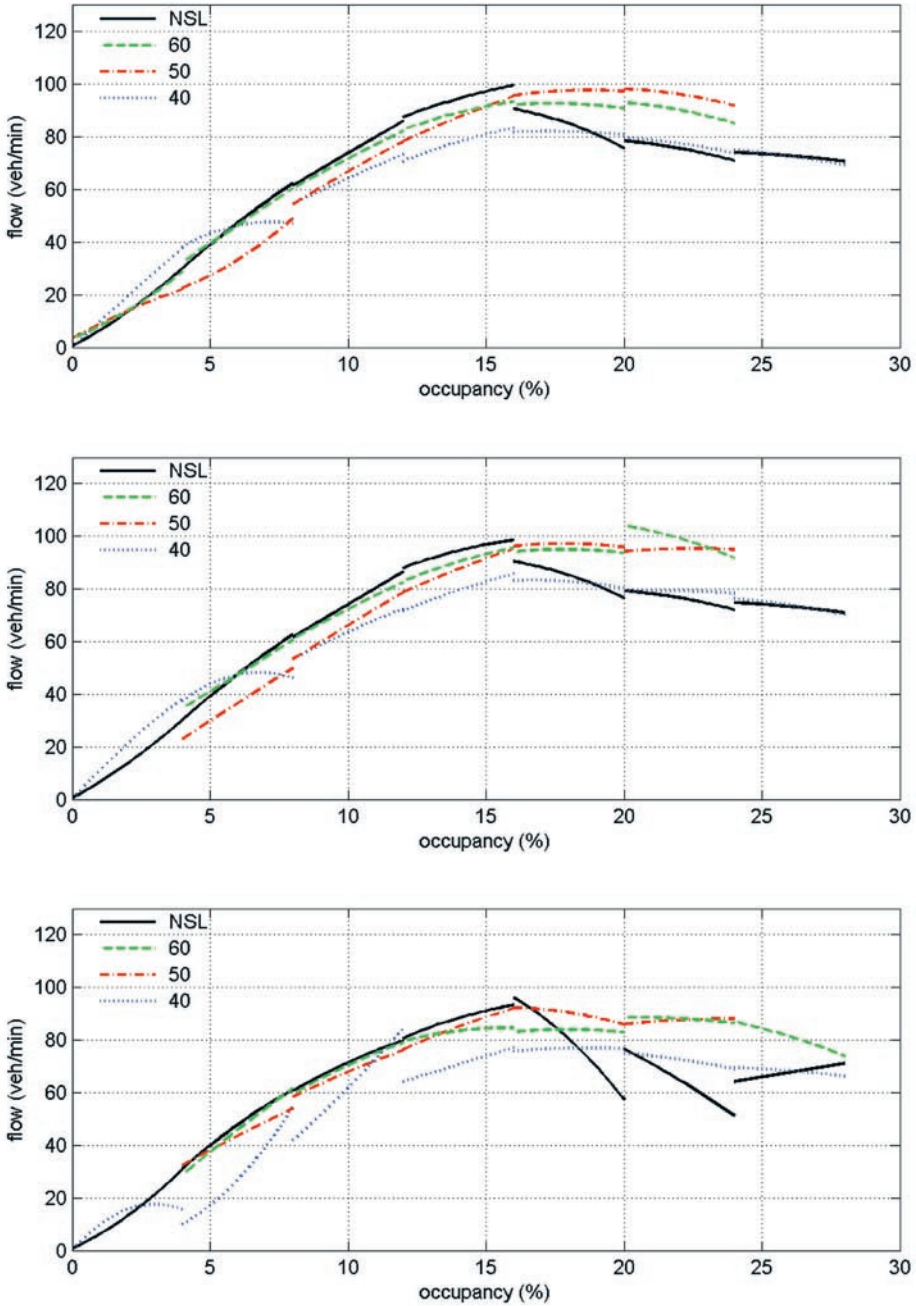


Figure 6: curve-fitting results for all 27 representative days (upper plot), “clear” representative days (middle plot) and “rainy” representative days (lower plot). Location 325B

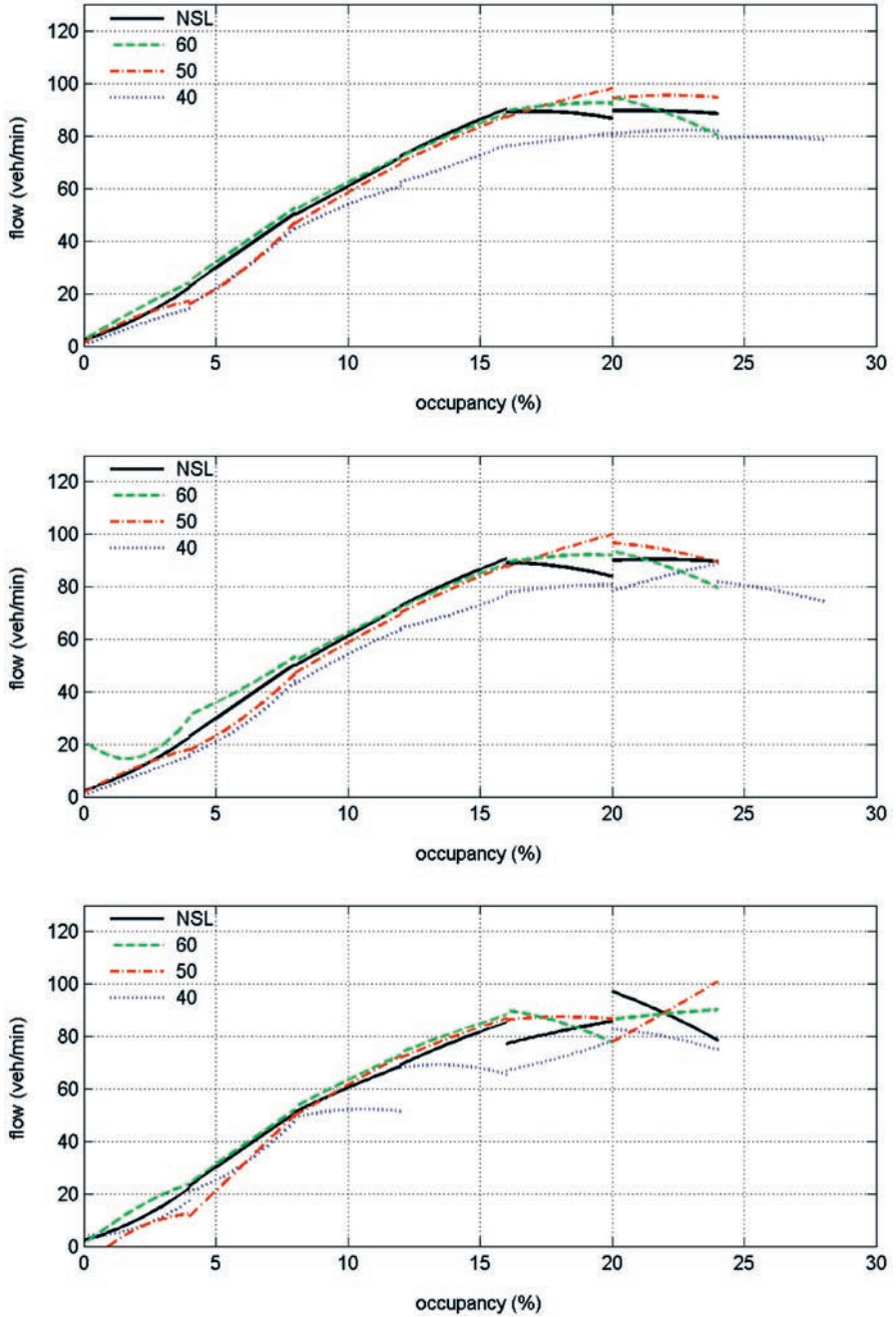
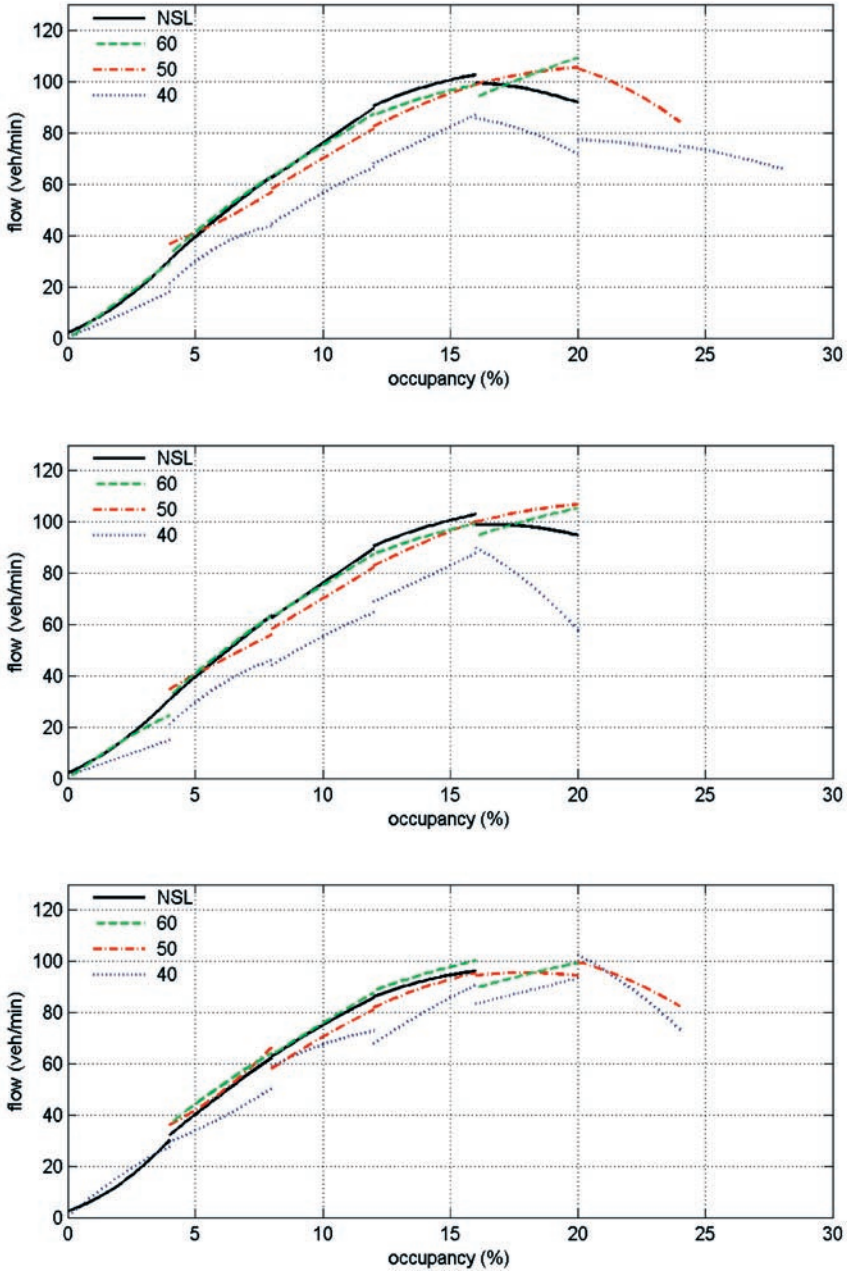


Figure 7: curve-fitting results for all 27 representative days (upper plot), “clear” representative days (middle plot) and “rainy” representative days (lower plot). Location 345B



- The average slope of the flow-occupancy diagram at clearly undercritical occupancies ($0 \leq o \leq 10\%$) is around 8 (veh/min/%) for most locations and does not change on “rainy” days. The latter observation indicates that any change of the flow-occupancy diagram due to weather conditions occurs only around the critical occupancy region; in other words, on “rainy” days, the flow-occupancy curve flattens more strongly near the (same) critical occupancy (leading to lower flow capacity) than on “clear” days.

Impact on slope at undercritical occupancies

Generally speaking, all diagrams confirm that the average impact of VSL on the flow–occupancy diagram at undercritical occupancies is a visible slope decrease, the intensity of which increases with decreasing VSL. This is in accordance with the expectations expressed earlier.

The following more detailed observations may be made regarding the VSL impact on the fitted curve’s slope at undercritical occupancies:

- The slope decrease appears to be slightly different at different locations; this might partly be attributed to the distance of the VSL-display gantry from the specific data location as well as to various stochastic effects.
- The slope decrease appears similar on “clear” and “rainy” days.
- The slope decrease due to VSL of 60 mph is virtually negligible at occupancies lower than 10%; for the critical occupancy region, the impact of VSL = 60 mph is relatively small.
- The slope decrease due to VSL of 50 mph or 40 mph is clearly visible at virtually all locations and all occupancy values. For undercritical occupancies in the range $0 \leq o \leq 10\%$, this decrease is from 8 (no-VSL case) to around 7 for VSL = 50 mph and around 6 for VSL = 40 mph.

Cross-points of curves

Various cross-points of fitted curves with and without VSL are observed at most displayed locations. All appearing cross-points are located at or beyond the critical occupancy of 15%. More specifically:

- Cross-points of no-VSL with 60 mph VSL curves (if any) do not seem to be of major significance for VSL control efficiency, even for direction A where the impact of 60 mph VSL is more visible in the critical occupancy area. It seems that 60 mph VSL could be considered as a preparatory step towards lower VSL, as it has limited impact on the aggregate traffic flow behavior for both “clear” and “rainy” days. Thus, a VSL of 60 mph could be switched on just before the critical occupancy is reached (e.g. at $o = 13\%$). There may be some merit though in setting 60 mph purely to reduce accidents, emissions or noise in certain areas.
- The 50 mph VSL seems to play a crucial role for VSL control efficiency. In fact, most displayed diagrams indicate a clear cross-point of the corresponding fitted curve with the non-VSL curve on both “clear” and “rainy” days. The

cross-point occurs generally at or slightly beyond the critical occupancy. At the right of the cross-point, the VSL-affected flows are clearly higher (for the same occupancy values) than the non-VSL flows, although this is not always clearly visible on “rainy” days due to stronger scatter of the related data.

- The 40 mph VSL is probably more important for traffic safety than for traffic flow efficiency. Although some cross-points may be observed at some diagrams, there is no clear flow increase visible at the right of those cross-points. Therefore, 40 mph VSL could be switched on at clearly overcritical occupancies (e.g. at $o = 25\%$) for traffic safety considerations.
- Regarding the potential increase of flow capacity due to VSL impact (as claimed in [1]), the analysis of the data is rather inconclusive. Some of the displayed diagrams (particularly for direction B) indicate a possible slight increase of capacity while others do not.
- The critical occupancy (at which we attain a zero-slope on the flow-occupancy diagram) is shifted to higher values under the VSL impact. In particular for VSL = 50 mph, the average critical occupancy is shifted quite consistently to around 20% (from 15% without VSL).

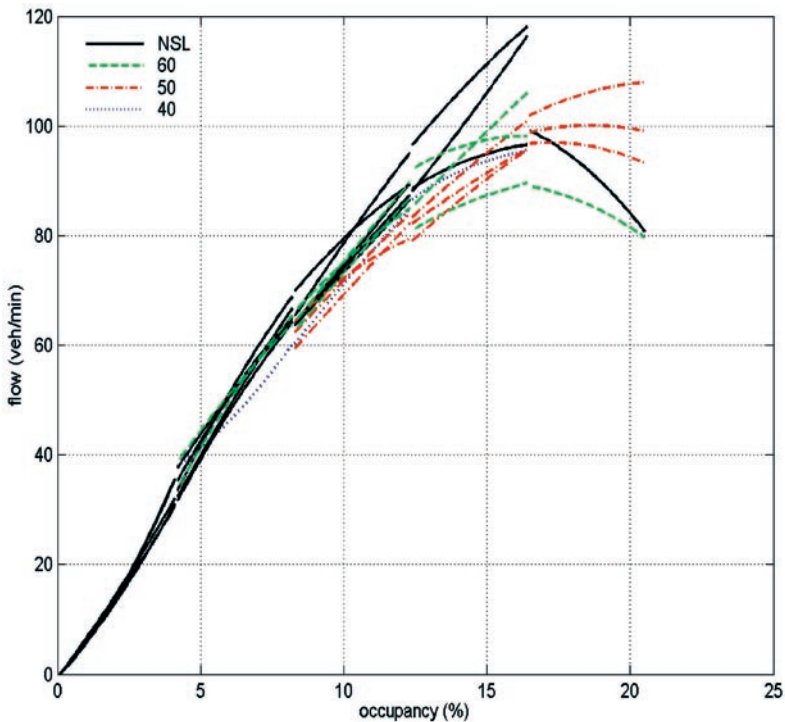
3.2. Daily impact consideration

As noticed earlier, the VSL-impact analysis based on individual daily data would suffer from fluctuations in aggregate traffic behavior from day to day which cannot be attributed to weather or other measurable conditions and is therefore classified as stochastic. This was the main reason for proceeding to an **average** impact analysis, whereby 27 days were considered simultaneously at each motorway location examined. It is, however, quite important to also examine the nature of the mentioned stochastic variations from day to day, along with their possible implications for the reliability of VSL control strategies. To this end, Figure 8 displays simultaneously various curves, each fitted individually on the basis of single-day data (“clear” days only), for the main bottleneck location 441A.

The following observations can be derived from Figure 8:

- Similar fitted curves on different days are observed for each VSL-value (including no-VSL) at clearly undercritical occupancies, i.e. $o \leq 10\%$, which indicates that the mentioned stochastic effects are not significant in this low-occupancy region. In contrast, the fitted curves for the same VSL-value but for different days are seen to diverge at occupancies higher than 10%.
- Flow capacity q_{cap} and critical speed v_{cr} are seen to vary from day to day. As a consequence, the usage of absolute values of flow and mean speed for VSL-control is not the best choice in the interest of control reliability and efficiency, as it may lead to “too early” or “too late” switching on and off of VSL, according to the day of application. Some operational control strategies rely mainly on flow thresholds with supplementary information provided by speed thresholds. These control strategies may need to be revisited in the light of these findings.

Figure 8: curve-fitting results for three individual “clear” days (location 441A)



- The critical occupancy is seen to be much less affected by stochastic effects; in fact $o_{cr} \approx 15\%$ is observable on all days displayed. Thus, if the critical occupancy for a motorway location is known, decision-making based on occupancy measurements and thresholds appears to be more reliable and efficient.

3.3. Implications for VSL control strategies

The variability of aggregate traffic flow behavior in the critical occupancy area, which was seen to occur even on days with similar weather and traffic demand characteristics, renders many currently operational VSL control strategies sensitive to the choice of the utilized flow and speed thresholds. In fact, due to the variability of traffic flow behavior, any choice of flow/speed thresholds may turn out, on different days, to be either aggressive or conservative with the risk of “too-early” or “too-late” activation.

Of course, an occupancy-based control strategy requires knowledge or accurate estimation of the location’s critical occupancy. This has the disadvantage – as in the case of flow and speed threshold selection – of requiring a considerable

effort of estimating the critical occupancy for each motorway location; moreover, since the critical occupancy, despite being more stable, may also change due to adverse weather conditions or truck percentage changes or incidents, etc., this approach might also be somewhat sensitive.

A real-time approach was proposed and evaluated recently for adaptively estimating the critical occupancy based on traffic measurements [8, 9]. The approach focuses on critical occupancy estimation for ramp metering purposes and was shown to estimate quite accurately the critical occupancy at locations with significantly different critical occupancies (ranging from 18% to 27%); the adaptive approach employs a slope estimator for the flow-occupancy diagram and uses the slope estimation to eventually come up with critical occupancy estimates.

The same approach – appropriately modified – could be used by a VSL control scheme to activate VSL around the critical occupancy. Since the focus of VSL control would be on occupancy-flow slope estimation at the **current traffic conditions** rather than on the critical occupancy estimation as for ramp metering purposes, only the slope estimation module of the algorithm presented in [8, 9], – appropriately modified – should be used within a VSL control strategy. Some encouraging preliminary testing results for a potential VSL control strategy based on real-time slope estimation are presented in [2]. The results indicate a pertinent decision making for VSL switching (not “too early”, not “too late”), virtually without a need for calibration.

The reason why a slope-based decision procedure does not require a tedious threshold calibration for different sites, different weather conditions, etc. is that, whatever the site, weather, and further (stochastic) conditions, when the real traffic flow approaches the critical occupancy (or flow capacity) area, the slope of the flow-occupancy diagram will approach zero; thus the specification of thresholds for the slope appears easier and more general than the specification of thresholds for the absolute values of the traffic flow variables. In other words, whatever the values of q_{cap} , o_{cr} , v_{cr} of a specific location and VSL-curve, the slope of the flow-occupancy diagram in that area will necessarily approach zero.

Conclusions

The effect of Variable Speed Limits (VSL) on the aggregate traffic flow behavior (in form of the flow-occupancy diagram) was investigated on the basis of traffic data from a European motorway where a flow/speed threshold based VSL control algorithm is currently used. A main focus of the reported work is on verifying some long-held conjectures regarding the VSL impact on the shape of the flow-occupancy diagram.

The main findings of the reported investigations may be summarized as follows:

- Speed limits – when applied at undercritical occupancies – have the effect of decreasing the slope of the flow-occupancy diagram. Moreover, the smaller the imposed speed limit, the larger the decrease in the slope of the flow-occupancy diagram.
- The VSL-affected flow-occupancy curve crosses (at least for some VSL) the non-VSL curve, shifting the critical occupancy to higher values in the flow-occupancy diagram. The major cross-points were found to lie on or slightly beyond the non-VSL critical occupancy.
- Regarding the potential increase of flow capacity, the data analysis was rather inconclusive, as a slight increase is indeed visible at some locations while at other locations no increase can be observed.
- On “rainy” days, the non-VSL flow capacity and the critical speed are reduced by some 10%. The same quantities were observed to change from day to day (even for the same location) without any obvious reason (stochastic effects).
- In contrast, the non-VSL critical occupancy was found to be less sensitive to weather conditions and stochastic effects.
- The speed limit of 50 mph was found to be the main contributor to modified aggregate traffic flow behavior that could be exploited towards more efficient traffic flow. The 60 mph VSL was found to have a moderate impact; while the 40 mph VSL could be used at high occupancies in the interest of traffic safety rather than traffic flow efficiency.
- In view of changing real absolute thresholds of flow, mean speed, and, to a lesser extent, occupancy, a more robust and efficient algorithm might result if the decisions on VSL-activation are based on real-time estimated **slopes** of the flow-occupancy diagram. In other words, while, e.g. the motorway’s flow capacity value may vary from day to day, the fact that flow capacity values occur near a zero slope does not change.
- Quite importantly, the real-time slope estimation algorithm is expected to reduce the effort of fine-tuning the location-dependent VSL control algorithm thresholds.

The reported results are likely to be representative for other sites in a qualitative sense as they are largely conform to sparse previous observations and expectations. On the other hand, quantitative figures may be different at different sites, similarly to the no-VSL fundamental diagrams.

Generally speaking, the investigation of the VSL impact on aggregate traffic behavior is a difficult endeavor because data collected during specific VSL value applications may not cover the whole range of possible occupancies. Moreover, the critical occupancy area may be covered with data for all applied VSL values, but these data may reflect different traffic conditions; for example, no-VSL data in the critical occupancy area are more likely to correspond to non-congested conditions; while 40 mph VSL data in the same area are likely to correspond to

overcritical traffic conditions, simply because 40 mph VSL are usually not displayed at undercritical traffic conditions. In other words, the specific VSL control strategy that was active during data collection may have some impact on the derived results.

References

- [1] Zackor, H. Speed Limitation on Freeways: Traffic-Responsive Strategies. In *Concise Encyclopedia of Traffic and Transportation Systems*, M. Papageorgiou, Editor, Oxford, UK, 1991, pp. 507–511.
- [2] Papageorgiou, M., E. Kosmatopoulos, M. Protopappas, and I. Papamichail. *Effects of Variable Speed Limits (VSL) on Motorway Traffic*. Internal Report No. 2006-25, Dynamic Systems and Simulation Laboratory, Technical University of Crete, Chania, Greece, 2006.
- [3] Zackor, H. Beurteilung verkehrsabhängiger Geschwindigkeitsbeschränkungen auf Autobahnen. *Strassenbau und Strassenverkehrstechnik*, Vol. 128, 1972, pp. 1–61.
- [4] Cremer, M. *Der Verkehrsfluß auf Schnellstrassen*. Springer, Berlin, Germany, 1979.
- [5] Smulders, S. Control of Freeway Traffic Flow by Variable Speed Signs. *Transportation Research Part B*, Vol. 24, 1990, pp. 111–132.
- [6] Hegyi, A. *Model Predictive Control for Integrating Traffic Control Measurers*. PhD Thesis, TRAIL Thesis Series T2004/2, Delft University of Technology, The Netherlands, 2004.
- [7] Allaby, P., B. Hellinga, and M. Bullock. Variable speed limits: Safety and operational impact of a candidate control strategy for an urban freeway. *Proc. IEEE Conference on Intelligent Transportation Systems*, Toronto, Canada, 2006, pp. 897–902.
- [8] Smaragdis, E., M. Papageorgiou, and E. Kosmatopoulos. A Flow-Maximizing Adaptive Local Ramp Metering Strategy. *Transportation Research Part B*, Vol. 38, 2004, pp. 251–270.
- [9] Kosmatopoulos, E., M. Papageorgiou, D. Manolis, J. Hayden, R. Higginson, K. McCabe, and N. Rayman. Real-Time Estimation of the Critical Occupancy for Maximum Motorway Throughput”, In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1959, TRB, National Research Council, Washington, D.C., 2006, pp. 65–76.

Real-time merging traffic control with applications to toll plaza and work zone management

**Markos Papageorgiou, Ioannis Papamichail, Anastasia Dimitriou
Spiliopoulou, Antonis Fanouriou Lentzakis**

Dynamic Systems and Simulation Laboratory

Technical University of Crete

73100 Chania, Greece

markos@dssl.tuc.gr; ipapa@dssl.tuc.gr; natasa@dssl.tuc.gr;

aflentzakis@gmail.com

From Transportation research. Part C, Emerging technologies, 2008, vol. 16, n°5, pp. 535-553; reprinted with the permission of Transportation Research part C.

Abstract

There is a number of highway transportation infrastructures where traffic flow arriving on a higher number of lanes must merge into a lower number of lanes within a limited space, e.g., merging highways, merging on-ramps, toll plazas, work zones. When the arriving traffic flow exceeds the downstream capacity, congestion is formed, and the outflow is reduced due to the congestion-induced capacity drop. A general conceptual framework for real-time merging traffic control is developed in the paper, aiming at throughput maximization or, equivalently, delay minimization. To this end, an algorithm known from local ramp metering operations is employed (ALINEA). The proposed framework is applied to new and innovative traffic control applications for real-time toll plaza and work zone management. Microscopic simulation results demonstrate the potentially achievable benefits. Further significant applications of the developed concept are outlined.

Introduction

There is a number of highway transportation infrastructure facilities that call for merging of traffic flow from a higher number of lanes into a lower number of lanes within a limited space, including:

- (i) Merging of two highways or freeways

- (ii) Merging of on-ramps into the freeway mainstream
- (iii) Merging within a toll plaza infrastructure
- (iv) Merging of freeway mainstream lanes due to road works.

When the total arriving merging flow is lower than the flow capacity of the downstream reduced-lane highway, the merging efficiency is usually satisfactory and there is no need for external intervention. If the total arriving flow reaches or exceeds the downstream capacity, congestion is created in the merging area and the exiting flow is reduced below the downstream capacity due to the capacity-drop phenomenon. Under these conditions, real-time control of the arriving flow streams may be employed to guarantee a maximum exit flow from the merge area. This is indeed practiced in ramp metering or freeway-to-freeway control installations (Papageorgiou and Kotsialos, 2002) where a part of the arriving traffic flow is controlled in order to maximize the merge area throughput.

This paper proposes a general conceptual framework for real-time merging traffic control, whereby metering may be applied to all or a part of the arriving streams, via suitable modification and use of ramp-metering like algorithms, in particular ALINEA (Papageorgiou, *et al.*, 1991, 1997). The proposed framework is then applied to new and innovative traffic control applications, e.g., for real-time toll plaza management or real-time work zone management on freeways. Section 1 presents the details and options of the general real-time merging traffic control framework, while sections 2 and 3 present specific instances of the general framework for toll plaza and work zone management, respectively, along with microscopic simulation results that demonstrate the efficiency of the proposed approach. Section 4 summarizes the conclusions and outlines a further, potentially most significant application of the general concept to combined on-ramp/mainline metering.

1. Real-time merging traffic control framework

1.1. Structure and elements

A general real-time merging traffic control system (Figure 1) comprises a number of general elements, each of which may take a number of different forms or options:

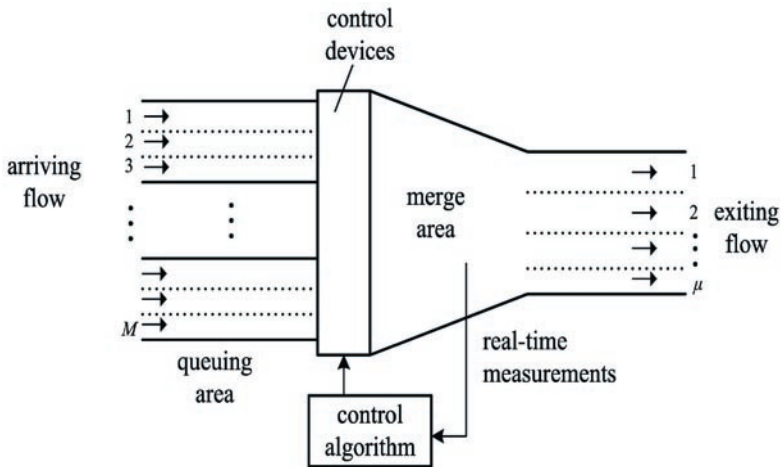
- (i) Arriving flow and queuing infrastructure (total of M lanes)
- (ii) Exiting flow infrastructure (μ lanes, $\mu < M$)
- (iii) Merge area
- (iv) Control devices
- (v) Real-time measurements or estimates
- (vi) Control algorithm.

Each of these elements will be discussed in the following with regard to its characteristics and possible options.

1.2. Arriving flow and queuing infrastructure

The arriving traffic flow approaches the merge area on a total number M of lanes which may belong to one single highway (as, e.g., in work zones) or to more than one highway (as, e.g., at freeway on-ramp merges or highway convergence areas).

Figure 1: a general real-time merging traffic control system framework with M entering and μ exiting lanes, $\mu < M$



If the total arriving traffic flow occasionally exceeds the downstream infrastructure capacity (see section 1.3), then queuing is inevitable. Within the queuing area (Figure 1), lane changing may be:

- Fully allowed, in which case it is reasonable to assume that the individual lane queues will tend to be balanced, provided the individual lane flows (and hence the queue waiting times) are similar to each other;
- Not allowed, in which case individual lanes may have queues (and waiting times) that are different from each other;
- Partly allowed, e.g., in the case of traffic arriving from more than one highways; vehicles may be allowed to change lanes within each arriving highway but cannot switch to lanes of another highway. In this case, queues (and waiting times) may be different on the lanes of each arriving highway.

Restrictions may be present with regard to the maximum admissible queue on some lanes. For example, ramp metering is usually not allowed to create over-long ramp queues that would interfere with the adjacent street traffic. Note that generalized restrictions for all M lanes may not be feasible if the arriving traffic flow is sufficiently high.

1.3. Exiting flow infrastructure

Although a most general assumption would be to allow for the possibility of more than one highways also at the exit of the merge area, this may lead to some additional problems that would need to be addressed; since we cannot think of any real merging infrastructure of this kind, we restrict our framework to one single exit highway with μ lanes and corresponding flow capacity q_{cap} . Obviously we must have $\mu < M$ for merging. Moreover, it is assumed that the exit flow capacity is (occasionally) lower than the maximum total arriving flow; otherwise merging traffic control is not needed.

1.4. Merge area

The vehicles arriving from a total of M lanes, change lanes appropriately within the merge area in order to fit into the μ lanes of the exit. The merging procedure may be quite complex, and drivers adapt their lane-changing behavior to the specific requirements of each kind of merging infrastructure. For example, merging conditions may be “symmetric” (if all vehicles have the same lane-changing rights and opportunities) or (more or less) “asymmetric”. “Asymmetries” may arise, e.g., because:

- Some arriving lanes have priority over others, e.g., freeway mainstream lane vehicles have priority over on-ramp vehicles (or vice versa in some infrastructures as the Paris Boulevard Périphérique); nevertheless, it is quite typical that some mainstream vehicles change from the right-most lane to other mainstream lanes in order to avoid potential conflicts with vehicles merging from the on-ramp.
- Some arriving lanes may prolong (explicitly or implicitly) into exit lanes, hence vehicles on these lanes may not need to be actively involved in lane-changing maneuvers.

Asymmetries may also characterize the infrastructure geometry; while toll plaza merging areas are typically symmetric (trapezoidal, see Figure 3), freeway on-ramp merging or work zone merging areas are usually geometrically asymmetric. Further differences concern the size (surface) of the merge area in different infrastructure types, e.g., toll plaza merge areas are typically bigger than freeway work zone merge areas.

The variability and complexity of diverse merging infrastructure characteristics that are reflected in correspondingly diverse driver lane-changing behaviors, is a challenge for microscopic models and simulators. Indeed, the lane-changing component is a notorious weak point of microscopic simulators when applied to various merging traffic infrastructures (Kondyli, *et al.*, 2007).

Fortunately, the complexity and variability of microscopic merging behaviors does not reflect on the real-time merging control actions envisaged here, because the basic assumptions that justify a common real-time control approach for any merging infrastructure are of a macroscopic nature and are related to the well-known notion of a fundamental diagram. Figure 2 displays a typical flow-density

Figure 2: fundamental diagram of a merging area

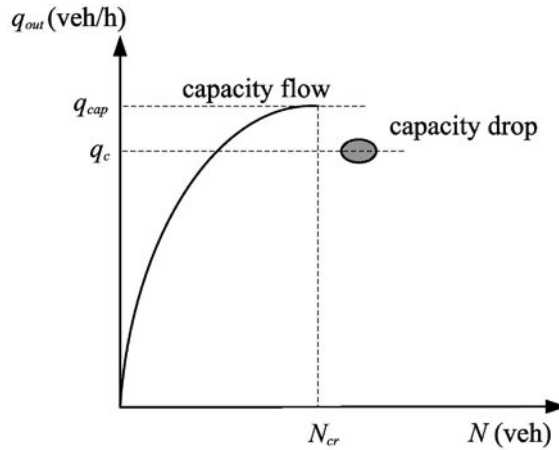


diagram for the merge area where the flow q_{out} is the exit flow and N is the number of vehicles included in the merge area (or in the downstream part of the merge area, see section 1.6). When N is small, merging conflicts are scarce and swift, while the exit flow is correspondingly low. As N increases, merging conflicts may increase, but q_{out} increases as well until, for a specific value N_{cr} , the exit flow reaches the downstream capacity q_{cap} . If N increases beyond N_{cr} merging conflicts become more serious, leading to substantial vehicle decelerations and eventual accelerations that reduce the exit flow to lower values q_c , where $q_{cap} - q_c$ is the capacity drop due to congestion.

The phenomenon of the capacity drop has been repeatedly observed and reported, mostly for freeway/on-ramp merge areas, see, e.g., Cassidy and Rudjanakanoknad, 2005. More specifically, it was observed that traffic flow in merge areas may increase up to some capacity value q_{cap} at relatively high mean speeds; if the total arriving flow continues to increase beyond this value, a traffic breakdown occurs with dropping speeds and increased occupancy. The measured flow q_c downstream of the formed congestion head (congestion discharge rate) was then found to be lower than q_{cap} by some 5-20% in different investigations. It should be noted that the precise reasons behind (and factors affecting the extent of) the capacity drop are not well understood, see, e.g., Treiber, *et al.*, 2006, for a potential explanation.

In our view, the capacity drop appears as a result of vehicle acceleration at the congestion head, i.e., from lower speeds within the congestion to higher speeds downstream of the congestion head. More specifically, individual vehicle accelerations are quite different from each other (due to different driver attitudes or different vehicle capabilities, e.g., trucks etc.); as a consequence, “over-long” gaps are likely to be created in front of low-acceleration vehicles due to faster accelerating vehicles ahead; “over-long” in the sense that the same vehicles with the

same speed would apply shorter gaps under stationary (non-accelerating) conditions. Filling of the over-long gaps via lane-changing from other lanes on multilane highways is unlikely; first, because lane-changing of vehicles is less likely in the acceleration mode; and, second, because over-long gaps are more probable on the slow lanes that are not a target for overtaking vehicles. In conclusion, longer gaps (lower density) at the same speed corresponds to lower flows during acceleration than at stationary conditions, and this may indeed be a reason for the observed capacity drop. If this view is correct, then the level of the capacity drop may depend on the speed gradient (upstream vs. downstream of the congestion head). Moreover, this view suggests that a capacity drop should be observable even at bottlenecks without lane drop, e.g., in front of tunnels etc., and there are some indications in the literature that this is actually happening.

Real data analysis from freeway on-ramp merge areas revealed that traffic breakdown may occur at different flow values q_{cap} on different days, even under similar environmental conditions (Eleftheriadou, *et al.*, 1995; Lorenz and Eleftheriadou, 2001; Cassidy and Rudjanakanoknad, 2005). Naturally, flow capacity differences become even more pronounced in case of adverse environmental conditions (Keen, *et al.*, 1986). In contrast, the critical occupancy (at which capacity flow occurs) was found to be fairly stable (Cassidy and Rudjanakanoknad, 2005) even under adverse environmental conditions (Keen, *et al.*, 1986; Papageorgiou, *et al.*, 2006). These findings suggest the introduction of a control scheme that would act on the arriving traffic flow in order to maintain (whenever necessary) the number of vehicles N in the merge area close to its critical value N_{cr} , thus maximizing the exit flow q_{out} , i.e., optimizing the merge infrastructure efficiency. This is a usual ramp metering practice, e.g., by use of the local ramp metering strategy ALINEA (Papageorgiou, *et al.*, 1991; 1997).

It should be emphasized that the phenomenon of capacity drop is crucial for throughput increase via the proposed merging traffic control measures. If no capacity drop appears in a merging traffic infrastructure, the only justification for applying merging traffic control might be the potentially improved traffic safety due to more orderly vehicle merging; improved traffic safety was indeed observed in some ramp metering installations, see, e.g., Papageorgiou *et al.*, 1997.

1.5. Control devices

There are several control devices that may be used to regulate the flows arriving in the merge area. Each control device may apply to one individual lane or to a subset of lanes:

- Traffic lights are the most popular devices to regulate arriving flows. Traffic lights are used for ramp metering, but also in some toll plaza management systems, e.g., at the San Francisco-Oakland Bay Bridge, California. Traffic lights may be applied to individual lanes or groups of lanes.
- Physical moving barriers may be used for the control of individual lanes, particularly at toll plaza booths.

- Variable speed limits may be used to regulate the mainstream traffic flow in freeway on-ramp merge areas. This may be achieved via ordinary road-side variable message signs or via emerging dual vehicle-to-infrastructure communication systems that act directly on individual vehicle speeds.

Whenever the existing infrastructure allows, the control devices (e.g., traffic lights) should be positioned sufficiently far from the actual merging area in order to allow for vehicles to accelerate sufficiently for efficient merging.

It should be emphasized that the control algorithm for merging traffic control (section 1.7) calls for implementation of a specific flow q_i for each controlled entering lane i , and the control devices should be operated appropriately to this end. Section 2 presents a possible way of achieving this by use of traffic lights. On-going research work considers the appropriate application of variable speed limits or emerging dual vehicle-to-infrastructure systems.

In some cases, control devices may be installed only on a part of the entering lanes (e.g., in the case of ramp metering) while traffic flow on other lanes is allowed to enter the merge area freely (e.g., by-pass lanes for HOV, buses, emergency vehicles). In other cases, different control devices may be applied to different groups of entering lanes, e.g., traffic lights at the on-ramp and variable speed limits for the mainstream of a freeway on-ramp merge area.

1.6. Real-time measurements or estimates

Application of feedback control in order to maintain the number of vehicles N close to N_{cr} (section 1.4) calls for availability of real-time measurements or estimates of N . Although this quantity may be directly measurable by use of video sensors, this possibility may be costly or difficult, particularly if the merge area is extended, as in toll plazas. An alternative possibility would be the real-time estimation of N by use of a limited number of ordinary loop detectors (Vigos, *et al.*, 2008).

Yet another possibility would be to consider the easily measurable occupancy o , rather than the number of vehicles N , as the variable under control. As a matter of fact, the diagram of Figure 2 is valid for occupancy as well (and occupancy rather than N is typically used for ramp metering operation). This raises the question on the specific location where the occupancy should be measured. The occupancy measurement should best be placed at or just upstream from the location where serious vehicle decelerations (congestion) appear first. This is because congestion, once occurred, propagates upstream, thus the measurements feeding the control algorithm should be placed:

- not downstream of the location where congestion appears first, otherwise the congestion is not visible and the control goal $o \approx o_{cr}$ may not be achievable;
- as close upstream from that location as possible, in order to minimize delays in the triggered control reaction.

Usually, merge congestions in freeway on-ramp merge areas appear first a few hundred meters downstream of the ramp nose, which is, in most cases, downstream of the acceleration lane drop. This is probably due to the fact that drivers are very attentive while merging from the acceleration lane onto the mainstream, thus tolerating relatively short inter-vehicle gaps at high speeds (which corresponds to accordingly high flows); after the acceleration lane drop, when drivers relax and attempt to restore their usual gaps, a traffic breakdown first occurs. In contrast, microscopic simulators create the merge congestion typically at the acceleration lane-drop location.

Placing the measurement detectors at or upstream from the typical location of first congestion appearance is crucial in order to enable the system to monitor any occurring merge congestion; related misunderstandings may lead to misapplication and control failure (Papageorgiou, *et al.*, 2007).

In conclusion, the functional merge area for pertinent merging traffic control may be different than the physical merge area. More specifically, the functional merge area for control may extend downstream to cover the location of first congestion appearance, which, in some cases, may be few hundred meters within the downstream highway, as in freeway ramp metering merges.

1.7. Control algorithm

The control algorithm receives the real-time measurements or estimates of N or o and drives the control devices in order to maintain $N \approx N_{cr}$ or $o \approx o_{cr}$ which maximizes the merge area exit flow. This is the major control goal, but there may be alternative and secondary control goals as discussed in this section. The overall control algorithm may be decomposed into three distinct parts, each with individual duties:

- a) Feedback control for exit flow regulation (ALINEA)
 - b) Distribution of entering flows
 - c) Translation of control decisions
- to be discussed in the following.

1.7.1. Feedback control

This is the central task of the control algorithm that is essentially identical for any kind of merge infrastructure. Feedback control is activated at each time interval T , whose value may be selected within the range [20 s, 60 s]. More specifically, at the end of each running period T , time-averaged measurements of occupancy o from the ending period (or the latest measurements or estimates of vehicle-number N) are used to calculate the entering flow to be implemented (via the control devices) in the next period in the aim of maintaining $o \approx o_{cr}$ (or $N \approx N_{cr}$). This may be achieved by use of the well-known integral-feedback (I-type) ALINEA regulator (Papageorgiou, *et al.*, 1991, 1996):

$$q(k) = q(k-1) + K_R[\hat{o} - o(k-1)] \quad (1)$$

where:

- $k = 1, 2, \dots$ is the discrete time index;
- $q(k)$ is the entering flow (veh/h) to be implemented during the new period k ;
- $o(k-1)$ is the last measured occupancy (%) averaged over all exit lanes;
- $K_R > 0$ is a regulator parameter;
- \hat{o} is the set (desired) value for the occupancy which may be set equal to o_{cr} for maximum exit flow.

The regulator (1) apparently integrates (accumulates) the regulation error $\hat{o} - o(k-1)$ from time-step to time-step. As a consequence, the flow $q(k)$ resulting from (1) could increase (or decrease) indefinitely under certain circumstances due to the well-known wind-up effect of I-type regulators that may be illustrated as follows. If $o < \hat{o}$ holds for a long period (e.g., during off-peak), then the regulation error $\hat{o} - o(k-1)$ in (1) is persistently positive and drives $q(k)$ to accordingly high values, while the actual inflow q_{in} is much lower due to lack of demand. When o eventually exceeds \hat{o} (e.g., at the start of the peak hour), the regulation error becomes negative, and (1) will gradually reduce q , starting from the highest value Q that it had reached during the period of positive regulation errors; but an actual reduction of the entering flow is only achieved when q becomes smaller than the current inflow q_{in} , and this may take many time-steps, whose number is obviously larger for larger Q (and lower K_R in (1)); during this transient period, o may continue to increase and an overshooting (i.e., a persisting negative regulation error $\hat{o} - o(k-1)$) is created, whose magnitude and duration increase with increasing Q , due to the corresponding increase of the transient period.

Thus, in order to avoid this wind-up effect, the flow $q(k)$ resulting from (1) is truncated if it takes values outside a pre-specified range $[q_{min}, q_{max}]$ and the truncated value is used in (1) as $q(k-1)$ in the next time step. The range $[q_{min}, q_{max}]$ should be sufficiently wide to enable flexible control around the usual operation value $q(k) \approx q_{cap}$ without frequent bound-hitting during the peak hours. On the other hand:

- q_{max} should not be too large in order to limit the duration of the aforementioned transient period and the magnitude of the resulting overshooting
- q_{min} should not be too small to avoid (temporarily) strong strangulations of the entering traffic flow.

The usual control goal in case of local control is outflow maximization (i.e. $q_{out} \approx q_{in} \approx q_{cap}$) which is achieved by the choice of $\hat{o} = o_{cr}$ (or $\hat{N} = N_{cr}$) in (1). It should be noted that occupancy, rather than flow, measurements are used to feed the regulator (1) because, as Figure 2 suggests, measured flows do not uniquely characterize the prevailing traffic state, since the same flow value may reflect either free or congested traffic conditions. In addition, according to the discussion of section 1.4, capacity values may differ from day to day; hence any selected

target \hat{q}_{cap} would either underload (on days where the actual capacity happens to be higher than \hat{q}_{cap}) or overload (on days where the actual capacity happens to be lower than \hat{q}_{cap}) the infrastructure.

In contrast, the critical occupancy o_{cr} (or N_{cr}), at which capacity flow occurs, appears to be more stable (section 1.4) and may be selected as a target, i.e., $\delta = o_{cr}$. The precise value of o_{cr} (or N_{cr}) may depend on the specific infrastructure and the measuring devices and can be identified experimentally. Alternatively, one may employ a real-time estimator for these quantities, e.g., as proposed in Smaragdis, *et al.*, 2004; Kosmatopoulos, *et al.*, 2006.

The algorithm (1) leads to a stable closed-loop behavior (Papageorgiou, *et al.*, 1991; Kosmatopoulos and Papageorgiou, 2003) for a broad range of positive K_R values. In the steady-state (stationary conditions), we have in (1) $q(k) = q(k-1)$ and hence $o(k-1) = \delta$, i.e., exact regulation. Due to the integral-feedback character of (1), stationary accuracy is achieved even in case of uncontrolled entering lanes or by-pass lanes for buses or HOV etc. Stationary accuracy is also guaranteed if there are errors in the implementation of the calculated q (e.g., due to traffic light operation inaccuracies, red-light violations), see Papageorgiou, *et al.*, 1991; Kotsialos, *et al.*, 2006.

It is interesting to note that the aforementioned wind-up effect could also be circumvented if the last calculated (and truncated) flow value $q(k-1)$ in (1) would be replaced by the real measured inflow $q_{in}(k-1)$. It may be shown, however, that the stationary regulator accuracy cannot be guaranteed in this case if there are errors in the implementation of the calculated q . Moreover, some small but unnecessary metering delays may appear during light traffic if the inflow values q_{in} vary strongly over time, see Papageorgiou, *et al.*, 1991; Kotsialos, *et al.*, 2006. On the other hand, replacing $q(k-1)$ in (1) by $q_{in}(k-1)$ would utterly suppress the transient period and the related overshooting.

The control results are little sensitive to the specific value of K_R within a broad range of values. As a rule-of-thumb, K_R in (1) should be higher than $25\text{-}\mu$ veh/h/%. Extremely high values of K_R may lead to unstable closed-loop behavior.

In some cases, the merging traffic control problem may be a component of a more comprehensive network-wide control scheme that prescribes specific outflows \hat{q}_{out} to the local control components. In this case, if the desired flow setpoint is sufficiently smaller than capacity flow (e.g., $\hat{q}_{out} \leq 0.9q_{cap}$), then the problem of ambiguous characterization of the traffic state is not an issue, and hence one may use the following flow-based regulator (FL-ALINEA; see Smaragdis and Papageorgiou, 2003) instead of (1)

$$q(k) = q(k-1) + K_F[\hat{q}_{out} - q_{out}(k-1)] \quad (2)$$

with $K_F = 1$.

In some cases, a portion of the M arriving lanes may not be controllable, e.g., the mainstream lanes in ramp metering or some by-pass lanes for HOV or buses. In these cases, q in (1) or (2) reflects the flow on the controllable arriving lanes only, while the uncontrollable flows act as disturbances that do not need to be taken into account in the control algorithm thanks to its feedback character.

1.7.2. Distribution of entering flow

The feedback regulators (1) or (2) deliver, at each time step T , the total flow values q to be implemented at the exit of the controllable entering lanes. How should this total flow be distributed among individual (or among groups of) controllable entering lanes ? This question cannot be answered in a general way as it may depend on various characteristics of the merging infrastructure and the pursued control goals. Here are some examples:

- i. Consider a toll plaza or work zone infrastructure where lane-changing is possible in the queuing area (section 1.2); an equal distribution of q among the entering lanes seems reasonable, as arriving drivers are likely to join the shorter queues, thus leading to a “spontaneous” equalization of the individual lane queues.
- ii. Consider a toll plaza or work zone infrastructure where lane-changing is not allowed or not possible in the queuing area (e.g., due to physical separation); it may then be reasonable to actively distribute q among the entering lanes in a way that equalizes the lane queues or the corresponding waiting times. This calls for a corresponding feedback control by use of real-time measurements or estimates of individual lane queue lengths.
- iii. Consider a freeway on-ramp merge area where the on-ramp flow is controlled via traffic lights (ramp metering) while the mainstream flow is controlled via variable speed limits. The waiting queues in the ramp are not allowed to exceed an upper limit (to avoid interference with the adjacent street traffic). Under these conditions there may be a variety of different policies on how to distribute q among the mainstream and the on-ramp, taking into account also the limitation of the ramp queue length. Clearly, each potential policy must be implemented via a suitable distribution algorithm.

1.7.3. Translation of control decisions

At the end of parts a) and b) at each time step T , each controllable entering lane (or each group of lanes with one common control device) has an assigned subflow q_i that needs to be implemented via appropriate activation of the corresponding control device. The specifics of this translation of the control decisions into control device actions cannot be specified in a general form. Even for the same control device, there may be different possible operational policies; e.g., for traffic lights we may have one-car-per green, n -cars-per-green, full traffic cycle, discrete release rates etc., see Papageorgiou and Papamichail, 2008.

This completes the presentation of the general framework for real-time merging traffic control. Two instances of the framework, one for real-time toll plaza

management and another for real-time work zone management will be presented in the following. More detailed results for both examples are reported in corresponding dedicated papers (Spiliopoulou, *et al.*, 2008; Lentzakis, *et al.*, 2008).

2. Real-time toll plaza management

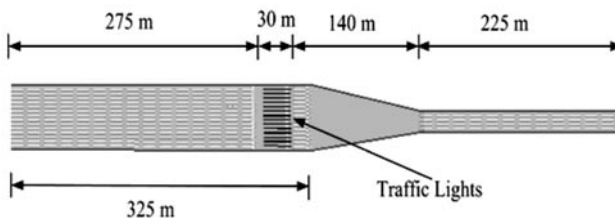
2.1. Infrastructure description

Toll plaza management aims at smooth, safe and efficient servicing of vehicles in the related facilities. Proposed management measures for toll plazas address lane configuration, signing, markings, geometric design, toll collection technologies (FHWA, 2006). In some rare cases, traffic lights may be also installed (e.g., in the toll plaza of the San Francisco-Oakland Bay Bridge, California). In this section, a simulated case study will be presented in order to demonstrate the potential benefits of real-time merging traffic control for maximum throughput.

Toll plazas contain usually a higher number of toll-paying booths than of exiting lanes. Thus, vehicles exiting the booths need to merge before reaching the exiting highway lanes. During peak hours, the rate of booth-exiting vehicles may be higher than the downstream infrastructure capacity q_{cap} , in which case a congestion may form that reduces the exit flow below q_{cap} (capacity drop). Based on the reasoning presented in section 1, it may be worthwhile to apply merging traffic control in order to maximize the toll plaza throughput, or, equivalently, minimize the average vehicle delays, or, equivalently, minimize the upstream extent of the formed queue.

The particular infrastructure considered in order to demonstrate, via microscopic simulation, the potential of real-time toll plaza management, mimics the Bay Bridge (California) toll plaza with $M = 15$ entering (booth) lanes and is sketched in Figure 3. Traffic arrives on 15 lanes and reaches, after 275 m, the booth area (30 m) which is not explicitly modeled (no booth delays) because it is of little relevance in the present context. After the booth area there is a trapezoidal merging area followed by the exit highway with 5 lanes and capacity $q_{cap} \approx 10\,500$ veh/h. Lane changing is possible anywhere except in the booth area.

Figure 3: a toll plaza infrastructure mimicking the Bay Bridge (California) toll plaza



The specific arrival demand scenario considered is stochastic; the average demand is increased in small steps, starting with 3000 veh/h (in an empty system) and reaching 12000 veh/h (which exceeds the downstream capacity q_{cap}) within the first 10 minutes; during the next 10 minutes, the average demand is maintained constant at this high value (peak hour); during time $t \in [20 \text{ min}, 30 \text{ min}]$, the demand is reduced step-wise back to 3000 veh/h and then drops to zero until the end of the investigation (40 min) in order to have an empty system again. This scenario includes a total of about 4400 veh served.

The feedback algorithm (1) is used for merging traffic control, albeit by use of real-time measurements or estimates of the number N of vehicles included in the trapezoidal merge area (Figure 3) rather than occupancy o . The regulator is activated every $T = 30 \text{ s}$, and the required bounds are set $q_{min} = 4\,500 \text{ veh/h}$ et $q_{max} = 13\,000 \text{ veh/h}$. The flow value q delivered by (1) is eventually distributed evenly among the 15 entering lanes by use of 15 traffic lights (one per lane) that are placed at the exit of the booth area, just upstream from the merge area (Figure 3). Each traffic light has a constant green phase of 4 s that allows in average some 2 vehicles to pass. A signal cycle c (in s) consists of a (constant) green phase and a (variable) red phase and allows some 2 vehicles to pass. Thus, the implementation of a flow q (veh/h) calls for a traffic cycle (equal for all lanes) that satisfies

$$q = (2 \cdot M / c) \cdot 3600 \Rightarrow c = 7200 \cdot M / q. \quad (3)$$

The red phase is not allowed to be lower than 2 s, hence there is a minimum cycle length $c_{min} = 6 \text{ s}$ and c resulting from (3) is truncated if it is lower than c_{min} . Note that c_{min} is never reached here since $q_{max} = 13\,000 \text{ veh/h}$ delivers $c = 8,3 \text{ s}$ in (3).

To avoid bursty entering flows that may appear if all 15 lanes assume their green and red phases simultaneously, the 15 lanes are subdivided in three groups of five lanes each, namely $\{1,4,7,10,3\}$ $\{2,5,8,11, 14\}$ $\{3,6, 9,12,15\}$, respectively, and a shift (offset) is introduced for the cycle start of each group relative to the cycles of the other groups.

Note that the regulator time step T is not an integer-multiple of the cycle c in general, which creates a synchronization problem that is solved as follows. As soon as $q(k)$ is delivered by the feedback algorithm, the corresponding cycle $c(k)$ is calculated via (3). This cycle is held for application at the end of the last (running) cycle $c(k - 1)$ of the previous period $c(k - 1)$. The resulting implementation error is deemed minor, particularly in view of the automatic disturbance-rejection property of (1) (see section 1.7).

The described infrastructure was simulated by use of the microscopic simulator AIMSUN NG Professional Edition 5.0.10 using the simulator's default parameters and a simulation time step of 0.75 s. The trapezoidal merging area (Figure 3) is modeled as a "junction", and this is actually a first reason for using measurements of N rather than of occupancy, as AIMSUN does not allow the placement of loop

detectors inside “junctions”. On the other hand, this allows us to test the feedback algorithm (1) with this measurement option as well.

The simulated merging control software is interfaced with the simulator via the AIMSUN API (Application Programming Interface) that allows us to emulate a real-time closed-loop operation. More specifically, the simulator delivers every T the current number N of vehicles in the junction, based on which the control software calculates the corresponding traffic light settings and returns them to the simulator for application.

The merging behavior of vehicles is a well-known weak point of microscopic simulators (Kondyli, *et al.*, 2007), particularly in the case of somewhat unusual merging infrastructures. Therefore, the produced simulation results should be understood as a demonstration of potential benefits achievable via merging traffic control, rather than as a quantitatively reliable assessment of achievable benefits. It should be emphasized that the macroscopic fundamental diagram “emerges” from the microscopic simulation of individual vehicle movements and is not imposed in any way in our simulations.

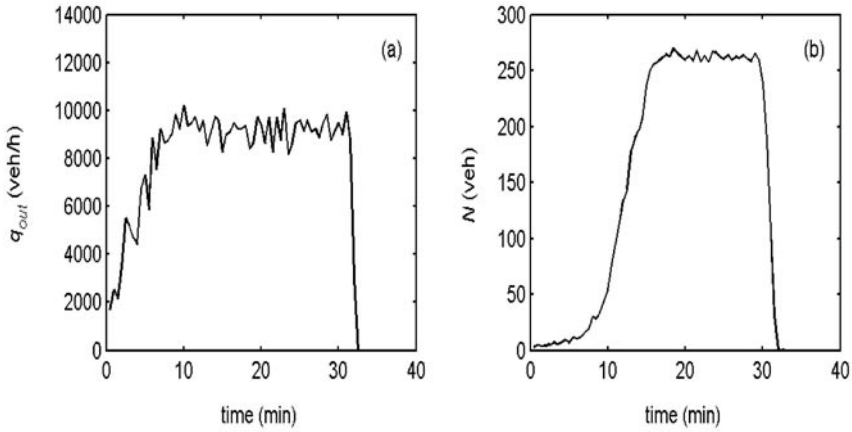
Since AIMSUN is a stochastic simulator, different runs with different seeds may produce quite different results. To address this issue, the usual practice is to run several simulations, each with different seed, for each simulated scenario. This practice is followed in the reported investigations with 10 replications per scenario. The main evaluation criterion used is the average vehicle delay (AVD) (in s/veh/km) that is delivered by AIMSUN. Note that vehicle delay expresses the actual travel time minus the (minimum possible) travel time that would be needed for a vehicle to travel a distance at free speed. The average vehicle delay (AVD) expresses the average delay-per-kilometer of a vehicle, taking into account all vehicles that have completed their trip during the simulation.

2.2. Simulation results

In the no-control case, the arriving vehicles continue their trip in the merge area and exit without noteworthy problems as long as the arriving demand is lower than the exit capacity q_{cap} . When the arriving demand approaches q_{cap} , serious vehicle merging conflicts with strong vehicle decelerations are observed that lead to the formation of a congestion close to the exit of the merge area; the congestion spills back onto the 15 entering lanes without ever reaching the simulated system entrance. The resulting mean AVD (in s/veh/km) for 10 replications is 164 with a standard deviation of 15.5, and highest/lowest AVD values of 190/137, respectively.

Figures 4a and 4b display the trajectories of the outflow q_{out} and of the number of vehicles N , respectively, for one particular simulation run with an individual AVD = 168 s/veh/km close to the mean AVD of the corresponding 10 replications. The outflow q_{out} is roughly following the arriving demand increase until about $t = 10$ min where it shortly approaches the capacity q_{cap} . The vehicle number N is increasing slowly (due to increasing flow) during the same period. After $t = 10$ min, the vehicle accumulation N in the merge area has a steep increase (due to the formed congestion) and stabilizes eventually at a value around 260 veh; at that

Figure 4: outflow q_{out} (a) and vehicle-number N (b) in the no-control case for the toll plaza example



state, the merge area is full of vehicles due to congestion. Soon after the first appearance of the congestion (e.g., after $t = 10$ min) the outflow reduces (capacity drop) to values around 9500 veh/h. When the demand is decreased and the formed queue is served, the system is emptied (at around $t = 32$ min).

When merging traffic control is applied, the traffic lights are operated with a short red phase (corresponding to $q_{max} = 13\ 000$ veh/h) for as long as N in the merge area is lower than the set value \hat{N} , according to (1). Note that some slight delays may occur to some vehicles that encounter the red phase, but there are no real queues in front of the traffic lights and the system outflow is mainly determined by the arriving demand (no real metering). The slight delays could be avoided if the traffic lights are switched on only when needed (i.e., when N approaches \hat{N}). As the arriving demand increases, N increases, and, when $N(k) > \hat{N}$, the regulator (1) starts its actual entering flow control operation aiming at maintaining $N(k)$ near \hat{N} . At this time, a queue is formed just upstream from the traffic lights which increases, without ever reaching the simulated system entrance.

As mentioned earlier, the ultimate goal of merging traffic control is outflow maximization which leads to minimization of the average delay and minimization of the upstream queue extent. Outflow maximization is achieved via an appropriate specification of \hat{N} in (1). To investigate this issue, a series of simulation runs were carried out with different \hat{N} -values and 10 replications for each such value; the utilized K_R value in (1) is set equal to $500\ h^{-1}$. Figure 5 displays some of the obtained results in terms of the average vehicle delay (AVD); more specifically Figure 5 displays, for each investigated \hat{N} -value in the range $[16, 50]$, the AVDs of the 10 simulation replications and the corresponding mean AVD value. For comparison, Figure 5 also displays (horizontal lines) the mean, highest and lowest AVD values of the 10 replications of the no-control case.

Figure 5: average vehicle delay versus \hat{N} with and without merging traffic control for the toll plaza example

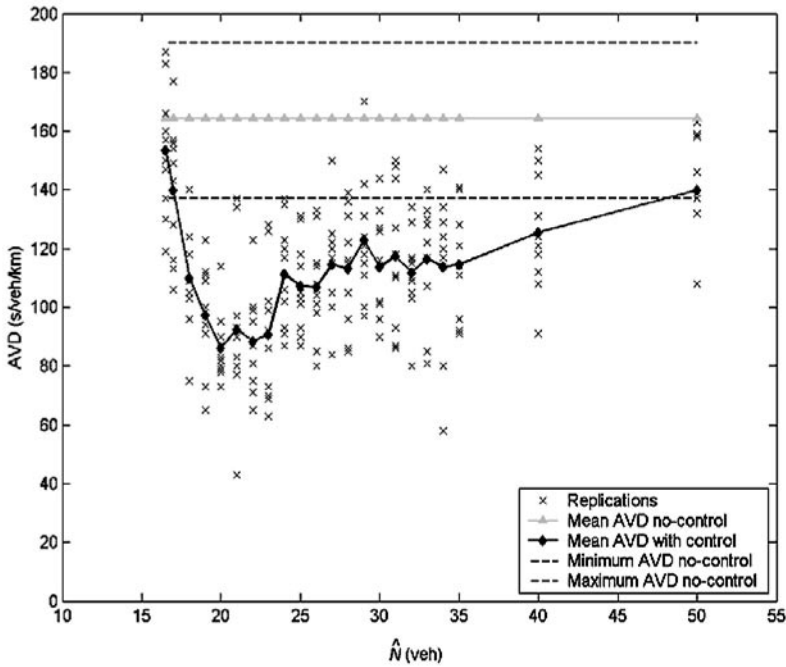
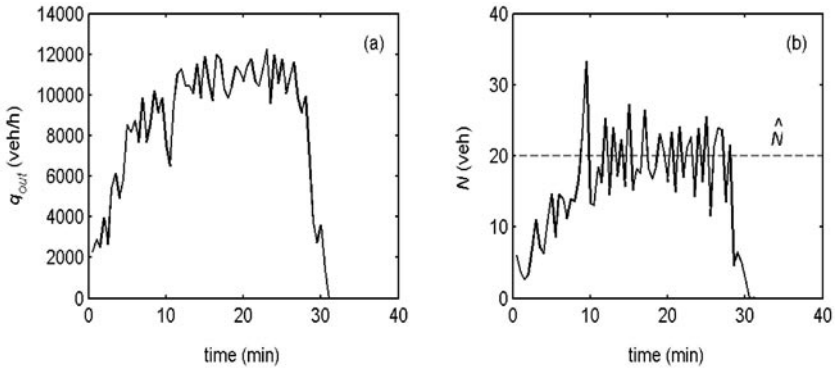


Figure 5 gives rise to the following observations:

- For \hat{N} values around 20 veh, the mean AVD is minimized, apparently due to outflow maximization. This corresponds to the N_{cr} -value mentioned earlier.
- For \hat{N} lower than $N_{cr} = 20$ veh, the mean AVD increases because the downstream highway “starves for flow”, i.e., the resulting q_{out} is (unnecessarily) restricted to values lower than $q_{cap}\pi$, and hence, the downstream highway capacity is underutilized.
- For \hat{N} higher than $N_{cr} = 20$ veh, the mean AVD increases due to lower q_{out} , but the latter has now a different background than for $\hat{N} < N_{cr}$ the higher \hat{N} (beyond N_{cr}), the more frequent the appearance of vehicle merging conflicts that lead to strong decelerations and reduced outflows.
- Although the mean AVD in the control case is clearly lower than the mean AVD of the no-control case for all $\hat{N} \in [16, 50]$, individual control replications may be worse than individual no-control replications. Merely in the range $\hat{N} \in [18, 26]$, there is virtually no control replication providing worse results than the best no-control replication.

Figure 6: outflow q_{out} (a) and vehicle-number N (b) in the merging control case with $\hat{N} = 20$ veh, $K_R = 500$ h⁻¹ for the toll plaza example



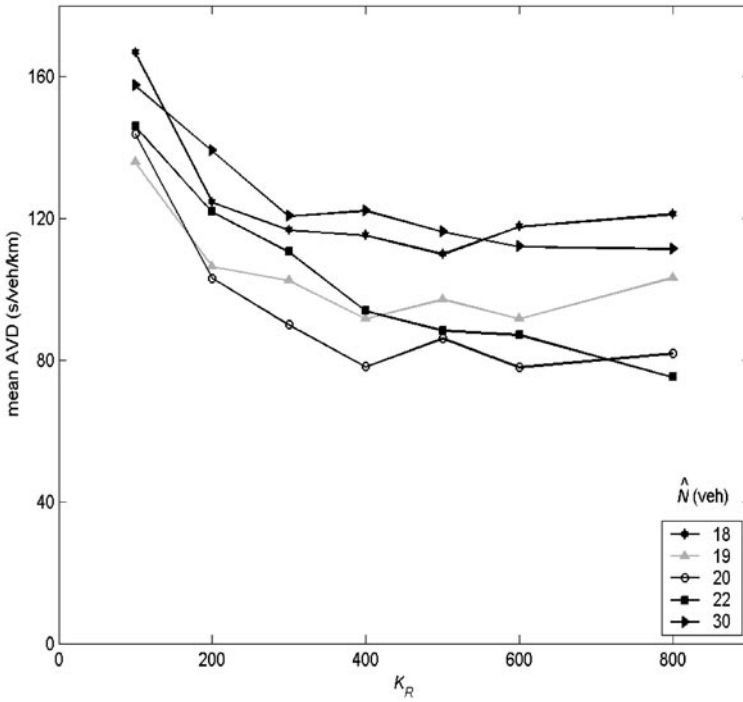
- The mean AVD for $\hat{N} = N_{cr}$ is 90 s/veh/km, which is 45% lower than the mean AVD of the no-control case.

Figures 6a and 6b display the trajectories of the outflow q_{out} and of the vehicle-number N , respectively, for one particular simulation run with $\hat{N} = 20$ veh and an individual AVD = 90 s/veh/km close to the mean AVD of the corresponding 10 replications. The outflow q_{out} is roughly following the arriving demand increase until about $t = 8$ min; at this time, N is seen to exceed \hat{N} (Figure 6b), and the regulator (1) is activated; for reasons mentioned in section 1.7, N exhibits an overshooting (up to $N \approx 35$ veh), followed by a short undershooting, before being stabilized by the regulator action to values around 20 veh. The relatively strong N -variations during the transient period $t \in [8 \text{ min}, 12 \text{ min}]$ reduce temporarily the outflow q_{out} ; in contrast, q_{out} is seen during the stationary control period $t \in [13 \text{ min}, 26 \text{ min}]$ to attain values around 10 500 veh/h, i.e., around the downstream highway capacity q_{cap} .

Comparing Figures 4 and 6, merging traffic control increases the toll plaza outflow by some 1000 veh/h or 10.5% compared to the no-control case during the stationary period. As a result, the formed queue for the system under control is emptied at around $t = 27$ min, i.e., some 5 minutes earlier than in the no-control case.

The reported merging traffic control simulation results were produced with a value $K_R = 500$ h⁻¹ in (1). It is important to investigate the sensitivity of the control results with respect to the value of K_R , since high sensitivity may imply a high effort of fine-tuning in field applications. Figure 7 displays the mean AVD values resulting from merging traffic control with several \hat{N} values, against K_R within the range [100, 800]. Note that each point in Figure 7 corresponds to a specific couple

Figure 7: mean AVD versus K_R for different set-values \hat{N} for the toll plaza example



of utilized K_R , \hat{N} values, and that 10 simulation replications were carried out for each such couple.

Figure 7 shows that the control results (in terms of the mean AVD) for all considered \hat{N} -values are virtually identical for $K_R \geq 300$. For $K_R < 300$, the initial overshooting/undershooting observed in Figure 6b become more pronounced (for reasons detailed in section 1.7), leading to temporarily lower outflows q_{out} during the transient period and hence increased mean AVD. In contrast, the control behavior for $K_R < 300$ is similar to the one displayed in Figure 6 after the transient period (i.e., in the stationary phase). Thus, if the transient period duration (some 5 minutes) is negligible compared to the whole peak period duration, then K_R values of 100 or 200 are virtually equally efficient as higher values. On the other hand, it may be shown that the closed-loop system becomes unstable beyond some (very high) value of K_R .

In conclusion, there is a broad range of K_R values, for which the merging traffic control behavior is virtually equally efficient, and hence the need for fine-tuning K_R in field applications is expected to be minor.

More details on toll plaza merging traffic control and related simulation results may be found in Spiliopoulou (2007) and Spiliopoulou *et al.* (2008).

3. Real-time work zone management

3.1. Infrastructure description

Work zones on freeways necessitate the drop of one or more lanes at the work zone entrance which may create traffic flow disruptions and congestion upstream from the work zone due to reduced capacity. Work zone management aims at smooth, safe and efficient passage of vehicles, with particular focus on the work zone entrance where vehicles need to merge into fewer lanes. Proposed management measures for work zones address speed limitations as well as signing, markings, geometric design (FHWA, 2005). In this section, a simulated case study of a hypothetical work zone will be presented in order to demonstrate the potential benefits of real-time merging traffic control for maximum throughput.

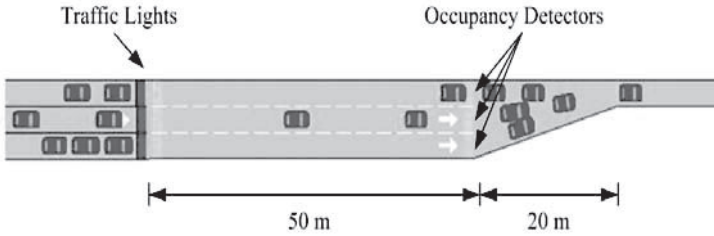
Vehicles approaching the work zone need to merge into a lower number of lanes. If the rate of arriving vehicles exceeds the capacity q_{cap} of the remaining lanes in the work zone, a congestion will form and the entering flow will be reduced below q_{cap} (capacity drop). Based on the reasoning of section 1, it may be worthwhile to apply merging traffic control in order to maximize the work zone throughput. Clearly, a higher throughput through the work zone is associated with lower vehicle delays and reduced congestion extent upstream from the work zone entrance.

The particular hypothetical infrastructure considered in order to demonstrate, via microscopic simulation, the potential of real-time work zone management, has $M = 3$ entering and $\mu = 1$ leaving lanes as sketched in Figure 8. The total length of simulated freeway is 750 m while the trapezoidal merging area is 20 m long, followed by the work-zone restricted one lane freeway portion with a capacity $q_{cap} \approx 2300$ veh/h.

The specific arrival demand scenario is stochastic; the average demand is increased in small steps, starting at level zero (in an empty freeway) and reaching 2500 veh/h (which exceeds q_{cap}) within the first 10 minutes; during the next 10 minutes, the demand is maintained constant at this high value; while, in the third 10-min period, the demand decreases step-wise back to zero and is maintained there until the end of the investigation (40 min) in order to have an empty system again. This scenario includes a total of about 830 veh passing.

The feedback algorithm (1) is used for merging traffic control with $T = 30$ s, $q_{min} = 1000$ veh/h, $q_{max} = 3000$ veh/h. The feedback algorithm is fed with occupancy measurements collected just upstream from the trapezoidal merge area (Figure 8); note that possible congestion due to merging conflicts starts within the merging area in this simulated environment; thus, any appearing congestion moves upstream and is captured soon by the emulated loop detectors. In a real environment it would be probably better to place the loop detectors within the

Figure 8: a hypothetical work zone entrance



merging area (upstream from the typical location of first congestion appearance) for faster congestion capture and corresponding reaction; however, the merging area was modeled here as a “junction”, and AIMSUN does not allow for the placement of loop detectors within “junctions”.

The flow value q delivered by (1) is eventually distributed evenly among the three entering lanes by use of three respective traffic lights that are placed 50 m upstream from the merge area (Figure 8). The traffic lights are operated as described in section 2.1 with a constant green phase of 4 s and a variable red phase. A shift (offset) is applied to the signal cycle start of each traffic light relative to the cycles of the other traffic lights in order to minimize simultaneous vehicle departures. The use of traffic lights in work zones was recently proposed by Wei and Pavithran (2006), albeit with fixed signal settings rather than the real-time control considered here.

The real-time work zone management example considered assumes that all upstream lanes are signal-controlled for reasons of equity and safety. In case of partial metering (e.g., metering of the merging lanes only), traffic safety could be affected due to unbalanced upstream queues motivating drivers to frequent lane changes towards the free lanes; while drivers preferring to stay in their (metered) lanes would suffer higher delays than drivers on free lanes. Anyhow, as mentioned earlier, the presence of uncontrolled lanes (e.g., for buses, HOV etc.) does not pose any problem to the presented method.

Simulations are carried out with AIMSUN NG Professional Edition 5.1.1 as described in section 2.1, with 10 simulation replications for each investigated scenario, and the average vehicle delay (AVD, in s/veh/km) is used as the main evaluation criterion.

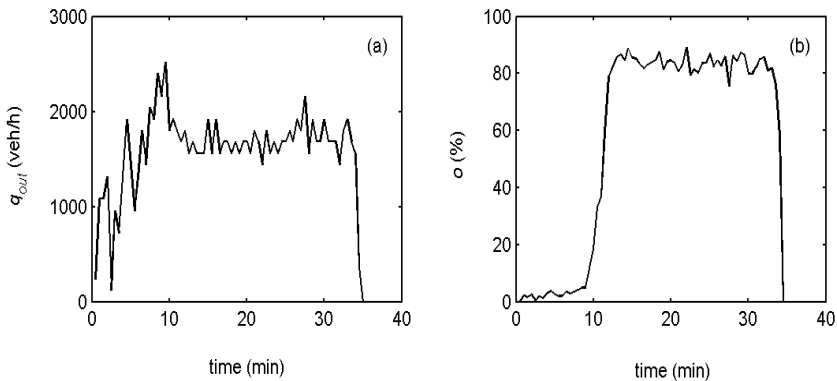
3.2. Simulation results

In the no-control case, arriving vehicles continue their trip in the merge area and the work zone area without noteworthy problems as long as the arriving demand is lower than q_{cap} . When the arriving demand approaches q_{cap} , serious vehicle merging conflicts lead to the formation of a congestion in the merge area; the congestion spills back onto the three entering lanes without ever reaching the simulated system entrance. The congestion is captured by the occupancy detec-

tors when it reaches the corresponding location (Figure 8). The resulting mean AVD (in s/veh/km) for 10 replications is 217 with a standard deviation of 38.8 and highest/lowest AVD values of 269/158, respectively.

Figures 9a and 9b display the trajectories of the merge area outflow q_{out} and occupancy o , respectively, for one particular simulation run with $AVD = 217$ s/veh/km (i.e., equal to the mean AVD of 10 replications). The outflow q_{out} is roughly following the arriving demand increase until about $t = 10$ min where it shortly approaches the capacity q_{cap} . The occupancy is increasing slowly (due to increasing flow) during the same period. After $t = 10$ min, the occupancy is seen to increase sharply (due to the formed congestion) and stabilizes eventually at a value around 85%. The sharp occupancy increase is accompanied by an outflow decrease (capacity drop) to values around 1800 veh/h. When the demand is decreased and the congestion is dissolved, the system is emptied (at around $t = 35$ min).

Figure 9: merge area outflow q_{out} (a) and occupancy o (b) in the no-control case for the work zone example

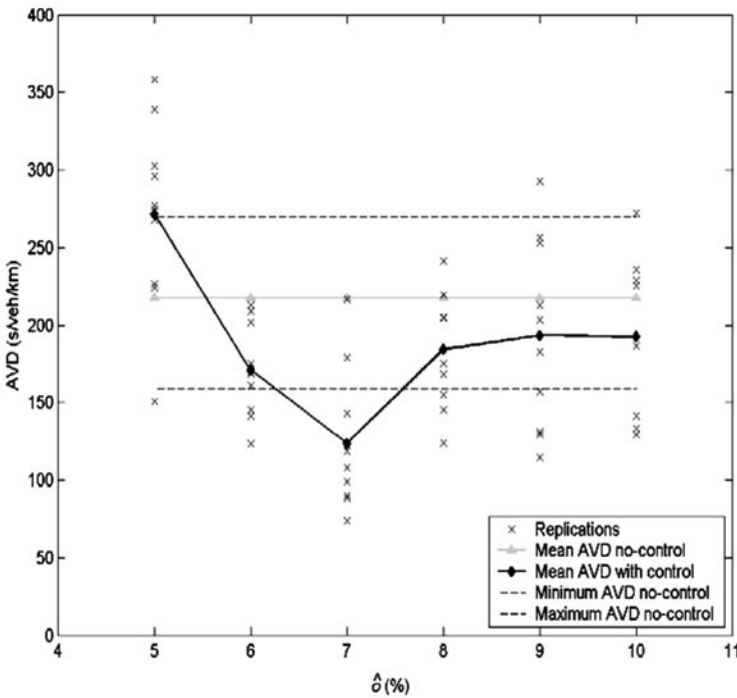


When merging traffic control is applied, the traffic lights are operated with a short red phase (corresponding to $q_{max} = 3000$ veh/h) for as long as o is lower than the set value \hat{o} , according to (1). As the arriving demand increases, o increases, and, when $o > \hat{o}$, the regulator (1) starts its actual operation aiming at maintaining $o(k)$ near \hat{o} ; this leads to a queue upstream from the traffic lights which increases, without ever reaching the simulated system entrance.

To determine the optimal set-point \hat{o} for throughput maximization, a series of simulation runs were carried out with different \hat{o} -values and 10 replications for each value; the utilized K_R value in (1) is set equal to 100 (veh/h/%). Figure 10 displays the AVDs of the 10 replications and the mean AVD for each investigated $\hat{o} \in [5\%, 10\%]$; along with the mean, highest and lowest AVD values of the 10 replications for the no-control case (horizontal lines).

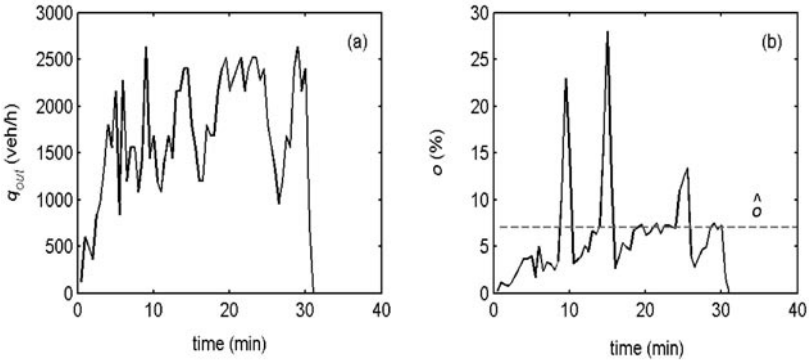
It may be seen in Figure 10 that the mean AVD under control is minimized for $\delta = 7\%$. At first view, this value may appear very low for a critical occupancy; recall, however, that o is measured on three lanes (Figure 8) to maximize the flow of one single lane (downstream). The mean AVD is lower than its no-control counterpart for all δ except $\delta = 5\%$; the mean AVD for $\delta = 7\%$ is 123 s/veh/km which is 43% lower than the mean AVD of the no-control case. However, in contrast to section 2.2, there are some control replications with a higher AVD than the best no-control AVD, even for the optimal set-point $\delta = 7\%$.

Figure 10: average vehicle delay versus δ with and without merging traffic control for the work zone example



Figures 11a and 11b display the trajectories of q_{out} and o , respectively, for one particular simulation run with $\delta = 7\%$; and AVD = 118 s/veh/km which is close to the mean AVD of the corresponding 10 replications. The outflow q_{out} is roughly following the arriving demand increase until about $t = 9$ min, at which time o exceeds δ and the regulator (1) is activated; for reasons mentioned in section 1.7, o exhibits a strong overshooting (reaching 20%), followed by an undershooting.

Figure 11: outflow q_{out} (a) and occupancy o (b) in the merging control case with $\hat{\delta} = 7\%$, $K_R = 100$ veh/h/% for the work zone example



In contrast to section 2.2 (Figure 6b), the period after the initial overshooting is characterized by occasional, randomly appearing, significant departures of the occupancy under control to values that may reach 30% or more. The observed occupancy spikes are due to occasionally appearing (simulated) vehicle merging conflicts that lead to provisionally strong vehicle decelerations in the merge area, i.e., small congestions that are captured by the occupancy detectors; this triggers a corresponding regulator reaction that drives the occupancy back to values close to $\hat{\delta}$, albeit with a corresponding undershooting. Thus, the regular merging control period (after the transient period) comprises a number of short-lived occupancy spikes and eventual undershootings, during which the outflow is seriously reduced. Between the spikes, there are subperiods of proper regulation, with $o(k) \approx \hat{\delta}$ and $q_{out} \approx q_{cap} \approx 2300$ veh/h.

It is interesting to note that the manifest probability of occurrence of the spikes increases in our simulations with increasing $\hat{\delta}$ beyond 7%. In contrast, for $\hat{\delta} = 6\%$ there are virtually no spikes (other than the initial overshooting) observed in the corresponding 10 simulation replications, i.e., the occupancy is maintained around 6% while $q_{out} \approx 2100$ veh/h, i.e., lower than capacity. However, the average outflow for $\hat{\delta} = 7\%$ is apparently higher than $q_{out} \approx 2100$ veh/h of the case $\hat{\delta} = 6\%$, hence the higher mean AVD of the latter in Figure 10.

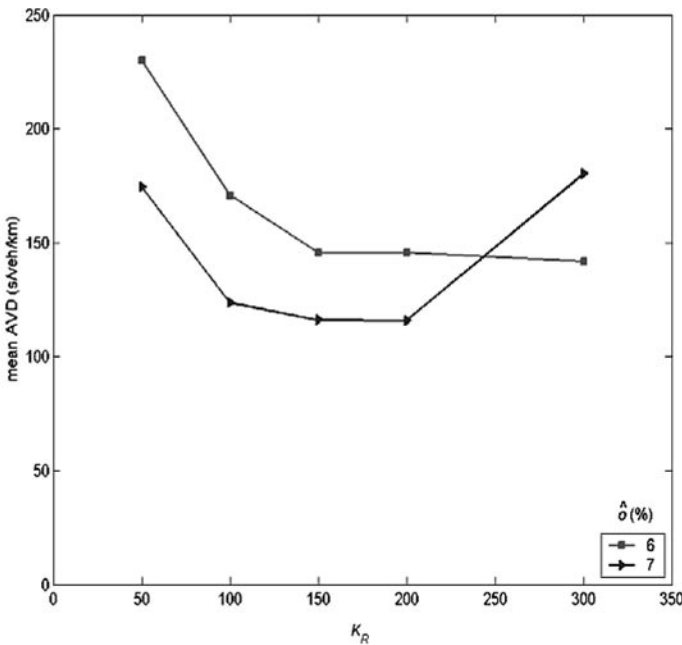
It is difficult to judge whether the observed occupancy spikes are due to the geometry of the particular merge area or due to the typically limited accuracy of the vehicle-merging behavior by the microscopic simulator (this is what we suspect) or both. Related field data are probably the most reliable means to answer this question. In any case, even in the presence of occupancy spikes, the merge area outflow is increased in average with merging control and $\hat{\delta} = 7\%$, leading to the mentioned mean AVD decrease of 43% against the no-control case, while the formed queue in the control case ($\hat{\delta} = 7\%$) is emptied at around $t = 31$ min, i.e., 4 min earlier than in the no-control case.

Figure 12 displays the mean AVDs for merging traffic control with δ equal to 6% and 7%, against K_R within the range [50, 300]. It may be seen that AVD is improved for $K_R > 100$ in the case $\delta = 6\%$, due to better handling of the initial overshooting as explained in section 2.2 (Figure 7). For $\delta = 7\%$, best results are obtained in the range $K_R \in [100, 200]$; for $K_R < 100$, the control performance deteriorates due to slower handling of the occurring spikes by the regulator; for $K_R > 200$, the control performance deteriorates due to the regulator's overreaction to the occurring occupancy spikes.

In conclusion, the selection of optimal values for δ and K_R appears somewhat more cumbersome in this application compared to section 2.2, but the range of values leading to significant improvements compared to the no-control case is still reasonably broad. Field trials would provide more evidence about the actual occupancy behavior under control and, most importantly, about the achievable level of benefits via merging traffic control.

More details on work zone merging traffic control and related simulation results may be found in Lentzakis, *et al.*, 2008.

Figure 12: mean AVD versus K_R for different set values δ for the work zone example



4. Conclusions and future work

A general framework was developed for merging traffic control in a variety of highway transportation infrastructures where vehicles are merging from a higher to a lower number of lanes. The control framework employs the well-known local ramp metering strategy ALINEA and aims at throughput maximization (or, equivalently, delay minimization) in the corresponding merging traffic facilities. The concept is demonstrated via microscopic simulation for the particular innovative applications of real-time toll plaza management and real-time work zone management with considerable achievable benefits.

A further innovative (and potentially more significant) application of the proposed concept may be the combined on-ramp and mainline metering at freeway junctions. Ramp metering may be applied at freeway on-ramp merges in order to maximize the freeway throughput downstream of the on-ramp. However, the available storage space of on-ramps is usually limited, e.g., because the ramp queue should not intervene with the adjacent street traffic; hence, ramp metering is usually released if the ramp storage space is fully covered with vehicles, and, as a consequence, a merge area congestion is created and the freeway throughput is reduced due to capacity drop. The proposed framework could be used for sustained throughput maximization via combined on-ramp and mainline metering (upstream from the ramp), whereby mainline metering may be effectuated by use of any of the control devices mentioned in section 1.5. This approach bears some similarities with the one proposed independently by Chang, *et al.* (2007).

Extending this idea further, the local on-ramp/mainline metering system could be understood as a component of a more comprehensive network-wide freeway traffic control system that decides about the (optimal) outflows to be realized by each local component by use of (2). The ultimate vision would be a comprehensive freeway traffic control system that exploits not only the (usually limited) ramp storage space but also the mainline towards an orderly, efficient and safe operation that would replace the chaotic traffic conditions observed daily on most metropolitan freeways around the world. Corresponding investigations are in progress and will be reported in due time.

As a final remark, the presented general concept for merging traffic control may be also applied for throughput maximization in transportation infrastructures where no actual merging takes place because $M \leq \mu$, i.e., the number M of arriving lanes is equal or smaller than the number μ of exiting lanes, provided the downstream capacity q_{cap} is lower than the arriving traffic flow. This may be the case at the entrance of tunnels or other facilities or highway parts where the downstream the capacity q_{cap} is lower than the capacity of the upstream infrastructure parts. Traffic control at the entrance of tunnels has in fact been practiced occasionally, albeit sometimes with somewhat vague goals and control algorithms.

References

- Cassidy, M.J. and Rudjanakanoknad, J., 2005. Increasing the capacity of an isolated merge by metering its on-ramp. *Transportation Research Part B*, 39, 896–913.
- Chang, H., Wang, Y., Zhang, J. and Ioannou, P., 2007. An integrated roadway controller and its evaluation by microscopic simulator VISSIM. CD-ROM European Control Conference 2007, Kos, Greece, 2436–2441.
- Elefteriadou, L., Roess, R.P. and Mc Shane, W.R., 1995. Probabilistic nature of breakdown at freeway merge junctions. *Transportation Research Record*, 1484, 80–89.
- FHWA, 2005. Developing and Implementing Transportation Management Plans for Work Zones. Report of the Federal Highway Administration, U.S. Department of Transportation.
- FHWA, 2006. State of the Practice: Traffic Control Strategies for Toll Plazas. Report of the Federal Highway Administration, U.S. Department of Transportation.
- Keen, K.G., Schofield, M.J. and Hay, G.C., 1986. Ramp metering access control on M6 freeway. In *Proc. 2nd IEE International Conference on Road Traffic Control*, London, UK.
- Kondyli, A., Duret, A. and Elefteriadou, L., 2007. Evaluation of CORSIM and AIMSUN for freeway merging segments under breakdown conditions. CD-ROM 86th Annual Meeting of the Transportation Research Board, Washington, D.C., U.S.A.
- Kosmatopoulos, E.B. and Papageorgiou, M., 2003. Stability analysis of the freeway ramp metering control strategy ALINEA. CD-ROM 11th IEEE Mediterranean Conference on Control & Automation, Rhodes, Greece.
- Kosmatopoulos, E., Papageorgiou, M., Manolis, D., Hayden, J., Higginson, R., McCabe, K. and Rayman, N., 2006. Real-time estimation of the critical occupancy for maximum motorway throughput. *Transportation Research Record*, 1959, 65–76.
- Kotsialos, A., Papageorgiou, M., Hayden, J., Higginson, R., McCabe, K. and Rayman, N., 2006. Discrete release rate impact on ramp metering performance. *IEE Proc. Intelligent Transportation Systems*, 153, 85–96.
- Lentzakis, A.F., Spiliopoulou, A.D., Papamichail, I., Papageorgiou, M. and Wang, Y., 2008. Real-time work zone management for throughput maximization. 87th Annual Meeting of the Transportation Research Board, Washington, D.C., U.S.A., paper 08-0772.
- Lorenz, M. and Elefteriadou, L., 2001. Defining highway capacity as a function of the breakdown probability. *Transportation Research Record*, 1776, 43–51.

- Papageorgiou, M., Hadj-Salem, H. and Blosseville, J.M., 1991. ALINEA: a local feedback control law for on-ramp metering. *Transportation Research Record*, 1320, 58–64.
- Papageorgiou, M., Haj-Salem, H. and Middelham, F., 1997. ALINEA local ramp metering: Summary of field results. *Transportation Research Record*, 1603, 90–98.
- Papageorgiou, M., Kosmatopoulos, E., Papamichail, I. and Wang, Y., 2007. ALINEA maximises freeway throughput – An answer to flawed criticism. *Traffic Engineering and Control*, 48, 271–276.
- Papageorgiou, M., Kosmatopoulos, E., Protopapas, M. and Papamichail, I., 2006. Evaluation of the Effects of Variable Speed Limits on Freeway Traffic using M42 Traffic Data. Internal Report 2006–25, Dynamic Systems and Simulation Laboratory, Technical University of Crete, Chania, Greece.
- Papageorgiou, M. and Kotsialos, A., 2002. Freeway ramp metering: An overview. *IEEE Trans. on Intelligent Transportation Systems*, 3, 271–281.
- Papageorgiou, M. and Papamichail, I., 2008. Overview of traffic signal operation policies for ramp metering. 87th Annual Meeting of the Transportation Research Board, Washington, D.C., U.S.A., paper 08-0777.
- Smaragdis, E. and Papageorgiou, M., 2003. A series of new local ramp metering strategies. *Transportation Research Record*, 1856, 74–86.
- Smaragdis, E., Papageorgiou, M. and Kosmatopoulos, E., 2004. A flow-maximizing adaptive local ramp metering strategy. *Transportation Research Part B*, 38, 251–270.
- Spiliopoulou, A.D., 2007. Vehicle Flow Control in a Freeway Toll Plaza (in greek). Diploma Thesis, Dynamic Systems and Simulation Laboratory, Technical University of Crete, Chania, Greece.
- Spiliopoulou, A.D., Papamichail, I. and Papageorgiou, M., 2008. Real-time toll plaza management for throughput maximization. 87th Annual Meeting of the Transportation Research Board, Washington, D.C., U.S.A., paper 08-0771.
- 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* 74, 016123.
- Vigos, G., Papageorgiou, M., Wang, Y. (2008) Real-time estimation of vehicle-count within signalized links. *Transportation Research Part C*, 16, in press.
- Wei, H. and Pavithran, M., 2006. Concept of dynamic merge metering approach for work zone traffic control. In *Preprints 11th IFAC Symposium on Control in Transportation Systems*, Delft, The Netherlands, 374–379

Fiche bibliographique

Unité de recherche GREZIA	Projet n°	Actes INRETS n° 123
Titre Modélisation du trafic		
Sous-titre Actes du groupe de travail, 2007		Langue Français, anglais
Auteur(s) Jean-Patrick Lebacque, Maurice Aron Vincent Aguiléra		Organismes UPE/INRETS-GREZIA UPE/ ENPC-LVMT
Nom adresse financeur, co-éditeur		N° contrat, conv.
		Date de publication Décembre 2009
Remarques		
Résumé <ul style="list-style-type: none"> • Salim Mammam <i>et al.</i> analysent les points fort du modèle ARZ d'écoulement du trafic du second ordre (comprenant une équation de la vitesse). Ils donnent une solution analytique dans certains cas, et un schéma de résolution numérique pour le cas général. • Anthony Busson <i>et al.</i> calculent un réseau dont les nœuds sont des véhicules communiquant avec une carte réseau sans fil, capables de relayer une alerte. Ils donnent des résultats théoriques (analytiques) et pratiques (par simulation). • Aurélien Duret <i>et al.</i> estiment le modèle de poursuite de Newell, soit classiquement, soit en calculant les paramètres (vitesse de remontée de congestion et distance à l'arrêt) du diagramme fondamental triangulaire. • L'objectif du contrôle d'accès est d'agir sur le débit de la bretelle d'accès afin de maintenir la densité dans la voie principale au-delà d'un seuil critique. Dans ce contexte, H. Abouaïssa et V. Iordanova propose une nouvelle loi de commande basée sur le concept de la platitude différentielle. L'approche développée permet, grâce à l'inversion du système, d'obtenir la loi de commande en fonction de l'état du système (la densité) et de sa dérivée et ce sans intégration d'équation différentielle. • Markos Papageorgiou <i>et al.</i> calculent l'impact sur la circulation et notamment sur le diagramme fondamental de la régulation des vitesses sur autoroute: hors congestion la vitesse est moins élevée, plus homogène; la densité critique est plus élevée; ils proposent une analyse en temps réel du diagramme fondamental pour optimiser la régulation. • Markos Papageorgiou <i>et al.</i> traitent de la diminution du nombre de voies sur autoroute, qu'elle soit due à l'existence de travaux, d'une zone de péage, à la convergence de deux sections autoroutières. Les auteurs montrent comment la commande optimale type ALINEA permet l'optimisation de la capacité de l'ouvrage et des temps de parcours. 		
Mots clés écoulement du trafic, modèle du second ordre, modèle de poursuite, routage, simulation, calcul analytique, platitude, régulation de vitesse, régulation d'accès, réseau de communication, convergent, zone de péage, zone de travaux, alinea.		
Nb de pages 130	Prix 15,24 euros	Bibliographie oui

Publication data form

Research Unit Transportation Network Engineering and advance Informatics (GRETIA)	Projet No	INRETS proceeding No 123
Title Traffic Modelling		
Subtitle Proceedings of the seminar, 2007		Language French, English
Author(s) Jean-Patrick Lebacque, Maurice Aron Vincent Aguiléra		Affiliation UPE/INRETS/GRETIA UPE/ENPC/LVMT
Sponsor, co-editor, name and address		Contract, conv. N°
		Publication date December 2009
Notes		
Summary <ul style="list-style-type: none"> • Salim Mammar, J-P. Lebacque and H. Haj-Salem highlight the main points of the ARZ traffic flow model (a second order model, including a speed equation). They provide an analytic solution in a few cases, and a numerical scheme for solving the model in the general case. • A. Busson, M Mabilia and V. Vèque define and compute a network, the nodes of which being vehicles communicating thanks to a wireless network card, able to relay an alert or other pieces of information. The authors give some theoretical (analytical) results and some practical results (within a simulation). • A. Duret, C. Buisson, N. Chiabaut estimate the Newell car-following model first directly from empirical data, and second by computing the parameters of the triangular fundamental diagram, i.e. the congestion propagation speed, and the spatial headway observed after the stopping of the vehicles. • The main objective of ramp metering is to act on the traffic flow at the on-ramp in order to maintain the traffic density of the freeway segment below a critical threshold. In this context, H. Abouaïssa and V. Iordanova propose a new control law based on the differential flatness concept. This allows, thanks to the system's inversion to obtain the control variable as function of the state (traffic density) and its first time-derivative without integration of any differential equation. • M. Papageorgiou, E Kosmatopoulos and I. Papamichail compute the impact of the speed control on the motorway traffic flow and noticeably on the fundamental diagram: Beneath the congestion, the speed is lesser, more homogeneous, thanks to the speed control. The authors propose a real time analysis of the fundamental diagram for optimizing the speed control strategy. • M. Papageorgiou, I. Papamichail, A. D. Spiliopoulou, A. F. Lentzakis model the merging problem on motorway, whatever its source is – merging highways, work area, toll plaza... The drop of capacity and the congestion due to the merging are minimized thanks to an optimal control algorithm based on ALINEA. 		
Key words traffic flow, second order model, car-following, simulation, analytic calculation, routing, communication network, flatness, speed control, ramp metering, merging, work area, toll plaza, Alinea		
Nb of pages 130	Price 15,24 euros	Bibliography yes

Imprimé en France
JOUVE - 1, rue du Docteur Sauvé - 53101 Mayenne
N° 499488S - Dépôt légal : Décembre 2009

Les avancées scientifiques exposées en 2007 concernent les modèles d'écoulement du trafic, leur lien avec la régulation du trafic, et la régulation proprement dite.

On apprécie le calibrage du modèle de poursuite de Newell, ainsi que la résolution du modèle macroscopique d'écoulement ARZ, et on comprend l'impact sur le diagramme fondamental de la régulation de vitesse sur autoroute.

Une nouvelle possibilité théorique de régulation apparaît au niveau des convergents autoroutiers : la réduction de capacité au convergent est minimisée grâce à la commande optimale ALINEA.

A noter un nouveau calcul de la régulation d'accès. Pour maintenir la concentration au-dessous d'une valeur critique, la « platitude différentielle » consiste à inverser le système étudié pour obtenir directement la loi de commande en fonction de la concentration et de sa dérivée.

Enfin, on mesure l'intérêt du trafic « coopératif », grâce au calcul des performances théoriques et pratiques d'un réseau de véhicules communiquant entre eux par une carte réseau sans fil.

Jean-Patrick Lebacque est ingénieur général des Ponts et Chaussées à l'INRETS-GRETIA

Maurice Aron est chargé de recherche à l'INRETS-GRETIA

Vincent Aguiléra est chargé de recherche au LVMT-ENPC

Actes n° 123

Décembre 2009

15,24 €



Institut national de recherche
sur les transports et leur sécurité

ISSN 0769-0266

ISBN 978-2-85782-676-7