



HAL
open science

Modélisation statistique et probabiliste du TEMPS INTER-VEHICULAIRE aux différents niveaux de trafic

Duy-H Ha

► **To cite this version:**

Duy-H Ha. Modélisation statistique et probabiliste du TEMPS INTER-VEHICULAIRE aux différents niveaux de trafic. Modélisation et simulation. Université Paris-Est, 2011. Français. NNT: . tel-01006504

HAL Id: tel-01006504

<https://theses.hal.science/tel-01006504>

Submitted on 23 Jul 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Paris-Est
GRETTIA/IFSTTAR

**Modélisation statistique et probabiliste
du TEMPS INTER-VEHICULAIRE
aux différents niveaux de trafic**

Duy-Hung HA

Directeur de thèse : Simon Cohen, GRETTIA/IFSTTAR
Président du jury : Fabien Laurent, Ecole des Ponts ParisTech
Rapporteurs : Guillaume Saint-Pierre
Gérard Louah
Saïd Mammar,
Nour-Eddin El-Faouzi

Noisy-Champs, Mai 2011

À ma chère famille au Vietnam

Modélisation statistique et probabiliste du TIV aux différents niveaux de trafic

Duy-Hung HA (duy-hung.ha@hotmail.fr)

Résumé

Le Temps Inter-véhiculaire (TIV) est une variable microscopique fondamentale dans la théorie du trafic, et a été étudié depuis le début du développement de cette théorie, vers 1930. La distribution de probabilité du TIV décrit la répartition des arrivées des véhicules en un point donné et reflète dans une certaine mesure le comportement de conduite. Beaucoup d'applications en ingénierie du trafic viennent de la connaissance fine de cette variable. La thèse a pour but d'approfondir cette connaissance en modélisant la distribution du TIV dans différents contextes selon différents points de vue. Tout d'abord, deux méthodes d'échantillonnage, la méthode de groupement et la méthode de raffinement sont considérées. L'application numérique concerne deux bases de données, celle de la route nationale RN118 et celle de l'autoroute A6. Ensuite, trois types de modèles probabilistes sont analysés et classifiés. Une comparaison exhaustive des modèles et des méthodes d'estimation est réalisée ce qui conduit à considérer que le modèle gamma-GQM est supérieur aux autres modèles en matière de performance statistique et en efficacité de calcul. Une procédure d'estimation est proposée qui favorise la stabilité des paramètres estimés. Six nouveaux modèles de TIV sont proposés, calibrés, analysés. Mis à part deux modèles de performance inférieure aux autres et au modèle gamma-GQM, quatre modèles sont équivalents voire meilleurs que le modèle gamma-GQM. Pour une raison pratique, le modèle Double Gamma est choisi à côté du modèle gamma-GQM, comme modèle de comparaison, dans toute la modélisation des TIV. Le calibrage des modèles et l'analyse des paramètres des modèles sont menés, à partir des données réelles, en considérant trois dimensions d'étude du trafic : les échelles macroscopique, mésoscopique et microscopique. Une quatrième dimension d'étude des TIV est constituée des facteurs exogènes au trafic. La prise en compte de ces facteurs exogènes, à chaque échelle macroscopique entraîne la distinction de deux types de facteur exogène : « empêchant » et « impulsif ». Finalement, différentes approches de validation sont testées. L'approche proposée par « enveloppe des distributions » semble prometteuse pour le futur.

Abstract

Time Headway (TH) is a microscopic variable in traffic flow theories that has been studied since the 1930s. Distribution of this fundamental variable describes the arrival pattern of vehicles in traffic flow, so probabilistic modeling is the main approach to study TH and represents driving behavior. The applications of the variable in traffic engineering are varied ; include capacity calculation, microscopic simulation, traffic safety analysis, etc. This dissertation aims at modeling the TH distribution in different contexts. Firstly, the short-time sampling method and long-time sampling method are applied to obtain TH samples from the two data bases (the RN118 national roadway and the A6 motorway). Then, three probabilistic TH model types are analyzed and classified. An exhaustive comparison between the existing models and between the corresponding estimation methods leads to consider that the gamma-GQM is the best TH model in the literature. An estimation process is also proposed in order to obtain good and stable estimated results of the parameters. After that, the TH probabilistic modeling is developed by six new models. Except for the two ones which are worse, the four other models are statistically equivalent and/or better than the gamma-GQM. For practical reason, the Double Gamma model is selected, as a comparison model, with the gamma-GQM to calibrate all TH samples. Three traffic levels are considered : macroscopic, mesoscopic and microscopic. The effects of exogenous factors are also examined. Examining this factor in each macroscopic variable level leads to distinguish two following factor types : "impeding" factor and "propulsive" factor. Finally, different approaches for TH validation are tested. The proposed approach of "envelope of distributions" seems to be promising for future applications.

Remerciements

Cette page est dédiée aux professeurs, collègues, amis, parents sans qui cette thèse n'aurait pas vu le jour. Pendant cette période de thèse, j'ai eu l'occasion d'échanger, de partager des idées, des sentiments avec de nombreuses personnes. Je les remercie du fond du coeur pour la contribution, directe ou indirecte, qu'ils m'ont apportée.

En premier lieu, je tiens à remercier tout particulièrement mon encadrant, co-directeur de thèse, Maurice Aron, chargé de recherche au GRETTIA/IFSTTAR. Sa bonne humeur, sa disponibilité, sa générosité formidable, la pertinence de ses idées ont été un soutien solide pendant les années de mon travail au laboratoire. Il a suivi au quotidien l'avancée de mon travail, corrigé et commenté en détail les résultats et les interprétations. Il m'a aussi appris le sens critique, indispensable pour la recherche, utile aussi bien sûr pour la vie quotidienne.

Mes remerciements vont en deuxième lieu à Simon Cohen, directeur de thèse, directeur de recherche au GRETTIA/IFSTTAR. Sous sa direction, sa générosité et sa rigueur, la thèse a pris une bonne direction. Je le remercie aussi pour ses cours de Master grâce auxquels le domaine de l'ingénierie du trafic m'a fortement marqué. C'est aussi grâce à lui que je dois de travailler à l'IFSTTAR.

Je tiens à exprimer ma gratitude envers Saïd Mammar, Professeur de l'Université d'Evry et envers Nour-Eddin El-Faouzi, Directeur de Recherche de l'Université de Lyon qui ont accepté d'être les rapporteurs de cette thèse. Je les remercie pour le temps qu'ils ont consacré et pour leurs commentaires précieux sur mes travaux, commentaires grâce auxquels la thèse a, tout au moins je l'espère, gagné en profondeur et en clarté.

Je tiens aussi à remercier Fabien Laurent, Professeur de l'Ecole des Ponts ParisTech pour avoir accepté de prendre le temps nécessaire, sur son emploi du temps déjà très chargé, pour exercer le rôle de président du jury et d'examinateur. Merci également à Guillaume Saint-Pierre et à Gérard Louah d'avoir accepté d'examiner le document de thèse et de venir à ma soutenance, merci aussi pour les critiques, remarques et compléments qu'ils ne manqueront pas d'apporter, grâce à leur compétence dans la science de la statistique et dans le domaine des temps intervéhiculaires.

Mes remerciements ne seraient pas complets si je ne parlais pas de mes collègues au GRETTIA. J'ai eu l'opportunité de travailler dans un laboratoire formidable et fort. J'ai pu apprendre de chacun des membres et acquérir beaucoup d'expérience professionnelle et de la vie. Vous avez toujours été gentils et avez su créer une ambiance de travail de recherche conviviale et performante. Merci à Gérard Scemama, l'ex-directeur du GRETTIA, pour son leadership calme et efficace, pour sa générosité; merci à Régine, à Neïla, à Sophie, à Ruth, à Mahdi, à Nathalie, à Joëlle et aux autres chercheurs dont je retiens le nom dans mon coeur et ma tête. Merci beaucoup à Mustapha pour sa disponibilité, son appui technique permanent irremplaçable aux doctorants.

Je saisis cette occasion pour remercier mes jeunes ami-collègues. Je pense à Hilal, à Florence, à Taiyu, à Yousef, à Moussa, à Sofiane (doctorant) et à Fabien dont j'ai partagé le bureau pendant cette période importante de la vie. Je voudrais dire bonne continuation

aux anciens stagiaires du GRETTIA parmi lesquels Sofiane (stagiaire), Lucas, Viet-Phuong et Mohamed avec qui nous avons passé de très bons moments à l'INRETS. Je pense aussi à Judith, ma camarade de Master dans le bureau voisin, à sa force et à ses efforts en particulier après le tremblement de terre en Haïti, et qui m'a beaucoup appris. Merci en particulier à Krystyna qui, par son talent et son expérience, m'a donné beaucoup de conseils et pas seulement sur le plan scientifique. Je pense aussi aux amis d'autres unités de recherche parmi lesquels Thai-Phu, Mathilde, Daphna, Faicel, Marouan, Thanh-Tu, Zohra, et aussi aux amis que j'ai rencontrés durant des conférences, à mes amis au quotidien, à la cité universitaire et à toutes les personnes que j'ai eu l'occasion de rencontrer et avec qui j'ai échangé, et qui donnent de la couleur à la vie. Merci à vous tous.

Surtout, je tiens à remercier ici l'ensemble de ma famille au Vietnam où vivent ma mère et mon père, ma soeur, mon frère, mes deux neveux Son, Giang et ma petite nièce Phuong Anh. Ils sont pour moi la source d'énergie inépuisable qui m'a permis d'accomplir ce mémoire. Que j'aurais aimé voir leurs sourires, notamment celui de ma mère, à la nouvelle de la bonne fin de la thèse. Merci tout particulièrement à mon frère Long et ma soeur Thuy pour leur soutien permanent durant mon parcours en France.

Val de Fontenay, 04 mars 2011

Table des matières

1	Bases de données et échantillonnage	37
1.1	Pré-échantillonnage	37
1.2	Échantillonnage	38
1.2.1	Méthode de groupement	38
1.2.2	Méthode de raffinement	41
1.3	Analyses statistique et spectrale	45
1.3.1	Analyse statistique	45
	Les statistiques descriptives	45
	Distribution empirique des TIV	47
1.3.2	Analyse spectrale	48
	Les propriétés statistiques du second ordre	49
	Estimateurs des statistiques du second ordre	50
	Estimateurs du spectre des intervalles	51
1.3.3	Run test	52
1.3.4	Test de renouvellement	53
	Corrélogramme et Test de renouvellement sur les ρ_i	53
	La statistique R_1	53
	Périodogramme	53
	Test de la loi géométrique de la longueur des pelotons	54
1.4	Les données réelles	55
1.4.1	Les données de l'Autoroute A6	55
	Présentation du site	55
	Trafic aux différentes périodes	55
	Analyse statistique	56
	Analyse spectrale	57
	Tests de tendance, de groupement et de renouvellement	59
1.4.2	Les données de la Route Nationale RN118	61
	Présentation du site	61
	Présentation des données	61
	Analyse statistique des échantillons groupés de la RN118	63
	Analyse statistique des échantillons raffinés de la RN118	71
	Analyse spectrale des échantillons raffinés	71
1.5	Conclusion du chapitre	74
2	Modélisation probabiliste du TIV	75
2.1	Approches stochastiques	75
2.1.1	Approche de Heidemann	75
	Le modèle	75
	Résultats	78
	Discussion	79

2.1.2	Modèle des automates cellulaires de Nagel – Schreckenberg	81
	Le modèle	81
	Résultats et Discussion	85
2.2	Approche probabiliste	85
2.2.1	Les modèles simples	86
	1. Modèle Exponentiel	86
	2. Modèle Gamma	88
	3. Modèle Log-normal	91
	4. Autre modèles	93
2.2.2	Modèle combiné	94
	1. Modèle M3 de Cowan	95
	2. Modèle Hyperexponentiel et modèle de Schuhl	96
	3. Modèle Hypergamma et modèle Hyperlang	98
	3. Modèle Hypernormal	100
2.2.3	Modèle de mélange	100
	1. Modèle semi-Poisson (SPM)	101
	2. Modèle de file d'attente généralisée (GQM)	103
2.2.4	Méthode d'estimation des paramètres	105
	Méthode des moments (MM)	105
	Méthode du minimum du χ^2 (MC)	105
	Méthode du minimum de l'intégrale de la moyenne des carrés des erreurs (MISE)	105
	Méthode du Maximum de Vraisemblance (MV)	107
	Les autres méthodes	108
2.2.5	Les tests d'ajustement	109
	Le test du χ^2	110
	Le test de Kolmogorov–Smirnov (K–S)	110
	Le test d'Anderson–Darling (A–D)	110
2.3	Comparaison des modèles	111
2.3.1	L'état de l'art de la comparaison des modèles pour le TIV	111
	Chronologie	111
	Discussion	112
2.3.2	Ensemble test des échantillons	113
2.3.3	Le processus d'estimation	114
2.3.4	Modèles simples	115
2.3.5	Modèles combinés	116
2.3.6	Modèles de mélange	117
	Les modèles g -GQM	119
	Les méthodes d'estimation pour le modèle Γ -GQM	120
2.3.7	Paramètre de décalage	122
2.4	Discussion et développement des modèles	124
2.4.1	Modèle simple et modèle combiné	124
2.4.2	Modèle de mélange	126
2.4.3	Développement des modèles pour les TIV	129
	A. Développement des modèles de mélange	129
	B. Amélioration des modèles combinés	130
2.4.4	Méthode d'estimation	131
2.4.5	Comparaison des nouveaux modèles	131
	Comparaison du modèle HyperGQM avec le modèle Γ -GQM	131
	Les autres modèles développés	132

2.5	Conclusion du chapitre	135
3	Les paramètres des modèles Γ-GQM et DG	137
3.1	Paramètre de décalage τ	137
3.2	Le modèle Γ -GQM	138
3.2.1	Les propriétés statistiques	138
3.2.2	Influence de $\theta, \lambda, \alpha, \beta$ du modèle Γ -GQM	138
	Le paramètre θ du modèle Γ -GQM	138
	Le paramètre λ du modèle Γ -GQM	139
	Le rôle des paramètres α et β du modèle Γ -GQM	141
	La relation entre α et β	142
3.3	Le modèle DG	143
3.3.1	Les propriétés statistiques	143
3.3.2	Influence des paramètres du modèle DG	144
	Le paramètre θ du modèle DG	144
	Les rôles de α_1 et β_1 du modèle DG	145
	La relation entre α_i et β_i	146
3.4	Estimation par la méthode du maximum de vraisemblance	146
3.4.1	Les fonctions de vraisemblance des modèles	146
3.4.2	La méthode du Bootstrap	147
3.5	Conclusions du chapitre	150
4	Niveau macroscopique et conditions exogènes	151
4.1	État de l'art	151
4.1.1	Les facteurs	151
4.1.2	État de l'art	153
4.1.3	Discussion	155
4.2	Distribution des TIV à l'échelle macroscopique (les échantillons groupés RN118)	156
4.2.1	Débit 6 minutes Q_6	156
	Le modèle Γ -GQM	156
	Le modèle Double Gamma DG	157
4.2.2	Taux d'occupation 6 minutes \mathcal{T}_6	158
	Le modèle Γ -GQM	158
	Le modèle DG	159
4.2.3	Vitesse moyenne harmonique 6 minutes \mathcal{V}_6	159
	Le modèle Γ -GQM	160
	Le modèle DG	160
4.2.4	Le Diagramme fondamental	160
	Les états élémentaires	160
	La sensibilité des paramètres	162
4.2.5	Discussion sur les paramètres du modèle Γ -GQM	164
4.3	Les effets exogènes sur la distribution des TIV	166
4.3.1	Conditions de météo sur la VL	167
4.3.2	Le modèle Γ -GQM selon les conditions de météo	167
4.3.3	Le modèle DG selon les conditions de météo	168
4.3.4	Comparaison entre les voies par temps sec de jour	168
4.3.5	Le modèle Γ -GQM selon les voies par temps sec de jour	169
4.3.6	Le modèle DG selon les voies par temps sec de jour	169
4.3.7	Discussion sur les paramètres du modèle Γ -GQM	170

4.4	Modélisation du TIV sur différentes bases de données	172
4.4.1	Les échantillons raffinés de la RN118	172
	L'ajustement des modèles Γ -GQM et DG	172
	Le modèle Γ -GQM	173
	Le Modèle DG	173
	Discussion sur les résultats des échantillons raffinés	174
4.4.2	Soixante-quinze échantillons horaires de l'autoroute A6	175
	L'ajustement des modèles Γ -GQM et DG	175
	Le modèle Γ -GQM	175
	Le modèle Double Gamma	176
	Étude des périodes de même tranche d'heure	177
	L'effet de la réduction de la largeur de voie sur l'autoroute l'A6	178
	Discussion	179
4.5	Les conditions plus complexes et les cas particuliers	181
4.5.1	Les échantillons ayant des paramètres d'échelle similaires	181
4.5.2	La discontinuité du paramètre θ du modèle Γ -GQM	182
4.5.3	Les échantillons de débit très faible	183
4.5.4	Les échantillons de très haut débit	185
4.5.5	L'existence de distributions bi-modales des TIV	186
4.5.6	Les échantillons dans la période du soir	187
4.6	Modélisation du comportement standard	188
4.6.1	Schéma de principe	188
4.6.2	Régression des paramètres des modèles en fonction du débit Q_6	188
4.7	Conclusions du chapitre	192
5	Niveaux mésoscopique et microscopique	193
5.1	L'échelle mésoscopique	193
5.1.1	Le groupement des variables microscopiques	193
	État de l'art	193
	Application sur les données de la RN118	196
5.1.2	Résultats	197
	Groupement en fonction de la DIV d_i	197
	Groupement en fonction du TIV	200
	Groupement en fonction de la vitesse	203
	Discussion	204
5.1.3	Distribution des TIV et l'échelle mésoscopique	205
	La position du véhicule en peloton	205
	La configuration	209
5.1.4	Distribution des TIV selon le TIV du véhicule précédent	213
5.2	L'échelle microscopique	214
5.2.1	Distribution des TIV et de la vitesse instantanée	214
	Discussion	218
5.2.2	La variable RHIS	218
5.3	Conclusion du chapitre	221
6	Validation	223
6.1	Retour sur les buts de la modélisation probabiliste	223
6.2	Méthodes de validation	224
6.2.1	Validation selon les relations établies sur les différents sites de la RN118 - Approche de régression	225

6.2.2	Validation de statistiques des échantillons - Approche de stockage . .	226
	Validation de statistiques partielles (type I) sur les différents sites de la RN118	226
	Validation d'une statistique <i>totale</i>	227
6.2.3	Statistiques partielles (type II) sur les différents échantillons glissants du Site N°1 de la RN118	228
	Le critère de comparaison entre deux échantillons	229
6.2.4	Validation au sein du Site N°1	230
6.3	Proposition d'une approche de validation et d'application	231
6.3.1	RN118 sur A6	233
6.3.2	Validation sur différents sites de la RN118	234
6.4	Conclusion du chapitre	235
7	Conclusion générale	237
A	Les Figures et Tableaux supplémentaires	249
B	Les résultats numériques de l'A6 et de la RN118	255

Liste des tableaux

1.1	Écart-type de $2 I_n/s^2$ sur la voie normale et la voie de largeur réduite	59
1.2	Moyenne et écart-type de $2 I_n/s^2$ des échantillons de l'A6	60
2.1	L'ensemble test	113
2.2	Propriétés statistiques des TIV dans les échantillons de l'ensemble test . . .	113
2.3	Le processus d'estimation pour les modèles complexes	114
2.4	Les statistiques K-S des modèles simples, Numéro 1	115
2.5	Les statistiques K-S des modèles simples, Numéro 2	116
2.6	Les statistique K-S des modèles combinés, Numéro 1	117
2.7	Les statistiques K-S des modèles combinés, Numéro 2	117
2.8	Statistique K-S des modèles Γ -GQM et Γ -SPM	118
2.9	Valeurs de la statistique du test du χ^2 des modèles Γ -GQM et Γ -SPM . . .	119
2.10	Valeurs de la statistique K-S du modèle Γ -GQM utilisant les méthodes MV, MISE pour f et F	121
2.11	Densité des modèles simples ou fonction g des modèles combinés	126
2.12	Modèles de mélange	127
2.13	Les statistiques K-S des modèles Hyper-GQM, Γ -GQM	132
3.1	La sensibilité des paramètres du modèle Γ -GQM dans les échantillons boots- trap	148
3.2	La sensibilité des paramètres du modèle DG dans les échantillons bootstrap	149
4.1	P-valeurs du test K-S du modèle Γ -GQM du jour	171
4.2	Les échantillons sur les trois voies pour deux périodes horaires du modèle Γ -GQM	177
4.3	Les échantillons sur les trois voies pour deux périodes horaires du modèle DG	178
4.4	Deux échantillons (L1 et R3c) ayant des débits horaires Q_h similaires sur les VL et VR, modèle Γ -GQM	178
4.5	Deux échantillons (L1 et R3c) ayant des débits horaires Q_h similaires sur les VL et VR, modèle DG	179
4.6	Les paramètres du modèle Γ -GQM de deux échantillons (L3b et L3c) de la VL au même niveau de Q_h	179
4.7	Les paramètres du modèle Γ -GQM pour quatre couples d'échantillons sur deux sections de l'A6	179
4.8	Les paramètres du modèle DG pour quatre couples d'échantillons sur deux sections de l'A6	181
4.9	Deux couples d'échantillons ((L6c,L2) et (L4, L11)) avec des valeurs de α, β équivalentes	182
4.10	Résultats numériques des cinq échantillons classés selon le débit Q_6	183

4.11	Paramètres du modèle Γ -GQM pour les trois échantillons de débit Q_6 très faible	184
4.12	Paramètres du modèle Γ -GQM pour les trois échantillons de très haut débit Q_6	186
4.13	Paramètres du modèle Γ -GQM pour les 4 échantillons bi-modaux	186
4.14	Le schéma de principe de l'influence des effets exogènes \mathcal{F}_1	189
4.15	Les entrées utilisées dans les régressions des paramètres des modèles Γ -GQM et DG pour le comportement standard	190
4.16	Les régressions des paramètres des modèles Γ -GQM et DG pour le comportement standard, $R^2 > 0,9$	191
5.1	La moyenne, le CV des TIV selon le rang dans le peloton et l'ajustement des modèles calibrés	210
5.2	La moyenne, le CV des TIV et l'ajustement des modèles calibrés selon la configuration	211
5.3	Les distributions des TIV selon le TIV du véhicule précédent	212
5.4	Le nombre d'échantillons des TIV groupés selon la vitesse instantanée v	214
6.1	Validation selon l'approche de régression sur les différents sites de la RN118	225
6.2	Validation d'une statistiques partielles (type I) pour les échantillons des différents sites de la RN118, débits faibles	226
6.3	Validation d'une statistiques partielles (type I) pour les échantillons des différents site de la RN118, débits forts	227
6.4	Tableau de contingence du test du χ^2	229
6.5	La validation sur les données de l'autoroute A6	234
6.6	Le niveau de validation dans différents sites de la RN118	235
A.1	Fonction de pondération spectrale	250
A.2	L'information supplémentaire sur les sites de mesure de la RN118	251
B.1	Les propriétés statistiques du TIV de l'autoroute A6	256
B.2	Les résultats de calibrage sur le Site N°1, VL, jour-sec	257
B.3	Les résultats de calibrage sur le Site N°1, VL, jour-pluie	258
B.4	Les résultats de calibrage sur le Site N°1, VL, soir-sec	259
B.5	Les résultats de calibrage sur le Site N°1, VL, soir-pluie	260
B.6	Les résultats de calibrage sur le Site N°1, VR, jour-sec	261
B.7	Les résultats de calibrage sur le Site N°1, VR, jour-pluie	262
B.8	Les résultats de calibrage sur le Site N°1, VR, soir-sec	263
B.9	Les résultats de calibrage sur le Site N°1, VR, soir-pluie	264

Table des figures

1	Les trajectoires des véhicules dans le plan spatio-temporel	22
2	Les définitions du Temps Inter-Véhiculaire	28
1.1	Méthode de raffinement des échantillons de TIV	43
1.2	Relations $K-S^2$ du modèle log-normal, de la famille de modèles de Pearson et empirique	47
1.3	L'effet de lissage de l'histogramme par l'utilisation de la fonction Épanechnikov	48
1.4	Logarithme de la fonction de survie des données empiriques	49
1.5	Les valeurs des TIV sur l'autoroute A6	55
1.6	Débits horaires sur trois voies selon l'heure de la journée	56
1.7	Les propriétés statistiques des TIV sur l'A6	57
1.8	Coefficient d'auto-corrélation d'ordre i sur les trois voies	57
1.9	Valeurs des estimateurs de la somme des carrés des ρ_i	58
1.10	Les valeurs estimées de la densité spectrale dans les échantillons	58
1.11	Les valeurs de densité spectrale estimées des échantillons L1, M1, R1, de l'A6	59
1.12	Les valeurs des TIV sur la VL de la RN118	62
1.13	Les valeurs des TIV sur la VR de la RN118	62
1.14	Profils particuliers des valeurs de TIV de la RN118, VL	62
1.15	Moyenne et écart-type des TIV du jour en fonction de Q_6	63
1.16	Relation entre la moyenne et l'écart-type des TIV, du jour	64
1.17	CV des TIV du jour en fonction de Q_6	64
1.18	Distribution des vitesses des TIV des échantillons ayant des valeurs de CV élevées	65
1.19	Histogrammes des échantillons ayant des valeurs de CV élevées	65
1.20	Coefficient de symétrie S_3 et d'aplatissement K_4 des TIV du jour en fonction de Q_6	66
1.21	Relation $K-S^2$ et le <i>mode</i> des TIV du jour en fonction de Q_6	67
1.22	Les échantillons particuliers et leurs échantillons voisins	67
1.23	Moyenne et écart-type des TIV du jour en fonction du taux d'occupation 6 minutes \mathcal{T}_6	68
1.24	CV et Mode des TIV du jour en fonction de \mathcal{T}_6	68
1.25	S_3 et Relation $K-S^2$ des TIV du jour, en fonction de \mathcal{T}_6	69
1.26	Différences des pourcentages des TIV courts en fonction de Q_6	69
1.27	Les pourcentages des TIV courts (TIV < 1 s, 2 s et 6 s)	70
1.28	Les statistiques descriptives des TIV des échantillons raffinés	71
1.29	Coefficient d'auto-corrélation d'ordre i sur les VL et VR	72
1.30	Valeurs des estimateurs de la somme des carrés des ρ_i	72
1.31	Valeurs estimées de la densité spectrale f_0	72
1.32	Valeurs estimées de la densité spectrale f_1	73
1.33	Le périodogramme	73

1.34	Moyenne, Distance à $1/\pi$ des moyennes et écart-type des $2I_n(\omega_p)/s^2$	73
1.35	Les échantillons particuliers du point de vue d'analyse spectrale	74
2.1	Les distributions des TIV empiriques et le modèle de Heidemann	78
2.2	Les distributions des TIV empiriques et le modèle de Heidemann avec différentes entrées de \mathcal{P}_0	79
2.3	Les distributions des TIV empiriques et le modèle de Heidemann avec différentes valeurs de t_a, t_p	79
2.4	Les distributions empiriques et le modèle de Heidemann	80
2.5	Distribution des TIV utilisant le modèle N-S avec $v_{\max} = 1$	85
2.6	Le modèle exponentiel décalé et les données de la RN118	87
2.7	Le modèle Pearson III et les données de la RN118	90
2.8	Le modèle Log-Normal et les données de la RN118	91
2.9	Le modèle Hyperexponentiel et les données de la RN118	97
2.10	Le modèle Hypergamma et les données de la RN118	99
2.11	Le modèle SPM et les données de la RN118	102
2.12	Le modèle gamma-GQM et les données de la RN118	104
2.13	La densité des modèles combinés	118
2.14	Le modèle lognormal-GQM	119
2.15	La comparaison entre les modèles Γ -GQM et le modèle \mathcal{R} -GQM	120
2.16	Comparaison entre les méthodes d'estimation pour le modèle Γ -GQM	120
2.17	Comparaison entre les méthodes MV et MISE pour F	121
2.18	Les statistiques K-S des modèles Pearson-III Γ_τ et Gamma Γ	122
2.19	Les statistiques K-S des modèles Γ_τ^* et Γ^*	123
2.20	Les statistiques K-S pour les modèles $\ln \mathcal{N}_\tau^*$ et $\ln \mathcal{N}^*$	123
2.21	(a) Les statistiques K-S, (b) Temps d'estimation, (c) Temps d'ajustement des modèles Γ -GQM décalé et Γ -GQM	123
2.22	Le modèle Γ -GQM décalé estimé par la méthode MISE pour F et le modèle Γ -GQM estimé par la méthode MV	124
2.23	La dérivée de la densité au TIV zéro	125
2.24	Les cas particuliers au niveau du TIV zéro des modèles combinés aux débits faibles	126
2.25	Les cas particuliers au niveau du TIV zéro des modèles combinés aux débits denses	127
2.26	Les temps d'estimation et d'ajustement du modèle Hyper-GQM	132
2.27	L'ajustement du modèle DG	132
2.28	Les ajustements des modèles DL et LG	133
2.29	Les ajustements du modèle GQM mixte $\ln \mathcal{N}$ -GQM- Γ	133
2.30	Les temps t_e et t_a des modèles développés	133
2.31	L'ajustement graphique des meilleurs nouveaux modèles de TIV	134
3.1	La qualité du modèle Γ -GQM sans paramètre de décalage au niveau des TIV très courts ($TIV < 5$ s)	137
3.2	Les distributions du modèle Γ -GQM selon le paramètre θ pour différentes valeurs de α, β	139
3.3	Les coefficients statistiques du modèle Γ -GQM en fonction de θ	139
3.4	Les distributions du modèle Γ -GQM selon le paramètre λ pour différentes valeurs de α, β	140
3.5	Les coefficients statistiques du modèle Γ -GQM en fonction de λ	140

3.6	Les distributions du modèle Γ -GQM selon les paramètres α, β pour différentes valeurs de θ, λ	141
3.7	Les coefficients statistiques du modèle Γ -GQM en fonction de α et β	141
3.8	Densité du modèle Γ -GQM selon différentes relations entre α et β	142
3.9	La distribution du modèle Γ -GQM en fonction du paramètre α	142
3.10	Les distributions du modèle DG selon le paramètre θ	144
3.11	Les coefficients statistiques du modèle DG en fonction de θ	144
3.12	Les distributions du modèle DG selon le paramètre α_1, β_1 pour deux couples (α_2, β_2)	145
3.13	Les coefficients statistiques du modèle DG en fonction de α_1	145
3.14	Densité du modèle DG selon différentes relations entre les paramètres α_1 et les paramètres β_1, α_2 et β_2	146
3.15	Les cas bi-modaux de la distribution du modèle DG	147
3.16	Les distributions des paramètres du modèle Γ -GQM dans les échantillons bootstrap	149
3.17	Les distributions des paramètres du modèle DG dans les échantillons bootstrap	150
4.1	Variations des paramètres du modèle Γ -GQM en fonction de Q_6	156
4.2	Variation de λ du modèle Γ -GQM en fonction de Q_6 et de \mathcal{T}_6	157
4.3	Variations des paramètres du modèle DG en fonction de Q_6	158
4.4	Variations des paramètres du modèle Γ -GQM en fonction de \mathcal{T}_6	158
4.5	Variations des paramètres du modèle DG en fonction de \mathcal{T}_6	159
4.6	Variations des paramètres du modèle Γ -GQM en fonction de \mathcal{V}_6	160
4.7	Variations des paramètres du modèle DG en fonction de \mathcal{V}_6	161
4.8	Variations de θ du modèle DG en fonction de Q_6, \mathcal{T}_6 , et \mathcal{V}_6	161
4.9	Le diagramme fondamental et l'ajustement du modèle Γ -GQM pour les états élémentaires du trafic	162
4.10	Le diagramme fondamental et l'ajustement du modèle DG pour les états élémentaires du trafic	162
4.11	La sensibilité des paramètres du modèle Γ -GQM, selon \mathcal{T}_6	163
4.12	La sensibilité des paramètres du modèle Γ -GQM, selon Q_6	163
4.13	La sensibilité des paramètres du modèle DG, selon \mathcal{T}_6	164
4.14	La sensibilité des paramètres du modèle DG, selon Q_6	165
4.15	Comparaison des statistiques par temps sec et par temps de pluie	166
4.16	Comparaison des statistiques des TIV sur la VL et sur la VR	166
4.17	Les paramètres du modèle Γ -GQM selon les conditions de météo	167
4.18	Les paramètres du modèle DG selon les conditions de météo	168
4.19	Le paramètre θ du modèle DG sous influence des conditions exogènes (météo, type de voie)	168
4.20	Les paramètres du modèle Γ -GQM selon le type de voie, et la modalité jour-sec	169
4.21	Les paramètres $\alpha_1, \alpha_2, \beta_1, \beta_2$ du modèle DG selon le type de voie, et la modalité jour-sec	170
4.22	P-valeur du test K-S du modèle DG	170
4.23	P-valeur des modèles Γ -GQM et DG sur les échantillons raffinés de la RN118	172
4.24	Les paramètres du modèle Γ -GQM dans les échantillons raffinés	173
4.25	Les paramètres du modèle DG, dans les échantillons raffinés	174
4.26	P-valeur sur les données de l'A6	175
4.27	Variations des paramètres du modèle Γ -GQM en fonction du débit horaire Q_h	176
4.28	La relation entre les paramètres α et β du modèle Γ -GQM	176
4.29	Variations des paramètres du modèle DG en fonction du débit horaire Q_h .	177

4.30	Deux échantillons (L1 et R3c) ayant des débits horaires Q_h similaires sur la VL et la VR	178
4.31	Deux échantillons (L3b et L3c) de la VL au même niveau de Q_h	180
4.32	L'effet de la réduction de la largeur de voie sur la distribution des vitesses	180
4.33	Deux couples d'échantillons ((L6c,L2) et (L4, L11)) avec des valeurs de α, β équivalentes	182
4.34	Les ajustements du modèle Γ -GQM aux données empiriques pour cinq échantillons de débit Q_6 intermédiaire	183
4.35	Les distributions des TIV selon le modèle Γ -GQM ajusté pour cinq échantillons de débit Q_6 intermédiaire	184
4.36	Les distributions des TIV selon le modèle Γ -GQM ajusté pour trois échantillons de débit Q_6 très faible	184
4.37	L'ajustement du modèle Γ -GQM pour trois échantillons de débit Q_6 très faible	185
4.38	L'ajustement du modèle Γ -GQM aux données empiriques pour cinq échantillons de débit Q_6 élevé	185
4.39	La distribution des TIV selon le modèle Γ -GQM ajusté pour cinq niveaux de débit Q_6 élevé	186
4.40	L'ajustement du modèle Γ -GQM aux données empiriques pour les échantillons ayant un histogramme bi-modal	187
4.41	Les échantillons ayant un histogramme bi-modal	187
4.42	L'ajustement des modèles Γ -GQM et DG le soir en fonction du débit Q_6 et selon la condition météorologique	188
5.1	Les paramètres d'intensité ρ et λ des modèles \mathcal{E} des TIV normaux et des TIV en groupe en fonction du débit Q (Source : Athol [1])	194
5.2	Moyenne \bar{L} et écart-type σ_L des longueurs en fonction de la DIV	197
5.3	Moyenne \bar{V} et écart-type σ_V des vitesses en fonction de la DIV	198
5.4	Moyenne \bar{V}_r et écart-type σ_{V_r} des vitesses relatives en fonction de la DIV	198
5.5	Les coefficients ρ_{LL} , ρ_{LV} et ρ_{LV_r} en fonction de la DIV	199
5.6	Les coefficients ρ_{VV} et ρ_{VV_r} en fonction de la DIV	200
5.7	Moyenne \bar{L} et écart-type σ_L des longueurs en fonction du TIV	200
5.8	Moyenne \bar{V} et écart-type σ_V des vitesses en fonction du TIV	201
5.9	Moyenne \bar{V}_r et écart-type σ_{V_r} des en fonction du TIV	201
5.10	Les coefficients ρ_{LL} , ρ_{LV} et ρ_{LV_r} en fonction du TIV	202
5.11	Les coefficients ρ_{VV} et ρ_{VV_r} en fonction de la TIV	202
5.12	Moyenne \bar{L} et écart-type σ_L en fonction de la vitesse	203
5.13	Moyenne \bar{H} et écart-type σ_H des TIV en fonction de la vitesse	203
5.14	Moyenne \bar{V}_r et écart-type σ_{V_r} en fonction de la vitesse	204
5.15	Le coefficient ρ_{HV_r} en fonction de la vitesse	204
5.16	Les distributions des TIV en peloton dans différentes périodes, empiriques et selon le modèle Γ -GQM	206
5.17	Les distributions des TIV sur les deux voies (VL et VR), empiriques et selon le modèle Γ -GQM	207
5.18	Différents modèles calibrés et les distributions empiriques des TIV en peloton, VR, jour-sec	208
5.19	Les distributions des TIV, empiriques et selon le modèle Γ -GQM, selon les configurations, modalité jour-pluie sur la VL	209
5.20	La supériorité des ajustements du modèle DG sur ceux du modèle Γ -GQM pour la configuration (PL-PL) sur la VL	213

5.21	L'ajustement du modèle Γ -GQM et les distributions des TIV sachant le TIV du véhicule précédent, jour-sec, VL	213
5.22	Moyennes des TIV dans les échantillons groupés par vitesse instantanée v	215
5.23	Écarts-types des TIV dans les échantillons groupés par vitesse instantanée v	215
5.24	CV des TIV dans les échantillons groupés par vitesse instantanée v	216
5.25	Les paramètres α_1 et α_2 du modèle DG	216
5.26	Les paramètres β_1 et β_2 et θ du modèle DG	217
5.27	Les paramètres du modèle Γ -GQM dans la période du jour, VL	217
5.28	Les paramètres du modèle Γ -GQM dans la période du soir, VL	218
5.29	P-valeur du test K-S et du test χ^2 , modèle Γ -GQM de la variable RHIS	219
5.30	Les P-valeurs de test K-S du modèle Γ -GQM pour la variable RHIS en fonction du débit Q_6	219
5.31	Les densités de probabilité du modèle Γ -GQM pour la variable RHIS de l'ensemble test	220
5.32	Les paramètres du modèle Γ -GQM pour la variable RHIS en fonction du débit Q_6 selon les conditions de météo	220
5.33	La densité de probabilité du modèle Γ -GQM pour la variable RHIS, sur la VL de la RN118	221
6.1	La validation de statistiques partielles, type I et type II	225
6.2	Validation d'une statistique totale sur les échantillons glissants	228
6.3	Validation d'une statistique partielle de type II pour les échantillons glissants, critère de similitude p_2 des TIV $< 2 \text{ s}$	230
6.4	Validation d'une statistique partielle de type II pour les échantillons glissants, critère de similitude p_7 des TIV $\geq 7 \text{ s}$	231
6.5	Validation d'une statistique partielle de type II pour les échantillons glissants, critère de similitude $2 \leq \text{TIV} < 7 \text{ s}$	232
6.6	Les ensembles de courbes de la fonction de répartition dans les classes de débit	232
6.7	Les ensembles de courbes de la fonction de densité dans les classes de débit	233
6.8	Les ensembles de courbes de la fonction de répartition (A6 et RN118)	234
7.1	Les valeurs des TIV sur une route départementale des Côtes d'Armor	239
A.1	Les échantillons sur les 3 voies dans la même période, Autoroute A6	252
A.2	Relation K-S ² des TIV du soir en fonction de Q_6 , RN118, Site N°1	253
A.3	CV des TIV du soir en fonction de \mathcal{T}_6 , RN118, Site N°1	253
A.4	Validation d'une statistique totale sur les échantillons glissants, RN118, Site N°1	254

Notations

TIV	Temps InterVéhiculaire
h_i	Le TIV normal (time headway) du véhicule i
g_i	Le TIV net (time gap) du véhicule i
p_i	Le TIV tracé du véhicule i
e_i	Le TIV embarqué du véhicule i
$F(h)$	La fonction de répartition de probabilité des TIV
$f(h)$	La fonction de densité de probabilité des TIV, PDF
f^L	La transformation de Laplace de la fonction f
f^F	La transformation de Fourier de la fonction f
f^K	Fonction de noyau de la fonction f
$f_1(t), F_1(t)$	Les fonctions de répartition et de densité de probabilité des TIV contraints
$f_2(t), F_2(t)$	Les fonctions de répartition et de densité de probabilité des TIV libres
K-S ²	La relation entre le coefficient d'aplatissement (Kurtosis) et le carré du coefficient de symétrie (Skeness)
$\mathcal{P}(\alpha, \beta, \dots)$	Le modèle \mathcal{P} (en général) des paramètres α, β, \dots
\mathcal{P}_τ	Le modèle \mathcal{P} décalé de τ secondes à droite de l'abscisse
\mathcal{P}^*	Le modèle Hyper- \mathcal{P} basé sur le modèle \mathcal{P}
\mathcal{E}	La loi exponentielle
Γ	La loi Gamma
Γ_τ	Le modèle Pearson III
\mathcal{B}	La loi Bêta
\mathcal{N}	La loi Normale
$\log \mathcal{N}$	La loi log-normale
\mathcal{E}^*	Le modèle Hyperexponentiel (ou bien le modèle de Schuhl)
\mathcal{N}^*	Le modèle Hypernormal
\mathcal{G}^*	Le modèle Hypergamma
$\log^* \mathcal{N}$	Le modèle Hyperlognormal
\mathcal{C}	Le modèle M3 de Cowan
GQM, \mathcal{G}	Le modèle de file d'attente généralisée
SPM, \mathcal{S}	Le modèle semi-Poisson
$\mathcal{S}_\mathcal{P}$	Le modèle semi-Poisson utilisant la loi \mathcal{P} pour la fonction $f_1(h)$
$\mathcal{Q}_\mathcal{P}$	Le modèle GQM utilisant la loi \mathcal{P} pour la fonction $f_1(h)$
test K-S	Test de Kolmogorov-Smirnov
test χ^2	Test du chi-deux
test A-D	Test d'Anderson-Darling
VP	Véhicule particulier
PL	Poids lourds
θ	La part des véhicules contraints
α	Paramètre d'échelle de la loi Gamma

β	Paramètre de forme de la loi Gamma
λ	Paramètre d'intensité d'une loi exponentielle
τ ou τ_i	Paramètres de décalage
μ, σ	Moyenne et écart-type des lois normale et lognormale
MM	Méthode des Moments
MV	Méthode du Maximum de Vraisemblance
MC	Méthode du Minimum du χ^2
véh/h/voie	Véhicule par heure et par voie
véh/s/voie	Véhicule par seconde et par voie
véh/h	Véhicule par heure
véh/s	Véhicule par seconde
Q	Le débit (par défaut sur une voie) en véh/h
q	Le débit (par défaut sur une voie) en véh/s
q_t, Q_t	Débit sur la totalité des voies
Q_h	Débit horaire
Q_r	Débit d'un échantillon raffiné
K	La concentration ou la densité du trafic
\mathcal{T}	Le taux d'occupation
\mathcal{V}	La vitesse moyenne harmonique d'un flux
v_r	La vitesse relative d'un véhicule par rapport au véhicule précédent
v_i	La vitesse instantanée mesurée d'un véhicule i
ρ_k	L'auto-corrélation d'ordre k
.ang	Mot anglais correspondant
t_e	Temps d'estimation
t_a	Temps d'ajustement
DG	Modèle Double Gamma
DL	Modèle Double Lognormal
LG	Modèle Lognormal Gamma
p_g	Pourcentage des véhicules en groupe
δ_H	Seuil critique selon TIV de peloton et/ou de véhicule libre
δ_d	Seuil critique selon DIV de peloton et/ou de véhicule libre
α_l	Pourcentage des véhicules libres
n_v	Nombre de voies de la route ou d'une section de la route
R^2	Coefficient de régression linéaire
\mathcal{C}_h	Groupement des véhicules selon les valeurs de TIV
\mathcal{C}_d	Groupement des véhicules selon les valeurs de DIV
\mathcal{C}_g	Groupement des véhicules selon les valeurs de TIV net
\mathcal{C}_v	Groupement des véhicules selon les valeurs de vitesse instantanée
P_ℓ	Véhicule libre
P_n	Véhicule suiveur en n -ème position en peloton
VP-VP	Configuration de poursuite d'un VP suivant un VP
VP-PL	Configuration de poursuite d'un VP suivant un PL
PL-VP	Configuration de poursuite d'un PL suivant un VP
PL-PL	Configuration de poursuite d'un PL suivant un PL
p_2	Pourcentage des TIV inférieurs à 2 secondes
p_7	Pourcentage des TIV supérieurs à 7 secondes

Introduction

La motorisation et le développement de l'utilisation des véhicules ainsi que l'extension de différents réseaux routiers font du trafic en général et plus particulièrement de l'ingénierie du trafic un sujet central dans la recherche sur les transports et leur sécurité. Deux principales approches sont possibles : la première consiste à considérer qu'un ensemble de véhicules forme un flux de trafic. Ce flux est caractérisé par des grandeurs macroscopiques comme le débit, la vitesse moyenne, le taux d'occupation, etc. Ce point de vue permet d'aborder le trafic via certains indicateurs concis et faciles à établir. La seconde approche regarde le trafic à un niveau plus détaillé, jusqu'aux véhicules individuels, disons microscopique. Dans cette optique, les véhicules eux-mêmes et leurs interactions sont l'objet d'étude. Les valeurs des variables attachées à chacun des véhicules varient donc de l'un à l'autre, dépendant des comportements de conduite et des facteurs exogènes. On peut considérer que ces variables sont des variables aléatoires, notamment le temps intervéhiculaire (TIV). Le TIV est l'intervalle en temps entre deux véhicules routiers consécutifs sur la même voie. Son caractère aléatoire est confirmé et étudié depuis plusieurs années.

Revue globale sur la modélisation du trafic

Le but de la modélisation du trafic est de proposer des modèles mathématiques qui sont susceptibles de décrire les phénomènes observés dans la réalité. Pour cela, le trafic doit être caractérisé par des variables mesurables qui sont accessibles sous forme de données réelles. Étant donné que les problématiques du trafic sont différentes selon le niveau du trafic, les variables fondamentales prises en compte varient en fonction de l'échelle du trafic considérée. En général, il y a deux approches principales décrivant les phénomènes du trafic. La première approche est macroscopique basée sur les variables agrégées à partir d'un ensemble de véhicules, donc un flux du trafic. Les trois variables macroscopiques fondamentales sont le débit Q (véh/h), la concentration K (véh/km) et la vitesse moyenne harmonique \mathcal{V} (m/s). Les relations entre ces variables sont régies par analogie avec l'hydrodynamique, principalement par des équations aux dérivées partielles de type *hyperbolique* dans lesquelles la loi de conservation des véhicules est respectée. Selon l'état du trafic considéré *en équilibre* ou *hors équilibre*, le modèle macroscopique du premier ordre de Lightwill-Whitham-Richards (LWR) ou les modèles macroscopiques d'ordre supérieur sont respectivement proposés. Il est à noter que dans le régime d'équilibre, il y a une *relation d'état* entre la vitesse \mathcal{V} et la concentration K , $\mathcal{V} = \mathcal{V}_e(\rho)$, tandis que dans le régime hors équilibre, l'accélération due aux anticipations et aux relaxations est prise en compte en complément. L'approche macroscopique permet de décrire des phénomènes du trafic à l'échelle supérieure à quelques minutes et de s'adapter à la représentation des réseaux de grande taille.

La deuxième approche est microscopique, réalisée à travers des variables associées à chaque véhicule individuel. La variable fondamentale microscopique de cette approche est la position $x_i(t)$ à tout instant t considéré du véhicule i . Dans le cas où toutes les positions x_i sont connues à tout instant t , les trajectoires des véhicules i sont obtenues et peuvent

être tracées sur le plan spatio-temporel (x, t) , comme le montre la Figure 1. Dans ce cas -là, les vitesses instantanées $v_i(t)$ et les accélérations $a_i(t)$ de tous les véhicules i à tout instant t sont aussi déductibles des valeurs de $x_i(t)$. On a donc toutes les informations de la réalité du trafic.

Le mouvement des véhicules se réalise dans le temps et l'espace. Comme les véhicules ont toujours des vitesses v_i positives, ceci leur permet d'avancer sans arrêt ou sans marche arrière, les trajectoires dans la Figure 1 sont par conséquent toujours des courbes croissant de façon monotone. Le cas d'une information totale précitée n'existe que dans la théorie car les courbes de trajectoire ne sont pas généralement modélisées simplement par une fonction $f(x, t)$, d'autant plus qu'il faut mettre en jeu plusieurs véhicules et que le temps est une variable continue. En conséquence, les données de trafic sont recueillies soit à un instant donné, soit à un point donné de la route, soit dans une *fenêtre* en discrétisant l'espace et le temps en des pas spatio-temporels prédéterminés. Notons dans la Figure 1 que le véhicule $i - 1$ précède le véhicule i , et x_* signifie le point de mesure en espace selon le temps, tandis que t_* est l'instant de mesure des données sur l'espace.

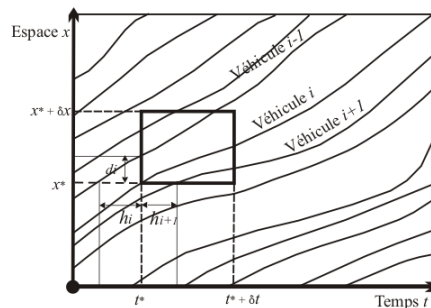


FIGURE 1 : Les trajectoires des véhicules dans le plan spatio-temporel

Selon la méthode de recueil de données, les variables correspondantes se caractérisent différemment. Les deux premières méthodes sont plus utilisées que la troisième :

1. En un point de mesure x_* , les vitesses v_i , les accélérations a_i des véhicules i et d'autres variables sont mesurées pendant une période de temps I . Elles sont appelées les *variables temporelles*,
2. À un instant donné t_* , sur un tronçon de route, les positions $x_i(t_*)$, les vitesses $v_i(t_*)$, les accélérations $a_i(t_*)$ et d'autres variables sont calculées. Les variables sont appelées les *variables spatiales*.

Adams [2] considère le mouvement des véhicules sur les routes comme celui des *points* le long d'une *ligne*. Les notions de *point* et de *ligne* peuvent se comprendre différemment en fonction du point de vue de l'observateur. Comme dans la méthode de recueil de données, deux points de vue sont considérés : soit la *ligne* représente la longueur de la route sur laquelle les *points* sont les positions des véhicules à un instant t_* donné, soit elle représente un emplacement déterminé durant une période de temps, les *points* étant des instants de passage des véhicules. Parmi ces deux points de vue, le deuxième est celui adopté dans notre étude.

Dans la littérature, cette deuxième méthode de mesure est la plus utilisée puisque les études théoriques et pratiques sont souvent fondées sur une base de données locales, recueillies via les stations de mesure installées au bord de la route. Ces stations fixes enregistrent automatiquement des données avec une précision dépendant des méthodes et techniques utilisées. Dans ces stations, le système ne mesure pas seulement les instants

d'arrivée (donc le TIV), mais aussi d'autres variables des véhicules telles que : la longueur et la vitesse instantanées de chacun. De plus, certains systèmes proposent également le calcul et l'enregistrement des caractéristiques du trafic comme le débit agrégé Q_I , le taux d'occupation agrégé \mathcal{T}_I , la vitesse agrégée \mathcal{V}_I , etc.

Il est à noter en premier lieu que la longueur l_i des véhicules est, elle aussi, une variable microscopique. Pourtant, cette variable est constante pour chaque véhicule tant en fonction de temps qu'en fonction d'espace. En deuxième lieu, à la différence du modèle macroscopique, le modèle microscopique décrit des interactions entre les véhicules ce qui permet d'approcher des éléments plus ponctuels du réseau plutôt que l'ensemble global du réseau. Par conséquent, l'approche microscopique est liée davantage aux *comportements individuels* des conducteurs dans des contextes précis plutôt qu'aux *comportements collectifs* des flux du point de vue macroscopique. La distribution des TIV peut jouer le rôle d'un des liens de ces deux types de comportement selon différents contextes.

Par ailleurs, on parle aussi d'un troisième niveau d'agrégation, intermédiaire entre les échelles macro et microscopique, ce que l'on appelle l'échelle *mésoscopique*. Cette approche concerne par exemple le *peloton* défini comme un petit groupe de véhicules. En pratique, il semble possible d'y ajouter aussi la configuration du trafic définie comme l'ensemble des catégories des véhicules successifs constituant le peloton. Plus précisément, la définition de peloton n'est pas unique, et varie en fonction de la définition quantitative du *véhicule libre* et du *véhicule contraint*. Qualitativement, un véhicule libre n'est influencé par aucun autre, au contraire des véhicules *contraints*. Cette approche est importante dans les études de trafic.

L'échelle microscopique

Si les sorties des modèles macroscopiques sont les états Q, K, \mathcal{V} du trafic en fonction du temps et de l'espace une fois que les conditions initiales du réseau sont connues, les sorties des modèles microscopiques sont les interactions entre les véhicules consécutifs. Comme la variable TIV étudiée est microscopique, et est le résultat des arrivées successives des véhicules, les modèles microscopiques sont passés en revue plus en détail dans cette section.

Les modèles microscopiques sont principalement les modèles *de poursuite* car à l'échelle microscopique, la source d'influence d'un véhicule est supposée venir uniquement du véhicule *en tête*. Soit $i + 1$ le véhicule derrière le véhicule considéré i , à l'instant t quelconque, $v_i(t), a_i(t)$ sont la vitesse et l'accélération du véhicule i . Dans ce cas, l'influence du véhicule i sur le véhicule $i + 1$ est traduite soit par la distance intervéhiculaire (DIV) $s_i(t)$, soit par l'interdistance $d_i(t)$, soit par la vitesse relative $v_{i+1}(t) - v_i(t)$. Selon les approches des différents auteurs, cette influence est caractérisée soit par des contraintes (une distance à respecter), soit par un modèle cinématique (l'accélération), soit par des modèles de comportement ou de décision.

Brackstone *et al.* [3] ont récapitulé les modèles de poursuite à partir de l'article de Pipes en 1953 [4] jusqu'aux années 1998. Les modèles de poursuite ont un rôle central dans les simulations microscopiques du trafic et dans des systèmes de transports intelligents. En effet, la puissante capacité de calcul numérique permet d'une part d'examiner les liens entre les phénomènes au niveau du véhicule individuel et des comportements à une échelle plus macroscopique, d'autre part d'intégrer les modèles microscopiques dans des dispositifs d'assistance et d'automatisation de la conduite (par exemple AVCSS, ACC), ou d'évaluer des stratégies de contrôle local. Les modèles de poursuite sont les suivants :

Les modèle de poursuite à distance de sécurité

Une approche tout à fait naturelle est de supposer que les conducteurs prennent conscience de la sécurité représentée par une distance nécessaire à garder par rapport au véhicule en tête [5]. Basée sur l'équation de mouvement de Newton, la distance intervéhiculaire (DIV) à respecter en ajustant un temps de relaxation T est la suivante :

$$s_i(t - T) = \alpha v_{i-1}^2(t - T) + \beta_1 v_i^2(t) + \beta v_i(t) + b_0 \quad (1)$$

Cette approche a ensuite été développée par Gipps [6] qui considérait un temps supplémentaire de réaction de sécurité égal à $T/2$ et qui a inséré un terme de décélération de freinage b_i dont la valeur maximale est de l'ordre de -3 m/s^2 . Malgré des problèmes encore existants, le modèle est simple, pratique et utilisé dans plusieurs logiciels de simulation microscopique.

Les modèle de poursuite à fonction d'accélération

En réalité, l'influence du véhicule en tête $i-1$ sur le véhicule considéré i joue le rôle d'un *stimulus* qui entraîne des changements de vitesse, donc une accélération ou une décélération a_i . Cette approche à *stimuli* a été proposée pour la première fois par Chandler *et al.* [7]. Le modèle le plus simple est le modèle linéaire proposé par Helly [8].

$$\begin{aligned} a_i(t + T) &= C_1(v_{i-1} - v_i) + C_2([x_{(i-1)}(t) - x_i(t)] - D_i(t + T)) \\ \text{où } D_i(t + T) &= \alpha + \beta v_i(t) + \gamma a_i(t) \end{aligned} \quad (2)$$

Le modèle le plus renommé dans les années 60–70s est celui de Gazis-Herman-Rothry (GHR) dont la formule est la suivante :

$$a_i(t + T) = c v_i^m(t + T) \frac{v_{i-1}(t) - v_i(t)}{[x_{i-1}(t) - x_i(t)]^l} \quad (3)$$

La *sensibilité* du modèle est représentée par le coefficient c qui dépend de la phase d'accélération ou de décélération. Elle dépend aussi de la DIV s_i et de la vitesse du véhicule i tandis que le *stimulus* est la vitesse relative. Dans les modèles de poursuite, le temps *de réaction* T a un rôle important car il reflète en réalité le temps *de perception*, *de décision* et *d'application* [9].

L'idée de la régulation de vitesse selon différentes stratégies en poursuivant le véhicule en tête permet aux modèles de poursuite de décrire le mécanisme de façon dynamique dans le trafic, de modéliser les comportements des conducteurs et donc d'obtenir les TIV et les vitesses en fonction du temps. Une fois les conditions initiales données, le TIV peut être déterminé de façon dynamique sous la forme d'une fonction du temps qui varie dans la *phase de transition* et tend vers une valeur TIV *d'équilibre*.

Malgré des efforts pour combiner le modèle linéaire de Helly et le modèle GHR [10], de nombreuses faiblesses de cette approche ont été montrées dans [9, 11, 12] ; elles portent notamment sur :

1. Le modèle n'est applicable que dans la *phase de poursuite* mais le seuil entre cette phase et la *phase de conduite libre* n'est pas bien indiqué. En effet, chaque paramétrage du modèle GHR correspond à une relation fondamentale $Q-K$. Le modèle rencontre des problèmes lorsque la concentration est faible ou très élevée.
2. La description du phénomène de dépassement est limitée.

3. Selon Leutzbach, la sensibilité du modèle en *phase d'accélération* doit différer de celle de la *phase de décélération*.
4. Bien que le TIV *d'équilibre* existe, cette valeur semble varier selon la condition initiale tandis que le modèle ne montre pas explicitement une telle *fonction d'équilibre* de sécurité comme dans le modèle à distance de sécurité.
5. Dans la réalité, la réaction de poursuite n'est pas déterministe, précise ; elle ne change pas de façon continue car il existe un seuil de perception face aux changements de distance et de vitesse.

Le modèle à vitesse optimale développé par Bando *et al.* [13] considère que l'état d'équilibre se produit lorsque le véhicule *suiveur* adapte sa vitesse à celle du *leader* . En conséquence, la régulation de l'instabilité lors d'une poursuite est ajustée par la différence entre la vitesse pratiquée et la vitesse *désirée*. Cette différence est une notion représentant l'état stable de sécurité et est une fonction de l'interdistance :

$$a_i(t) = \frac{1}{\tau} [v_e(d_i(t)) - v_i(t)] \quad (4)$$

La fonction $v_e(d_i(t))$ est appelée la fonction de vitesse *optimale* et peut adopter différentes formes. Le modèle de vitesse optimale (OVM) est ensuite amélioré par Lenz *et al.* [14] et Juang [15] en introduisant des temps de relaxation.

Un modèle de poursuite largement étudié est le second modèle de Newell [16] qui porte sur la poursuite selon les trajectoires de deux véhicules consécutifs. Le modèle des conducteurs intelligents (IDM) est développé par Treiber *et al.* [17]. Dans ce modèle, la loi d'accélération est une fonction de la vitesse, de l'interdistance d_i et de la vitesse relative. La manipulation des relations entre les composantes permet au modèle IDM de s'adapter dans différentes situations du trafic telles que : (1) *En conduisant librement*, (2) *En trafic dense*, (3) *En approchant un obstacle de loin*, (4) *En freinant en urgence*.

Récemment, une autre approche intéressante proposée par Tordeux *et al.* [12] a incorporé le TIV (sous forme du TIV net) dans le processus de poursuite. Une équation différentielle d'adaptation est ajustée :

$$\frac{dT_i}{dt} = \lambda_i(t) [\mathcal{T}(v_i(t)) - T_i(t)] + \sigma \epsilon_i(t) \quad (5)$$

où $\mathcal{T}(v_i(t))$ est une fonction d'équilibre de sécurité de la vitesse. Le TIV est supposé s'ajuster vers la fonction d'équilibre de façon exponentielle. Le terme $\epsilon_i(t)$ représente un *biais aléatoire*.

Les autres modèles

L'approche psychologique du problème de modélisation microscopique du trafic est abordée pour la première fois par Michaels [18]. Wiedemann [19] propose ensuite de définir quatre situations de conduite dont (1) *Conduite libre* (2) *Approche* (3) *Freinage* et (4) *Poursuite stable*. Différents modèles *psycho-physiques* sont utilisés dans quelques logiciels en Grande Bretagne et en Allemagne. L'approche psycho-physique prend en compte des seuils de perception des conducteurs [20].

Le modèle en logique floue est initialement proposé par Kikuchi *et al.* [21]. À l'entrée du modèle, des ensembles flous sont tout d'abord définis. Ensuite, les sorties des actions, représentées aussi par des ensembles flous, sont obtenues via les opérations logiques correspondantes.

Le modèle des automates cellulaires se base sur un algorithme incorporant des éléments aléatoires pour simuler le trafic. L'algorithme utilise une discrétisation du temps et de

l'espace et essaie de décrire les comportements des conducteurs. Cette approche est analysée plus en détail dans un autre chapitre.

La modélisation du trafic a aussi connue l'approche de discrétisation correspondant à une chaîne de Markov à état discret [22]. Cette méthode et son estimation ont été récemment appliquées en employant une matrice de transition de dimension plus grande que celle de Dunne *et al.* [23].

En somme, les modèles microscopiques sont peu ou prou liés à la variable TIV. Pourtant, le lien est indirect et on tire peu de résultats directs pour les TIV.

Approche statistique et l'importance du TIV dans le domaine du trafic

Les modèles dynamiques et déterministes présentés ci-dessus ont connu un plein essor jusqu'à nos jours grâce notamment au développement de la capacité de calcul numérique ; celui-ci permet de simuler les données, de discrétiser les équations continues dans les théories du trafic. Par contre, l'approche statistique générale porte sur la distribution des valeurs des variables fondamentales en considérant éventuellement les paramètres de type *conducteur-véhicule-infrastructure*. En conséquence, les données recueillies jouent un rôle essentiel dans cette approche tant dans les phases de modélisation que dans les phases de validation statistique et probabiliste. Malgré le développement des méthodes et des techniques d'enregistrement, ainsi que de la capacité de stockage, l'approche statistique attire cependant moins l'attention des chercheurs.

Les trois types principaux de modèle statistique cités ici sont le modèle de Prigogine-Herman, le modèle de Heidemann et les modèles probabilistes.

L'approche de Prigogine-Hermann [24, 25, 26] correspond aux modèles cinétiques. Elle utilise l'équation de Boltzmann pour expliquer la distribution microscopique $f(x, v, t)$ de la vitesse selon le temps et l'espace. À l'origine, le modèle de Prigogine-Hermann se limite dans les cas de faible concentration du trafic car l'analogie du flux de trafic avec les comportements des gaz était supposée. Le modèle est présenté comme suit :

$$\partial_t f + v \partial_x f = \frac{1}{\tau} (f - f_0) + \rho(\bar{v} - v)(1 - P)f \quad (6)$$

où $f(x, v, t)$ est la fonction de densité des véhicules dans une phase espace, c.à.d que N (cf. l'équation 7) est le nombre de véhicules à l'instant t se trouvant entre les positions x_1 et x_2 ayant la vitesse de v_1 à v_2 .

$$N = \int_{v_1}^{v_2} \int_{x_1}^{x_2} f(x, v, t) dv dx \quad (7)$$

En conséquence, la concentration à l'instant t au point x est :

$$\rho(x, t) = \int_0^{\infty} f(x, v, t) dv \quad (8)$$

et $\bar{v}(x, t)$ est la vitesse moyenne des véhicules, τ est le temps de relaxation et f_0 est la concentration à vitesse désirée sans oublier que P est la possibilité de dépassement.

La fonction de densité des probabilités ici du point de vue ingénierie du trafic est considérée sur un échantillon en espace à un instant donné t .

Nelson [27] a montré deux propriétés importantes du modèle Prigogine-Herman en accord avec les données réelles :

1. La solution de l'équation de Prigogine-Herman de vitesse à l'état d'équilibre est bimodale
2. En état d'équilibre du trafic à haute concentration, le débit du trafic est *éparpillé*.

L'approche de Prigogine-Herman est critiquée et améliorée par Munjal et Pahl [28] dans les situations où la file d'attente ou le phénomène de **stop-and-go** se présentent. Pavari-Fontana [29] critique le modèle sur le fait d'utiliser la vitesse désirée d'équilibre comme une variable. Des équations cinétiques alternatives ont ensuite été proposées. Cependant, les modèles développés par Pavari-Fontana, par Phillips [30] et par Klar & Wegener (1999) [31] ne satisfont pas les deux caractéristiques (*benchmarks*) citées par Nelson. Nelson (1995) [27] a lui aussi proposé un modèle satisfaisant une seule condition. En employant la proposition de Pavari-Fontana, Hoogendoorn et Bovy ont calculé la solution numérique de l'équation cinétique en utilisant la méthode de Monte-Carlo. Le lien entre les échelles micro et macroscopique est aussi réalisé via le modèle cinétique par Helbing [32, 33] utilisant le modèle développé par Pavari-Fontana. Nelson et Sopasakis (1999) [34] ont aussi développé des modèles macroscopiques basés sur le modèle cinétique et utilisés des expansions de Hilbert et de Chapman-Enskog.

L'approche cinétique est compliquée. Comme elle n'a pas de lien direct avec le TIV, elle n'est pas abordée dans la thèse.

Le modèle de Heidemann a porté sur l'approche statistique. Plus simple que les modèles cinétiques, le modèle se base sur des hypothèses stochastiques entre les états du trafic et des conditions initiales. Les états du trafic sont caractérisés par un couple de valeurs (ρ, i) , de concentration et de classe de vitesse i . Cette approche sera abordée plus en détail dans une section d'un chapitre de la thèse.

L'approche générale probabiliste recherche des distributions directes des valeurs supposées aléatoires du trafic. Les variables concernées sont microscopiques et observables dans les expérimentations. C'est aussi l'approche principale de cette thèse. La question permanente de l'approche probabiliste est : *Quelle est la probabilité d'une telle valeur microscopique T du trafic dans des contextes différents ?*. Pour répondre à cette question, beaucoup de données et de modèles doivent être utilisés et testés, d'autant plus que les contextes pris en compte dans le trafic sont très nombreux.

$$P(t) = \Pr\{ T \leq t \mid \text{contexte donné} \} \quad (9)$$

La variable Temps Inter-Véhiculaire (TIV)

Complément sur la définition

L'intervalle de temps entre deux véhicules successifs sur une route, la définition précise du TIV dépend des moyens de mesure :

- Lorsqu'un instrument mesurant la distance est embarqué à l'avant d'un véhicule (radar, lidar, télémètre), le TIV est calculable à tout instant pour ce véhicule. Il est égal au rapport entre la distance du véhicule précédent et la vitesse du véhicule.
- Lorsqu'un capteur routier est installé en bord de route, au point x_* , l'instant de passage de chaque véhicule à ce capteur est enregistré ; il s'agit du passage soit de l'avant, soit de l'arrière du véhicule ; le TIV est obtenu comme différence des temps de passage de deux véhicules successifs (avec éventuellement une correction tenant compte de la longueur et de la vitesse de l'un de véhicules).

En conséquence, les notions de TIV suivantes sont distinguées, illustrées par la Figure 2 :

Le TIV normal, TIV (*time headway*) h_i , est l'intervalle en temps entre les arrivées des véhicules consécutifs sur un point donné sur la même voie de la route. Comme la dimension des véhicules l_i n'est pas négligeable, les instants considérés dans la notion de TIV h_i sont mesurés aux mêmes points de référence de tous les véhicules, souvent, de l'avant d'un véhicule à l'avant du véhicule suivant.

Le **TIV net** (*time gap*) g_i , est aussi une variable temporelle, mesurée par la différence en l'instant de l'arrivée de l'arrière d'un véhicule au point de mesure à l'instant d'arrivée de l'avant du véhicule suivant en le même point de mesure.

Le **TIV à bord, TIV embarqué** $e_i(t)$, est une notion dynamique et une variable *spatiale*. Le TIV e_i est associé à un véhicule particulier i et varie en fonction du temps. L'interdistance entre l'avant du véhicule i et l'arrière du véhicule précédent $i - 1$ est facilement mesurée par un radar par exemple. Le TIV $e_i(t)$ est déduit par la prise en compte de la vitesse relative entre les deux véhicules.

Le **TIV tracé** p_i , est une variable temporelle concernant particulièrement les véhicules équipés de dispositifs permettant aux systèmes d'identification de les reconnaître. En un point donné du bord de la route multi-voies, p_i est mesurée par la différence de temps des arrivées des avants de deux véhicules traceurs successifs (quelle que soient leur voies). Cette notion permet donc des valeurs nulles du TIV.

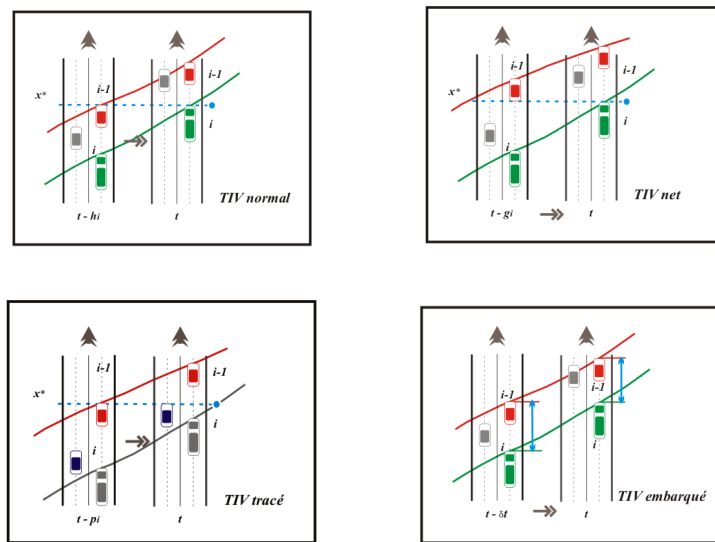


FIGURE 2 : Les définitions du Temps Inter-Véhiculaire

Le fait de choisir la distance en temps plutôt qu'en espace est donné par les trois raisons suivantes (cf. Luttinen [11]) :

1. Le TIV est relatif au débit alors que la DIV concerne la concentration. Selon Luttinen, le débit est plus usuel, a plus de sens en trafic et est plus facile à mesurer.
2. En peloton ou dans les situations de poursuite, le TIV net et le TIV ne sont pas aussi sensibles à la vitesse des véhicules que la DIV
3. Pour estimer la DIV, plus de calculs sont demandés. Il faut aussi connaître la vitesse et supposer que cette vitesse est constante durant la période de mesure.

Luttinen se place du point de vue temporel en regardant le trafic en un point fixé. Dans cette optique, la troisième raison n'est autre qu'une partie de la première, c. à. d la facilité de mesure de TIV. La deuxième remarque est renforcée par le fait que le temps de réaction - étant une variable très importante dans les modèles microscopique, est un composant du TIV.

En fonction du but de la recherche et selon la base de données disponible, la variable TIV appropriée sera considérée. Il est à noter que le TIV embarqué est pour un seul véhicule tandis que le TIV tracé concerne seulement une partie de tous les TIV du flux du trafic. Par

ailleurs, le TIV embarqué et le TIV tracé demandent des équipements techniques spéciaux pour diffuser et recevoir des informations.

Enfin, on met l'accent sur deux avantages de l'utiliser le TIV (normal) par rapport à l'utilisation de la notion du TIV net :

1. Bien que le TIV net soit lié plus directement à certains indicateurs de sécurité routière, cette variable est l'une des deux composantes du TIV. La deuxième composante est le temps pendant lequel les véhicules passent sur le capteur. Le calcul du TIV net est donc plus complexe et comporte plus de risque d'erreur sur l'amplitude de quelques secondes du TIV. Des valeurs négatives peuvent être proposées si l'on sur-estime par exemple la longueur du véhicule précédent et la vitesse du véhicule en considération.
2. La relation directe entre le TIV et le débit permet une première connection entre les échelles micro et macroscopique.

Par ailleurs, les distributions des deux variables sont en fait similaires malgré le biais de la composante temps *d'occupation*. En conséquence, cette étude se concentre sur la variable TIV normale - sujet intéressant dès le début de la modélisation du trafic dans les années 1930.

Les approches d'étude du TIV

Parmi les variables du trafic, le TIV est l'une des variables microscopiques essentielles ; elle caractérise directement et intuitivement les comportements au niveau du véhicule individuel. En général, l'instant de passage d'un véhicule en un point donné (le point de mesure) est considéré comme une variable aléatoire. Ceci veut dire que dans les mêmes conditions de trafic et les mêmes conditions environnementales, l'arrivée d'un véhicule ne peut pas être exactement prédite. À partir de l'instant 0 (considéré comme point de repère), si on désigne par N_t le nombre de véhicules arrivés jusqu'à un instant t quelconque, ce nombre dépend des instants d'arrivée T_i des véhicules :

$$N_t = \max \{i | T_i \leq t\} \quad (10)$$

Les deux possibilités d'analyse de ce type de processus sont d'une part l'analyse des nombres de passage N_t (*analyse de dénombrement*), et d'autre part l'étude des caractéristiques statistiques des intervalles temporels entre les T_i , donc des TIV. Cette seconde approche concerne directement les TIV et est justifiée notamment lorsque les TIV peuvent être supposés indépendants et distribués identiquement (i.i.d). Quand une série d'événements possède cette propriété, le processus est dit processus de *renouvellement* - une généralisation du processus de Poisson. Il est à noter que, dans un processus d'événements, le choix de l'instant de début est primordial. Cet instant est soit arbitraire, soit coïncide avec un événement quelconque.

Dans les modèles simples comme ceux de Poisson ou d'Erlang, le passage d'une approche à l'autre est possible. Dans le cas général, c'est l'approche par intervalle qui facilite la modélisation et la représentation mathématique. En conséquence, on privilégie le deuxième cas, relatif aux TIV étudiés.

Intérêt de la variable TIV

L'étude des TIV est considérée comme essentielle dans les théories et les applications du trafic routier. Cette variable est liée au débit à l'échelle macroscopique et concerne le comportement des conducteurs.

Comme le TIV comprend le TIV net, il est aussi une variable importante dans certaines études de sécurité routière. Des indicateurs individuels du risque sont souvent calculés par l'intermédiaire de cette variable comme la distance d'arrêt de sécurité, le temps de collision, etc. Lorsque le TIV est couplé avec la vitesse, le résultat donne aussi un indicateur important mesurant le risque de collision par l'arrière. En général, les TIV faibles peuvent être liés à un risque accru de collision. Aussi, les pourcentages des TIV courts sont des indicateurs de sécurité souvent utilisés. Ils sont aussi une méthode d'approximation des niveaux de service dans le HCM (1985).

Les TIV intéressent aussi les études d'exploitation de la route, de prédiction du flux de trafic et également les études de conception de l'infrastructure puisque le TIV est lié au débit et à la capacité du trafic. La régulation des TIV peut aussi faire partie des mesures d'exploitation routière soit pour améliorer le débit en harmonisant les arrivées, soit pour minimiser les temps d'attente au niveau des intersections signalisées [35, 36, 37, 11].

Le TIV peut être utilisé dans les systèmes de transport intelligent. Dans ce cas, les TIV embarqués et les TIV tracés semblent plus concernés car ils sont des informations que le système de transport intelligent traite (parmi d'autres) pour évaluer de façon dynamique les conditions de trafic et de sécurité.

Au niveau du comportement, le TIV est un facteur direct à prendre en compte lorsqu'un conducteur décide de réaliser des opérations comme, changer de voie sur une section droite, intégrer un flux à l'intersection, au rond-point [38, 39].

Au niveau des modèles microscopiques, les modèles déterministes permettent de simuler les mouvements des véhicules consécutifs en fonction du temps. Ils sont aussi intégrés dans des modèles macroscopique d'ordre supérieur. Pourtant, les mécanismes déterministes dans les modèles microscopiques sont *a priori* peu ou prou subjectifs, ce qui fait naître un doute sur la validation des résultats simulés par rapport à la réalité. En conséquence, une des applications de la recherche de la variable TIV est de vérifier la sortie de ces modèles; les distributions de TIV sont aussi utilisées à l'entrée d'un système simulé. L'étude du TIV est indispensable pour les simulations microscopiques du trafic tant pour générer des distributions initiales crédibles que pour valider la simulation.

Une longue histoire des études du TIV

May [40] a résumé assez complètement les travaux portant sur le TIV jusqu'aux années 1964 tant au niveau du recueil des données qu'au niveau de la modélisation probabiliste. Selon May, le processus de Poisson est le noyau des modèles pendant 30 années, interrompues d'ailleurs quelques années par la Deuxième Guerre Mondiale.

Les premiers travaux sont de Kinser [41], Rader [42], Adams [2] et Garwood [43] dont l'article d'Adams proposant le modèle exponentiel est le plus cité. Le modèle est ajusté aux données réelles de débit faible d'environ 500 véh/h. Les applications et le développement du modèle exponentiel sont ensuite réalisés notamment par Greenshields [44], Tanner [45], Gerlough [46] jusqu'en 1955.

En 1955, Schuhl [47] a proposé une combinaison de deux modèles exponentiels dont chacun représente soit la distribution des véhicules contraints soit la distribution des véhicules libres.

Les années 1950 sont aussi marquées par l'essor de la capacité de calcul numérique des ordinateurs. À cette époque, le besoin de génération des données du trafic est également un sujet important. Par conséquent, les distributions mathématiques ont joué un rôle important. Gerlough [48] et après Kell [49, 50], Schuhl [47] ont traité ce sujet de simulation numérique dans leur travaux. Gerlough a ainsi montré l'importance de l'utilisation du paramètre de décalage τ dans le modèle exponentiel du TIV. Kell a utilisé le modèle de

Schuhl et proposé que les paramètres soient fonction du débit.

Lewis [51] a proposé un modèle qui modifie la loi binomiale et divisé le modèle en deux cas. Ce modèle a été utilisé par l'auteur [52] dans la simulation des intersections mais il n'est plus développé et appliqué depuis.

Hight *et al.* [53] ont proposé une généralisation du modèle exponentiel en utilisant la loi d'Erlang. Le paramètre de cette dernière avait été espéré un indicateur de l'état du trafic fluide à régulier. Par la suite, des généralisations du modèle d'Erlang - le modèle gamma et le modèle Pearson - III, sont utilisées.

Jusqu'aux années 1960, deux catégories de modèles probabilistes sont proposées : Dans la première, la distribution des nombres de passages en un point donné est fournie en fonction de la période T ; la deuxième concerne la distribution des intervalles temporels entre les passages des véhicules. Bien que les deux types de distributions décrivent la même situation physique, la relation entre elles n'est pas toujours évidente, d'autant plus que le passage mathématique de l'une à l'autre est difficile et compliqué pour l'état du trafic qualifié d'*intermédiaire*, c. à. d ni fluide ni congestionné [53]. En effet, la distribution du nombre de passages est discrète alors que la distribution des intervalles est de nature continue. Par conséquent, cette dernière est de plus en plus utilisée. Par ailleurs, le noyau des modèles développés et modifiés dans cette période est le modèle exponentiel qui est théoriquement justifié pour le trafic en état aléatoire en temps et en espace [54].

Greenberg en 1966 [55] a proposé le modèle Log-normal et l'a mis en relation avec des lois de poursuite. En 1968, Dawson et Chimini [56] ont proposé le modèle Hyperlang. Buckley en 1968 [57] a proposé le modèle Semi-Poisson. Branston [58] et Cowan [59] ont proposé le modèle de file d'attente généralisée en 1976. Entre 1965 et 1975, de nombreux modèles puissants pour le TIV et de nombreux travaux comparatifs portant sur ces modèles et les méthodes numériques d'estimation correspondantes sont proposés (cf. Grecco [60], Tolle [61], Ashton [62], Wasielewski [63, 64]).

Dans les années 1980, peu de recherches sur le TIV parmi lesquelles on peut compter Aly [65], Khasnabis [66]. Ces deux derniers se sont focalisés sur le côté pratique du problème de la distribution des TIV. Khasnabis continue les travaux de Grecco en mettant en relation les paramètres du modèle de Schuhl avec le débit du trafic. Aly propose une combinaison pratique des différentes composantes des véhicules distinguées par la prise en compte d'un seuil de vitesse moyenne des flux du trafic. Dans la thèse d'Aly, on trouve aussi la prise en compte des facteurs influençant la distribution des TIV. Chishaki *et al.* [67] ont proposé un modèle par distinction des distributions des véhicules *leaders* et *suiveurs* du point de vue pratique.

Entre les années 1980 et 1990 on remarque l'approche de Heidemann [68, 69, 70]. En critiquant l'approche probabiliste manquant de raison analytique, Heidemann a utilisé l'approche stochastique avec l'ambition d'obtenir les distributions des vitesses et des TIV seulement à partir de connaissances faibles et limitées du trafic. Ces connaissances initiales sont celles de la distribution des vitesses à concentration très faible, proche de zéro, ainsi que la probabilité de dépassement des véhicules. Pourtant, ce modèle n'a pas été développé dans les années suivantes.

Les années 1990 sont la période des applications des distributions de TIV aux intersections et aux rond-points. Pour cette raison, les TIV longs ont plus intéressé les chercheurs. Par conséquent, les modèles plus simples sont utilisés parmi lesquels le modèle M3 de Cowan est le plus étudié. Sur le thème, des recherches continuent ainsi jusqu'après l'an 2000. Les chercheurs importants sont Sullivan [38] (1994), [71] (1997), Troutbeck [39] (1997), Akcelik et Chung [72] (1994), Luttinen [73] (1999), Hagring [74] (1996), [75] (2000), [76, 77] (2002). Cette période est particulièrement marquée par la thèse de Luttinen en 1996 [11]. Cet important travail s'est focalisé sur une démarche commune proposée pour les études

statistiques des TIV. Luttinen récapitule les modèles utilisés auparavant et les méthodes d'estimation correspondantes. Il propose également la technique de glissement des échantillons et l'utilisation de la notion de *probabilité de puissance* afin de pouvoir combiner et comparer les données d'origines diverses. Luttinen défend enfin le modèle semi-Poisson de Buckley. Quelques années plus tard, Hoogerdoorn [78, 79, 80] a proposé une nouvelle méthode d'estimation réservée au modèle GQM et a montré la puissance du modèle GQM.

Parallèlement, à partir de 1992 jusqu'au début des années 2000, il y a des approches physiques sur la distribution des TIV. La première est celle du modèle à automates cellulaires. Ce modèle est appliqué et calculé pour les cas de DIV et de TIV par Chowdhury *et al.*, Ghosh *et al.* [81] (1997), [82] (1997), [83] (1998). La deuxième approche se fait par l'analogie des phénomènes physiques comme les gaz de Dyson proposé par Krbalek [84, 85, 86, 87]. Pourtant, cette dernière approche revient aux modèles simples - le modèle Gamma à un seul paramètre en trafic intermédiaire. Krbalek aborde aussi la méthode de matrice aléatoire de covariance. Cette approche est à l'origine celle de Dunne *et al.* [22] et a récemment été utilisée sous forme d'une chaîne de Markov [23].

Très récemment, les effets des nouvelles infrastructures routières et/ou des nouveaux équipements sont pris en compte dans les études des TIV portant sur les voies HOV (réservées aux véhicules à haute occupation) (cf. Zhang [37] (2007)), ou sur les véhicules *traceurs* (cf. Pueboobpaphan [88] (2008)). Certaines études portent sur le côté purement pratique et local dans un but de simulation comme l'approche de Helmut [89] par utilisation de formules hyperboliques pour la fonction de répartition des TIV.

Les idées retenues

De nombreux chercheurs ont travaillé sur la question de la distribution de TIV. Chacun propose et/ou conforte les modèles qu'il considère comme les plus adéquats en fonction des données disponibles. En résumé, deux types de modèles ont été utilisés en vue de modéliser la distribution des TIV : le *modèle simple* et le *modèle combiné*.

Le *modèle simple* utilise directement des lois de probabilité capables de décrire la répartition des TIV. Une hypothèse implicite propose que les véhicules se déplacent de la même façon dans tous les cas. Le modèle simple ne distingue donc pas les catégories de conducteurs. Les lois les plus utilisées sont la loi exponentielle, la loi d'Erlang, la loi Gamma, la loi Log-normale.

Le *modèle combiné* prend en compte différents comportements des conducteurs dans des conditions de conduite différentes. Deux conditions de conduite sont toujours distinguées dans ce modèle : la condition de conduite *libre* et la condition de conduite *contrainte* (ou *gênée*). Les autres facteurs sont en fait reflétés via des paramètres du modèle.

Les modèles combinés, d'une manière générale, sont plus explicatifs que les modèles simples. Les notions de condition *libre* et *contrainte* jouent un rôle essentiel et entraînent les notions de *véhicule libre* et de *véhicule contraint*. Dans la première condition, le véhicule est dit *libre* car il est capable de se déplacer à sa vitesse désirée. Cette condition prévaut notamment par débit faible avec peu de véhicules sur la route. Le véhicule dit *contraint*, dans la condition de conduite *contrainte*, est en revanche influencé ou gêné par d'autres véhicules en particulier par des véhicules situés devant lui sur la même voie.

Si le *peloton* est défini par une série de véhicules contraints, on distingue également deux catégories de véhicules selon leur place dans le peloton : le *suiveur* et le *leader*. Le *leader* est en tête du peloton, *libre* et suivi par des *suiveurs* qui sont des véhicules *contraints*.

Le modèle combiné est une combinaison des modèles simples faisant intervenir un paramètre de participation du type de véhicules dans le trafic.

Grosso modo, les modèles combinés sont plus compliqués que les modèles simples et

plus complexes au niveau du nombre de paramètres utilisés. Les méthodes d'estimation des paramètres jouent un rôle important. Gerlough 1975, [90] a mis l'accent sur le fait que la méthode d'estimation utilisée pour un modèle doit être fiable et facile.

Problématiques et objectifs de la thèse

Le cadre de la recherche et les données

Les études de statistique descriptive des TIV des chercheurs du laboratoire GRETTIA ont été effectuées depuis les années 2002, à l'occasion de la nouvelle réglementation du TIV net. Ce sujet se trouve dans la thématique *Exploitation & Sécurité* qui se développe au GRETTIA. Les études sont concentrées surtout sur le TIV net, correspondant à la durée de passage de l'arrière du véhicule précédent à l'avant du véhicule considéré. Trois indicateurs de sécurité sont définis, des seuils sur ces indicateurs délimitant des comportements à risque [91]. Des études sur les différences de comportement selon les conditions météorologiques et dans différents contextes ont aussi été effectuées. Ces études montrent la nécessité d'une connaissance plus fine des situations exogènes et la nécessité de tenir en compte de la diversité des réseaux. C'est la raison pour laquelle une étude approfondie du TIV a été entreprise.

Problématique

1. Face aux nombreux modèles proposés, un classement des modèles est nécessaire pour non seulement suivre l'évolution des propositions mais aussi pour fournir des idées lors des développements et des extensions des modèles probabilistes.
2. La distribution des TIV est un indicateur du *comportement de l'ensemble* des conducteurs qui varie en fonction de différents contextes. Le lien entre la distribution des TIV et les contextes a jusqu'à maintenant été peu étudié, d'une part à cause de la complexité du recueil de données nécessaires dans les conditions intéressées, d'autre part à cause du choix entre une large gamme des modèles probabilistes.
3. Les contextes exogènes ont un impact sur les TIV indépendamment du trafic, mais le trafic lui-même influence l'interaction et les comportements des conducteurs. Par conséquent, le niveau et l'état du trafic sont aussi des contextes à considérer. Dans la littérature, les travaux se sont concentrés principalement sur l'influence du débit sur les distributions de TIV.
4. Un des points faibles limitant les applications de la distribution de TIV est le manque de méthode ou d'approche de validation. Il semble « *fragiles* » d'utiliser sans validation des régressions ou bien des abaques et cette phase importante de validation est souvent oubliée dans les études du TIV.
5. Luttinen a montré aussi l'intérêt de l'étude du modèle de Heidemann; Athol [1] a étudié les propriétés statistiques des TIV en fonction de la position en peloton; la méthode d'échantillonnage de Breiman est intéressante, etc. Toutes ces idées initiales sont à étoffer et à approfondir dans la recherche sur les TIV.

Objectifs

Le champ de recherche de la thèse est l'étude statistique et probabiliste du TIV (normal) avec application sur l'autoroute A6 à 3 voies et la route nationale RN118 à 2 voies. Les objectifs sont :

1. Utiliser différentes méthodes d'échantillonnage des TIV pour sélectionner les échantillons *représentatifs*. En suite, élargir les études statistiques des TIV en approfondissant les propriétés du seconde ordre pour les échantillons des TIV consécutifs.
2. Analyser le modèle de Heidemann, classer et comparer les modèles probabilistes, ainsi que les méthodes d'estimation existantes pour les TIV.
3. Développer, proposer et comparer de nouveaux modèles probabilistes pour les TIV.
4. Une question centrale de la thèse est de connaître les distributions des TIV à différentes échelles du trafic macroscopique, mésoscopique et microscopique. Les facteurs d'influence doivent être analysés de façon hiérarchique en tenant en compte des effets des conditions exogènes comme la météo et le type de voie.
5. Proposer différentes approches pour la phase de validation des modèles de TIV.

Les objectifs sont réalisés en établissant de manière exhaustive les distributions des TIV en fonction des catégories d'analyse disponibles. La modélisation est établie en fonction du débit, du taux d'occupation, de la vitesse moyenne, de la position des véhicules en peloton et de la configuration du trafic en distinguant le type de voie, la condition météorologique, et la période de la journée. Le débit et le taux d'occupation sont agrégés sur 6 minutes. Les vitesses sont regroupées en vitesse moyenne harmonique à chaque pas de 6 minutes.

Au niveau mésoscopique, une étude générale sur le comportement selon les groupes de TIV est réalisée afin de donner une définition « *brute* » pour le peloton. Elle est nécessaire car aucune définition précise du peloton dans la communauté scientifique ne fait unanimité.

Au niveau microscopique, la distribution des TIV est étudiée soit par groupement des valeurs de vitesses instantanées, soit en fonction du TIV précédent, soit par combinaison du TIV et de la vitesse instantanée pour former l'indicateur RHIS.

Les points apportés par la thèse intéressent principalement :

- Les gestionnaires de la route qui veulent plus d'information sur le trafic à partir d'indicateurs simples en vue de mieux réguler le trafic et de mieux gérer la voirie. Rappelons aussi que la réglementation sur le TIV net minimum de 2 secondes sera éclairée par une meilleure connaissance des TIV pratiqués.
- Le conducteur est sensibilisé par les indicateurs de sécurité utilisant des variables microscopiques parmi lesquelles le TIV joue un rôle essentiel.
- D'autres systèmes demandent une connaissance plus fine des TIV, par exemple les modèles de simulation microscopique du trafic (le TIV est une entrée et une sortie du modèle) ou certaines aides à la conduite comme le régulateur d'allure.

Plan

Le document de la thèse est composé de six principaux chapitres :

Dans le premier chapitre, les méthodes d'échantillonnage sont présentées et appliquées aux deux bases de données. Ensuite, les analyses statistiques et les comparaisons des caractéristiques des TIV sont faites.

Le Chapitre 2 traite de la modélisation du TIV. Un état de l'art exhaustif est présenté. Le modèle de Heidemann et le modèle à automates cellulaires sont examinés en privilégiant l'approche probabiliste. Les modèles et les méthodes d'estimation sont ensuite comparés à l'aide d'un *ensemble test* représentatif de différentes conditions du trafic. La fin du chapitre concerne les développements des nouveaux modèles pour le TIV parmi lesquels le modèle Double Gamma est sélectionné pour être étudié dans les chapitres en raison de sa simplicité.

Le Chapitre 3 présente les caractéristiques des modèles choisis, à savoir le modèle gamma-GQM et le modèle Double Gamma, en fonction de leurs paramètres. L'étude de la sensibilité de ces paramètres est réalisée par la méthode du Bootstrap.

Le Chapitre 4 commence par l'étude des effets des variables macroscopiques et des conditions exogènes sur la distribution des TIV. L'étude mène aussi à regarder l'effet du diagramme fondamental empirique sur la distribution des TIV et la sensibilité des paramètres des modèles choisis selon le débit et le taux d'occupation.

Le Chapitre 5 traite de la distribution des TIV à l'échelle mésoscopique et microscopique du trafic. Les échantillons sont classés selon les modalités correspondant à ces échelles.

Le dernier chapitre aborde la validation à partir de différentes approches. Les méthodes principales sont d'une part le glissement des échantillons pour comparer les résultats au sein des échantillons du même site, et d'autre part l'utilisation des échantillons des sites différents.

La thèse comporte aussi des annexes.

Cette thèse a donné lieu à des communications dans les conférences suivantes :

[92] D. H. Ha, M. Aron, and S. Cohen, *Variation of the parameters of the Time Headway Models according to Macroscopic Variables, Fundamental Diagram and Exogenous Effects*, accepté pour une publication sur le Transportation Reserch Record - revue du 90th Transportation Reserch Board, 2011.

[93] D. H. Ha, M. Aron, and S. Cohen, *Time Headway Variable and Probabilistic Modeling*, Traffic Flow Theory Committee, Juillet 2010, Annecy, France, en révision à Transportation Research Part C, Octobre 2010.

[94] D. H. Ha, M. Aron, and S. Cohen, *Comparison of Time Headway Distribution in different traffic contextes*, World Conference of Transportation Research, Juillet 2010, Lisbon, Portugal, Paper Number 1261.

[95] D. H. Ha, M. Aron, and S. Cohen, *Exogenous Effects on the Microscopic Distributions*, Transportation Reserach Arena, YEAR 2010, Bruxelles, Belgique, Poster.

[96] D. H. Ha, M. Aron, and S. Cohen, *Modélisation du Temps Intervéhiculaire sur l'autoroute en France*, Séminaire SPI-STIC 2009, Acte N°134/INRETS.

Chapitre 1

Bases de données et échantillonnage

Circuler fait partie des activités principales de l'homme. Cette activité évolue en conséquence au fil du temps à long terme, et comprend en même temps des phénomènes périodiques à court terme. Le trafic routier a beaucoup varié selon le temps : à long ou moyen et même à court terme, il y a les changements de technologie dont ceux du moyen de transport routier et de l'infrastructure, de l'utilisation des systèmes intelligents et d'information avancés, etc. ; à chaque instant, il y a la condition de météorologie, l'ambiance environnementale (parfois appelée la lisibilité de la route), l'information survenue, etc. Le trafic routier varie aussi selon l'espace, par exemple la zone d'aménagement (urbain, péri-urbain, commerciale, rural, etc.), le type du réseau routier (autoroute, route nationale, route départementale, etc.), la position du tronçon (en section droite et plate, à proximité d'une intersection, en virage, etc.) et la localisation latérale de la voie (sur la voie rapide, voie lente, etc.)... Toutes ces influences et plusieurs autres sont prises de façon consciente ou inconsciente par le conducteur. Elles participent de toute manière à la décision finale au volant qui représente un *comportement individuel* dans le trafic. Ce comportement une fois enregistré est ensuite dans une *base de données* du trafic. Enfin, l'analyse statistique va traiter la base de données pour fournir des connaissances sur le *comportement de l'ensemble* (ou bien comportement collectif).

Comme plusieurs facteurs sont mis en jeu dans la réalité du trafic, la constitution de la base de données précitée devrait dans une certaine mesure faciliter la connaissance des facteurs principaux en précisant *où, quand et dans quelles conditions* les informations ont été obtenues. Le but est d'approcher et d'interpréter le mieux possible le comportement collectif des conducteurs. Par conséquent, l'identification des échantillons dans la base de données constitue une étape préliminaire et indispensable pour le calibrage, la modélisation probabiliste et la validation.

Ce chapitre présente les bases de données utilisées et les méthodes d'échantillonnage appliquées sur ces bases. Les résultats statistiques sont ensuite étudiés sur les différentes sources d'influence du trafic.

1.1 Pré-échantillonnage

La première étape consiste à enregistrer de manière automatique les données individuelles au moyen des stations de mesure. La durée de l'enregistrement peut varier de quelques heures à quelques jours pendant lesquelles les informations de passage véhicule par véhicule sont transmises à la station. Trois types de données de base dans l'étude de TIV sont traités : l'instant de passage (normalement en centièmes de seconde), la vitesse instantanée (en m/s) et la longueur (en dm) des véhicules.

Ensuite, les données sont sauvegardées dans des fichiers .TXT qui représentent un *échantillon primitif*. Dans les échantillons primitifs, chaque ligne d'information comporte des caractéristiques individuelles d'un véhicule. Cependant, ces données sont encore à l'état brut et comprennent éventuellement certaines erreurs. Sur les données dont on dispose, les deux principaux types d'erreur suivants sont préalablement traités :

- *TIV zéro* : Des boucles magnétiques reconnaissent deux arrivées des véhicules en même temps sur la même voie. Ce phénomène *multiple* est évidemment impossible car la longueur des véhicules est positive. Il semble que sur certaines stations de mesure, le même événement soit enregistré plusieurs fois.
- *Arrivées irréalistes* : Ce sont des arrivées de véhicules avec un TIV très court (disons moins de 0,05 s) d'autant plus que ces arrivées sont accompagnées par une vitesse relative (par rapport au véhicule en tête) positive, même élevée. Il peut s'agir de remorques de camion identifiées à tort comme véhicules ou bien une faute de mesure de vitesse.

Pour faciliter le traitement des données tout en respectant l'information privilégiée de la variable TIV, les enregistrements du premier type sont éliminés tandis que dans le cas des erreurs du second type, la vitesse et la longueur du véhicule considéré sont ajustées à celles correspondantes du véhicule précédent. Cet ajustement du deuxième type d'erreur n'entache pas en revanche la distribution des TIV très courts obtenue car peu d'erreurs dans la base de données sont survenues au final. D'autant que l'assimilation d'une remorque par un véhicule particulier, si tel était le cas, peut être considérée comme une définition élargie du TIV ce qui est aussi intéressant à interpréter.

1.2 Échantillonnage

Les données brutes comportent éventuellement des variations non-aléatoires, c. à. d que les caractéristiques des données peuvent comprendre des *tendances statistiques*. Il faut soit tenir compte spécifiquement de ces tendances dans les modèles, soit se restreindre aux *échantillons stationnaires* possédant des propriétés stables et n'évoluant pas en fonction du temps. Deux méthodes sont utilisées pour éliminer les *échantillons non-stationnaires* de TIV : la première consiste à regrouper les données selon quelques critères prédéterminés tandis que la deuxième consiste à échantillonner en se basant sur la stabilité du nombre de véhicules arrivés dans une période.

La première méthode appelée méthode *de groupement* a besoin généralement d'une base de données assez importante afin d'éviter la fragmentation et la stratification - c. à. d un trop faible nombre de données dans chaque « case » issue du croisement des modalités des variables d'influence. La deuxième méthode étant la méthode *de raffinement* assure la stabilité du débit. Pour les deux méthodes d'échantillonnage, l'idée commune est d'utiliser des *critères statistiques* pour obtenir une stabilité au niveau des données réelles.

Une troisième méthode utilise un processus d'*analyse de tendance* en vue de choisir les échantillons des TIV [11]. Dans notre étude, les tests de tendance de cette méthode sont employés en complément après que la méthode de raffinement a été utilisée. Les deux méthodes précitées sont décrites en détail.

1.2.1 Méthode de groupement

Notons \mathcal{A} un critère statistique à prendre en considération, \mathcal{I}_i un individu i contient d'une information D_k de l'ensemble d'informations \mathcal{D} en général. Le principe de la *méthode de groupement* consiste à classer deux individus \mathcal{I}_i et \mathcal{I}_j , où $i \neq j$, comportant le même

critère \mathcal{A} dans un échantillon \mathcal{E} , c.à.d :

$$\mathcal{I}_i \in \mathcal{E} \text{ et } \mathcal{I}_j \in \mathcal{E} \quad \text{si} \quad D_{k_i} \in \mathcal{A} \text{ et } D_{k_j} \in \mathcal{A} \quad (1.1)$$

Le critère \mathcal{A} peut être simple ou complexe et doit être relié aux informations D dont on dispose. Il est *simple* lorsqu'il ne met en jeu qu'une information parmi d'autres dans l'ensemble d'information \mathcal{D} , et considéré comme *complexe* dans le cas contraire où plus d'une information est mise en jeu. Particulièrement, dans le cas où $\mathcal{A} = \Omega$, l'ensemble des données enregistrées de la base constitue un échantillon unique.

En réalité, le trafic est soumis à des influences de nature très différentes, et l'on doit considérer \mathcal{A} à différents niveaux. Cela veut dire que l'information \mathcal{D} est à différentes échelles, non nécessairement microscopique mais aussi macroscopique ou mésoscopique. Autrement dit, les informations D_k attachées à un véhicule peuvent être dérivées soit sur le véhicule lui-même, soit agrégées à partir d'un groupe de véhicules dans lequel le véhicule considéré participe. Il est à noter que le critère \mathcal{A} ne fait pas uniquement référence à des variables quantitatives, au contraire, celles-ci peuvent être qualitatives, le véhicule appartient à une certaine catégorie de gabarit par exemple.

Dans le cas où l'information D_k est agrégée, deux variantes peuvent être appliquées. La première méthode consiste à grouper les véhicules sur une période temporelle fixe Δ . On a l'information D_k^Δ . La deuxième prend un nombre constant N de véhicules consécutifs. On a l'information D_k^N .

Pour des raisons pratiques en traitant des données temporelles, l'information D_k^Δ est la plus fréquente dans la réalité et est employée dans cette étude. Plus souvent, des courts périodes de temps (au plus de 10 minutes) sont pratiquées (*ang short-time period*). On a donc :

$$D_k^\Delta = f(I_j) \quad \text{avec} \quad (j-1)\Delta \leq I_j < j\Delta \quad (1.2)$$

où j est l'indice de la période considérée, Δ est la période temporelle fixée. La fonction f indique que l'information D_k^Δ dépend de la période considérée.

Dans le cas où un très faible nombre d'individus est observé dans un échantillon \mathcal{E} , le groupement avec des échantillons voisins ayant une moyenne de TIV similaire est nécessaire pour assurer une taille suffisante. En conséquence, le choix d'une période convenable Δ ne doit pas être trop court notamment en trafic fluide. Néanmoins, la pratique du groupement des échantillons augmente généralement la variance de l'échantillon final [97]. Pour les données de cette étude, la période Δ sur l'autoroute A6 en France est une **heure** ce qui assure la stabilité statistique des échantillons, et sur la route nationale RN118, elle est de **6 minutes**. D'autres systèmes utilisent différentes *périodes d'agrégation* Δ telles que **5 minutes** aux États-Unis, **10 voire 15 minutes** dans certaines expérimentations ailleurs. Le choix de Δ dépend donc aussi du système de mesure du trafic.

Dans une *période temporelle* I , à partir d'une série des instants d'arrivée T_i des véhicules i , la variable microscopique TIV, notée H , et la quantité macroscopique *débit*, noté Q_I peuvent être calculées directement. En connaissant de plus les vitesses des véhicules, on peut obtenir la vitesse moyenne harmonique \mathcal{V}_I , et donc la concentration du trafic \mathcal{K}_I par l'intermédiaire de la relation fondamentale $Q = \mathcal{V} \times \mathcal{K}$. Ensuite, si les informations des longueurs de véhicule sont disponibles, les variables microscopiques comme la *distance inter-véhiculaire* (DIV), le TIV *net* G et la variable macroscopique *taux d'occupation* \mathcal{T}_I peuvent être calculées.

En conséquence, la fonction f dans l'équation 1.2 peut être déterminée par les quantités comme le débit Q_I , le taux d'occupation \mathcal{T}_I , la vitesse moyenne \mathcal{V}_I ou la concentration \mathcal{K}_I .

Soit T_i, l_i, v_i avec $i = 0, 1, \dots, n$ l'instant d'arrivée de l'avant, la longueur et la vitesse instantanée du véhicule i . L'indice $i = 0$ indique le premier véhicule de la période I_j

considérée où j est l'indice de la période. Les variables microscopiques attachées à chaque véhicule individuel sont calculées comme suit ¹ :

$$\begin{aligned} \text{TIV } H : \quad & h_i = T_{i-1} - T_i \quad \text{avec } i = 1, 2, \dots \\ \text{TIV net } G : \quad & g_i = h_i - \frac{l_{i-1}}{v_{i-1}} \\ \text{Vitesse relative } V_r : \quad & v_i^r = v_i - v_{i-1} \\ \text{DIV } S : \quad & s_i = v_{i-1}h_i - l_{i-1} \end{aligned}$$

Les quantités macroscopiques sont calculées dans la période I_j comme suit :

$$\begin{aligned} Q_I : \quad q_j &= \frac{N_j}{\Delta} \quad \text{avec } N_j = \sum_k \mathbb{1}_{\{T_k \in I_j\}} \\ \mathcal{T}_I : \quad \tau_j &= \frac{1}{\Delta} \sum_k \frac{l_k + L}{v_k} \times \mathbb{1}_{\{T_k \in I_j\}} \\ \mathcal{V}_I : \quad 1/\bar{v}_j &= \frac{1}{N_j} \sum_k \frac{1}{v_k} \times \mathbb{1}_{\{T_k \in I_j\}} \end{aligned}$$

où L est la longueur des boucles magnétiques. L'indice $(i - 1)$ dans les calculs de TIV *net* G et de DIV S signifie que les valeurs prises en compte sont considérées au moment du passage du véhicule i sur le point de mesure. Il est à noter dans certaines études en France que les calculs de G et S ont été effectués en utilisant l'indice i de la longueur et de la vitesse. Cette considération mesure donc les informations « passées » par rapport à l'instant T_i . Dans les deux cas, la supposition d'une vitesse constante entre T_{i-1} et T_i , du véhicule i ou de son précédent $(i - 1)$ est demandée ce qui fait croître l'erreur des estimations du TIV *net* et de la DIV. Dans cette étude, ces deux variables ne sont utilisées que dans la définition mésoscopique du *peloton*, ce dernier ayant un sens encore vague qui accepte une certaine tolérance. Par ailleurs, la forme de la distribution des TIV *net* est très proche de celle des TIV. Au final, les deux considérations ci-dessus sont secondaires.

La concentration du trafic \mathcal{K}_I est par définition mesurée en espace. Si les vitesses du premier véhicule dans la période considérée sont constantes, la concentration \mathcal{K}_I est donnée par :

$$\mathcal{K}_I : \quad k_j = \frac{N_j}{v_{(1)}^{I_j} (T_{(1)}^{I_j} - T_{(f)}^{I_j})} \quad (1.3)$$

où $v_{(1)}^{I_j}$ et $T_{(1)}^{I_j}$ sont la vitesse et l'instant d'arrivée du premier véhicule entrant dans la période I_j , $T_{(f)}^{I_j}$ est l'instant d'arrivée du dernier véhicule sortant de la période I_j .

Pour que les calculs des variables agrégées soient totalement corrects, une hypothèse importante implicitement adoptée est qu'aucun dépassement du premier véhicule ni aucun changement de voie ne se produit dans la période I_j . De toute façon, le système des boucles magnétiques utilisé n'est pas en mesure de détecter ces manoeuvres qui concernent la dynamique entre les deux voies. Le taux d'occupation \mathcal{T}_I peut représenter la concentration \mathcal{K} du trafic [98, 99]. Enfin, le système de mesure en France utilise $\Delta = 6$ minutes et $L = 1$ mètre.

Une autre approche du critère \mathcal{A} pourrait être proposée en se basant sur des informations D_k à l'échelle *mésoscopique*. Des informations dites « mésoscopique » sont celles liées au *peloton* - un groupe particulier des véhicules dans lequel le véhicule considéré est inséré. À ce niveau du trafic, ce sont les véhicules au voisinage du véhicule considéré qui déterminent l'information D_k . Les trois principales informations à considérer sont la configuration de poursuite selon le type de véhicule \mathcal{C} , le rang du véhicule \mathcal{P} dans le peloton et la

1. lorsque les capteurs enregistrent l'instant de passage de l'avant des véhicules

valeur du TIV du véhicule précédent T_p^k . Les deux premières informations sont qualitatives et l'on distingue les cas suivants :

1. \mathcal{C} : VP-VP, VP-PL, PL-VP et PL-PL
2. \mathcal{P} : en-tête ou libre, 1p - première position en peloton, 2p - deuxième position en peloton, 3p - troisième position en peloton, et queue signifie tous les positions de rang supérieur à trois.
3. T_p^k : inférieur à 1 s, (1-2) s, (2-3) s, ... etc.

La définition du peloton a connu de longues discussions qui ne sont pas encore terminées. Les définitions tendent généralement à identifier un seuil de TIV ou une autre variable comme la vitesse relative. Les seuils obtenus dans la littérature varient en fonction évidemment de la définition mais une fois celle ci faite, dépendent aussi des données traitées. En conséquence, une étude supplémentaire a été faite pour considérer les informations de \mathcal{C} et de \mathcal{P} en tant qu'information D_k . Dans le cas où les définitions sont clairement énoncées, l'information mésoscopique D_k peut être obtenue à partir des données microscopiques initiales T_i, l_i, v_i .

Aussi, d'autres informations peuvent être reconnues comme la condition de météorologie \mathcal{M} (par temps de pluie, par temps sec), le type du réseau (autoroute, route nationale, route départementale) ou le type de voie \mathcal{L} (voie rapide (VR), voie lente (VL)), etc. Ces informations dans le cas plus simple sont enregistrées par une variable logique (0 ou 1) ce qui permet de détecter l'apparition de la circonstance concernée.

En résumant les points de vue ci-dessus, une approche unifiée est que chaque véhicule individuel i est dorénavant caractérisé non seulement par ses informations microscopiques directement tirées de l'enregistrement, mais aussi des informations mésoscopiques et macroscopiques. Finalement, on a l'approche suivante :

$$\mathcal{D} := \{H, G, S, V_r, Q_I, \mathcal{T}_I, \mathcal{V}_I, \mathcal{K}_I, \mathcal{C}, \mathcal{P}, \mathcal{C}, \mathcal{P}, \dots\} \quad (1.4)$$

$$D_k \in \mathcal{D} \text{ avec } k = 1, 2, \dots \quad (1.5)$$

$$\mathcal{A} := f^*(D_k, k = 1, 2, \dots) \quad (1.6)$$

où la fonction f^* est une fonction représentant la relation logique entre les informations à prendre.

Cette méthode de groupement est facile à réaliser. Elle n'élimine aucune donnée des échantillons primitifs. Une base de donnée suffisamment large est nécessaire pour éviter des échantillons vides ou comprenant trop peu d'individus sur lesquels le caractère aléatoire du TIV ne pourrait pas être analysé. En effet, dans le cas de \mathcal{A} complexe, Greenberg (1966) [55, 11] a montré que la stratification des données de TIV pouvait engendrer de graves troncatures. Une autre lacune de cette méthode d'échantillonnage réside dans le fait que les échantillons finals ne sont pas généralement ceux comportant des série d'informations consécutives. Cette dernière faiblesse entraîne donc l'utilisation en parallèle d'une deuxième méthode d'échantillonnage.

1.2.2 Méthode de raffinement

Dans la méthode de groupement, on s'aperçoit que tous les individus se trouvant dans la même période I_j appartiendront à un même échantillon \mathcal{E} car ils ont la même information agrégée. Si la période I_j est courte comme dans la majorité de cas, les échantillons \mathcal{E} sont généralement formés des courtes périodes non consécutives ; par conséquent, les TIV groupés ne sont pas consécutifs.

Pour obtenir des échantillons de TIV consécutifs, Cox et Lewis [100] ont proposé une méthode d'échantillonnage basée sur la stationnarité du nombre cumulé des véhicules arrivés. Autrement dit, la méthode suppose la stabilité du débit de l'échantillon constitué, dans cette étude on appelle cette méthode *méthode de raffinement*. Comme une série de TIV consécutifs constitue chaque *échantillon raffiné*, la possibilité d'étudier les caractéristiques du second ordre d'une telle série chronologique pourrait fournir des informations complémentaires sur le comportement collectif.

La méthode de raffinement proposée s'appuie sur une statistique U calculée à partir du nombre des véhicules arrivés en un point donné sur l'intervalle de temps $I_j = \varepsilon$. La procédure de calcul basée sur l'étude de Breiman [101] est réalisée dans cette étude selon les étapes suivantes :

1. Élargir la période ε par un intervalle δ (une **minute** dans le cas des TIV). Il y a deux possibilités :
 - (a) À partir de l'instant de départ $t = 0$ ou de l'instant final de la période de l'échantillon précédemment obtenu, prolonger l'intervalle δ donne la première période de l'échantillon $\varepsilon \mid 0 < \varepsilon \leq \delta$,
 - (b) En m -ième étape, la période obtenue est $\varepsilon \mid 0 < \varepsilon \leq t$, avec $t = m \delta$,
2. Calculer le nombre de véhicules $N(\varepsilon)$ dans cette période ε sous l'hypothèse que le débit calculé est constant,
3. Compter le nombre réel des véhicules arrivés $N_\varepsilon(t)$ dans cette période,
4. Calculer l'intégrale des différences entre ces deux quantités pour obtenir la statistique $Z(\varepsilon)$ en normalisant par l'écart-type σ des nombres de véhicules dans chaque intervalle ajouté δ :

$$Z(\varepsilon) = \frac{1}{\sigma} \left[\int_0^\varepsilon N(t) dt - \frac{\varepsilon N(\varepsilon)}{2} \right] \quad (1.7)$$

$$= \frac{1}{\hat{\sigma}} \left[\sum_0^m N(m \delta) - \frac{\varepsilon N(\varepsilon)}{2} \right] \quad (1.8)$$

Miller [102, 65] a proposé d'estimer l'écart-type $\hat{\sigma}$ dans l'équation 1.8 selon la formule suivante :

$$\hat{\sigma}^2 = \frac{1}{2(n-1)} \sum_{m=1}^{m-1} (N_{k+1} - N_k)^2 \quad (1.9)$$

Où N_k est le nombre des véhicules dans le k -ième intervalle d'une minute δ . Si l'on pose $U(\varepsilon) = Z(\varepsilon) \sqrt{\frac{\varepsilon^3}{12}}$, cette statistique a été prouvée asymptotiquement normale, c. à d $U \sim \mathcal{N}(0, 1)$ lorsque $n \rightarrow \infty$,

5. Comparer $U(\varepsilon)$ à un paramètre de contrôle α , deux cas suivants sont possibles :
 - (a) Si $|U(\varepsilon)| < \alpha$, retourner à l'**Étape.1**. Breiman *et al* [101] ont proposé de choisir $\alpha = 3, 5$,
 - (b) Dans le cas où $|U(\varepsilon)| \geq \alpha$, raccourcir la période ε en question par un intervalle δ , et retourner à l'**Étape.2**. Ensuite, comparer $U(\varepsilon)$ à un deuxième paramètre de contrôle β . Breiman *et al* [101] proposent de choisir $\beta = 1$,
6. Le processus d'échantillonnage se terminera lorsqu'une valeur ε_0 satisfait l'inégalité $|U(\varepsilon_0)| < \beta$.

Dans la réalisation de la méthode de raffinement, on compare simultanément $|U(\varepsilon)|$ à α et β en l'**Étape.4**. Ensuite, toutes les valeurs ε_i telle que $|U(\varepsilon_i)| < \beta$ sont sauvegardées. Enfin, la valeur maximale parmi les valeurs ε_i obtenues tout au long du processus du début jusqu'au dépassement de $|U(\varepsilon)|$ du seuil α sera considérée comme la période finale de

l'échantillon choisi. On a donc :

$$\varepsilon_0 = \max_i \varepsilon_i \quad (1.10)$$

La Figure 1.1-a illustre le principe de la méthode de raffinement tandis que la Figure 1.1-b représente un résultat sur la base de données RN118 où la courbe est un échantillon primitif. Dans cette figure, il est observé que la méthode a bien distingué différentes tendances dans la courbe du nombre cumulé des véhicules. Il est à noter qu'une ligne droite sur la courbe signifie un débit assez stable de manière générale.

Par la méthode de raffinement appliquée aux données du Site N°1, 242 échantillons sur la voie lente (VL) et 148 échantillons sur la voie rapide (VR) sont obtenus. La figure à droite de la Figure 1.1 représente un échantillon primitif (un fichier d'enregistrement). Le point rond et le carré signaient le début et la fin d'une période d'un échantillon.

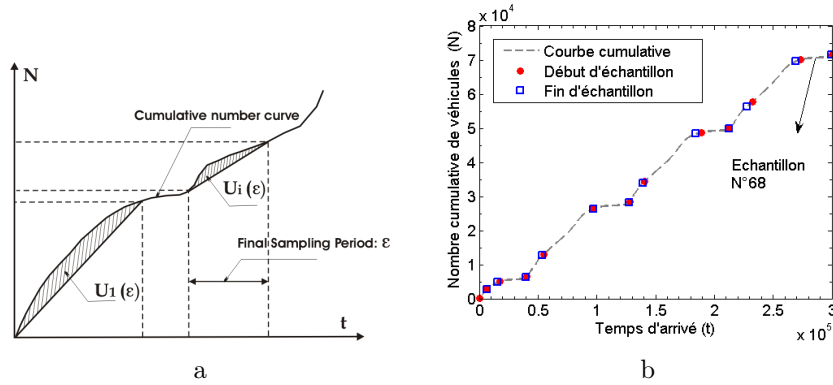


FIGURE 1.1 : Méthode de raffinement des échantillons de TIV

En étudiant les différents profils des courbes cumulées de véhicules, il semble que les paramètres α, β proposés par Breiman *et al.* ne soient pas sévères, car des échantillons de longue période sont aussi acceptés d'autant plus que certains ont un histogramme de TIV atypique. En conséquence, une analyse spectrale présentée plus tard pourrait être utilisée pour détecter ce type d'échantillon. Pourtant, dans la phase d'échantillonnage, des tests de tendance sont préférables pour exclure des échantillons inappropriés. Les trois tests suivants ont été utilisés : le test de signe pondéré (.ang *Weighted sign test*), le test de corrélation des rangs de Kendall (.ang *Kendall's rank correlation test*), et le test de score exponentiellement ordonné (.ang *Exponential ordered scores test* [103, 104, 100]).

1. Les deux premières statistiques utilisent une variable h_{ij} résultant de la comparaison entre deux TIV différents h_i et h_j :

$$h_{ij} = \begin{cases} 1 & \text{si } h_i > h_j \\ 0 & \text{si } h_i < h_j \end{cases} \quad (1.11)$$

Deux statistiques correspondantes sont :

(a) Statistique S :

$$S = \sum_{k=1}^{k=n/2} (n - 2k + 1) h_{k, n-k+1} \quad (1.12)$$

$$\mathbb{E}[S] = n^2 / 8 \quad (1.13)$$

$$\text{Var}[S] = \frac{n(n^2 - 1)}{24} \quad (1.14)$$

(b) Statistique τ :

$$\tau = 1 - \frac{4}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n h_{ij} \quad (1.15)$$

$$\mathbb{E}[\tau] = 0 \quad (1.16)$$

$$\text{Var}[\tau] = \frac{2(2n+5)}{9n(n-1)} \quad (1.17)$$

Sous l'hypothèse nulle qu'il n'y a pas de tendance statistique, les statistiques S proposées par Cox & Stuart et τ proposée par Kendall tendent vers une loi normale \mathcal{N} de moyenne et d'écart-type calculés dans les équations 1.13, 1.14, 1.16 et 1.17.

2. La troisième statistique proposée par Cox & Lewis [100] semble être la plus sévère dans le sens du taux de rejet d'un échantillon [11]. Soit une série des TIV (h_1, h_2, \dots, h_n) , notons r le numéro d'ordre d'un TIV h_k est sa position dans la série ordonnée $(h_{(1)}, h_{(2)}, \dots, h_{(n)})$. Le numéro r sert à calculer une pondération (ou un score) du TIV h_i :

$$s_{r,n} = \sum_{i=1}^r \frac{1}{n-i+1} \quad (1.18)$$

Ensuite, un facteur $(i - \frac{n+1}{2})$ est ajouté en vue de faire une liaison entre les TIV et les pondérations correspondantes. Enfin, la statistique V est prouvée tendre vers une loi normale $\mathcal{N}(0, \sigma_V^2)$ lorsque $n \rightarrow \infty$:

$$V = \sum_{r=1}^n s_{r,n} \left(r - \frac{n+1}{2} \right) \quad (1.19)$$

Et

$$\sigma_V^2 = \sum_{r=1}^n s_{r,n} \left(r - \frac{n+1}{2} \right)^2 K_{2,n} \quad (1.20)$$

Ici, $K_{2,n}$ est le second semi-invariant de la population constituée par les pondérations :

$$K_{2,n} = 1 - \frac{1}{n-1} \left(\frac{1}{n} + \frac{1}{n-1} + \dots + \frac{1}{2} \right) \quad (1.21)$$

Il est à noter que Luttinen [11] de façon directe a employé les mêmes principes d'élargissement et de raccourcissement des intervalles en vue d'échantillonner des TIV consécutifs. Dans le cadre de notre étude, les tests de tendance sont utilisés pour vérifier les échantillons raffinés obtenus par la méthode de raffinement de Breiman.

Pour résumer, on appelle *échantillons groupés* et *échantillons raffinés* les *échantillons fins* correspondant respectivement à la méthode de groupement et de raffinement. L'adjectif *fin* souligne qu'un traitement ou une procédure ont été appliqués sur les échantillons *primitifs*. Les échantillons fins sont établis par divers critères \mathcal{A} , soit des statistiques calculées soit les informations attachées aux différents niveaux du trafic. L'étape d'échantillonnage a pour but de constituer des échantillons *stables* et *représentatifs* de la population étudiée. Cette étape permet aussi de distinguer *a priori* les différents contextes du trafic, et constitue la phase préparatoire à la phase d'analyse suivante.

1.3 Analyses statistique et spectrale

Dans cette étape, seuls les échantillons fins sont considérés. En dépendant de la méthode d'échantillonnage et de la période temporelle choisie, les échantillons constituent un ensemble de véhicules soit consécutifs soit non consécutifs. Dans le dernier cas, seule l'analyse statistique est appliquée afin d'exploiter les informations. Le premier cas permet en complément l'utilisation de l'analyse spectrale qui est en mesure d'approcher des caractéristiques statistiques de seconde ordre de l'échantillon.

1.3.1 Analyse statistique

Supposons un échantillon (fin) contenant un ensemble de TIV (non nécessairement consécutifs) : $H = \{h_1, h_2, \dots, h_n\}$, on pose $H_{or} = \{h_{(1)}, h_{(2)}, \dots, h_{(n)}\}$ la version ordonnée croissante de H , où h_i est le TIV du véhicule i . Les TIV sont ensuite groupés dans m classes C_j avec $j = \overline{1, m}$. Appelons N_j le nombre de TIV dans la classe C_j et a_j la largeur de la classe C_j (en \mathbf{s}). Les a_j sont en général choisis identiques ce qui permet d'omettre l'indice j ; la largeur de toutes les classes est donc notée communément a . On a directement les relations suivantes :

$$N_j = \text{Card } C_j \quad (1.22)$$

$$\sum_{j=1}^m N_j = n \quad (1.23)$$

Les échantillons se caractérisent par des statistiques décrivant l'échantillon de façon globale. Les statistiques suivantes fournissent des connaissances de base que l'on les considère comme les caractéristiques de premier ordre d'un échantillon des TIV. Dans tous les cas étudiés dorénavant, notons H la variable aléatoire réelle (v.a.r) non négative, notons (H_1, H_2, \dots, H_n) un échantillon des TIV avec les valeurs empiriques correspondantes (h_1, h_2, \dots, h_n) , où les H_i sont indépendants et identiquement distribués i.i.d.

Les statistiques descriptives

1. *L'espérance et la moyenne* : L'espérance de la variable H est calculée comme suit :

$$\mathbb{E}[H] = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n h_i \quad (1.24)$$

Pour un échantillon de taille n , l'estimateur de l'espérance $\mathbb{E}[H]$ est la moyenne arithmétique des TIV :

$$\mu_H = \frac{1}{n} \sum_{i=1}^n h_i \quad (1.25)$$

On calcule aussi le *mode* (la classe de TIV le plus fréquemment rencontrée dans l'échantillon) et la *médiane* qui indique la valeur en deçà (ou au-delà) laquelle, les 50 % de TIV les plus faibles (ou les plus grands) sont constatés. Le triplet (*moyenne*, *mode*, *médiane*) est particulièrement utile pour donner une première esquisse de la distribution des TIV avant l'emploi d'autres outils statistiques plus complets comme la *boîte à moustache*, le *taux de survie*, etc.

2. *La variance et l'écart-type* : La variance de la v.a.r H fournit une connaissance de la dispersion des valeurs des TIV observées autour de la moyenne.

$$\text{Var}[H] = \sigma^2 = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n (h_i - \mu_H)^2 \quad (1.26)$$

où σ est l'écart-type de H . L'estimateur s_H^2 de la variance σ^2 d'un échantillon de taille n est obtenu en divisant par $(n - 1)$ la somme des n carrés des écarts entre les valeurs empiriques et la moyenne empirique. La division par $n - 1$ au lieu de n permet d'éviter le *biais* de l'estimation d'échantillonnage.

$$s_H^2 = \frac{1}{n-1} \sum_{i=1}^n (h_i - \mu_H)^2 \quad (1.27)$$

3. *Le coefficient de variation (CV) :*

$$CV = \frac{s_H}{\mu_H} \quad (1.28)$$

Étant le quotient entre l'écart-type s_H et la moyenne μ_H , le CV joue un rôle important dans l'analyse statistique des TIV à comparer avec le CV d'une distribution exponentielle qui vaut toujours 1. En conséquence, le CV empirique fournit une première estimation sur le caractère exponentiel ou non de la répartition des TIV.

Le CV est aussi un indicateur de l'état général du trafic, par exemple un trafic régulier sera associé à un CV faible. Le CV varie également en fonction du débit et est influencé par plusieurs facteurs comme le type de la route, etc. En effet, Luttinen [11] a constaté, sur les routes finlandaises, que les CV sont supérieurs à 1 pour les débits moyens, qu'ils sont proches de la valeur 1 pour les débits très faibles et qu'ils sont enfin inférieurs à cette valeur unitaire pour les hauts débits. En revanche, Al-Ghamdi [105] a obtenu des CV entre 0,32 à 0,82 sur toutes les routes étudiées en Arabie Saoudite tandis que d'autres auteurs ont trouvé différentes valeurs en fonction de l'infrastructure, par exemple le CV semble plus élevé sur les autoroutes.

4. *Le coefficient de symétrie, le coefficient d'aplatissement et leur relation :*

Le coefficient de symétrie S_3 mesure le niveau de symétrie d'une distribution tandis que le coefficient d'aplatissement K_4 mesure le rôle des grandes valeurs par rapport aux autres valeurs de la distribution ainsi que leur groupement autour du centre de la distribution. Ces coefficients sont estimés selon les deux formules suivantes :

$$\hat{S}_3 = \frac{1}{ns^3} \sum_{i=1}^n (h_i - \mu_H)^3 \quad (1.29)$$

$$\hat{K}_4 = \frac{1}{ns^4} \sum_{i=1}^n (h_i - \mu_H)^4 \quad (1.30)$$

Le coefficient $\hat{S}_3 = 0$ signifie une distribution symétrique. Il est normalement positif pour les distributions des TIV. Le niveau d'aplatissement reflété par le coefficient \hat{K}_4 est souvent comparé à celui de la loi normale $\mathcal{N}(0, 1)$. La relation entre le coefficient d'aplatissement et le carré du coefficient de symétrie (appelée relation K-S² avec K et S désignant les K_4 et S_3 correspondent en anglais aux *Kurtosis* et *Skewness*) est souvent représentée graphiquement.

La Figure 1.2 représente les relations K-S² de différents modèles et des échantillons des bases de données utilisées. Cette représentation est souvent utilisée pour avoir de premières idées sur les modèles probabilistes convenables pour des TIV. La figure montre préalablement que le modèle Pearson III, la loi log-normale, voire le modèle Pearson VI pourront être utilisés. Parmi les modèles de la famille Pearson, les relations K-S² empiriques obtenues se trouvent dans la majorité des cas près de la ligne du modèle de Pearson III - une généralisation de la distribution Gamma ce qui laisse prévoir un rôle important de ce modèle dans la modélisation du TIV.

Distribution empirique des TIV

Soit H la v. a. r TIV prenant des valeurs aléatoires h . Les fonctions de répartition de probabilité théorique et empirique de H sont respectivement calculées comme suit :

$$F(h) = \Pr \{H \leq h\} \tag{1.31}$$

$$F_n(h) = \frac{j}{n} \quad \text{Avec } h_{(j-1)} < h \leq h_{(j)} \quad \text{où } j = 1, 2, \dots, n \tag{1.32}$$

Les fonctions de densité de probabilité théorique et empirique de H sont respectivement :

$$f(h) = \frac{dF}{dh} \tag{1.33}$$

$$f_{n,a}(h) = \frac{N_i}{n} \times \mathbb{1}_{\{h \in C_i\}} \tag{1.34}$$

Si l'on connaît les fonctions $f(h)$ et $F(h)$, l'espérance, la variance et les autres coefficients statistiques sont calculés par :

$$\mathbb{E} [H] = \int_0^\infty h f(h) dh \tag{1.35}$$

$$\text{Var} [H] = \int_0^\infty (h - \mathbb{E} [H])^2 f(h) dh \tag{1.36}$$

$$S_3 = \frac{1}{\sigma^3} \int_0^\infty (h - \mathbb{E} [H])^3 f(h) dh \tag{1.37}$$

$$K_4 = \frac{1}{\sigma^4} \int_0^\infty (h - \mathbb{E} [H])^4 f(h) dh \tag{1.38}$$

Au niveau empirique, les valeurs de la fonction $f_{n,a}$ servent à construire un *histogramme*. Cette méthode de représentation dépend fortement du choix de l'intervalle a de groupement et entraîne évidemment des discontinuités au niveau du diagramme. Par conséquent, rendre le diagramme plus *lisse* donne naissance à la méthode du *noyau* qui prend en compte des valeurs voisines dans chaque calcul de la fréquence d'un TIV h_i . En utilisant la fonction du *noyau* $K(h)$, la méthode modifie la densité empirique dans l'équation 1.34 dans le sens représenté par la formule suivante :

$$f_{n,b}^K(h) = \frac{1}{nb} \sum_{i=1}^n K \left(\frac{h - h_i}{b} \right) \tag{1.39}$$

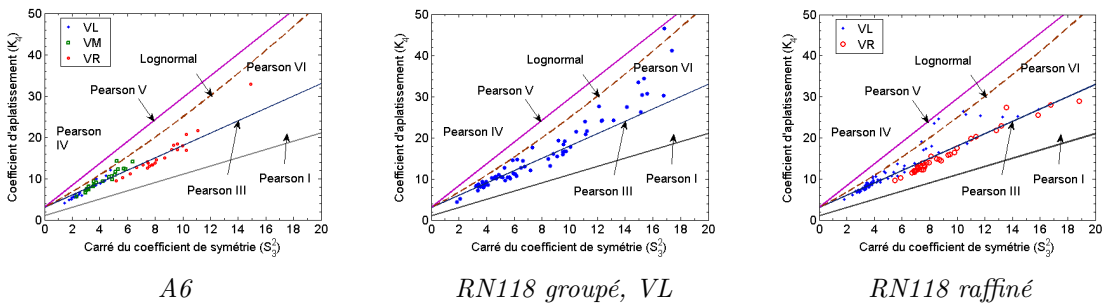


FIGURE 1.2 : Relations K - S^2 du modèle log-normal, de la famille de modèles de Pearson et empirique

où le paramètre b représente le niveau de lissage. La fonction $K(x)$ est une *fonction noyau* qui a différentes formules. La fonction considérée comme une des plus efficaces est la fonction noyau d'Épanechnikov [106] :

$$K(x) = \frac{3}{4\sqrt{5}} \left[1 - \frac{1}{5}x^2 \right] \times \mathbb{1}_{\{|x| < \sqrt{5}\}} \quad (1.40)$$

Dans le cadre de cette étude, l'intervalle $a = 0,1$ s convient très bien pour l'histogramme des TIV. On constate que le choix de $b = 0,3$ s assez faible fait déjà perdre certaines informations comme montré sur la Figure 1.3. Bien que les valeurs du noyau peuvent être utilisées dans l'estimation des paramètres en vue d'obtenir plus aisément les résultats numériques, cette technique n'est pas utilisée dans notre recherche.

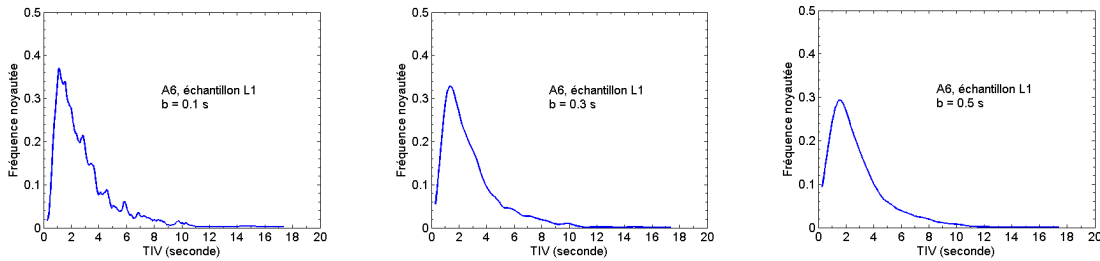


FIGURE 1.3 : L'effet de lissage de l'histogramme par l'utilisation de la fonction Épanechnikov

Parmi les fonctions décrivant la distribution des TIV, la *fonction aléatoire* \mathcal{H} et la fonction *de survie* empirique R_n sont parfois utilisées. Sachant que cette dernière fonction est constante dans le cas d'une distribution exponentielle, elle est largement utilisée dans les études sur la queue de distribution (c.à.d la partie de distribution où se trouvent les TIV élevés). Les fonctions aléatoires théorique et empirique sont respectivement formulées comme ci-dessous :

$$\mathcal{H}(h) = \frac{f(h)}{1 - F(h)} \quad (1.41)$$

$$\mathcal{H}_{n,c}(h) = \frac{\text{Card} \{h_i \mid h \leq h_i < h + c\}}{\text{Card} \{h_i \mid h \leq h_i\}} \quad (1.42)$$

où c est l'intervalle de calcul.

La fonction de survie représente la proportion $R_n(h)$ des valeurs supérieures à une valeur h donnée. Sachant que la distribution des TIV est de forme exponentielle pour les TIV élevés, le logarithme de la fonction de survie est souvent utilisé :

$$\log R_n(h) = \log \left[1 - \frac{j}{n+1} \right] \quad \text{Avec } h_{(j-1)} < h \leq h_{(j)} \quad \text{et } j = 1, 2, \dots, n \quad (1.43)$$

où l'usage de $n+1$ à la place de n dans l'Équation 1.43 permet d'éviter l'indéfini $\log 0$ lorsque $j = n$. Dans la Figure 1.4, la tendance exponentielle des TIV élevés est bien représentée par la linéarité des courbes empiriques. Les fonctions \mathcal{H} et $\log R_n$ ne sont pas utilisées dans notre étude.

1.3.2 Analyse spectrale

L'analyse spectrale n'est valable que sur les échantillons ayant des TIV consécutifs. Soit un échantillon raffiné $H^r = (H_1, H_2, \dots, H_n)$ de valeurs (h_1, h_2, \dots, h_n) où les H_i

sont i.i.d et h_i est le i ième TIV de la série des TIV. Comme la suite H^r est un processus ponctuel, les deux méthodes d'analyse suivantes sont applicables en complément de l'analyse statistique descriptive :

1. L'analyse du nombre de passages (ou l'analyse de dénombrement)
2. L'analyse des intervalles.

Une des propriétés importantes pour une série d'événements est la *stationnarité*. Une suite dite *stationnaire*, dans le cas unidimensionnel des TIV, signifie que la loi du nombre d'événements dans un intervalle déterminé $(t_1, t_2]$ est invariante par translation ; autrement dit, la loi en question reste aussi la même pour l'intervalle $(t_1 + \Delta t, t_2 + \Delta t]$ quel que soit Δt . En conséquence, la définition donne de l'importance au choix de l'instant de début à partir duquel on commence à observer la série.

Le choix de l'instant de début peut être pris de deux façons : soit il est choisi de manière arbitraire (dans ce cas l'intervalle entre l'instant choisi et l'événement le plus proche précédemment est noté par W), soit cet instant coïncide avec un des instants d'arrivée.

Le processus de *dénombrement* se base sur les nombres de passages dans la période durant t secondes à partir de l'instant initial. Le processus d'*intervalles* quant à lui s'appuie sur les écarts temporels consécutifs entre les instants d'arrivée. Étant donné que les TIV ne sont autres que les intervalles en temps, notre étude est focalisée sur le processus des intervalles. Par ailleurs, la distribution de la variable W est différente de celle des intervalles entre événements. C'est la raison pour laquelle, le processus commençant par un événement arbitraire fait l'objet de la recherche.

En général, étant donné une série (H_1, H_2, \dots) stationnaire, il existe toujours une relation entre le processus de dénombrement et le processus d'intervalles. En effet, désignons par N_t le nombre d'événements observés dans une période temporelle $(0, t]$ commençant lors d'un événement quelconque et qui ne comprend pas cet événement, on a :

$$\Pr\{N_t < r\} = \Pr\{H_1 + H_2 + \dots + H_r > t\} \quad \text{Avec } r = 1, 2, \dots \quad (1.44)$$

L'équation 1.44 peut être éventuellement dérivée si les H_i sont i.i.d d'autant plus si la loi commune des H_i est connue et simple.

Les propriétés statistiques du second ordre

Possédant un échantillon raffiné, on peut aborder les statistiques du second ordre. Les propriétés statistiques du second ordre d'un processus s'appuient notamment sur les calculs d'*auto-covariance* et d'*auto-corrélation*. Il s'agit des moments mixtes de second ordre entre les v.a.r H_i d'une suite chronologique. Ces propriétés caractérisent statistiquement les

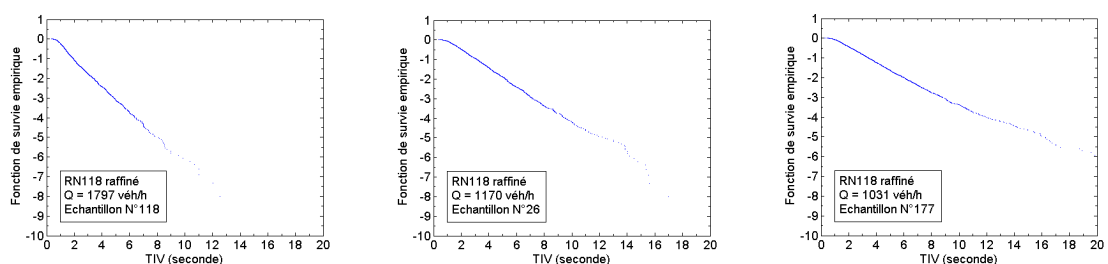


FIGURE 1.4 : Logarithme de la fonction de survie des données empiriques

liaisons entre les composantes de la suite d'événements. La fonction d'auto-covariance d'ordre k , noté C_k , est donnée par :

$$\begin{aligned} C_k &= \text{cov}(H_i, H_{i+k}) \\ &= \int_0^\infty (H_i - \mathbb{E}[H]) (H_{i+k} - \mathbb{E}[H]) f(h) dh \end{aligned} \quad (1.45)$$

La fonction d'auto-corrélation d'ordre k correspond à la formule :

$$\begin{aligned} \rho_k &= \frac{\text{cov}(H_i, H_{i+k})}{\text{Var}[H]} \\ &= \text{Corr}(H_i, H_{i+k}) \end{aligned} \quad (1.46)$$

Wold [107, 100] a montré qu'une suite $\{\rho_k\}_k$ quelconque est une série de fonctions d'auto-corrélation d'une suite aléatoire stationnaire $\{H_i\}$ si et seulement si chaque ρ_k peut être mis sous la forme suivante :

$$\rho_k = \int_{-\pi}^{\pi} \cos(k\omega) dF(\omega) \quad \text{Avec } k = 0, \pm 1, \pm 2, \dots \quad (1.47)$$

Où $F(\omega)$ est une fonction de répartition symétrique continue, dérivable et appelée *fonction de répartition spectrale*. La dérivée de $F(\omega)$ donne une autre fonction $f(\omega)$ appelée la *densité de probabilité spectrale* qui s'écrit selon les ρ_k :

$$\begin{aligned} f(\omega) &= \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \rho_k e^{-ik\omega} \\ &= \frac{1}{2\pi} \left[1 + 2 \sum_{k=1}^{\infty} \rho_k \cos(k\omega) \right] \quad \text{Avec } -\pi \leq \omega \leq \pi \end{aligned} \quad (1.48)$$

où $\rho_k = \rho_{-k}$ sont donc les coefficients de Fourier de $f(\omega)$:

$$\rho_k = \int_{-\pi}^{\pi} f(\omega) \cos(\omega) d\omega \quad \text{Avec } k = 0, \pm 1, \pm 2, \dots \quad (1.49)$$

La fonction \mathcal{P} appelée le *spectre de puissance* est aussi importante dans l'étude de la structure d'un processus :

$$\mathcal{P}(\omega) = \sigma^2 f(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} C_k \cos(k\omega) \quad \text{Avec } -\pi \leq \omega \leq \pi \quad (1.50)$$

Où σ est l'écart-type de la v.a.r H

Estimateurs des statistiques du second ordre

Étant donné des observations $\{h_1, h_2, \dots, h_n\}$, l'auto-covariance d'ordre k est estimée en tenant compte de la valeur moyenne de la série comme suit :

$$\hat{C}_k = \frac{1}{n-k} \sum_{i=1}^{n-k} (h_i - \bar{H}'_k) (h_{i+k} - \bar{H}''_k) \quad (1.51)$$

$$= \frac{1}{n-k} \sum_{i=1}^{n-k} h_i h_{i+k} - \bar{H}'_k \bar{H}''_k \quad (1.52)$$

où :

$$\bar{H}'_k = \frac{1}{n-k} \sum_{i=1}^{n-k} h_i, \quad \bar{H}''_k = \frac{1}{n-k} \sum_{i=1}^{n-k} h_{i+k} \quad (1.53)$$

Par ailleurs, on obtient l'approximation ci-dessous lorsque n est grand :

$$\hat{C}_k \sim C_k - \frac{\sigma^2}{n-k} \sum_{i=-\infty}^{\infty} \rho_i \quad (1.54)$$

L'estimateur d'auto-corrélation d'ordre k est de la forme :

$$\hat{\rho}_k = \frac{\hat{C}_k}{\sqrt{\hat{C}'_{0,k} \hat{C}''_{0,k}}} \quad (1.55)$$

où :

$$\hat{C}'_{0,k} = \frac{1}{n-k} \sum_{i=1}^{n-k} (h_i - \bar{H}'_k)^2, \quad \hat{C}''_{0,k} = \frac{1}{n-k} \sum_{i=1}^{n-k} (h_{i+k} - \bar{H}''_k)^2 \quad (1.56)$$

Cet estimateur est *sans biais* et sa variance dépend de la quantité Δ suivante :

$$\text{Var} [\hat{\rho}_k] \approx \Delta = \frac{1}{n-k} \sum_{i=-\infty}^{\infty} \rho_i^2 \quad (1.57)$$

La quantité Δ est estimée par l'approximation suivante [100] :

$$\hat{\Delta} = \frac{1}{2} \left[1 + 2 \sum_{i=1}^{n-1} \left(1 - \frac{i}{n} \right) \rho_i^2 \right] \quad (1.58)$$

En posant $\theta = m/n$, si l'on ne fait intervenir que m coefficients $\hat{\rho}_1, \hat{\rho}_2, \dots, \hat{\rho}_m$, l'estimation pourrait prendre la valeur $\hat{\Delta}_\theta$:

$$\hat{\Delta}_\theta = \frac{1}{(1+\theta)^2} \left[1 + 2 \sum_{i=1}^m \left(1 - \frac{i}{n} \right) \rho_i^2 \right] \quad (1.59)$$

Dans les équations (1.51) et (1.55), $\hat{C}_{n-1} = \hat{C}'_{0,n-1} = \hat{C}''_{0,n-1} = 0$, le calcul de $\hat{\rho}_{n-1}$ n'a plus de sens. De même, dans certains cas où $k = n-2$, $\hat{\rho}_k$ n'a pas non plus de sens lorsque les deux TIV sont identiques. Pour éviter ces cas, on recalcule dans cette étude les estimateurs d'auto-covariance avec la moyenne commune de l'échantillon et avec $\hat{C}'_{0,n-1} = \hat{C}''_{0,n-1} = \text{Var} [H]$. Par ailleurs, les calculs de $\hat{\Delta}_\theta$ sont réalisés en complément avec $m = 100$.

Estimateurs du spectre des intervalles

Des estimateurs de $f(\omega)$ et de $\mathcal{P}(\omega)$ sont calculés en se limitant à un certain nombre défini de ω sur l'intervalle $[-\pi, \pi]$. La périodicité de la fonction **cosinus** permet de limiter l'analyse aux valeurs ω de la forme :

$$\omega_p = \frac{2\pi p}{n} \quad \text{Avec } p = 1, 2, \dots, [n/2] \quad (1.60)$$

Un estimateur sans biais de $\mathcal{P}(\omega)$ s'écrit comme suit [100] :

$$\begin{aligned} I_n(\omega_p) &= \frac{1}{2\pi} \sum_{i=-n+1}^{n-1} \left(1 - \frac{|i|}{n} \right) \hat{C}_i \cos(i\omega_p) \\ &= \frac{1}{\pi} \left[\frac{C_0}{2} + \sum_{i=1}^{n-1} \left(1 - \frac{i}{n} \right) \hat{C}_i \cos(i\omega_p) \right] \end{aligned} \quad (1.61)$$

Il est à noter que les distributions de $I_n(0)$ et $I_n(\pi)$ sont légèrement différentes de celles des autres $I_n(\omega_p)$. De plus, la statistique $I_n(\omega_p)$ n'est pas un estimateur *exhaustif*. Il sert au test de renouvellement et à vérifier le caractère aléatoire d'une série de TIV. En conséquence, l'étude des caractéristiques d'échantillonnage de $I_n(\omega_p)$ est tout à fait utile.

La fonction de densité spectrale $f(\omega)$ peut être calculée directement :

$$\hat{f}(\omega_p) = \frac{1}{2\pi} \sum_{i=-n+1}^{n-1} \hat{\rho}_i \cos(i\omega_p) \quad (1.62)$$

Les études de Daniel, de Barlett, de Tukey et de Parzen (cf. Cox & Lewis [100]) fournissent des estimateurs exhaustifs de $f(\omega)$ sous la forme générale suivante :

$$\hat{f}_k(\omega) = \int_{-\pi}^{\pi} \hat{h}_k(\omega - y) \hat{f}(y) dy \quad (1.63)$$

$$= \frac{1}{2\pi} \sum_{i=-\infty}^{\infty} \lambda_i^{(k)} \hat{\rho}_i \cos(i\omega) \quad (1.64)$$

Les estimateurs sont calculés sur un nombre fini l de l'indice i . Le biais asymptotique des estimateurs est réduit en augmentant l , et par conséquent en diminuant l'intervalle de variation de la fonction $\hat{h}_k(\omega)$. Les fonctions $\hat{h}_k(\omega)$ appelées fonctions de *pondération spectrales* sont listées dans le Tableau A.1 de l'Annexe. L'analyse peut être limitée à l'intervalle fermé $[0, \pi]$. Après avoir développé en série en choisissant l , $(l + 1)$ valeurs de $\hat{f}_k(\omega_p)$ avec $\omega_p = \pi p/l$ suffisent pour calculer les estimateurs puisque toutes les valeurs de $\hat{f}_k(\omega)$ appartiennent à un polynôme trigonométrique de degré $l + 1$. En conséquence, plus l est grand, plus l'on est capable de détailler le spectre des TIV.

Dans notre étude, l est égal à 20 et le paramètre ϵ de la fonction $\hat{f}_0(\omega)$ est égal à 6 (cf. Tableau A.1)

1.3.3 Run test

Le *run test* vérifie le phénomène de groupement des données, c. à. d il vérifie si les TIV longs et les TIV courts sont distribués de façon aléatoire. Le test utilisé est celui de Wald et Wolfowitz [108] s'appuyant sur la comparaison entre les TIV h_i et la *médiane* M_d de l'échantillon. La variable x_i suivante est associée au TIV h_i :

$$x_i = \begin{cases} 1 & \text{si } h_i > \widetilde{M}_d \\ 0 & \text{si } h_i < \widetilde{M}_d \end{cases} \quad (1.65)$$

où \widetilde{M}_d est l'estimateur de la *médiane* de l'échantillon.

On établit une variable R qui signifie le nombre de groupes dans lesquels les valeurs consécutives de x_i sont identiques. Cette variable est appelée un « *run* » et sa distribution tend approximativement vers la distribution normale $\mathcal{N}(\mu, \sigma)$ quand n est grand.

$$\mu = \frac{2r(n-r)}{n} + 1 \quad (1.66)$$

$$\sigma^2 = \frac{2r(n-r)[2r(n-r) - n]}{n^2(n-1)} \quad (1.67)$$

Luttinen [97] (1992) a obtenu des groupements significatifs dans ses données pour tous les débits sur les routes où la limite de vitesse est comprise entre 80 et 100 km/h, et pour les débits supérieurs à 700 véh/h par voie sur les routes où la limite de vitesse est comprise entre 50 et 70 km/h.

1.3.4 Test de renouvellement

L'hypothèse de *renouvellement* est l'hypothèse d'un processus sans mémoire supposant que les TIV sont indépendants et identiquement distribués (i. i. d). L'hypothèse nulle est :

$$H_0 : \Pr \{H_i \leq h \mid H_1, \dots, H_{i-1}\} = \Pr\{H_i \leq h\} \quad (1.68)$$

L'hypothèse alternative est que les TIV H_i sont corrélés mutuellement.

Corrélogramme et Test de renouvellement sur les ρ_i

Le *corrélogramme* est établi par la représentation de toutes les valeurs estimées $\hat{\rho}_i$ des coefficients d'auto-corrélation ρ_i avec $i = 0, 1, \dots$. Le test de renouvellement peut être basé sur les coefficients d'auto-corrélation en utilisant la statistique $\hat{\rho}_1 \sqrt{n-1}$ qui suivra la loi normale centrée réduite $\mathcal{N}(0, 1)$ lorsque n est grand. Par conséquent, le test au niveau de signification de α consiste à rejeter l'hypothèse nulle H_0 si :

$$|\hat{\rho}_1 \sqrt{n-1}| > c_{\alpha/2} \quad \text{où} \quad |\hat{\rho}_1| > \frac{c_{\alpha/2}}{\sqrt{n-1}} \quad (1.69)$$

où $c_{\alpha/2}$ est le quantile de la loi normale centrée réduite au delà duquel la probabilité est $(1 - \alpha/2)$ que les TIV sont supérieurs.

La statistique R_1

Supposons que l'on a n estimations $\hat{\rho}_i$ des $\rho_1, \rho_2, \dots, \rho_n$. On ordonne cet ensemble de sorte que $r_j = \hat{\rho}_{(j)}$. La statistique R_1 est définie par :

$$R_1 = \sum_{j=1}^{n-1} r_{j+1} r_j \quad (1.70)$$

Sous l'hypothèse de l'indépendance des coefficients de covariances sériales, la statistique R_1 tend vers une loi normale avec la moyenne et la variance de la forme :

$$\mathbb{E} [R_1] = \frac{1}{12} (n-1)(n+1)(3n+2) \quad (1.71)$$

$$\text{Var} [R_1] = \frac{5n^6 + 16n^5 - 14n^4 - 80n^3 - 35n^2 + 64n + 44}{720(n-1)} \quad (1.72)$$

Périodogramme

Le *périodogramme* est établi par la représentation des valeurs estimées des $I_n(\omega_p)$ dans l'intervalle $[0, \pi]$. Par commodité, les valeurs de $I_n(0)$ (dans le cas n est pair) et de $I_n(\pi)$ sont négligées (cf. [100]). Le test d'un processus de renouvellement peut être basé sur les valeurs de $I_n(\omega)$ en sachant que les $I_n(\omega_p)$ suivent à la limite des lois exponentielles indépendantes ayant la même moyenne de $\sigma^2 f(\omega)$. Notons ainsi que le spectre de puissance d'un processus de renouvellement est constamment égal à $\sigma^2/2\pi$.

La première possibilité du test de renouvellement consiste à utiliser la statistique s_i^2 pour tester l'homogénéité des quantités $I_n(\omega_p)/f(\omega_p)$.

$$s_i^2 = \sum_{p=(i-1)v+1}^{iv} \frac{I_n(\omega_p)}{f(\omega_p)} \quad (i = 1, 2, \dots, k) \quad (1.73)$$

où v est la taille de k groupes contigus des valeurs de $I_n(\omega_p)$ sachant que le nombre total des $I_n(\omega_p)$ est $l = kv$. Dans ce cas, la statistique H_0 suivante représente l'homogénéité de la suite des TIV :

$$H_0 = \frac{\left[2l \log \left(\frac{1}{2l} \sum_{i=1}^k s_i^2 \right) - \sum_{i=1}^k 2v \log \left(\frac{s_i^2}{2v} \right) \right]}{(6v-2)(6v-3)} \quad (1.74)$$

Si $2v > 5$, H_0 suit approximativement la loi du χ^2 à $k-1$ degrés de liberté. La valeur de $f(\omega_p)$ dans l'équation 1.73 peut être prise égale à la valeur constante $1/2\pi$.

La deuxième possibilité est de tester l'uniformité de la statistique suivante :

$$U_{(i)} = \frac{\sum_{p=1}^i I_n(\omega_p) / f(\omega_p)}{l} \quad (1.75)$$

$$\sum_{p=1} I_n(\omega_p) / f(\omega_p)$$

Test de la loi géométrique de la longueur des pelotons

Si le peloton est défini au moyen d'un seuil sur le TIV δ_p , et si l'hypothèse de renouvellement est vérifiée, alors la longueur du peloton (en nombre de véhicules) suit une loi géométrique. L'hypothèse de renouvellement peut aussi être examinée par un test sur la distribution géométrique de la taille de peloton.

Tout d'abord un seuil δ est défini servant à distinguer si un véhicule est un *leader* (il est codé par conséquent une valeur 0) ou un *suiveur* (la valeur associée vaut 1) :

$$y_i = \begin{cases} 1 & \text{si } h_i > \delta \\ 0 & \text{si } h_i \leq \delta \end{cases} \quad (1.76)$$

On nomme ℓ la longueur d'un peloton, ℓ étant le nombre des suiveurs consécutifs plus 1 (c. à d le peloton comprend aussi le leader). Si le processus des TIV est de renouvellement, la probabilité qu'un véhicule soit suiveur et la probabilité que la longueur ℓ vaille k sont données respectivement par les formules :

$$p = \Pr \{y_i = 0\} = \Pr \{h_i \leq \delta\} = F(\delta) \quad (1.77)$$

$$p_k = \Pr \{\ell = k\} = p^{k-1} (1 - p) \quad (1.78)$$

La v.a.r H étant i.i.d et F étant la fonction de répartition des TIV, les hypothèses nulle et alternative sont énoncées comme suit :

$$H_0 : p_k = p^{k-1} (1 - p) \quad (1.79)$$

$$H_1 : p_k \neq p^{k-1} (1 - p) \quad (1.80)$$

Avec k connu, l'hypothèse nulle peut être testée par le test du χ^2 à $k-2$ degré de liberté. Après avoir utilisé ce test, Luttinen [97] (1992) a rejeté H_0 pour les échantillons ayant un débit supérieur à 500 véh/h par voie sur les routes où la limite de vitesse est comprise entre 80 et 100 km/h. Luttinen n'a pas pu rejeter H_0 pour les échantillons sur les routes où la limite de vitesse est comprise entre 50 et 70 km/h.

1.4 Les données réelles

1.4.1 Les données de l'Autoroute A6

Présentation du site

Les données sont collectées sur l'autoroute A6 dans le sud de Paris en 1996. Le site de mesure est rectiligne et loin des échangeurs. Le volume moyen de trafic journalier est environ de 70 000 véhicules dont 5 % de Poids Lourds (PL). La vitesse limite autorisée de la section étudiée était 110 km/h. Par ailleurs, les données ont aussi été enregistrées sur le tronçon où le profil en travers était réduit sur 2 km de long par des lignes de marquage. Cette réduction de voie n'a cependant pas été communiquée en avance aux conducteurs par exemple par une signalisation. La largeur de voie de 3,25 m a été réduite de 13 % sur toutes les voies.

Les périodes mesurées étaient le matin, l'après midi et le soir en semaine. Le système SIREDO (Système Informatisé de REcueil de DONnées) a enregistré les données véhicule par véhicule. Trois informations principales sont enregistrées : le temps d'arrivée (en heures **h**, en minutes et en secondes **s**, en de centièmes d'une seconde), la vitesse instantanées (en km/h) et la longueur des véhicules (en mètre **m**). De plus, les données collectées sont qualifiées par des caméras fixées en vue d'observer les événements réels. Ce système de caméras aide à vérifier les informations obtenues par le système de boucles magnétiques SIREDO.

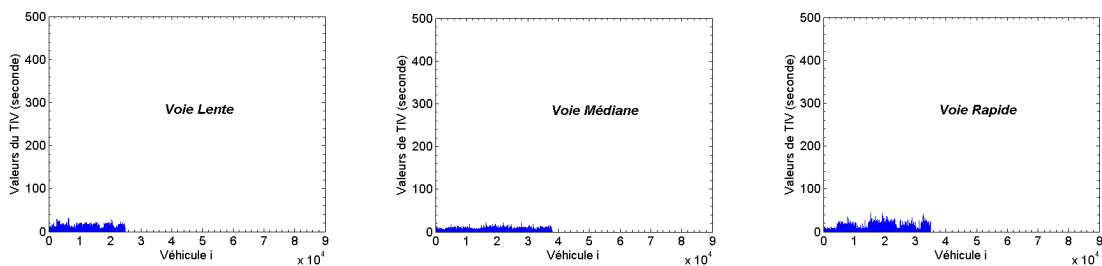


FIGURE 1.5 : Les valeurs des TIV sur l'autoroute A6

Les données ont été collectées pendant 25 heures, sur les trois voies séparément. Au total, l'expérience a été effectuée pendant 10 jours non consécutifs, entre 8h : 00 et 15h : 00 et aussi de 18h : 00 à 21h : 00.

La Figure 1.5 donne une vue globale des valeurs des TIV sur 3 les voies de l'A6. Les silhouettes des valeurs des TIV montrent *a priori* la différence en général des TIV sur les trois voies : voie lente (VL), voie médiane (VM) et voie rapide (VR).

Chaque échantillon de la base de données de l'A6 correspond à une heure d'enregistrement. Le choix que $\Delta = 1$ heure s'est fait car le nombre de véhicules observés n'a pas été suffisamment élevé pour appliquer la méthode de groupement utilisant Δ de quelques minutes et la méthode de raffinement. Pourtant, les analyses statistiques dans la section suivantes montreront que ces échantillons sont assez stables.

Trafic aux différentes périodes

Le Figure 1.6 représente les débits horaires sur chaque voie en fonction des heures de la journée. Chaque petite ligne horizontale dure une heure et correspond à un échantillon. On observe tout d'abord que les débits de l'échantillon (8h – 9h) étaient nettement plus élevés que ceux des échantillons de 10h à 15h. De plus, dans la période (8h – 9h), il y a plus de véhicules sur la VR que sur les VM et VL. Au contraire, entre 10h à 15h (période

où le niveau du trafic est intermédiaire), plus de conducteurs ont choisi la VM au lieu de la VR. L'utilisation de la VM semble permettre aux conducteurs de conduire avec plus de flexibilité en terme de changement de voie.

Dans les échantillons examinés de 8h à 15h, les débits horaires les plus faibles se trouvent sur la VL. En général, plus la voie est lente, plus les débits horaires sont stables. Pourtant, les débits horaires des trois voies de la période (18h – 21h) sont similaires. Il faut noter que les données de cette période étaient collectées dans les sections de profil réduit, et correspondaient dans la majorité des cas à l'état de *bouchon* du trafic.

Les remarques précédentes sont valides à l'exception des cinq échantillons notés sur la Figure 1.6 qui se trouvent soit sur la VL (l'échantillon L8 de période (11h – 12h) et l'échantillon L11 de période (14h – 15h)), soit sur la VR (R8, R11, et R7c). Ils coïncident par ailleurs au phénomène de bouchon dans la section de profil en travers réduit.

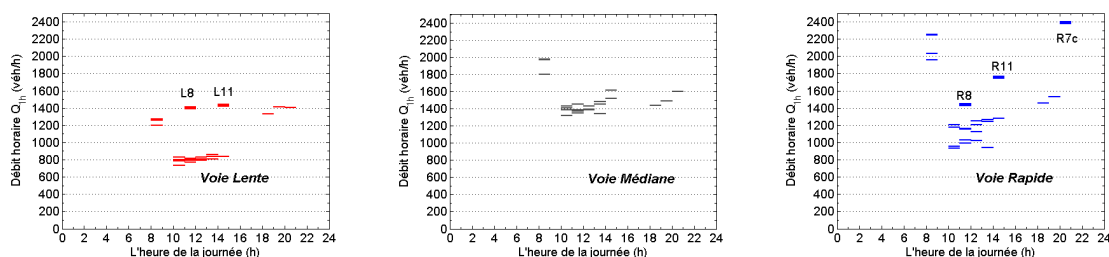


FIGURE 1.6 : Débits horaires sur trois voies selon l'heure de la journée

Analyse statistique

La proportion des PL - les véhicules de longueur $l_i > 7$ m - est particulièrement élevée sur la VL dans la période (10h – 15h). Cette haute proportion atteint près de 30% à midi. Les pourcentages des PL sur la VM sont très faibles (inférieur à 2 %), voire nuls sur la VR dans la plupart des échantillons (cf. Tableau B.1, Annexe).

La moyenne et l'écart-type des TIV diminuent en fonction du débit horaire (cf. Figure 1.7). L'écart-type des TIV des échantillons de la VR est nettement supérieur à ceux des échantillons de la VL et de la VM. Les variations des valeurs d'écart-type des TIV en fonction du débit horaire sur la VM et la VL semblent être similaires.

Les valeurs de CV sont stables et inférieures à 1 sur la VL et la VM. Sur la VR, le CV est supérieur à 1 dans les échantillons de débit horaire autour de 1000 véh/h, et inférieur à 1 quand le trafic est plus dense (plus de 1500 véh/h).

Les valeurs des coefficients de symétrie S_3 des échantillons des VL et VM sont similaires et proches de 2. Ces valeurs sont plus élevées dans les échantillons de la VR. De plus, la relation $K-S^2$ semble être linéaire sur les trois voies. De ce point de vue, la Figure 1.7 montre encore une fois la similarité statistique des échantillons de la VL et de la VM légèrement différente de celle des échantillons de la VR.

Les autres propriétés statistiques sont récapitulées dans le Tableau B.1 dans l'Annexe où ρ_1 est le coefficient d'auto-corrélation entre un véhicule et le véhicule précédent et p est le pourcentage des PL. L'interaction entre véhicules représentée par ρ_1 est plus significative pendant l'heure de pointe (de 8h à 9h). Des valeurs négatives sont apparues sur la VL qui correspondent à un taux élevé du pourcentage p .

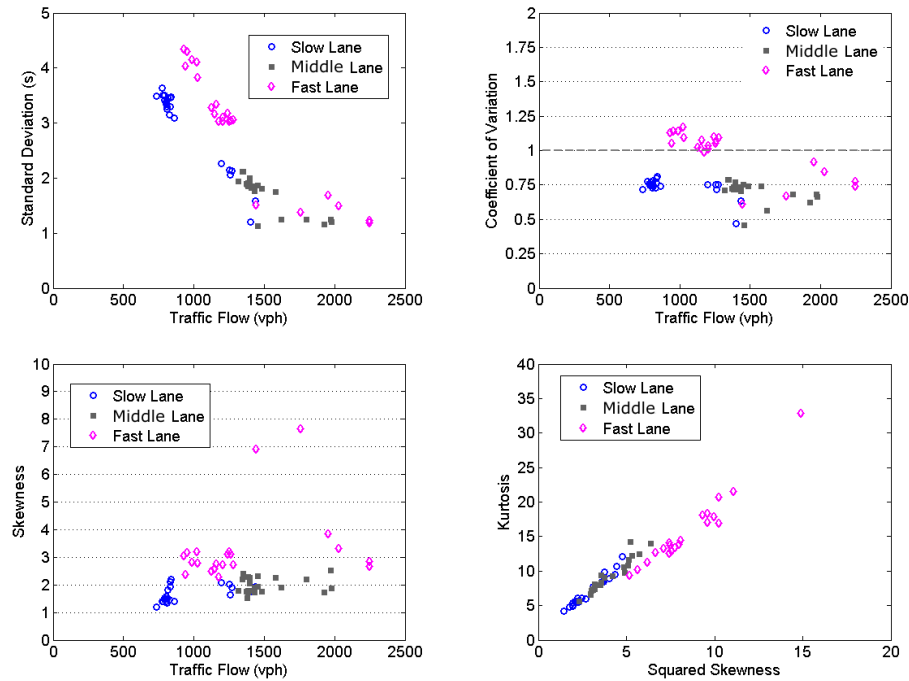
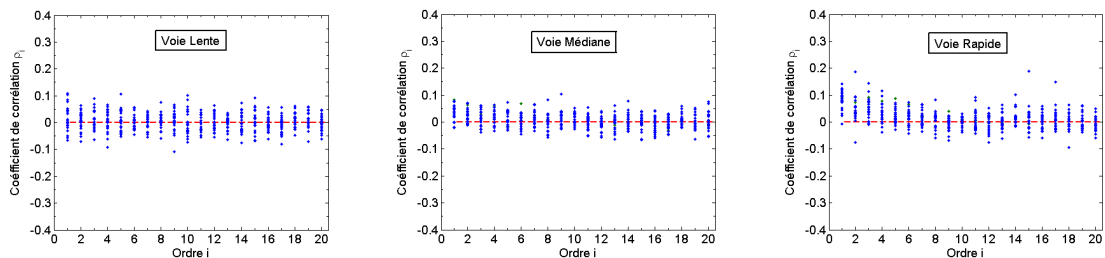


FIGURE 1.7 : Les propriétés statistiques des TIV sur l'A6

Analyse spectrale

La Figure 1.8 donne les valeurs calculées des coefficients d'auto-corrélation des TIV sur les trois voies VL, VM et VR de l'A6. Les valeurs des ρ_i sur la VM sont généralement plus petites que celles sur les VL et VR. Les valeurs ρ_1 et ρ_2 sur la VM sont majoritairement positives, il en est de même pour les cinq premiers ρ_i (avec $i = 1, 2, \dots, 5$) des échantillons de la VR. Sur les 20 premiers ρ_i des échantillons de la VL, les valeurs sont réparties également entre les valeurs positives et négatives. De plus, les coefficients des échantillons de la VL et de la VM sont généralement compris entre -0.1 et 0.1 tandis qu'il existe plusieurs échantillons de la VR ayant des coefficients d'ordre de 1 à 4 supérieurs à 0.1 .

FIGURE 1.8 : Coefficient d'auto-corrélation d'ordre i sur les trois voies

La Figure 1.9 fournit les valeurs calculées des estimateurs Δ et Δ_θ . On remarque qu'en faisant intervenir $\theta = 100$ premières valeurs de ρ_i , le Δ_θ remonte d'environ 50 %. Dans tous les échantillons, l'estimateur Δ est de l'ordre de 1 et Δ_θ est de l'ordre de 1,5. Par conséquent, la variance des $\hat{\rho}_i$ est très faible lorsqu'on prend en compte un grand nombre d'individus n dans chaque échantillon.

La Figure 1.10 trace les valeurs estimées des densités spectrales \hat{f}_k avec $k = 0, 1$. L'estimateur \hat{f}_0 est calculé en utilisant $\epsilon = 6$ et l'estimateur \hat{f}_1 est calculé en utilisant $l = 100$. Les valeurs estimées de f_1 et f_2 sont très proches l'une de l'autre. La Figure

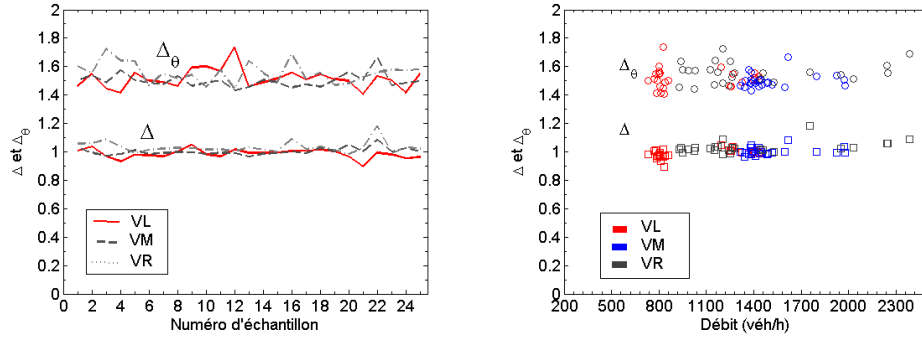


FIGURE 1.9 : Valeurs des estimateurs de la somme des carrés des ρ_i

1.10 montre aussi que les processus des TIV sur la VR sont ceux qui sont les plus éloignés d'un processus de renouvellement. Aussi, il semble que le comportement dans les cas de bouchon sur la VR est différent de ceux sur les VL et VM alors que ces deux derniers sont assez similaires. Sur les trois voies, le processus des TIV à l'heure de pointe est légèrement différent de ceux des autres heures (cf. Figure 1.10, première ligne).

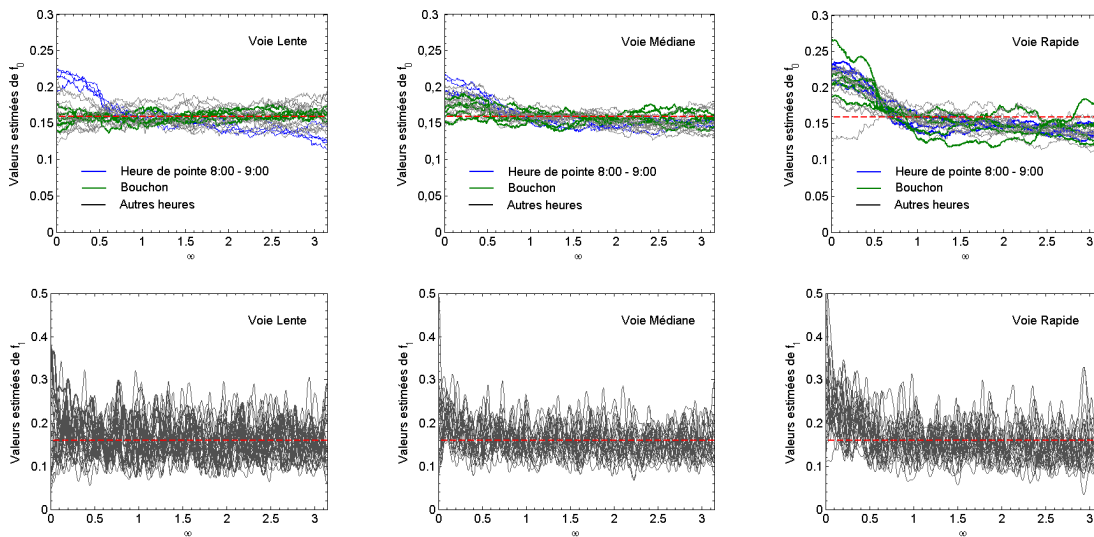


FIGURE 1.10 : Les valeurs estimées de la densité spectrale dans les échantillons

Il semble plus commode d'étudier l'estimateur $I_n(\omega_p)$ que d'étudier les \hat{f}_k car ces derniers sont très sensibles aux paramètres de lissage l et ϵ . Le périodogramme des estimateurs $I_n(\omega_p)$ sur l'intervalle $(0, \pi)$ de ω est tracé pour chaque échantillon. La Figure 1.11 présente les périodogrammes des échantillons L1, M1 et R1 respectivement dans la même période d'enregistrement.

Le Tableau 1.2 représente les valeurs des moyennes et des écarts-types de $2I_n/s^2$ pour tous les échantillons de l'A6. Dans l'ensemble des échantillons étudiés, les moyennes des $2I_n/s^2$ sur les trois voies sont très proches de la valeur $1/\pi$ d'un processus de renouvellement. Cependant, l'écart-type de $2I_n/s^2$ est assez grand, du même ordre que sa moyenne. Il semble être plus correct de comparer les valeurs des écarts-types de $2I_n/s^2$ pour des échantillons ayant les débits similaires.

Sur la même voie, les écarts-types des deux échantillons L6a et L6b sont très proches, respectivement de 0,3226 et 0,3284 (leur différence $\delta = 0,0058$). Pour les échantillons L3b et L3c, les écarts-types sont respectivement 0,2912 et 0,3287 ($\delta = 0,0375$). Il est à noter

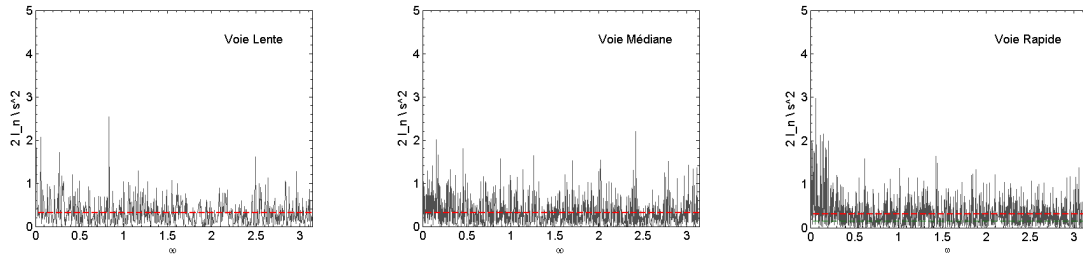


FIGURE 1.11 : Les valeurs de densité spectrale estimées des échantillons L1, M1, R1, de l'A6

que dans la dernière comparaison, on a aussi obtenu de grandes différences de distribution des TIV et de distribution des vitesses entre les deux échantillons L3b et L3c (cf. Chapitre 4). En comparant deux échantillons L1 et R3c correspondant à deux voies différentes avec débits similaires, les écarts-types sont respectivement 0,3056 et 0,3265 ($\delta = 0,0109$).

Écart-type de $2I_n/s^2$				
	Période de pointe	Période creuse		
Section	VL4 & VL1	VL6c & VL3a	VM6a & VM3a	VR6c & VR9c
Normale	0.3056	0.3102	0.3146	0.3163
Réduite	0.3078	0.3414	0.3113	0.3043

TABLE 1.1 : Écart-type de $2I_n/s^2$ sur la voie normale et la voie de largeur réduite

Lorsque l'on étudie l'effet de réduction de la voie sur les arrivées des véhicules, on obtient des comparaisons dans le Tableau 1.1 relatives à quatre couples d'échantillons sur la voie normale et sur la voie réduite. Pour chaque couple, les débits horaires sont équivalents. Le Tableau 1.1 montre que l'écart-type de $2I_n/s^2$ change plus significativement pendant la période creuse sur la VL et la VR. Cependant, les tendances de changement d'écart-type de $2I_n/s^2$ sont inverses entre ces deux voies. De plus, le changement de l'écart-type de $2I_n/s^2$ est faible entre les échantillons correspondants de la VM pendant la période creuse. Étant donné que le pourcentage des PL est élevé dans la période creuse sur la VL, une interprétation est de dire que la forte inertie des PL sur la VL pourrait obliger les conducteurs des PL à faire plus attention à la largeur de la voie en vue de stabiliser leur trajectoire. Par conséquent, ce comportement d'ajustement influence le comportement de l'ensemble des conducteurs. Quant au changement sur la VR, la réduction de largeur de voie pourrait influencer particulièrement les conducteurs qui se déplacent à grande vitesse et qui doivent concilier le champ de vision et l'espace de la voie. Enfin, l'effet de la réduction semble n'apparaître que dans la période creuse pendant laquelle le changement de largeur de voie perçu par des marquages est plus visible aux yeux des conducteurs.

Tests de tendance, de groupement et de renouvellement

Dans cette section, les trois tests de tendance, le *run* test et le test de renouvellement utilisant les statistiques ρ_1 et R_1 sont utilisés. On remarque que :

- Sur les VL et VM, on ne peut pas rejeter l'hypothèse nulle (qu'il n'y a pas de tendance) au seuil de risque à 5 % seulement dans les échantillons L8, M8, L12b, M12b en utilisant les statistiques S et V , et dans les échantillons L8, M8, L12b, M12b, L2, M2 en utilisant la statistique τ . Le *run* test ne rejette l'hypothèse nulle (qu'il n'y a pas de groupement) que pour les échantillons L1, M1, L4, M4 relatifs à

l'heure de pointe.

- Sur la VR, on rejette l'hypothèse nulle dans six échantillons (R2, R6d, R7c, R8, R9d, R12c) en utilisant la statistique S , six échantillons (R2, R3c, R6b, R7c, R8, R9d) en utilisant la statistique τ , et huit échantillons (R2, R3b, R6d, R7b, R7c, R8, R9d, R12c) en utilisant la statistique V . Sauf pour quatre échantillons (R8, R9c, R11, R12a), on rejette l'hypothèse nulle (qu'il n'y pas de groupement) sur la VR.
- En utilisant la statistique ρ_1 sur la VL, le test de renouvellement ne rejette l'hypothèse nulle (que la série des TIV soit i. i. d) que dans quatre échantillons (L1, L2, L4, L5). Le test rejette 19 échantillons sur 25 sur la VM et 8 échantillons sur la VR. La statistique R_1 semble être plus sélective en rejetant 8 échantillons (les quatre précédents sur la VL et les L7a, L7c, L8, L11), 20 échantillons sur la VM et tous les échantillons sur la VR. Il est à noter que les échantillons L1, L2, L4 et L5 sont relatifs à l'heure de pointe et les échantillons L7a, L7c, L8, L11 sont dans les périodes où les bouchons sont observés.

Le test de renouvellement utilisant la statistique d'homogénéité H_0 n'est pas employé car les résultats sont sensibles à la pratique de groupement des valeurs $I_n(\omega_p)$.

Statistique $2 I_n/s^2$						
Échantillon	Moyenne			Écart-type		
	VL	VM	VR	VL	VM	VR
1	0.3189	0.3183	0.3184	0.3056	0.3178	0.3328
2	0.3176	0.3184	0.3170	0.3249	0.3182	0.3253
3a	0.3167	0.3179	0.3182	0.3102	0.3021	0.3554
3b	0.3178	0.3179	0.3164	0.2912	0.3270	0.3493
3c	0.3185	0.3179	0.3183	0.3287	0.3185	0.3265
3d	0.3167	0.3180	0.3181	0.3090	0.3078	0.3289
3e	0.3182	0.3179	0.3184	0.3116	0.3119	0.3439
4	0.3181	0.3181	0.3184	0.3078	0.3159	0.3125
5	0.3181	0.3180	0.3183	0.3246	0.3007	0.3328
6a	0.3186	0.3174	0.3187	0.3226	0.3113	0.3199
6b	0.3184	0.3183	0.3183	0.3284	0.3111	0.3318
6c	0.3175	0.3180	0.3155	0.3414	0.2901	0.3043
6d	0.3169	0.3175	0.3149	0.2875	0.2925	0.3255
7a	0.3185	0.3173	0.3182	0.3149	0.3269	0.3197
7b	0.3180	0.3181	0.3168	0.3276	0.3017	0.3102
7c	0.3174	0.3178	0.3164	0.3019	0.3092	0.3483
8	0.3170	0.3175	0.3176	0.3224	0.3078	0.3038
9a	0.3171	0.3181	0.3181	0.3256	0.3199	0.3257
9b	0.3167	0.3174	0.3175	0.3189	0.3146	0.3098
9c	0.3166	0.3177	0.3182	0.3087	0.3281	0.3163
9d	0.3184	0.3165	0.3174	0.2928	0.3165	0.3100
11	0.3181	0.3166	0.3185	0.3272	0.3455	0.3273
12a	0.3179	0.3173	0.3179	0.3180	0.2911	0.3181
12b	0.3155	0.3182	0.3171	0.2869	0.3188	0.3119
12c	0.3167	0.3182	0.3177	0.3122	0.2992	0.3364

TABLE 1.2 : Moyenne et écart-type de $2 I_n/s^2$ des échantillons de l'A6

1.4.2 Les données de la Route Nationale RN118

Présentation du site

La route nationale RN118 se trouve au sud-ouest de Paris. Elle appartient au réseau de voies rapides urbaines (VRU) en Île-de-France. La RN118 relie la RN10 (au niveau de Sèvres) à l'autoroute A10 (au niveau des Ulis). L'utilisation de la RN118 est gratuite, comme toutes les routes nationales françaises. La RN118 comporte seize échangeurs principaux. Elle permet également la desserte de la zone commerciale et industrielle de Vélizy - Villacoublay ainsi que du Parc d'activités de Courtaboeuf.

Côté enregistrement des données, un pluviomètre est installé à Trappes grâce auquel les informations de pluviométrie sont jointes aux fichiers de données de trafic. Côté trafic, il est à noter que les états du trafic des deux sens vers Province (sens Y) et vers Paris (sens W), sont significativement différents [109]. Par ailleurs, les états du trafic entre deux parties, l'une de la RN10 à l'A86 et l'autre de l'A86 à l'A10, sont nettement distincts. Par exemple, le trafic sur la première partie sens Y était toujours fluide tandis que le trafic sur le deuxième tronçon sens Y était souvent saturé ou ralenti à l'heure de pointe. Côté infrastructure, la RN118 est à deux fois deux voies de largeur de 3,5 m (sauf sur deux sections à 2x3 voies pour lesquelles nous ne disposons pas des données de trafic). La limite de vitesse légale est de 110 km/h ou moins selon la localisation et les périodes par temps sec ou de pluie.

Présentation des données

Cette base de données est la base principale de cette étude. Des millions de véhicules ont été enregistrés sur la RN118. Au total, neuf stations de mesure tout au long de la route ont été mises en service en 2005. De plus, la durée d'expérience s'étale de Janvier à Novembre 2005. Les localisations des neuf stations sont représentées dans le Tableau A.2 de l'Annexe. Elles sont équipées des boucles magnétiques capables de capturer les instants d'arrivée, et par différence de mesurer les longueurs et vitesses des véhicules qui passent. En outre, les boucles installées sous chacune des deux voies permettent en conséquence d'étudier séparément les comportements différents des conducteurs des deux voies.

La base de données de la RN118 comporte des informations sur les passages individuels consécutifs (instant d'arrivée, vitesse, longueur) avec une haute précision (de 0,01 s pour l'instant de passage, d'un cm pour la longueur). Des connaissances météorologiques (l'état de la pluie) y ont aussi été ajoutées. Trois niveaux de pluie sont enregistrés : **non-pluie**, **pluie**, **pluie forte** selon le niveau d'eau mesuré par le pluviomètre.

Correspondant à chaque voie dans chacun des 9 sites de mesure, plusieurs fichiers sont enregistrés. L'ensemble de ces fichiers représente les données du trafic d'un site. Chaque fichier correspond à quelques jours d'un mois. Finalement, les informations de chaque véhicule sont écrites sur une ligne dans les fichiers .TXT.

La Figure 1.12 donne une vue globale sur les valeurs consécutives des TIV de la RN118. En étudiant différents profils des valeurs des TIV, on remarque que :

1. Le trafic de **jour** correspond en général aux TIV faibles représentés par les parties plates entre les pics dans les figures,
2. Les valeurs de pic correspondent au trafic de **nuit** pendant lequel les passages des véhicules se produisent beaucoup plus rarement,
3. À l'intermédiaire entre ces deux types de profils, il existe des *zones transitaires* qui se présentent avant la période de **nuit**, plus remarquablement sur la VL. Cette zone intermédiaire correspond principalement à la période de **soir**,

4. Il existe éventuellement des périodes où les TIV sont *irréguliers*. Dans ces périodes, les TIV sont par exemple particulièrement longs pendant le jour, ou les *zones transitoires* sont plus larges (cf. Figure 1.14).

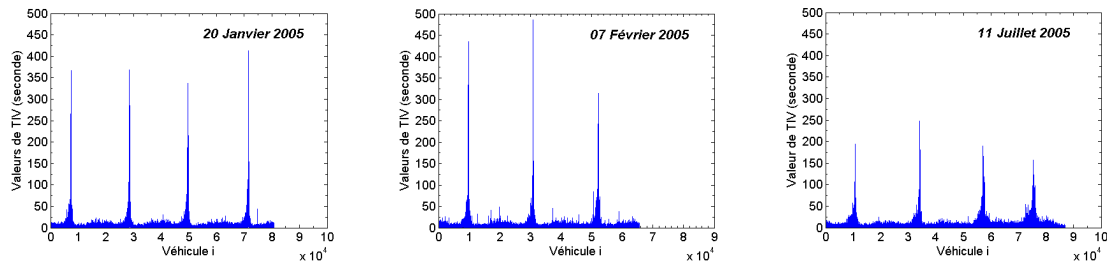


FIGURE 1.12 : Les valeurs des TIV sur la VL de la RN118

La Figure 1.13 illustre les valeurs des TIV sur la VR. Les TIV dans la zone correspondant à la zone plate de la VL ont des profils assez différents.

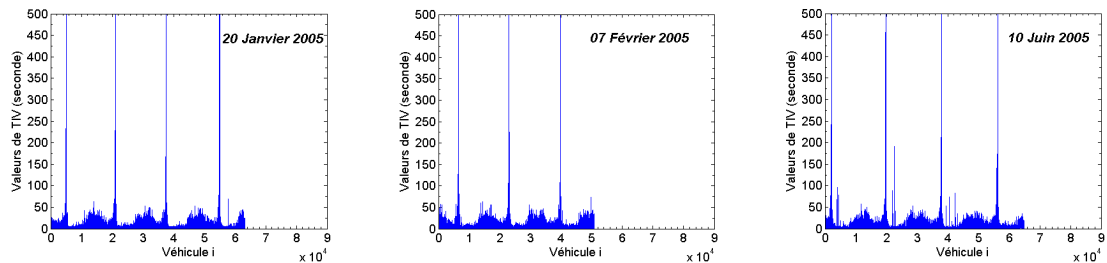


FIGURE 1.13 : Les valeurs des TIV sur la VR de la RN118

En conséquence, en tenant compte les phénomènes périodiques de la journée, les trois périodes principales suivantes sont considérées dans la recherche :

1. jour : (08 :00 - 17 :59)
2. soir (18 :00 - 22 :59)
3. nuit : (01 :00 - 05 :59)

De plus, comme la condition météo est aussi considérée, les cinq modalités suivantes sont prises en compte dans notre étude :

1. jour-sec
2. jour-pluie
3. soir-sec
4. soir-pluie
5. nuit

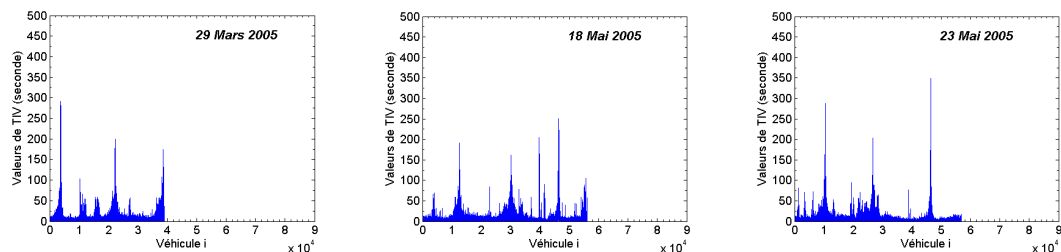


FIGURE 1.14 : Profils particuliers des valeurs de TIV de la RN118, VL

La condition de *pluie* incluant tous les états d'apparition de pluie et une seule modalité de *nuit* sont considérées car il y a eu peu de données. Cette modalité n'est pas le sujet de notre recherche. On se concentre plus particulièrement sur les deux premières modalités.

Finalement, compte tenu de la différence des valeurs des TIV entre les voies VL et VR, les quatre modalités suivantes sont les noyaux de l'exploitation de la base de données RN118 :

1. jour-sec-VL
2. jour-sec-VR
3. jour-pluie-VL
4. jour-pluie-VR

Analyse statistique des échantillons groupés de la RN118

Les propriétés statistiques étudiées dans cette section sont celles des TIV du Site N°1 de la RN118. Les échantillons correspondants des Sites N°1 à N°5 seront utilisés dans le dernier chapitre portant sur la validation des modèles probabilistes.

Les propriétés statistiques sont analysées en fonction des quantités macroscopiques agrégées sur une période Δ de 6 minutes. Les deux variables macroscopiques mises en jeu dans cette section sont le débit et le taux d'occupation.

1. En fonction du débit 6 minutes Q_6

Théoriquement, la moyenne des TIV μ_H d'un échantillon est l'inverse du débit Q_I calculé sur la période I pendant laquelle l'échantillon est collecté. Dans ce cas, le débit Q_I est obtenu après avoir constitué l'échantillon. Au contraire, en utilisant la méthode de groupement, les débits Q_6 sont calculés avant et utilisés comme l'information D_k du critère \mathcal{A} . En conséquence, les valeurs empiriquement obtenues ne se trouvent pas exactement sur la courbe hyperbolique comme le montre la Figure 1.15. L'écart entre les moyennes empiriques et théoriques est significatif pour les trafics fluides parce que, pour une largeur de classe de débit donnée, les TIV correspondant à une classe de trafic fluide ont un écart-type nettement plus élevé que les écarts-types des TIV correspondant à une classe de débit élevé.

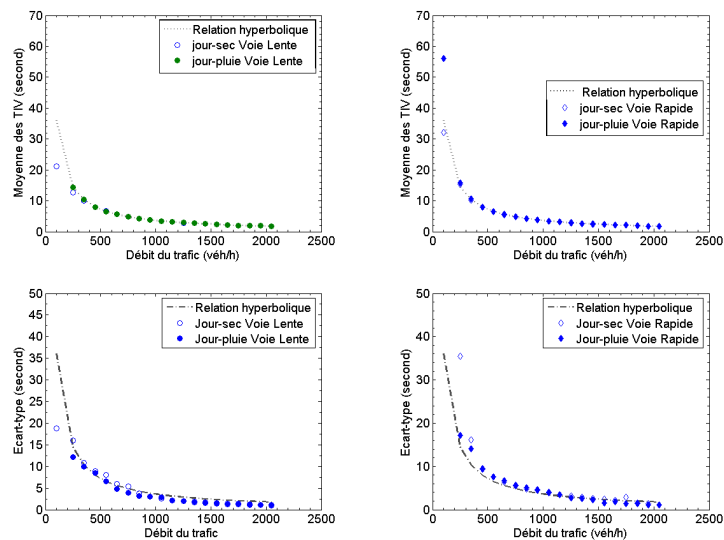


FIGURE 1.15 : Moyenne et écart-type des TIV du jour en fonction de Q_6

Pour les débits supérieurs à 400 véh/h, les différences de moyenne des TIV selon les conditions de météo et selon les voies sont très faibles dû à l'écart-type faible des TIV. Par ailleurs, il n'y a pas d'échantillon de modalité **soir-pluie** pour les débits Q_6 supérieurs à 1800 véh/h.

L'écart-type σ_H des TIV des échantillons diminue en fonction du débit 6 minutes Q_6 . La ligne pointillée dans la Figure 1.15 représente encore une relation hyperbolique entre σ_H et Q_6 . On observe pour les échantillons ayant les débits Q_6 , disons inférieurs à 1000 véh/h, que

les valeurs de σ_H sur la VR se trouvent au dessus, tandis que les valeurs correspondantes de σ_H sur la VL par temps de pluie sont au dessous de la ligne hyperbolique. En général, l'écart-type σ_H est plus grand sur la VR que sur la VL, cette différence devenant évidemment négligeable lorsque le débit Q_6 augmente.

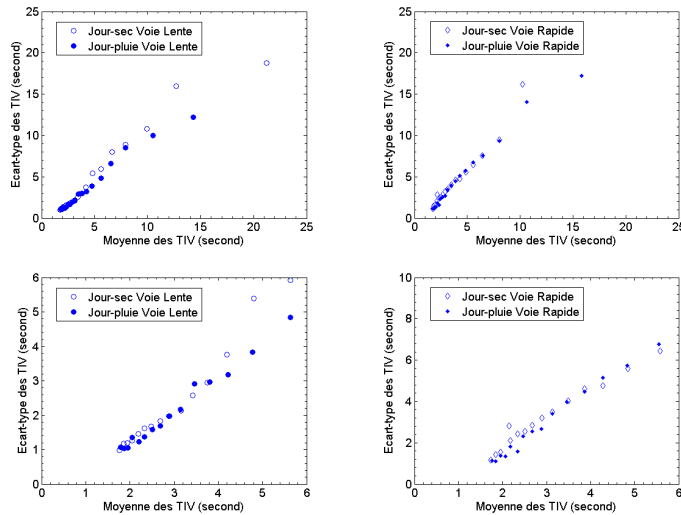


FIGURE 1.16 : Relation entre la moyenne et l'écart-type des TIV, du jour

Pour les échantillons de même débit élevé, la différence entre les écarts-types σ_H des deux conditions météorologiques semble être très faible d'autant plus que les valeurs de σ_H observées sont près de la ligne hyperbolique. Cette dernière remarque indique donc une relation étroite entre la moyenne μ_H et l'écart-type σ_H . La Figure 1.16 l'illustre nettement et de plus montre qu'aux mêmes valeurs de μ_H , l'écart-type σ_H est généralement plus grand dans les échantillons par temps sec que dans les échantillons par temps de pluie. Par ailleurs, la relation entre μ_H et σ_H est assez linéaire.

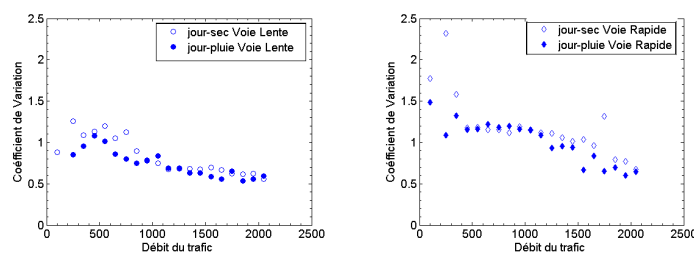


FIGURE 1.17 : CV des TIV du jour en fonction de Q_6

La relation entre μ_H et σ_H peut être aussi représentée par le coefficient CV. Pour les échantillons de même débit Q_6 , le CV par temps de pluie est généralement inférieur au CV par temps sec. De plus, pour les échantillons de même débit Q_6 , le CV sur la VL est significativement inférieur au CV sur la VR. Dans les échantillons de débit faible, le CV est généralement supérieur à 1. Lorsque le débit Q_6 augmente, le CV des échantillons diminue et cette diminution est assez régulière pour les échantillons sur la VL.

Sur la VL, le CV des échantillons devient inférieur à 1 lorsque le débit dépasse 800 véh/h. Le CV des échantillons sur la VR diminue plus lentement en fonction du débit Q_6 et reste supérieur à 1 même dans des cas où $Q_6 \approx 1400$ véh/h. La Figure 1.17 montre clairement la différence significative de CV entre les VR et VL.

Pourtant, la Figure 1.17 représente aussi des échantillons *particuliers* ayant des valeurs

de CV significativement différentes de celles des échantillons de classes de Q_6 voisines (même voie et même condition météorologique). Ils sont par exemple :

1. (0-200), (200-300) et (1700-1800) véh/h, jour-sec VR
2. (300-400), et (1600-1700) véh/h soir-pluie VR
3. (400-500) véh/h soir-sec VR.

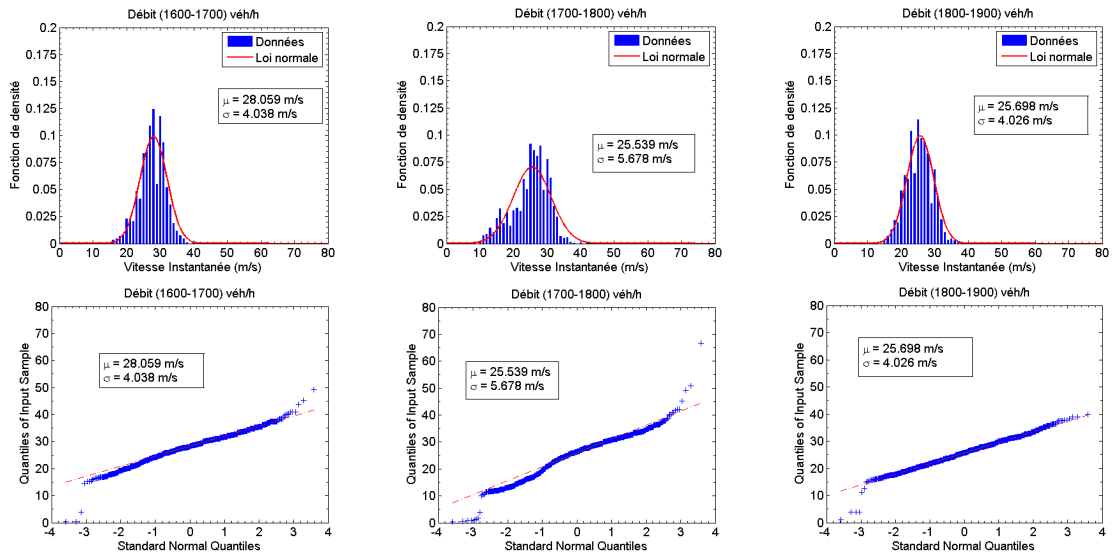


FIGURE 1.18 : Distribution des vitesses des TIV des échantillons ayant des valeurs de CV élevées

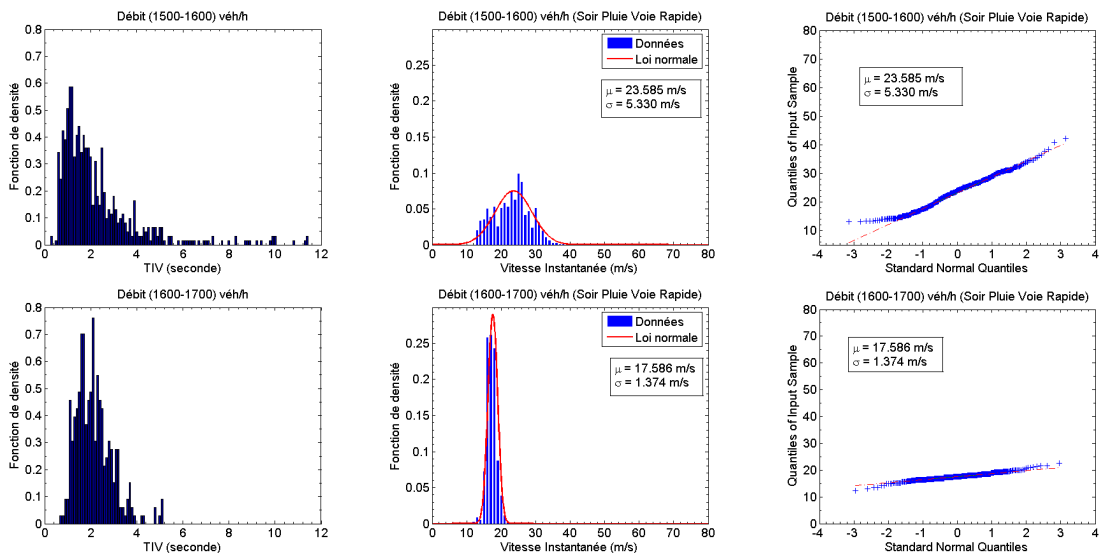


FIGURE 1.19 : Histogrammes des échantillons ayant des valeurs de CV élevées

Dans la Figure 1.18, la distribution des vitesses de l'échantillon *particulier* est au centre en vue de la comparer avec ses deux échantillons voisins. Dans la Figure 1.19, les distributions des vitesses et des TIV de l'échantillon particulier est au deuxième rang afin de les comparer avec celles de l'échantillon voisin. Les figures en bas de la Figure 1.18 et à droite de la Figure 1.19 illustrent le rapprochement des distributions des vitesses empiriques à la loi normale correspondante. En étudiant les échantillons particuliers ci-dessus, on constate que :

1. Les échantillons particuliers de CV élevé sont principalement observés sur la VR,
2. Pour les échantillons de débit faible, disons $Q_6 < 500$ véh/h, il n'y a pas de différence remarquable au niveau des *histogrammes* des TIV ou de la distribution des vitesses en comparant avec les autres échantillons voisins. Cela est dû au fait que les TIV dans des échantillons en trafic fluide peuvent être élevés et entraînent donc un fort écart-type σ et finalement un CV particulièrement élevé.
3. En débit élevé, disons $Q_6 > 1500$ véh/h, un échantillon ayant une valeur particulièrement élevée de CV possède une particularité au niveau de la distribution des vitesses. Cette dernière n'entraîne pas pourtant une particularité au niveau de la distribution des TIV. Il est possible que la valeur du CV des TIV soit proportionnelle à l'écart-type des vitesses instantanées σ_v de l'échantillon.

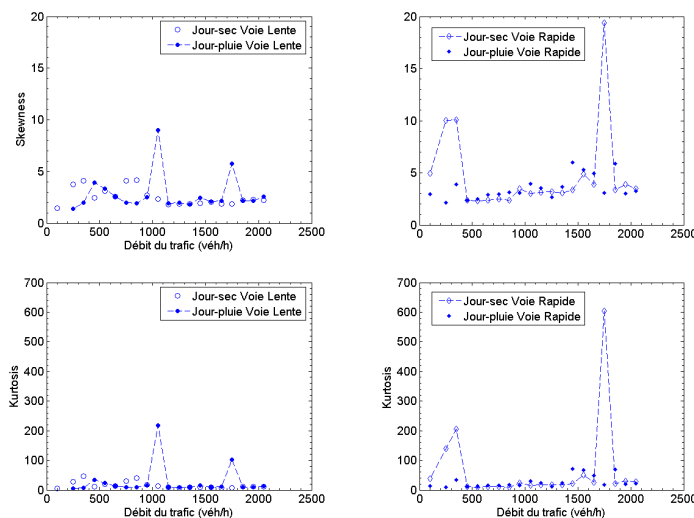


FIGURE 1.20 : Coefficient de symétrie S_3 et d'aplatissement K_4 des TIV du jour en fonction de Q_6

Le coefficient de symétrie S_3 est stable et inférieur à 5 dans la majorité des échantillons, tandis que le coefficient d'aplatissement K_4 est aussi stable (cf. Figure 1.20). Les deux coefficients sont corrélés comme la Figure 1.21 illustre la relation $K-S^2$. Cette corrélation est forte notamment pour les échantillons de K_4 inférieur à 40, et de S_3^2 inférieur à 15. Une fois encore, les différences sur les résultats empiriques dues aux conditions de météo sont moins marquées que celles dues aux types de voie. La Figure A.2 de l'Annexe confirme également ce point en montrant la différence de la relation $K-S^2$ entre les modalités **soir** sur VL et **soir** sur VR.

Comme dans le cas des valeurs de CV, il y a des échantillons particuliers ayant des valeurs de S_3 et K_4 particulièrement élevées par rapport à celles des échantillons voisins :

1. Le cas de S_3 : 1. E°1 (1000-1100) véh/h : jour-pluie VL ; 2. (1600-1700) véh/h : jour-pluie VL ; 3. (200-300) véh/h : jour-sec VR ; 4. (400-500) véh/h : jour-sec VR ; 5. E°2 (1700-1800) véh/h : jour-sec VR ; 6. E°3 (300-400) véh/h : soir-pluie VR ; 7. E°4 (1200-1300) véh/h : soir-pluie VR ; 8. E°5 (400-500) véh/h : soir-sec VR ;
2. Le cas de K_4 : 1. E°1 (1000-1100) véh/h : jour-pluie VL ; 2. E°2 (1700-1800) véh/h : jour-pluie VL ; 3. E°3 (300-400) véh/h : soir-pluie VR ; 4. E°4 (1200-1300) véh/h : soir-pluie VR ; 5. E°5 (400-500) véh/h : soir-sec VR ;

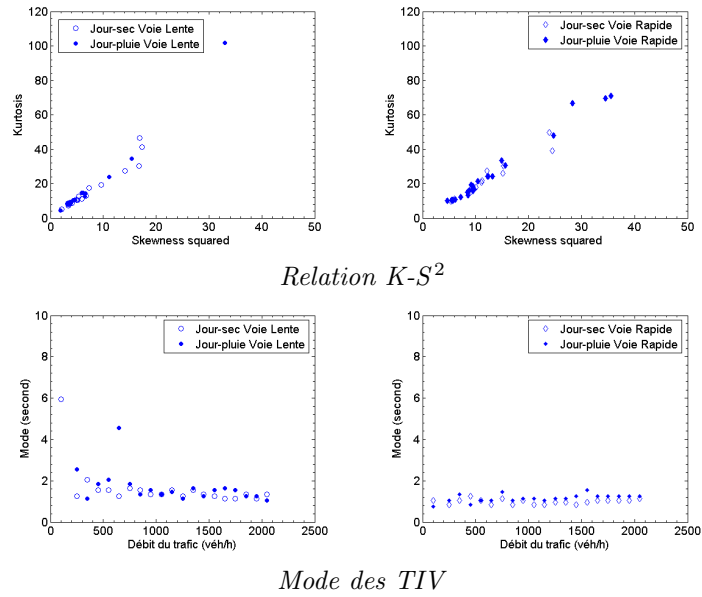


FIGURE 1.21 : Relation K-S² et le *mode* des TIV du jour en fonction de Q_6

La Figure 1.22 étudie les cinq échantillons notés de E°1 à E°5 où *avant* ou *après* signifient les échantillons ayant le débit Q_6 inférieur ou supérieur respectivement au débit de l'échantillon considéré. Les échantillons ayant les valeurs de K_4 particulièrement élevées coïncident avec les échantillons des valeurs S_3 particulièrement élevées. En analysant ces échantillons, on remarque que les distributions des TIV de ces échantillons comportent des valeurs extrêmement grandes en comparant avec les distributions des TIV des échantillons voisins. Ces résultats indiquent que la relation K-S² dépend forcément des grandes valeurs de TIV et que la méthode des moments semble très fragile en estimant les paramètres d'un modèle à plusieurs paramètres.

Le *mode* des distributions des TIV est assez stable dans tous les échantillons sur la VR et dans les échantillons de débit $Q_6 > 700$ véh/h sur la VL. Dans ces échantillons, le *mode* est inférieur à 2 s. La Figure 1.21 montre que sauf le jour sur la VL, à débit Q_6 équivalent, le mode par temps de pluie est plus grand que le mode par temps sec.

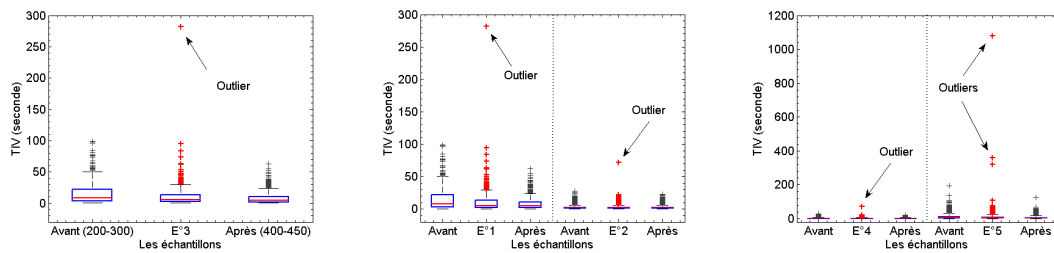


FIGURE 1.22 : Les échantillons particuliers et leurs échantillons voisins

2. En fonction du taux d'occupation 6 minutes \mathcal{T}_6

Pour les échantillons de taux d'occupation $\mathcal{T}_6 > 8\%$, la moyenne μ_H des TIV semble indépendante du taux d'occupation et des conditions de météo. Il n'y a pas un nombre suffisant de véhicules observés pour former les échantillons de $\mathcal{T}_6 > 18\%$ de la modalité *soir-pluie* sur la VL, de (20-22) et (22-24) % de la modalité *jour-pluie* sur la VR et de $\mathcal{T}_6 > 14\%$ de la modalité *soir-pluie* sur la VR. Les moyennes des TIV pour deux classes distinctes de taux d'occupation, ainsi que les moyennes des TIV pour deux conditions

météo distinctes ne sont significativement différentes que dans les échantillons de taux d'occupation correspondant à un trafic fluide \mathcal{T}_6 inférieur à 6%.

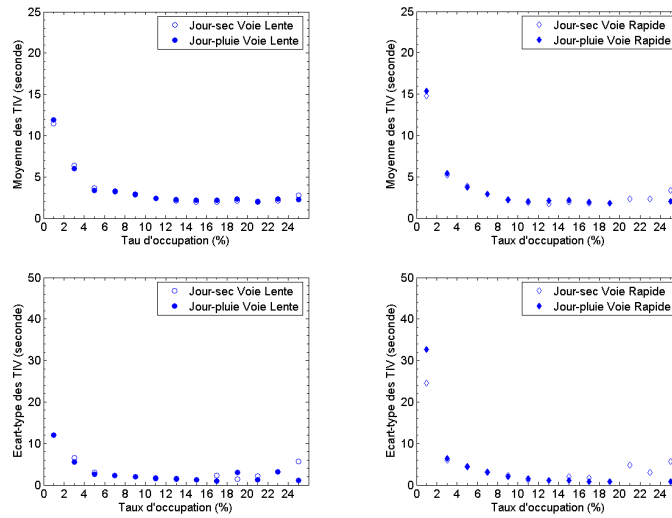


FIGURE 1.23 : Moyenne et écart-type des TIV du jour en fonction du taux d'occupation 6 minutes \mathcal{T}_6

Pour les échantillons de taux d'occupation $\mathcal{T}_6 < 18\%$, l'écart-type σ_H des TIV diminue en fonction de taux d'occupation \mathcal{T}_6 . Dans ce cas, la différence de l'écart-type entre les deux conditions de météo n'est remarquable que dans l'échantillon (0-2) % de la période jour sur la VR, et dans les échantillons de la période soir sur la VR. Pour les échantillons de taux d'occupation $\mathcal{T}_6 > 18\%$, l'écart-type σ_H n'est pas inférieur aux écart-types des échantillons de taux d'occupation plus faibles comme (14-16) et (16-18) %.

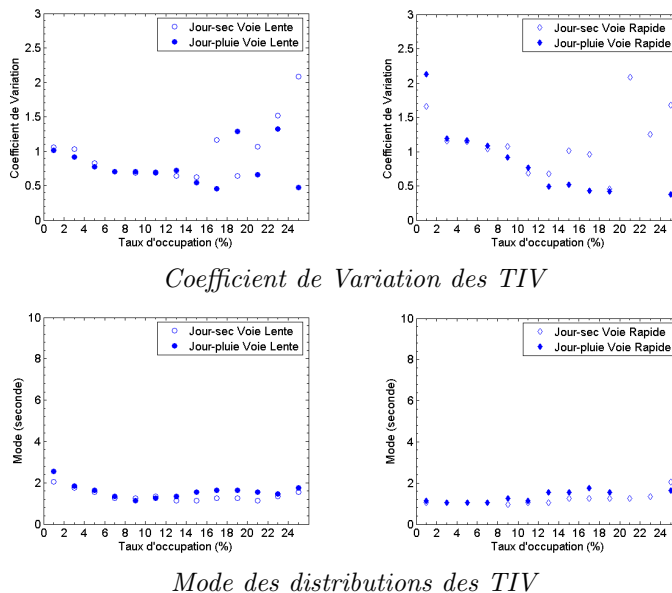


FIGURE 1.24 : CV et Mode des TIV du jour en fonction de \mathcal{T}_6

Le CV des TIV diminue en fonction du taux d'occupation \mathcal{T}_6 tant que celui-ci est inférieur à un seuil, 16 % pour la modalité jour-sec sur la VL, 18 % pour la modalité jour-pluie sur la VL, 14 % pour la modalité soir-sec sur la VL, 18 % pour la modalité soir-pluie sur la VL, 16 % pour la modalité jour-sec sur la VR, 22 % pour la modalité

jour-pluie sur la VR et 14 % pour la modalité jour-sec sur la VR. Pour les échantillons de taux d'occupation supérieur à ces seuils par temps *sec* (ou par temps de pluie, le jour seulement, sur la VL), le CV fluctue et prend de plus grandes valeurs. Ses valeurs pour les échantillons du *soir* sur la VR sont les plus dispersées (cf. Figure A.3 et l'Annexe).

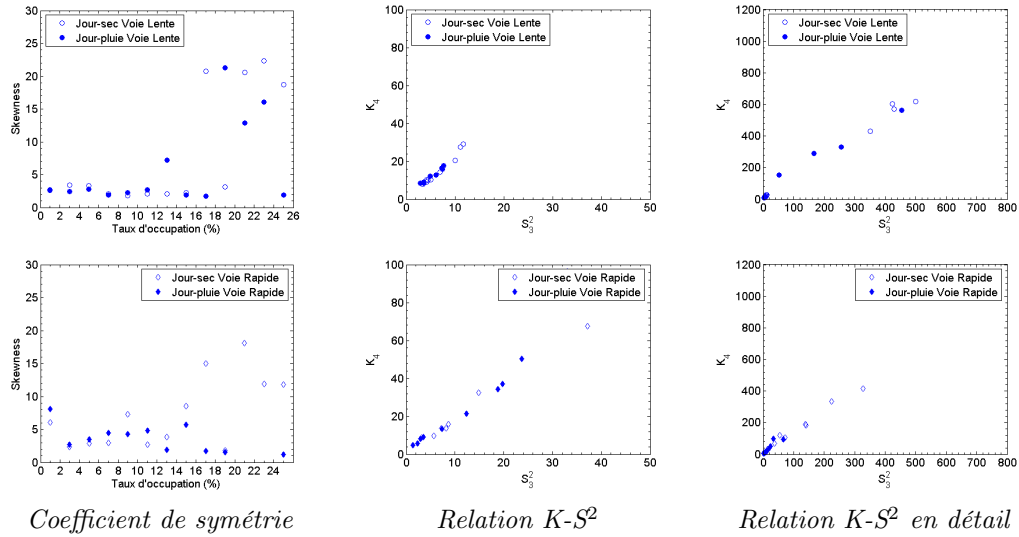


FIGURE 1.25 : S_3 et Relation $K-S^2$ des TIV du jour, en fonction de \mathcal{T}_6

Au contraire de cette forte variation des CV, le *mode* des distributions des TIV dans tous les échantillons reste stable. Pour les échantillons de même taux d'occupation \mathcal{T}_6 , le *mode* des TIV par temps de pluie est légèrement plus élevé que le mode des TIV par temps *sec*.

La Figure 1.25 illustre la variation du coefficient de symétrie S_3 en fonction de taux d'occupation \mathcal{T}_6 et représente la relation $K-S^2$. Sur la VL, le coefficient S_3 est stable dans les échantillons de taux d'occupation \mathcal{T}_6 inférieur à 18 %. Pourtant, ce coefficient dans les échantillons de taux d'occupation \mathcal{T}_6 supérieur à 18 % sur la VL est particulièrement élevé. Le coefficient S_3 sur la VR varie fortement selon les modalités.

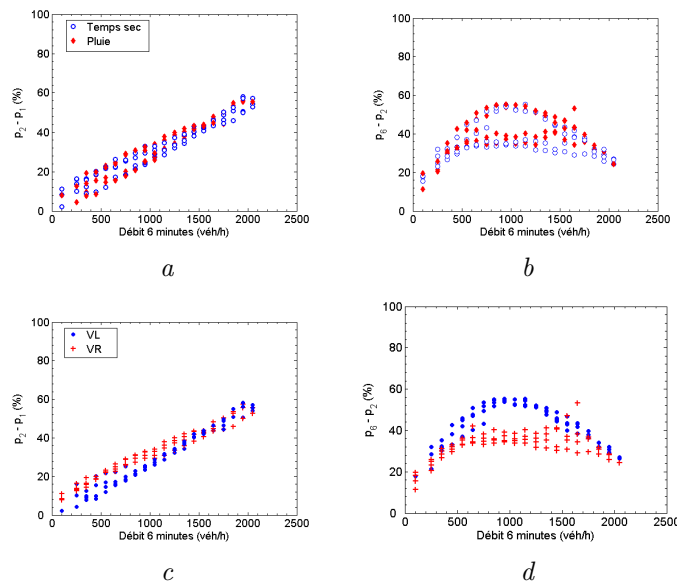


FIGURE 1.26 : Différences des pourcentages des TIV courts en fonction de Q_6

Il y a des échantillons ayant les valeurs empiriques des S_3 et K_4 très élevés. Cependant, quel que soit le niveau de S_3^2 ou K_4 , la relation entre le carré de S_3 et le coefficient K_4 est toujours significative (cf. la colonne au milieu de la Figure 1.25).

3. TIV courts

La Figure 1.27 montre qu'en général, les pourcentages de TIV courts (p_1 (TIV < 1 s), p_2 (TIV < 2 s), et p_6 (TIV < 6 s)) augmentent en fonction du débit Q_6 . L'augmentation de p_2 est assez constante dans tous les cas, tandis que p_6 croît fortement pour les $Q_6 \leq 1000$ véh/h et plus légèrement pour $Q_6 > 1000$ véh/h. Le pourcentage p_1 ne dépasse pas 20 % sauf pour les échantillons de $Q_6 > 1100$ véh/h sur la VR, et p_2 ne dépasse pas 80 %.

Au même débit Q_6 , les pourcentages p_1 et p_2 par temps sec sont plus grands que par temps de pluie. De plus, ces écarts de p_1 selon les conditions météorologiques sont plus grands le jour et aux débits Q_6 supérieurs à 1100 véh/h. La différence ($p_2 - p_1$) augmente selon le débit Q_6 mais cette différence selon les conditions météorologiques n'est pas évidente. Sauf le jour sur la VL aux débits $Q_6 < 1000$ véh/h et pour quelques échantillons, la plupart des échantillons par temps de pluie ont un pourcentage p_6 plus grand que par temps sec (au même débit Q_6). De plus, la différence ($p_6 - p_2$) varie en fonction du débit Q_6 comme le montre la Figure 1.26-b. On observe un petit décalage de l'ensemble des points dans la condition de pluie vers la droite aux points par temps sec.

Au même débit Q_6 , les pourcentages p_1 et p_2 sur la VR sont nettement plus grands que sur la VL. De plus, ces écarts de p_1 et p_2 entre les voies sont plus importants pour les débits Q_6 inférieurs à 1500 véh/h. Sur la VR, la différence ($p_2 - p_1$) augmente en fonction du débit Q_6 et est plus grande dans les échantillons de débit $Q_6 < 1500$ véh/h. Le pourcentage p_6 sur la VL est plus faible que sur la VR pour les débits $Q_6 < 700$ véh/h et plus grand que sur la VR pour les débits $Q_6 > 1000$ véh/h. La Figure 1.26-d illustre la différence ($p_6 - p_2$) en fonction du débit Q_6 selon les voies.

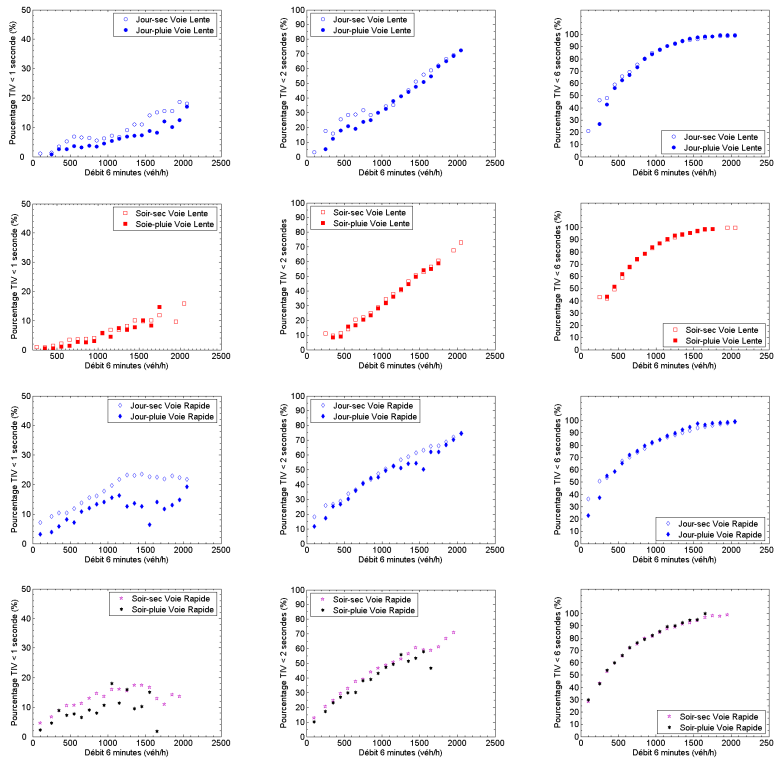


FIGURE 1.27 : Les pourcentages des TIV courts (TIV < 1 s, 2 s et 6 s)

Analyse statistique des échantillons raffinés de la RN118

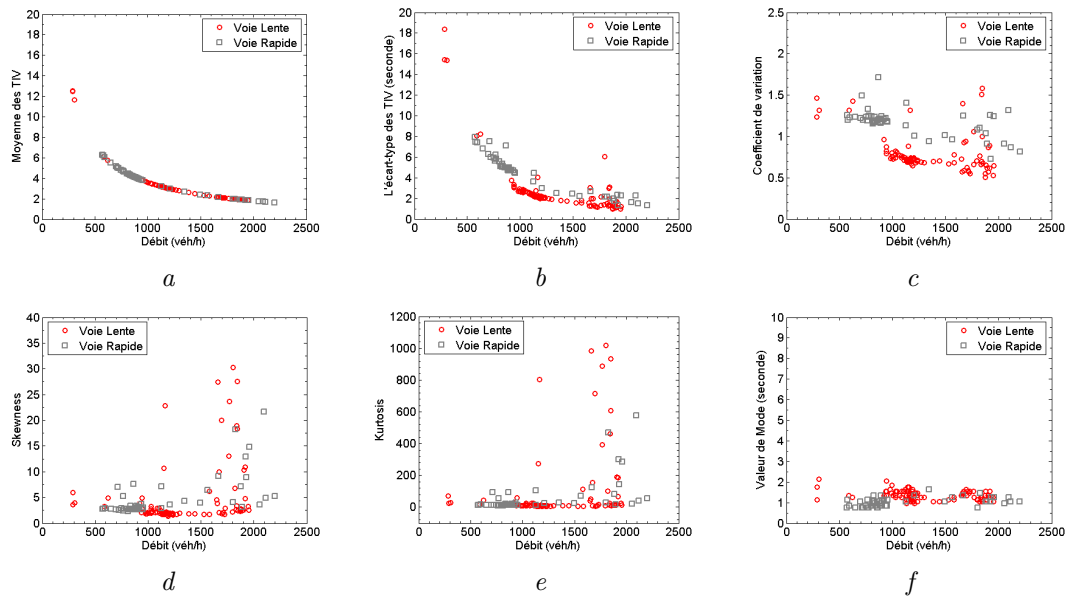


FIGURE 1.28 : Les statistiques descriptives des TIV des échantillons raffinés

La Figure 1.28 représente les statistiques de base des échantillons sur les VL et VR en fonction du débit Q_r calculé sur la période d'échantillonnage. En comparant avec les statistiques descriptives des données de l'autoroute A6, on retrouve en général, les mêmes caractéristiques des VL et VR.

Tout d'abord, la moyenne des TIV dans les échantillons raffinés est exactement l'inverse du débit Q_r des échantillons correspondants. Cette relation est représentée par une courbe hyperbolique dans la Figure 1.26-a, et elle est identique sur les deux voies. L'écart-type des TIV des échantillons raffinés diminue en fonction du débit comme illustré dans la Figure 1.26-b. Au même niveau du débit Q_r qui n'est pas très fluide, l'écart-type des TIV des échantillons raffinés sur la VL est inférieur à celui sur la VR. En conséquence, le coefficient de variation (CV) des échantillons sur la VL est aussi inférieur à celui sur la VR. De plus, la Figure 1.26-c montre que la plupart des échantillons sur la VL ont un CV inférieur à 1 alors que les CV de la plupart des échantillons sur la VR sont supérieurs à 1.

Les formes des nuages de points représentant les coefficients de symétrie S_3 et d'aplatissement K_4 sont assez semblables. Le coefficient S_3 est toujours positif et les deux coefficients sont généralement stables en fonction du débit Q_r . Il semble qu'un échantillon ayant une haute valeur de S_3 possède aussi une haute valeur de K_4 . Ces deux coefficients de la distribution des TIV sont fortement liés (cf. la relation K-S² dans la Figure 1.2).

Les échantillons de hautes valeurs de S_3 et K_4 incluent normalement des valeurs particulières (ou des *outliers*). Cependant, ces valeurs ont peu d'influence sur la distribution globale des TIV. Le *mode* reste stable quelque soit l'échantillon. Les valeurs obtenues du *mode* des TIV se trouvent entre 0,5 à moins de 2 s.

Analyse spectrale des échantillons raffinés

La Figure 1.29 représente les valeurs calculées des coefficient d'auto-corrélation des TIV dans les échantillons raffinés sur les voies VL et VR de la RN118. Dans la majorité des échantillons, les coefficients ρ_i sont positifs sur les deux voies, particulièrement pour les ρ_1 et ρ_2 sur la VR. Les valeurs des ρ_i sont généralement comprises entre -0,05 et 0,1 (cf.

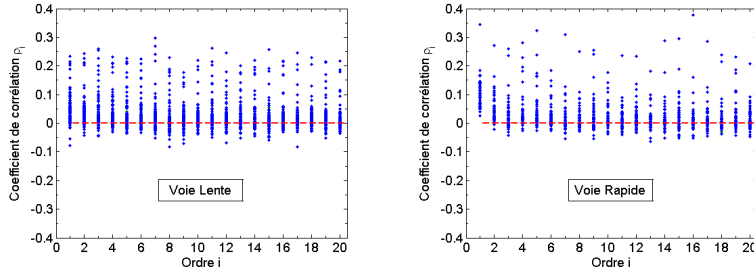


FIGURE 1.29 : Coefficient d'auto-corrélation d'ordre i sur les VL et VR

Figure 1.30). En comparant avec les données de l'A6, les valeurs des ρ_i sur la RN118 sont beaucoup plus grandes.

La Figure 1.30 fournit les valeurs calculées des estimateurs Δ et Δ_θ . Sauf pour des échantillons particuliers ayant des valeurs Δ et Δ_θ singulières, la plupart des échantillons ont les mêmes ordres de valeurs Δ et Δ_θ , que celles des échantillons de l'A6, c.à.d respectivement d'environ 1 et 1,5. Pourtant, la différence entre les deux voies ne peut pas être clairement remarquée dans la Figure 1.30.

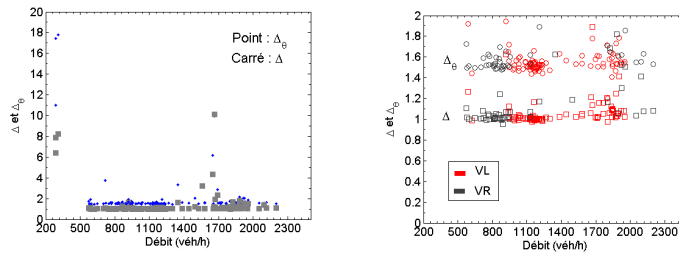


FIGURE 1.30 : Valeurs des estimateurs de la somme des carrés des ρ_i

Les valeurs estimées des densités spectrales \hat{f}_k avec $k = 0, 1$, sont fournies dans la Figure 1.31 ci-dessous. Comme dans le cas des données d'autoroute A6, \hat{f}_0 est calculée en utilisant $\epsilon = 6$ et \hat{f}_1 en utilisant $l = 100$. Par ailleurs, les valeurs estimées de f_1 et f_2 sont très proches l'une de l'autre.

La Figure 1.31 représente les valeurs estimées de f_0 selon deux échelles, l'une pour la comparaison avec la Figure 1.10 précédente de l'A6 et l'autre pour illustrer aussi les valeurs de f_0 estimée des échantillons particuliers. La forme des courbes est apparemment similaire entre les échantillons sur la VR de l'A6 et de la RN118.

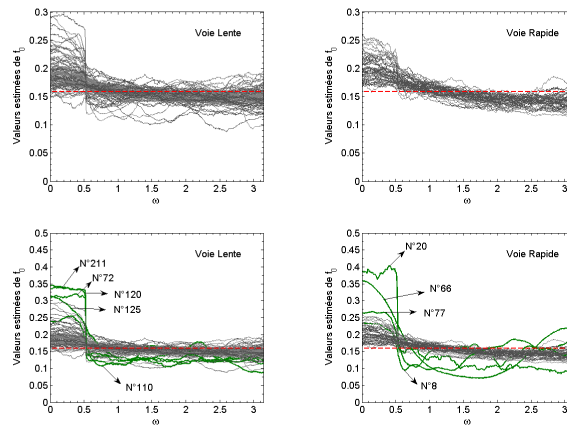


FIGURE 1.31 : Valeurs estimées de la densité spectrale f_0

La Figure 1.32 représente aussi les valeurs estimées de f_1 en deux échelles, l'une pour comparer avec la Figure 1.10, et la seconde pour montrer qu'il y a des échantillons ayant des valeurs estimées de f_1 très grandes (de l'ordre de 4).

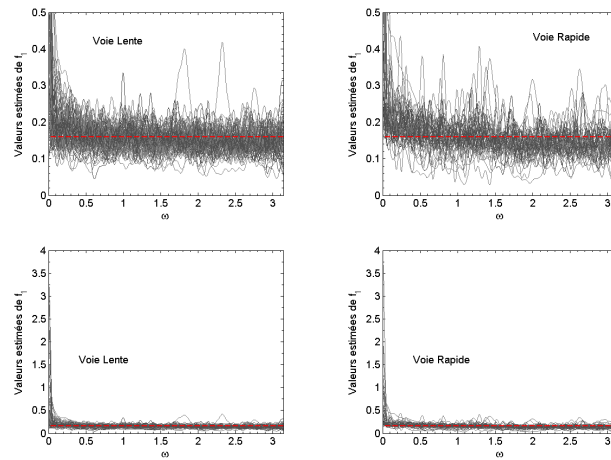


FIGURE 1.32 : Valeurs estimées de la densité spectrale f_1

Les *périodogrammes* des TIV sont représentés dans la Figure 1.33. Dans les cas non particuliers, les valeurs de la statistique $2I_n(\omega_p)/s^2$ fluctuent autour d'une valeur fixe de la valeur $1/\pi$ d'un processus de renouvellement. La Figure 1.34-a représente donc les moyennes des $2I_n(\omega_p)/s^2$ tandis que la Figure 1.34-b représente la distance de ces moyennes par rapport à la valeur de renouvellement $1/\pi$.

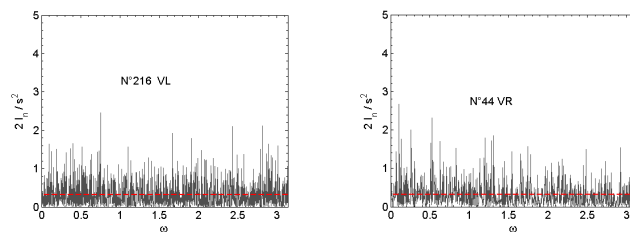


FIGURE 1.33 : Le périodogramme

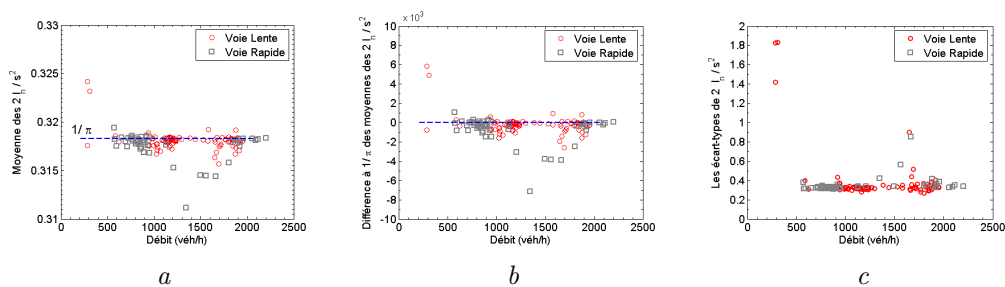


FIGURE 1.34 : Moyenne, Distance à $1/\pi$ des moyennes et écart-type des $2I_n(\omega_p)/s^2$

La Figure 1.34-c montre que les écarts-types des $2I_n(\omega_p)/s^2$ des échantillons des TIV sont très stables, de l'ordre de 0.3. Certains échantillons particuliers ayant un écart-type plus grand ou beaucoup plus élevé ont été trouvés, leurs périodogrammes sont très loin d'être celui d'un processus renouvellement. Quelques échantillons particuliers du point de vue de l'analyse spectrale sont représentés sur la Figure 1.35.

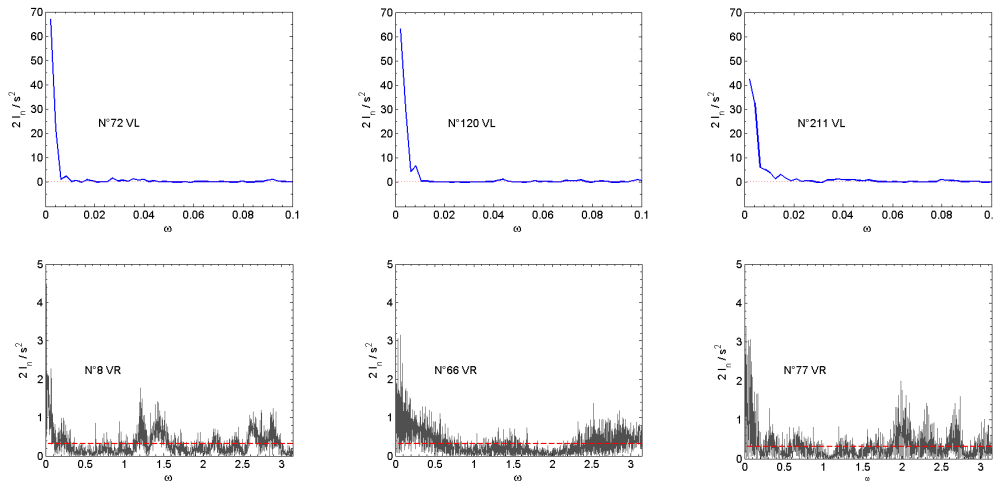


FIGURE 1.35 : Les échantillons particuliers du point de vue d'analyse spectrale

1.5 Conclusion du chapitre

Dans ce chapitre, deux bases de données de l'A6 et de la RN118 sont analysées. Elles sont tout d'abord échantillonnées par deux méthodes d'échantillonnage : la méthode de groupement et la méthode de raffinement. Les caractéristiques statistiques sont ensuite comparées entre les échantillons représentant différentes modalités dont la condition météo, le type de voie et la période de la journée.

Les caractéristiques statistiques sont aussi analysées en fonction de la demande de trafic (débit agrégé, taux d'occupation agrégé, débit horaire pour la méthode de groupement, et débit de l'échantillon pour la méthode de raffinement). Les résultats montrent les caractéristiques typiques de l'écart-type, du coefficient de variation des TIV tandis que les coefficients de symétrie et d'aplatissement sont particulièrement sensibles à des valeurs extrêmes de TIV. Bien que la relation forte entre ces deux derniers coefficients soit une des caractéristiques typiques de la distribution des TIV et sert à choisir préalablement certains modèles, la sensibilité de ces deux coefficients souligne la fragilité de la méthode des moments lorsque celle-ci est utilisée pour estimer les modèles probabilistes à plusieurs paramètres.

De plus, au même niveau de débit et de taux d'occupation 6 minutes, les effets des facteurs exogènes sont observés notamment sur l'écart-type et le coefficient de variation (CV) des TIV dans les échantillons groupés de la RN118. La tendance et le niveau d'effet dépendent du type de facteurs (soit de la condition météorologique, soit de la voie). L'effet de la voie est confirmé pour les échantillons raffinés de la RN118 et les échantillons horaires de l'A6. Dans ces derniers, les caractéristiques statistiques des TIV sur la voie lente (VL) sont similaires à celles des TIV de la voie médiane (VM).

L'analyse spectrale est utilisée pour les échantillons de TIV consécutifs (les échantillons raffinés de la RN118 et les échantillons horaires de l'A6). Les résultats montrent que cette analyse fine fournit des connaissances supplémentaires sur les caractéristiques des échantillons, voire des réseaux. Elle met en évidence la différence entre les voies et semble être capable de détecter des comportements différents à cause de facteurs exogènes (la réduction de largeur de voie par exemple) et de reconnaître les échantillons qui sont loin de correspondre à un processus de renouvellement. Pour les échantillons horaires, les tests de renouvellement et de groupement montrent l'existence en général des groupements de véhicules et les processus d'arrivée ne sont pas *i. i. d* dans la majorité des cas, en particulier sur la VL.

Chapitre 2

Modélisation probabiliste du TIV

Un modèle probabiliste a essentiellement pour but de contribuer à l'analyse, voire de prédire la fréquence d'apparition d'une v. a. C'est le cas des modèles du TIV qui est une v. a réelle continue non négative. Pour cela, deux approches sont possibles. La première se base sur des informations initiales du trafic et des hypothèses de comportement des conducteurs. Cette approche est donc souvent de caractère stochastique et les modèles sont généralement implicites sauf dans des cas simplifiés. La deuxième approche est au contraire plus directe. Sachant que si les lois physiques gouvernent les fonctionnements des phénomènes, les lois de probabilité ne sont révélées qu'*a posteriori* à partir des observations. C'est la raison pour laquelle, cette deuxième approche probabiliste met l'accent sur les données réelles enregistrées.

Dans ce chapitre, les deux approches sont abordées en se concentrant plus particulièrement sur l'approche probabiliste directe.

2.1 Approches stochastiques

2.1.1 Approche de Heidemann

Le modèle

Heidemann [68, 69, 70] a établi un modèle des TIV basé sur des connaissances initiales du trafic. L'idée est de déterminer la distribution des TIV à partir des trois informations suivantes :

1. La distribution des vitesses des véhicules à une concentration du trafic très faible (\mathcal{H}_1).
2. Une fonction modélisant la probabilité de dépassement (\mathcal{H}_2).
3. La concentration du trafic en congestion (\mathcal{H}_3).

Le but est de trouver la distribution des TIV en fonction de la concentration du trafic. Pour cela, un système d'équations différentielles issu des hypothèses sur le changement d'état du trafic doit être établi. La notion d'*état du trafic* est tout d'abord déterminée par deux informations fondamentales : la concentration et la classe de vitesse à laquelle la vitesse des véhicules appartient. Les vitesses sont classées dans n groupes déterminés par leurs seuils $\mathcal{V}_0, \mathcal{V}_1, \dots, \mathcal{V}_n$; un véhicule appartient à la classe i si sa vitesse v est telle que $\mathcal{V}_{i-1} \leq v < \mathcal{V}_i$, où $\mathcal{V}_0 = 0$ et $\mathcal{V}_n = \infty$. À un instant donné, l'état d'un véhicule est caractérisé par la classe de vitesse i et la concentration K du flux au sein duquel le véhicule se déplace, et représenté par le couple $\eta = (K, i)$. Finalement, l'état du trafic n'est autre que l'ensemble des états des véhicules individuels.

La distribution de la vitesse en fonction de la concentration doit être tout d'abord obtenue. Pour cela, la probabilité des changements d'état des véhicules est nécessaire afin d'induire la distribution des vitesses à partir de la condition (\mathcal{H}_1) à une concentration quelconque :

$$\mathcal{P}_{K+dK}(j) = \sum_{i \geq j} \mathcal{P}_K(i) \times \mathcal{P}_{K,K+dK}(i,j) \quad (2.1)$$

Où :

$$\Pr \{ \eta = (K, i) \} = \mathcal{P}_K(i) \quad (2.2)$$

$$\mathcal{P}_{K,K+dK}(i,j) = \Pr \{ (K, i) \longrightarrow (K + dK, j) \} \quad (2.3)$$

La probabilité de passage de la vitesse i à la vitesse j lorsque la concentration passe de K à $K + dK$ dépend de l'état actuel du véhicule, donc $\mathcal{P}_K(i)$, et aussi d'autres fonctions concernant le dépassement et la proximité du véhicule en tête. Heidemann suppose que cette probabilité de *transmission* ne se présente que dans le cas où le véhicule précédent se déplace dans un groupe de vitesse j plus lente que dans le groupe i du véhicule considéré :

$$\mathcal{P}_{K,K+dK}(i,j) = \begin{cases} 0 & \text{si } i < j \\ \mathcal{P}_K(j) \times \bar{h}'(K,i) dK & \text{si } i > j \\ 1 - \sum_{i > r} \mathcal{P}_K(r) \bar{h}'(K,i) dK & \text{si } i = j \end{cases} \quad (2.4)$$

La formule dans le cas $(i = j)$ entraîne bien que $\sum_j \mathcal{P}_{K,K+dK}(i,j) = 1$. La fonction de Heidemann $\bar{h}(K,i)$ est supposée avoir la forme suivante :

$$\bar{h}(K,i) = \frac{(1 - \mathbb{E}[V_1 | v \in \mathcal{V}_i]) g(K,i)}{1 - p(k,i)} \quad (2.5)$$

Où

$p(k,i) = (1 - \mathbb{E}[V_2 | v \in \mathcal{V}_i]) \times [\mathcal{P}_K(i)(1 - g(K)) + g(K)]$ signifie la probabilité qu'un véhicule en état (K,i) est suiveur dans un *peloton*.

$V_1 \sim \exp[-v t_a K_*]$ et $V_2 \sim \exp[-v t_p K_*]$

$K_*^{-1} = K^{-1} - K_{\max}^{-1}$, K_*^{-1} : la différence des inverses des concentrations courante et maximum K_{\max} . Ces inverses correspondent aux moyennes des DIV *nettes* entre l'avant d'un véhicule et l'arrière du véhicule précédent.

$t_a = t_a(K,i)$: le seuil de TIV en dessous duquel le véhicule considéré va adapter (diminuer) sa vitesse à celle du véhicule plus lent précédent.

$t_p = t_p(K,i)$: le seuil de TIV en dessous duquel on considère le véhicule en *peloton*. Les véhicules dans ce peloton sont supposés se déplacer à la même vitesse v .

$g(K,i)$: est une fonction de *dépassement*, prédéfinie dans la condition \mathcal{H}_2 .

La relation (2.1) réagit comme un processus de Markov et la solution de ce système stochastique est obtenue numériquement pour les $\mathcal{P}_K(j)$ comme suit :

$$\frac{\mathcal{P}_K(j)}{A(j)} = \frac{1}{\left(\frac{A(j)}{\mathcal{P}_0(j)} - 1 \right) e^{\bar{h}(K,j)} + 1} \quad (2.6)$$

Où

$$A(j) = 1 - \sum_{i > j} \mathcal{P}_K(j) \quad (2.7)$$

$$\tilde{h}(K, j) = \frac{(1 - \mathbb{E}[V_1 | v \in \mathcal{V}_j]) g(K, j)}{1 - (1 - \mathbb{E}[V_2 | v \in \mathcal{V}_j]) \left[\frac{\mathcal{P}_K(j)}{A(j)} (1 - g(K, j)) + g(K, j) \right]} \quad (2.8)$$

Les calculs sont exécutés dans l'ordre de $\mathcal{P}_K(n)$ jusqu'à $\mathcal{P}_K(0)$ en utilisant successivement $A(j)$ dans l'Équation 2.7. La distribution de $\mathcal{P}_K(n)$ est le résultat d'une équation différentielle :

$$\mathcal{P}_K(n) = \frac{1}{\left(\frac{1}{\mathcal{P}_0(n)} - 1 \right) e^{h(K, n)} + 1} \quad (2.9)$$

où les valeurs $\mathcal{P}_0(j)$ sont connues par l'hypothèse \mathcal{H}_1 .

Jusqu'ici, les calculs permettent en principe d'accéder aux probabilités $\mathcal{P}_K(i)$ des vitesses en fonction de la densité K . Pour obtenir les distributions des TIV selon la concentration, le même processus de déduction est utilisé en ajoutant une condition sur la DIV nette K_*^{-1} , il s'agit que :

$$\mathcal{P}_{s, K}(i) = \Pr \{ \eta = (K, i) | K_*^{-1} \leq s \} \quad (2.10)$$

Par la loi marginale de probabilité, on a :

$$\mathcal{P}_{s, K}(i) = \sum_{j=1}^n \mathcal{P}_K(j) \times \mathcal{P}_{s, K}(i | j) \quad (2.11)$$

où $\mathcal{P}_{s, K}(i | j)$ est la probabilité qu'un véhicule se déplace en classe de vitesse \mathcal{V}_i à concentration K , à condition que sa DIV nette K_*^{-1} soit inférieure à s et que le véhicule précédent roule en classe de vitesse \mathcal{V}_j . Cette probabilité est obtenue par les formules suivantes :

$$\mathcal{P}_{s, K}(i | j) = \begin{cases} \mathcal{P}_0(i) & \text{si } i < j \\ \mathcal{P}_0(j) e^{h_s(K, i)} & \text{si } i > j \\ 1 - \sum_{r < j} \mathcal{P}_0(r) - \sum_{r > j} \mathcal{P}_0(r) e^{h_s(K, r)} & \text{si } i = j \end{cases} \quad (2.12)$$

Avec

$$\tilde{h}_s(K, i) = \frac{(1 - \mathbb{E}[V_3 | v \in \mathcal{V}_i]) g(K, i)}{1 - e^{-sK_*}} + \quad (2.13)$$

$$+ g(K, i) [1 - \mathbb{E}[V_1 | v \in \mathcal{V}_i]] \times \frac{p_s(K, i)}{1 - p(K, i)} \quad (2.14)$$

Où :

$$V_3 = \min \{ v t_a K_*, s K_* \} \quad (2.15)$$

$$p_s(K, i) = \frac{1 - \mathbb{E}[V_4 | v \in \mathcal{V}_i]}{1 - e^{-sK_*}} \times [\mathcal{P}_K(i) (1 - g(K)) + g(K)] \quad (2.16)$$

Il est à noter que $p_s(K, i)$ n'est autre que $p(K, i)$ en ajoutant la condition $K_*^{-1} \leq s$.

$$V_4 = \min \{ v t_p K_*, s K_* \} \quad (2.17)$$

En utilisant la probabilité conditionnelle, les probabilités $\mathcal{P}_{s,K}(i)$ peuvent être calculées. Finalement, la distribution des TIV en fonction de la concentration K est obtenue comme suit :

$$\Pr \{H \leq h\} = \int_0^\infty \mathcal{P}_K \left(v \geq \frac{s + 1/K_{\max}}{h} \mid s \leq s(i, h) \right) dF_k(s) \quad (2.18)$$

$$= \sum_{i=1}^{n_0} \mathcal{P}_{s(i,h),K}(i) \times \left[1 - e^{-s(i,h)K_*} \right] \quad (2.19)$$

Avec :

$$s(i, h) = \mathbb{E} [V \mid v \in \mathcal{V}_i] h - 1/K_{\max} \quad (2.20)$$

Le passage entre l'Équation (2.18) et l'Équation (2.19) est effectué sous l'hypothèse que la distribution des DIV nettes K_*^{-1} suit une loi exponentielle.

Résultats

La modélisation est réalisée en utilisant les paramètres suivants :

- La moyenne μ_v et l'écart type σ_v des vitesses de l'échantillon ayant le taux d'occupation (0-2) % sont calculés. Ensuite, la loi normale $\mathcal{N}(\mu_v, \sigma_v)$ est appliquée pour la distribution des vitesses à concentration très faible selon la condition \mathcal{H}_1 .
- Pour comparer avec les résultats de Heidemann, les seuils $t_a, t_p, g(k)$ sont pris comme dans son article [69], soit $t_a = 2$ s pour $0 < \mathcal{V}_0 \leq 5$ m/s et $t_a = 8$ s pour $140 < \mathcal{V}_n < \infty$ m/s, t_a est proportionnel à la valeur de classe de vitesse; $t_p = 2$ s est constant, et $g(k, i) = g(k) = (K/K_j)^{2/3}$.
- Sachant que $K = \frac{10\tau}{l+L}$ (véh/km) où l est la longueur de la boucle magnétique et L la longueur moyenne des véhicules, on prend un incrément de la concentration K de 4 véh/km correspondant à l'incrément du taux d'occupation de 2 %. Ici $l = 1$ m et $L = 4$ m. En conséquence, la concentration en congestion est $K_{\max} = 48$ véh/km.

La Figure 2.1 montre la différence entre la distribution empirique et la distribution prédite par Heidemann. Du plus bas au plus haut sont représentées les courbes de répartition des TIV des échantillons de taux d'occupation allant de (0-2) % à (> 24 %) sur la base de données de la RN118.

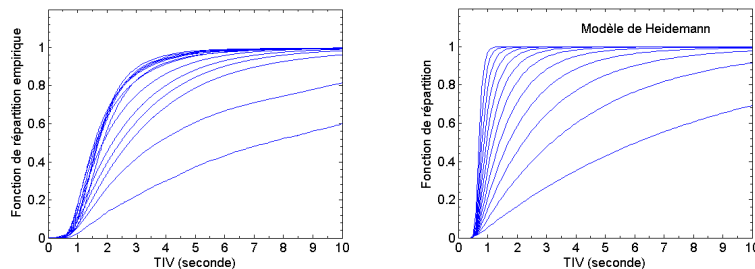


FIGURE 2.1 : Les distributions des TIV empiriques et le modèle de Heidemann

On remarque que le modèle de Heidemann ne représente pas bien les fréquences des TIV courts. De plus, l'approximation et/ou la superposition des courbes en congestion ne sont pas retrouvées par le modèle. Par ailleurs, le modèle surestime les probabilités des TIV courts aux trafics denses.

La Figure 2.2 ci-dessous est obtenue en modifiant la distribution des vitesses à concentration très faible dans la condition initiale \mathcal{H}_1 . Les distributions des résultats semblent ne pas changer beaucoup, ce qui signifie que le modèle n'est pas suffisamment sensible à cette condition initiale.

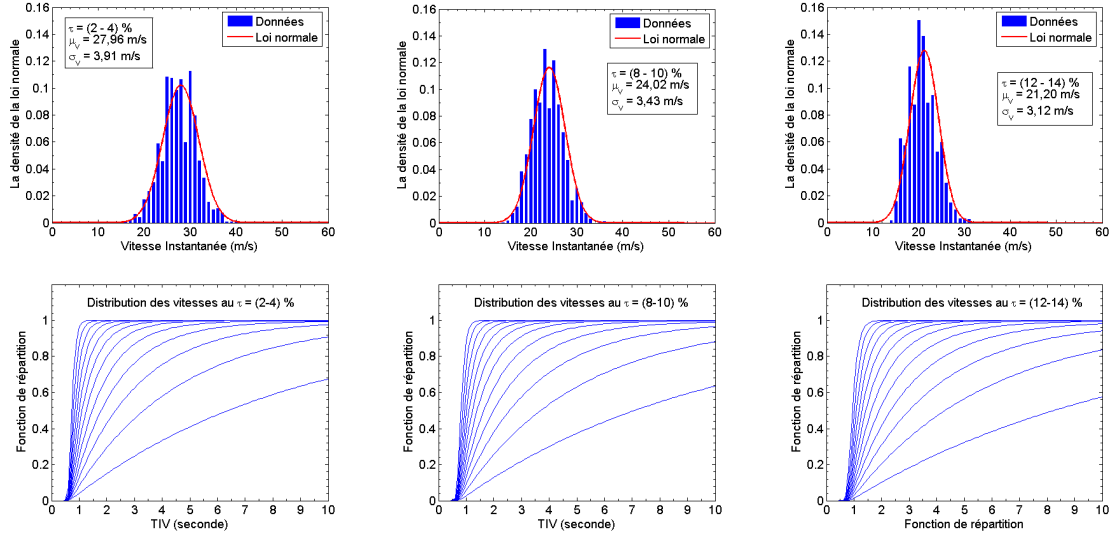


FIGURE 2.2 : Les distributions des TIV empiriques et le modèle de Heidemann avec différentes entrées de \mathcal{P}_0

En modifiant le temps t_p ou en prenant t_a constant, les résultats changent également peu. Aussi, le changement de la fonction de dépassement $g(k)$ ne donne pas non plus de changements significatifs des distributions résultats des TIV, comme le montré la Figure 2.3. Il semble donc que le modèle de Heidemann n'est pas flexible.

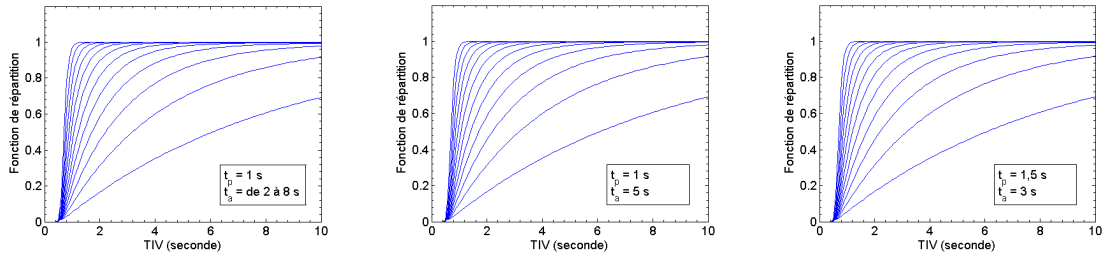


FIGURE 2.3 : Les distributions des TIV empiriques et le modèle de Heidemann avec différentes valeurs de t_a, t_p

Discussion

On remarque sur le modèle de Heidemann que :

1. L'équation 2.4 dans le modèle de Heidemann exclut qu'un véhicule de vitesse i plus faible rattrape son précédent de vitesse j plus élevée dans la phase de *transmission*, bien que cette manoeuvre conduise aussi à une augmentation de concentration du trafic $K \rightarrow K + dK$. En effet, on a $\mathcal{P}_{K, K+dK}(i, j) = 0$ si $i < j$. Cette hypothèse semble entraîner que l'accélération (positive) n'est pas envisagée dans le modèle et que le trafic devient de plus en plus dense par la réduction de vitesse et décélération. Cela pourrait expliquer la raison pour laquelle les fréquences des TIV courts en trafic dense sont trop élevées.

- Généralement, la distribution des TIV est calibrée en fonction du débit Q car il y a un lien direct entre la moyenne des TIV et le débit. L'idée de Heidemann est d'utiliser plus d'informations fondamentales macroscopiques en utilisant K et V pour établir la distribution des TIV, d'autant plus que seules les conditions initiales $\mathcal{H}_1, \mathcal{H}_3, \mathcal{H}_3$ sont utilisées. Il est à noter que la relation fondamentale du trafic est $Q = K \times V$. Cependant, le concept de concentration K utilisé dans le modèle a un caractère microscopique plutôt que macroscopique car elle est calculée directement par l'inverse de l'interdistance (la DIV nette) de chaque véhicule. En outre, la vitesse utilisée dans le modèle n'est pas non plus la moyenne harmonique se trouvant dans la relation fondamentale.
- En vue d'obtenir la distribution des TIV, l'étape intermédiaire obligatoire est d'établir les distributions des vitesses en fonction de la concentration K . La Figure 2.4 donne une comparaison entre la distribution des vitesses réellement obtenue sur les données de la RN118 et celle prévue par le modèle de Heidemann. Visiblement, les résultats ne sont pas compatibles. Il faut ajouter cependant que la distribution des vitesses n'est pas la cause de la distribution des TIV même si ces deux informations se produisent en même temps et sont corrélées.

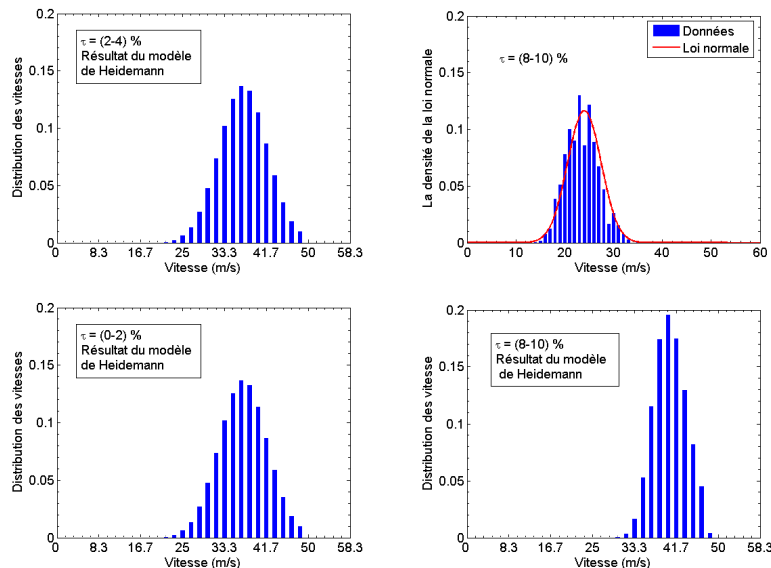


FIGURE 2.4 : Les distributions empiriques et le modèle de Heidemann

- L'hypothèse que les DIV suivent une loi exponentielle joue un rôle central dans le modèle lorsqu'on estime les espérances conditionnelles des vitesses et la distribution des TIV. Pourtant, cette hypothèse ne semble raisonnable qu'en trafic faible.
- Basé sur les trois conditions initiales, le modèle semble simplifier les phénomènes du trafic. Pour maîtriser les résultats du modèle et rapprocher plus les comportements de conduite de la réalité, plusieurs autres paramètres à valeurs prédéterminées sont mis en jeu, tels que $t_a, t_p, K_j, g(K)$. Cependant, ces paramètres ne fournissent pas la flexibilité nécessaire et ne rendent pas le modèle plus compatible avec la réalité. Par ailleurs, la fonction $g(K)$ peut ne pas être continue mais plutôt discontinue comme une *Z-ligne* dans la théorie des *trois-phases* du trafic de Kerner [110].
- Au niveau pratique, l'approche de Heidemann est compliquée à réaliser car les fonctions de densité ou de répartition de probabilité ne sont pas explicites. En outre, le champs d'application du modèle n'est pas encore clair : en espace, sur un tronçon/un

réseau à un instant donné, ou à un point fixé en fonction du temps, ou bien en espace durant une période de temps donnée.

2.1.2 Modèle des automates cellulaires de Nagel – Schreckenberg

Le modèle

L'approche de Heidemann est analytique et basée sur des conditions initiales. Nagel et Schreckenberg [111] ont développé un algorithme de simulation des états du trafic routier en discrétisant l'espace et le temps. Un des résultats de cette approche est d'obtenir la distribution des TIV discrétisés.

Tout d'abord, la route est divisée en L cellules identiques. L'état σ_i de la i ième cellule quelconque est soit *vide* (s'il n'y a aucun véhicule) soit *occupé* (le cas contraire). Chaque cellule est capable d'accueillir un seul véhicule au maximum :

$$\sigma_i = \begin{cases} 0 & \text{si } i \text{ est vide} \\ 1 & \text{si } i \text{ est occupée} \end{cases} \quad (2.21)$$

Ensuite, les vitesses sont fixées à des valeurs entières, de 1 à v_{\max} . Et le temps doit s'écouler par pas discrets. Un système de L cellules dans lequel les états de toutes les cellules et les vitesses associées aux cellules occupées sont connus à un moment peut être représenté par une *configuration* du système. Une configuration de n - cellules (aussi dit n - *cluster*) est un n - *uplet*, sa probabilité est désignée par $\mathcal{P}_n(\sigma_1, \sigma_2, \dots, \sigma_n)$. Si la i ième cellule est occupée, on appelle v_i sa vitesse et d_i sa distance (calculée en nombre de cellules) au véhicule précédent.

L'évolution du système se fait après chaque pas de temps durant lequel les vitesses et les positions des cellules sont mises à jour, conformément à un algorithme en 4 étapes :

1. **Accélération** : Si $(v_i < v_{\max}, d_i > v_i + 1)$, alors $v_i \rightarrow v_i + 1$
2. **Décélération** : Si la cellule $(i + j)$ est occupée et $j \leq v_i$, alors $v_i \rightarrow j - 1$
3. **Étape aléatoire** : $\Pr\{v_i \rightarrow v_i - 1 \mid v_i > 0\} = p$
4. **Mouvement des véhicules** : Le véhicule de la cellule i avance v_i cellules.

Nagel et Schreckenberg mettent l'accent sur l'importance de la troisième étape où l'aléa intervenu dans le processus le rend réaliste. De plus, Schreckenberg [112] a proposé d'utiliser le processus *2-3-4-1* plutôt que celui *1-2-3-4* en vue d'éviter une vitesse nulle possible qui surviendrait immédiatement à l'étape 1. Plusieurs approches numériques conduisent à utiliser en pratique la valeur $v_{\max} = 5$, une longueur d'une cellule de l'ordre de 7,5 m (notamment en congestion) et un pas temporel de 0,5 à 1 s.

À l'échelle macroscopique, les définitions du débit et de la concentration sont construites en un point donné i ; pour la lisibilité, l'indice i est omis dans les expressions du débit et de la concentration. À partir d'un moment $t_0 = 0$, la concentration et le débit du trafic sont calculés en un point i sur une période T comme suit :

$$\bar{\rho}^T = \frac{1}{T} \sum_{t=1}^T \sigma_i(t) \quad (2.22)$$

$$\bar{q}^T = \frac{1}{T} \sum_{t=1}^T \sigma_{i,i+1}(t) \quad (2.23)$$

où

$$\sigma_{i,i+1}(t) = \begin{cases} 1 & \exists \text{ un mouvement } i \rightarrow i+1 \\ 0 & \text{sinon} \end{cases} \quad (2.24)$$

Les études théoriques de l'algorithme N-S sont effectuées en particulier sur un système fermé, un anneau par exemple, dans lequel la concentration globale du trafic est constante. En effet :

$$\rho = \frac{\text{Card}\{i | \sigma_i = 1\}}{L} = C^{\text{te}} \quad (2.25)$$

Au niveau macroscopique, les résultats de Nagel et Schreckenberg semblent être proches des phénomènes observés dans réalité, par exemple :

- Pour un débit faible, la simulation des flux connaît des *flux laminaires*, c. à. d des trajectoires parallèles sur le diagramme spatio - temporel. Ce phénomène a été observé sur des autoroutes américaines.
- Pour un haut débit, une bande de vitesses nulles apparaît signifiant bien une congestion.
- Au niveau du diagramme fondamental, la forme de la courbe simulée débit - concentration ressemble à celle établie à partir des données réelles, bien que le débit maximal simulé dépende également du choix du pas de temps de la simulation.

Au niveau microscopique, Chowdhury *et al.* (1997) [82] approfondissent le modèle N-S en étudiant la distribution des DIV d_i dans un système circulaire fermé. Les résultats numériques montrent que :

- Lorsque ($v_{\max} = 5$ ou $v_{\max} = 3$) et $p = 0,5$:
 1. Quand ρ faible, la distribution des DIV est unimodale. Dans certains cas, $\Pr\{d_i \leq v_{\max}\} = 0$, ceci est expliqué par le phénomène de maintien des DIV de l'ordre de v_{\max} .
 2. Quand ρ approche d'une valeur ρ_m , disons de *concentration optimale*, la distribution devient bi-modale. En particulier, quand $\rho \rightarrow 1$, la distribution des DIV tend vers la fonction *delta* - Dirac δ .
- Quand ($v_{\max} = 2$ et $p = 0,5$) :
La distribution des DIV est unimodale pour toutes les valeurs de ρ . La probabilité $\Pr\{d_i = 0\}$ est positive jusqu'à la concentration $\rho \approx \rho_m$. Néanmoins, la forme bi-modale réapparaît pour p plus élevé, par exemple lorsque $p = 0,9$.

La formule analytique de la distribution des DIV n'est obtenue que dans le cas $v_{\max} = 1$. Pour cela, Chowdhury *et al.* désignent tout d'abord \mathcal{P}_j la probabilité qu'une DIV soit de j cellules :

$$\mathcal{P}_j = \mathcal{P}(1 | \underbrace{0 \dots 0}_j 1) \quad (2.26)$$

Afin de calculer l'expression à droite, deux approximations sont considérées :

- L'approximation selon un processus d'exclusion asymétrique simple (ASEP) : Le processus suppose que les corrélations entre les véhicules soient négligées. Par conséquent, les DIV suivront une loi binomiale : $\mathcal{P}_j = c(1-c)^j$ où $c = \Pr\{\sigma_i = 1\}$ est la concentration constante du trafic dans l'anneau.
- L'approximation par fenêtre de n - cellules : Cette approche considère que l'ensemble des véhicules sont composés de groupes de n cellules consécutives. De plus, dans chaque groupe, les corrélations d'un véhicule avec ses $(n-1)$ voisins les plus proches sont non négligeables. Deux groupes voisins ont alors $(n-1)$ cellules en commun. Le

cas étudié par Chowdhury correspond à $n = 2$ et on a :

$$\mathcal{P}_2(0) = \mathcal{P}_2(\underline{1} | 1) \quad (2.27)$$

$$\mathcal{P}_2(1) = \mathcal{P}_2(\underline{1} | 0) \mathcal{P}_2(\underline{0} | 1) \quad (2.28)$$

$$\mathcal{P}_2(j > 1) = \mathcal{P}_2(\underline{1} | 0) [\mathcal{P}_2(\underline{0} | 0)]^{j-1} \mathcal{P}_2(\underline{0} | 1) \quad (2.29)$$

La résolution d'une équation différentielle de la probabilité $\mathcal{P}_2(1 | 0)$ donne le résultat (cf. [112]) :

$$\mathcal{P}_2(\underline{1} | 0) = \frac{1 - \sqrt{4qc(1-c)}}{2q} = y \quad (2.30)$$

En conséquence :

$$\mathcal{P}_2(\underline{1} | 1) = c - y \quad (2.31)$$

$$\mathcal{P}_2(\underline{0} | 1) = y \quad (2.32)$$

$$\mathcal{P}_2(\underline{0} | 0) = 1 - c - y \quad (2.33)$$

Où $q = 1 - p$.

La distribution des DIV dans le cas $v_{\max} = 1$ est unimodale et de la forme :

$$\mathcal{P}_2(j) = \begin{cases} 1 - y/c & \text{pour } j = 0 \\ \frac{y^2}{c(1-c)} \left(1 - \frac{y}{1-c}\right)^{j-1} & \text{pour } j \in \mathbb{N}^* \end{cases} \quad (2.34)$$

Ghosh, Majumdar et Chowdhury (1997) [81] continuent à approfondir l'algorithme N-S pour étudier la distribution des TIV. La notion de TIV est tout d'abord définie par le choix au préalable d'une cellule *détecteur* que l'on l'étiquette $j = 0$. Les cellules suivantes sont successivement $j = 1, j = 2, \dots$ etc. Le comptage du temps doit être remis à 0 chaque fois qu'un véhicule passe le détecteur.

Comme pour la détermination de la distribution des DIV, la distribution des TIV n'est obtenue analytiquement que dans le cas $v_{\max} = 1$. Tous les cas où $v_{\max} > 1$ ne peuvent être résolus que par simulation numérique.

Lorsque $v_{\max} = 1$, $\mathcal{P}(t)$ désigne la probabilité qu'un TIV soit égal à t . On a :

$$\mathcal{P}(t) = \sum_{t_1=1}^{t-1} \mathcal{P}(t_1) Q'(t - t_1 | t_1) \quad (2.35)$$

où la probabilité $Q'(t - t_1 | t_1)$ est la probabilité conditionnelle que le véhicule arrivé à t_1 périodes après le précédent, s'installe au détecteur $j = 0$ pendant une période de $t - t_1$.

Deux possibilités pour calculer la probabilité $Q'(t - t_1 | t_1)$:

1. La première approche :

La relation entre le temps d'arrêt après l'arrivée et l'instant d'arrivée est négligée. Cela veut dire que :

$$Q'(t - t_1 | t_1) = Q(t - t_1) \quad (2.36)$$

La configuration intéressante dans ce cas est celle $(\underbrace{100 \dots 0}_n | \underline{0})$ avec $n = t_1$ (où le chiffre souligné signifie la vacuité du détecteur $j = 0$). La probabilité d'obtenir une telle configuration est :

$$\Pi(n) = \mathcal{P}(1 | \underline{0}) \times [\mathcal{P}(0 | \underline{0})]^{n-1} \quad (2.37)$$

En conséquence, la probabilité d'obtenir un TIV de valeur t_1 est :

$$\begin{aligned}\mathcal{P}(t_1) &= \sum_{n=1}^{t_1} \Pi(n) q^n p^{t_1-n} C_{t_1-1}^{n-1} \\ &= \mathcal{P}(1|0) q \times [\mathcal{P}(0|0) q + p]^{t_1-1}\end{aligned}\quad (2.38)$$

Où $q = \Pr\{\sigma_j = 1\}$ et $p = \Pr\{\sigma_j = 0\} = 1 - q$.

Ensuite, la configuration $(\underline{1}|\underbrace{11\dots 1}_m|0)$ signifie une file d'attente de $(m - 1)$ véhicules.

La probabilité d'obtenir une telle configuration est :

$$\Psi(m) = [\mathcal{P}(\underline{1}|1)]^{m-1} \times \mathcal{P}(\underline{1}|0) \quad (2.39)$$

Donc, la probabilité Q peut être calculée selon la formule :

$$\begin{aligned}Q(t_2) &= \sum_{m=1}^{t_2} \Psi(m) q^m p^{t_2-m} C_{t_2-1}^{m-1} \\ &= \mathcal{P}(1|0) q \times [\mathcal{P}(1|1) q + p]^{t_2-1}\end{aligned}\quad (2.40)$$

La distribution des TIV dans le cas $v_{\max} = 1$ est finalement fournie par la formule suivante :

$$\mathcal{P}(t) = \left(\frac{yq}{c-d}\right) \left[\left(1 - \frac{yq}{c}\right)^{t-1} - \left(1 - \frac{yq}{d}\right)^{t-1} \right] \quad (2.41)$$

Où $d = 1 - c$.

2. La deuxième approche :

À la différence de la première approche, Ghosh *et al.* prennent cette fois en compte la relation entre le temps d'arrêt après l'arrivée et l'instant d'arrivée. La fonction Q' a une forme différente :

$$Q'(t - t_1 | t_1) = \sum_{t_1=1}^{t-1} q [p^{t-t_1} S(t_1) + q R(t_1, t - t_1)] \quad (2.42)$$

Où :

$$S(t_1) = \sum_{t_2=1}^{t_1} Q(t_2) \quad (2.43)$$

$$R(t_1, t_3) = \sum_{t_2=t_1+1}^{t_1+t_3-1} Q(t_2) p^{t_2+(t_1-t_3)-1} \quad (2.44)$$

La distribution finale des TIV est :

$$\begin{aligned}\mathcal{P}(t) &= \left(\frac{yq}{c-y}\right) \left(1 - \frac{yq}{c}\right)^{t-1} + \left(\frac{yq}{d-y}\right) \left(1 - \frac{yq}{d}\right)^{t-1} - \\ &\quad - \left(\frac{yq}{c-y} - \frac{yq}{d-y}\right) p^{t-1} - q^2(t-1)p^{t-2}\end{aligned}\quad (2.45)$$

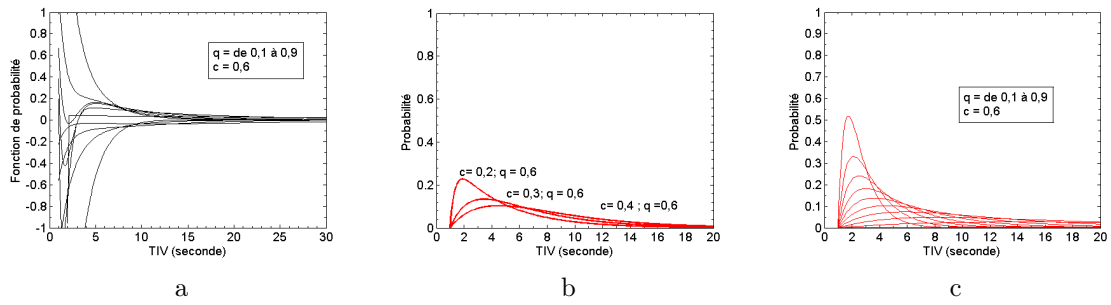


FIGURE 2.5 : Distribution des TIV utilisant le modèle N–S avec $v_{\max} = 1$

Résultats et Discussion

Les distributions obtenues en utilisant l'équation 2.45 ne sont pas réalistes car elles comportent des cas dans lesquels la probabilité des TIV est négative (cf. Figure 2.5-a). En principe, le modèle 2.41 donne des courbes de densité apparemment convenables comme le montrent dans les Figures 2.5-b,c.

Le modèle N–S est très intéressant à plusieurs titres. L'algorithme proposé se base sur des hypothèses simples de comportements individuels et facilite les calculs numériques. Cependant, le modèle pour les TIV ne peut être obtenu que dans le cas très simple où $v_{\max} = 1$, ce qui signifie que pour étudier les cas plus sophistiqués (par exemple la vitesse v_{\max} plus grande ou bien des conditions aux limites ouvertes plutôt que fermées comme dans un anneau circulaire), un modèle analytique n'est pas possible.

En outre, la variable TIV dans le modèle N–S n'est pas continue. C'est une variable aléatoire discrète car le pas de temps dans le modèle ne peut pas être trop court. En effet, le pas de temps est au moins suffisamment long pour qu'un véhicule de taille non négligeable change de cellule. Cette imprécision du modèle ne permet pas d'obtenir une distribution des TIV avec la granularité nécessaire aux études de sécurité routière ou de capacité de la route.

2.2 Approche probabiliste

À la différence des approches stochastiques précitées qui s'appuient sur le comportement des conducteurs face à des circonstances dans le trafic, l'approche probabiliste analyse directement la caractéristique aléatoire de la variable TIV H , et en proposant des hypothèses, établit des lois \mathcal{P} auxquelles H pourrait obéir, soit $H \sim \mathcal{P}$.

On considère que le TIV est une variable H à valeur dans $\mathbb{R}_+ = \mathbb{R} \cup +\infty$, $H : \Omega \rightarrow \mathbb{R}_+$ est une variable aléatoire réelle (v. a. r) non négative et sur un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$, où \mathcal{A} est la tribu et la mesure \mathbb{P} est la probabilité. Supposons en plus que H est une variable à densité, c.à.d qu'il existe une fonction $f(h)$ positive d'intégrale 1 correspondant à H . La question de l'approche probabiliste, et aussi du chapitre, est de trouver quelle est la fonction f en général.

Dans l'approche en considération, les données portent un rôle essentiel car les lois probabilistes doivent être confrontées aux données réelles avant d'être utilisées. En conséquence, les modèles sont généralement proposés *a priori* mais déterminés *a posteriori* tout en dépendant également de contextes dans lesquels les données sont recueillies. On peut aussi étudier H de façon globale sur l'ensemble des données enregistrées au sens où on ne distingue pas les contextes du trafic ou d'autres conditions hors trafic affectant les comportements de conducteur. Pourtant, cette dernière considération dépend fortement de la période d'enregistrement, et ne rend pas utile dans la pratique et ne fournit pas

de connaissance sur les comportements de l'ensemble. C'est ici, la question d'obtention des données représentatives est primordiale. Par conséquent, l'étape d'échantillonnage du chapitre précédent est très importante pour avoir des échantillons de TIV.

La fonction de densité f choisie doit satisfaire des critères objectifs afin de la comparer aux autres modèles et/ou juger les modèles par rapport aux données dans les échantillons. Les critères statistiques unanimement utilisés sont les tests de Kolmogorov–Simirnov (K–S), du χ^2 et d'Anderson–Darling (A–D). Par ailleurs, à côté de ce point de vue statistique, il y en a d'autres qui sont aussi importants tels que :

1. **Le niveau des paramètres** : les paramètres du modèle doivent être explicables [59],
2. **Le niveau de la structure** : les composants du modèle doivent être explicatifs [11],
3. **Le niveau d'interprétation** : le modèle doit être capable de représenter le comportement du conducteur et compatible avec les phénomènes du trafic (e.x. la capacité) [76].

Les sections suivantes parleront des modèles probabilistes utilisés et de leur structure.

2.2.1 Les modèles simples

1. Modèle Exponentiel

Si H suit une loi exponentielle décalée de paramètres λ et τ , $H \sim \mathcal{E}_\tau(\lambda)$, la densité f et la fonction de répartition F sont :

$$f(h) = \lambda e^{-\lambda(h-\tau)} \times \mathbb{1}_{\{h \geq \tau\}} \quad (2.46)$$

$$F(h) = [1 - e^{-\lambda(h-\tau)}] \times \mathbb{1}_{\{h \geq \tau\}} \quad (2.47)$$

Les transformations de Laplace et de Fourier de la fonction f sont :

$$f^L(s) = \frac{\lambda}{s + \lambda} e^{-s\tau} \quad (2.48)$$

$$f^F(\omega) = \frac{\lambda}{i\omega + \lambda} e^{-i\omega\tau} \quad (2.49)$$

Les principales statistiques de H sont :

$\mathbb{E}[H]$	$=$	$\tau + \lambda^{-1}$	Md	$=$	$\tau + \ln 2 \lambda^{-1}$	Mo	$=$	τ
Var $[H]$	$=$	λ^{-2}	S_3	$=$	2	K_4	$=$	9

$$CV = \frac{1}{1 + \lambda\tau} \quad (2.50)$$

Le paramètre $\lambda > 0$ est le paramètre d'intensité du trafic, et le paramètre $\tau \geq 0$ est un paramètre de décalage.

La Figure 2.6 montre l'ajustement du modèle \mathcal{E}_τ aux données réelles de la RN118. On remarque que la distribution modélise apparemment bien les TIV longs.

La raison d'utiliser le modèle $\mathcal{E}(\lambda)$ est sa liaison étroite avec le processus de Poisson des événements ponctuels aléatoires. En effet, les arrivées des véhicules en un point donné peuvent être considérées comme un tel processus dans certaines conditions. Le processus de Poisson décrit une suite d'événements *complètement aléatoires* ayant une intensité de survenue constante. Si l'on appelle $N_{t,t+h}$ le nombre d'événements qui se produisent dans

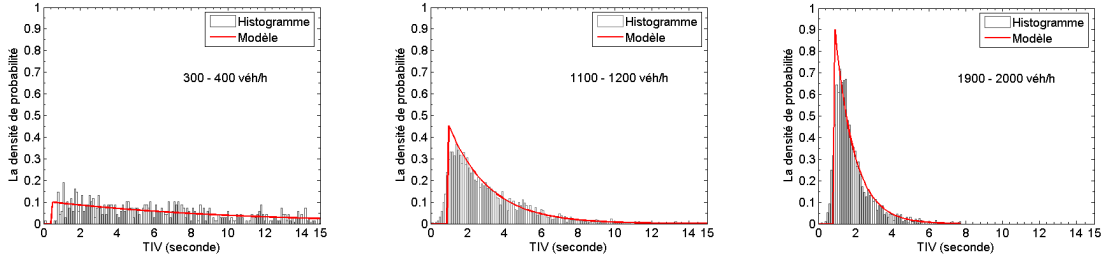


FIGURE 2.6 : Le modèle exponentiel décalé et les données de la RN118

l'intervalle de temps $(t, t + h]$, avec $h > 0$, le processus de Poisson est caractérisé par les deux propriétés suivantes :

$$1. \quad \Pr \{N_{t,t+h} = 0\} = 1 - \Lambda h + o(h) \quad (2.51)$$

$$2. \quad \Pr \{N_{t,t+h} = 1\} = \Lambda h + o(h) \quad (2.52)$$

Ces deux propriétés impliquent que :

$$a. \quad \Pr \{N_{t,t+h} \geq 2\} = o(h) \quad (2.53)$$

$$b. \quad \Pr \{N_{t,t+h} = r\} = \frac{(\Lambda h)^r}{r!} e^{-\Lambda h} \quad \text{Avec } r \in \mathbb{N} \quad (2.54)$$

Le paramètre Λ dans les équations 2.51 et 2.52 désigne le nombre moyen d'événements dans une période de temps unité. La propriété (a) dans l'équation 2.53 consiste à exclure les événements multiples tel que le passage de deux véhicules au point de référence en même temps (quand $h \rightarrow 0$) sur la même voie. La propriété (b) dans l'équation 2.54 montre que $N_{t,t+h}$ ne dépend que de h , non de t , et on peut donc écrire $N_{t,t+h} = N_h$ - une v. a à valeurs discrètes. Il est prouvé que la propriété (b) est équivalente à la fonction f dans (2.46), et dans ce cas-là, on a $\Lambda = \lambda$.

En trafic, le phénomène *complètement aléatoire* pourrait se produire dans les conditions de débit très faible où les interactions entre véhicules sont négligeables. Dans ce cas-là un processus de Poisson peut être obtenu. Si l'on regarde le trafic d'un point de vue spatial, Breiman [54] a aussi prouvé l'existence d'un processus de Poisson des arrivées des véhicules. Sous l'hypothèse que les vitesses sont associées aux véhicules de façon indépendante, le nombre de véhicules trouvés dans l'intervalle I de la route suit en théorie une loi exponentielle. On a alors :

$$\lim_{T \rightarrow +\infty} \Pr \{N_I(T) = r\} = \frac{\Lambda_I^r}{r!} e^{-\Lambda_I} \quad (2.55)$$

où T est la période durant laquelle les véhicules sont comptés sur la section de longueur I . Le paramètre Λ indique le taux d'arrivée des véhicules et ne dépend que de T . Le théorème de Breiman peut être généralisé sur une union des intervalles I_i disjoints [113].

En somme, tant du point de vue temporel en un point donné que du point de vue spatial sur un intervalle de la route, le caractère aléatoire des arrivées a un lien mathématique avec le processus de Poisson, autrement dit avec le modèle \mathcal{E} . Ce constat est aussi conforté par le fait que le modèle \mathcal{E} représente très bien la distribution des TIV élevés. En conséquence, ce modèle exponentiel joue toujours un rôle essentiel et central dans la modélisation probabiliste des TIV.

\mathcal{E} est le premier modèle de TIV. Il est tout d'abord utilisé en tant qu'exemple d'application de la théorie des probabilités [41, 42, 43]. Adams l'a ensuite testé sur les données réelles [2]. Pour cela, les données sont regroupés en classes de débit de 70 à 1400 véh/h

et le nombre de véhicules est estimé et comparé avec le nombre de véhicules réellement passés :

$$\text{Card} \{H > h\} = N e^{-ah} \quad (2.56)$$

où N est le nombre total de véhicules considérés, a est le nombre moyen de véhicules arrivant à chaque seconde.

Afin de mettre en relation avec le modèle $\mathcal{E}_{\tau=0}$, on a :

$$\frac{\text{Card} \{H > h\}}{N} = \Pr \{H > h\} = 1 - \Pr \{H \leq h\} = 1 - F(h) \quad (2.57)$$

En comparant avec la formule (2.47), le paramètre a coïncide donc avec le paramètre λ .

Par graphique, Adams (1936) a obtenu une proximité des deux courbes empirique et théorique qui l'amène à défendre le modèle \mathcal{E} dans le cas des débits faibles.

Greenshields (1947) [44] a obtenu la même conclusion en étudiant les échantillons de débit de 100 à 1800 véh/h. Ensuite, Ricker (1948) sur les données de différentes voies a proposé l'utilisation du modèle \mathcal{E} dans les cas suivants, avec Q_t le débit de l'ensemble de la route :

- Pour une route à 1 voie : $Q_t \leq 200$ véh/h
- Pour une route à 2 voies : $Q_t \leq 400$ véh/h
- Pour une route *multivoies* : $Q_t \leq 1000$ véh/h

Gerlough (1955) [46] a décrit les caractéristiques et les applications du modèle \mathcal{E} pour le trafic et proposé en 1959 pour la première fois d'introduire le paramètre de décalage τ dans le but d'éviter une sur-estimation de la probabilité des TIV courts. C'est donc le modèle \mathcal{E}_τ .

Récemment, Pueboobpaphan *el al.* [88] (2008) utilisent le modèle \mathcal{E} pour modéliser les arrivées des véhicules *traceurs* (*ang. probe vehicle*). Les véhicules traceurs sont ceux qui sont munis d'instruments modernes d'identification (systèmes AVI - *Automatic Vehicle Identification*), de positionnement (GPS - *Global Positioning System*), de transmission, etc...Le TIV dans ce contexte est défini par le TIV *tracé* - la différence minimale en temps entre les instants d'arrivée des véhicules traceurs en un point donné, quelque soit la voie de la route. Cette définition permet des événements multiples et une probabilité positive pour le TIV nul. Pueboobpaphan *el al.* observent que :

1. Si la distribution de tous les TIV est exponentielle, la distribution des TIV *tracés* l'est aussi,
2. Si le débit des véhicules traceurs est faible, la distribution des TIV *tracés* sera exponentielle quelque soit le débit de l'ensemble des véhicules.

En somme, le modèle \mathcal{E} semble fournir de bons ajustements pour la distribution de TIV de l'ensemble des véhicules lorsque le débit est faible, ou pour la distribution des TIV d'une partie particulière de l'ensemble des véhicules lorsque le débit cette partie est faible. La partie particulière en question est soit l'ensemble des véhicules ayant un TIV long, soit l'ensemble des véhicules traceurs participant au trafic.

2. Modèle Gamma

Soient $H_1, H_2 \dots H_n \sim \mathcal{E}(\lambda)$ des v.a.r indépendantes et identiquement distribuées (i.i.d). La loi de la somme $H = H_1 + H_2 + \dots + H_n$ suit une loi Gamma de paramètres n et λ , $H \sim \Gamma(n, \lambda)$. Les fonctions de densité f et de répartition F d'une loi $\Gamma_\tau(\beta, \alpha)$ décalé avec le paramètre de décalage τ sont données ci-dessous :

$$f(h) = \frac{\alpha^\beta (h - \tau)^{\beta-1}}{\Gamma(\beta)} e^{-\alpha(h-\tau)} \times \mathbb{1}_{\{h \geq \tau\}} \quad (2.58)$$

$$F(h) = \frac{\gamma[\beta, \alpha(h - \tau)]}{\Gamma(\beta)} \times \mathbb{1}_{\{h \geq \tau\}} \quad (2.59)$$

avec $\alpha, \beta \in \mathbb{R}_+$ et $\tau \geq 0$.

Les transformations de Laplace et de Fourier de la fonction f sont :

$$f^L(s) = \left(\frac{\alpha}{s + \beta} \right)^\beta e^{-s\tau} \quad (2.60)$$

$$f^F(\omega) = \left(\frac{\alpha}{i\omega + \beta} \right)^\beta e^{-i\omega\tau} \quad (2.61)$$

Les statistiques descriptives de H sont :

$\mathbb{E}[H]$	$= \tau + \frac{\beta}{\alpha}$	$\text{Mo} = \left[\tau + \frac{\beta - 1}{\alpha} \right] \times \mathbb{1}_{\{\beta < 1\}} + \tau \times \mathbb{1}_{\{\beta \geq 1\}}$
$\text{Var}[H]$	$= \frac{\beta}{\alpha^2}$	$S_3 = \frac{2}{\sqrt{\beta}} \quad K_4 = 3 + \frac{6}{\beta}$

$$CV = \frac{\sqrt{\beta}}{\beta + \alpha\tau} \quad (2.62)$$

Les paramètres α, β sont respectivement le paramètre d'échelle et le paramètre de forme du modèle Γ . On remarque que le modèle Γ est une généralisation du modèle \mathcal{E} et de la loi du χ^2 , en effet : $\mathcal{E}(\lambda) \equiv \Gamma(1, \lambda)$ et $\chi^2(n) \equiv \Gamma\left(\frac{n}{2}, \frac{1}{2}\right)$. De plus, lorsque $\beta \in \mathbb{N}^*$, le modèle Γ devient un modèle d'Erlang. Lorsque $\tau > 0$ le modèle Γ n'est autre que le modèle Pearson Type-III.

Dans l'équation 2.58, $\Gamma(\cdot)$ est la *fonction gamma* d'Euler :

$$\Gamma(\beta) = \int_0^\infty x^{\beta-1} e^{-x} dx \quad (2.63)$$

Dans l'équation 2.59, la fonction gamma incomplète $\gamma(\cdot, \cdot)$ est définie par :

$$\gamma(\beta, y) = \int_0^y x^{\beta-1} e^{-x} dx \quad (2.64)$$

On remarque aussi que la transformation de Laplace (ou de Fourier) de la convolution des fonctions est égale à la multiplication des transformations de Laplace (ou de Fourier) des fonctions facteurs :

$$\text{Si les } H_i \text{ sont i.i.d : } f^F\left(\sum_{i=1}^l H_i\right) = \prod_{i=1}^l f_i^F(H_i)$$

où f et f_i sont les densités de probabilité de $H = H_1 + \dots + H_l$ et des H_i respectivement.

La Figure 2.7 représente comment le modèle Γ ajuste les données des TIV de la RN118. On remarque que ce modèle est unimodal et positivement asymétrique comme les caractéristiques des distributions des TIV réels.

Le modèle Γ est proposé pour la première fois par Haight *et al* [53] sous forme la loi d'Erlang. En étudiant les états du trafic où la caractéristique aléatoire est remplacée par des interactions non négligeables entre les véhicules, Haight *et al* constatent qu'en utilisant

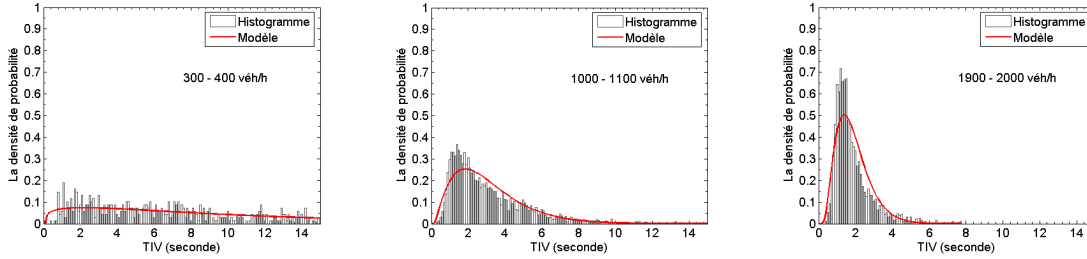


FIGURE 2.7 : Le modèle Pearson III et les données de la RN118

le modèle \mathcal{E} , le nombre de véhicules est sous-estimé par rapports aux données. Il suppose donc que la probabilité d'observer n véhicules dans une période donnée T est égale à une somme de k lois de Poisson de même paramètre λ .

$$\begin{aligned} \Pr \{N_T = n\} &= \sum_{i=1}^k \frac{(\lambda T)^{nk+i-1}}{(nk+i-1)!} e^{-\lambda T} \\ &= \sum_{r=nk}^{(n+1)k-1} \frac{(\lambda T)^r}{r!} e^{-\lambda T} \end{aligned} \quad (2.65)$$

Où N_T est le nombre de véhicules observés dans l'intervalle de temps T , et k est le nombre de composants de la somme. L'équation 2.65 est en fait la probabilité de dénombrement auquel le modèle d'Erlang $\Gamma(k, \lambda)$ en tant que la probabilité des intervalles correspond [53]. Le paramètre k est généralement prédéfini et dépend de l'état du trafic à modéliser. En constatant que lorsque $k = 1$, on obtient la loi de Poisson ayant la caractéristique de maximiser l'entropie qui quantifie l'aléatoire, et que, lorsque $k \rightarrow +\infty$, le modèle proposé tend vers l'état de régularité du trafic, Haight *et al* espèrent que le coefficient $1/k$ pourrait décrire différents états du trafic.

Dans son livre, May [114] a calibré et comparé le modèle Γ et Pearson III Γ_τ pour le TIV. Il a obtenu des formes raisonnables de la distribution pour les TIV courts mais le modèle sous-estime souvent les fréquences des TIV inférieurs à 1 s ou des TIV supérieurs à 4 s ce qui montre une capacité limitée du modèle Γ .

L'utilisation du modèle Γ est à mettre en perspective avec celle proposée par Krbalek *et al.* (2001) [84] qui étudie la distribution des TIV en appliquant des résultats sur les particules du gaz de Dyson des systèmes unidimensionnels en physique. Le système comporte N particules détenues dans un anneau en potentiel harmonique. Une particule caractérisée par la position x_i peut interagir avec r particules voisines au moyen des forces de répulsion. Dans ce gaz, on constate une distribution exponentielle des interdistances $s = \Delta x$ entre les particules à température élevée. En outre, les interdistances suivent une loi Γ à température normale. Il est à noter que la température est le résultat de l'ensemble des mouvements des particules elles-mêmes. Enfin, la distribution de s est comme suit avec un seul paramètre β positif représentant l'effet de la température.

$$f(s) = \frac{(\beta + 1)^{\beta+1} s^\beta}{\Gamma(\beta + 1)} e^{-(\beta+1)s} \quad (2.66)$$

Krbalek *et al.* appliquent donc cette même loi au trafic pour modéliser les données d'une autoroute en Allemagne et obtiennent de bons ajustements pour les trafics intermédiaires, c.à.d ni très faibles ni très élevés, avec un coefficient β assez stable, de l'ordre 3,24. Toutefois, on remarque que le modèle Γ ne prédit pas bien la fréquence des TIV aux alentours du *mode* de la distribution.

L'utilisation du modèle Γ proposé par Haight *et al.*, par May [114], et par Krbalek *et al.* a pour but de distinguer les principaux régimes du trafic. Ils ont classé le trafic en trois états différents. Le modèle Γ s'attache particulièrement à l'état intermédiaire du trafic :

1. *Le trafic libre* correspond aux TIV distribués complètement aléatoirement, $H \sim \mathcal{E}$ (T1),
2. *La congestion* correspond à une régularité dans laquelle les TIV sont stables, avec des phases *démarrage-arrêt* (*stop and go*) alternées (T2),
3. *Le trafic synchronisé* est l'état intermédiaire entre les deux états précédents (T3).

3. Modèle Log-normal

Soit $X \sim \mathcal{N}(\mu, \sigma)$ donc $H = \exp(X)$ suit la loi Log-normale $\ln \mathcal{N}(\mu, \sigma)$.

La densité f et la fonction de répartition F de H sont :

$$f(h) = \frac{1}{h - \tau} \phi\left(\frac{\ln(h - \tau) - \mu}{\sigma}\right) \times \mathbb{1}_{\{h \geq \tau\}} \quad (2.67)$$

$$F(h) = \Phi\left(\frac{\ln(h - \tau) - \mu}{\sigma}\right) \times \mathbb{1}_{\{h \geq \tau\}} \quad (2.68)$$

Les transformations de Laplace et de Fourier du modèle $\ln \mathcal{N}(\mu, \sigma)$ n'ont pas d'écriture analytique. Les principales statistiques descriptives de H sont :

$\mathbb{E}[H]$	$= \tau + e^\mu \sqrt{\xi}$	Md	$= \tau + e^\mu$	Mo	$= \tau + \frac{1}{\xi} e^\mu$
Var $[H]$	$= \xi e^{2\mu} (\xi - 1)$	S_3	$= (\xi + 2) \sqrt{\xi - 1}$	K_4	$= \xi^4 + 2\xi^3 + 3\xi^2 - 6$

$$CV = \frac{\sqrt{\xi - 1}}{1 + \tau e^{-\mu} \sqrt{\xi - 1}} \quad (2.69)$$

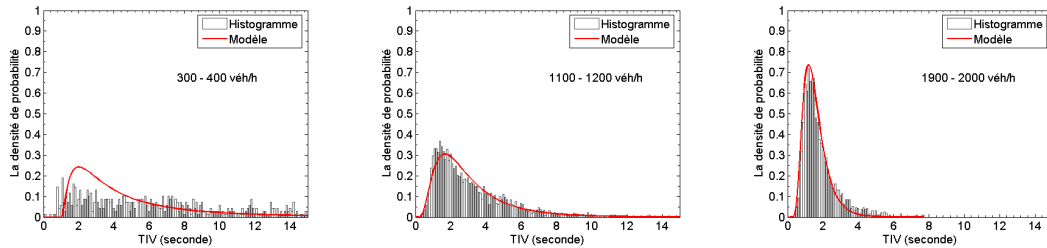


FIGURE 2.8 : Le modèle Log-Normal et les données de la RN118

où $\xi = e^{\sigma^2}$. Les paramètres μ et σ sont positifs. Le paramètre $\tau \geq 0$ est le paramètre de décalage. Les fonctions ϕ et Φ sont :

$$\phi(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x^2}{2}} \quad (2.70)$$

$$\Phi(x) = \int_0^x \phi(t) dt \quad (2.71)$$

D'après le théorème central limite, la loi $\ln \mathcal{N}(\mu, \sigma)$ peut être obtenue comme la loi d'une $v.a$ résultant d'une multiplication d'un grand nombre n de $v.a$ indépendantes, quand $n \rightarrow +\infty$. L'interprétation des paramètres μ et λ est que si $H \sim \ln \mathcal{N}(\mu, \sigma)$, alors la variable

$Y = \ln(H) \sim \mathcal{N}(\mu, \sigma)$, c.à.d que μ et σ sont respectivement la moyenne et l'écart-type des logarithmes des TIV.

Greenberg (1966) [55] fut le premier à proposer le modèle $\ln \mathcal{N}$ pour les TIV. À partir de quatre hypothèses initiales, il a montré que les $\ln(H)$ suivent une loi normale. Les quatre hypothèses de Greenberg sont :

1. Sur l'intervalle du temps $(t, t + \Delta t]$:

$$h_i(t + \Delta t) - h_i(t) = u_i(t) h_i(t)$$

où $u_i(t)$ est indépendant de $u_i(t \pm k\Delta t)$, $k = 1, 2, \dots$; i est l'indice de véhicule.

2. $\text{Var}(\ln(h_i(t))) = \text{Var}(\ln(h_i(t + \Delta t)))$
3. $\ln(1 + u_i(t)) = -a_i \ln(h_i(t)) + b_i + y_i(t)$ où a_i, b_i sont constants et $y_i(t)$ et une v.a indépendante de la v.a $\ln(h_i(t))$ et de $y_i(t \pm k\Delta t)$, $k = 1, 2, \dots$
4. $\mathbb{E}[\ln(h_i)(t)]$ est constant.

Si on divise t en N intervalles suffisamment petits Δt , c.à.d $t = N \Delta t$, les hypothèses donnent enfin que :

$$\ln h_i(t) = \ln(h_i(N \Delta t)) = \sum_{k=1}^N N(1 - a_i^2)^{N-k} (b_i + y_i(k \Delta t)) + o(1/N) \quad (2.72)$$

D'après le théorème centrale limite, si les y_i sont indépendantes et uniformément bornées, les composantes à droite tendent approximativement vers une loi normale \mathcal{N} quand $N \rightarrow +\infty$, autrement dit h_i suit une loi Log-normale $\ln \mathcal{N}$.

L'idée principale vient de l'hypothèse N°1 et est renforcée par l'hypothèse N°3. Les hypothèses N°2 et N°4 garantissent plutôt les stabilités de l'espérance et de la variance de H . L'hypothèse N°1 veut dire que si le changement de TIV d'un véhicule à l'instant $t + \Delta t$ est aléatoirement proportionnel au TIV de lui même à l'instant t , c.à.d Δt secondes juste avant, la suite des TIV (H_0, H_1, \dots, H_n) converge en probabilité vers la loi Log-Normale.

Zhang *et al.* [37] supposent que deux étapes $(i - 1)$ et i sont séparées par un intervalle de temps très court pour dériver la loi Log-Normale. En effet, on a :

$$H_i - H_{i-1} = H_{i-1} \lambda_i \quad (2.73)$$

L'accumulation des n intervalles donne :

$$\sum_{i=1}^n \lambda_i = \sum_{i=1}^n \frac{H_i - H_{i-1}}{H_i} \approx \int_{H_1}^{H_n} \frac{dH}{H} = \ln H_n - \ln H_1 \quad (2.74)$$

L'équation 2.74 montre aussi que $H_n \sim \ln \mathcal{N}$, à condition que les coefficients λ_i soient des v.a indépendantes et uniformément bornées.

Il est à noter que la notion de TIV de Greenberg est différente de celle considérée ici. Les hypothèses de Greenberg calculent le TIV h_i indirectement par les positions $X_i(t)$ des véhicules successifs et la vitesse $\dot{X}_i(t)$ du véhicule i à l'instant t . Il s'agit de :

$$h_i(t) = \frac{X_{i-1}(t) - X_i(t)}{\dot{X}_t} \quad (2.75)$$

De plus, les hypothèses concernent un processus de poursuite. Pour satisfaire les hypothèses de Greenberg, la présence du véhicule précédent semble être nécessaire comme *stimulus*, obligeant le véhicule courant à changer son TIV. Ceci n'est pas le cas en trafic fluide, et le modèle $\ln \mathcal{N}$ n'englobe pas cet état du trafic. Par ailleurs, lorsque le trafic

devient très dense, il n'est pas évident que les coefficients $u_i(t)$ soit des variables aléatoires et indépendantes.

Tolle (1971) [61] a calibré le modèle $\ln\mathcal{N}$ décalé sur de données réelles. Le modèle a fourni de bons ajustements pour la plupart des échantillons de débits entre 800 à 1900 véh/h au sens du test K-S. Faute probablement d'un nombre suffisant d'observations dans quelques cas, le test du χ^2 rejette l'hypothèse nulle de l'ajustement dans plusieurs échantillons.

He *et al.* (2002) [115] ont utilisé le modèle $\ln\mathcal{N}$ pour la distribution des TIV *désirés* (*ang. Preferred Time Headway (PTH)*). Cette notion de TIV est la différence entre les temps de passage de deux véhicules successifs qui se déplacent à peu près à la même vitesse ($|\Delta v| \leq 5 \text{ feet/s}$). Le test du χ^2 n'a pas rejeté l'hypothèse nulle de l'ajustement dans les 16 échantillons des TIV désirés autoroutiers.

Finalement, le modèle $\ln\mathcal{N}$ modélise bien statistiquement les données du trafic, et il est interprétable au niveau du comportement de conduite. Néanmoins, le manque d'une transformation explicite de Laplace et de Fourier rend difficile la combinaison de ce modèle avec les autres. Employé tout seul, le modèle $\ln\mathcal{N}$ est simple à calibrer, à utiliser, et les paramètres sont corrélés aux états du trafic.

4. Autre modèles

En général, le TIV est une variable décrivant l'interaction entre les composants dans un système normalement uni-dimensionnel. Passant du niveau véhicule particulier dans le transport routier au niveau des métros ou des trains dans le transport ferroviaire, la définition des TIV ne change pas mais l'échelle des valeurs des TIV est totalement transformée car les mécanismes ne sont pas les mêmes. Sur une section sans signalisation, les TIV routiers entre les véhicules sont auto-organisés, tandis que les TIV des métros sont régulés par des agents intermédiaires pour assurer la sécurité du trafic. Résultat, les arrivées des véhicules sont prévisibles et déterministes [116].

Cela veut dire que les caractéristiques du système dépendent du mécanisme de l'interaction entre les composants et entraîne par conséquent différents résultats. Pour plus de clarté, prenons un cas intermédiaire entre les deux cas du véhicule particulier et du métro précités : Le système d'organisation des autobus dans le transport en commun en Amérique Latine est un sujet très intéressant. L'arrivée des bus est sujette à prévision car le système ne s'appuie pas sur un horaire fixe. Les autobus appartiennent chacun à une société privée indépendante. En conséquence, l'arrivée des bus et l'intervalle de temps TIV entre deux arrivées successives dépendent plutôt d'un partage raisonnable de la clientèle. En effet, l'information sur l'arrivée du bus précédent est accessible au conducteur du bus suivant et le conducteur décide de partir en fonction du nombre de voyageurs à bord. L'information partagée ci-dessus représente la corrélation entre les composants du système et bouleverse complètement la régularité d'un système de transport en commun.

En étudiant les données des autobus au Mexique, Krbalek *et al.* (2001) [84, 87, 86] proposent la loi de Wigner pour modéliser la distribution des TIV :

$$f(h) = A h^2 e^{-B h^2} \quad (2.76)$$

Dans certains cas, le modèle de Rayleigh peut être utilisé :

$$f(h) = C h e^{-C h^2/2} \quad (2.77)$$

où A , B et C sont les paramètres positifs des modèles.

2.2.2 Modèle combiné

Partant du constat, dans les sections précédentes, que les TIV sont distribués différemment en fonction de l'état du trafic, certains auteurs ont pensé à combiner différents modèles en introduisant une pondération des modèles pour bien modéliser les cas intermédiaires. Supposons qu'une v.a.r H soit composée par l composantes exclusives et complémentaires H_i . Chaque composante H_i caractérise l'une des propriétés, désignée par p_i , de la variable H . La probabilité que H se réduise à sa composante H_i est notée π_i . De plus, les H_i suivent des distributions quelconques, pas nécessairement différentes. Soit les l fonctions de répartition F_1, F_2, \dots, F_l et les l fonctions de densité f_1, f_2, \dots, f_l des composantes H_1, H_2, \dots, H_l . Si F et f sont la fonction de répartition et de densité de H , elles sont respectivement les combinaisons linéaires des fonctions F_i et f_i . En effet :

$$\Pr \{p = p_j\} = \pi_j \quad (2.78)$$

Où π_j la part de participation du composante H_j dans la v.a.r H . L'exhaustivité donne :

$$\sum_{j=1}^l \pi_j = 1 \quad (2.79)$$

Ensuite :

$$\begin{aligned} \Pr \{H \leq h\} &= \sum_{j=1}^l \Pr \{p = p_j\} \Pr \{H \leq h | p = p_j\} \\ &= \sum_{j=1}^l \pi_j \Pr \{H \leq h | p = p_j\} \end{aligned} \quad (2.80)$$

$$= \sum_{j=1}^l \pi_j \Pr \{H_j \leq h\} \quad (2.81)$$

Donc :

$$F(h) = \sum_{j=1}^l \pi_j F_j(h) \quad (2.82)$$

Comme π_j ne dépend pas de h :

$$f(h) = \sum_{j=1}^l \pi_j f_j(h) \quad (2.83)$$

Le cas $l = 2$ est toujours utilisé pour la variable TIV. La première composante H_1 caractérise normalement les TIV *constraints*, c. à d. $\{p_1 := \text{TIV constraint}\}$ et la deuxième H_2 caractérise le TIV des véhicules libres, c. à d. $\{p_2 := \text{TIV libre}\}$. La combinaison linéaire des deux composantes est considérée comme un modèle combiné résultant d'un mélange linéaire de deux modèles simples de probabilité.

La fonction f s'écrit donc :

$$F(h) = \theta F_1(h) + (1 - \theta) F_2(h) \quad (2.84)$$

$$f(h) = \theta f_1(h) + (1 - \theta) f_2(h) \quad (2.85)$$

Les propriétés statistiques et les transformations de la densité des TIV sont induites par celles des composantes. En effet :

Les transformations de la fonction de densité :

$$f^L(s) = \theta f^L(s) + (1 - \theta) f^L(s) \quad (2.86)$$

$$f^F(\omega) = \theta f^F(\omega) + (1 - \theta) f^F(\omega) \quad (2.87)$$

L'équation 2.88 montre que les moments non centrés d'ordre k peuvent être obtenus par les dérivées de la transformée de Laplace en supposant que la fonction f^L appartient continûment à C^k à $s = 0$:

$$\mathbb{E}(H^k) = (-1)^k \frac{d^k f^L(s)}{ds^k} \Big|_{s=0} \quad (2.88)$$

De plus, les moments d'ordre k de H du modèle combiné sont :

$$\mathbb{E}[H^k] = \theta \mathbb{E}[H_1^k] + (1 - \theta) \mathbb{E}[H_2^k] \quad k = 1, 2, \dots \quad (2.89)$$

En conséquence, les principales propriétés statistiques sont calculées comme suit :

$$\mathbb{E}[H] = \theta \mathbb{E}[H_1] + (1 - \theta) \mathbb{E}[H_2] \quad (2.90)$$

$$\begin{aligned} \text{Var}[H] = \sigma^2 &= \mathbb{E}[H - \mathbb{E}[H]]^2 \\ &= \mathbb{E}[H^2] - \mathbb{E}[H]^2 \end{aligned} \quad (2.91)$$

$$\begin{aligned} S_3 &= \frac{1}{\sigma^3} \mathbb{E}[H - \mathbb{E}[H]]^3 \\ &= \frac{1}{\sigma^3} [\mathbb{E}[H^3] - 3\mathbb{E}[H] \mathbb{E}[H^2] + 2\mathbb{E}[H]^3] \end{aligned} \quad (2.92)$$

$$\begin{aligned} K_4 &= \frac{1}{\sigma^4} \mathbb{E}[H - \mathbb{E}[H]]^4 \\ &= \frac{1}{\sigma^4} [\mathbb{E}[H^4] - 4\mathbb{E}[H] \mathbb{E}[H^3] + 6\mathbb{E}[H]^2 \mathbb{E}[H^2] - 3\mathbb{E}[H]^4] \end{aligned} \quad (2.93)$$

Le modèle combiné est caractérisé par ses composantes. Vue la forme exponentielle de la distribution des TIV longs, l'une des deux composantes employées est normalement la loi exponentielle. Dans ce cas-là, le modèle combiné est appelé *Hyper-* modèle selon le nom du modèle constitué par la seconde composante. En pratique, la distinction entre les parts apportées par chacune des deux composantes n'est pas précise quantitativement mais l'indépendance et l'exhaustivité des priorités p_i des composantes sont implicitement respectées.

1. Modèle M3 de Cowan

Cowan (1975) [59] a proposé la combinaison des deux modèles suivants :

$$f_1 = \delta_\tau \quad (2.94)$$

$$f_2 = \lambda e^{-\lambda(h-\tau)} \times \mathbb{1}_{\{h \geq \tau\}} \quad (2.95)$$

où la fonction δ de Dirac représente une impulsion d'unité à $h = \tau$. La fonction de répartition est donc de la forme suivante :

$$F(h) = [1 - (1 - \theta)e^{-\lambda(h-\tau)}] \times \mathbb{1}_{\{h \geq \tau\}} \quad (2.96)$$

L'utilisation de la fonction δ signifie que $\Pr\{H \leq \tau\} = \Pr\{H = \tau\} = \theta$ et que les véhicules contraints ont un TIV égal à τ - le TIV minimum. Dans le cas contraire où les TIV sont supérieurs à τ , les véhicules sont libres.

Troutbeck (1986) [117] a utilisé le modèle M3 de Cowan pour la distribution des TIV dans les flux principaux aux carrefours sans feux. Le modèle M3 de Cowan permet de calculer facilement différentes caractéristiques du trafic en intersection comme le retard moyen du trafic, la capacité d'entrée au carrefour, etc. Aussi, Sullivan et Troutbeck (1994) [38] ont employé ce modèle pour les TIV aux giratoires et aux carrefours sans feux et confirmé encore une fois la facilité de calcul obtenue. Ce modèle se révèle bien plus précis que le modèle \mathcal{E} , et son ajustement est très significatif au niveau des TIV supérieurs à 4 s.

Hagring (1998) [118] a généralisé la distribution M3 de Cowan pour l'axe principal à plusieurs voies d'une intersection. Le modèle permet de formuler mathématiquement la capacité de l'intersection sur l'axe secondaire lorsque le flux de cet axe rencontre le flux de l'axe principal. Luttinen (1999) [73] a approfondi le modèle M3 de Cowan quant à la forme de la distribution, le CV, la relation K-S² et les méthodes d'estimation des paramètres du modèle. Les résultats d'estimation de la capacité d'un carrefour sans feux en utilisant le modèle M3 de Cowan sont bons, équivalant à ceux utilisant le modèle semi-Poisson.

Dans le modèle M3 de Cowan, le paramètre θ joue un rôle primordial qui conduit à étudier la proportion α des véhicules libre. Le paramètre α étant directement lié à θ par la relation $\alpha = 1 - \theta$, est souvent estimé en fonction du débit du trafic Q_I . Tanner [119] a proposé une relation linéaire tandis qu'une relation exponentielle est étudiée par Akçelik & Chung [72] (cf. Chapitre 5). Cette relation dépend aussi du nombre de voies de la route. Cependant, il faut noter que le modèle ne fournit pas de bons ajustements pour les TIV courts et n'est utile que dans les cas où les TIV longs jouent un rôle dominant, comme dans les interactions aux giratoires ou aux intersections sans feux, etc.

2. Modèle Hyperexponentiel et modèle de Schuhl

Le modèle Hyperexponentiel \mathcal{E}^* est une combinaison de deux modèles exponentiels de paramètres différents comme suit :

$$f_1 = \lambda_1 e^{-\lambda_1(h-\tau_1)} \times \mathbb{1}_{\{h \geq \tau_1\}} \quad (2.97)$$

$$f_2 = \lambda_2 e^{-\lambda_2(h-\tau_2)} \times \mathbb{1}_{\{h \geq \tau_2\}} \quad (2.98)$$

L'idée d'utiliser deux modèles exponentiels est proposée pour la première fois par Schuhl (1955) [47]. Dans le modèle, une composante caractérise les véhicules *retardés* et l'autre représente les véhicules libres. Selon Schuhl, le retard d'un véhicule retardé est induit par les véhicules précédents.

Pour les véhicules retardés, Schuhl suppose l'existence d'un TIV minimum τ . Il propose aussi de prendre en compte le dépassement des véhicules, ce qui entraîne la possibilité d'une probabilité positive pour un TIV nul. À l'origine, le modèle de Schuhl exprimait la probabilité qu'aucun véhicule se trouve dans un intervalle de temps valant h :

$$\begin{aligned} \mathcal{P}_0(h) = & \left[N\gamma\eta_1 e^{-\frac{h-\tau}{\eta_1}} + N(1-\gamma)\eta_2 e^{-\frac{h}{\eta_2}} \right] \times \mathbb{1}_{\{h > \tau\}} + \\ & + \left[N\gamma(\eta_1 - h) + N(1-\gamma)\eta_2 e^{-\frac{h}{\eta_2}} \right] \times \mathbb{1}_{\{0 \leq h \leq \tau\}} \end{aligned} \quad (2.99)$$

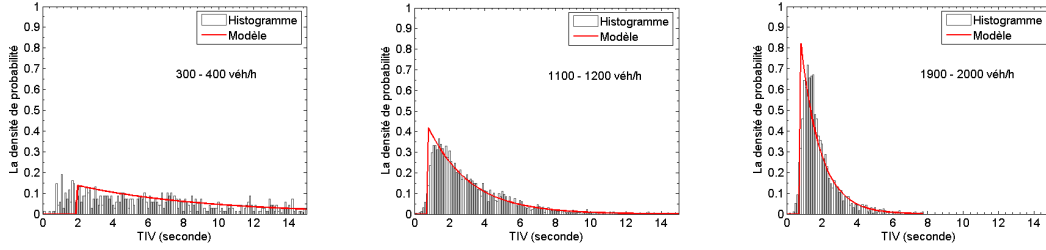


FIGURE 2.9 : Le modèle Hyperexponentiel et les données de la RN118

où N est le nombre total des véhicules dans la période unité. Les paramètres η_1 et η_2 sont respectivement les moyennes des TIV retardés et des TIV libres. Le paramètre γ est la proportion de nombre des véhicules retardés. La relation entre les paramètres est :

$$N\gamma\eta_1 + N(1 - \gamma)\eta_2 = 1 \quad (2.100)$$

On remarque dans le cas $h > \tau$ que le modèle de Schuhl n'est autre qu'un cas particulier du modèle Hyperexponentiel \mathcal{E}^* avec $\tau_1 = \tau$ et $\tau_2 = 0$. En effet, dans le cas où $h > \tau$, on a :

$$\mathcal{P}_0(h) = \Pr\{H > h\} \quad (2.101)$$

$$= 1 - \Pr\{H \leq h\} \quad (2.102)$$

$$= 1 - F(h) \quad (2.103)$$

$$= 1 - [\theta F_1(h) + (1 - \theta) F_2(h)] \quad (2.104)$$

$$= 1 - \theta \left[1 - e^{-\lambda_1(h-\tau)}\right] - (1 - \theta) \left[1 - e^{-\lambda_2 h}\right] \quad (2.105)$$

$$= \theta e^{-\lambda_1(h-\tau)} + (1 - \theta)e^{-\lambda_2 h} \quad (2.106)$$

L'équation (2.106) est identique à l'équation (2.99) si on pose :

$$\begin{aligned} \theta &\longleftrightarrow N\gamma\eta_1 \\ 1 - \theta &\longleftrightarrow N(1 - \gamma)\eta_2 \\ \lambda_1 &\longleftrightarrow 1 / (\eta_1 - \tau) \\ \lambda_2 &\longleftrightarrow 1 / \eta_2 \end{aligned}$$

Les paramètres λ_1 et λ_2 représentent les inverses des moyennes des TIV, et sont donc liées à l'intensité du trafic. Il est à noter que Schuhl avait soustrait la valeur τ dans le dénominateur de la fraction correspondante à λ_1 . Le produit $N\gamma\eta_1$ représente la part de la période unité occupée par les véhicules retardés, le produit $N(1 - \gamma)\eta_2$ est la part de durée occupée par les véhicules libres. En conséquence, Schuhl a utilisé le proportion θ au sens du temps et pas au sens du nombre de véhicules. On obtient donc la relation (2.100)

En outre, le modèle Hyperexponentiel et le modèle de Schuhl diffèrent dans le cas où $0 \leq h \leq \tau$. Schuhl a supposé qu'il y a un seul véhicule retardé qui a le TIV $h \leq \tau$, tandis que le modèle Hyperexponentiel en prédit zéro. Cette différence entraîne le remplacement de la probabilité $\exp\left[-\frac{h-\tau}{\eta_1-\tau}\right]$ par la probabilité $\left(1 - \frac{h}{\eta_1}\right)$ (cf. l'équation (2.99)). Utiliser le modèle \mathcal{E}^* provoque une petite surestimation de l'ordre de $N\gamma h$ sur la probabilité qu'il y ait 0 événement durant la période $(0, h]$; le développement suivant le montre (cf. l'équation 2.99 et l'équation 2.113) :

$$\mathcal{P}_0(h) = \Pr \{H > h\} \quad (2.107)$$

$$= 1 - \Pr \{H \leq h\} \quad (2.108)$$

$$= 1 - F(h) \quad (2.109)$$

$$= 1 - [\theta F_1(h) + (1 - \theta) F_2(h)] \quad (2.110)$$

$$= 1 - (1 - \theta) \left[1 - e^{-\lambda_2 h}\right] \quad (2.111)$$

$$= \theta + (1 - \theta) e^{-\lambda_2 h} \quad (2.112)$$

$$= N\gamma\eta_1 + N(1 - \gamma)\eta_2 e^{-\lambda_2 h} \quad (2.113)$$

Kell (1962) [49], Grecco (1968) *et al.* [60], et Khasnabis & Heimbach (1980) [66] ont proposé de modifier le modèle de Schuhl en enlevant les paramètres de décalage au dénominateur. Un paramètre de décalage τ_2 pour les véhicules libres peut être éventuellement introduit. Le modèle s'écrit sous la forme suivante :

$$\Pr \{H \geq h\} = \theta e^{-\frac{t-\tau_1}{\eta_1}} + (1 - \theta) e^{-\frac{t-\tau_2}{\eta_2}} \quad (2.114)$$

Le paramètre θ est la proportion des véhicules contraints; τ_1 et τ_2 sont respectivement les paramètres de décalage des véhicules contraints et libres; η_1 et η_2 sont les moyennes des TIV correspondant à ces deux types de véhicules.

Grecco *et al.* ont proposé de prendre $\tau_1 = 1$ s et $\tau_2 = 0$ pour un meilleur ajustement aux données réelles, et établi une abaque dans laquelle les paramètres du modèles sont calibrés en fonction du débit. Khanabis *et al.* quant à eux ont rencontré des problèmes de sous-estimation des fréquences des TIV courts.

Griffiths (1991) [120] a étudié les TIV en milieu urbain et proposé d'utiliser un modèle qu'il a nommé DDNED (Double Displaced Negative Exponential Distribution). Ce modèle est équivalent à un modèle Hyperexponentiel \mathcal{E}^* où $\tau_1 = \tau_2 = \tau$. La fonction de densité de probabilité est écrite comme suit :

$$f(h) = \left[\theta \lambda_1 e^{-\lambda_1 (h-\tau)} + (1 - \theta) \lambda_2 e^{\lambda_2 (h-\tau)} \right] \times \mathbb{1}_{\{h \geq \tau\}} \quad (2.115)$$

Où θ est la proportion des véhicules contraints; λ_1, λ_2 sont des paramètres d'échelle; τ le paramètre de décalage. Griffiths a obtenu de bons ajustements visuels du modèle pour les échantillons de débit allant de 172 à 1351 véh/h/voie.

Zhang *et al.* (2007) [37] ont examiné plusieurs modèles sur les données de la voie HOV (ang. High Occupancy Vehicle), qui réserve la voie pour les véhicules comportant un nombre suffisant de personnes. Il a constaté que le modèle Hyperexponentiel \mathcal{E}^* calibre bien la distribution des TIV sur la HOV.

3. Modèle Hypergamma et modèle Hyperlang

Le modèle Hypergamma Γ^* est le mélange entre un modèle Gamma Γ et un modèle exponentiel \mathcal{E} :

$$f_1 = \frac{\alpha^\beta (h - \tau_1)^{\beta-1}}{\Gamma(\beta)} e^{-\alpha(h-\tau_1)} \times \mathbb{1}_{\{h \geq \tau_1\}} \quad (2.116)$$

$$f_2 = \lambda e^{-\lambda(h-\tau_2)} \times \mathbb{1}_{\{h \geq \tau_2\}} \quad (2.117)$$

Dawson & Chimini (1968) [56] ont proposé que f_1 soit sous forme d'une loi d'Erlang représentant la distribution des véhicules contraints. Selon les auteurs, un phénomène non-aléatoire quelconque que l'on pourrait considérer comme intermédiaire entre deux phénomènes extrêmes, l'un complètement aléatoire et l'autre parfaitement régulier peut se

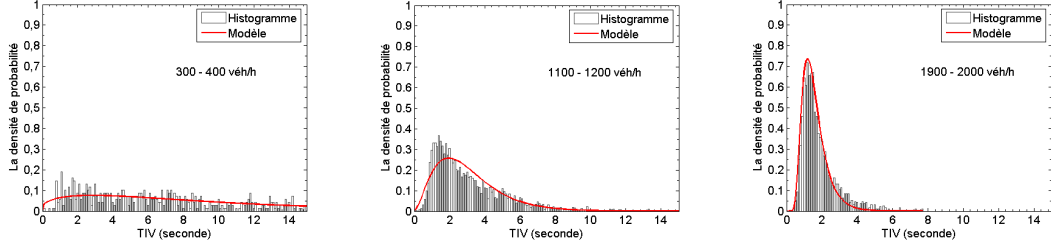


FIGURE 2.10 : Le modèle Hypergamma et les données de la RN118

modéliser par une fonction d'Erlang. Par conséquent, le modèle Hyperlang est proposé sous la forme :

$$\begin{aligned} \Pr \{H \geq h\} &= \alpha_1 \left\{ \left[e^{-\kappa \frac{h-\tau_1}{\eta_1-\tau_1}} \sum_{r=0}^{\kappa-1} \frac{1}{r!} \frac{(\kappa(h-\tau_1))^r}{\eta_1-\tau_1} \right] \times \mathbb{1}_{\{h \geq \tau_1\}} + \mathbb{1}_{\{h < \tau_1\}} \right\} + \\ &+ \alpha_2 \left[e^{-\frac{h-\tau_2}{\eta_2-\tau_2}} \times \mathbb{1}_{\{h \geq \tau_2\}} + \mathbb{1}_{\{h < \tau_2\}} \right] \end{aligned} \quad (2.118)$$

où les paramètres α_1, α_2 sont respectivement les proportions des véhicules contraints et des véhicules libres, et $\alpha_1 + \alpha_2 = 1$. Les paramètres η_1, η_2 sont respectivement les moyennes des TIV contraints et des TIV libres ; les paramètres τ_1, τ_2 sont respectivement le TIV contraint minimum et le TIV libre minimum. On obtient facilement la relation entre les paramètres de Dawson et le modèle Hypergamma Γ^* :

$$\Pr \{H \geq h\} = \Pr \{H > h\} \quad (2.119)$$

$$= 1 - \Pr \{H \leq h\} \quad (2.120)$$

$$= 1 - F(h) \quad (2.121)$$

$$= 1 - [\theta F_1(h) + (1 - \theta) F_2(h)] \quad (2.122)$$

$$= \theta \left\{ \left[e^{-\beta(h-\tau_1)} \sum_{r=0}^{\beta-1} \frac{1}{r!} (\beta(h-\tau_1))^r \right] \times \mathbb{1}_{\{h \geq \tau_1\}} + \mathbb{1}_{\{h < \tau_1\}} \right\} + \quad (2.123)$$

$$+ (1 - \theta) \left\{ \left[e^{-\lambda(h-\tau_2)} \right] \times \mathbb{1}_{\{h \geq \tau_2\}} + \mathbb{1}_{\{0 \leq h < \tau_2\}} \right\} \quad (2.124)$$

En comparant avec l'équation 2.118, on obtient les correspondances entre les paramètres :

$$\begin{aligned} \alpha &\longleftrightarrow \frac{1}{\eta_1 - \tau_1} \\ \beta &\longleftrightarrow \kappa \\ \lambda &\longleftrightarrow \frac{1}{\eta_2 - \tau_2} \end{aligned}$$

Il est à noter que Dawson *et al.* ont utilisé les moyennes empiriques des TIV pour représenter les débits des véhicules (libres ou contraints) du trafic. De plus, sur une base de données, le paramètre β a été pris égal à 2 pour tous les échantillons de 150 véh/h à 1050 véh/h (incrément de 100 véh/h) tandis que sur une seconde base de données, le paramètre $\beta \in \mathbb{N}$ varie de 1 à 6. Par ailleurs, sur des données d'Allemagne et d'Égypte, Aly (1989) [65] a obtenu des valeurs de β composées entre 1,65 et 2,27, ce qui incite à prendre pour β la valeur entière 2 dans le cas du modèle Hyperlang. Ce dernier est utilisé par Aly

pour modéliser les TIV des véhicules ayant une vitesse supérieure ou égale à 25 km/h, et pour établir une abaque comparant les distributions des TIV entre les routes (autoroutes et routes en rase campagne) en Allemagne et les routes en Égypte.

Il faut noter que le modèle Hypergamma Γ^* est une généralisation du modèle Hyperlang (dans lequel $\beta \in \mathbb{N}$), et du modèle Hyperexponentiel \mathcal{E}^* (dans lequel $\beta = 1$).

3. Modèle Hypernormal

May (1990) [121] modélise les TIV par une combinaison entre une loi normale et une loi exponentielle :

$$f_1 = \frac{1}{\sqrt{2\pi}\sigma^2} \phi\left(\frac{h-\mu}{\sigma}\right) \times \mathbb{1}_{\{h \geq 0\}} \quad (2.125)$$

$$f_2 = \lambda e^{-\lambda(h-\tau)} \times \mathbb{1}_{\{h \geq \tau\}} \quad (2.126)$$

où ϕ est la densité de probabilité de la loi normale centrée réduite.

En supposant qu'un véhicule suiveur poursuit un autre véhicule en régulant son TIV autour d'une valeur stable, May a essayé d'attribuer une loi normale (la première composante) aux TIV des véhicules suiveurs en peloton. La seconde composante représente les TIV des véhicules libres. Résultat : de faibles adéquations du modèle aux données notamment pour les TIV courts de 1 à 2,5 s. De plus, le modèle sur-estime les probabilités des TIV supérieurs à 4 s et sous-estime les fréquences des TIV de 2,5 à 4 s.

2.2.3 Modèle de mélange

Dans les modèles combinés, les caractères p_i des composantes sont statiques et prédéfinis explicitement (le cas d'Aly) ou implicitement (dans la plupart des cas). L'essentiel est d'assurer l'indépendance, la disjonction et l'exhaustivité entre les composantes. En conséquence, il n'y a pas de relation au moins mathématique entre les fonctions f_1 et f_2 .

Le modèle *de mélange* se base aussi sur la composition de deux composantes indépendantes f_1 et f_2 mais propose en plus une relation mathématique entre elles. Dans les modèles combinés, la densité f_2 est toujours exponentielle mais dans le modèle de mélange, elle est déduite de la densité f_1 tout en dépendant de l'hypothèse sur la relation entre les deux composantes. On a :

$$f(h) = \theta f_1(h) + (1 - \theta) f_2(h) \quad (2.127)$$

$$= \theta f_1(h) + (1 - \theta) \mathcal{F}[f_1(h)] \quad (2.128)$$

où \mathcal{F} signifie une relation en général plutôt qu'une application directe. Cependant, la tendance exponentielle des TIV longs représentée par f_2 doit être satisfaite ce qui implique que la fonction \mathcal{F} contient également un composant sous forme $e^{-\lambda h}$ avec λ positif.

Nous disposons jusqu'à présent de deux modèles de mélange : le modèle semi-Poisson (**ang.** Semi-Poisson Model (SPM)) et le modèle de file d'attente généralisée (**ang.** Generalized Queuing Model (GQM)), proposés respectivement par Buckley [57] et Branston & Cowan [58, 59].

Comme il y a mélange des distributions, l'interprétation du paramètre θ n'est pas la même que celle du modèle combiné. La frontière entre les deux composantes est plus abstraite et vraisemblablement aléatoire.

1. Modèle semi-Poisson (SPM)

Buckley (1968) [57] a proposé le modèle SPM dont le concept central est la variable aléatoire Z , dite *zone vide*. La variable Z d'un véhicule est un intervalle en temps devant le véhicule et qui empêche le processus d'arrivée de Poisson. Cet intervalle est différent du TIV minimum car il varie de façon aléatoire et est muni d'une fonction de densité $f_1(z)$. Ensuite, Buckley suppose que si le TIV réel du véhicule est supérieur à Z , ce TIV est une autre v.a V qui suit une loi exponentielle, $H_2 \sim \mathcal{E}$.

Soit une v.a $Y \sim \mathcal{E}$ de densité $g(y) = \lambda e^{-\lambda y}$, si $H > Z$, alors $H = H_2$ et $H_2 \stackrel{\mathcal{L}}{=} Y$. Supposons que H_2 est v.a de densité f_2 . À partir de l'hypothèse de distinction Z et H_2 , on a :

$$f(h) = \theta f_1(h) + (1 - \theta) f_2(h) \quad (2.129)$$

avec $\theta = \Pr\{H = Z\}$. La fonction $f_1(h)$ est supposée connue. La relation entre H_2 et Z entraîne la fonction de densité f_2 comme suit :

$$\begin{aligned} f_2(h) &= \frac{d}{dh} \Pr\{H_2 \leq h\} \\ &= \frac{d}{dh} [\Pr\{H_2 \leq h \mid H_2 > Z\} \Pr\{H_2 > Z\} + \Pr\{H_2 \leq h \mid H_2 \leq Z\} \Pr\{H_2 \leq Z\}] \\ &= \frac{d}{dh} \Pr\{H_2 \leq h \mid H_2 > Z\} \\ &= \frac{d}{dh} \Pr\{Y \leq h \mid Y > Z\} \\ &= \frac{d}{dh} \left[\frac{\Pr\{Y \leq h, Z < Y\}}{\Pr\{Y > Z\}} \right] \\ &= \frac{1}{\Pr\{Y > Z\}} \times \frac{d}{dh} \Pr\{Z < Y \leq h\} \end{aligned} \quad (2.130)$$

On a :

$$\begin{aligned} \Pr\{Y \leq h, Z < Y\} &= \int_0^h \int_0^y g(y) f_1(z) dz dy \\ &= \int_0^h g(y) \int_0^y f_1(z) dz dy \\ &= \int_0^h g(y) F_1(y) dy \end{aligned} \quad (2.131)$$

$$\begin{aligned} \Pr\{Z < Y\} &= \int_0^\infty \int_0^y g(y) f_1(z) dz dy \\ &= \int_0^\infty \lambda e^{-\lambda y} \int_0^y f_1(z) dz dy \\ &= \lambda F_1^L(\lambda) \\ &= f_1^L(\lambda) \end{aligned} \quad (2.132)$$

Finalement, la fonction de densité de probabilité des TIV est écrite comme suit :

$$f(h) = \theta f_1(h) + (1 - \theta) \frac{\lambda e^{-\lambda h}}{f_1^L(\lambda)} F_1(h) \quad (2.133)$$

La transformée de Laplace de f est :

$$f^L(s) = \theta f_1^L(s) + (1 - \theta) \frac{\lambda}{s + \lambda} \frac{f_1^L(s + \lambda)}{f_1^L(\lambda)} \quad (2.134)$$

Si l'on pose $c(h) = \frac{f_1(h)}{f_1^L(h)} e^{-\lambda h}$, les moments μ_n d'ordre n quelconque de la fonction f sont calculés grâce aux moments de f_1 et de c (cf. [57]) :

$$\begin{aligned} \mu_n(f) &= \theta \mu_n(f_1) + (1 - \theta) \mu_n(f_2) \\ &= \theta \mu_n(f_1) + (1 - \theta) \left[n! \sum_{j=0}^n \frac{\lambda^{j-n}}{j!} \mu_j(c) \right] \end{aligned} \quad (2.135)$$

Les formules (2.133, 2.135) signifient que la connaissance de la fonction f_1 est suffisante pour déterminer toute la distribution des TIV. Dans le modèle, le processus de Poisson est implicitement mis en jeu pour les TH longs et explicitement représenté dans l'équation (2.133).

Buckley suppose que f_1 est soit le modèle Γ , soit la loi normale \mathcal{N} . Le modèle SPM est d'ailleurs une généralisation du modèle Hyperexponentiel quand f_1 est une fonction exponentielle.

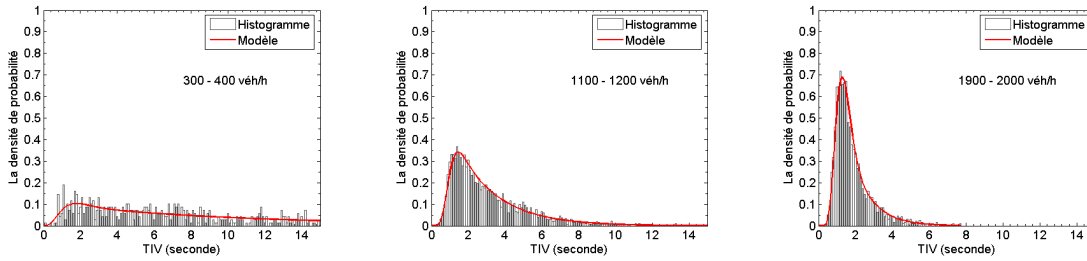


FIGURE 2.11 : Le modèle SPM et les données de la RN118

Visiblement, le modèle SPM convient très bien aux données de la RN118 dans la Figure 2.11. Luttinen (1999) [73] défend le modèle SPM dans sa thèse et l'utilise pour simuler des TIV. La fonction f_1 utilisée est le modèle $\Gamma(\beta, \alpha)$ de paramètres α, β . Dans ce cas, la densité et la fonction de répartition sont les suivantes :

$$f(h) = \theta \frac{\alpha^\beta h^{\beta-1}}{\Gamma(\beta)} e^{-\alpha h} + (1 - \theta) \frac{\lambda}{\Gamma(\beta)} \left(1 + \frac{\lambda}{\alpha}\right)^\beta \gamma(\beta, \alpha h) e^{-\lambda h} \quad (2.136)$$

$$F(h) = \theta \frac{\gamma(\beta, \alpha h)}{\Gamma(\beta)} + (1 - \theta) \frac{\lambda}{\Gamma(\beta)} \left(1 + \frac{\lambda}{\alpha}\right)^\beta \int_0^h \gamma(\beta, \alpha u) e^{-\lambda u} du \quad (2.137)$$

où $\alpha, \beta, \lambda \geq 0$, $0 \leq \theta \leq 1$, $\Gamma(\beta)$ et $\gamma(\beta, \alpha t)$ sont respectivement la fonction gamma et gamma incomplète (cf. Équations (2.62-2.64)).

Il est ensuite nécessaire de poser le problème inverse qui est d'identifier la distribution de f_1 à partir des données réelles, autrement dit à partir des estimations de la fonction $\tilde{f}(h)$. Wasielewski (1974) [63] a posé cette même question en étudiant la capacité du trafic car cette dernière, du point de vue de Wasielewski, est calculée par moyenne seulement des TIV courts. Pour résoudre cette question, Wasielewski a abouti à une équation intégrale résolue numériquement pour la fonction f_1 :

$$\theta f_1(h) = f(h) - A \lambda e^{-\lambda h} \frac{\int_0^h \theta f_1(u) du}{\int_0^\infty \theta f_1(u) du} \quad (2.138)$$

où les paramètres A et λ sont estimés à partir des équations suivantes :

$$\hat{\lambda}^{-1} = \frac{1}{n} \sum_i (h_i - \delta) \times \mathbb{1}_{\{t_i \geq \delta\}} \quad (2.139)$$

$$\hat{A} = \frac{n\delta}{N} e^{\hat{\lambda}\delta} \quad (2.140)$$

Le seuil δ est défini comme celui au delà duquel l'interaction entre les véhicules est très faible ; il est normalement déterminé *a priori* à 4 s ou à 6 - 8 s, donc les TIV suivent alors une loi exponentielle :

$$f(h) = (1 - \theta) f_2(h) = A \lambda e^{-\lambda h} \quad \text{Pour } h > \delta \quad (2.141)$$

Le paramètre N est le nombre total de TIV observés tandis que n_δ est le nombre des TIV $h < \delta$.

Un corollaire tiré en établissant le modèle SPM est que la fonction f_1 soit indépendante de débit du trafic. Wasielewski (1979) [64] a aussi testé l'hypothèse nulle en utilisant le test du χ^2 . Résultat : Avec les données de la voie centrale d'une autoroute urbaine à 6 voies, réparties en 12 groupes, $\delta = 3,5$ s pris *a priori*, l'hypothèse d'indépendance de la fonction de densité f_1 par rapport au débit moyen agrégé sur 10 min n'est pas rejetée au seuil de risque de 5 %.

2. Modèle de file d'attente généralisée (GQM)

Cowan (1975) [59] a proposé un système de quatre modèles pour les distributions des TIV. Les deux premiers modèles sont les modèles exponentiels sans et avec le paramètre de décalage τ . Le troisième est un modèle combiné étudié dans la section 2.2.2, tandis que le quatrième est écrit sous forme d'une fonction de répartition comme suit :

$$F(h) = \left[\theta G(h) + (1 - \theta) \int_0^h G(h - u) \lambda e^{-\lambda u} du \right] \times \mathbb{1}_{\{h \geq 0\}} \quad (2.142)$$

où G est la fonction de répartition des TIV contraints.

Disposant d'un échantillon de plus de 1300 TIV (précision de l'ordre 0,2 s) recueillis sur une section de 3,2 km d'un trafic stable pendant 2 h en Australie, Cowan en utilisant la loi Gamma pour la fonction G , a montré graphiquement une adéquation convenable. De plus, le modèle fournit des estimations de la moyenne, du *mode*, de la *médiane* proches des statistiques empiriques.

Branston (1976) a proposé un modèle en utilisant la théorie des files d'attente, et donné le nom du modèle de file d'attente généralisée, le modèle GQM (**ang.** *Generalized Queuing Model*). L'idée résulte de la similitude entre la sortie d'un système de files d'attente à entrée aléatoire et les mouvements des véhicules contraints passant en un point donné. Les considérations de sécurité, de stabilité de conduite et la restriction d'opportunité d'un dépassement peuvent prendre le rôle d'un serveur dans un système de file d'attente qui à son tour modifie le processus de Poisson et fait un véhicule libre devenant un véhicule contraint.

Deux hypothèses sont supposées pour le modèle GQM :

1. Les TIV contraints sont *i.i.d.*,
2. Les TIV contraints à la sortie sont indépendants de l'aléatoire à l'entrée.

Notons U la v.a à densité f_1 représentant les TIV contraints à l'aval du système dans le cas où le serveur fonctionne. Et ce serveur fonctionne aléatoirement avec une probabilité θ de disponibilité. Dans le cas contraire, une v.a H_3 représentant les TIV à la sortie doit comprendre un temps d'attente supplémentaire V jusqu'à ce que le serveur se

remette en service. Il est supposé que la variable $V \sim \mathcal{E}(\lambda)$ munie une fonction de densité $g(v) = \lambda e^{-\lambda v}$:

$$H_3 = U + V \quad (2.143)$$

La densité de probabilité du modèle GQM est donc de la forme :

$$\begin{aligned} f(h) &= \theta f_1(h) + (1 - \theta) f_3(h) \\ &= \theta f_1(h) + (1 - \theta) (f_1 * g)(h) \\ &= \theta f_1(h) + (1 - \theta) \int_0^h f_1(u) \lambda e^{-\lambda(h-u)} du \\ &= \theta f_1(h) + (1 - \theta) \lambda e^{-\lambda h} \left[\int_0^h f_1(u) e^{\lambda u} du \right] \end{aligned} \quad (2.144)$$

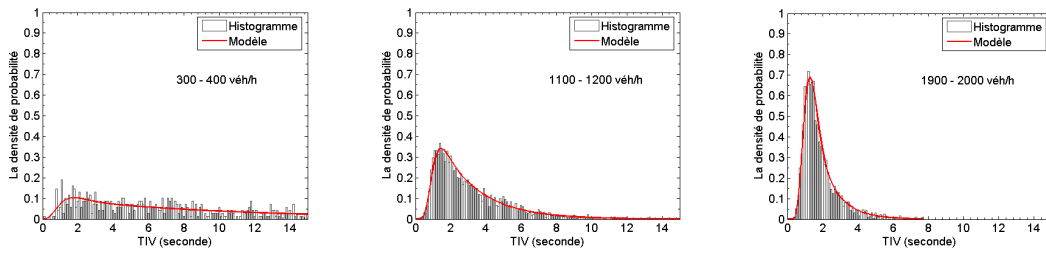


FIGURE 2.12 : Le modèle gamma-GQM et les données de la RN118

Branston a utilisé les lois log-normale, normale, Gamma et Bêta pour la densité f_1 ; la loi log-normale est le meilleur choix au sens du test χ^2 .

Zhang *et al.* ont clairement montré que le modèle M4 de Cowan n'est autre que le modèle GQM. En effet, dériver la fonction de répartition dans 2.142 donne :

$$\begin{aligned} \frac{dF}{dh} &= \theta g(h) + (1 - \theta) \left[\lambda G(h - h) e^{-\lambda h} + \int_0^h \frac{\partial}{\partial h} \left[G(h - u) \lambda e^{-\lambda u} \right] du \right] \\ &= \theta g(h) + (1 - \theta) \int_0^h g(h - u) \lambda e^{-\lambda u} du \\ &= \theta g(h) + (1 - \theta) \lambda e^{-\lambda h} \left[\int_0^h g(v) e^{\lambda v} dv \right] \end{aligned} \quad (2.145)$$

Notons que $g = f_1$ et $G(0) = 0$.

Pour mieux estimer la capacité du trafic en tenant compte des types de véhicules, Hoogendoorn (1998) *et al.* [80] ont proposé d'appliquer le modèle GQM pour chaque type p de véhicules, les reconstituer ensuite par combinaison linéaire pour de former la distribution des TIV finale.

$$F(h) = \sum_p \psi_p \left[\theta_p G_p(t) + (1 - \theta_p) \int_0^t G_p(t - u) \lambda_p e^{-\lambda_p u} du \right] \quad (2.146)$$

$$\text{Avec } \sum_p \psi_p = 1 \text{ et } \sum_p \psi_p \theta_p = \theta \quad (2.147)$$

où θ_p est la proportion des véhicules contraints parmi les véhicules de type p , ψ_p est la probabilité qu'un véhicule soit de type p .

2.2.4 Méthode d'estimation des paramètres

Supposons que l'échantillon $H = (H_1, H_2, \dots, H_n)$ ayant des variables aléatoires $H_i : \Omega \rightarrow \mathbb{R}_+$ soient i. i. d et que la loi commune des TIV H_i soit dans une famille de probabilités $\mathcal{P} = \{P_\theta \mid \theta \in \Theta \subset \mathbb{R}^l\}$. Soient $(h_1, h_2, \dots, h_n) \in \mathbb{R}_+^n$ des réalisations des H_i . Les paramètres θ , ou une fonction $g(\theta)$ de Θ sont généralement estimés par des statistiques S , dit *estimateurs* calculés à partir des données (h_1, h_2, \dots, h_n) . Les méthodes suivantes fournissent des estimateurs pour les modèles des TIV :

Méthode des moments (MM)

La fonction de densité $f(h, \Theta)$ permet en principe de calculer l moments pas nécessairement consécutifs pour la distribution des TIV, et de former un vecteur des moments $\mathcal{M} = (\mu_1, \mu_2, \dots, \mu_l)^t \in \mathbb{R}_+^{l \times 1}$, avec $\mu_i = \mu_i(\Theta)$. Par ailleurs, les moments ont aussi leur estimateurs calculés à partir de données réelles : $\hat{\mathcal{M}} = (\hat{\mu}_1, \hat{\mu}_2, \dots, \hat{\mu}_l)$, où $\hat{\mu}_i = \mu_i(h_i, i = 1, 2, \dots, n)$. Les estimateurs des l paramètres θ_i de Θ sont les composantes de la solution de l'équation vectorielle ci-dessous :

$$\mathcal{M}(\Theta) = \hat{\mathcal{M}} \quad (2.148)$$

La MM est souvent utilisée pour fournir des premières approximations numériques des paramètres. Elle est efficace pour les distributions symétriques en particulier pour celles ressemblant à la loi normale. En revanche, la méthode est moins robuste dans les cas d'estimation des distributions asymétriques comme celles des TIV. Le calcul de $\hat{\mathcal{M}}$ dans (2.148) demande d'ailleurs un grand nombre de données pour assurer une certaine précision. Les estimateurs de la MM sont asymptotiquement normaux lorsque n est grand mais ils sont *biaisés*.

Dans plusieurs cas où la résolution du système (2.148) est impossible, les paramètres pourraient être estimés de façon approximative en suivant le processus d'optimisation ci-dessous (cf. Ashton [62]) :

$$\hat{\Theta} = \text{Arg} \min_{\theta_i \in \Theta} \sum_{i=1}^l [\mu_i(\theta) - \hat{\mu}_i]^2 \quad (2.149)$$

Méthode du minimum du χ^2 (MC)

Les TIV (h_1, h_2, \dots, h_n) sont divisés en m groupes mutuellement exclusifs et exhaustifs. Les estimateurs des paramètres de la MMC sont ceux trouvés en minimisant la statistique suivante :

$$\chi^2(\Theta) = \sum_{i=1}^m \frac{[N_i - n \Pr\{h \in C_i \mid \Theta\}]^2}{n \Pr\{h \in C_i \mid \Theta\}} \quad (2.150)$$

où $\Pr\{h \in C_i \mid \Theta\}$ est la fréquence des TIV dans le groupe C_i ; elle dépend du modèle spécifique; elle est comparée à N_i le nombre d'observations dans le groupe C_i . La méthode n'est pas très efficace car elle perd normalement des informations du fait de la classification des données. La MMC demande d'ailleurs que le nombre des TIV dans chaque groupe C_i soit suffisamment grand.

Méthode du minimum de l'intégrale de la moyenne des carrés des erreurs (MISE)

Hoogendoorn (1997) *et al.* [79] ont proposé une méthode basée sur l'analyse harmonique de Fourier. La méthode permet d'estimer les paramètres des modèles, particulièrement ceux

de GQM et ceux des modèles qui disposent d'une fonction explicite de transformation de Fourier. La méthode s'appuie sur la minimisation d'une statistique D_n sous forme d'intégrale dans le domaine de fréquence. La statistique est la différence quadratique entre la fonction théorique et la fonction empirique toutes deux transformées par la formule Fourier :

$$D_n(\Theta) = \int_{-\infty}^{+\infty} |F(h|\Theta) - F_n|^2 dh \quad (2.151)$$

$$d_n(\Theta) = \int_{-\infty}^{+\infty} |f(h|\Theta) - f_n|^2 dh \quad (2.152)$$

où F_n, f_n sont respectivement la fonction de répartition empirique et la densité empirique, δ est la fonction de Dirac.

$$F_n = \frac{1}{n} \sum_{j=1}^n \mathbb{1}_{\{H_j | h_j \leq t\}} \quad (2.153)$$

$$f_n = \frac{1}{n} \sum_{j=1}^n \delta(h_i - H_i) \quad (2.154)$$

L'utilisation de ces statistiques dans le domaine de Fourier facilite le calcul numérique. Par conséquent, minimiser les statistiques $D_n^F(\Theta)$ et/ou $d_n^F(\Theta)$ des fonctions transformées est préférable :

$$D_n^F(\Theta) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} |F^F(\xi|\Theta) - F_n^F(\xi)|^2 d\xi \quad (2.155)$$

$$d_n^F(\Theta) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} |f^F(\xi|\Theta) - f_n^F(\xi)|^2 d\xi \quad (2.156)$$

où les transformations des fonctions empiriques F_n et f_n sont obtenues par :

$$F_n^F(\xi) = \frac{1}{i n \xi} \sum_{j=1}^n e^{i \xi H_j} \quad (2.157)$$

$$f_n^F(\xi) = \frac{1}{n} \sum_{j=1}^n e^{i \xi H_j} \quad (2.158)$$

En toute rigueur, on ne peut pas écrire $F^F(\xi)$ car F n'est pas une fonction périodique. Pour cela, une fonction périodique $\tilde{F}(h)$ est utilisée pour remplacer $F(h)$ en élargissant le domaine de définition :

$$\tilde{F}(h + k T_0 | \Theta) = F(h | \Theta) \quad \text{Avec } -\frac{T_0}{2} < h \leq \frac{T_0}{2}, \forall k \in \mathbb{N} \quad (2.159)$$

On peut écrire maintenant :

$$F^F(\xi | \Theta) = \tilde{F}^F(\xi | \Theta) \quad (2.160)$$

Comme \tilde{F} est une fonction - T_0 périodique, le développement de Fourier de cette fonction est :

$$\tilde{F}(h | \Theta) = \sum_{k=-\infty}^{+\infty} a_k(\Theta) e^{i \xi k h} \quad \text{Avec } \xi = \frac{2\pi}{T_0} \quad (2.161)$$

Où

$$a_k(\Theta) = \tilde{F}^F(k\xi | \Theta)$$

En outre, la relation de Parseval fournit :

$$\frac{1}{T_0} \int_{T_0} |\tilde{F}(\xi | \Theta)|^2 d\xi = \sum_{k=-\infty}^{+\infty} |a_k(\Theta)|^2 \quad (2.162)$$

Si l'on pose $F_n^F(k\xi) = b_k$, la statistique de distance quadratique intégrale est approchée par l'approximation suivante :

$$\begin{aligned} D_n^F(\Theta) &= \sum_{k=-\infty}^{+\infty} |a_k(\Theta) - b_k|^2 \\ &\approx \sum_{k=-N}^N |a_k(\Theta) - b_k|^2 \end{aligned} \quad (2.163)$$

$$= D_n^F(\Theta, N, T_0) \quad (2.164)$$

Idem pour la statistique $d_n^F(\Theta)$. L'approximation $D_n^F(\Theta)$ est toujours bornée tandis que $d_n^F(\Theta)$ pourrait tendre vers l'infini dans certains cas particuliers. On a également la limite suivante :

$$\lim_{N, T_0 \rightarrow \infty} D_n^F(\Theta, N, T_0) = D_n^F(\Theta)$$

Pour améliorer éventuellement la méthode, il est possible d'ajouter une fonction de pondération $q(\xi)$ à la fonction (2.155). Hoogendoorn propose que $q(\xi) \geq \sqrt{\xi}$ pour que la statistique d_n^* soit bornée :

$$d_n^*(\Theta, q(\xi)) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left| \frac{f^F(\xi | \Theta) - f_n^F(\xi)}{q(\xi)} \right|^2 d\xi \quad (2.165)$$

On s'aperçoit que $d_n^*(\Theta, 1) = d_n^F(\Theta)$ et $d_n^*(\Theta, \xi) = D_n^F(\Theta)$. Enfin, il semble que pour les $T_0 \geq 2 \max_i \{H_i\}$, l'approximation de \tilde{F} à F ne produit pas de grandes erreurs. L'idée de la méthode est la même que celle dans la méthode consistant à minimiser les statistiques qui servent à tester l'ajustement des modèles aux données. Cependant la méthode du minimum de la statistique K-S n'est pas utilisée ici.

Méthode du Maximum de Vraisemblance (MV)

On définit tout d'abord une fonction dite *fonction de vraisemblance* \mathcal{L} pour une distribution continue des TIV :

$$\begin{aligned} \mathcal{L} : \Theta &\longrightarrow \mathbb{R} \\ \Theta &\longmapsto f(h_1, h_2, \dots, h_n; \Theta) \end{aligned}$$

où (h_1, h_2, \dots, h_n) sont des réalisations de (H_1, H_2, \dots, H_n) . Du fait de l'indépendance entre (H_i) et de l'identité des distributions des H_i , la fonction $f(\Theta)$ est écrite :

$$f(h_1, h_2, \dots, h_n; \Theta) = \prod_{i=1}^n f_i(h_i; \Theta) \quad (2.166)$$

$$= \prod_{i=1}^n f(h_i; \Theta) \quad (2.167)$$

Alors :

$$\mathcal{L}(\Theta; h_1, h_2, \dots, h_n) = \prod_{i=1}^n f(h_i; \Theta) \quad (2.168)$$

Un estimateur $\hat{\Theta}$ de Θ est appelé *estimateur du maximum de vraisemblance* s'il maximise la fonction (2.168)

$$\hat{\Theta}(h_1, h_2, \dots, h_n) = \text{Arg} \max_{\Theta} \mathcal{L}(\Theta; h_1, h_2, \dots, h_n) \quad (2.169)$$

Il est à noter que l'estimation donne les mêmes résultats en utilisant le logarithme $\ln \mathcal{L}(\hat{\Theta})$ au lieu de $\mathcal{L}(\hat{\Theta})$:

$$\ln \mathcal{L}(\Theta; h_1, h_2, \dots, h_n) = \sum_{i=1}^n \ln f(t_i; \Theta) \quad (2.170)$$

L'estimateur MV est une estimation *fortement convergente, asymptotiquement normale* quand n devient grand et *asymptotiquement efficace*. La dernière propriété signifie que la variance des estimateurs obtenus est minimale parmi les autres estimateurs. La méthode MV fonctionne bien dans la majorité des cas à l'exception des modèles ayant un paramètre de décalage τ car il semble exister de grandes fluctuations des valeurs de $\ln \mathcal{L}$ au voisinage de la valeur optimale du paramètre de décalage (cf. Hoogendoorn [79]).

Par ailleurs, la méthode MV n'est cependant pas capable d'estimer globalement efficacement les paramètres du modèle $\ln \mathcal{N}$ décalé (LNM). Pour contourner la difficulté, les deux méthodes suivantes sont utilisées, et montrent toutes les deux des performances significatives [122] :

1. La méthode MV modifiée (MMVM),
2. La méthode MV locale (MMVL).

Les autres méthodes

A. Méthode des moindres carrés :

Avec le même principe de répartition en classes que celui utilisé dans la méthode MC, les paramètres des modèles de TIV sont estimés par le processus de minimisation de la somme des carrés suivante :

$$\hat{\Theta} = \text{Arg} \min \sum_{i=1}^n [F(h_i | \Theta) - F_n(h_i)]^2 \quad (2.171)$$

Dans le cas de l'utilisation de la densité f au lieu de la fonction de répartition F dans (2.171), le groupement préalable en m classes C_j est nécessaire et par conséquent, semblable à celui de la méthode MC 2.150 [37].

Khasnabis *et al.* [66] utilisent la méthode des moindres carrés pour estimer les cinq paramètres du modèle de Schuhl. La méthode fournit de bonnes adéquations au sens du test du χ^2 . À savoir, seuls trois échantillons sur 18 ont été rejetés au seuil de risque de 5 %.

B. Méthode de Bartholomew :

Bartholomew (1959) [123, 62] a proposé une méthode pour estimer le modèle Hyperexponentiel \mathcal{E}^* , sans paramètre de décalage, c. à d. $\tau_1 = \tau_2 = 0$. La méthode n'est utilisée que dans des cas très simples et limités. Il s'agit de prendre trois valeurs (arbitraires) de TIV h_1, h_2 et h_3 pour estimer les trois paramètres restant du modèle, de préférence : h_1 arbitraire, $h_2 = 2h_1$ et $h_3 = kh_1$. Vu que l'on peut calculer les valeurs empiriques de $p_i = 1 - F(h_i)$ avec $i = 1, 2, 3$, le système de trois équations pourrait être résolu soit analytiquement soit par la méthode numérique de Newton - Raphson (dans le cas du modèle Hyperexponentiel \mathcal{E}^*).

C. Combinaison des méthodes :

Dans plusieurs cas, notamment pour les modèles combinés, l'utilisation successive de plusieurs méthodes est préférée. Des contraintes sur une partie ou sur l'ensemble des paramètres à estimer sont éventuellement ajoutées. Un des exemples est celui qui consiste à poser une contrainte sur le TIV minimum $h_{(1)}$:

$$F^{-1} \left(\frac{1}{n+1} \mid \hat{\Theta} \right) = h_{(1)} \quad (2.172)$$

Un exemple typique est celui de l'estimation des paramètres du modèle M3 de Cowan, qui a fait l'objet en effet de nombreuses études. Tout d'abord, la méthode MV est utilisée pour estimer le paramètre λ , puis la méthode des moindres carrés pour obtenir l'estimateur du paramètre θ (cf. Luttinen [11]). Un autre exemple est le cas de Griffiths qui a estimé les paramètres du modèle Hyperexponentiel \mathcal{E}^* en employant la MM pour estimer les paramètres λ_1 et λ_2 et la méthode MC pour les autres paramètres. La combinaison des méthodes a fourni de bons ajustements aux échantillons pour les débits de 172 à 1351 véh/h/voie.

D. Méthode directe

Méthode d'utilisation d'une abaque : La méthode d'estimation des paramètres de Haight *et al.* par exemple s'est appuyée sur une abaque où les relations entre les paramètres sont prédéterminées. Haight l'applique pour l'estimation du modèle Hyperlang dans lequel deux paramètres λ et κ sont reliés par une liaison sur l'abaque :

$$\lambda = \mu \kappa + \frac{1}{2(\kappa - 1)} \quad (2.173)$$

où le paramètre μ est la moyenne des TIV dans l'échantillon étudié.

Aly [65] a estimé les paramètres du modèle Hyperlang en utilisant des estimateurs obtenus par des interprétations attribuées aux paramètres lors de la construction du modèle. Cette méthode représente bien l'indépendance et l'exhaustivité des composantes dans un modèle combiné car les seuils entre les composantes sont prédéterminés quantitativement. Cependant, la flexibilité du modèle sera restreinte.

2.2.5 Les tests d'ajustement

Les tests d'ajustement sont des tests de signification utilisés pour vérifier l'adéquation d'un modèle aux données réelles. Supposons que la variable H suive une loi de fonction de répartition $F_H(h)$, et que $F_0(h)$ est la fonction de répartition spécifiée. L'hypothèse nulle suivante est considérée :

$$H_0 : F_H = F_0$$

contre trois alternatives différentes :

- (a) $H_1^a : F_H \neq F_0$
- (b) $H_1^b : F_H \geq F_0$ et $F_H \neq F_0$
- (c) $H_1^c : F_H \leq F_0$ et $F_H \neq F_0$

Le test de H_0 contre H_1^a s'appelle un test d'ajustement bilatéral ; les tests (b) et (c) sont dits simples.

En général, on ne connaît pas la loi de F_H , donc F_H est pris à partir des données réelles. Les tests d'ajustement suivants sont utilisés pour mesurer le niveau d'ajustement d'un modèle par rapport aux données réelles. Par défaut, la P-valeur est employée pour signifier le rejet ou l'incapacité de rejet de l'hypothèse.

$$\text{P-valeur} = \Pr \{ H > h_{\text{observé}} \} \quad (2.174)$$

Dans le cas de $P\text{-valeur} \geq 0,05$, l'ajustement est confirmé et on ne peut pas rejeter l'hypothèse nulle au niveau de signification de 5%.

Le test du χ^2

La statistique χ^2 mesure la distance globale entre les valeurs estimées par un modèle et les valeurs correspondantes observées réellement. Le calcul est fait sur la fonction de densité de probabilité et la fonction de densité empirique. Pour cela, on divise tout d'abord les valeurs des observations des TIV en m classes, la statistique χ^2 est ensuite calculée par la formule (2.150).

La statistique χ^2 suit une loi du χ^2 à $m - k - 1$ degrés de liberté où k est le nombre de paramètres du modèle spécifique estimés à partir des données réelles. Plus elle est petite, plus l'ajustement est bon.

Le test de Kolmogorov–Smirnov (K–S)

Le test de Kolmogorov–Smirnov se base sur les statistiques suivantes :

$$\begin{aligned} D_n^+ &= \sup_{-\infty < h < \infty} \{F_n(h) - F_0(h)\} = \max_{1 \leq i \leq n} \{i/n - F_0(H_{(i)})\} \\ D_n^- &= \sup_{-\infty < h < \infty} \{F_0(h) - F_n(h)\} = \max_{1 \leq i < n} \{F_0(H_{(i)}) - (i-1)/n\} \\ D_n &= \sup_{-\infty < h < \infty} |F_n(h) - F_0(h)| = \max\{D_n^+, D_n^-\} \end{aligned}$$

Quand n est suffisamment grand, les statistiques D_n^+ , D_n^- suivent approximativement la loi suivante :

$$\lim_{n \rightarrow \infty} \Pr \{D_n^+ \sqrt{n} \leq d\} = 1 - e^{-2d^2} \quad 0 \leq d \leq \infty \quad (2.175)$$

En utilisant la statistique D_n , le test d'ajustement bilatéral consiste à rejeter l'hypothèse nulle au niveau de signification α si la valeur observée D_n est supérieure à $d_{n,\alpha}$ qui est défini comme suit :

$$\Pr \{D_n > d_{n,\alpha}\} = \alpha \quad (2.176)$$

Le test d'Anderson–Darling (A–D)

Le test K–S basé sur la statistique D_n est connu pour sa sensibilité aux différences entre F_n et F_0 au niveau des TIV ayant la valeur de fonction de répartition $F(h)$ au milieu entre les valeurs 0 et 1. Pour tenir compte aussi les distances entre F_n et F_0 en queue de la distribution, Anderson et Darling proposent d'ajouter une fonction de pondération dans la statistique de Cramer–von–Mise¹ :

$$A_n^2 = n \int_0^1 \left[\frac{F_n(h) - F_0(h)}{\sqrt{h(1-h)}} \right]^2 dh \quad (2.177)$$

$$= -n - \sum_{j=1}^n \frac{2j-1}{n} [\log U_{(j)} + (2n-2j+1) \log(1-U_{(j)})] \quad (2.178)$$

Où $U_{(i)} = F_H(H_{(i)})$. La loi de la statistique A_n^2 dépend du modèle F_0 .

1. La statistique de Cramer–von–Mise :

$$\omega_n^2 = \int_{-\infty}^{\infty} [F_n(h) - F_0(h)]^2 dF_0(h)$$

2.3 Comparaison des modèles

2.3.1 L'état de l'art de la comparaison des modèles pour le TIV

Chronologie

La comparaison entre les modèles de probabilité des TIV fait toujours l'objet d'études. Le but de la comparaison est de choisir les meilleurs modèles selon divers critères. Étant donné que les TIV sont très variables en fonction du contexte, la valeur d'un modèle dépend des données réelles. Les données dont les chercheurs disposent diffèrent au niveau de la précision, de la taille d'échantillon, de la continuité, etc.

Greenberg (1966) [55] compare le modèle $\ln\mathcal{N}$ et le modèle \mathcal{S} sur les données d'une autoroute en Californie (État-Unis). Le modèle \mathcal{S} fournit une valeur plus faible du χ^2 que le modèle $\ln\mathcal{N}$. Si ces deux valeurs de χ^2 sont supposées indépendantes, le test de Fisher qui prend en compte en plus les degrés de liberté atteste une adéquation équivalente des deux modèles, au niveau de signification de 40 %.

Grecco (1968) *et al.* [60] analysent des données recueillies voie par voie sur des contournements en Indiana (États-Unis). Les détecteurs pneumatiques permettent une précision de 0,01 s. La vitesse moyenne est environ de 80 km/h et le pourcentage de PL est de 13%. En outre, les débits moyens sont classés par tranche de 100 véh/h tandis que les TIV sont regroupés dans des classes d'amplitude d'une seconde, la dernière classe englobant tous les TIV supérieurs à 20 s. Grecco *et al.* comparent le modèle Γ_τ (comprenant la loi Gamma et la loi d'Erlang) et le modèle \mathcal{E}^* pour des débits de moins de 700 véh/h. Seul le modèle de \mathcal{E}^* n'est pas rejeté par le test du χ^2 au niveau de signification de 5%.

Buckley (1968) [57] étudie les modèles des TIV sur les données d'une autoroute à 6 voies en périphérie de Sydney (Australie). Les débits sont classés en 7 classes de 1 à plus de 31 véh/min (équivalent à 1860 véh/h). La largeur de chaque classe est de 5 véh/min (équivalant à 300 véh/h). Les paramètres sont estimés par la MM et les modèles sont testés par le test du χ^2 . Buckley compare les modèles $\mathcal{S}_\mathcal{N}$, $\mathcal{S}_\mathcal{G}$ et les modèles simples \mathcal{E} , \mathcal{E}_τ , Γ et Γ_τ . Il a constaté que :

- Seul le modèle $\mathcal{S}_\mathcal{G}$ convient aux TIV courts dans les cas où les débits ne sont pas faibles.
- Les deux modèles $\mathcal{S}_\mathcal{N}$, $\mathcal{S}_\mathcal{G}$ conviennent très bien pour les débits supérieurs à 21 véh/min (équivalant à 1200 véh/h).
- Les modèles simples ne conviennent que pour les débits faibles. Les modèles simples avec un paramètre de décalage \mathcal{G}_τ et \mathcal{E}_τ sont meilleurs que ceux qui n'en incluent pas (respectivement \mathcal{G} et \mathcal{E}).

Ashton (1971) [62] examine trois distributions : \mathcal{E}_τ , \mathcal{E}^* et \mathcal{S}_Γ . Deux paramètres du modèle \mathcal{E}_τ sont estimés par la MM et la méthode MV en même temps. Les trois paramètres du modèle \mathcal{E}^* (car $\tau_1 = \tau_2 = 0$) sont estimés par la méthode de Bartholomew [123]. Les paramètres du modèle $\mathcal{S}_\mathcal{G}$ sont estimés par la MM en utilisant le processus (2.149). Les données sont distinguées par voie et par période de temps (heure de pointe ou heure creuse). Le test du χ^2 montre de meilleurs ajustements pour le modèle \mathcal{E}^* par rapport aux autres modèles. Le modèle $\mathcal{S}_\mathcal{G}$ est moins bon à l'heure de pointe. En observant la figure fournie par Ashton dans son article, on constate nettement que les données étudiées sont distribuées de façon exponentielle. Lorsqu'on analyse les valeurs estimées, les paramètres du modèle \mathcal{E}^* et les paramètres du modèle $\mathcal{S}_\mathcal{G}$ sont très proches dans le cas où ces modèles \mathcal{E}^* , \mathcal{S}_Γ dégénèrent en une loi exponentielle.

Branston (1976) [58] compare les modèles suivants : $\mathcal{Q}_{\ln\mathcal{N}}$, $\mathcal{S}_\mathcal{N}$, \mathcal{S}_Γ , $\mathcal{Q}_\mathcal{N}$, $\mathcal{Q}_\mathcal{G}$ et \mathcal{L}_τ ($\tau = 0, 3$). Les TIV sont groupés en fonction du débit. La méthode d'estimation est une combinaison des méthodes MM et MC. Le test d'ajustement du χ^2 est utilisé en obligeant des

P-value minimales acceptables de 0,1. Branston conseille le modèle $\mathcal{Q}_{\ln\mathcal{N}}$ qui est le plus adéquat aux données avec des P-value de 0,3.

Khasnabis et Heimbach (1980) [66] étudient les TIV sur les routes à 2×2 voies en rase campagne. Neuf sites correspondant à 18 échantillons couvrent les débits de 80 à 632 véh/h/voie (dont 50% entre 100 et 150 véh/h/voie). Trois modèles sont examinés : \mathcal{E} , \mathcal{G}_τ et \mathcal{E}^* . Le modèle \mathcal{E}^* (modèle de Schuh) en s'appuyant sur les estimations des paramètres proposées par Grecco [60], est le meilleur au sens du test du χ^2 . Le modèle exponentiel convient bien pour le trafic fluide particulièrement pour les trois débits 122, 143 et 166 véh/h/voie.

Sullivan et Troutbeck (1994) [38] comparent les modèles \mathcal{C} et \mathcal{E}^* à partir des données routières en Australie. Le modèle \mathcal{E}^* est statistiquement plus adéquat notamment pour les TIV courts. Le modèle \mathcal{C} présente une facilité de calcul particulièrement appréciable pour les TIV élevés.

Luttinen (1996) [11] évalue plusieurs modèles parmi lesquels \mathcal{E} , \mathcal{E}_τ , \mathcal{G}_τ , \mathcal{L}_τ , Γ^* , \mathcal{C} , \mathcal{Q} et \mathcal{S} . Le modèle \mathcal{S} est le plus performant car il montre la flexibilité avec ses 5 paramètres et s'adapte le mieux aux données réelles (en Finlande) à tous niveaux : de la fonction de répartition, de la fonction aléatoire, du coefficient CV et de la relation de K-S².

Après avoir proposé la méthode MISE, Hoogendoorn (1997) et al. [47] la comparent avec les autres méthodes d'estimation. Les coefficients T_0 et N sont pré-choisis : $T_0 = 160$ s et $N = 800$. Le modèle Γ_τ est choisi pour la fonction f_1 . Les remarques suivantes sont mises en avant :

- La méthode MV est moins bonne dans l'estimation du paramètre de décalage et dans la reconstitution des débits.
- Par contre, si le paramètre de décalage est nul, la méthode MV est absolument la meilleure.
- La MM est moins bonne que les méthodes MV et MISE.
- La méthode MISE présente des avantages par rapport à la MM au niveau de la qualité des paramètres et par rapport à la méthode MV au niveau du temps du calcul.

Zhang *et al.* (2007) [37] évaluent la performance de différents modèles à partir des données sur une autoroute (États-Unis). Les TIV sont observés en différentes périodes (matin, midi, soir et nuit) et sur différents types de voies dont une voie HOV qui est dédiée aux véhicules comportant au moins 2 personnes. Un système de boucle inductive permet d'enregistrer les TIV avec une précision de l'ordre de 0,03 s. Les débits sont rangés entre 24 et 1680 véh/h/voie. Zhang *et al.* comparent les modèles suivants : Γ_τ , $\ln\mathcal{N}_\tau$, \mathcal{E}^* , \mathcal{Q}_Γ , $\mathcal{Q}_\mathcal{N}$. La méthode d'estimation est la MV pour les modèles simples et la méthode des moindres carrés pour les modèles *complexes* (combiné et de mélange). Ils constatent que :

- Le modèle LNM $\ln\mathcal{N}_\tau$ est meilleur que le modèle Γ_τ au niveau du test K-S. Cependant, la qualité de l'ajustement des modèles diminue considérablement sur les voies HOV.
- Les modèles \mathcal{Q}_Γ et $\mathcal{Q}_\mathcal{N}$ produisent des valeurs élevées du test K-S particulièrement sur les voies HOV. Le modèle \mathcal{E}^* convient pour tout type de voies.

La stabilité du trafic, beaucoup plus accentuée sur la voie HOV par rapport aux autres voies, explique des ajustements modestes de la majorité des modèles utilisés.

Discussion

- La performance des modèles de mélange \mathcal{Q} , \mathcal{S} est généralement confirmée,
- Le modèle $\ln\mathcal{N}$ peut être considéré comme le meilleur modèle simple,
- Toutefois, il manque une comparaison complète entre tous les modèles et entre les méthodes d'estimation.

2.3.2 Ensemble test des échantillons

Basé sur l'échantillonnage défini dans le chapitre précédent, un *ensemble test* d'échantillons est choisi pour tester la performance, la flexibilité des modèles pour la variable TIV. Le Tableau 2.1 représente les 25 échantillons sélectionnés à partir de différentes bases de données de la RN118 et de l'A6, de différentes méthodes d'échantillonnage sous des conditions exogènes différentes (période de la journée, condition de météo).

RN118 ¹					A6
Groupement (VL)			Raffinement		
jour-pluie	soir-sec	jour-sec	VL	VR	Horaire
véh/h	véh/h	%	véh/h	véh/h	véh/h
(300-400)	(200-300)	(2-4)	226 (N°104 ²)	1346 (N°52)	772 (L9b ³)
(700-800)	(700-800)	(8-10)	1820 (N°14)	2200 (N°46)	1291(M9b)
(1100-1200)	(1100-1200)	(14-16)	991 (N°149)	920 (N°129)	2251(R2)
(1500-1600)	(1500-1600)	(20-22)	1544 (N°114)	1207 (N°70)	
(1900-2000)	(1900-2000)				

1. Les échantillons de N°1 à N°25 sont placés de haut à bas, puis de gauche à droite,
2. L'ordre de l'échantillon raffiné obtenu du Site N°1 de la base de donnée de la RN118,
3. Le nom de l'enregistrement commence par la lettre L, M et R pour indiquer les voies Lente, Médiane et Rapide respectivement. Les lettres et numéros suivants indiquent l'ordre d'enregistrement.

TABLE 2.1 : L'ensemble test

Le Tableau 2.2 illustre les statistiques descriptives des échantillons de l'*ensemble test*. Il montre que l'ensemble test représente tous les états du trafic, de fluide à très dense qui sont caractérisés par le débit et le taux d'occupation.

Propriétés statistiques									
N°	Mean (s)	STD (s)	CV	Skewness	N°	Mean (s)	STD (s)	CV	Skewness
N°1	10.492	10.015	0.955	1.949	N°15	15.910	24.204	1.521	3.858
N°2	4.775	3.838	0.804	1.973	N°17	3.633	2.736	0.753	1.834
N°3	3.139	2.176	0.693	1.923	N°18	2.331	1.507	0.647	1.864
N°4	2.332	1.375	0.590	2.087	N°16	1.979	1.507	0.761	6.784
N°5	1.880	1.044	0.555	2.158	N°21	3.911	4.728	1.209	3.160
N°6	12.342	14.514	1.176	3.839	N°22	2.982	2.994	1.004	3.630
N°7	4.832	4.151	0.860	4.135	N°19	2.673	2.518	0.942	4.341
N°8	3.143	2.268	0.722	2.009	N°20	1.636	1.339	0.818	5.205
N°9	2.344	1.496	0.638	1.920	N°23	4.666	3.632	0.778	1.397
N°10	1.861	0.891	0.479	1.884	N°24	2.601	1.867	0.718	2.290
N°11	6.367	6.564	1.031	3.430	N°25	1.598	1.185	0.742	2.664
N°12	2.870	1.960	0.683	1.841					
N°13	1.945	1.220	0.628	2.255					
N°14	1.984	2.117	1.067	20.575					

TABLE 2.2 : Propriétés statistiques des TIV dans les échantillons de l'ensemble test

2.3.3 Le processus d'estimation

Les modèles simples sont estimés plus facilement par des dérivations analytiques (les modèles exponentiel et de Rayleigh) ou par résolution d'un système d'équations simples (la MMVM, MMVL pour le modèle Log-normal $\ln \mathcal{N}$, quelques méthodes d'estimation pour le modèle Γ). Cependant, pour les modèles complexes (combiné et de mélange), le nombre de paramètres demande toujours un processus d'optimisation.

Dans cette étude, un processus réalisé par programmation de Matlab 7.0 est proposé pour les modèles complexes. Il peut être utilisé avec les méthodes d'estimation différentes : la méthode MV, la méthode MC, la méthode MISE. Le programme est divisé en trois parties comme montré dans le cadre suivant :

1. **Étape 1. Initialisation** : Les valeurs initiales des paramètres sont tout d'abord estimées localement. Pour les modèles de mélange utilisant le modèle $\Gamma(\beta, \alpha)$ pour la fonction de densité g , l'ordre d'estimation des paramètres est : $\alpha \rightarrow \beta \rightarrow \theta \rightarrow \lambda \rightarrow \alpha$. Un exemple du processus d'estimation utilisant la fonction de vraisemblance $\ell(\Theta, h)$ est montré comme suit :

$$\begin{aligned}\alpha_* &= \arg \min_{\alpha} [-\ell(\alpha)] \\ \beta_0 &= \arg \min_{\beta} [-\ell(\beta | \alpha_*)] \\ \theta_0 &= \arg \min_{\theta} [-\ell(\theta | \alpha_*, \beta_0)] \\ \lambda_0 &= \arg \min_{\lambda} [-\ell(\lambda | \alpha_*, \beta_0, \theta_0)] \\ \alpha_0 &= \arg \min_{\alpha} [-\ell(\alpha | \beta_0, \theta_0, \lambda_0)] \\ \Theta_0 &= [\alpha_0, \beta_0, \theta_0, \lambda_0]\end{aligned}$$

Où

$$\ell(\Theta, h) = \log \left[\prod_{i=1}^n f(h_i | \Theta) \right] = \sum_{i=1}^n \log f(h_i | \Theta)$$

avec h_i les valeurs du TIV du véhicule i , et $\Theta = [\alpha, \beta, \theta, \lambda]$

2. **Étape 2. Optimisation** :

$$\Theta = \arg \min_{\Theta} [-\ell(\Theta | \Theta_0 = [\alpha_0, \beta_0, \theta_0, \lambda_0])]$$

$$\Delta = |\Theta_0 - \Theta|$$

3. **Étape 3. Comparaison** :

Si $\exists i \mid \Delta(i) > 0,0003$ alors $\Theta_0 = \Theta \rightarrow$ Étape 2
Si $\forall i \mid \Delta(i) < 0,0003$ alors Θ

TABLE 2.3 : Le processus d'estimation pour les modèles complexes

La performance et la facilité sont deux critères essentiels de choix d'un bon modèle [90]. En général, la performance concerne l'ajustement obtenu en utilisant les valeurs estimées des paramètres. Le niveau d'ajustement est mesuré principalement par le résultat des tests statistiques comme K-S et χ^2 . La facilité est la rapidité ou l'accessibilité de la méthode pour réaliser ou calculer l'estimation, donc, le temps d'exécution est un paramètre important. En conséquence, à chaque modèle de TIV correspondent quelques bonnes méthodes d'estimation dont les qualités dépendent aussi des données empiriques.

Malgré la différence des échelles de valeur des paramètres des modèles, le critère d'arrêt de 0,0003 semble convenir pour la modélisation du TIV. L'étape 1 est nécessaire car la méthode MISE fournira de mauvais résultats à l'ajustement des TIV longs si les valeurs initiales des paramètres sont choisies de façon arbitraire. Dans le cas où l'étape 1 est intégrée, l'étape 3 est utilisée afin de mesurer l'efficacité des méthodes d'estimation. Prenons un exemple, la méthode MV donne l'estimation finale généralement après 4 itérations comparatives. Il est à noter que le temps d'estimation des paramètres dépend aussi des données.

Le processus de comparaison entre les modèles se base tout d'abord sur l'ajustement, c. à. d le modèle acceptable doit nécessairement passer le test K-S et le test du χ^2 . D'autres critères comme le temps d'estimation t_e , le temps d'ajustement t_a , le test A-D seront ensuite pris en compte pour mesurer l'efficacité du modèle et de la méthode d'estimation correspondante.

2.3.4 Modèles simples

Cinq modèles simples sont comparés, à savoir le modèle LNM (le modèle Log-normal $\ln \mathcal{N}_\tau$ avec le paramètre de décalage τ), le modèle de Rayleigh \mathcal{R} , le modèle exponentiel décalé \mathcal{E}_τ (DNED), le modèle Gamma Γ et le modèle Pearson-III Γ_τ . Ils ont soit un paramètre (Rayleigh), soit deux paramètres (DNED, Gamma) ou trois paramètres (LNM et Pearson-III).

Les valeurs $\mathcal{D}_1 = D_1/\sqrt{n}$ and $\mathcal{D}_2 = D_2/\sqrt{n}$ sont les seuils au delà desquels l'ajustement du modèle sera rejeté aux niveaux de signification respectifs de 1 % et 5 %.

Statistique K-S $\times 10^{-4}$												
N°	N°1	N°2	N°5	N°6	N°10	N°14	N°15	N°19	N°20	N°21	N°23	N°24
\mathcal{Q} ou \mathcal{T}_6	350	750	1950	250	1950	21 %	291	330	823	836	772	1291
LNM	3098	861	1027	4196	5517	5403	2686	613	701	600	1222	2277
\mathcal{R}	1459	1302	1969	2862	2562	5048	5124	4906	4388	5296	1689	1401
\mathcal{E}_τ	308	258	547	805	872	2830	1314	2086	1849	1166	817	499
Γ	523	657	860	790	597	956	1028	1199	1196	1359	464	578
Γ_τ	484	545	780	501	596	927	958	1163	1170	1237	341	314
$\mathcal{D}_1, 1\%$	621	502	372	530	325	297	393	297	297	297	586	438
$\mathcal{D}_2, 5\%$	518	419	310	442	271	248	328	248	248	248	489	365

TABLE 2.4 : Les statistiques K-S des modèles simples, Numéro 1

Le Tableau 2.4 et le Tableau 2.5 fournissent les valeurs de statistique K-S, ainsi que les seuils d'ajustement. Les deux tableaux présentent respectivement les cas où le modèle LNM donne ou non de bons ajustements aux données. La grande différence entre les valeurs des tests K-S dans les deux tableaux montre que le modèle LNM ne peut fournir de bons résultats que dans des échantillons particuliers.

Dans les échantillons pour lesquels le modèle LNM donne de faibles valeurs de la statistique K-S, inférieures aux seuils $\mathcal{D}_1, \mathcal{D}_2$, le modèle LNM donne de bons ajustements pour les échantillons aux trafics intermédiaires, de débit entre 800 et 1700 véh/h dans la période du jour, de débit de 700 à 1600 véh/h dans la période du soir, aux alentours du taux d'occupation de 10% ou du débit de 1000 à 1500 véh/h dans les échantillons raffinés. Pour les états du trafic plus dense ou plus fluide, le modèle LNM est le plus mauvais des modèles simples.

Le modèle Rayleigh \mathcal{R} ne convient à aucun échantillon tandis que le modèle \mathcal{E}_τ n'est pas rejeté seulement pour les échantillons N°1 et N°2. Le modèle \mathcal{E}_τ ne fournit pas non plus

un bon ajustement pour l'échantillon N°6 qui représente un trafic très faible de 250 véh/h. Pourtant, dans la majorité des cas, le modèle \mathcal{E}_τ modélise bien la pente descendante de la distribution des TIV longs (cf. Figure 2.13-b).

Il n'y a pas une distinction claire entre le modèle \mathcal{E}_τ et le modèle Γ au niveau du test statistique K-S. Cependant, le modèle Pearson-III Γ_τ est toujours meilleur que le modèle Γ . Il est à noter que le premier est une généralisation du deuxième en ajoutant un paramètre de décalage τ . Les valeurs de la statistique K-S du modèle Γ sont inférieures au seuil \mathcal{D}_2 dans deux échantillons (N°1 et N°23) alors que les valeurs de la statistique K-S du modèle Pearson-III Γ_τ sont inférieures à ce seuil dans quatre échantillons (N°1, N°6, N°23 et N°24). Ces échantillons correspondent soit à des trafics fluides, soit à des trafics intermédiaires.

	Statistique K-S $\times 10^{-4}$											
N°	N°3	N°4	N°7	N°8	N°9	N°11	N°12	N°13	N°16	N°17	N°18	N°22
\mathcal{Q} ou \mathcal{T}_6	1150	1550	750	1150	1550	3%	9%	15%	714	1035	1485	1207
LNM	280	180	145	250	208	346	293	346	404	229	220	333
\mathcal{R}	3401	3276	3935	3602	3477	4514	3412	3491	3991	3512	3440	4558
\mathcal{E}_τ	486	702	597	387	464	683	523	533	1511	585	470	1089
Γ	640	577	562	606	751	757	652	960	891	466	790	1063
Γ_τ	580	523	469	535	559	619	605	770	827	415	598	892
$\mathcal{D}_1, 1\%$	297	297	297	297	297	297	297	297	297	297	297	317
$\mathcal{D}_2, 5\%$	248	248	248	248	248	248	248	248	248	248	248	265

TABLE 2.5 : Les statistiques K-S des modèles simples, Numéro 2

En somme, le modèle LNM est considéré comme le meilleur modèle simple malgré sa capacité limitée. Le modèle Pearson-III, étant une généralisation du modèle Gamma Γ , du modèle \mathcal{E}_τ , et du modèle exponentiel \mathcal{E} , fournit de bons ajustements pour quelques échantillons. En général, les modèles simples ne conviennent pas aux échantillons de TIV dans beaucoup d'états du trafic.

2.3.5 Modèles combinés

Les quatre modèles combinés examinés sont le modèle Hyperexponentiel \mathcal{E}^* , le modèle Hyperayleigh \mathcal{R}^* , le modèle Hypergamma Γ^* et le modèle Hyperlognormal $\ln\mathcal{N}^*$. Ces modèles ont un paramètre λ de plus par rapport au modèle f_1 avant la combinaison. En particulier, le modèle f_1 avant la combinaison n'est pas décalé par le paramètre τ . En conséquence, les modèles Hyperexponential \mathcal{E}^* et Hyperayleigh \mathcal{R}^* ont deux paramètres tandis que les modèles Hypergamma Γ^* et Hyperlognormal $\ln\mathcal{N}^*$ ont trois paramètres à estimer.

Sauf pour l'échantillon N°11 de taux d'occupation \mathcal{T}_6 très bas, le modèle Hyperexponentiel \mathcal{E}^* est rejeté dans tous les autres échantillons. La forme de la distribution du modèle est similaire à celle du modèle \mathcal{E}_τ (cf. Figure 2.13-a). En combinant une deuxième distribution exponentielle à un modèle exponentiel, l'ajustement global du modèle Hyperexponentiel \mathcal{E}^* est considérablement amélioré; cependant, l'ajustement des TIV longs est moins bon que celui obtenu par le modèle DNED \mathcal{E}_τ . Il est noté que la méthode MC est meilleure que la méthode MV pour l'estimation des paramètres du modèle Hyperexponentiel \mathcal{E}^* .

Dans 21 sur 25 échantillons, le modèle Hyperayleigh \mathcal{R}^* donne un meilleur ajustement que le modèle Hyperexponentiel \mathcal{E}^* . En outre, le modèle Hyperayleigh \mathcal{R}^* n'est pas rejeté dans quatre échantillons de TIV (N°1, N°6, N°23 and N°24) au niveau de signification

de 5% (cf. Table 2.7), et dans cinq échantillons au niveau 1% (quatre au niveau 5% et l'échantillon N°2).

En utilisant la méthode MV, le modèle Pearson-III Γ_τ est moins bon que le modèle Hypergamma Γ^* seulement dans quatre échantillons (N°6, N°17, N°23 and N°24). Notons que pour deux (N°6 et N°17) de ces quatre échantillons, les différences entre les deux modèles sont très faibles. En utilisant la méthode MC, le modèle Pearson-III Γ_τ est moins bon que le modèle Hypergamma Γ^* dans sept échantillons (N°3, N°7, N°9, N°17, N°18, N°23 et N°24).

Le modèle Hyperlognormal $\ln\mathcal{N}^*$ est étudié pour la première fois. Il est meilleur que le modèle LNM dans les cas où le LNM est rejeté par les test K-S. Dans le cas contraire, le modèle Hyperlognormal $\ln\mathcal{N}^*$ est moins bon que le modèle LNM.

Modèle	Statistique K-S $\times 10^{-4}$											
	N°2	N°4	N°5	N°8	N°9	N°10	N°14	N°16	N°17	N°19	N°20	N°25
\mathcal{E}^*	618	1183	1247	844	536	1903	1456	1110	466	779	1212	999
\mathcal{R}^*	491	594	863	467	488	1244	1046	773	417	732	1055	393
Γ^*	484	396	531	533	454	447	380	387	424	232	390	363
$\ln\mathcal{N}^*$	310	213	495	396	309	288	276	345	228	212	237	292
$\mathcal{D}_1, 1\%$	502	297	372	297	297	325	297	297	297	297	297	343
$\mathcal{D}_2, 5\%$	419	248	310	248	248	271	248	248	248	248	248	286

TABLE 2.6 : Les statistique K-S des modèles combinés, Numéro 1

Le modèle Hypergamma Γ^* est meilleur que le modèle Hyperlognormal $\ln\mathcal{N}^*$ seulement dans quatre échantillons (N°1, N°6, N°13 and N°15) (cf. Table 2.7). Cependant, les différences entre les valeurs de la statistique K-S dans les échantillons N°6 et N°13 entre les deux modèles sont très faibles (respectivement de 0,0005 et 0,0002). Le modèle Hyperlognormal $\ln\mathcal{N}^*$ n'est pas rejeté au niveau de signification de 1% et 5% respectivement dans 15 et 11 échantillons. De plus, les valeurs de la statistique K-S des autres échantillons sont proches du seuil \mathcal{D}_1 .

En bref, le modèle Hyperlognormal $\ln\mathcal{N}^*$ est considéré comme le meilleur modèle combiné et la combinaison d'un modèle avec la tendance exponentielle améliore toujours l'ajustement global par rapport au modèle originel f_1 .

Modèle	Statistique K-S $\times 10^{-4}$											
	N°1	N°3	N°6	N°7	N°11	N°12	N°13	N°15	N°18	N°21	N°22	N°23
\mathcal{E}^*	1237	523	585	345	297	534	1123	1131	848	360	460	801
\mathcal{R}^*	404	614	224	397	420	531	532	487	447	421	351	390
Γ^*	323	578	515	390	509	517	501	240	566	447	410	464
$\ln\mathcal{N}^*$	377	322	520	182	373	319	521	348	442	368	265	455
$\mathcal{D}_1, 1\%$	621	297	530	297	297	297	297	393	297	297	317	586
$\mathcal{D}_2, 5\%$	518	248	442	248	248	248	248	328	248	248	265	489

TABLE 2.7 : Les statistiques K-S des modèles combinés, Numéro 2

2.3.6 Modèles de mélange

Les modèles GQM \mathcal{Q} et SPM \mathcal{S} utilisant la fonction g comme distribution des TIV suiveurs sont appelés les modèles g -GQM et g -SPM, par exemple les modèles gamma-

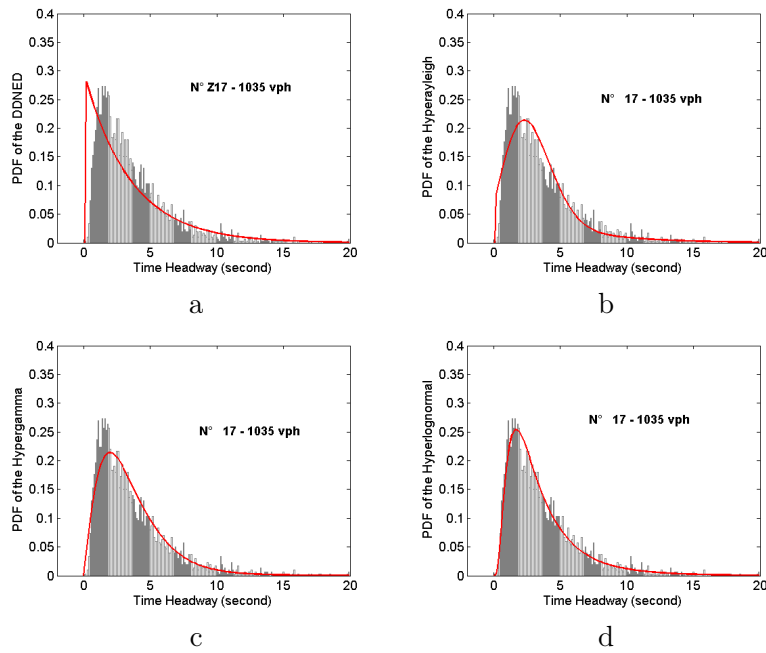


FIGURE 2.13 : La densité des modèles combinés

GQM (Γ -GQM), bêta-SPM, etc. Les deux modèles Γ -GQM et Γ -SPM fournissent de bons ajustements pour tous les échantillons au niveau de signification de 1 % et pour les 24 sur 25 échantillons au niveau de signification de 5 %. Le Tableau 2.8 montre que les modèles gamma-SPM et Γ -GQM sont équivalents au niveau du test statistique K-S.

Le Tableau 2.9 montre aussi l'équivalence entre le modèle Γ -GQM et gamma-SPM au niveau du test statistique χ^2 . Il est clair que les valeurs de la statistique, le degré de liberté et les P-valeurs du test χ^2 sont très proches entre ces deux modèles de mélange.

Valeurs de la Statistique K-S $\times 10^{-4}$											
N°	Γ -GQM	Γ -SPM	5%	N°	Γ -GQM	Γ -SPM	5%	N°	Γ -GQM	Γ -SPM	5%
1	292	292	518	10	144	144	271	19	199	199	248
2	<u>201</u>	<u>147</u>	419	11	<u>200</u>	<u>201</u>	248	20	197	197	248
3	141	141	248	12	87	87	248	21	210	210	248
4	108	108	248	13	132	132	248	22	<u>229</u>	<u>225</u>	265
5	101	101	310	14	285	285	248	23	348	348	489
6	213	213	442	15	252	252	328	24	169	169	365
7	135	135	248	16	155	155	248	25	227	227	343
8	<u>108</u>	<u>109</u>	248	17	149	149	248				
9	109	109	248	18	87	87	248				

TABLE 2.8 : Statistique K-S des modèles Γ -GQM et Γ -SPM

Pourtant, il y a deux faiblesses du modèle Γ -SPM par rapport au modèle Γ -GQM :

- Le temps d'ajustement t_a pour le modèle Γ -SPM est beaucoup plus long que celui du modèle Γ -GQM. Ce temps est le temps nécessaire pour calculer les test K-S, χ^2 et A-D après avoir estimé les paramètres,
- Les valeurs du test A-D dans les échantillons (N°1, N°2, N°6, N°10, N°15, N°22, N°23, N°24, N°25) sont élevées pour le modèle Γ -SPM tandis que dans les autres échantillons, elles sont proches des valeurs correspondantes du modèle Γ -GQM.

Valeurs de la statistique du test χ^2											
N°	χ_{GQM}^2 ¹	χ_{SPM}^2	P_{GQM} ²	P_{SPM}	DL ³	N°	χ_{GQM}^2 ¹	χ_{SPM}^2	P_{GQM} ²	P_{SPM}	DL
1	1.01	1.01	- ⁴	-	0	14	115.23	115.26	<u>0.0000</u>	<u>0.0000</u>	34
2	46.13	46.29	0.7681	0.7626	54	15	59.09	59.11	0.1775	0.1772	50
3	75.91	75.88	0.2939	0.2946	70	16	57.12	57.10	0.1480	0.1485	47
4	50.18	50.13	0.4662	0.4683	50	17	66.54	66.55	0.8925	0.8922	82
5	24.17	24.16	0.9157	0.9160	35	18	41.83	41.83	0.8657	0.8656	53
6	0.38	0.38	-	-	0	19	71.21	71.23	<u>0.0129</u>	<u>0.0128</u>	47
7	93.41	93.43	0.6661	0.6655	100	20	79.28	79.34	0.0003	0.0003	41
8	83.08	83.06	0.1546	0.1550	71	21	103.74	103.76	0.1895	0.1891	92
9	70.39	70.43	0.0552	0.0548	53	22	77.37	77.46	0.1820	0.1795	67
10	46.55	46.29	0.0917	0.0961	35	23	43.97	43.78	0.3071	0.3141	40
11	158.01	158.06	<u>0.0157</u> ⁵	<u>0.0156</u>	122	24	37.09	37.09	0.8494	0.8732	47 (48) ⁶
12	63.98	63.94	0.4771	0.4787	64	25	60.18	60.26	<u>0.0124</u>	<u>0.0122</u>	38
13	65.91	65.93	<u>0.0227</u>	<u>0.0226</u>	45						

1. Valeur de la statistique χ^2 du modèle Γ -GQM

2. P-valeur du modèle Γ -GQM

3. DL : degré de liberté

4. - : Aucun résultat obtenu car DL=0

5. Chiffre souligné : le modèle est rejeté au niveau de signification de 5 %

6. 47 et 48 sont les DL des modèles Γ -GQM et gamma-SPM respectivement.

TABLE 2.9 : Valeurs de la statistique du test du χ^2 des modèles Γ -GQM et Γ -SPM

Les modèles g -GQM

Les modèles normal-GQM et normal-SPM ont été étudiés par Branston [58] et Buckley [124]. Cependant, l'utilisation de la loi normale pour la fonction g n'a pas donné de meilleurs résultats que l'utilisation des modèles Gamma Γ et log-normal $\ln \mathcal{N}$. Le modèle bêta-GQM a été étudié par Hoogendoorn [78], mais ce modèle n'était pas un bon choix car des erreurs de calcul non négligeables interviennent lorsqu'on manipule des fonctions utilisant le modèle Bêta. De plus, les modèles normal et Bêta ont une capacité limitée à décrire tout l'intervalle de définition de la variable TIV ; la loi normale en effet accepte des valeurs négatives tandis que la loi Bêta n'est définie que dans un intervalle défini.

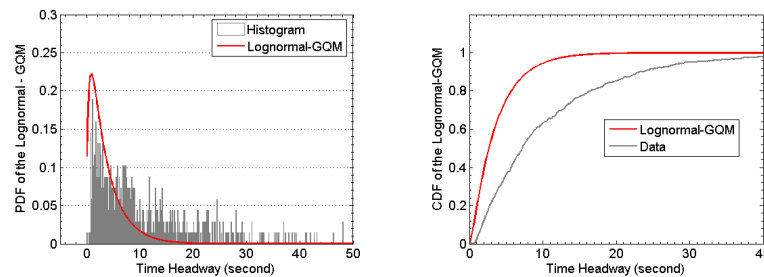


FIGURE 2.14 : Le modèle lognormal-GQM

Le modèle lognormal-GQM est calibré par la méthode MV mais le temps d'estimation t_e est très long. Le nombre d'itérations est élevé et les paramètres obtenus, par exemple θ , coïncident les valeurs extrêmes des bornes. Par conséquent, la méthode MV semble n'être pas faisable, non seulement pour le modèle LNM [11, 125] mais aussi pour le modèle

lognormal-GQM. En utilisant la méthode MC, le temps d'estimation t_e est aussi exorbitant mais les résultats ne sont pas bons (cf. Figure 2.14).

Il reste à comparer les modèles rayleigh-GQM et Γ -GQM. Il est à noter que le modèle rayleigh-GQM ne fournit aucun résultat pour l'échantillon N°17. La Figure 2.15 illustre clairement la supériorité du modèle Γ -GQM sur le modèle rayleigh-GQM qui est rejeté au niveau de signification de 5%, et même au niveau de 1% dans la majorité des échantillons.

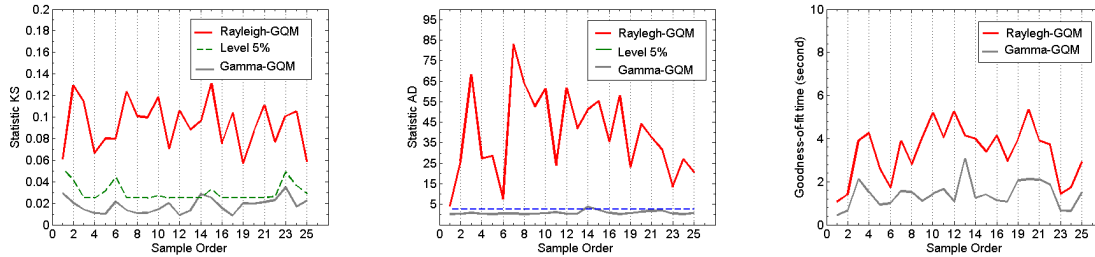


FIGURE 2.15 : La comparaison entre les modèles Γ -GQM et le modèle \mathcal{R} -GQM

En somme, le modèle Γ -GQM est considéré comme le meilleur modèle parmi tous les modèles étudiés jusqu'ici pour le TIV. Ce modèle de mélange fournit de très bons ajustements pour tous les échantillons d'autant plus qu'il ne présente pas de difficulté d'estimation.

Les méthodes d'estimation pour le modèle Γ -GQM

Dans cette section, huit méthodes d'estimation pour le modèle Γ -GQM sont comparées : la méthode MV (ang. MLE), la méthode MISE pour la fonction de répartition F (appelé « Min F » dans les figures), la méthode MISE pour la densité f (appelé « Min f » dans les figures), la méthode MISE pondérée pour la fonction F , la méthode MISE pondérée pour la densité f , la méthode MC (ang. MCS), la MM et la MM modifiée.

Les méthodes MISE pondérées utilisent la fonction de pondération $q = \sqrt{\xi}$ dans le domaine de fréquence caractérisé par la variable ξ . De plus, toutes les méthodes MISE sont réalisées en utilisant les paramètres clefs $T_0 = 200$ et $N = 1000$ (cf. [78] où $T_0 = 160$ s et $N = 800$ sont utilisés).

Tout d'abord, la MM et la MM modifiée ne conviennent pas parce que les résultats obtenus sont inadéquats, comme le paramètre θ au delà de l'intervalle $[0, 1]$, ou des valeurs négatives des autres paramètres. La méthode MISE pondérée pour F est la moins bonne des méthodes MISE. La Figure 2.16 montre les comparaisons entre les méthodes au niveau de la valeur de la statistique K-S.

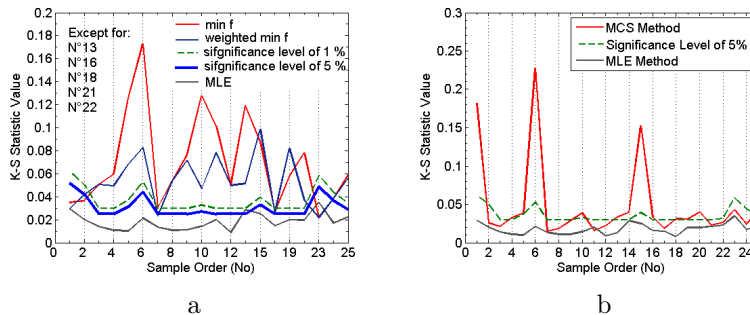


FIGURE 2.16 : Comparaison entre les méthodes d'estimation pour le modèle Γ -GQM

Statistique K-S $\times 10^{-4}$							
N°	MISE f	MISE F	MLE	N°	MISE f	MISE F	MLE
1	351	201	292	14	1194	155	285
2	366	198	201	15	878	369	252
3	507	172	141	16	1016	156	155
4	597	119	108	17	271	165	149
5	1259	92	101	18	<u>736</u>	<u>2061</u>	80
6	<u>1730</u>	<u>9268</u>	213	19	<u>575</u>	<u>10000</u>	199
7	303	92	135	20	779	166	197
8	537	105	108	21	893	273	210
9	<u>756</u>	<u>850</u>	109	22	<u>971</u>	<u>2487</u>	229
10	1275	101	144	23	213	256	348
11	1011	255	200	24	386	131	169
12	499	86	87	25	590	201	227
13	1075	121	132				

TABLE 2.10 : Valeurs de la statistique K-S du modèle Γ -GQM utilisant les méthodes MV, MISE pour f et F

On constate que :

- La fonction de pondération $q = \sqrt{\xi}$ améliore l'ajustement de la méthode MISE pour f . Cependant, cette amélioration ne suffit pas pour éviter le rejet par le test K-S au niveau de signification de 5 %. Aussi, la méthode MV est meilleure que les méthodes MISE pour f et MISE pondérée pour f (cf. Figure 2.16-a).
- Sauf sur les cinq échantillons (N°6, N°9, N°18, N°19 and N°22), la méthode MISE pour F est bien meilleure que les méthodes MISE pour f et MISE pondérée pour f (cf. Figure 2.16-b).
- La méthode MC est une bonne méthode d'estimation pour le modèle Γ -GQM. En général, les valeurs du test K-S obtenues par cette méthode sont proches du seuil \mathcal{D}_2 de signification de 5 %. Cependant, par rapport à la méthode MV, la méthode MC est toujours moins bonne (cf. Figure 2.16-c). De plus, le temps d'estimation t_e de la méthode MC est plus long que celui de la méthode MV (cf. Figure 2.17).
- Dans 15 des 25 échantillons examinés, la méthode MISE pour F est légèrement meilleure que la méthode MV. Cependant, dans quatre autres échantillons, les résultats du test K-S obtenus par la méthode MISE pour F sont au contraire très mauvais (cf. Tableau 2.10).

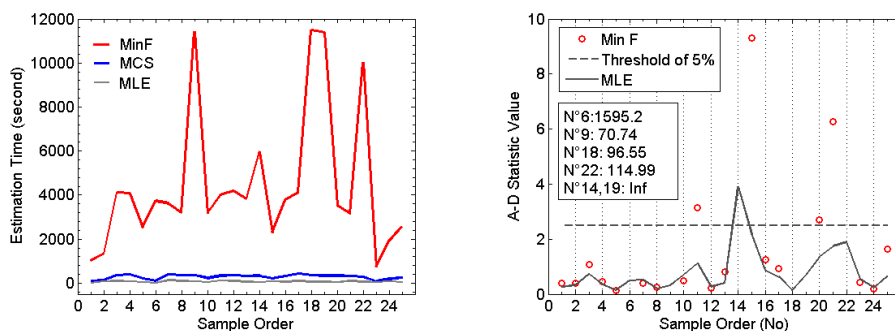


FIGURE 2.17 : Comparaison entre les méthodes MV et MISE pour F

La Figure 2.17 illustre la comparaison entre la méthode MV et la méthode MISE pour F avec d'autres critères de comparaison. Il est observé que le temps d'estimation t_e de la méthode MISE pour F est très long. Ce résultat se produit parce que la phase d'initialisation de l'étape 1 du processus d'estimation proposé (cf. Tableau 2.3.3) et le seuil d'arrêt Δ dans l'étape 3 sont ajoutées. L'étape 3 est importante et ne se termine que lorsque les résultats sont stables. Le temps d'estimation dans l'étape 2 de la méthode MISE pour F est plus court que celui de la méthode MV mais les résultats obtenus ne sont pas consistants. En conséquence, la méthode MV est finalement plus rapide et plus robuste que la méthode MISE pour F . Les valeurs des tests K-S et A-D confirment aussi la performance de la méthode MV dans tous les cas.

En conclusion, la méthode MV est considérée comme la meilleure pour estimer les paramètres du modèle gamma-GQM. Cette méthode utilisée avec le processus d'estimation proposé est rapide, efficace, fournit de très bons ajustements pour la majorité des échantillons des TIV et induit aussi des bons résultats au niveau des autres critères statistiques.

2.3.7 Paramètre de décalage

Le paramètre de décalage τ joue un rôle très important pour tous les modèles simples de TIV. Ce paramètre améliore considérablement la performance du modèle exponentiel \mathcal{E} car il évite une sur-estimation de la fréquence des TIV aux alentours de la valeur zéro. L'amélioration de l'ajustement en utilisant τ pour le modèle exponentiel \mathcal{E} a été confirmée par Gerlough [90].

Au niveau du trafic intermédiaire, May and Krbalek [114, 84] ont proposé d'utiliser le modèle Pearson-III Γ_τ avec le paramètre de décalage τ . La section précédente a montré que le modèle Γ_τ était meilleur que le modèle Γ dans la majorité des échantillons. De plus, en utilisant les deux méthodes MV et MC, le modèle Pearson-III Γ_τ reste meilleur que le modèle Gamma Γ dans la majorité des cas (cf. Figure 2.18). Dans ces cas, la méthode MC donne de meilleurs résultats que la méthode MV sauf seulement pour les deux échantillons N°1 et N°6.

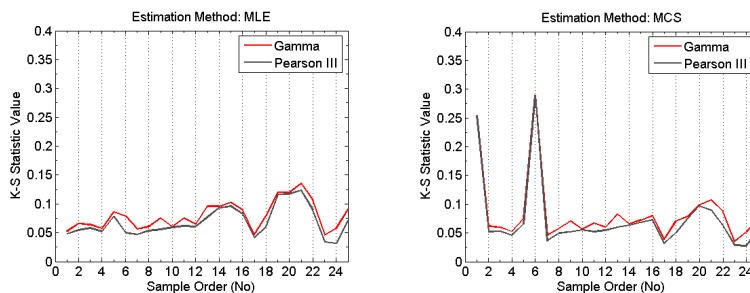


FIGURE 2.18 : Les statistiques K-S des modèles Pearson-III Γ_τ et Gamma Γ

Le modèle log-normal $\ln \mathcal{N}$ (sans paramètre τ) ne représente pas bien les données. Cependant, l'ajout du paramètre τ transforme ce modèle : il devient le meilleur des modèles simples, le modèle LNM $\ln \mathcal{N}_\tau$, meilleur même que les modèles combinés.

Le paramètre τ dans les modèles combinés et les modèles de mélange est utilisé pour décaler toute la courbe de distribution à la droite de l'abscisse. Schuhl [47], Grecco [60], Griffiths [120] ont utilisé deux paramètres de décalage respectivement pour les deux composantes d'un modèle combiné. Ici, seul un paramètre de décalage est utilisé pour toute la distribution après avoir combiné ou mélangé les composantes (non décalées). L'addition du paramètre τ dans les trois modèles Γ -GQM, Hypergamma Γ^* et Hyperlognormal $\ln \mathcal{N}^*$ les transforme respectivement en modèles Γ -GQM décalé, Hypergamma décalé Γ_τ^*

et HyperLNM $\ln \mathcal{N}_\tau^*$ qui ont maintenant cinq paramètres à estimer.

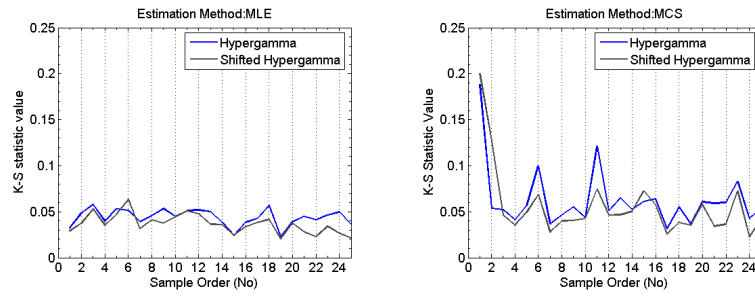


FIGURE 2.19 : Les statistiques K-S des modèles Γ_τ^* et Γ^*

Tant avec un calibrage issu de la méthode MV qu’avec celui issu de la méthode MC, les valeurs de la statistique K-S sont plus petites pour le modèle Hypergamma décalé Γ_τ^* que pour le modèle Hypergamma Γ^* (cf. Figure 2.19). Le paramètre τ a donc amélioré l’ajustement du modèle Hypergamma dans la majorité des échantillons (sauf seulement pour l’échantillon N°6 lorsque le calibrage est effectué avec la méthode MV). De plus, la méthode MV est meilleure que la méthode MC pour l’estimation des paramètres du modèle Hypergamma décalé Γ_τ^* .

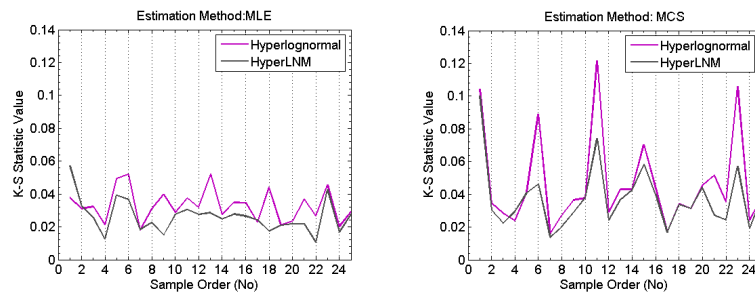


FIGURE 2.20 : Les statistiques K-S pour les modèles $\ln \mathcal{N}_\tau^*$ et $\ln \mathcal{N}^*$

En utilisant les deux méthodes MV et MC, les valeurs de la statistique K-S obtenues pour le modèle HyperLNM $\ln \mathcal{N}_\tau^*$ sont inférieures à celles du modèle Hyperlognormal $\ln \mathcal{N}^*$. Dans ce cas aussi, le paramètre de décalage τ améliore la performance du modèle Hyperlognormal $\ln \mathcal{N}^*$. De plus, la méthode MV est préférable pour estimer les paramètres des modèles Hyperlognormal $\ln \mathcal{N}^*$ et HyperLNM $\ln \mathcal{N}_\tau^*$.

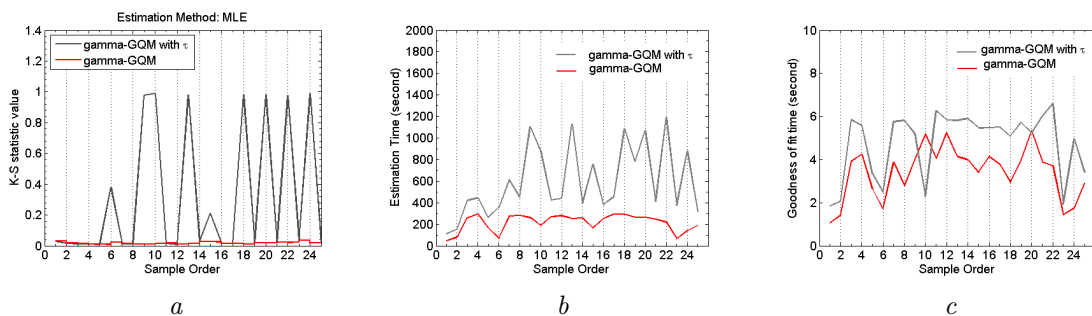


FIGURE 2.21 : (a) Les statistiques K-S, (b) Temps d’estimation, (c) Temps d’ajustement des modèles Γ -GQM décalé et Γ -GQM

Toutefois, en utilisant la méthode MV, l’utilisation du paramètre τ réduit la performance du modèle Γ -GQM. De plus, le temps d’estimation t_e et le temps d’ajustement t_a

pour le modèle Γ -GQM décalé sont considérablement plus longs (cf. Figure 2.21 et Figure 2.22).

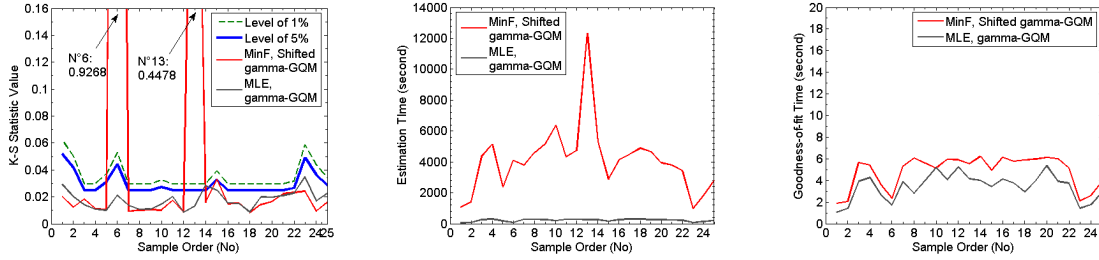


FIGURE 2.22 : Le modèle Γ -GQM décalé estimé par la méthode MISE pour F et le modèle Γ -GQM estimé par la méthode MV

Dans tous les cas, le modèle Γ -GQM (sans paramètre de décalage τ) associé avec le processus d'estimation proposé, est le meilleur choix pour calibrer les distributions de TIV.

2.4 Discussion et développement des modèles

2.4.1 Modèle simple et modèle combiné

Tandis que le modèle combiné est plus facile à exécuter que le modèle de mélange, le premier rencontre deux problèmes majeurs lors du calibrage :

1. Quand $\theta = 0$, le modèle dégénère en une distribution exponentielle dans laquelle les densités des TIV courts sont extrêmement sur-estimées.
2. La combinaison de deux composantes dans les modèles combinés entraîne des valeurs de la densité anormales qui ne conviennent pas pour le TIV *zéro* ou le TIV minimum.

Rappelons qu'en réalité, la distribution de tous les TIV est de forme exponentielle seulement pour les TIV longs. Dans tous les états du trafic, même ceux de très faible débit, des TIV très courts sont aussi obtenus et leur fréquence devrait être aux alentours de zéro, idéalement pour un modèle de TIV soit :

$$\begin{cases} f(0) = 0 \\ f(\epsilon) > 0 \quad \forall \epsilon > 0 \end{cases} \quad (2.179)$$

Soit la fonction f une fonction de densité d'un modèle combiné, on a :

$$f(h) = \theta g(h) + (1 - \theta) \lambda e^{-\lambda h} \quad (2.180)$$

Alors que

$$f(0) = \theta g(0) + (1 - \theta) \lambda \quad (2.181)$$

Sauf dans le cas $\theta = 1$ où le modèle combiné dégénère en modèle de densité $g(h)$ et $g(0) = 0$, la valeur de $f(0)$ d'un modèle combiné est différente de zéro. Dans le cas d'utilisation d'une fonction g telle que $g(0) > 0$ comme dans le modèle Hyperexponentiel \mathcal{E}^* , la fonction $f(0)$ n'est jamais égale à 0. Cette estimation est donc inacceptable tant dans la réalité que dans la théorie. La Figure 2.24 illustre les inconvénients de certains modèles au niveau TIV zéro. On observe que le modèle Hyperlognormal $\ln \mathcal{N}^*$ est le moins affecté, mais au contraire, pour les modèles Hypergamma Γ^* et Hyperrayleigh \mathcal{R}^* , les courbes de densité sont visiblement discontinues au niveau du TIV zéro.

La deuxième condition « idéale » dans la remarque 2.179 incite également à ne pas utiliser le paramètre de décalage τ . En effet, pour les TIV ϵ tels que $0 < \epsilon < \tau$, la densité $f(\epsilon) = 0$.

La Figure 2.23 montre les diverses formes de distribution aux alentours du TIV zéro. Les différences entre elles proviennent de la valeur de la dérivée df/dh . En raison de la sécurité routière et du fait qu'un TIV zéro signifie une réelle collision, il semble exister un intervalle $(0, I_*)$ dans lequel la présence d'un TIV est rare, d'autant plus qu'il faut compter le temps de passage sur un point donné de la longueur du véhicule lui-même. En conséquence, l'évolution de la densité de TIV au voisinage de la valeur de TIV zéro est peu importante, ce qui conduit à la condition préférable pour les fonctions de densité de TIV que :

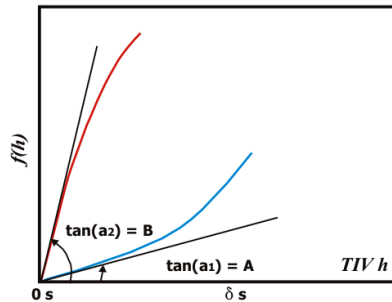


FIGURE 2.23 : La dérivée de la densité au TIV zéro

$$\frac{df}{dh} \Big|_{h=0} \leq A < +\infty \quad (2.182)$$

Plus particulièrement que la dérivée de la fonction de densité g des TIV contraints soit :

$$\frac{dg}{dh} \Big|_{h=0} \leq B < +\infty \quad (2.183)$$

Où A, B sont certains seuils non négatifs. En réalité, il semble que $A = 0$ et $B = 0$ reflètent bien les distributions des TIV. Dans le cas du modèle combiné, la dérivée de la fonction f est comme suit :

$$\frac{df}{dh} \Big|_{h=0} = \theta B - (1 - \theta) \lambda^2 \quad (2.184)$$

Les inconvénients analysés ci-dessus causent de mauvais ajustements principalement dans les échantillons de trafic faible ou très faible (cf. la comparaison des deux Figures 2.24 et 2.25). On a aussi remarqué que l'utilisation du paramètre de décalage τ pourrait modifier la dérivée $df/dh \Big|_{h=0}$, et par conséquent changer les valeurs des paramètres de forme et/ou d'échelle des modèles.

Le tableau suivant représente les valeurs de $f(0)$ (ou de $g(0)$) et de la dérivée $df/dh \Big|_{h=0}$ pour différents modèles simples et combinés.

On constate que :

- Pour le modèle exponentiel, les deux valeurs $f(0)$ et $df/dh \Big|_{h=0}$ ne sont pas égales à 0, la dérivée prend même une valeur négative.
- Pour le modèle Rayleigh, la moitié des conditions « idéales » est remplie. La dérivée $df/dh \Big|_{h=0}$ est positive mais varie selon un fonction quadratique du paramètre σ ; elle tend vers 0 lorsque σ tend vers l'infini et vice versa.

Modèle	$f(0)$	$df/dh _{h=0}$
$\mathcal{E}(\lambda)$	λ	$-\lambda^2$
$\Gamma(\beta, \alpha)$	0	$\begin{cases} 0 & \text{si } \beta > 2 \\ \alpha^2 & \text{si } \beta = 2 \\ +\infty & \text{si } 1 < \beta < 2 \\ -\alpha^2 & \text{si } \beta = 1 \\ -\infty & \text{si } 0 < \beta < 1 \end{cases}$
$\ln \mathcal{N}(\mu, \sigma)$	0	0
$\mathcal{R}(\sigma)$	0	$1/\sigma^2$

TABLE 2.11 : Densité des modèles simples ou fonction g des modèles combinés

- Au niveau des TIV très courts, les modèles Γ et $\ln \mathcal{N}$ sont les plus compatibles avec la condition « idéale » pour le TIV. En conséquence, ce résultat montre l'avantage de l'utilisation de ces deux modèles pour la fonction g des TIV contraints dans les modèles complexes.

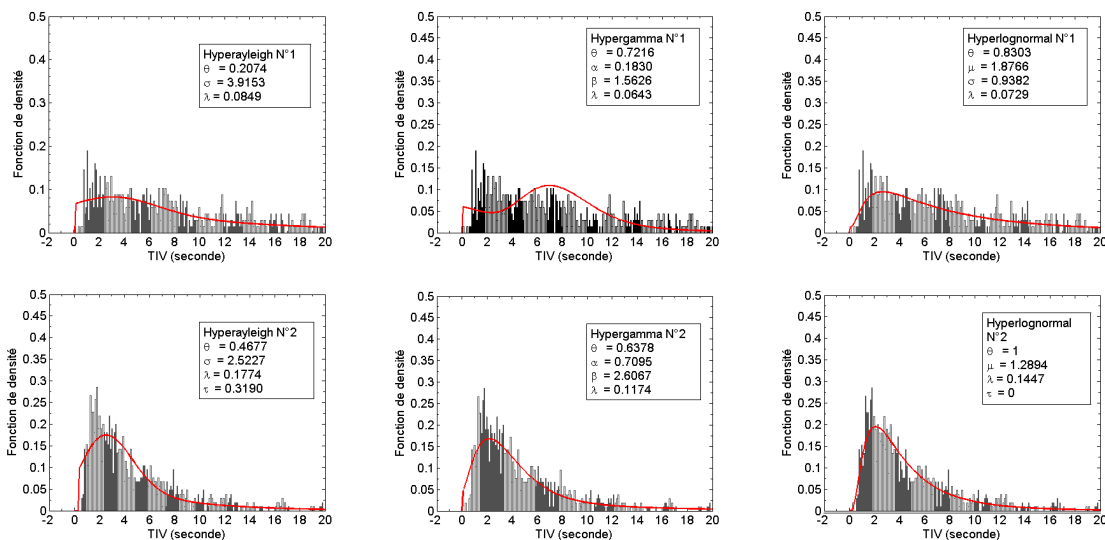


FIGURE 2.24 : Les cas particuliers au niveau du TIV zéro des modèles combinés aux débits faibles

Par ailleurs, le modèle Hyperlognormal $\ln \mathcal{N}^*$ est ici utilisé pour la première fois dans le but de calibrer les distributions des TIV. À la différence des modèles LNM et lognormal-GQM, le modèle Hyperlognormal $\ln \mathcal{N}^*$ peut être estimé par la méthode MV ce qui constitue un grand avantage pour ce modèle combiné.

2.4.2 Modèle de mélange

Pour éviter les inconvénients des modèles combinés visibles sur les courbes de distribution, le modèle de mélange est utilisé en modifiant la variable V des véhicules libres avant la composition des composantes. Le Tableau 2.12 montre les principes de combinaison des composantes dans les modèles de mélange. En résultat, la fonction f_2 obtenue dans les modèles de mélange est différente de celle (exponentielle) des modèles combinés (cf. les Équation 2.133 et 2.144).

Modèle	Distribution g de U	Auteurs	Relation entre U et V
SPM	g prédéterminée	Buckley [124]	$\Pr\{V \geq t V > U\} \sim e^{-\lambda t}$
GQM	g prédéterminée	Cowan [59], Branston [58]	$V = U + Y$ où $Y \sim e^{-\lambda t}$

TABLE 2.12 : Modèles de mélange

Les fonctions de densité de V dans les modèles g -SPM et g -GQM sont respectivement comme suit :

$$\Xi_1(h) = \lambda e^{-\lambda h} \frac{G(h)}{g^L(\lambda)} \quad (2.185)$$

$$\Xi_2(h) = \lambda e^{-\lambda h} \left[\int_0^h g(u) e^{\lambda u} du \right] \quad (2.186)$$

On obtient évidemment que $\Xi_1(0) = 0$ et $\Xi_2(0) = 0$ ce qui donne aux modèles de mélange tous les avantages du modèle combiné en évitant des inconvénients de ces derniers. En effet, on a :

$$f(h) = \theta g(h) + (1 - \theta) \Xi_i(h) \quad \text{avec } i = 1, 2. \quad (2.187)$$

On remarque que le niveau de difficulté de calcul et d'estimation des modèles g -GQM dépend essentiellement de la fonction $I(h)$ suivante :

$$I(h) = \int_0^h g(u) e^{\lambda u} du \quad (2.188)$$

Pour le modèle g -SPM, la transformation de Laplace est requise :

$$g^L(\lambda) = \int_0^{+\infty} g(u) e^{-\lambda u} du \quad (2.189)$$

Comme il n'y a pas d'expression analytique de la transformation de Laplace du modèle log-normal $\ln \mathcal{N}$, l'emploi de modèle de mélange lognormal-SPM est déconseillé pour le TIV.

Dans les sections précédentes, les résultats obtenus montrent que les ajustements obtenus des deux modèles de mélange Γ -GQM et Γ -SPM sont quasiment équivalents et que les temps d'estimation et de calcul du modèle Γ -SPM sont exorbitants. Cette section va donc mettre en lumière la raison pour laquelle les calculs de la fonction de densité f et de la fonction de répartition F du modèle g -GQM sont généralement plus faciles à exécuter que ceux du modèle g -SPM.

En effet, selon Gross & Harris [126, 97], la fonction de répartition F des modèles g -GQM est écrite :

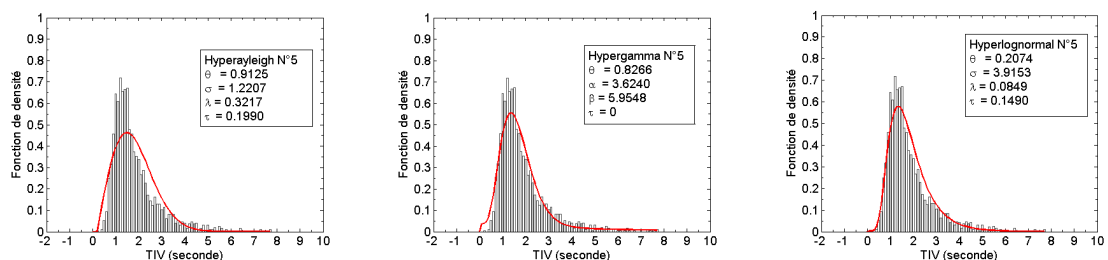


FIGURE 2.25 : Les cas particuliers au niveau du TIV zéro des modèles combinés aux débits denses

$$F(h) = \theta G(h) + (1 - \theta) \int_0^h G(h - u) \lambda e^{-\lambda u} du \quad (2.190)$$

Le deuxième terme de l'Équation 2.190 n'est autre que la convolution d'une variable exponentielle et d'une autre variable de fonction de répartition G . Ce terme est développé comme suit :

$$\int_0^h G(h - u) \lambda e^{-\lambda u} du = - \int_0^h G(h - u) d e^{-\lambda u} \quad (2.191)$$

$$= - G(h - u) e^{-\lambda u} \Big|_{u=0}^h + \int_0^h e^{-\lambda u} d G(h - u) \quad (2.192)$$

$$= G(h) - G(0) e^{-\lambda h} + \int_0^h e^{-\lambda u} d G(h - u) \quad (2.193)$$

$$\text{En posant } h - u = v \quad (2.194)$$

$$= G(h) - \int_0^h e^{-\lambda(h-v)} g(v) dv \quad (2.195)$$

La fonction de répartition des TIV est donc :

$$F(h) = \theta G(h) + (1 - \theta) \left[G(h) - \int_0^h e^{-\lambda(h-v)} g(v) dv \right] \quad (2.196)$$

$$= G(h) - (1 - \theta) e^{-\lambda h} \int_0^h g(u) e^{\lambda u} du \quad (2.197)$$

On remarque que la fonction F du modèle g -GQM est toujours inférieure à celle du modèle de densité g car le deuxième terme dans l'équation 2.197 n'est pas négatif. On obtient également que le terme important (en temps de calcul) du modèle g -GQM est celui de la fonction $I(h)$ dans l'équation 2.188, ce qui n'est pas le cas pour le modèle g -SPM. En effet, la fonction de répartition F du SPM est comme suit :

$$F(h) = \int_0^h f(u) du \quad (2.198)$$

$$= \int_0^h \left[\theta g(u) + (1 - \theta) \frac{\lambda e^{-\lambda u}}{g^L(\lambda)} G(u) \right] du \quad (2.199)$$

$$= \theta G(h) + \frac{(1 - \theta) \lambda}{g^L(\lambda)} \int_0^h G(u) e^{-\lambda u} du \quad (2.200)$$

Le calcul du terme $J(h) = \int_0^h G(u) e^{-\lambda u} du$ est généralement plus compliqué que celui de $I(h)$. En effet, lorsque les fonctions g et G sont celles du modèle Gamma $\Gamma(\beta, \alpha)$, on a :

$$I(h) \sim \gamma[\beta, (\alpha - \lambda)h] \quad (2.201)$$

$$J(h) \sim \int_0^h \gamma(\beta, \alpha u) e^{-\lambda u} du \quad (2.202)$$

Dans notre étude, pour calculer l'intégrale de type $I(h)$, deux méthodes sont comparées : la méthode de Monte-Carlo et la méthode classique utilisant la division linéaire de Simpson. Finalement, la méthode classique impliquant des calculs plus rapides et précis est choisie.

2.4.3 Développement des modèles pour les TIV

En prolongation des modèles examinés, les deux approches suivantes sont faites pour trouver de nouveaux modèles pour le TIV :

A. Développement des modèles de mélange

Notons $\mathcal{G}_*(\Theta)$ la fonction de densité du modèle Γ -GQM de paramètre $\Theta = \{\alpha, \beta, \theta, \lambda\}$.

1. Modèle Hyper-GQM

Le modèle Hyper-GQM utilise le modèle GQM comme la distribution f_1 d'un modèle combiné. Comme le modèle $\mathcal{G}_*(\Theta)$ convient très bien pour les distributions des TIV courts, l'ajustement du modèle Hyper-GQM pourrait être satisfaisant. La fonction de densité est donc :

$$f(h) = \theta_* \mathcal{G}_*(h) + (1 - \theta_*) \lambda_* e^{-\lambda_* h} \quad (2.203)$$

où les paramètres θ_*, λ_* sont ceux du modèles Hyper-GQM, et sont distincts des paramètres θ et λ de la composante GQM $\mathcal{G}_*(\alpha, \beta, \theta, \lambda)$.

Le modèle Hyper-GQM possède deux paramètres de plus que le modèle GQM. Pourtant, ce modèle ne satisfait pas à la première condition « *idéale* » concernant le TIV zéro. De plus, l'utilisation de deux paramètres λ, λ_* et deux paramètres θ, θ_* rend difficile leur interprétation.

2. Modèle de mélange d'ordre deux

Comme le modèle Hyper-GQM, le modèle GQM d'ordre deux réutilise le modèle Γ -GQM dans un modèle de mélange. En principe, on peut aussi combiner un modèle GQM dans un modèle SPM ou un modèle SPM dans un modèle GQM, ou bien de la même façon un modèle SPM d'ordre deux. Pourtant, une approche aussi complexe n'est pas nécessaire, d'autant plus que les calculs des fonctions $I(h)$ et $J(h)$ seules sont déjà très lourds.

3. Le GQM mixte - une modification du modèle de mélange de type GQM

Dans les modèles g -GQM, les TIV dits libres V sont aussi constitués deux parties, l'une liée à la contrainte de sécurité V_1 (on l'appelle le TIV *gardé*) et l'autre liée aux arrivées exponentielles V_2 (on l'appelle le TIV *de relaxe*). Le modèle GQM suppose implicitement que $V_1 \equiv U$ ce qui correspond à inclure la partie représentant la contrainte de sécurité pour tous les véhicules aussi bien dans la distribution des TIV dits *libres* que dans celle des TIV dits *constraints*.

Dans cette section, on suppose au contraire que le TIV V_1 de la partie contrainte des véhicules libres est différente du TIV U des véhicules contraints. Les variables (considérées comme aléatoires) V_1 et U ne suivent pas de la même loi (soit deux modèles différents, soit un même modèle avec des paramètres différents). Cette distinction est bien raisonnable parce que les comportements des conducteurs sont éloignés dans les deux situations qui ne sont pas du tout semblables. En effet, dans le premier cas des TIV *constraints* U , les conducteurs s'aperçoivent visiblement de leur situation contrainte tandis que dans le deuxième cas des TIV *gardé* V_1 , les conducteurs gardent une attention (quelque part peut être dans leur sub-conscient) aux problèmes de sécurité, autrement dit, les conducteurs dits *libres* ne sont jamais totalement libres au moins au niveau psychologique. En conséquence, les deux

contraintes sont de nature différence; les variables TIV correspondantes sont supposées suivre deux distributions différentes.

L'idée est d'utiliser deux modèles différents pour la distribution des TIV contraints (TIV contraint U et TIV gardé V_1) - le premier modèle est employé dans le terme principal de la fonction $f(h)$, alors que le deuxième modèle est utilisé dans la convolution. Le nouveau modèle est appelé modèle *mixte*.

Cette approche conduit à augmenter le nombre des paramètres des modèles mixtes par rapport aux modèles GQM originels. Deux possibilités de mixtion sont intéressantes :

1. Un modèle Γ pour U , et un modèle $\ln\mathcal{N}$ pour V_1 : le modèle mixte obtenu est Γ -GQM- $\ln\mathcal{N}$.
2. Un modèle $\ln\mathcal{N}$ pour U , et un modèle Γ pour V_1 : le modèle mixte obtenu est $\ln\mathcal{N}$ -GQM- Γ .

Comme la convolution entre un modèle $\ln\mathcal{N}$ et un modèle exponentiel \mathcal{E} cause des difficultés de calcul, le modèle GQM mixte Γ -GQM- $\ln\mathcal{N}$ n'est pas utilisé. Par conséquent, le modèle mixte $\ln\mathcal{N}$ -GQM- Γ est proposé et examiné. Sa fonction de densité est comme suit :

$$f(h) = \theta \frac{1}{h\sigma} \phi\left(\frac{\ln h - \mu}{\sigma}\right) + (1 - \theta) \left(\frac{\alpha}{\alpha - \lambda}\right)^\beta \frac{\gamma[\beta, (\alpha - \lambda)h]}{\Gamma(\beta)} \lambda e^{-\lambda h} \quad (2.204)$$

Où la fonction $\phi(\cdot)$ est la densité de la loi normale centrée réduite :

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (2.205)$$

Ce modèle mixte comporte au total six paramètres. En se basant sur l'équation 2.197, la fonction de répartition du modèle $\ln\mathcal{N}$ -GQM- Γ est :

$$F(h) = \theta L(h) + (1 - \theta) \left[G(h) - \int_0^h e^{-\lambda(h-v)} g(v) dv \right] \quad (2.206)$$

$$= \theta \Phi\left(\frac{\ln h - \mu}{\sigma}\right) \quad (2.207)$$

$$+ (1 - \theta) \frac{\gamma[\beta, \alpha(h - \tau)]}{\Gamma(\alpha)} \quad (2.208)$$

$$- (1 - \theta) \left(\frac{\alpha}{\alpha - \lambda}\right)^\beta \frac{\gamma[\beta, (\alpha - \lambda)h]}{\Gamma(\beta)} e^{-\lambda h} \quad (2.209)$$

Où $L(\cdot)$ et $G(\cdot)$ sont respectivement les fonctions de répartition des modèles $\ln\mathcal{N}$ et Γ .

B. Amélioration des modèles combinés

Comme les modèles de mélange évitent les problèmes de TIV zéro par des techniques supplémentaires pour que les fonctions $f_2(0)$ soient égales à zéro, il y a une autre « technique » plus simple qui peut aussi régler le problème de TIV zéro. De façon similaire, on combine deux modèles f_1 et f_2 où $f_1(0) = 0$ et aussi $f_2(0) = 0$.

Par ailleurs, la tendance exponentielle des TIV longs doit être respectée. Pour cela, il faut choisir comme fonction f_2 soit une généralisation du modèle exponentiel \mathcal{E} , soit au moins une fonction asymétrique positive. Finalement, en s'appuyant sur l'étude des modèles simples dans les sections précédentes, trois possibilités sont examinées :

1. Le modèle Double-Gamma (modèle DG)

Ce modèle comporte cinq paramètres résultant de deux lois Gamma, $\Gamma(\beta_1, \alpha_1)$ et $\Gamma(\beta_2, \alpha_2)$ plus le paramètre θ de la composition. La fonction de densité est comme suit :

$$f(h) = \theta \frac{\alpha_1^{\beta_1} h^{\beta_1-1}}{\Gamma(\beta_1)} e^{-\alpha_1 h} + (1 - \theta) \frac{\alpha_2^{\beta_2} h^{\beta_2-1}}{\Gamma(\beta_2)} e^{-\alpha_2 h} \quad (2.210)$$

2. Le modèle Double-Lognormal (modèle DL)

Le modèle comporte cinq paramètres résultant de deux lois log-normale, $\ln \mathcal{N}(\mu_1, \sigma_1)$ et $\ln \mathcal{N}(\mu_2, \sigma_2)$ plus le paramètre θ de la composition. La fonction de densité est comme suit :

$$f(h) = \theta \frac{1}{h\sigma_1} \phi\left(\frac{\ln h - \mu_1}{\sigma_1}\right) + (1 - \theta) \frac{1}{h\sigma_2} \phi\left(\frac{\ln h - \mu_2}{\sigma_2}\right) \quad (2.211)$$

3. Le modèle Lognormal-Gamma (modèle LG)

Ce modèle comporte aussi cinq paramètres résultant d'une loi Gamma $\Gamma(\beta, \alpha)$ et d'une loi log-normale $\ln \mathcal{N}(\mu, \sigma)$ en ajoutant le paramètre θ de la combinaison. La fonction de densité du modèle LG est comme suit :

$$f(h) = \theta \frac{1}{h\sigma} \phi\left(\frac{\ln h - \mu}{\sigma}\right) + (1 - \theta) \frac{\alpha^\beta h^{\beta-1}}{\Gamma(\beta)} e^{-\alpha h} \quad (2.212)$$

L'équation 2.209 montre aussi que la valeur de la fonction de répartition du modèle LG est toujours supérieure à celle du modèle GQM mixte $\ln \mathcal{N}$ -GQM- Γ (avec les mêmes paramètres $\theta, \alpha, \beta, \mu, \sigma$, quelque soit $\lambda > 0$).

2.4.4 Méthode d'estimation

Pour estimer les modèles développés, la méthode MV associée avec le processus proposé est utilisée. Le paramètre de décalage n'est pas utilisé dans les nouveaux modèles.

2.4.5 Comparaison des nouveaux modèles

Pour comparer les modèles développés par rapport aux modèles déjà connus dans les sections précédentes, l'ensemble test des échantillons est utilisé à nouveau.

Comparaison du modèle HyperGQM avec le modèle Γ -GQM

Le modèle Hyper-GQM est ici comparé, via le test K-S, au modèle Γ -GQM qui ressortait de la sélection de la section précédente. Les comparaisons systématiques directes avec tous les modèles analysés précédemment ne sont pas nécessaires.

Le Tableau 2.13 montre que le modèle Hyper-GQM ne fournit pas de meilleur ajustement que le modèle Γ -GQM. Par ailleurs, le modèle HyperGQM ne donne aucun résultat pour les échantillons N°1, N°14, N°20, N°21 et N°22. En général, le modèle Hyper-GQM n'est pas rejeté au niveau de signification de 5 %. Pourtant, ce modèle n'est pas stable car ses calibrages sont mauvais avec de très grandes valeurs de la statistique K-S dans les échantillons N°3, N°8, N°19 et N°23.

Valeurs de Statistiques K-S $\times 10^{-4}$							
N°	Hyper-GQM	Γ -GQM	\mathcal{D}_2 , 5 %	N°	Hyper-GQM	Γ -GQM	\mathcal{D}_2 , 5 %
N°02	0.0181	0.0201	0.0502	N°12	0.0093	0.0087	0.0297
N°03	0.2178	0.0141	0.0297	N°13	0.0132	0.0132	0.0297
N°04	0.0109	0.0108	0.0297	N°15	0.0268	0.0252	0.0393
N°05	0.0245	0.0101	0.0372	N°16	0.0106	0.0155	0.0297
N°06	0.0229	0.0213	0.0530	N°17	0.0161	0.0149	0.0297
N°07	0.0097	0.0135	0.0297	N°18	0.0177	0.0087	0.0297
N°08	0.2095	0.0108	0.0297	N°19	0.2451	0.0199	0.0297
N°09	0.0168	0.0109	0.0297	N°23	0.1248	0.0348	0.0586
N°10	0.0106	0.0144	0.0325	N°24	0.0169	0.0169	0.0438
N°11	0.0373	0.0200	0.0297	N°25	0.0254	0.0227	0.0343

TABLE 2.13 : Les statistiques K-S des modèles Hyper-GQM, Γ -GQM

La Figure 2.26 montre également que le temps d'estimation t_e du modèle Hyper-GQM est largement plus long et le temps d'ajustement t_a est légèrement plus long que ceux du modèle Γ -GQM. En conséquence, le modèle Hyper-GQM est moins bon que le modèle Γ -GQM.

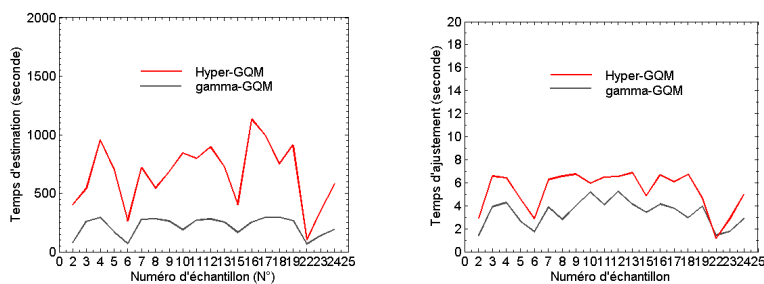


FIGURE 2.26 : Les temps d'estimation et d'ajustement du modèle Hyper-GQM

Les autres modèles développés

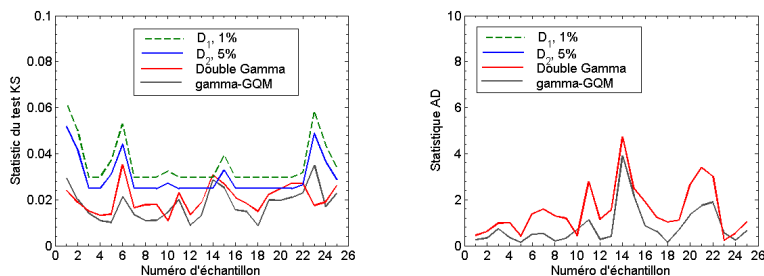


FIGURE 2.27 : L'ajustement du modèle DG

Tous les autres nouveaux modèles fournissent de bons résultats d'ajustement sur l'ensemble des 25 échantillons testés. Le modèle DG n'est pas rejeté dans la majorité des échantillons. La Figure 2.27 montre qu'il est meilleur que le modèle Γ -GQM pour certains échantillons. Au niveau du test A-D, il est légèrement moins bon que le modèle Γ -GQM. Le principal avantage du modèle DG par rapport au modèle Γ -GQM est au niveau du temps

d'estimation t_e . Cependant, le temps d'ajustement t_a de ce modèle est un peu plus long de quelques secondes par rapport à tous les autres nouveaux modèles développés (cf. Figure 2.30).

La Figure 2.28 montre que le modèle LG (le modèle Gamma-Lognormal dans les figures) est un bon modèle mais il est rejeté au niveau de signification de 5 % dans certains cas, par exemple dans les échantillons N°6, N°8, N°12 et N°18. Le temps d'ajustement t_a du modèle LG est plus court de quelques secondes mais le temps d'estimation t_e est toujours largement plus élevé que celui du modèle Γ -GQM.

À l'inverse, le modèle DL donne de très bons ajustements dans tous les échantillons testés. Ce modèle est meilleur que les modèles LG, DG et Γ -GQM.

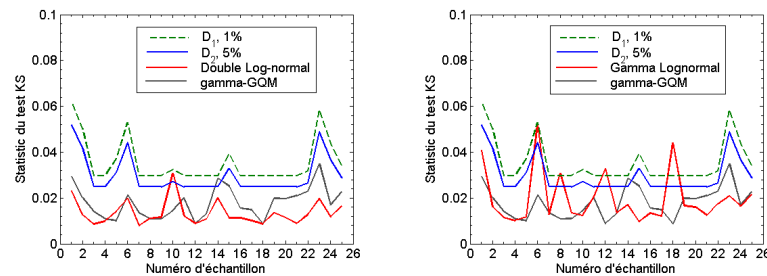


FIGURE 2.28 : Les ajustements des modèles DL et LG

Dans la Figure 2.29, les valeurs des tests K-S et A-D du modèle GQM mixte $\ln\mathcal{N}$ -GQM- Γ montrent clairement que ce modèle GQM mixte est absolument supérieur au modèle Γ -GQM. Le modèle $\ln\mathcal{N}$ -GQM- Γ est même meilleur que le modèle DL. De plus, le temps d'ajustement t_a de ce modèle mixte est le plus court (de moins de deux secondes). Pourtant, comme le modèle DL, ce modèle demande un temps d'estimation t_e considérablement plus long que les autres modèles (cf. Figure 2.30).

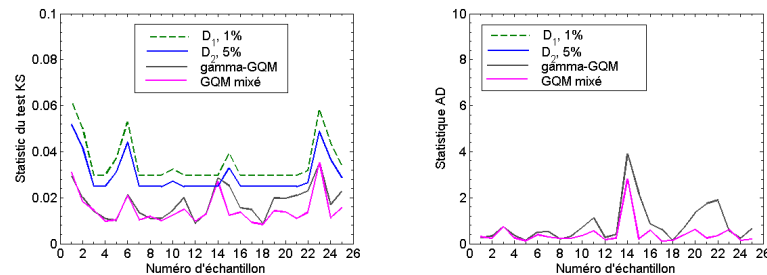


FIGURE 2.29 : Les ajustements du modèle GQM mixte $\ln\mathcal{N}$ -GQM- Γ

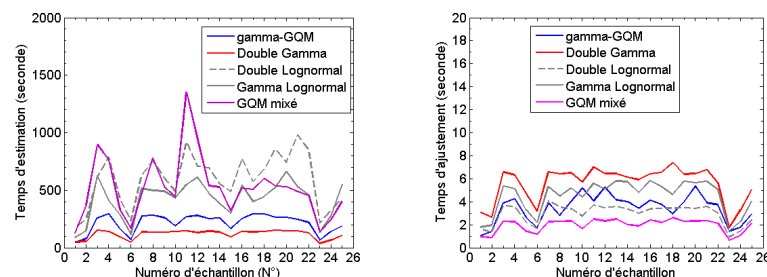


FIGURE 2.30 : Les temps t_e et t_a des modèles développés

La Figure 2.31 suivante illustre graphiquement l'adaptation des meilleurs modèles aux

données réelles :

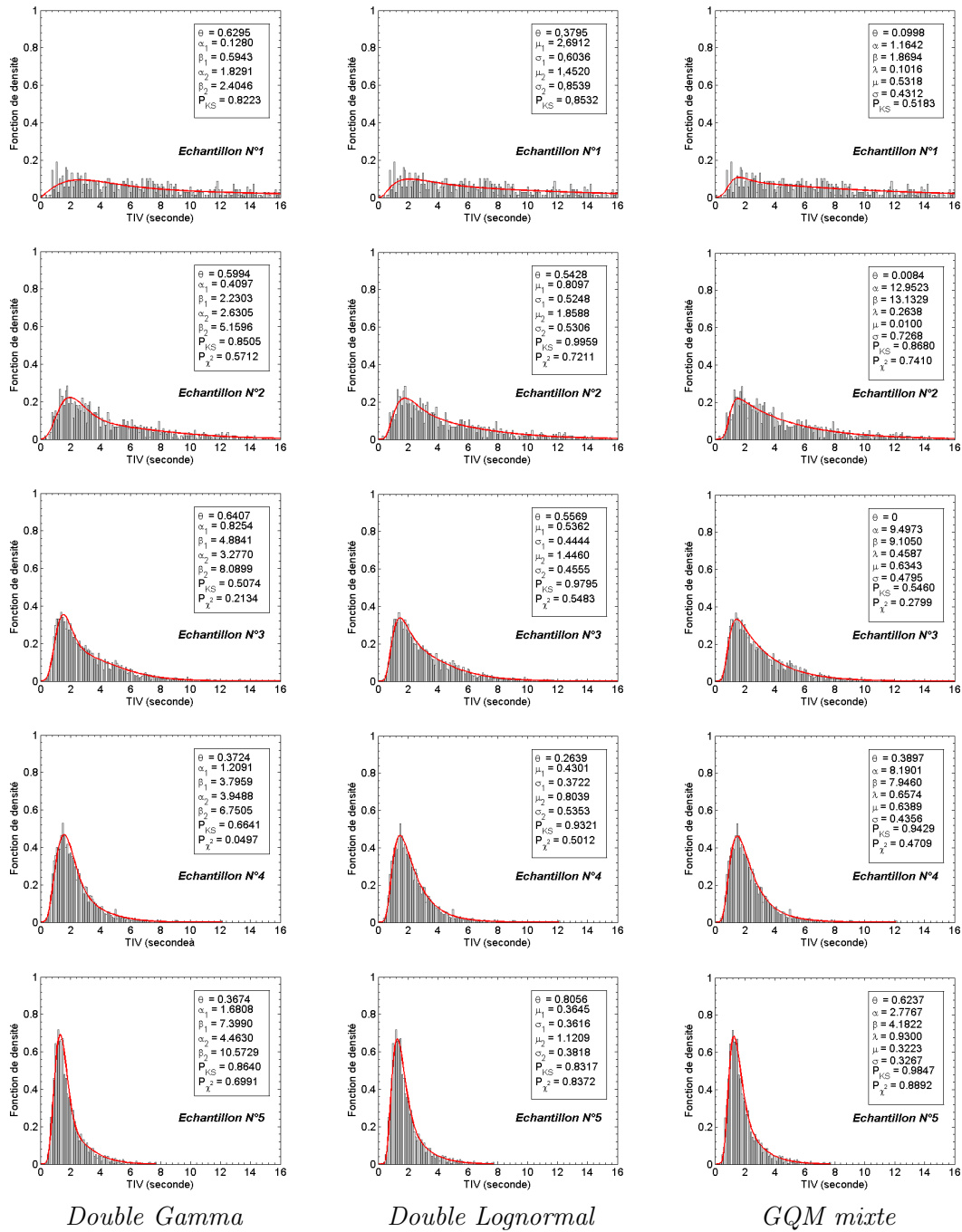


FIGURE 2.31 : L'ajustement graphique des meilleurs nouveaux modèles de TIV

2.5 Conclusion du chapitre

1. Le modèle de Heidemann et le modèle des automates cellulaires sont examinés. Leur limite pour la modélisation des TIV est discutée. Le modèle de Heidemann n'est pas flexible tandis que le modèle des automates cellulaires est trop compliqué même dans le cas le plus simple de $v_{\max} = 1$. En conséquence, l'approche probabiliste est utilisée pour la modélisation du TIV.
2. Trois types de modèles probabilistes de TIV sont distingués dans la littérature. Ce sont le modèle simple, le modèle combiné et le modèle de mélange. Les modèles sont distingués par les composantes considérées et le rapport entre les composantes au sein du modèle.
3. Un processus d'estimation pour les modèles complexes (combiné et de mélange) à partir d'un ensemble test est proposé pour examiner les performances des modèles. Au cours de cette analyse, les modèles Rayleigh, Hyperrayleigh, rayleigh-GQM et Hyper-LNM sont proposés. Le modèle Hyper-LNM est un bon modèle examiné pour la première fois. Les comparaisons exhaustives entre les modèles selon les méthodes d'estimation se terminent en faveur du modèle de mélange Γ -GQM (sans paramètre de décalage). Il est recommandé de calibrer ce modèle avec la méthode d'estimation du maximum de vraisemblance. Les deux modèles Γ -SPM et Γ -GQM sont équivalents tant au sens du test K-S qu'au sens du test du χ^2 .
4. Cinq nouveaux modèles de TIV sont proposés : le modèle Hyper-GQM, le modèle Double Gamma, le modèle Double Lognormal, le modèle Lognormal Gamma et en particulier le modèle GQM mixte. Le modèle Hyper-GQM est moins bon que le modèle Γ -GQM tandis que les quatre derniers sont soit équivalents, soit meilleurs que le modèle Γ -GQM au sens du test K-S. Pour des raisons pratiques, de la performance et du nombre de paramètres, le modèle Double Gamma et le modèle Γ -GQM sont sélectionnés dans les chapitres suivants pour modéliser les distribution des TIV.

Chapitre 3

Les paramètres des modèles Γ -GQM et DG

Dans le chapitre précédent, une étude exhaustive de la modélisation du TIV a été réalisée. Différentes approches et différents modèles probabilistes ont été approfondis. Une des conclusions essentielles est que le TIV peut être bien modélisé dans toutes les conditions réellement rencontrées par l'utilisation des modèles complexes, notamment le modèle Γ -GQM et le modèle DG. Une première question est de comprendre comment les paramètres de ces modèles s'adaptent aux changements de la distribution des TIV. Cette question est tout d'abord traitée par l'étude des paramètres des modèles eux-mêmes avant d'apprécier la flexibilité et l'ajustement des modèles en fonction des données extérieures. Ce chapitre est donc consacré à l'étude de la sensibilité des paramètres des deux modèles sélectionnés, - les plus pratiques -, les modèles Γ -GQM et DG.

3.1 Paramètre de décalage τ

En général, l'intervention d'un paramètre de décalage τ améliore l'ajustement d'un modèle probabiliste aux données réelles. C'est aussi le cas pour le modèle Γ -GQM. Pourtant, ce paramètre n'a pas de sens physique, dépend plutôt des TIV minima résultant de la précision de la méthode de mesure. D'autant plus que le TIV peut théoriquement prendre n'importe quelle valeur aussi petite qu'elle soit, le cas extrême étant une collision où le TIV est nul.

Les données ont montré aussi que sans le paramètre τ , l'ajustement est encore très bon. La Figure 3.1 représente la qualité du modèle Γ -GQM dans des cas où l'ajustement est très bon sans utilisation du paramètre de décalage τ .

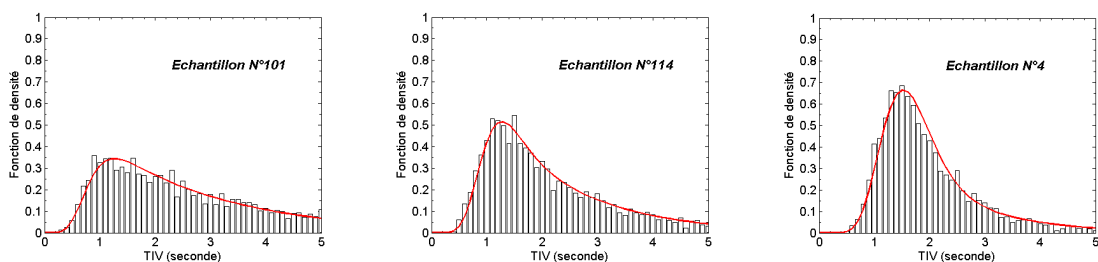


FIGURE 3.1 : La qualité du modèle Γ -GQM sans paramètre de décalage au niveau des TIV très courts (TIV < 5 s)

Un autre raison de ne pas utiliser le paramètre τ est qu'il modifierait considérablement

les valeurs des autres paramètres. En effet, on remarque que les dérivées de la fonction de densité df/dh au TIV minimum sont différentes et entraînent des valeurs des paramètres (notamment les α et β) largement différentes. En conséquence, si l'on considère que τ n'a pas de sens physique, il ne faut absolument pas l'utiliser puisqu'il altère les autres paramètres qui pourraient avoir des explications plus « physique » dans le modèle.

3.2 Le modèle Γ -GQM

3.2.1 Les propriétés statistiques

Rappelons la fonction de densité du modèle Γ -GQM :

$$f(h) = \theta g(h) + (1 - \theta) \lambda e^{-\lambda h} \left[\int_0^h g(v) e^{\lambda v} dv \right] \quad (3.1)$$

où g est le modèle Γ de paramètres α, β .

La transformation de Laplace de la fonction de densité du modèle Γ -GQM est :

$$f^L(s) = \left(\frac{\alpha}{s + \alpha} \right)^\beta \left[\theta + (1 - \theta) \frac{\lambda}{s + \lambda} \right] \quad (3.2)$$

Il est à noter que

$$\mathbb{E}[H^n] = (-1)^n \frac{d^n f^L}{dh^n} \Big|_{h=0} \quad (3.3)$$

Par conséquent, les statistiques de base du modèle Γ -GQM sont comme suit :

$$\mu(H) = \frac{\beta}{\alpha} + \frac{1 - \theta}{\lambda} \quad (3.4)$$

$$\sigma^2(H) = \frac{\beta}{\alpha^2} + \frac{1 - \theta^2}{\lambda^2} \quad (3.5)$$

$$CV = \frac{\sqrt{\beta \lambda^2 + \alpha^2 (1 - \theta)}}{\alpha \beta + \alpha (1 - \theta)} \quad (3.6)$$

$$S_3 = 2 \times \frac{[\alpha^3 (1 - \theta^3) + \beta \lambda^3]}{\sqrt{(\alpha^2 (1 - \theta^2) + \beta \lambda^2)^3}} \quad (3.7)$$

$$K_4 = 3 \times \frac{[\alpha^4 [4 - (1 + \theta^2)^2] + \beta \lambda^2 [2\alpha^2 (1 - \theta^2) + \lambda^2 (\beta + 2)]]}{(\alpha^2 (1 - \theta^2) + \beta \lambda^2)^2} \quad (3.8)$$

On constate que :

- Lorsque $\theta = 1$, le modèle Γ -GQM dégénère en modèle Γ .
- Lorsque $\theta = 0$, la distribution des TIV est une convolution entre un modèle Γ et une tendance exponentielle ce qui est totalement différent de la combinaison entre un modèle Γ et un modèle \mathcal{E} du modèle Hypergamma.
- La condition $\alpha > \lambda$ est absolument nécessaire pour que la fonction f prenne des valeurs réelles.

3.2.2 Influence de $\theta, \lambda, \alpha, \beta$ du modèle Γ -GQM

Le paramètre θ du modèle Γ -GQM

La Figure 3.2 représente l'effet de la variation du paramètre θ sur la forme de la distribution du modèle Γ -GQM lorsque $\lambda = 0,5$. Chaque petite figure correspond à une

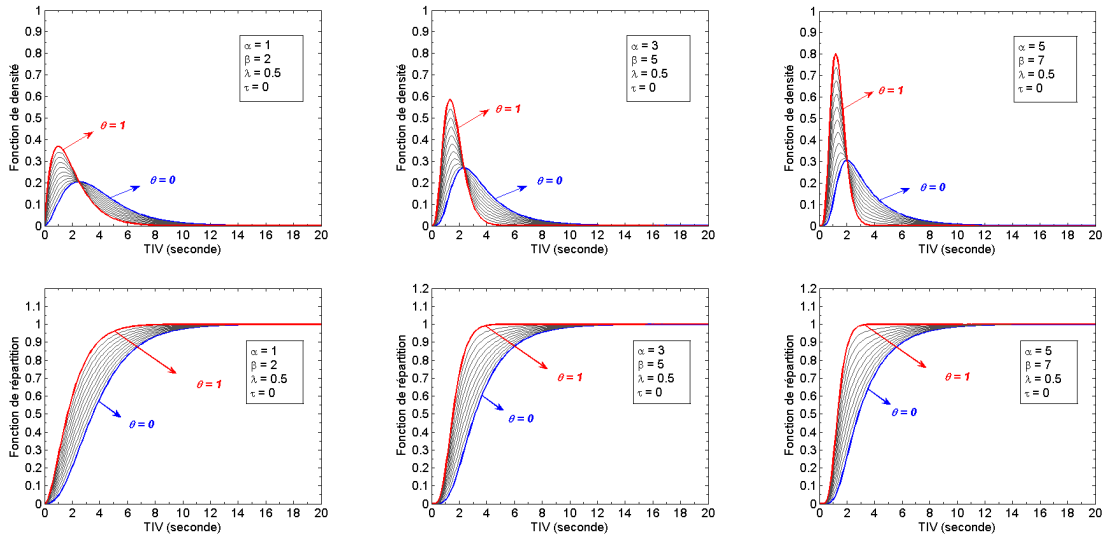


FIGURE 3.2 : Les distributions du modèle Γ -GQM selon le paramètre θ pour différentes valeurs de α, β

configuration des valeurs (α, β) . Les courbes de la fonction de répartition ne se croisent pas, et sont classées en ordre lorsque θ croît de 0 à 1. Les écarts entre les courbes de la fonction de répartition sont petits pour les TIV grands, disons supérieurs à 10 s.

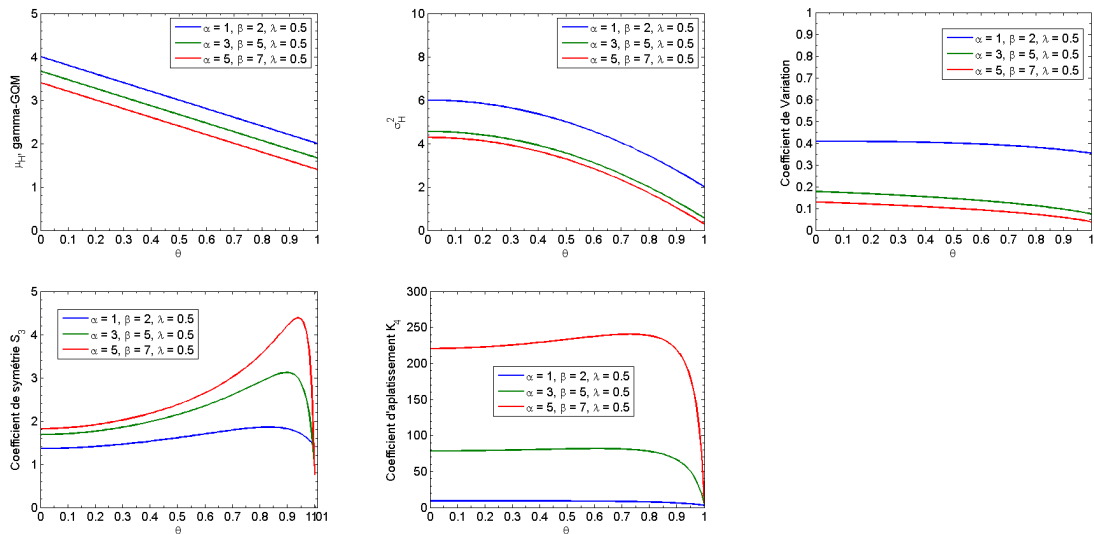


FIGURE 3.3 : Les coefficients statistiques du modèle Γ -GQM en fonction de θ

Lorsque les autres paramètres α, β, λ sont donnés, la moyenne et l'écart-type diminuent lorsque le paramètre θ croît. Le CV diminue légèrement en fonction du paramètre θ (cf. la Figure 3.3). La Figure 3.3 montre aussi que les coefficients de symétrie et d'aplatissement du modèle Γ -GQM sont sensibles pour les valeurs de θ grandes, disons supérieures à 0,8.

Le paramètre λ du modèle Γ -GQM

Dans le cas où $\lambda = 0$, la fonction de densité dégénère en une partie d'une loi Gamma. Dans le cas $\alpha = 1$ dans la Figure 3.4, le modèle Γ -GQM ne donne pas de valeurs réelles car $\lambda = 1 = \alpha$. On obtient que lorsque λ croît, les fréquences des TIV courts augmentent et les courbes de la fonction de répartition sont en ordre et ne se croisent pas. À la différence

du paramètre θ , les écarts entre les courbes de la fonction de répartition lorsque λ varie, sont petits pour les TIV courts, disons inférieurs à 2 s.

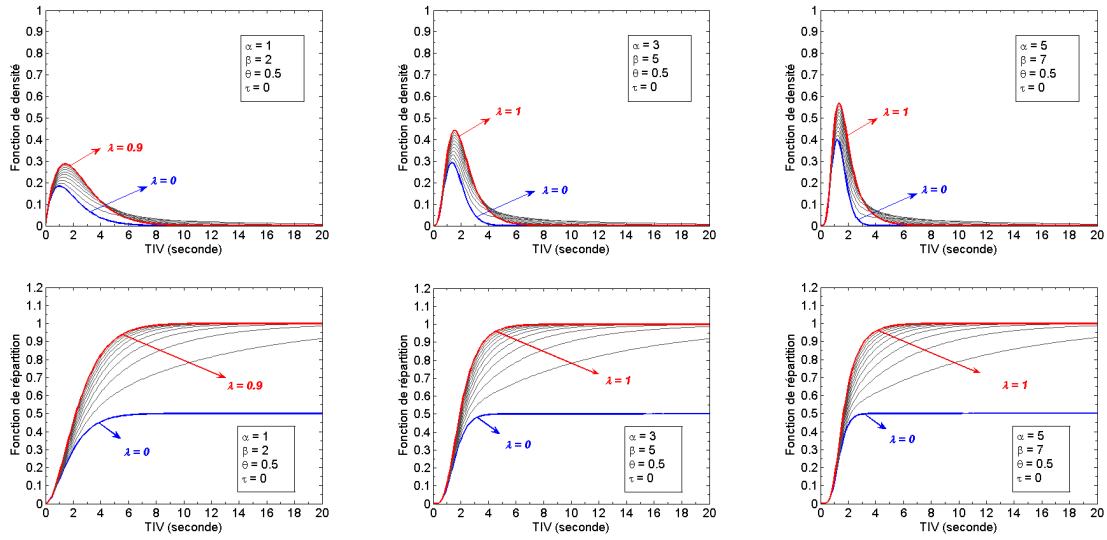


FIGURE 3.4 : Les distributions du modèle Γ -GQM selon le paramètre λ pour différentes valeurs de α, β

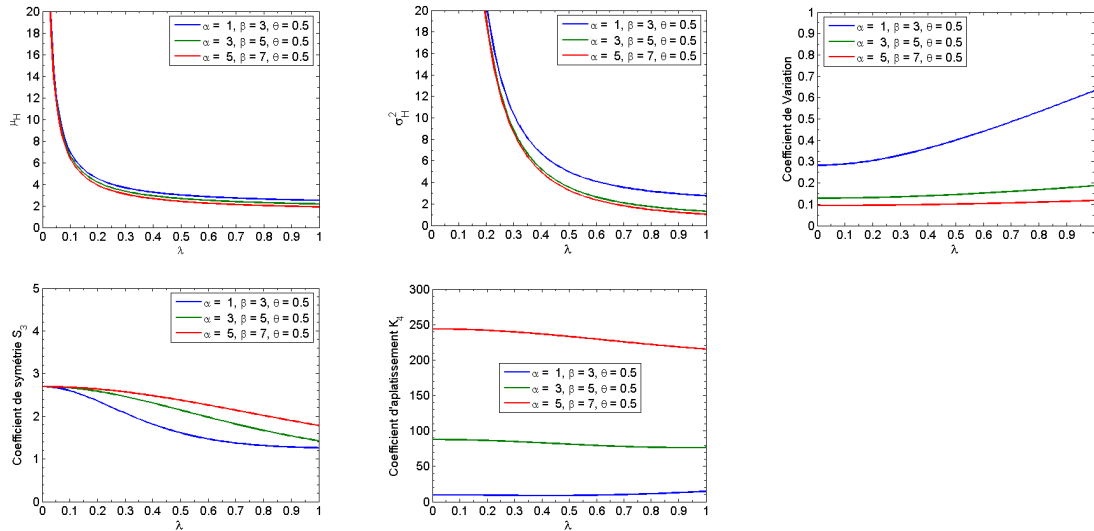


FIGURE 3.5 : Les coefficients statistiques du modèle Γ -GQM en fonction de λ

La Figure 3.5 illustre l'influence du paramètre λ sur les coefficients statistiques du modèle Γ -GQM lorsque les autres paramètres sont fixés. On constate que :

- La moyenne des TIV est très sensible aux petites valeurs de λ inférieures à 0,2 s. L'écart-type est aussi sensible aux petites valeurs de λ inférieures à 0,5 s. Le CV augmente légèrement en fonction de λ .
- Les valeurs du coefficient de symétrie S_3 varient entre 1 et 3. En général, ce coefficient diminue en fonction de λ .
- Le coefficient d'aplatissement semble stable en fonction de λ , il dépend considérablement des paramètres α et β .

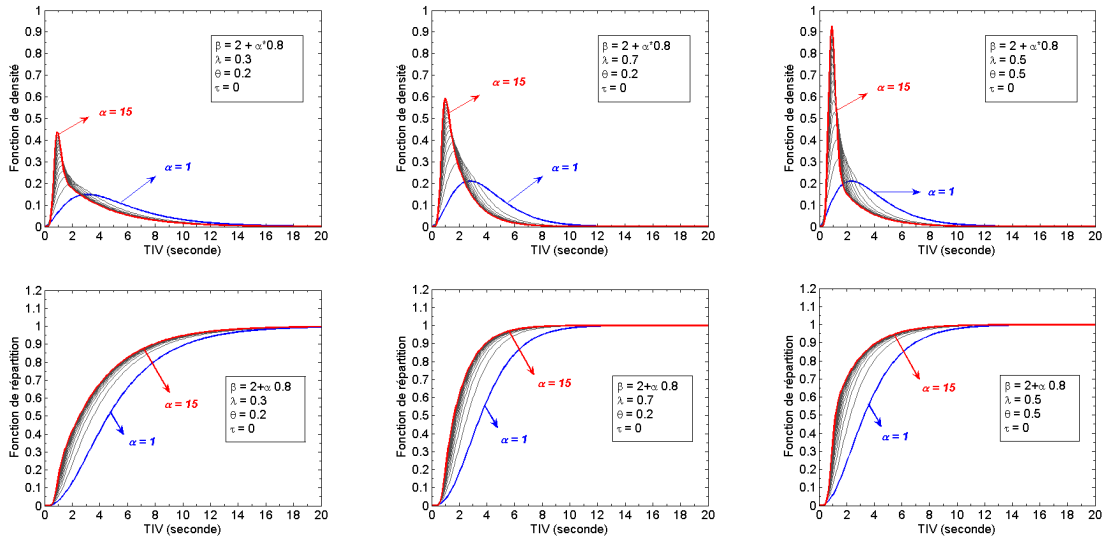


FIGURE 3.6 : Les distributions du modèle Γ -GQM selon les paramètres α, β pour différentes valeurs de θ, λ

Le rôle des paramètres α et β du modèle Γ -GQM

Les différentes formes de distribution des TIV en fonction des paramètres α, β sont représentées dans la Figure 3.6. On observe que :

- Les courbes de la fonction de répartition ne se croisent pas et coïncident aux TIV grands, disons supérieurs à 12 ou 15 s. La variation de α influence clairement la distribution des TIV courts.
- Les courbes de la fonction de répartition sont classées en ordre selon λ ou θ mais se concentrent de plus en plus lorsque α augmente.
- Plus α est grand, plus la distribution de densité est *pointue*, avec un densité élevée au *mode*.

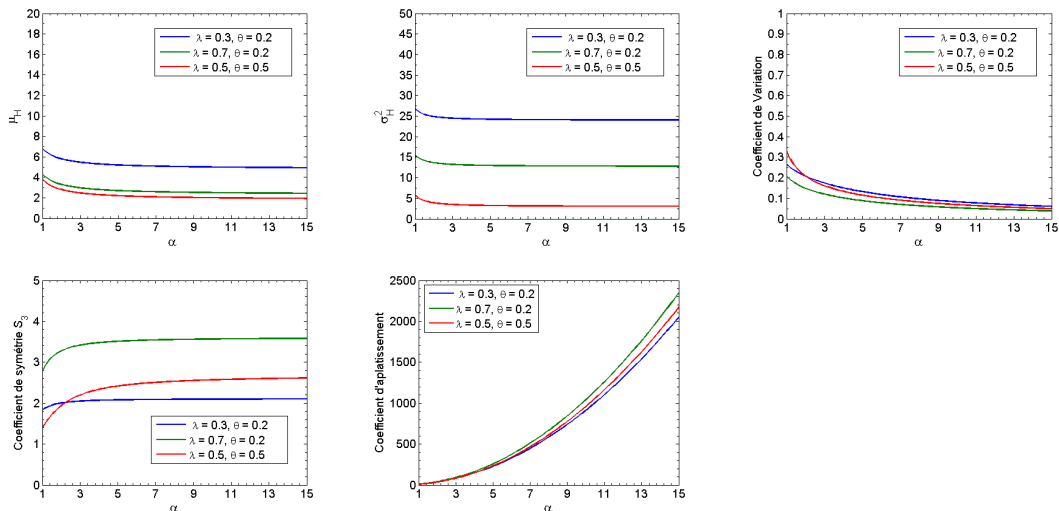


FIGURE 3.7 : Les coefficients statistiques du modèle Γ -GQM en fonction de α et β

La Figure 3.7 illustre comment les coefficients statistiques du modèle Γ -GQM varient en fonction de α (et donc β pris égal à $2 + 0,8\alpha$). Il est observé que la moyenne, l'écart-type

et le coefficient de symétrie des TIV sont légèrement dépendants des α et β , particulièrement lorsque α est supérieur à 3. Le CV diminue légèrement lorsque α augmente. Seul le coefficient d'aplatissement dépend fortement du paramètre α avec une augmentation quasi exponentielle.

La relation entre α et β

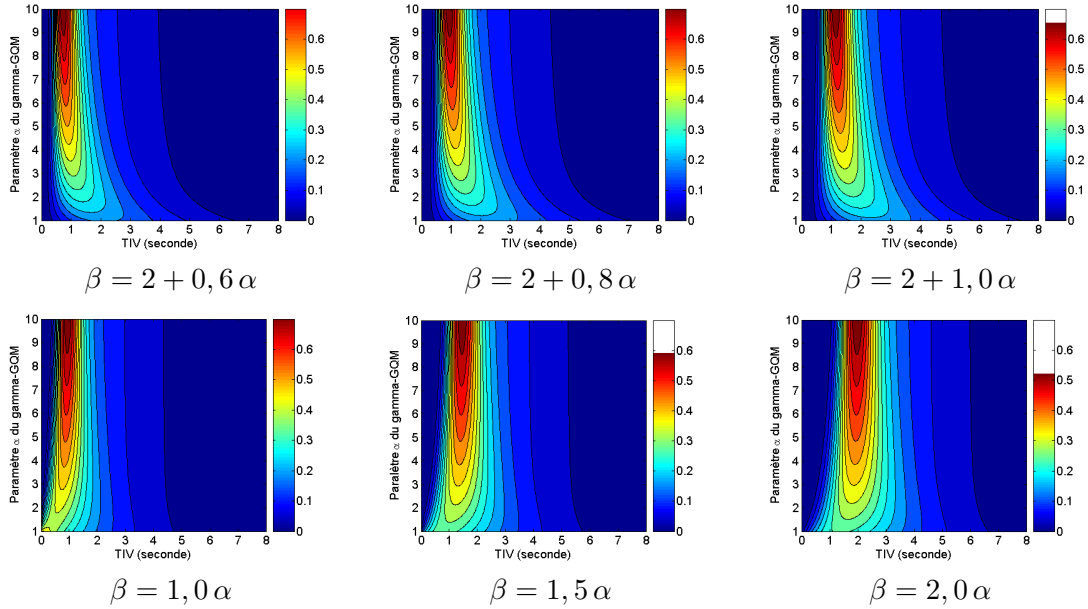


FIGURE 3.8 : Densité du modèle Γ -GQM selon différentes relations entre α et β

La Figure 3.8 montre l'influence de la relation entre α et β sur la distribution de la densité des TIV. Il est à noter que la relation proposée par Luttinen $\beta = 0,6 \alpha$ n'est pas valable pour des valeurs du paramètre α petites, disons inférieures à 1.

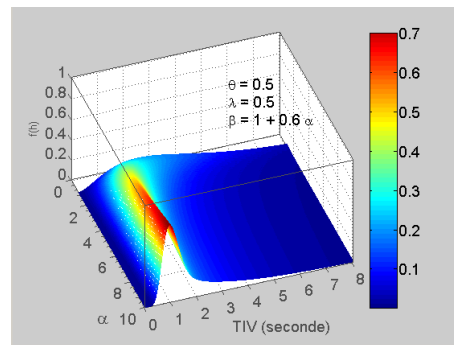


FIGURE 3.9 : La distribution du modèle Γ -GQM en fonction du paramètre α

3.3 Le modèle DG

3.3.1 Les propriétés statistiques

Rappelons, la fonction de densité du modèle DG :

$$f(h) = \theta \frac{\alpha_1^{\beta_1} h^{\beta_1-1}}{\Gamma(\beta_1)} e^{-\alpha_1 h} + (1 - \theta) \frac{\alpha_2^{\beta_2} h^{\beta_2-1}}{\Gamma(\beta_2)} e^{-\alpha_2 h} \quad (3.9)$$

La transformation de Laplace de la fonction de densité du modèle DG est :

$$f^L(s) = \theta \left(\frac{\alpha_1}{s + \alpha_1} \right)^{\beta_1} + \theta \left(\frac{\alpha_2}{s + \alpha_2} \right)^{\beta_2} \quad (3.10)$$

Par conséquent, les statistiques de base du modèle DG sont :

$$\mathbb{E}[H] = \theta \frac{\beta_1}{\alpha_1} + (1 - \theta) \frac{\beta_2}{\alpha_2} \quad (3.11)$$

$$\mathbb{E}[H^2] = \theta \frac{1}{\alpha_1^2} [\beta_1 (\beta_1 + 1)] + (1 - \theta) \frac{1}{\alpha_2^2} [\beta_2 (\beta_2 + 1)] \quad (3.12)$$

$$\mathbb{E}[H^3] = \theta \frac{1}{\alpha_1^3} [\beta_1 (\beta_1 + 1) (\beta_1 + 2)] + \theta \frac{1}{\alpha_2^3} [\beta_2 (\beta_2 + 1) (\beta_2 + 2)] \quad (3.13)$$

$$\begin{aligned} \mathbb{E}[H^4] &= \theta \frac{1}{\alpha_1^4} [\beta_1 (\beta_1 + 1) (\beta_1 + 2) (\beta_1 + 3)] + \\ &+ \theta \frac{1}{\alpha_2^4} [\beta_2 (\beta_2 + 1) (\beta_2 + 2) (\beta_2 + 3)] \end{aligned} \quad (3.14)$$

En conséquence, les coefficients statistiques du modèle DG sont comme suit :

$$\mu(H) = \theta \frac{\beta_1}{\alpha_1} + (1 - \theta) \frac{\beta_2}{\alpha_2} \quad (3.15)$$

$$\sigma^2(H) = \theta \frac{\beta_1}{\alpha_1^2} + (1 - \theta) \frac{\beta_2}{\alpha_2^2} + \theta (1 - \theta) \left(\frac{\beta_1}{\alpha_1} - \frac{\beta_2}{\alpha_2} \right)^2 \quad (3.16)$$

$$CV = \frac{\sigma(H)}{\mu(H)} \quad (3.17)$$

$$\begin{aligned} S_3 &= \frac{2}{(\sigma(H))^3} \times \left[\frac{\theta}{\alpha_1^3} \beta_1 (1 - \beta_1^2) + \frac{1 - \theta}{\alpha_2^3} \beta_2 (1 - \beta_2^2) + \left(\frac{\theta \beta_1}{\alpha_1} \right)^3 + \left(\frac{\theta \beta_2}{\alpha_2} \right)^3 \right] + \\ &= + 3 \theta (1 - \theta) \beta_1 \beta_2 \left[(1 - 2\theta) \left(\frac{\beta_2}{\alpha_1 \alpha_2^2} - \frac{\beta_1}{\alpha_2 \alpha_2^2} \right) - \left(\frac{1}{\alpha_1 \alpha_2^2} + \frac{1}{\alpha_2 \alpha_2^2} \right) \right] \end{aligned} \quad (3.18)$$

On constate que :

- Lorsque $\theta = 1$ ou $\theta = 0$, le modèle DG dégénère en un modèle Γ
- Les deux lois Gamma composantes sont équivalentes. En conséquence, les rôles des paramètres α_1, β_1 sont semblables à ceux des paramètres α_2, β_2 . Par ailleurs, le modèle dans le cas où θ se trouve entre 0 et 0,5 ne diffère pas du cas où θ est entre 0,5 et 1.
- Comme la convolution n'est pas utilisée, il n'y a pas de condition supplémentaire entre les paramètres du modèle comme dans le cas du modèle Γ -GQM.

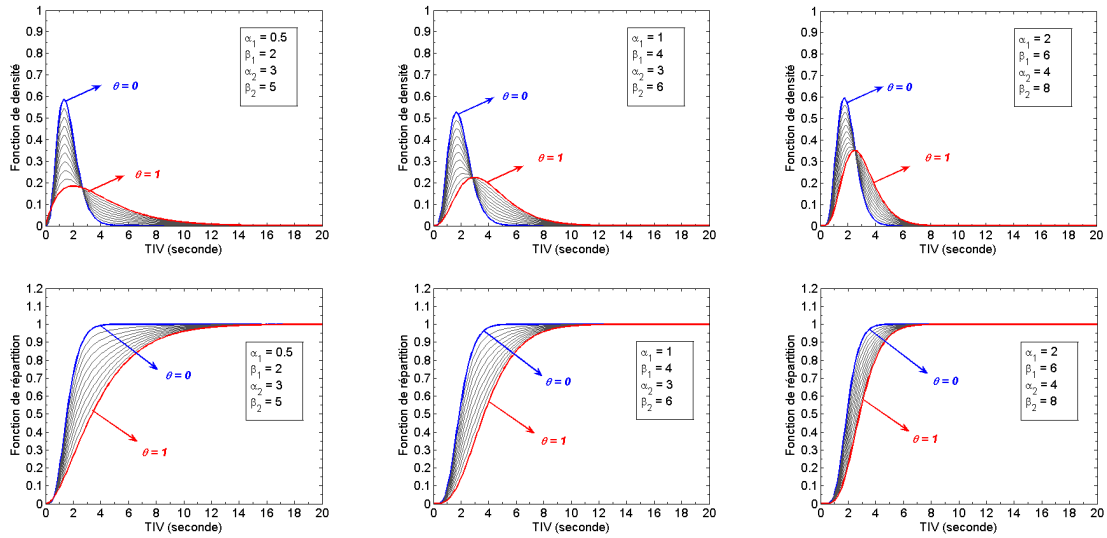


FIGURE 3.10 : Les distributions du modèle DG selon le paramètre θ

3.3.2 Influence des paramètres du modèle DG

Le paramètre θ du modèle DG

La Figure 3.10 illustre l'effet de la variation du paramètre θ sur la forme de la distribution du modèle DG en fixant les autres paramètres. Les courbes de la fonction de répartition ne se croisent pas et sont en ordre quand θ croît. Les écarts entre les courbes cumulatives sont petits pour les TIV grands, disons supérieurs à 10 s. Les écarts entre les courbes extrêmes de $\theta = 0$ et $\theta = 1$ dépendent fortement des valeurs des autres paramètres.

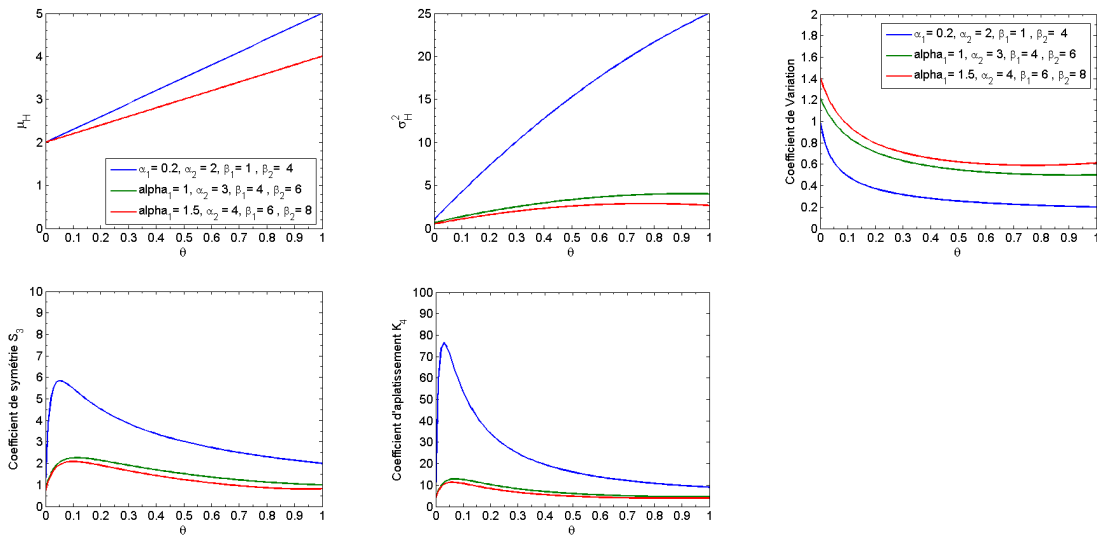


FIGURE 3.11 : Les coefficients statistiques du modèle DG en fonction de θ

La Figure 3.11 présente les variations des moments du modèle DG en fonction du paramètre θ . On observe que la moyenne augmente linéairement, et la variance des TIV croît en fonction du paramètre θ . Néanmoins, le CV est stable pour les θ supérieurs à 0,5. Aussi, dans les cas où θ est supérieur à 0,5, les coefficients de symétrie S_3 et d'aplatissement K_4 sont faibles. Les coefficients S_3 et K_4 sont maximaux pour θ autour de 0,05.

Les rôles de α_1 et β_1 du modèle DG

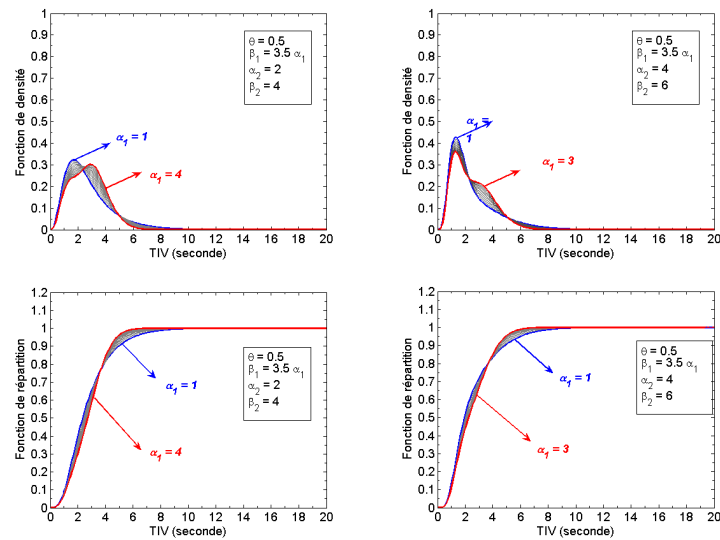


FIGURE 3.12 : Les distributions du modèle DG selon le paramètre α_1, β_1 pour deux couples (α_2, β_2)

Puisque les rôles des paramètres α_1, β_1 sont équivalents à ceux des paramètres α_2, β_2 , on n'étudie ici que les rôles de α_1 et β_1 . La Figure 3.12 représente les distributions des TIV selon la variation des paramètres α_1 en supposant qu'il y a une relation entre α_1 et β_1 . On constate que :

- Les courbes de la fonction de répartition se croisent aux valeurs des TIV intermédiaires. Les courbes ne varient pas régulièrement mais l'écart entre elles est faible.
- Le modèle DG peut produire des distribution bi-modales. Dans les données réelles des TIV, certains échantillons ont en effet un deuxième *mode* (local) qui se trouve après le *mode* global de la distribution (cf. Chapitre 4).

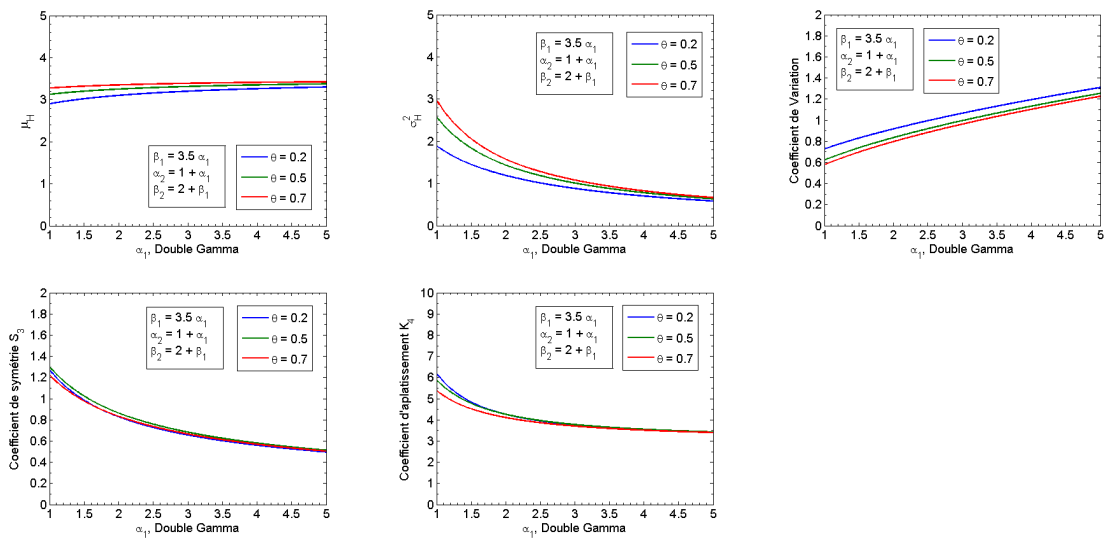


FIGURE 3.13 : Les coefficients statistiques du modèle DG en fonction de α_1

La Figure 3.13 illustre comment les coefficients statistiques du modèle DG évoluent en supposant θ constant et qu'il existe des relations entre α_1 et les autres paramètres β_1, α_2 et β_2 . En ajustant le modèle DG sur des données réelles, on obtient que dans tous les

échantillons, le paramètre α_1 est inférieur à 5. C'est la raison pour laquelle cette valeur est considérée comme la limite supérieure dans les figures. Il est à noter que les moments obtenus dépendent considérablement des relations proposées entre α_1 et les paramètres β_1, α_2 et β_2 . La Figure 3.13 montre plutôt que le paramètre θ a peu influence sur α_1 dans les calculs des coefficients statistiques.

La relation entre α_i et β_i

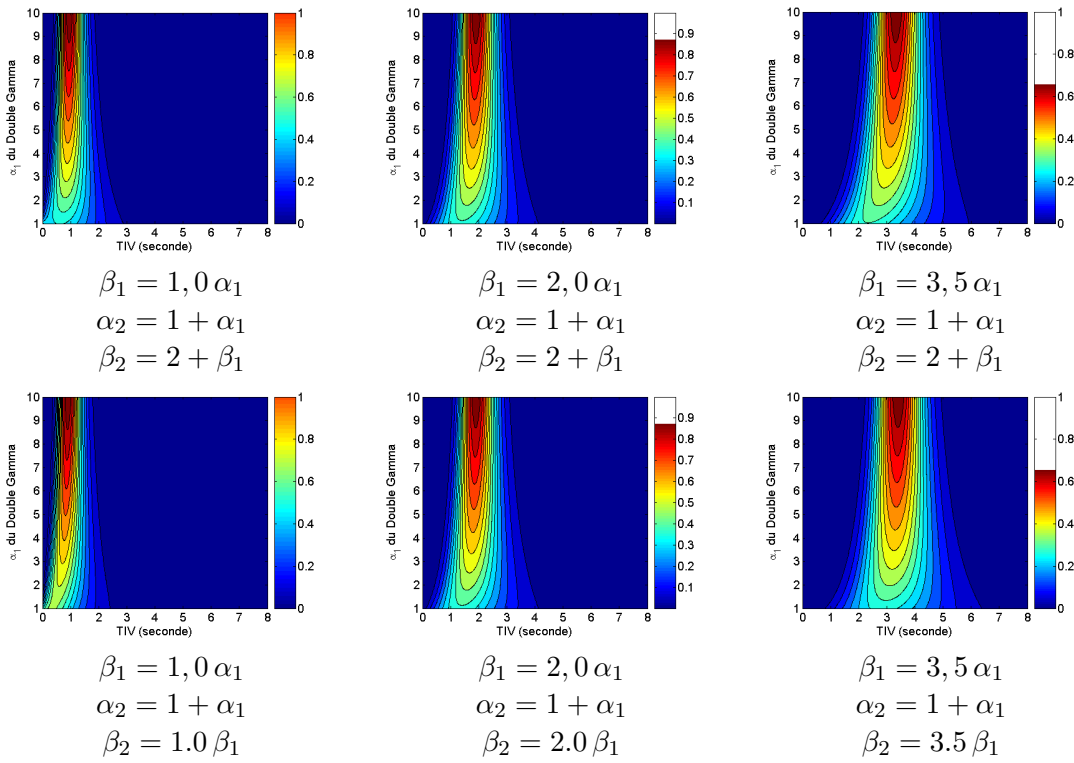


FIGURE 3.14 : Densité du modèle DG selon différentes relations entre les paramètres α_1 et les paramètres β_1, α_2 et β_2

La Figure 3.14 présente différents résultats des distributions des TIV obtenus en fonction des relations entre les paramètres. Le décalage des distributions confirme que l'utilisation du paramètre de décalage τ n'est pas nécessaire. La large variation des distributions montre la flexibilité du modèle DG.

Le modèle DG peut fournir des distributions bi-modales. La Figure 3.15 illustre ces cas particuliers en utilisant différentes relations entre les paramètres des lois Gamma.

3.4 Estimation par la méthode du maximum de vraisemblance

3.4.1 Les fonctions de vraisemblance des modèles

La fonction de Log-vraisemblance du modèle Γ -GQM est :

$$\begin{aligned} \ell_1(\Theta, \mathbf{h}) &= \beta \log \alpha - \log \Gamma(\beta) \\ &+ \log \left[\theta h^{\beta-1} e^{-\alpha h} + (1 - \theta) \left(\frac{1}{\alpha - \lambda} \right)^\beta \gamma[\beta, h(\alpha - \lambda)\lambda e^{-\lambda h}] \right] \end{aligned} \quad (3.19)$$

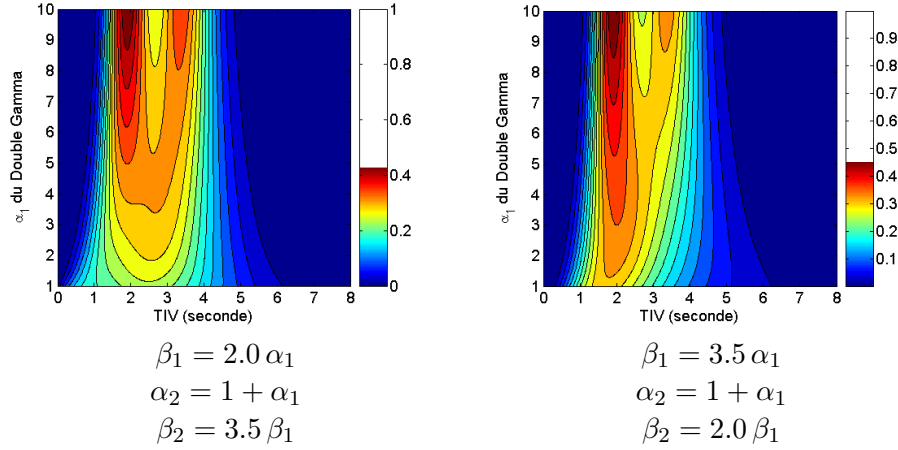


FIGURE 3.15 : Les cas bi-modaux de la distribution du modèle DG

La fonction de Log-vraisemblance du modèle DG est comme suit :

$$\ell_2(\Theta, \mathbf{h}) = \log \left[\theta \frac{\alpha_1^{\beta_1} h^{\beta_1 - 1}}{\Gamma(\beta_1)} + (1 - \theta) \frac{\alpha_2^{\beta_2} h^{\beta_2 - 1}}{\Gamma(\beta_2)} \right] \quad (3.20)$$

3.4.2 La méthode du Bootstrap

Les estimations des deux modèles Γ -GQM et DG sont réalisées par la méthode du maximum de vraisemblance (MV) dans cette étude. En conséquence, l'information de Fisher des deux modèles semble nécessaire pour mesurer la sensibilité des modèles à la variation des paramètres obtenus dans la phase de calibration.

Soit un modèle statistique $\mathcal{P} := f_H(h | \theta \in \Theta)$, avec la fonction de vraisemblance $\mathcal{L}(\theta, \mathbf{h})$. La méthode MV consiste à maximiser la fonction $\mathcal{L}(\theta, \mathbf{h})$ associée aux observations H . Cette estimation est équivalente à minimiser la fonction opposée de la log-vraisemblance $\ell(\theta, \mathbf{h})$.

Dans le cas où θ est un vecteur comme c'est le cas pour les modèles étudiés, c. à. d Θ est un ouvert dans l'espace $\mathbb{R}^{1 \times d}$, l'information de Fisher est représentée par une matrice de covariance de taille $d \times d$. Comme le domaine de définition de $H = (H_1, H_2, \dots, H_n) \in \mathcal{X}^n$ ne dépend pas de θ , et que les fonctions f étudiées satisfont la condition de régularité, on a :

$$I_n(\theta) = \mathbb{E} \left[\nabla_{\theta} \ell(\theta, \mathbf{h})^T \nabla_{\theta} \ell(\theta, \mathbf{h}) \right] \quad (3.21)$$

$$= -\mathbb{E}_{\theta} \left(\frac{\partial^2 \ell}{\partial \theta_i \partial \theta_j}(\theta, \mathbf{h}) \right) \quad 1 \leq i, j \leq d \quad (3.22)$$

La *variable score* est un vecteur $1 \times d$

$$\nabla_{\theta} \ell(\theta, \mathbf{h}) = \left(\frac{\partial \ell(\theta, \mathbf{h})}{\partial \theta_1}, \dots, \frac{\partial \ell(\theta, \mathbf{h})}{\partial \theta_d} \right) \quad (3.23)$$

L'inverse de la matrice $I_n(\theta)$ sert à déterminer la borne de Cramér - Rao exprimant la limite inférieure de la variance d'un estimateur *sans biais*.

Pourtant, les fonctions de log-vraisemblance des modèles Γ -GQM et DG sont très complexes à développer analytiquement comme il y a des termes de sommation dans les équations 3.19 et 3.20. En conséquence, les variations des paramètres estimés sont approximées par la méthode de *rééchantillonnage*. La méthode utilisée est celle du *Bootstrap*.

Soit un n -échantillon *initial* $H = (h_1, h_2 \dots h_n)$, la méthode du *Bootstrap* consiste à reconstituer un nombre B d'échantillons d'*image* (ou bien échantillon *bootstrap*) par tirage avec remise les données h_i de l'échantillon *initial*. Ces échantillons d'image sont notés H_k^* avec $k = 1, 2, \dots, B$.

Soit un paramètre $\hat{\theta}$ estimé par la statistique \mathcal{F} à partir des données de l'échantillon *initial* :

$$\hat{\theta} = \mathcal{F}(h_1, h_2 \dots h_n) = \mathcal{F}(H) \quad (3.24)$$

Dans le cas où il n'y a pas d'expression analytique pour $\hat{\theta}$, l'étude de cette statistique sur les échantillons *bootstrap* peut fournir certaines évaluations de $\hat{\theta}$. En effet, pour chaque échantillon obtenu par rééchantillonnage, le paramètre θ est recalculé en utilisant la même fonction \mathcal{F} :

$$\hat{\theta}_k^* = \mathcal{F}(H_k^*) \quad \text{avec } k = 1, 2, \dots, B \quad (3.25)$$

En fonction de la distribution empirique obtenue, la moyenne, l'écart-type et l'intervalle de confiance peuvent être estimés. On applique cette approche pour les paramètres estimés par la méthode de MV pour les modèles Γ -GQM et DG. Ici, on prend $B = 500$ et on considère que :

$$\mathcal{F}(H) = \arg \text{Max}_{\theta} \ell(H) \quad (3.26)$$

Le modèle Γ -GQM								
Θ	α		β		θ		λ	
N°	Var	CV	Var	CV	Var	CV	Var	CV
1	5.3243	1.1059	4.6036	0.7251	0.0926	0.7067	0.0055	0.0552
2	3.2139	0.4487	2.8358	0.3227	0.0914	1.0394	0.0108	0.0411
3	1.5409	0.2031	1.3536	0.1426	0.0441	0.2837	0.0100	0.0222
4	1.4389	0.2112	0.9821	0.1058	0.1105	0.3766	0.0226	0.0302
5	1.4589	0.1636	1.2528	0.1087	0.0762	0.1599	0.0506	0.0559
6	0.8159	0.3009	1.2183	0.1946	0.0371	0.1516	0.0040	0.0523
7	1.6000	0.2136	1.4786	0.1744	0.0297	0.8695	0.0062	0.0237
8	2.0550	0.2673	1.4672	0.1605	0.0728	0.5002	0.0110	0.0248
9	2.0614	0.2354	1.5295	0.1438	0.0880	0.3376	0.0238	0.0356
10	0.9842	0.1516	1.0381	0.1103	0.0859	0.1547	0.0650	0.0590
11	0.6307	0.1986	0.7231	0.1261	0.0394	0.1419	0.0058	0.0367
12	2.1625	0.2758	1.6479	0.1754	0.0653	0.3881	0.0121	0.0238
13	0.2313	0.0417	0.2204	0.0301	0.0190	0.0337	0.0156	0.0219
14	0.4927	0.1281	0.6135	0.0950	0.0508	0.0576	0.1308	0.3391

TABLE 3.1 : La sensibilité des paramètres du modèle Γ -GQM dans les échantillons bootstrap

Les Tableaux 3.1 et 3.2 représentent les résultats sur les 14 premiers échantillons de l'*ensemble test* (cf. Section 2.3.2) respectivement pour le modèle Γ -GQM et le modèle DG. On constate que :

- Dans le modèle Γ -GQM, le paramètre λ est celui qui a l'écart-type et le CV les plus petits.

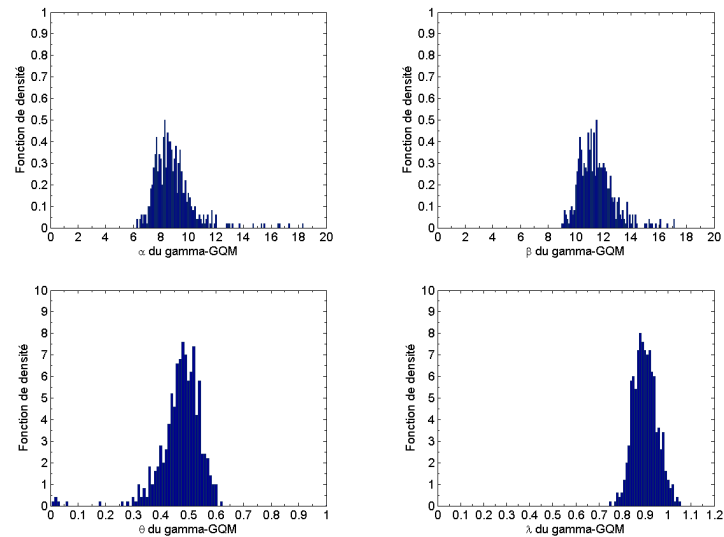


FIGURE 3.16 : Les distributions des paramètres du modèle Γ -GQM dans les échantillons bootstrap

- Ce sont les paramètres α et θ qui varient le plus tant au sens de l'écart-type que au sens du coefficient de variation (CV).
 - En général, les paramètres du modèle DG varient beaucoup moins que ceux du modèle Γ -GQM. Dans ce modèle, le paramètre β_1 semble être le paramètre le plus sensible.
- Les Figures 3.16 et 3.17 montrent les répartitions des valeurs des paramètres dans B échantillons. L'exemple est pris à partir de l'échantillon N°5. On constate que :
- Les distributions des paramètres obtenus par la méthode du bootstrap sont proches des lois normales ce qui pourrait faciliter le calcul de l'intervalle de confiance des paramètres
 - L'échelle du paramètre α_1 est différente de celles des paramètres α_2, β_1 et β_2

N°	Le modèle DG									
	α_1		α_2		β_1		β_2		θ	
	Var	CV	Var	CV	Var	CV	Var	CV	Var	CV
1	0.1675	1.1259	3.6168	1.5775	1.0917	1.1702	1.5839	0.5379	0.1522	0.2483
2	0.0294	0.0714	0.1616	0.0608	0.7970	0.3407	1.2089	0.2272	0.0623	0.1052
3	0.0368	0.0444	0.1381	0.0419	0.7514	0.1507	0.9408	0.1145	0.0331	0.0517
4	0.1058	0.0872	0.2706	0.0678	0.4348	0.1133	0.5269	0.0775	0.0605	0.1648
5	0.1462	0.0860	0.3944	0.0868	0.7309	0.0976	0.8532	0.0799	0.0436	0.1198
7	0.0329	0.0913	0.1413	0.0565	0.2384	0.1665	0.3654	0.0932	0.0596	0.1210
8	0.0396	0.0539	0.1381	0.0432	0.3607	0.1147	0.4467	0.0766	0.0406	0.0788
9	0.0587	0.0491	0.1476	0.0366	0.4243	0.0882	0.4973	0.0653	0.0337	0.0793
10	0.2049	0.1016	0.5552	0.1188	0.9176	0.1160	1.1530	0.0971	0.0581	0.1318
11	0.0117	0.0548	0.0758	0.0388	0.1804	0.1106	0.2781	0.0683	0.0273	0.0471
12	0.0425	0.0473	0.1532	0.0453	0.5421	0.1229	0.6545	0.0897	0.0383	0.0667
13	0.0866	0.0601	0.1537	0.0372	0.6420	0.0961	0.6755	0.0734	0.0348	0.0921
14	0.2318	0.5137	0.4967	0.2999	0.7121	0.1738	0.9756	0.1395	0.0649	0.4874

TABLE 3.2 : La sensibilité des paramètres du modèle DG dans les échantillons bootstrap

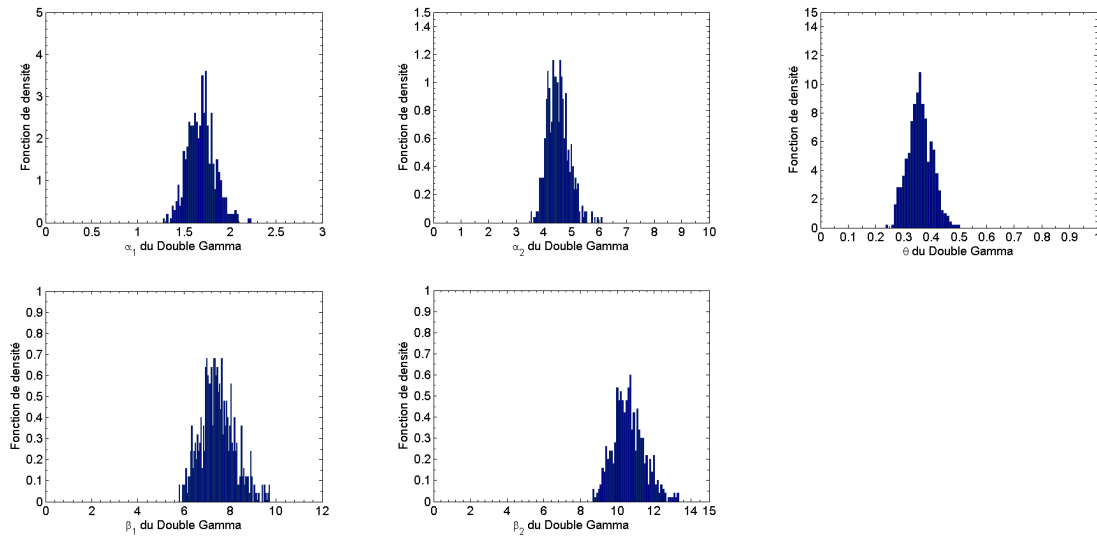


FIGURE 3.17 : Les distributions des paramètres du modèle DG dans les échantillons bootstrap

3.5 Conclusions du chapitre

Dans ce chapitre, l'emploi d'un paramètre de décalage est déconseillé car il modifie les valeurs des autres paramètres des modèles dont certains sont liés à des caractéristiques physiques du trafic. La variabilité des distributions des modèles Γ -GQM et DG est réalisée en fonction de leurs paramètres. Les caractéristiques statistiques sont aussi tracées selon les valeurs des paramètres des modèles. La sensibilité des modèles est obtenue par l'analyse de Bootstrap. Certains paramètres sont beaucoup plus sensibles que les autres au sein d'un même modèle probabiliste.

Chapitre 4

Niveau macroscopique et conditions exogènes

Dans les chapitres précédents, une étude exhaustive de la modélisation de la variable TIV a été réalisée. Différentes approches et différents modèles probabilistes ont été approfondis et développés. La modélisation du TIV peut être accomplie par l'utilisation des modèles complexes, le modèle Γ -GQM et les modèles prolongés comme le modèle Double Gamma (DG), le modèle Double Lognormal (DL), le modèle Lognormal Gamma (LG) et le modèle GQM mixte. Dans ce chapitre et le chapitre suivant, le modèle Γ -GQM et le modèle Double Gamma (DG) sont étudiés car le modèle Γ -GQM est assez reconnu tandis que la simplicité du modèle DG est un grand avantage dans la pratique. Malgré sa performance, le GQM mixte n'est pas appliqué pour l'instant car le nombre de paramètres alourdit le temps de calcul.

À la différence du Chapitre 3, la question centrale de ce chapitre est de savoir comment les paramètres des modèles choisis s'adaptent aux différentes distributions des TIV réels en fonction des variables macroscopiques, à savoir \mathcal{Q}_6 , \mathcal{T}_6 , \mathcal{V}_6 et selon différentes conditions exogènes comme la période dans la journée par temps **sec**, par temps de **pluie**, ou bien dépendent du type de voie sur différents réseaux de route. Pour cela, les paramètres des deux modèles prennent le rôle d'indicateurs mesurant la variation et la sensibilité des comportements de l'ensemble.

4.1 État de l'art

4.1.1 Les facteurs

Le TIV est une variable microscopique associée à chaque véhicule, et résulte directement en un point donné de la route des décisions du conducteur. Dans une section rectiligne et en un point donné, la décision d'un conducteur est un ensemble complexe *de choix* : choix de voie, choix de vitesse, choix d'accélération. Cet ensemble *de choix* dépend de beaucoup de *facteurs* lors de la prise de décision. La connaissance des distributions des TIV sous l'influence de ces divers facteurs doit être approfondie. Rappelons que la distribution des TIV est l'un des indicateurs importants pour décrire et interpréter le comportement de l'ensemble des conducteurs dans le trafic routier.

Aly [65] semble être le premier auteur ayant tenté de classer les *facteurs* affectant les distributions de TIV. En effet, il a récapitulé 5 facteurs principaux dans sa thèse comme suit :

1. Facteur lié à la circulation, \mathcal{A}_1
2. Le comportement du conducteur, \mathcal{A}_2

3. Facteur de la route, \mathcal{A}_3
4. Facteur de contrôle de trafic, \mathcal{A}_4
5. Facteur environnemental, \mathcal{A}_5 .

Les facteurs ont été listés et puis abordés par Aly mais ne sont pas analysés plus en détail. Le facteur lié à la circulation ou au trafic \mathcal{A}_1 , d'après Aly, englobe par exemple le pourcentage des PL dans le trafic, la vitesse moyenne globale, la catégorie de poursuite des véhicules comme véhicule *libre* ou véhicule *contraint*. La catégorie physique du véhicule est aussi considérée comme un facteur du trafic \mathcal{A}_1 , par exemple la distribution des TIV des VP est différente de celle des PL, etc.

Le deuxième facteur \mathcal{A}_2 indique l'état psychologique et/ou physique des conducteurs en fonction de leur possibilité de dépassement, de leur position dans le trafic. Prenons un exemple, un *suiveur* poursuivant un véhicule en tête se comporte différemment selon qu'il a, ou non l'intention de dépasser. Dans le premier cas, le conducteur, dit le *dépasseur* garde en permanence la volonté de doubler le véhicule en tête pour gagner en temps ou en liberté de mouvement. Le facteur \mathcal{A}_2 pourrait comporter le sexe, l'âge, l'expérience des conducteurs (l'accidentologie par exemple), etc.

Le facteur de la route \mathcal{A}_3 concerne l'infrastructure, la qualité de service, la limite de vitesse, le type du réseau routier (autoroute, route nationale, route départementale) ou bien le nombre de voies du tronçon, etc.

Pour le quatrième facteur \mathcal{A}_4 , Aly a indiqué la signalisation et aussi les informations de trafic fournies au conducteur lors du déplacement. Enfin, le facteur environnemental \mathcal{A}_5 est décrit comme la saison, l'environnement de la route, l'ambiance extérieure aperçue, etc.

Le classement d'Aly permet de comprendre la complexité des phénomènes du trafic et incite les chercheurs à préciser les informations dont ils disposent et qu'ils traitent. Le classement peut jouer aussi le rôle de guide pour des études futures en vue de mieux connaître les différents aspects des comportements des conducteurs.

Pourtant, le classement d'Aly est plutôt *qualitatif*. En effet, les facteurs d'Aly sont difficiles à quantifier alors que le TIV est une variable quantitative. Le caractère non mesurable des facteurs rend flous certains facteurs, par exemple celui du comportement de conduite \mathcal{A}_2 , car la distinction des types de conducteurs n'est pas facile à effectuer à partir des données empiriques disponibles. De plus, dans chaque facteur, les caractéristiques des « sous-facteur » sont très variées.

Pour éviter ces défauts du classement d'Aly, nous proposons dans cette étude, de prendre en compte les facteurs organisés de façon hiérarchique. Le classement consiste à distinguer quatre niveaux de facteurs comme ci-dessous :

1. Les facteurs exogènes, \mathcal{F}_1 .
2. Les facteurs macroscopiques, \mathcal{F}_2
3. Les facteurs mésoscopiques, \mathcal{F}_3
4. Les facteurs microscopiques, \mathcal{F}_4 .

Les facteurs exogènes \mathcal{F}_1 regroupent toutes les conditions environnementales à l'extérieur connues hors celles liées au flux du trafic. Il s'agit par exemple de la condition météorologique, du type de réseau routier (autoroute, route nationale, route départementale, etc.), de la saison, de la période dans la journée ou de l'environnement, de l'infrastructure. Le facteur \mathcal{F}_1 inclut également les contextes d'infrastructure comme la signalisation, l'information, les virages, le type de voie et le nombre de voies, etc. Autrement dit, en précisant ce facteur, on répondra aux questions *où* et *quand* la circulation a eu lieu. Ce sont les questions sur la localisation, les caractéristiques de la route, le temps et l'ambiance

de l'environnement. Le mot *exogène* utilisé ici qualifie des facteurs qui ne sont pas calculés à partir des données du trafic.

Les facteurs macroscopiques \mathcal{F}_2 sont des variables quantitatives des flux de trafic calculées par agrégation sur une période de temps I . Les trois variables principales sont le débit agrégé 6 minutes \mathcal{Q}_6 , le taux d'occupation agrégé 6 minutes \mathcal{T}_6 et la vitesse harmonique agrégée 6 minutes \mathcal{V}_6 . Le pourcentage des PL peut aussi être considéré dans ce facteurs.

Les facteurs mésoscopiques \mathcal{F}_3 ont un champ plus circonscrit que celui des facteurs macroscopiques \mathcal{F}_2 . Si les facteurs \mathcal{F}_2 prennent en compte l'ensemble du flux de trafic, les facteurs \mathcal{F}_3 se concentrent sur des relations à l'intérieur ou entre les *pelotons*, ou bien sur des interactions entre les véhicules voisins. Le cas le plus important est l'impact du véhicule de tête sur le comportement des véhicules *suiveurs*. La notion de *peloton* joue un rôle central dans cette optique et est l'objet d'une étude supplémentaire dans le chapitre suivant.

Enfin les facteurs microscopiques \mathcal{F}_4 affectant les véhicules sont liés aux autres informations du véhicule lui-même. La vitesse instantanée est l'une des trois informations individuelles obtenues par le système de mesure en un point donné (les deux autres sont la longueur et le temps d'arrivée du véhicule). Par conséquent, cette variable est étudiée par mise en relation et/ou par association avec le TIV.

On remarque que les facteurs \mathcal{F}_2 , \mathcal{F}_3 , \mathcal{F}_4 sont généralement quantitatifs. Un facteur macroscopique \mathcal{F}_2 important, le débit \mathcal{Q}_6 est lié à la demande du trafic. Cette demande dépend aussi de la période dans la journée, de la saison, etc., donc des facteurs \mathcal{F}_1 . Cela veut dire que dans certains cas, l'information macroscopique pourrait répondre à la question *quand* posée en cherchant les facteurs \mathcal{F}_1 .

Le classement hiérarchique proposé n'est pas complet. Il est focalisé particulièrement sur les informations de base couramment enregistrées. D'autres facteurs importants n'ont pas trouvé leurs place dans le classement comme le motif de déplacement, l'état psychologique ou l'expérience de conducteur, etc. Ces types d'information, n'étant pas disponibles au moins dans nos bases de données, sont exclus dans notre étude. Dans ce chapitre, les deux premiers facteurs \mathcal{F}_1 , \mathcal{F}_2 sont étudiés. Le chapitre suivant sera consacré aux deux derniers \mathcal{F}_3 , \mathcal{F}_4 .

4.1.2 État de l'art

Soit une période de mesure T dans laquelle les TIV H obtenus sont un n -échantillon (h_1, h_2, \dots, h_n) . Le débit Q_T calculé par unité de temps sur la période T est calculé comme suit :

$$Q_T = \frac{n}{T} \quad (4.1)$$

Par ailleurs, la moyenne des TIV dans la période T est :

$$\hat{\mu}_T = \frac{\sum_1^n h_i \times \mathbb{1}_{\{h_i \in T\}}}{n} \quad (4.2)$$

$$= \left[\frac{\sum_1^n h_i}{T} \right] \times \frac{1}{Q_T} \quad (4.3)$$

Dans le cas où $\sum_1^n h_i = T$, c. à. d que la période T est remplie entièrement tout d'abord par l'arrivée du véhicule 1 et en fin par l'arrivée du véhicule n au point de mesure, la

moyenne des TIV $\hat{\mu}_T$ est exactement l'inverse du débit Q_T . Pourtant, si la période T est fixée en avance, comme c'est le cas ($T = 6$ minutes) dans le système de mesure français, le cas idéal se produit rarement. On a donc en général $\mu_T \neq Q_T^{-1}$.

Les calculs dans le Chapitre 1 montrent par ailleurs que, malgré l'erreur numérique, la relation hyperbolique entre μ_T et Q_T est toujours marquée notamment lorsque le débit Q_T est élevé - le cas où plusieurs arrivées dans la période T sont réalisées. En conséquence, la relation entre le débit et la moyenne des TIV est naturellement acceptée dans tous les cas, on a $\mu_T \approx Q_T^{-1}$.

Dans la littérature, le débit Q_T a été considéré comme seul facteur décisif de la distribution des TIV. Deux principales approches concernant le débit sont appliquées. La première consiste en diviser les états du trafic en trois *régimes* : le trafic *fluide*, la *congestion* et le trafic *intermédiaire* (entre les deux premiers). Al-Ghamdi *et al.* en 2001 [105] ont proposé de distinguer les *régimes* du trafic par des seuils quantitatifs comme suit :

Trafic fluide : $Q_T < 400$ véh/h, des phénomènes aléatoires,

Trafic congestionné : $Q_T > 1200$ véh/h, les phénomènes de régularité,

Trafic intermédiaire : $400 \leq Q_T \leq 1200$ véh/h.

Dans cette *approche de classement*, différents types de distributions des TIV sont utilisés correspondant à chacun des régimes du trafic. On revoit donc les propositions de Haight *et al.*, de May [114], et de Krbalek *et al.*

La seconde approche tend à calibrer les TIV par l'utilisation d'un seul modèle *spécifique*. Les paramètres de ce modèle choisis sont ensuite jouent le rôle d'indicateurs pour mesurer la sensibilité du facteur de débit. Ensuite, les paramètres sont calibrés pour chaque niveau du débit et pourraient être représentés par des abaques. Ce sont Grecco *et al.* [60] et ensuite Khasnabis *et al.* [66] qui sont les premiers utilisateurs de cette approche. Cette *approche de calibrage* demande un modèle *spécifique* très performant ce qui n'est pas le cas du modèle Hyperexponentiel \mathcal{E}^* utilisé par Grecco *et al.* [60] et par Khasnabis *et al.* [66].

Grecco *et al.* en 1968 ont calibré les relations entre les paramètres du modèle Hyperexponentiel \mathcal{E}^* et le débit q_T en véh/s. La régression linéaire a été réalisée et une abaque des probabilités des TIV de type $\Pr\{H > h\}$ a été établie. La méthode a été répétée par Khasnabis *et al.* en 1980.

L'approche de calibrage est majoritaire parmi les auteurs dans la littérature. Dawson *et al.* [56] en 1968 ont appliqué le modèle Hyperlang Γ^* à deux bases de données, l'une était la base du HCM 1965 (Highway Capacity Manual) et l'autre était celle de l'Université Purdue, aux États-Unis. Ils constatent que certains paramètres sont sensibles au débit Q_T tandis que les autres ne le sont pas. Par ailleurs, la sensibilité des données est tout à fait différente entre les deux bases de données, par exemple le paramètre κ du modèle d'Erlang des TIV contraints a dépendu de Q_T pour la deuxième base mais en est resté indépendant pour la première base de données.

Les paramètres obtenus pour chaque niveau de débit Q_T dépendent aussi de la performance de la méthode d'estimation utilisée. Buckley en 1968 [57] a utilisé la MM et constaté que certains paramètres des modèles Γ_τ , Γ -SPM, \mathcal{N} -SPM, dont le paramètre β du modèle Γ_τ et les paramètres σ, λ du modèle \mathcal{N} -SPM, ont changé brusquement leur valeur dans l'échantillon ayant la moyenne de TIV de l'ordre de 4 s.

En 1975, des facteurs hors trafic de type \mathcal{F}_1 ont été constatés par Cowan [59] : le nombre de voies qui influe sur les possibilités de dépassement, influence aussi le processus d'arrivée des véhicules en un point donné. Si sur les routes à plusieurs voies, ce processus est un processus de Poisson, le processus des arrivées ne serait plus de Poisson sur les routes (à deux fois une voie par exemple) où le dépassement est restreint. Cette remarque était à l'origine du modèle M4 de Cowan - la première version des modèles g -GQM.

Branston [58] en 1976 a étudié la distribution des TIV sur l'autoroute péri-urbaine M4 à l'Ouest de Londres et sur les routes bidirectionnelles en Indiana aux États-Unis. Il a calibré le modèle $\ln \mathcal{N}$ -GQM et a constaté que les paramètres diminuaient en fonction du débit Q_T et dépendaient également de la vitesse moyenne du trafic. En conséquence, ces valeurs des paramètres ont été réutilisées dans la simulation.

Wasielewski en 1979 [64] a utilisé le modèle Γ -SPM pour estimer la capacité du trafic et tester l'indépendance de la distribution des TIV par rapport au débit Q_T . Wasielewski a constaté que le paramètre λ du modèle Γ -SPM variait linéairement en fonction du débit q (véh/s) :

$$\lambda = \lambda_0 + m q \quad \text{Avec} \quad \lambda_0 = -0,058, \quad m = 1,401 \quad (4.4)$$

Certains travaux traitent l'effet de la pluie sur la distribution des TIV comme celui de Seddiki [127] en 1993. Le travail s'est concentré sur l'ensemble des données en négligeant les différentes distributions de TIV en fonction de la demande du trafic.

Hoogendoorn *et al.* [79] en 1997 ont utilisé le modèle Γ -GQM. Ils ont proposé que tous les paramètres du modèle varient en fonction du débit Q_T . Hoogendoorn *et al.* [80] en 1998 ont aussi étudié l'influence de la période de la journée et du type des véhicules sur la distribution des TIV.

Michael *et al.* [128] en 2000 ont étudié l'influence d'un message de conseil envoyé aux conducteurs sur le TIV. Ce facteur est donc du type \mathcal{F}_1 . Les groupes de conducteurs sont divisés en deux, l'un avec information et l'autre sans information. Michael *et al.* constatent que le pourcentage des TIV courts (inférieurs à 2 s), la moyenne des TIV et la moyenne des DIV dans le groupe de conducteurs informés augmentent tandis que le pourcentage des TIV très courts (inférieurs à 1 s) diminue significativement.

Une autre approche pragmatique de Helmut *et al.* [89], proposée en 2007, consiste à estimer la fonction de répartition des TIV par l'utilisation de formules hyperboliques. L'idée est d'utiliser directement le débit horaire Q_h pour obtenir la distribution des TIV correspondante. Le résultat est ensuite entré dans la simulation du trafic. Cette approche est cependant moins précise et n'a eu que des applications locales limitées.

4.1.3 Discussion

En résumé, le problème de la distribution des TIV dans différents contextes et sous différentes conditions exogènes est encore peu étudié. En réalité, plusieurs distributions des TIV pourraient être obtenues pour un même niveau de débit. Pour interpréter ce phénomène, on est incité à en attribuer la cause aux conditions *exogènes* qui sont liées aux circonstances dans lesquelles les conducteurs conduisent les véhicules, ou lorsque les données de trafic sont collectées. Deux questions principales sont posées : *où* se trouvent les conducteurs des véhicules et *quand* font-ils (elles) partie du flux du trafic ? La question *où* vise le site de mesure, le type de voie alors que la question *quand* concerne la période de la journée et/ou la condition météorologique, etc. Ces deux questions ont pour but de préciser le facteur \mathcal{F}_1 précité.

Faute de données suffisantes dans la littérature, les facteurs macroscopiques du trafic tels que le débit Q_T , le taux d'occupation \mathcal{T}_T n'ont pas été séparés dans les études précédentes lorsque différentes situations exogènes sont mises en jeu. Par ailleurs, d'autres travaux ont montré des effets exogènes statistiquement significatifs sur les TIV courts (par exemple $TH < 2$ s, Aron *et al.* en 1997 [129]), et ces effets étaient eux-même fonction du débit. En conséquence, il est proposé d'étudier au niveau plus général toute la distribution des TIV en fonction des facteurs exogènes.

Peu d'auteurs jusqu'à maintenant ont tenté d'étudier l'impact de la concentration du trafic K sur la distribution des TIV. La concentration, mesurée en espace par exemple

à partir de photographies aériennes, est aussi une variable macroscopique essentielle du trafic. Son caractère spatial semble intuitivement empêcher les chercheurs de la mettre en relation avec la distribution du TIV. Pourtant, il est noté que la concentration peut être estimée par l'intermédiaire du taux d'occupation \mathcal{T}_T (cf. Cassidy [99]) qui est mesurable par des capteurs routiers en un point donné. En conséquence, le taux d'occupation \mathcal{T}_T sera un des objets d'étude de ce chapitre.

Dans la première partie de ce chapitre, la distribution des TIV est étudiée en prenant en compte le débit 6 minutes Q_6 , le taux d'occupation 6 minutes \mathcal{T}_6 et la vitesse moyenne harmonique 6 minutes \mathcal{V}_6 . Par ailleurs, les distributions des TIV sont aussi examinées dans différents états élémentaires e_{ij} du trafic dans le *diagramme fondamental* empirique. La notion d'état élémentaire correspond à un couple Débit–Taux d'occupation (Q_6, \mathcal{T}_6) représenté par un point du diagramme fondamental empirique (ou plutôt par un carré si le couple correspond à un couple de classes de débit et de taux d'occupation).

Dans la suite de ce chapitre, les effets de la condition météorologique et du type de voie sont examinés en même temps que ceux de la circulation Q_6 . Le modèle Γ -GQM et le modèle DG sont utilisés en tant que les modèles *spécifiques* de TIV. Tout au long de notre étude, ces deux modèles fournissent de très bons ajustements pour la majorité des échantillons au niveau de signification de 5 % du test K–S.

4.2 Distribution des TIV à l'échelle macroscopique (les échantillons groupés RN118)

4.2.1 Débit 6 minutes Q_6

Le modèle Γ -GQM

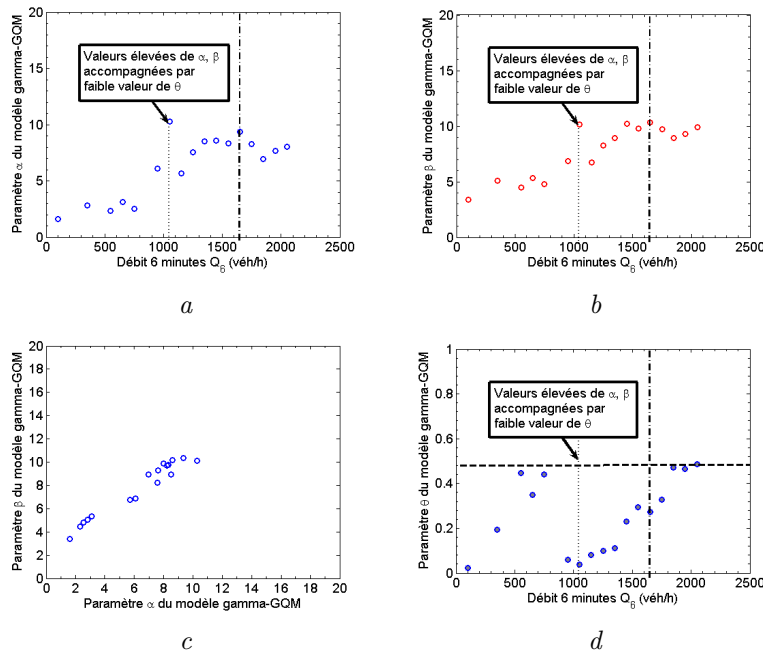


FIGURE 4.1 : Variations des paramètres du modèle Γ -GQM en fonction de Q_6

Les Figures 4.1-a,b,d illustrent les variations des paramètres α, β et θ du modèle Γ -GQM selon le débit Q_6 dans les échantillons ayant les P-valeurs supérieurs à 0,05. Dans ces cas-là, le modèle Γ -GQM n'est pas rejeté au niveau de signification de 5 % du test K–S.

Les données proviennent de la période jour non-pluie sur la VL. La Figure 4.1-c donne la relation entre les paramètres α et β .

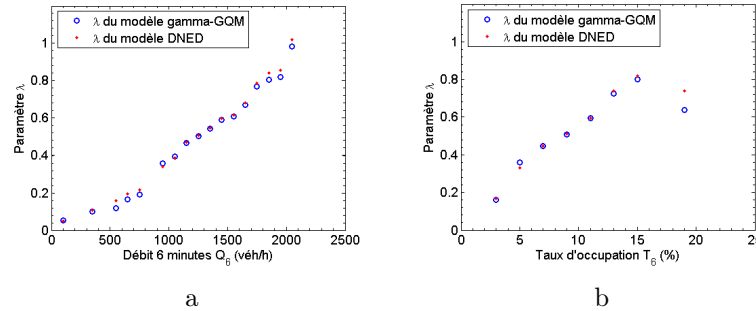


FIGURE 4.2 : Variation de λ du modèle Γ -GQM en fonction de Q_6 et de \mathcal{T}_6

Le modèle Γ -GQM fournit de bons ajustements pour les échantillons de l'état du trafic fluide. De faibles valeurs de α, β, θ (avec un paramètre θ convenable) suffisent pour que l'ajustement du modèle Γ -GQM aux TIV de l'échantillon (0-200) véh/h soit meilleur que celui du modèle \mathcal{E}_τ (les P-valeurs sont respectivement de 0,8588 et 0,8425). On remarque à partir des Figures 4.1 et 4.2-a que :

- Les paramètres α et β augmentent en fonction du débit Q_6 pour les échantillons de débit inférieur à 1700 véh/h . Ces deux paramètres varient de façon similaire et il y a une corrélation entre eux.
- Le paramètre θ varie en deux phases et le changement de phase se trouve au niveau du débit Q_6 intermédiaire entre 800 à 1000 véh/h . Toutes les valeurs obtenues de θ sont inférieures à 0,5.
- Le paramètre λ augmente en fonction du débit Q_6 pour tous les échantillons. Les valeurs estimées de λ du modèle Γ -GQM sont proches des paramètres estimés λ du modèle \mathcal{E}_τ .

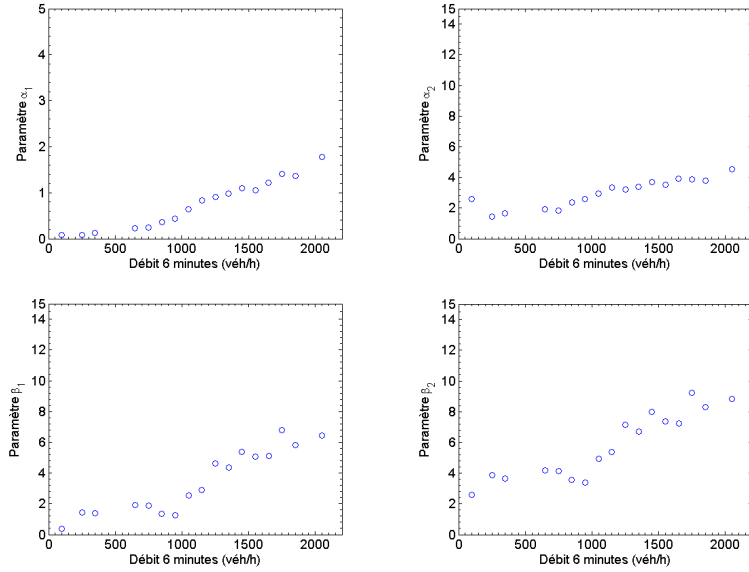
Les valeurs particulières de α, β et θ sont trouvées dans les échantillons de débit élevé ($Q_6 > 1700 \text{véh/h}$) et dans l'échantillon de Q_6 de (1000-1100) véh/h . Il est remarqué dans ces échantillons qu'une faible valeur de θ est associée à des valeurs élevées des paramètres α et β ; et pour les débits élevés, la réduction des paramètres α et β (c.à.d comparé aux valeurs correspondantes obtenues pour l'échantillon 1700 véh/h) est accompagnée par des valeurs élevées du paramètre θ (proche de 0,5).

Le modèle Double Gamma DG

La Figure 4.3 représente la variation des paramètres $\alpha_1, \alpha_2, \beta_1$ et β_2 du modèle DG dans les cas où ce modèle n'est pas rejeté par le test K-S au niveau de signification de 5%. On observe que :

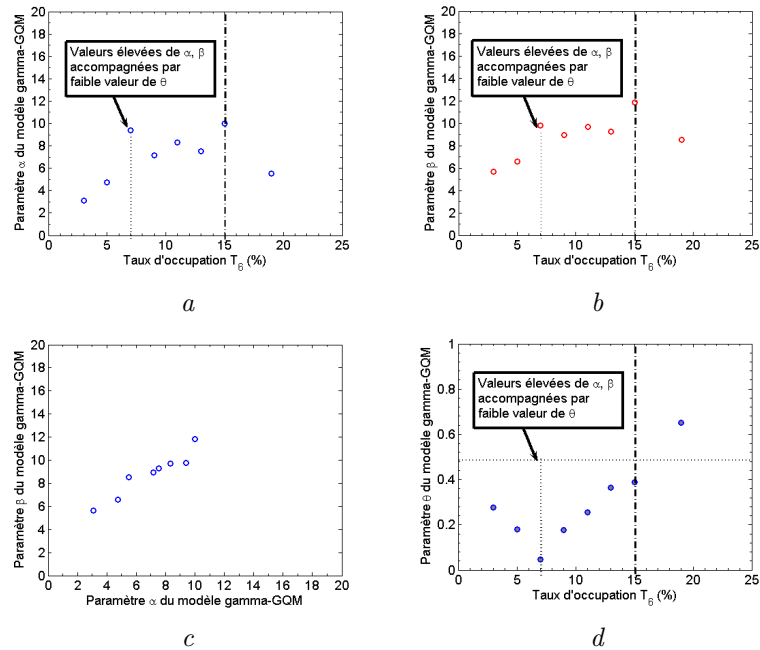
- Les quatre paramètres augmentent en fonction du débit Q_6 . Les variations de ces paramètres semblent ne pas avoir de point de discontinuité pour le trafic dense.
- Les paramètres β_1 et β_2 varient de la même façon.
- Le paramètre θ varie en deux groupes de deux phases de la même forme, chaque groupe comportant une phase de croissance suivie d'une phase de décroissance de θ en fonction du débit Q_6 .

À la différence du modèle Γ -GQM, les valeurs faibles du paramètre θ du modèle DG pour l'échantillon de débit 1000 véh/h correspondent plutôt aux valeurs faibles des paramètres β_1, β_2 (cf. Figure 4.8).

FIGURE 4.3 : Variations des paramètres du modèle DG en fonction de Q_6

4.2.2 Taux d'occupation 6 minutes \mathcal{T}_6

Le modèle Γ -GQM

FIGURE 4.4 : Variations des paramètres du modèle Γ -GQM en fonction de \mathcal{T}_6

Les Figures 4.4-a,b,d représentent les variations des paramètres α, β et θ en fonction du taux d'occupation \mathcal{T}_6 dans les cas où le modèle Γ -GQM n'est pas rejeté au niveau de signification de 5 % du test K-S. La Figure 4.4-c montre la corrélation entre les paramètres α et β . Les figures 4.4 et 4.2-b donnent que :

- Les paramètres α et β du modèle Γ -GQM augmentent en fonction de \mathcal{T}_6 dans les cas où \mathcal{T}_6 est inférieur à 16 %. Les deux paramètres varient de la même façon et sont fortement corrélés.

- Le paramètre θ varie en deux phases dont l'une est croissante et l'autre décroissante en fonction de \mathcal{T}_6 . Le seuil du changement de phase se trouve au niveau de \mathcal{T}_6 de 8 %. Les valeurs de θ sont inférieures à 0,5 dans la plupart des échantillons sauf celui du \mathcal{T}_6 de (18-20) %.
- Le paramètre λ croît linéairement en fonction de \mathcal{T}_6 tant que celui-ci reste inférieur à 16 %. Les valeurs estimées de λ du modèle Γ -GQM et du modèle \mathcal{E}_τ sont proches pour les échantillon de \mathcal{T}_6 inférieur à 16 %.

Les variations des paramètres du modèle Γ -GQM obtenues sont similaires à celles obtenues dans la section 4.2.1 pour le débit Q_6 . Les deux cas particuliers de α , β et θ sont aussi observés, l'un au niveau de \mathcal{T}_6 élevé et l'autre au niveau de \mathcal{T}_6 intermédiaire. Dans ces deux cas, la valeur relativement faible de θ est associée à des valeurs élevées des paramètres α et β et vice versa. L'échantillon de \mathcal{T}_6 (18-20) % a une valeur élevée de θ , $\theta = 0,6516$.

Le modèle DG

La Figure 4.5 illustre la variation des paramètres α_1 , α_2 , β_1 et β_2 du modèle DG dans les cas où ce modèle n'est pas rejeté par le test K-S au niveau de signification de 5 %. Comme dans la section 4.2.1 où l'analyse était faite en fonction du débit Q_6 , on remarque que :

- Les quatre paramètres α_1 , α_2 , β_1 et β_2 augmentent en fonction de \mathcal{T}_6 .
- Les paramètres β_1 et β_2 varie parallèlement, on a approximativement $\beta_2 = \beta_1 + 2$.
- Le paramètre θ diminue en fonction de \mathcal{T}_6 , il varie d'une valeur d'environ 0,8 à une valeur d'ordre 0,2.

Les tendances obtenues des paramètres du modèle DG sont visiblement nettes bien que le modèle n'ait pas donné de bons ajustements pour certains échantillons de TIV de \mathcal{T}_6 supérieur à 16 %.

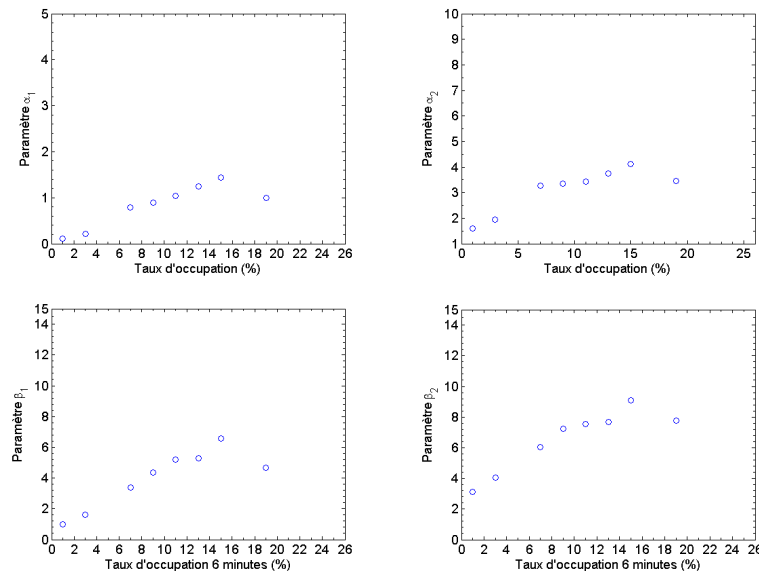


FIGURE 4.5 : Variations des paramètres du modèle DG en fonction de \mathcal{T}_6

4.2.3 Vitesse moyenne harmonique 6 minutes \mathcal{V}_6

L'échantillonnage des TIV selon la vitesse moyenne harmonique 6 minutes \mathcal{V}_6 a été réalisé, ainsi que le calcul des paramètres des modèles.

Le modèle Γ -GQM

Les Figures 4.6-a,b,d montrent les variations des paramètres α , β , θ et λ du modèle Γ -GQM en fonction de la vitesse moyenne \mathcal{V}_6 dans les cas où le modèle Γ -GQM ne peut pas être rejeté au niveau de signification de 5 % du test K-S. On observe que :

- Les paramètres α et β augmentent et varient de la même façon selon la vitesse moyenne \mathcal{V}_6 .
- Le paramètre θ décroît en fonction de la vitesse moyenne \mathcal{V}_6 , de 0,9 (dans l'échantillon de vitesse faible (2-4) m/s) à une valeur proche de zéro dans l'échantillon de vitesse \mathcal{V}_6 de (26 - 28) m/s. Le paramètre θ du modèle Γ -GQM semble avoir un lien direct avec vitesse \mathcal{V}_6 .
- Le paramètre λ du modèle Γ -GQM ainsi que sa variation en fonction de la vitesse \mathcal{V}_6 ne sont proches de celles du paramètre λ de modèle DNED τ que pour les échantillons de vitesse \mathcal{V}_6 supérieure à 16 m/s.

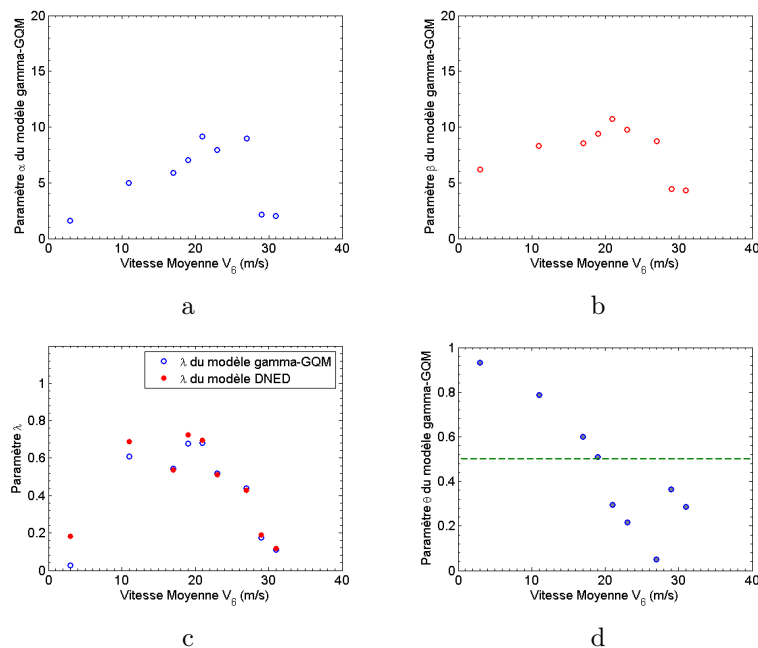


FIGURE 4.6 : Variations des paramètres du modèle Γ -GQM en fonction de \mathcal{V}_6

Le modèle DG

Le modèle DG fournit de bons ajustements pour les échantillons de TIV classés selon les valeurs de vitesse \mathcal{V}_6 . La Figure 4.7 montre que les paramètres α_1 , α_2 , β_1 et β_2 diminuent en fonction de la vitesse \mathcal{V}_6 pour les échantillons de vitesse \mathcal{V}_6 supérieure à 22 m/s dans lesquels le paramètre θ est de l'ordre 0,6 (cf. la Figure 4.8). Les paramètres α_1 , α_2 varient de la même façon. Idem pour les deux paramètres β_1 , β_2 .

4.2.4 Le Diagramme fondamental

Les états élémentaires

Les Figures 4.9 et 4.10 illustrent le diagramme fondamental empirique des données de la RN118. Cinq états élémentaires du trafic examinés sont représentés par des carrés. L'état N°1 représente le trafic fluide de débit Q_6 et de taux d'occupation \mathcal{T}_6 faibles. L'état N°2 représente un état intermédiaire du trafic. L'état N°3 se trouve juste avant la zone critique.

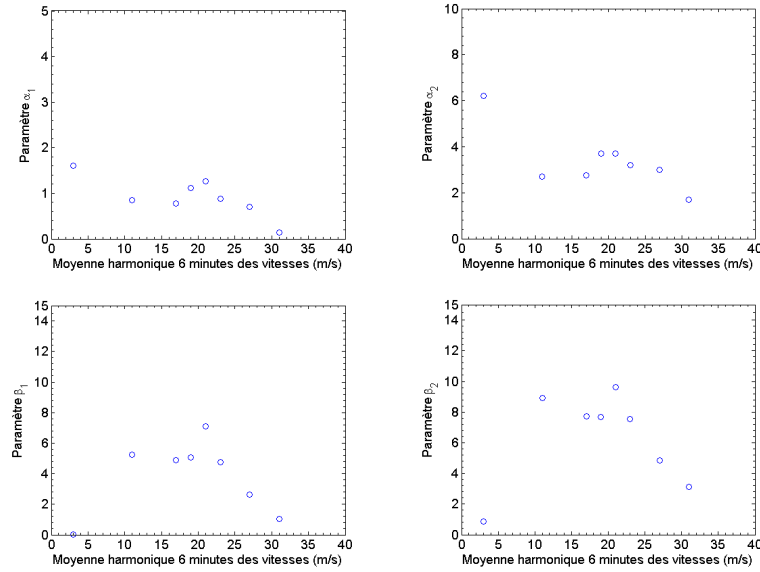


FIGURE 4.7 : Variations des paramètres du modèle DG en fonction de \mathcal{V}_6

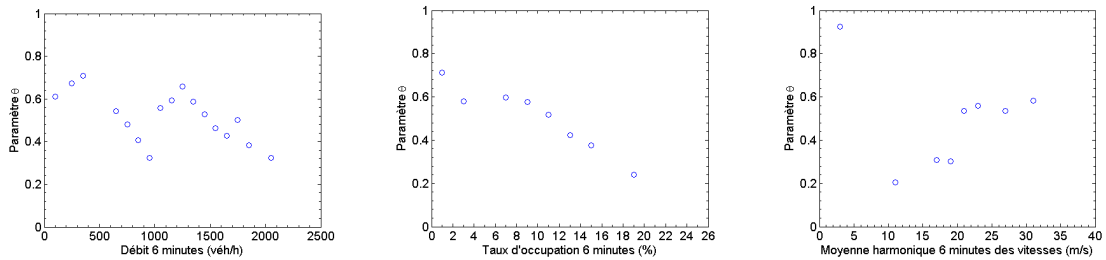


FIGURE 4.8 : Variations de θ du modèle DG en fonction de Q_6 , \mathcal{T}_6 , et \mathcal{V}_6

Et finalement, les états N°4 et N°5 sont dans la zone critique où la capacité de la voie est atteinte et où la congestion commence à se produire.

État N°1 $Q_6 = (500 - 600)$ véh/h, $\mathcal{T}_6 = (2 - 4)$ %

État N°2 $Q_6 = (1100 - 1200)$ véh/h, $\mathcal{T}_6 = (8 - 10)$ %

État N°3 $Q_6 = (1600 - 1700)$ véh/h, $\mathcal{T}_6 = (10 - 12)$ %

État N°4 $Q_6 = (> 2000)$ véh/h, $\mathcal{T}_6 = (16 - 18)$ %

État N°5 $Q_6 = (1900 - 2000)$ véh/h, $\mathcal{T}_6 = (20 - 22)$ %

De l'état N°1 à l'état N°3, les paramètres α , β et λ du modèle Γ -GQM augmentent. Le paramètre θ est égal à 0 dans l'état N°2 et augmente ensuite jusqu'à l'état N°4. Du N°3 au N°4, les paramètres α et β diminuent énormément, ensuite de l'état N°4 à l'état N°5, α et β cette fois-ci augmentent énormément bien que ces deux états sont très proches, tandis que λ augmente malgré une légère réduction au niveau de Q_6 . La valeur estimée du paramètre θ dans l'état N°4 est supérieure à 0,5 et la plus élevée des cinq états élémentaires.

Le paramètre λ semble être lié non seulement au débit Q_6 mais aussi au taux d'occupation \mathcal{T}_6 en trafic congestionné. De plus, ce paramètre influence fortement la pente descendante de la distribution des TIV longs, et croît tout au long du diagramme fondamental. En outre, des valeurs élevées de P-valeur du test K-S semblent indiquer que le modèle Γ -GQM aurait plus de difficulté pour modéliser des échantillons ayant un $CV > 1$ que ceux ayant un $CV < 0,6$.

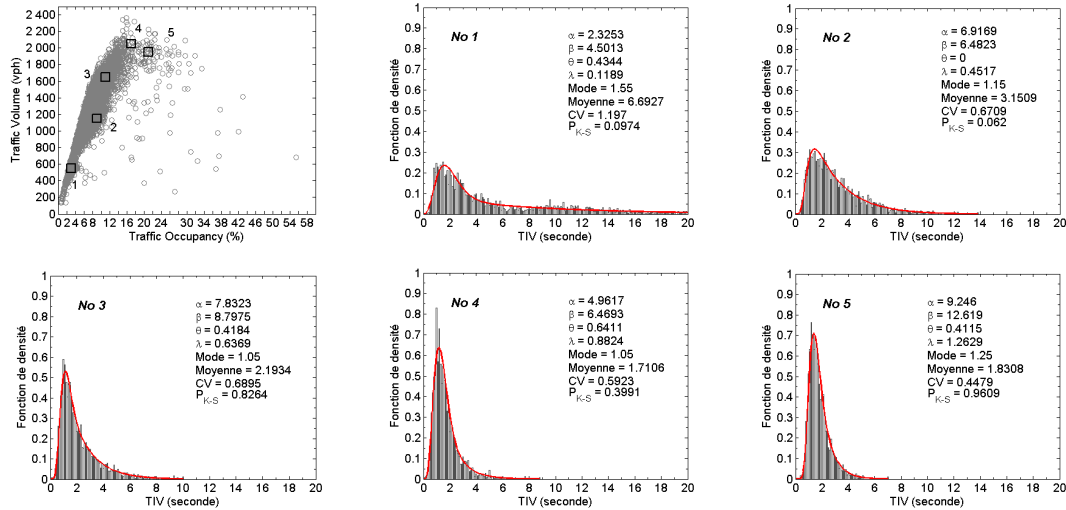


FIGURE 4.9 : Le diagramme fondamental et l'ajustement du modèle Γ -GQM pour les états élémentaires du trafic

La Figure 4.10 représente les ajustements du modèle DG pour les états élémentaires du trafic du diagramme fondamental. On observe que les paramètres α_1 , α_2 , β_1 et β_2 augmentent généralement tout au long du diagramme fondamental sauf le paramètre α_2 à l'état N°4. Cependant, le paramètre θ varie de 0,37 à 0,62 de façon irrégulière.

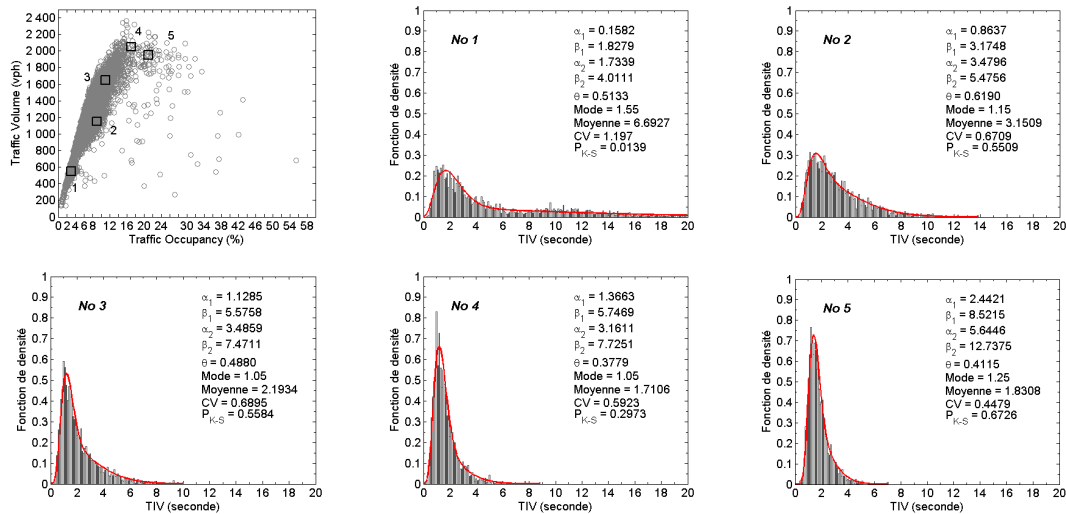


FIGURE 4.10 : Le diagramme fondamental et l'ajustement du modèle DG pour les états élémentaires du trafic

La sensibilité des paramètres

On remarque sur le diagramme fondamental empirique qu'il y a différents états élémentaires correspondant à une même valeur de débit Q_6 ou de taux d'occupation \mathcal{T}_6 . Cela incite à développer l'étude de la sensibilité des paramètres des modèles Γ -GQM et DG selon les Q_6 et \mathcal{T}_6 . Les résultats obtenus montrent que les modèles Γ -GQM et DG sont encore puissants pour la majorité des échantillons échantillonnés selon les états élémentaires du trafic. La Figure 4.11 et la Figure 4.12 présentent les variations des paramètres du modèle Γ -GQM pour les différents états élémentaires respectivement au même débit Q_6 ou taux

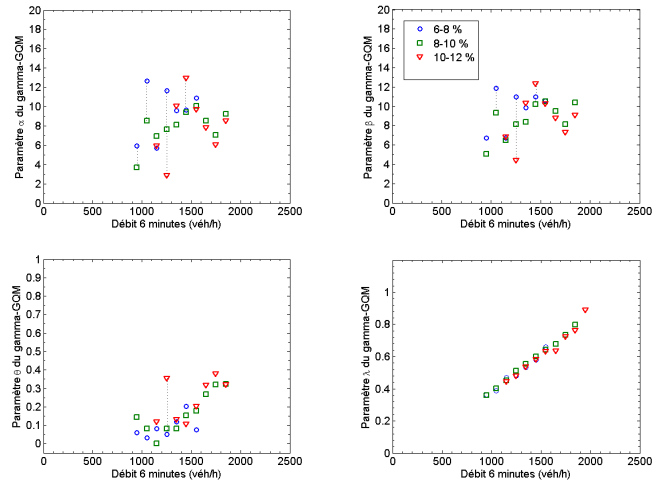


FIGURE 4.11 : La sensibilité des paramètres du modèle Γ -GQM, selon \mathcal{T}_6

d'occupation \mathcal{T}_6 . Idem pour la Figure 4.13 et la Figure 4.14 du modèle DG.

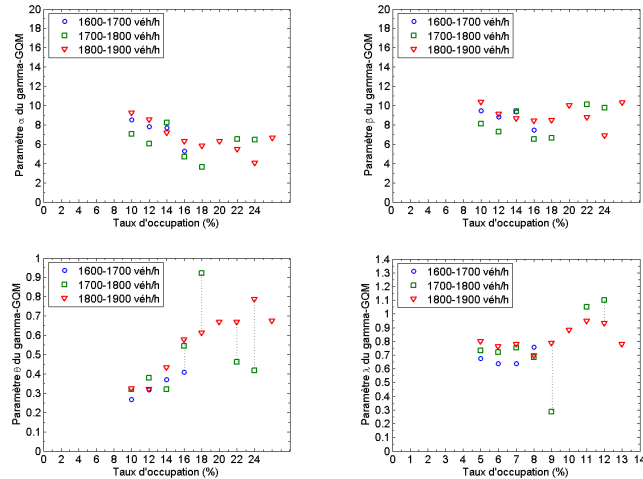


FIGURE 4.12 : La sensibilité des paramètres du modèle Γ -GQM, selon Q_6

La Figure 4.11 montre que les variations des paramètres du modèle Γ -GQM en fonction du débit Q_6 sont similaires à celles obtenues dans les sections précédentes concernant le débit Q_6 . Par exemple, de petites valeurs du paramètre θ au débit Q_6 intermédiaire; les variations similaires des paramètres α et β sont retrouvées. Cependant, α et β sont les paramètres du modèle Γ -GQM les plus sensibles car leurs variabilités selon le niveau de \mathcal{T}_6 sont élevées. Au contraire, le paramètre λ du modèle Γ -GQM semble ne pas dépendre de \mathcal{T}_6 .

La Figure 4.12 montre qu'en fonction du taux d'occupation \mathcal{T}_6 , les variations des paramètres du modèle Γ -GQM sont fortement différentes de celles obtenues précédemment dans les sections concernant le \mathcal{T}_6 . Par exemple, les paramètres α et β décroissent légèrement, le paramètre λ varie selon une courbe plutôt que selon une droite en fonction de \mathcal{T}_6 .

La Figure 4.13 montre qu'en fonction du débit Q_6 , les variations des paramètres $\alpha_1, \alpha_2, \beta_1, \beta_2$ et θ du modèle DG sont similaires à celles obtenues dans les sections précédentes concernant le débit Q_6 . Ces paramètres augmentent en fonction de Q_6 et leur sensibilité selon \mathcal{T}_6 est généralement moins élevée que celle des paramètres α, β du modèle Γ -GQM, en particulier les paramètres α_1 et α_2 semblent indépendants de \mathcal{T}_6 .

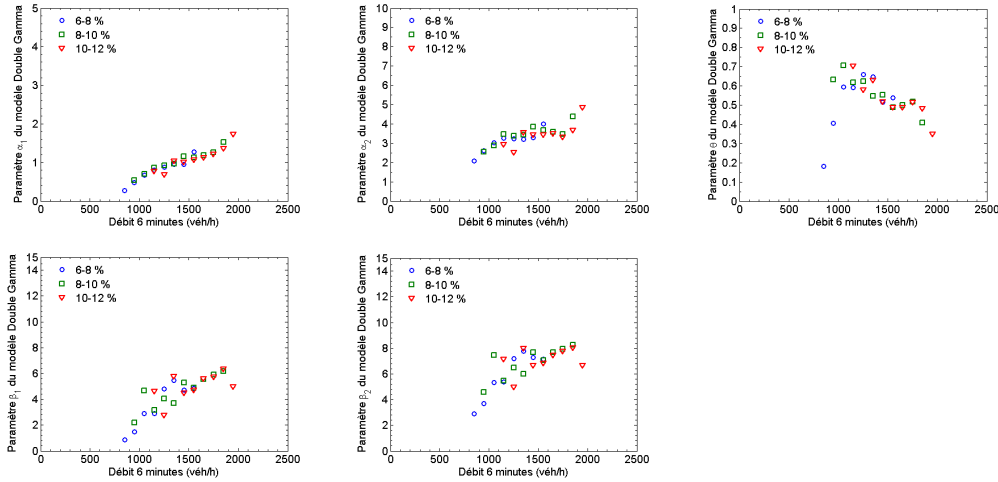


FIGURE 4.13 : La sensibilité des paramètres du modèle DG, selon \mathcal{T}_6

La Figure 4.14 montre qu'en fonction du taux d'occupation \mathcal{T}_6 , les variations des paramètres du modèle DG sont diverses et fortement différentes de celles obtenues précédemment dans les sections concernant \mathcal{T}_6 . Il semble confirmé que les paramètres $\alpha_1, \alpha_2, \beta_1, \beta_2$ dépendent peu du taux d'occupation \mathcal{T}_6 .

4.2.5 Discussion sur les paramètres du modèle Γ -GQM

La relation entre les paramètres α et β du modèle Γ -GQM est significative dans tous les échantillons pour lesquels le modèle Γ -GQM fournit de très bons ajustements. Le résultat est consistant avec ceux de Luttinen [11] et de Hoogendoorn [80]. Il implique que la moyenne des TIV *constraints* U ($\mathbb{E}[U] = \beta/\alpha$) est assez constante ce qui semble ne pas contredire en première analyse l'hypothèse de Branston que la distribution des TIV *constraints* U est indépendante du débit Q_T du trafic. Il est à noter pourtant que la variance $\text{Var}[U] = \beta/\alpha^2$ des TIV *constraints* U pourrait diminuer en fonction de Q_T . Par ailleurs, si la capacité du trafic est calculée par l'inverse de la moyenne des TIV *constraints* U , cette estimation semble constante et donc explicable parce que la capacité ne dépend pas de la demande Q_T du trafic. La relation entre α et β du modèle Γ -GQM est le résultat de l'adaptation de la fonction g du modèle Γ aux données de TIV.

En fonction du débit Q_6 et du taux d'occupation \mathcal{T}_6 , les paramètres α et β augmentent dans la partie gauche du diagramme fondamental. Dans cette partie, le débit Q_6 est proportionnel à \mathcal{T}_6 malgré de larges variations des données empiriques (cf. la Figure 4.9). Cela cause probablement les variations similaires de α et β observées dans les sections précédentes concernant Q_6 et \mathcal{T}_6 . Le résultat supplémentaire de la sensibilité de α et β renforce donc l'influence majeure du débit Q_6 sur la distribution des TIV.

Comme α et λ sont les paramètres d'échelle du modèle Γ -GQM, la comparaison des propriétés statistiques des états élémentaires du N°1 au N°3 de la Figure 4.9 conduit à proposer que l'augmentation de α et λ en fonction de Q_6 dans la majorité des échantillons en trafic fluide et intermédiaire est due à la réduction de la moyenne des TIV. Pour cette raison, le débit Q_T est confirmé comme le facteur macroscopique de type \mathcal{F}_2 le plus important affectant l'échelle et la forme de la distribution des TIV.

La relation entre θ du modèle Γ -GQM et la vitesse \mathcal{V}_6 (cf. la Figure 4.6), ainsi que la relation par morceaux de θ avec le débit Q_6 indiquent que θ concerne le pourcentage des véhicules contraints. En effet, une faible valeur de \mathcal{V}_6 survient en trafic dense qui comporte des véhicules contraints tandis qu'une valeur élevée de \mathcal{V}_6 signifie un trafic fluide avec plus

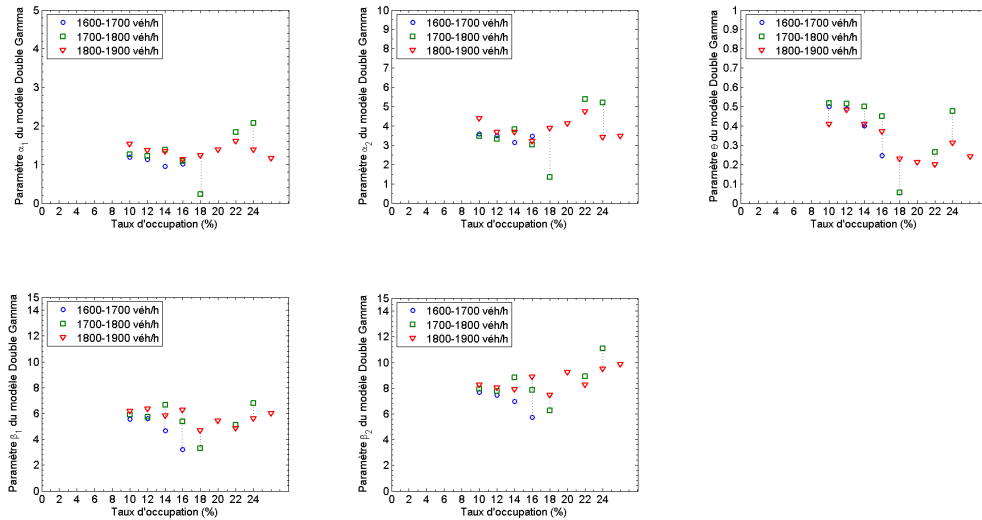


FIGURE 4.14 : La sensibilité des paramètres du modèle DG, selon Q_6

de véhicules libres, donc moins de présence des véhicules contraints. Il semble que θ est lié plus directement à la vitesse \mathcal{V}_6 qu'au débit Q_6 ou au taux d'occupation \mathcal{T}_6 . Pourtant, le paramètre θ peut prendre des valeurs faibles non seulement en trafic fluide mais aussi en trafic intermédiaire. Dans ce dernier cas, une valeur faible de θ est associée en général aux valeurs particulièrement élevées de α et β . De plus, dans la zone critique de trafic congestionné, les paramètres α et β décroissent et/ou fluctuent au lieu d'augmenter, et leurs valeurs sont associées à une valeur élevée de θ (aux alentours ou supérieure à 0,5). En conséquence, à côté de la relation entre α et β précitée, il semble exister une *interdépendance* entre les paramètres α , β et θ du modèle Γ -GQM. L'interprétation du paramètre θ comme la proportion des véhicules contraints n'est pas suffisante et il ne faut pas utiliser θ pour estimer le pourcentage d'un type spécifique de véhicules.

Le paramètre λ du modèle Γ -GQM a une relation étroite avec le trafic, particulièrement le débit Q_6 . Étant l'intensité des arrivées des véhicules libres dans le modèle de mélange GQM, λ joue en effet un rôle important dans l'ajustement de la distribution des TIV longs. En outre, la proximité des valeurs estimées de λ du modèle Γ -GQM et de celles de λ du modèle \mathcal{E}_T incite à aussi utiliser ce paramètre pour représenter (l'inverse du) débit Q_T du trafic.

En somme, l'existence de relations entre les paramètres du modèle Γ -GQM (et aussi du modèle DG) et les variables macroscopiques Q_6 , \mathcal{T}_6 , \mathcal{V}_6 ouvre la possibilité d'exprimer directement la distribution des TIV en fonction des facteurs globaux. Cependant, le lien n'est pas univoque, les interdépendances entre les paramètres entraînent éventuellement des perturbations numériques lors de l'estimation des paramètres à partir des données, ce qui empêche des interprétations pour tous les phénomènes liés aux TIV.

4.3 Les effets exogènes sur la distribution des TIV

La Figure 4.15 présente différentes statistiques descriptives selon les conditions de météo, à savoir par temps *sec* et par temps de *pluie*. La Figure 4.16 présente les comparaisons entre les statistiques descriptives selon le numéro de la voie, VL ou VR de la RN118. La condition météorologique et le type de voie sont les deux conditions *exogènes* considérées dans cette section.

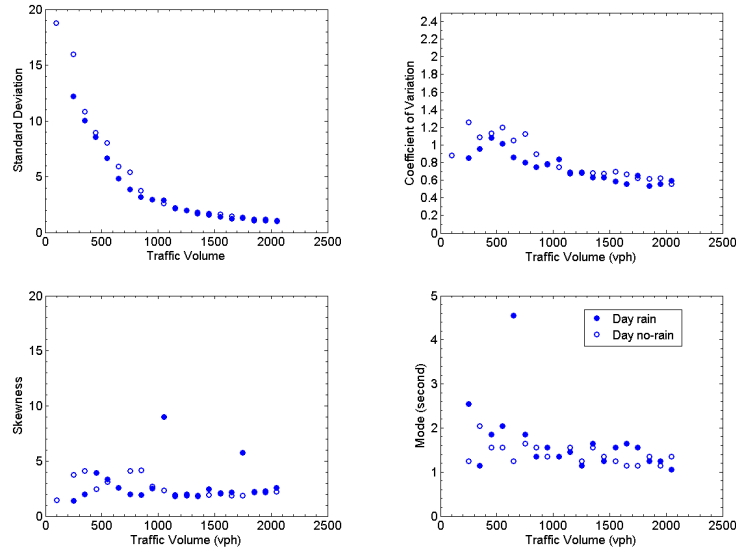


FIGURE 4.15 : Comparaison des statistiques par temps sec et par temps de pluie

À débit constant, la *pluie* entraîne en général un accroissement du *mode* des TIV. Comme la moyenne des TIV est peu changée à débit constant (cf. Équation 4.3), cet accroissement est lié à une diminution de l'écart-type des TIV. Ceci est plus visible dans la Figure 4.16 donnant le coefficient de variation (CV). À débit égal, le *mode* des TIV est plus élevé sur la VL ; l'écart-type et le coefficient de variation y sont moins élevés.

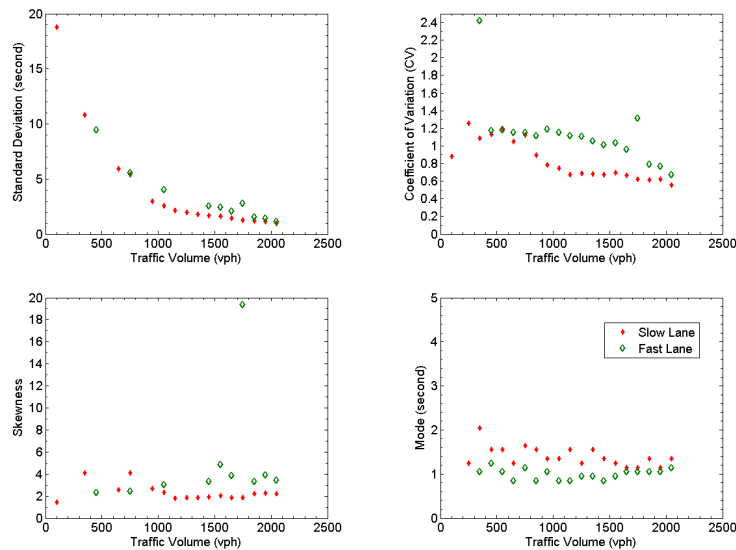


FIGURE 4.16 : Comparaison des statistiques des TIV sur la VL et sur la VR

4.3.1 Conditions de météo sur la VL

L'écart-type σ_H des TIV diminue en fonction du débit Q_6 . Au même niveau du débit Q_6 , σ_H dans la période **sec** est en général plus élevé que dans la période de **pluie** (sauf seulement dans les deux échantillons (1000-1100) et (>2000) véh/h). Au même niveau de débit Q_6 , la distance la plus élevée entre les écarts-types σ_H des TIV des deux conditions de météo est de 0,5 à 4 s dans les échantillons de débit Q_6 inférieur à 900 véh/h. Au même niveau de débit Q_6 , le CV dans la condition **sec** est plus élevé que dans la condition de **pluie**. Le CV est élevé en trafic fluide et diminue en fonction du débit Q_6 ce qui traduit la contrainte augmentant avec le débit Q_6 . Au même niveau de débit Q_6 , il n'y pas de différence significative du coefficient de symétrie S_3 entre les deux conditions de météo. Le coefficient S_3 est inférieur à 4,5 dans la majorité de cas et assez stable (entre 1,75 à 2,5) dans les échantillons de débit Q_6 supérieur à 900 véh/h. Le *mode* observé est situé entre 1 et 2 s dans la majorité des échantillons des TIV sur la VL.

4.3.2 Le modèle Γ -GQM selon les conditions de météo

Le modèle Γ -GQM fournit de bons ajustements pour quasiment tous les échantillons (cf. Table 4.1). La Figure 4.17 représente les variations des paramètres du modèle Γ -GQM selon les conditions de météo et en fonction du débit Q_6 . On constate que :

- Les tendances croissantes des paramètres α et β jusqu'à un certain niveau de débit Q_6 (représenté par les lignes l_1, l_2) sont obtenues. Dans cette partie, au même niveau de débit Q_6 , les valeurs estimées de α, β sous la condition de **pluie** sont en général plus élevées que celles dans la condition **sec**. La pente croissante de la tendance est aussi supérieure dans la condition de **pluie** que dans la condition **sec**.
- Le paramètre θ varie encore en deux phases dans la condition de **pluie**. Cependant, en deuxième phase, le paramètre θ est plus variable dans la condition de **pluie**. Le changement de phase par temps de **pluie** semble plus précoce (au débit Q_6 de 700 véh/h) que par temps **sec** (au débit Q_6 de 1000 véh/h).
- Le paramètre λ est fortement corrélé au débit dans les deux conditions de météo. Au même niveau de débit Q_6 , les valeurs estimées de λ sous la **pluie** ne sont pas plus petites que dans la condition **sec**.

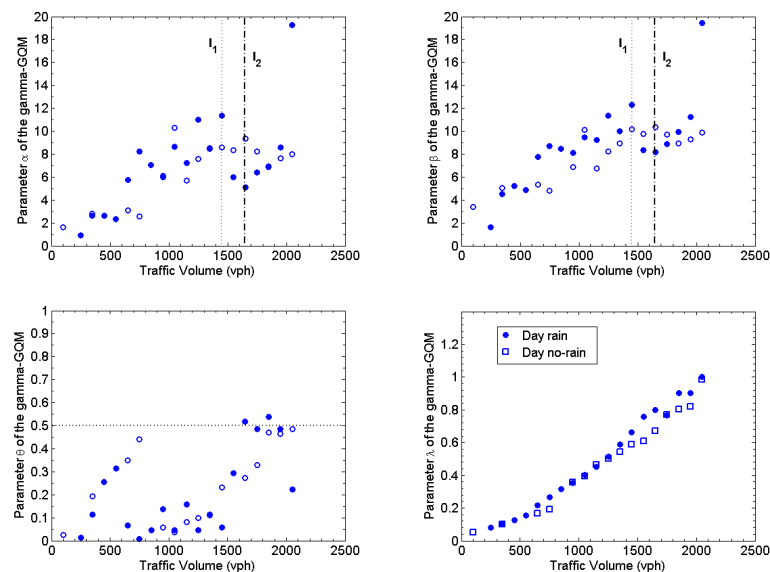


FIGURE 4.17 : Les paramètres du modèle Γ -GQM selon les conditions de météo

4.3.3 Le modèle DG selon les conditions de météo

On obtient que la pluie a une influence non négligeable sur les paramètres estimés du modèle DG. En général, les paramètres α_1 et surtout α_2 par temps de pluie sont supérieurs aux mêmes paramètres estimés par temps sec pour le même débit Q_6 . Pourtant, aucune tendance n'est observable sur les variations des paramètres β_1, β_2 entre les deux conditions de météo.

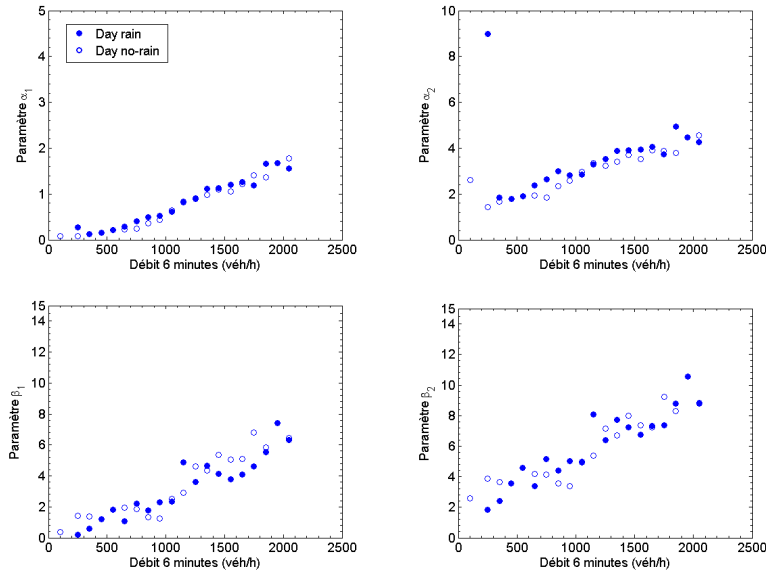


FIGURE 4.18 : Les paramètres du modèle DG selon les conditions de météo

La Figure 4.19 représente le paramètre θ en fonction du débit Q_6 selon les conditions de météo sur la VL (sec et pluie) et selon la voie (VL et VR) dans la période jour-sec. Par temps de pluie, le paramètre θ est inférieur à ce même paramètre par temps sec sur la VL, particulièrement pour les échantillons de débit Q_6 supérieur à 1000 véh/h.

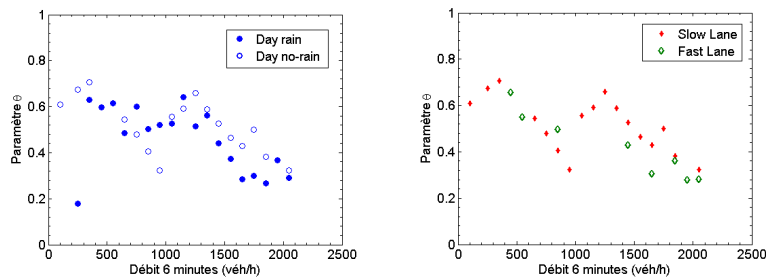


FIGURE 4.19 : Le paramètre θ du modèle DG sous influence des conditions exogènes (météo, type de voie)

4.3.4 Comparaison entre les voies par temps sec de jour

La Figure 4.16 montre que, au même niveau du débit Q_6 , l'écart-type σ_H des TIV sur la VR est supérieur à l'écart-type des TIV sur la VL, et leurs différences sont plus larges que celles correspondant aux conditions de météo. Au même niveau de débit Q_6 , le CV des TIV sur la VR est plus élevé que celui sur la VL. De plus, beaucoup d'échantillons de TIV sur la VR ont un CV supérieur à 1. Pour les échantillons de TIV ayant un débit Q_6 supérieur à 1000 véh/h, le coefficient de symétrie S_3 des TIV sur la VL est plus petit et

plus stable que celui sur la VR tandis que le *mode* des TIV sur la VL est plus grand que sur la VR.

4.3.5 Le modèle Γ -GQM selon les voies par temps sec de jour

Le Table 4.1 fournit les P-valeurs du test K-S du modèle Γ -GQM pour 20 échantillons de TIV sur la VL de modalité jour-sec et 20 échantillons de TIV sur la VR de modalité jour-sec. Pour respectivement 17 échantillons de TIV sur la VL et 12 échantillons de TIV sur la VR (tous de la même modalité jour-sec), les P-valeurs sont supérieures à 0,05, ce qui indique de bons ajustements au niveau de signification de 5 % au sens du test K-S. On constate à partir de la Figure 4.20 que :

- Les paramètres α et β sur la VR augmentent en fonction du débit Q_6 dans les cas où le débit Q_6 est inférieur à 1400 véh/h. Au même niveau de débit Q_6 , les valeurs estimées de α et β sur la VR sont en général plus petites que celles sur la VL. La tendance croissante sur la VR est aussi plus lente que sur la VL.
- Le paramètre θ ne prend pas de petite valeur au débit Q_6 en trafic intermédiaire. De plus, ce paramètre est supérieur à 0,5 en trafic dense de débit Q_6 supérieur à 1500 véh/h.
- Le paramètre λ augmente en fonction du débit Q_6 . Au même niveau de débit Q_6 , le paramètre λ sur la VR est plus petit que ce même paramètre sur la VR. La différence peut atteindre 0,35 véh/s.

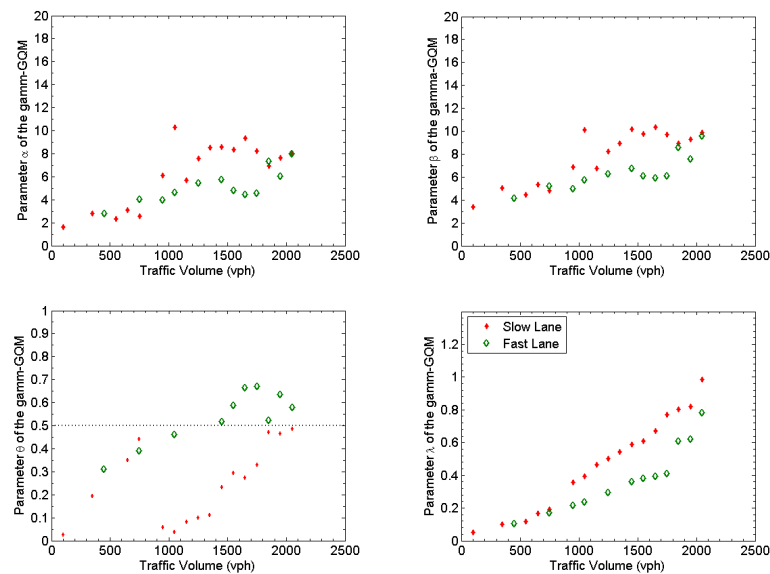


FIGURE 4.20 : Les paramètres du modèle Γ -GQM selon le type de voie, et la modalité jour-sec

4.3.6 Le modèle DG selon les voies par temps sec de jour

La Figure 4.22 montre la puissance de modélisation des distributions des TIV du modèle DG. Le modèle donne en particulier de bons ajustements pour les données de TIV sur la VL quelque soient les conditions de météo.

La Figure 4.21 montre clairement l'influence de la VR sur les paramètres d'échelle α_1 et α_2 du modèle DG. Le paramètre β_1 sur la VL est équivalent à ce même paramètre sur la VR tandis que le paramètre β_2 sur la VR est légèrement inférieur au paramètre β_2 sur la VL.

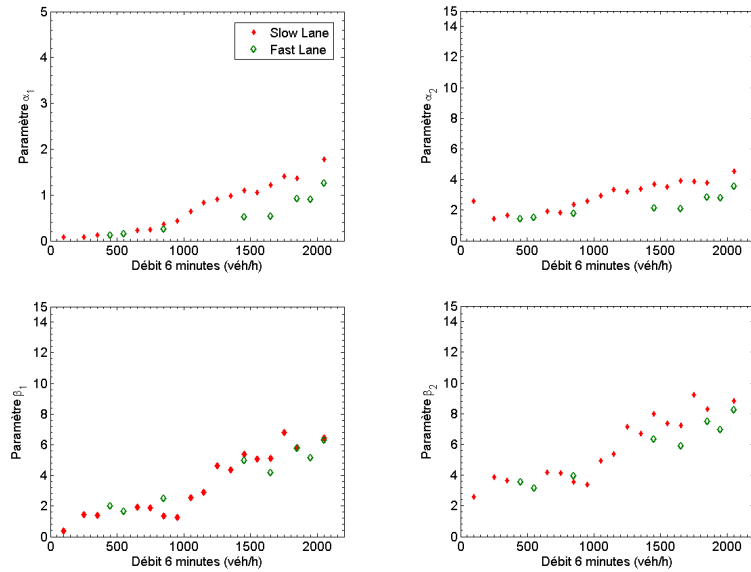


FIGURE 4.21 : Les paramètres $\alpha_1, \alpha_2, \beta_1, \beta_2$ du modèle DG selon le type de voie, et la modalité jour-sec

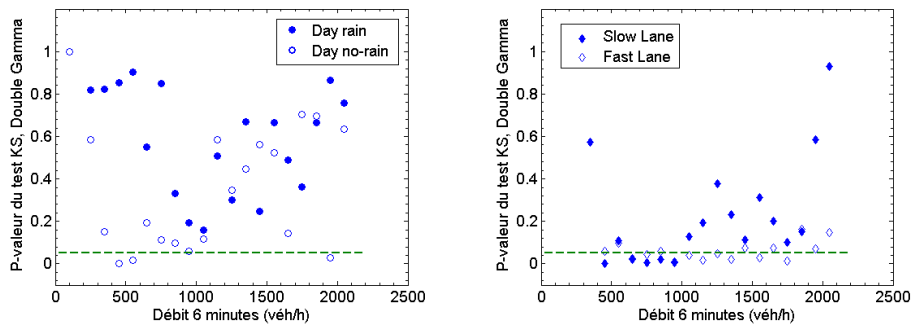


FIGURE 4.22 : P-valeur du test K-S du modèle DG

4.3.7 Discussion sur les paramètres du modèle Γ -GQM

Les changements de phase des paramètres α, β et θ du modèle Γ -GQM semblent plus précoces (c. à. d à partir d'une valeur de Q_6 plus faible) par temps de pluie que par temps sec. Par ailleurs, le TIV minimal obtenu sur la VR est supérieur à celui obtenu sur la VL, de même que le TIV minimal par temps de pluie est supérieur à celui par temps sec. En excluant les erreurs techniques, ces *décalages* observés sur les paramètres du modèle Γ -GQM et sur le TIV minimal semblent se produire probablement à cause de l'anticipation des conducteurs lors de déplacements plus rapides sur la VR et parce que les conducteurs font plus attention dans la condition de pluie.

Si conduire sur la VL par temps sec, le jour (modalité VL-jour-sec) est considéré comme un *comportement standard*, au même niveau de débit Q_6 , l'impact de la VR - un des facteurs exogènes - est plus net que celui du temps de pluie - un autre facteur exogène.

En comparant ces impacts sur les variations des paramètres du modèle Γ -GQM, il est proposé de distinguer deux types d'effet exogène. Le premier illustré par la condition de pluie est interprété comme une *gêne*. On l'appelle donc *facteur de gêne*. Les valeurs des paramètres α, β et λ du modèle Γ -GQM sont alors augmentées par rapport à celles du *comportement standard*. Au contraire, le second facteur illustré par la condition de circulation sur la VR, facilite le flux du trafic, et par conséquent, les valeurs des paramètres

Q_6 véh/h	Γ -GQM			\mathcal{T}_6 %	Γ -GQM SL sec
	VL sec	VL pluie	VR sec		
0 – 200	0.8588	–	–	0 – 2	0.0127
200 – 300	0.0280	0.9240	–	2 – 4	0.1779
300 – 400	0.5263	0.5937	0.0001	4 – 6	0.1153
400 – 500	0.0272	0.6994	0.3122	6 – 8	0.2172
500 – 600	0.0952	0.9855	0.0485	8 – 10	0.9769
600 – 700	0.3404	0.7512	0.0045	10 – 12	0.9701
700 – 800	0.3910	0.7848	0.1789	12 – 14	0.9163
800 – 900	0.0431	0.8110	0.0018	14 – 16	0.6709
900 – 1000	0.3583	0.7162	0.0693	16 – 18	0.0158
1000 – 1100	0.3456	0.6438	0.1129	18 – 20	0.2373
1100 – 1200	0.2288	0.5862	0.0288	20 – 22	0.0150
1200 – 1300	0.5761	0.7099	0.0720	22 – 24	0.0000
1300 – 1400	0.7573	0.6338	0.0244	> 24	0.0000
1400 – 1500	0.9805	0.7961	0.4717		
1500 – 1600	0.9701	0.8746	0.1118		
1600 – 1700	0.8648	0.7476	0.1877		
1700 – 1800	0.9708	0.8197	0.1201		
1800 – 1900	0.9892	0.7354	0.5145		
1900 – 2000	0.3547	0.9896	0.2950		
> 2000	0.8535	0.9908	0.6123		

TABLE 4.1 : P-valeurs du test K-S du modèle Γ -GQM du jour

α , β et λ du modèle Γ -GQM sont réduites par rapport à celles du *comportement standard*. On l'appelle le *facteur de propulsion*. La distinction entre le *facteur de gêne* et le *facteur de propulsion* est particulièrement nette dans la partie gauche du *diagramme fondamental* et se manifeste plus significativement dans le cas où l'impact est fort comme celui de la VR.

En distinguant les facteurs *exogènes*, les résultats de calibrage confirment l'existence de plusieurs distributions de TIV au même niveau de débit Q_6 . Cette co-existence est donc le résultat non simplement de variations numériques lors des calculs, mais plus profondément d'un changement de comportement des conducteurs adapté aux différents contextes réels. Notons sans changer le sens de cette analyse que la période d'agrégation de calcul joue aussi un rôle non négligeable.

Dans les sections précédentes, le débit Q_T est considéré et confirmé comme le facteur décisif sur l'échelle des valeurs de TIV. Les résultats dans cette section montrent qu'il existe aussi un deuxième niveau d'effet sur l'échelle des valeurs de TIV. En effet, au même niveau de débit Q_T , un facteur exogène va augmenter ou réduire les paramètres d'échelle α et λ du modèle Γ -GQM dans le comportement standard selon que le facteur est de type « *de gêne* » ou « *de propulsion* » respectivement. À cause de la contrainte posée *a priori* par la valeur du débit Q_T , il semble que la variation complexe du paramètre θ soit le résultat de la superposition des effets du débit Q_T et du facteur exogène.

La mise en évidence de ces deux facteurs implique que les paramètres d'échelle α et λ du modèle Γ -GQM ne dépendent pas seulement de la moyenne des TIV, mais aussi d'autres caractéristiques exogènes au trafic. Il semble que le faible effet de la *pluie* ne suffise pas pour réunir en une seule les deux phases de variation de θ . Il est à noter que tous les niveaux de *pluie* sont accumulés en une unique information de *pluie* ce qui pourrait être la cause de la variabilité élevée des paramètres estimés par temps de *pluie*. Notons aussi que les périodes de *pluie* ne représentant qu'une faible partie des périodes ; l'instabilité des paramètres pourrait être due à l'insuffisance des données. Une autre interprétation

possible de la grande différence des valeurs de θ du modèle Γ -GQM selon les voies (VL et VR) s'appuie sur la présence des PL sur la VL. En effet, les PL toujours sur la VL sont source de gêne du trafic. Pourtant, l'échelle de ce facteur est plus locale que les facteurs exogènes \mathcal{F}_1 abordés ici. Ce facteur du type mésoscopique \mathcal{F}_3 sera donc étudié dans le chapitre suivant.

Finalement, tant en débit fluide qu'en congestion, dans la condition de temps **sec**, les différences des valeurs des paramètres du modèle Γ -GQM selon les deux voies VL et VR sont faibles. Il semble que le type de voie influence le plus considérablement au niveau de trafic intermédiaire parce qu'en congestion, la restriction du trafic est dominante tandis qu'en trafic fluide, d'autres objectifs pourront arriver comme la rapidité, etc. Il semble que le TIV en trafic fluide soit principalement dû aux arrivées aléatoires, dont l'intensité ne dépend que du débit de classe identique sur les deux voies dans les comparaisons. Dans les deux cas, la perception des effets exogènes semble mineur et pas assez fort pour changer le comportement des conducteurs.

4.4 Modélisation du TIV sur différentes bases de données

4.4.1 Les échantillons raffinés de la RN118

L'ajustement des modèles Γ -GQM et DG

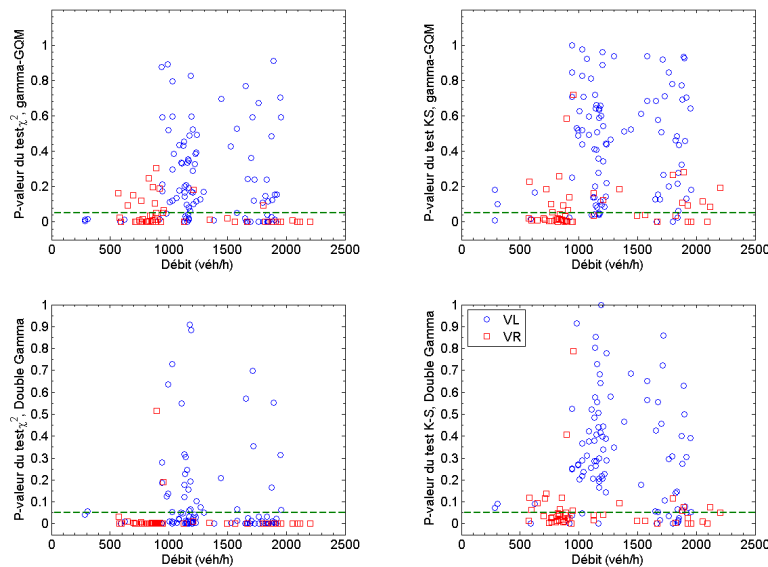


FIGURE 4.23 : P-valeur des modèles Γ -GQM et DG sur les échantillons raffinés de la RN118

La Figure 4.23 présente l'ajustement des modèles Γ -GQM et DG au sens des tests du χ^2 et K-S. On constate que :

- Les échantillons raffinés obtenus sur la VR ont soit un débit faible (moins de 1000 véh/h) soit un débit assez élevé alors que le débit des échantillons raffinés obtenus sur la VL se trouve principalement en intermédiaire. Cette observation montre que les états du trafic et de l'utilisation de la voirie des deux voies sont largement différents.
- Le test du χ^2 est plus « rigoureux » que le test K-S en rejetant les ajustements des modèles Γ -GQM et DG dans beaucoup d'échantillons raffinés de TIV. Il est à noter que ce test est sensible au groupement de données dans des classes de petite taille.

- Les modèles Γ -GQM et DG donnent de meilleurs résultats sur la VL que sur la VR. En particulier, le modèle DG fournit des P-valeurs supérieures à 0,05 pour la majorité des échantillons raffinés de TIV sur la VL au sens du test K-S.

Le modèle Γ -GQM

La Figure 4.23 montre les variations des paramètres du modèle Γ -GQM en fonction du débit Q_r pour les échantillons que le test K-S ne rejette pas au niveau de signification de 5 %.

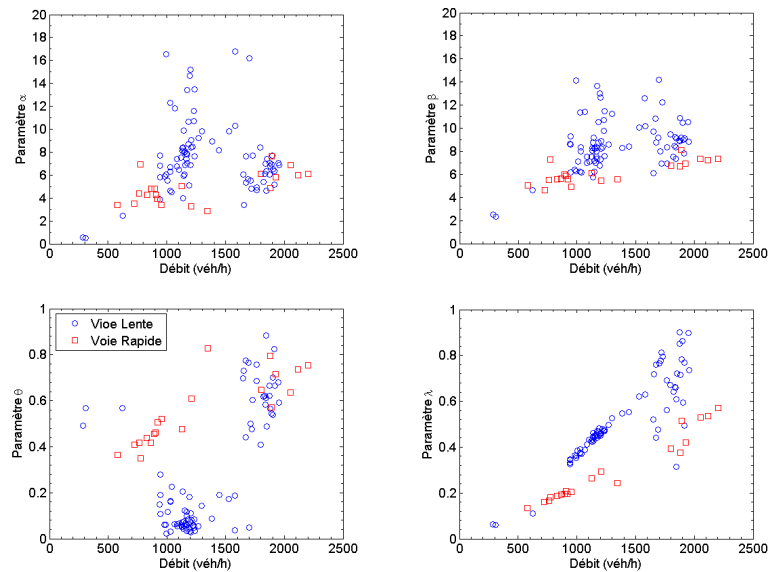


FIGURE 4.24 : Les paramètres du modèle Γ -GQM dans les échantillons raffinés

En comparant avec les variations étudiées dans les sections précédentes, on constate que :

- Les paramètres α , β et λ du modèle Γ -GQM sur la VR augmentent généralement en fonction du débit Q_r , sont plus faibles que ceux sur la VL. Aussi, les tendances croissantes de ces paramètres sur la VR sont plus faibles que celles sur la VL.
- Les paramètres α et β du modèle Γ -GQM sur la VL évoluent aussi en deux phases. La deuxième phase correspond aux débits Q_r supérieurs à 1500 véh/h.
- Le paramètre θ du modèle Γ -GQM sur la VL prend aussi des valeurs faibles en trafic intermédiaire de débit Q_r d'environ 1000 à 1200 véh/h.
- Pour certains niveaux de débit Q_r sur la VL, il y a plusieurs valeurs des paramètres α et β ce qui confirme la variabilité de ces paramètres.
- En trafic congestionné, plusieurs valeurs des paramètres λ et θ obtenus montrent que les distributions des TIV dans cette condition ne sont pas régulières.

Le Modèle DG

On constate en comparant avec les variations des paramètres du modèle DG étudiés précédemment que :

- Tous les paramètres α_1 , α_2 , β_1 et β_2 augmentent en fonction du débit Q_r . Au même niveau de débit Q_r , les paramètres d'échelle α_1 , α_2 sur la VR sont plus faibles que ceux sur la VL. De plus, les paramètres α_1 , α_2 sur la VR augmentent plus lentement en fonction du débit que sur la VL.

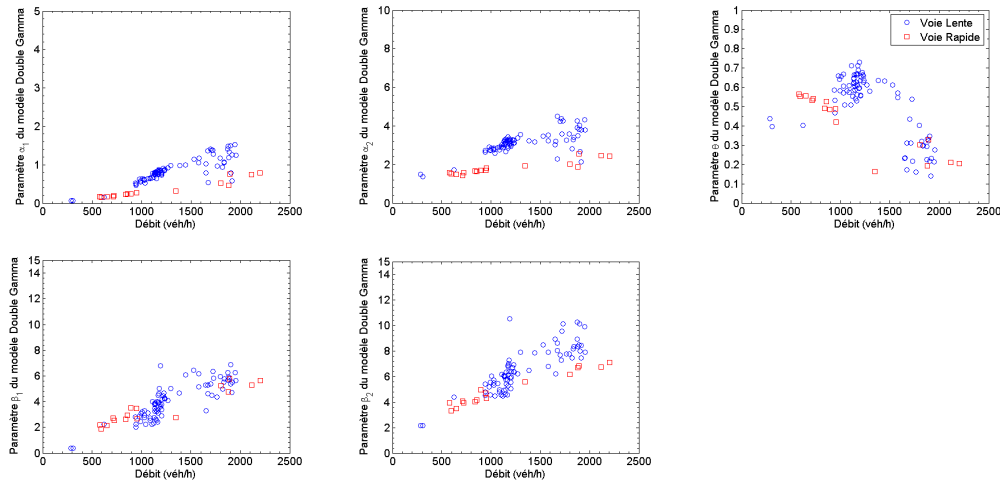


FIGURE 4.25 : Les paramètres du modèle DG, dans les échantillons raffinés

- La différence de tendance entre les voies du paramètre β_2 est moins nette, la différence de tendance entre les voies du paramètre β_1 est minime par rapport à celles des paramètres α_1 et α_2 .
- Le paramètre θ sur la VR diminue en fonction du débit Q_r . Pour les deux voies, ce paramètre atteint son maximum (de valeur environ 0,7) pour le débit Q_r de l'ordre de 1200 véh/h ; pour les débits faibles (de l'ordre de 300 véh/h), θ prend des valeurs d'environ 0,4 ; et pour les débits élevés (2000 véh/h), il est de l'ordre de 0,2.

Discussion sur les résultats des échantillons raffinés

En somme, on retrouve que les paramètres des modèles Γ -GQM et DG varient en fonction du débit de la même façon que dans les sections précédentes. On retrouve aussi quelque effets du facteur de la VR - un facteur exogène \mathcal{F}_1 comme :

1. La réduction des paramètres d'échelle α, λ et donc β du modèle Γ -GQM et les paramètres d'échelle α_1, α_2 du modèle DG. Les tendances croissantes des paramètres sont aussi plus lentes sur la VR.
2. Des valeurs faibles du paramètre θ du modèle Γ -GQM en trafic intermédiaire sur la VL tandis que θ augmente en fonction du débit sur la VR.

Les résultats de cette section montrent de plus que le régime « régularité » défini *a priori* précédemment - un des trois régimes dans l'état de l'art (cf. Section 4.1.2), n'est pas régulier. Comme les échantillons raffinés comportent des TIV consécutifs, les résultats irréguliers pourraient être la conséquence de phénomènes *d'accordéon* (*ang. stop and go*) ou d'autres phénomènes lorsque le trafic est très dense.

Finalement, on a remarqué que l'échantillonnage de raffinement entraînait plus d'instabilité des paramètres obtenus pour un niveau de débit donné. Autrement dit, un débit Q_r donné correspond à plusieurs comportements des conducteurs ce qui signifie que la méthode d'échantillonnage de raffinement ne donne guère des échantillons raffinés homogènes.

4.4.2 Soixante-quinze échantillons horaires de l'autoroute A6

L'ajustement des modèles Γ -GQM et DG

La Figure 4.26 présente les performances des modèles Γ -GQM et DG au sens des tests K-S et du χ^2 . Les deux modèles fournissent de bons ajustements pour la plupart des échantillons horaires au sens du test K-S. Le test du χ^2 est encore plus sélectif notamment pour le modèle DG en rejetant de plus nombreux échantillons.

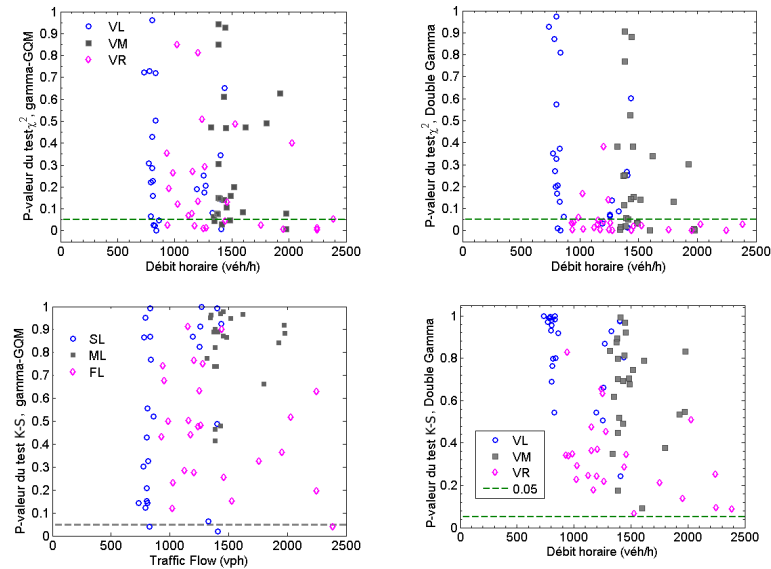


FIGURE 4.26 : P-valeur sur les données de l'A6

Le modèle Γ -GQM

La Figure 4.27 présente les variations des paramètres du modèle Γ -GQM en fonction du débit horaire Q_h . Les résultats sont analogues aux résultats correspondants des sections précédentes, et on constate en plus que :

- En général, les paramètres α , β et λ augmentent en fonction du débit horaire Q_h . L'augmentation de λ en fonction de Q_h est nette pour l'ensemble des voies ; les tendances de α et β sont plus nettes sur la VR que sur les VL et VM. Au même niveau du débit horaire Q_h , les valeurs estimées des paramètres α , β et λ sur la VR sont aussi plus faibles que sur la VL et VM.
- Le paramètre θ sur la VR augmente en une phase tandis que ce paramètre sur les VL et VM semblent varier en deux phases car θ prend aussi des valeurs faibles au débit intermédiaire Q_h d'environ 1200 - 1500 véh/h.
- Au débit horaire Q_h de 800 véh/h sur la VL, les paramètres α , β et θ varient largement (α et β de 2 à 6, θ de 0,05 à 0,25). En trafic plus dense de débit Q_h supérieur à 1400 véh/h, les valeurs correspondantes sont plus élevées mais plus diversifiées sur la VL et VM. Pourtant, les paramètres α , β et θ sont proches entre la VL et la VM ; ils sont différents de ceux sur la VR.
- Le paramètre λ est celui qui est le plus sensible au débit horaire Q_h dans les états du trafic fluide et intermédiaire car il varie aussi en fonction de petites variations de débit Q_h . Pour les débits plus élevés, l'évolution du paramètre λ est un peu perturbée dans les échantillons des VL et VM.

La Figure 4.28 montre la relation entre les paramètres α et β du modèle Γ -GQM. Cette corrélation est prévue par les variations similaires de α et β en fonction du débit Q_h dans

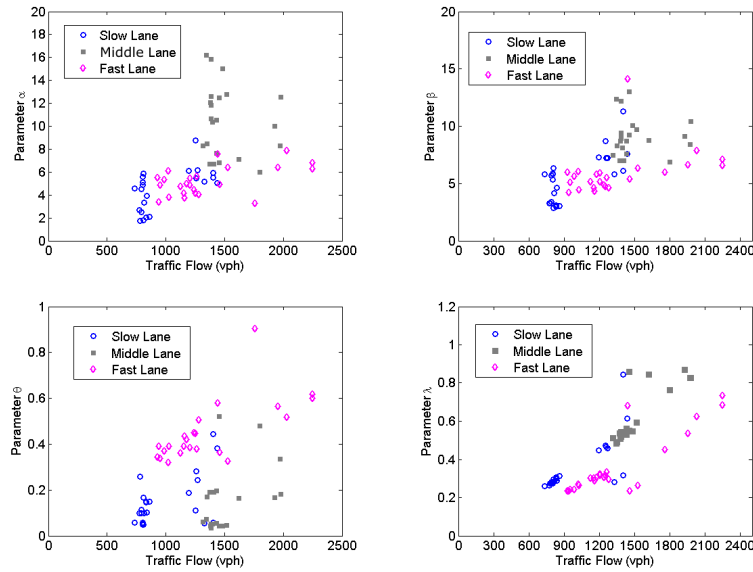


FIGURE 4.27 : Variations des paramètres du modèle Γ -GQM en fonction du débit horaire Q_h

la Figure 4.27.

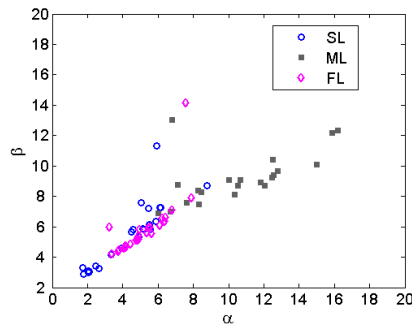


FIGURE 4.28 : La relation entre les paramètres α et β du modèle Γ -GQM

Le modèle Double Gamma

La Figure 4.29 présente les variations des paramètres du modèle DG en fonction du débit horaire Q_h . Ces variations sont analogues à celles mises en évidence dans les sections précédentes :

- En général, les paramètres $\alpha_1, \alpha_2, \beta_1$ et β_2 augmentent en fonction du débit Q_h . De plus, les paramètres d'échelle α_1 et α_2 sur la VL et VM sont nettement plus grands que ceux sur la VR. Ce même résultat est obtenu pour le paramètre β_2 mais la différence est moins significative tandis que le paramètre β_1 varie et prend des valeurs analogues sur les trois voies.
- Le paramètre θ diminue généralement en fonction du débit Q_h mais la variation des valeurs de θ est large.
- Au même niveau du débit Q_h sur la VL, les variations des paramètres β_1, β_2 et θ sont grandes, particulièrement autour du débit Q_h de 800 véh/h.

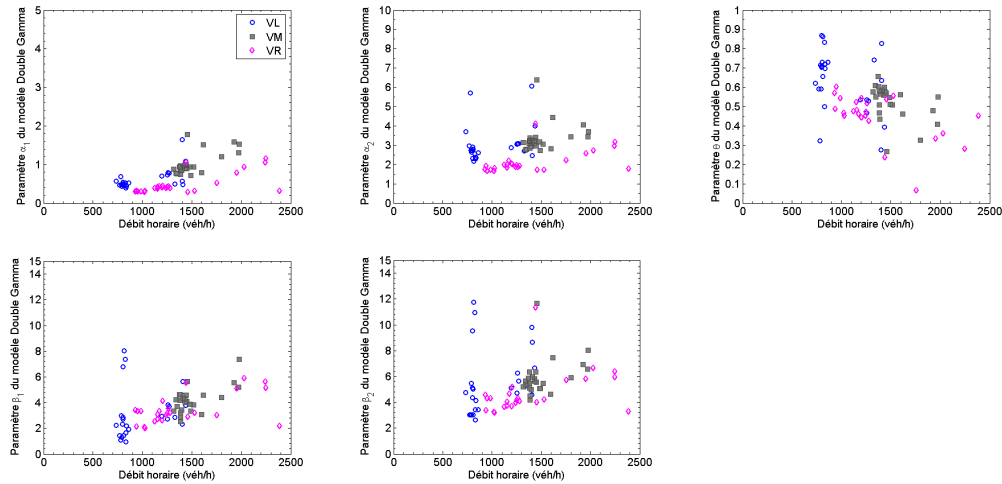


FIGURE 4.29 : Variations des paramètres du modèle DG en fonction du débit horaire Q_h

Étude des périodes de même tranche d'heure

Dans une même période horaire considérée, quelque soit cette période (période de pointe, période creuse), la vitesse est en moyenne plus élevée sur la VR, et plus basse sur la VL (cf. La Figure A.1 dans l'Annexe).

Dans la même période, le débit et la distribution des TIV sur les trois voies sont largement différents. Le Tableau 4.2 représente deux périodes étudiées, l'une à l'heure de pointe (8h-9h) et l'autre à l'heure creuse (10h-11h). Les résultats montrent la différence significative de distribution des TIV entre les trois voies. Le paramètre λ est le plus petit sur la VL et le plus grand sur la VM sur chacune des deux périodes. À l'heure de pointe, l'ajustement du modèle Γ -GQM sur la VR est le moins bon au sens du test K-S, probablement parce que les TIV courts sur la VR se concentrent plus autour du *mode* que sur la VL et la VM. On observe également que les valeurs du *mode* diminuent respectivement de la VL, la VM et à la VR tandis que les densités de probabilité du *mode* augmentent respectivement.

Voie	Q	Période	α	β	θ	λ	P_{χ^2}	P_{K-S}
VL	1272	8h - 9h	6.1610	7.2461	0.2437	0.4580	0.7223	0.9984
VM	1926	8h - 9h	10.00	9.1031	0.1659	0.8697	0.4370	0.8422
VR	2251	8h - 9h	6.2671	6.5915	0.5998	0.7331	0.4788	0.1953
VL	831	10h - 11h	2.0322	2.9788	0.1467	0.2982	0.6847	0.9932
VM	1411	10h - 11h	6.7236	6.9742	0.1901	0.5363	0.0736	0.8891
VR	1205	10h - 11h	5.4496	5.9030	0.3839	0.3235	0.8805	0.7668

TABLE 4.2 : Les échantillons sur les trois voies pour deux périodes horaires du modèle Γ -GQM

Pour le modèle DG, sur une même période horaire, les valeurs des paramètres α_1 , α_2 , β_1 et β_2 sont les plus grandes sur la VM. Le paramètre α_1 est faible, voire inférieur à 1 sur la VL en période de pointe et particulièrement sur l'ensemble des voies en période creuse.

Les Tableaux 4.4 et 4.5 présentent les paramètres des modèles sur la VL et la VR en comparant des échantillons de débits Q_h similaires. On retrouve les remarques qui ont été faites auparavant pour la VL (période de pointe) et la VR (période creuse).

On suppose que les différences de distribution des TIV selon les voies pour un même niveau de débit Q_h sont dues à des comportements de conduite différents. Il est donc

Voie	Q	Période	α_1	α_2	β_1	β_2	θ	P_{χ^2}	P_{K-S}
VL	1272	8h - 9h	0.7733	3.0687	3.6766	5.6578	0.5303	0.1386	0.8694
VM	1926	8h - 9h	1.5894	4.0485	5.5732	6.9295	0.4801	0.3005	0.5357
VR	2251	8h - 9h	1.1529	3.1574	5.1617	5.9446	0.2810	0.0024	0.0943
VL	831	10h - 11h	0.3862	2.2864	0.9630	2.6274	0.5012	0.3726	0.9846
VM	1411	10h - 11h	0.9597	3.2546	4.3787	5.9547	0.5846	0.0519	0.9919
VR	1205	10h - 11h	0.4565	2.0278	4.1381	5.2040	0.5433	0.3812	0.3698

TABLE 4.3 : Les échantillons sur les trois voies pour deux périodes horaires du modèle DG

Voie	Q	Période	α	β	θ	λ	P_{χ^2}	P_{K-S}
VL	1255	8h - 9h	8.7727	8.6785	0.1099	0.4738	0.5208	0.8248
VR	1253	12h - 13h	4.1709	4.7568	0.4480	0.3183	0.0634	0.6321

TABLE 4.4 : Deux échantillons (L1 et R3c) ayant des débits horaires Q_h similaires sur les VL et VR, modèle Γ -GQM

intéressant de savoir si ces différences sont aussi trouvées au niveau de la distribution des vitesses. La Figure 4.30 en est une preuve claire.

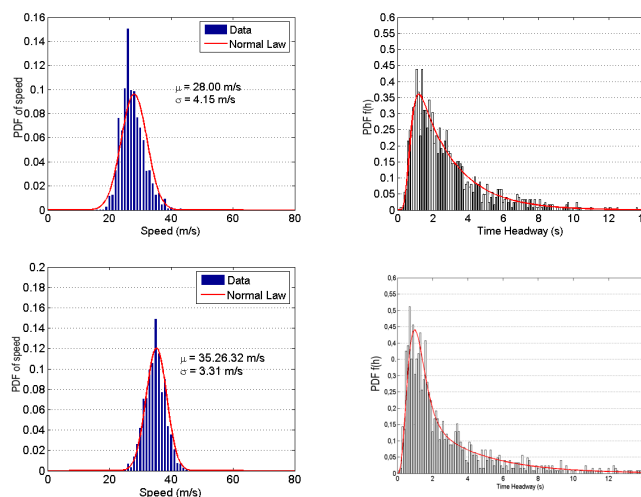


FIGURE 4.30 : Deux échantillons (L1 et R3c) ayant des débits horaires Q_h similaires sur la VL et la VR

Un autre exemple montre l'importance du rôle de la vitesse - cf. le Tableau 4.6 représentant deux échantillons aux débits Q_h similaires sur la VL. Ils sont aussi de la même section routière et ont seulement été enregistrés à deux périodes différentes. La Figure 4.31 montre que les paramètres α , β et θ du modèle Γ -GQM semblent être sensibles à la présence de TIV courts (disons inférieurs à 5 s) alors que le paramètre λ est très stable. Pour expliquer la différence entre les deux cas en négligeant l'existence de conditions exogènes dans les deux périodes étudiées, la Figure 4.31 montre que les distributions des vitesses des deux échantillons sont assez différentes.

L'effet de la réduction de la largeur de voie sur l'autoroute l'A6

Le Tableau 4.7 récapitule quatre couples d'échantillons horaires de débits Q_h similaires correspondant respectivement à deux sections de largeur de voie respectivement normale et réduite.

Voie	Q	Période	α_1	α_2	β_1	β_2	θ	P_{χ^2}	P_{K-S}
SL	1255	8h - 9h	0.7346	3.0550	2.7107	4.7110	0.4668	0.0738	0.5074
FL	1253	12h - 13h	0.4269	1.9618	3.1437	4.1737	0.4734	0.0027	0.6340

TABLE 4.5 : Deux échantillons (L1 et R3c) ayant des débits horaires Q_h similaires sur les VL et VR, modèle DG

Q	Period	α	β	θ	λ	P_{χ^2}	P_{K-S}	A^2
814	11h - 12h	3.3249	4.1521	0.1001	0.2839	0.1531	0.3253	0.7252
810	12h - 13h	1.7808	2.8602	0.1653	0.2943	0.2760	0.5575	0.5159

TABLE 4.6 : Les paramètres du modèle Γ -GQM de deux échantillons (L3b et L3c) de la VL au même niveau de Q_h

En comparant les paramètres du modèle Γ -GQM dans chaque couple d'échantillons, on observe que les paramètres λ sont plus bas dans la section réduite (non simplement à cause de la réduction de débit Q_h). Le paramètre change lors de la réduction de la largeur de voie, et change plus largement de la VR, la VM à la VL. À l'heure creuse et sur les trois voies, les paramètres α , β augmentent et θ diminue lorsque la voie est rétrécie. Dans ce cas c'est le paramètre θ sur la VL qui varie le plus. Cependant, les paramètres α , β et θ varient dans le sens inverse à l'heure de pointe dans le couple d'échantillons sur la VL bien que les formes des distributions de TIV soient très proches dans la Figure 4.32.

Pour le modèle DG, on observe que les paramètres α_1 , α_2 , β_1 , β_2 et θ sont plus grands sur les VL ou VM lorsque la largeur de voie est réduite. Pourtant, les paramètres α_1 , α_2 , β_1 et β_2 sur la VR réduite sont inférieurs à ceux de la VR normale, seul θ est supérieur. En particulier, deux échantillons L6c et L3a sont largement différents au niveau des paramètres du modèle DG.

En comparant les distributions des vitesses, on observe l'effet de la réduction de largeur de la voie sur la VL où la moyenne des vitesses est réduite et l'écart-type plus encore. Sous ce point de vue, la distribution des vitesses sur la VM est la moins affectée. (cf. Figure 4.32).

Discussion

En somme, on a retrouvé que les paramètres des modèles Γ -GQM et DG varient en fonction du débit de la même façon que ceux des sections précédentes. Les impacts des facteurs de la VR sont aussi observés sur les échantillons horaires. Par ailleurs, l'autoroute

Voie	Q	Période	α	β	θ	λ	P_{χ^2}	P_{K-S}
Largeur de voie réduite								
VL	1259	8h - 9h	5.4728	7.2098	0.2825	0.4659	0.3506	0.9117
VL	828	12h - 13h	6.1477	6.9136	0.0601	0.2937	0.0352	0.0387
VM	1384	10h - 11h	15.8654	12.1688	0.0332	0.5273	0.9978	0.9025
VR	1204	12h - 13h	6.1206	6.0662	0.3196	0.2693	0.8451	0.1183
Largeur de voie normale								
VL	1255	8h - 9h	8.7727	8.6785	0.1099	0.4738	0.5208	0.8248
VL	832	10h - 11h	2.0322	2.9788	0.1467	0.2982	0.6847	0.9932
VM	1384	11h - 12h	10.6731	9.0676	0.0507	0.5421	0.9897	0.8224
VR	1204	12h - 13h	4.8933	5.1749	0.3851	0.3182	0.3255	0.2765

TABLE 4.7 : Les paramètres du modèle Γ -GQM pour quatre couples d'échantillons sur deux sections de l'A6

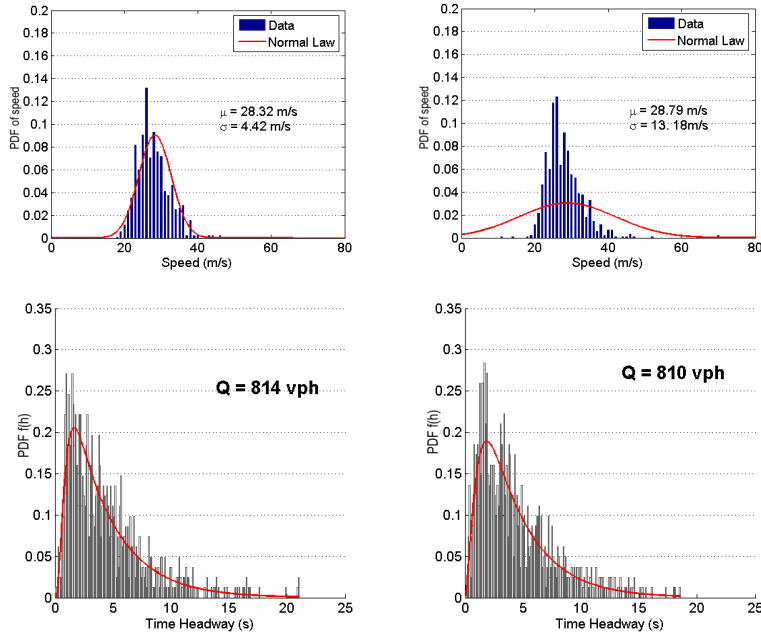
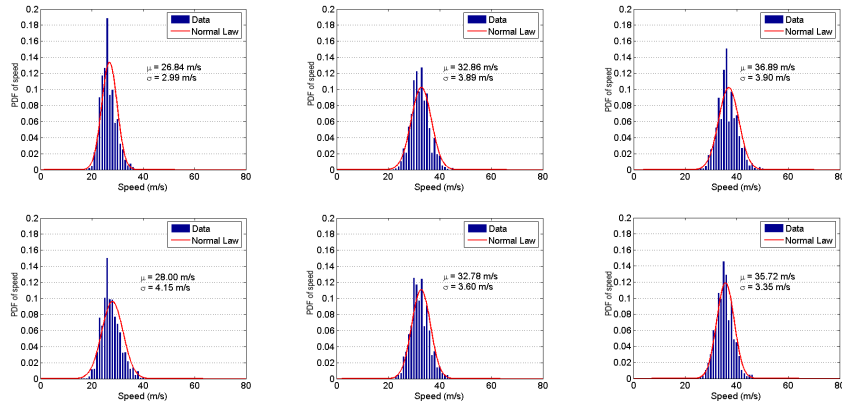


FIGURE 4.31 : Deux échantillons (L3b et L3c) de la VL au même niveau de Q_h

A6 comporte une voie de plus que la route RN118, c'est la VM. Les caractéristiques statistiques et probabilistes de cette voie sont plus proches de celles de la VL que de la VR. L'effet du facteur exogène de la VM est peu remarqué.

Les résultats montrent que la méthode d'échantillonnage par tranche horaire entraîne de grandes instabilités des paramètres obtenus pour un même niveau de débit horaire Q_h . Autrement dit, un débit Q_h donné correspond à plusieurs comportements des conducteurs ce qui signifie que la méthode d'échantillonnage utilisée ne donne guère des échantillons horaires représentatifs. Il semble qu'une heure est une assez longue période d'agrégation dans laquelle le débit n'est pas homogène. Par conséquent, plusieurs facteurs \mathcal{F}_1 pourraient émerger lors de l'enregistrement.

Si l'on considère que la réduction de largeur de voie est aussi un facteur exogène, plutôt du type de *gêne* sur le comportement standard sur la VL, les effets obtenus sur les paramètres



Distributions des vitesses (L4,L1), (M6a,M9b),(R6c,R9c)

FIGURE 4.32 : L'effet de la réduction de la largeur de voie sur la distribution des vitesses

Voie	Q	Période	α_1	α_2	β_1	β_2	θ	P_{χ^2}	P_{K-S}
Largeur de voie réduite									
VL	1259	8h - 9h	0.7878	3.0957	3.8445	6.2492	0.5347	0.0675	0.6610
VL	828	12h - 13h	0.5227	2.5583	7.3610	10.930	0.8329	0.1320	0.5441
VM	1384	10h - 11h	0.9697	3.3835	4.6060	6.3476	0.5791	0.9049	0.7981
VR	1204	12h - 13h	0.2908	1.6738	2.1173	3.2714	0.4685	0.1681	0.2275
Largeur de voie normale									
VL	1255	8h - 9h	0.7346	3.0550	2.7107	4.7110	0.4668	0.0738	0.5074
VL	832	10h - 11h	0.3862	2.2864	0.9630	2.6274	0.5012	0.3726	0.9846
VM	1384	11h - 12h	0.9044	3.2472	3.2278	5.1008	0.5075	0.7705	0.7030
VR	1204	12h - 13h	0.4123	2.0503	2.6521	3.7092	0.4452	0.0422	0.2446

TABLE 4.8 : Les paramètres du modèle DG pour quatre couples d'échantillons sur deux sections de l'A6

d'échelle α , β et λ du modèle Γ -GQM ne sont pas consistants avec les résultats des sections précédentes. Prenons par exemple, ces paramètres tantôt croissent tantôt diminuent sur la VL lors de la réduction de la largeur de voie, et de plus, les évolutions des α et β se font dans le sens inverse de celle du paramètre λ . Il semble être confirmé que pendant la longue période d'agrégation, d'autres facteurs exogènes plus forts soient intervenus et perturbent les distributions des TIV.

4.5 Les conditions plus complexes et les cas particuliers

L'ensemble des résultats de la modélisation dans les sections de ce chapitre entraîne les deux remarques principales suivantes :

- Il existe des relations entre les paramètres des modèles Γ -GQM et DG et des variables macroscopiques, particulièrement le débit.
- Les effets des conditions exogènes peuvent être détectés et/ou mesurés lorsque le comportement standard de référence a été identifié.

Les comparaisons entre les deux modèles montrent aussi que le modèle DG est plus facile à utiliser tant au niveau du calcul qu'au niveau de la sensibilité des paramètres. De plus, malgré un paramètre supplémentaire par rapport au modèle Γ -GQM, les relations entre les cinq paramètres du modèle DG et le débit Q sont plus nettes et quasiment linéaires (sauf θ). Pourtant, le modèle Γ -GQM a été longuement utilisé dans plusieurs recherches et ce d'autant plus que le paramètre λ a une interprétation physique pour le débit. Ce sont les raisons pour lesquelles, la première partie de cette section est consacrée à étudier des cas particuliers au point de vue de la modélisation employant le modèle Γ -GQM.

4.5.1 Les échantillons ayant des paramètres d'échelle similaires

Le Tableau 4.9 et la Figure 4.33 représentent deux couples d'échantillons horaires de l'A6 dans chacun desquels, les paramètres α et β du modèle Γ -GQM sont équivalents. De plus les échantillons à comparer ont des débits horaires Q_h largement différents.

Bien que le modèle Γ -GQM fournisse toujours de très bons ajustements dans tous les échantillons de TIV, les paramètres α , β et aussi θ sur la VL ne semblent pas avoir de lien significatif avec le débit horaire Q_h . En gardant constants quelques paramètres, le modèle Γ -GQM est tellement flexible qu'il peut modéliser d'autres distributions de TIV aux différents états du trafic. Néanmoins, cette flexibilité est un point faible du modèle

lorsque l'on l'utilise en sens inverse, c. à. d en sachant les états du trafic, le choix de bons paramètres pour simuler les données de TIV n'est pas évident.

Q	Period	α	β	θ	λ	P_{χ}^2	P_{K-S}	A^2
828	14h - 15h	6.1477	6.9136	0.0601	0.2937	0.0352	0.0387	1.2829
1272	11h - 12h	6.1610	7.2461	0.2437	0.4580	0.7223	0.9984	0.1642
1259	10h - 11h	5.4728	7.2098	0.2825	0.4659	0.3506	0.9117	0.4753
1436	14h - 15h	5.0604	7.6027	0.3832	0.6153	0.8875	0.9238	0.1338

TABLE 4.9 : Deux couples d'échantillons ((L6c,L2) et (L4, L11)) avec des valeurs de α, β équivalentes

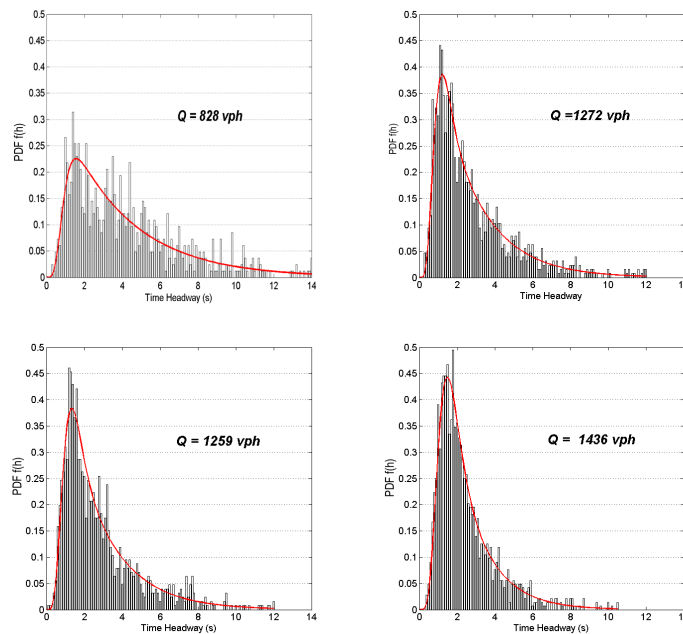


FIGURE 4.33 : Deux couples d'échantillons ((L6c,L2) et (L4, L11)) avec des valeurs de α, β équivalentes

4.5.2 La discontinuité du paramètre θ du modèle Γ -GQM

Cinq échantillons à cinq niveaux de débit Q_6 , ont été sélectionnés pour l'étude de la discontinuité du paramètre θ du modèle Γ -GQM. Ces niveaux de débit sont (700-800), (800-900), (900-1000), (1000-1100) et (1100-1200) véh/h (cf. Figure 4.1). L'échantillon (700-800) véh/h représente le cas avant la discontinuité avec $\theta = 0,4412$ alors que l'échantillon (900-1000) véh/h est l'échantillon après la discontinuité avec $\theta = 0,0382$. L'échantillon (1000-1100) véh/h associe à une très faible valeur de θ des valeurs élevées de α et β . Le dernier échantillon se trouve juste après l'échantillon particulier (1000-1100) véh/h. L'échantillon (800-900) véh/h est aussi étudié malgré un ajustement du test K-S légèrement moins bon, avec une P-valeur de 0,0431 mais avec un très bon ajustement pour le test du χ^2 avec une P-valeur de 0,1621.

Le Table 4.10 montre les valeurs estimées des paramètres et les ajustements du modèle Γ -GQM. La Figure 4.34 illustre aussi comment le modèle Γ -GQM s'adapte aux données réelles de TIV de la RN118. Les résultats numériques montrent que la moyenne μ_H , le CV

des TIV sont inversement proportionnels au débit Q_6 et que le paramètre λ est proportionnel au débit 6 minutes Q_6 .

Q_6	μ_H	CV	α	β	θ	λ	P_χ^2	P_{K-S}
700-800	4.808	1.1217	2.5441	4.7795	0.4412	0.1908	0.1849	0.3909
800-900	4.2024	0.8937	6.0259	7.5271	0.1116	0.3008	0.1621	0.0431
900-1000	3.7581	0.7854	6.1040	6.8713	0.0586	0.3576	0.0387	0.3583
1000-1100	3.4232	0.7513	10.2902	10.1242	0.0382	0.3943	0.1503	0.3456
1100-1200	3.1546	0.6737	5.6924	6.7503	0.0819	0.4663	0.1182	0.2288

TABLE 4.10 : Résultats numériques des cinq échantillons classés selon le débit Q_6

En observant les histogrammes de la Figure 4.34, on observe que la différence significative entre les échantillons (700-800) et (900-1000) véh/h réside dans les fréquences des TIV de l'intervalle [3, 6] s. Cette remarque est mise en évidence dans la Figure 4.36. On obtient également que les distributions des TIV de moins de 3 s des deux échantillons sont quasiment identiques.

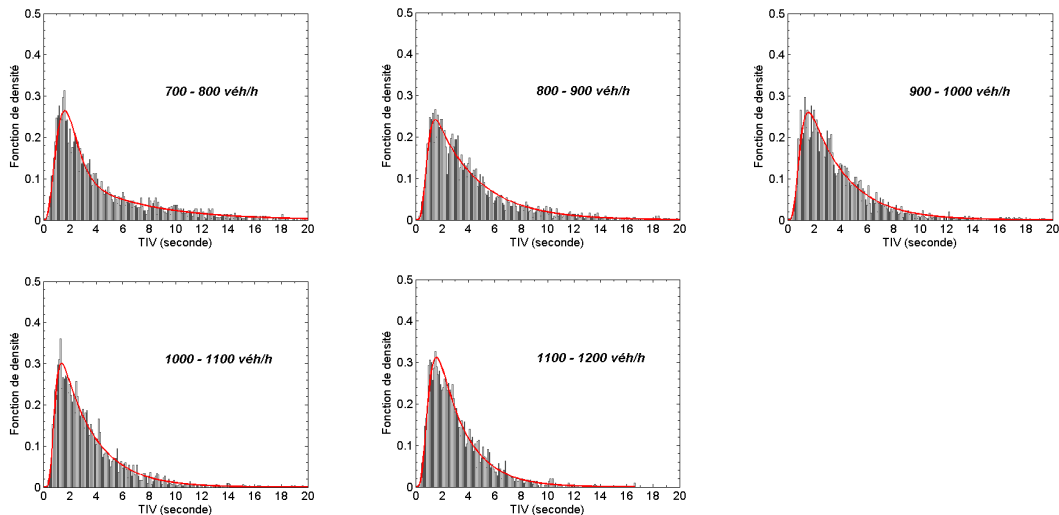


FIGURE 4.34 : Les ajustements du modèle Γ -GQM aux données empiriques pour cinq échantillons de débit Q_6 intermédiaire

Pour les échantillons à partir du débit (800-900) véh/h, les courbes de la fonction de densité et de la fonction de répartition varient régulièrement au niveau du graphique (cf. la Figure 4.35). En conséquence, on suppose que l'association des valeurs élevées des paramètres α et β à une valeur faible du paramètre θ n'est qu'un résultat purement numérique.

4.5.3 Les échantillons de débit très faible

Dans cette section, les trois échantillons de débits Q_6 les plus faibles sont étudiés. Les classes de débit sont (0-200), (200-300) et (300-400) véh/h. Bien que le deuxième échantillon soit rejeté par le test du χ^2 et le test K-S au niveau de signification de 5 %, sa P-valeur n'est pas faible, de 0,028. La Table 4.11 récapitule les résultats numériques obtenus pour ces trois échantillons en utilisant le modèle Γ -GQM.

La Figure 4.37 montre des différences significatives entre les distributions des TIV de ces échantillons de débits Q_6 très faibles. Il y a si peu de données dans l'échantillon (0-200) véh/h que son histogramme ressemble à celui d'une loi uniforme. L'écart entre la courbe

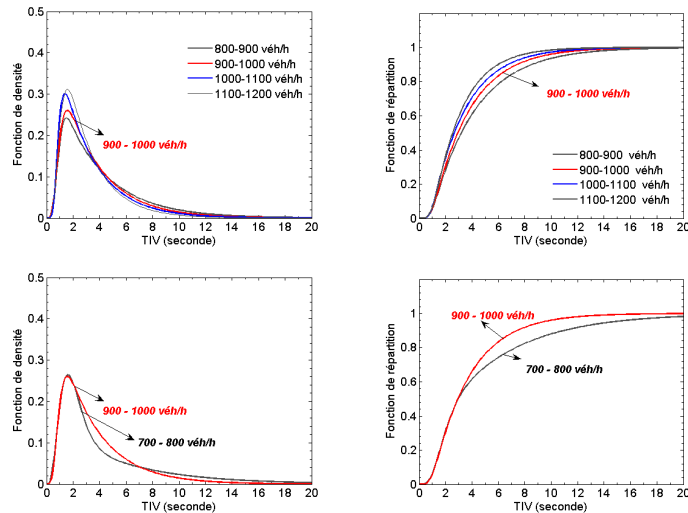


FIGURE 4.35 : Les distributions des TIV selon le modèle Γ -GQM ajusté pour cinq échantillons de débit Q_6 intermédiaire

Q_6	μ_H	CV	α	β	θ	λ	P^2_x	P_{K-S}
0-200	21.2119	0.8845	1.6210	3.4028	0.0239	0.0511	–	0.8588
200-300	12.6768	1.2602	5.2842	9.1940	0.2051	0.0727	–	0.0280
300-400	9.9280	1.0884	2.8094	5.0697	0.1940	0.0992	0.3946	0.5265

TABLE 4.11 : Paramètres du modèle Γ -GQM pour les trois échantillons de débit Q_6 très faible

de la fonction de répartition de l'échantillon (0-200) veh/h et les autres courbes est très grand (cf. Figure 4.36).

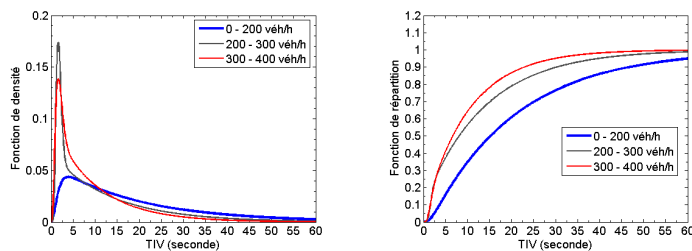


FIGURE 4.36 : Les distributions des TIV selon le modèle Γ -GQM ajusté pour trois échantillons de débit Q_6 très faible

Bien que ces courbes des fréquences cumulées soient ordonnées en fonction du débit Q_6 , le CV particulièrement élevé de l'échantillon (200-300) veh/h fait la différence au niveau des paramètres estimés parmi les trois échantillons. La Figure 4.37 montre, dans l'échantillon (200-300) veh/h, une concentration des TIV autour d'un *mode* de faible valeur, plus importante que celles des autres échantillons. De plus, la tendance décroissante des TIV longs dans l'échantillon (200-300) veh/h est plus faible que dans l'échantillon (300-400) veh/h. Ce sont probablement les raisons pour lesquelles les paramètres α et β du modèle Γ -GQM sont particulièrement élevés pour l'échantillon (200-300) veh/h tandis que λ de cet échantillon est encore proportionnel au débit Q_6 . Enfin, il est à noter que le paramètre θ est similaire dans les échantillons (200-300) et (300-400) veh/h.

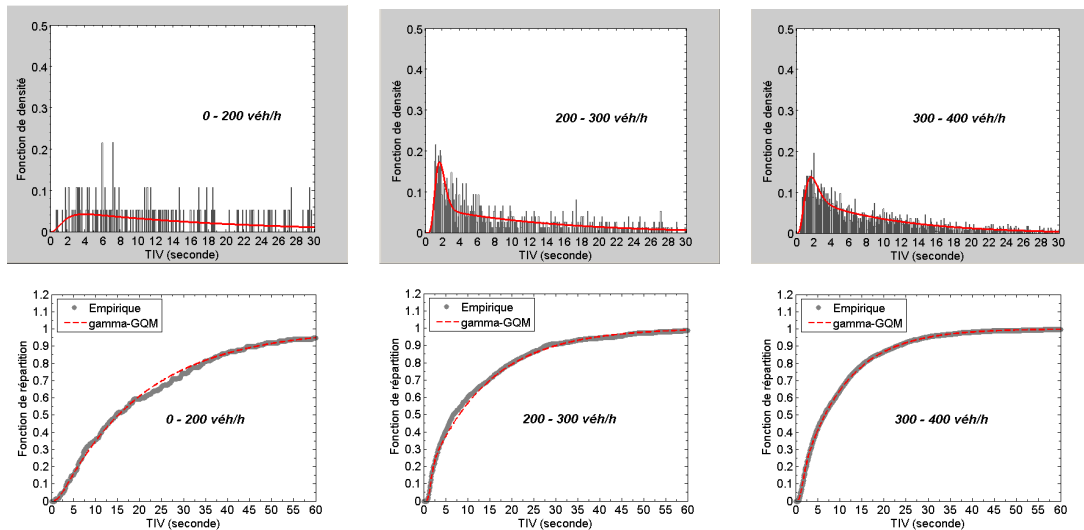


FIGURE 4.37 : L'ajustement du modèle Γ -GQM pour trois échantillons de débit Q_6 très faible

4.5.4 Les échantillons de très haut débit

Cinq échantillons sont étudiés dans cette section correspondant à des débits supérieurs au seuil à partir duquel les paramètres α et β perdent leur tendance croissante. Le modèle Γ -GQM fournit de très bons ajustements tant au sens du test du χ^2 que du test K-S. L'ajustement du test χ^2 est significativement amélioré car les valeurs observées des TIV sont inférieures à 10 s ce qui évite de constituer des classes parfois vides pour les TIV supérieurs.

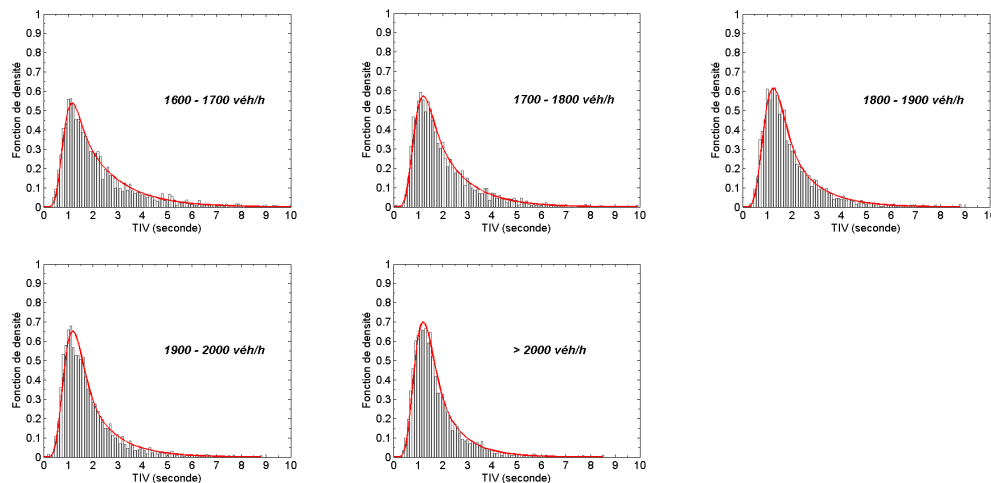


FIGURE 4.38 : L'ajustement du modèle Γ -GQM aux données empiriques pour cinq échantillons de débit Q_6 élevé

Dans la Figure 4.39, les courbes des fréquences cumulées sont ordonnées en fonction du débit Q_6 . On observe aussi que les courbes de la fonction de densité sont ordonnées en deux parties et se croisent ensemble au TIV d'environ 2,2 s. La valeur du *mode* de l'échantillon (1800-1900) véh/h est un peu plus grande que celles de ses deux échantillons voisins, le CV de l'échantillon (1800-1900) véh/h est en même temps plus faible que ceux de ses deux échantillons voisins (1700-1800) et (1900-2000) véh/h . Ce sont probablement les raisons pour lesquelles les paramètres α et β du modèle Γ -GQM sont particulièrement

faibles et le paramètre θ est élevé pour l'échantillon (1800-1900) véh/h par rapport aux paramètres des deux échantillons voisins.

Q_6	μ_H	CV	α	β	θ	λ	P_χ^2	P_{K-S}
1600-1700	2.1953	0.6686	9.3353	10.3531	0.2723	0.6699	0.0628	0.8648
1700-1800	2.0515	0.6196	8.2470	9.6999	0.3276	0.7682	0.0974	0.9708
1800-1900	1.9390	0.6136	6.9555	8.9148	0.4716	0.8039	0.8279	0.9892
1900-2000	1.8645	0.6266	7.6414	9.2790	0.4661	0.8212	0.1176	0.3548
> 2000	1.7568	0.5584	8.0099	9.8825	0.4854	0.9839	0.8598	0.8535

TABLE 4.12 : Paramètres du modèle Γ -GQM pour les trois échantillons de très haut débit Q_6

La moyenne des TIV est inversement proportionnelle au débit Q_6 alors que le paramètre λ est toujours proportionnel au débit Q_6 . Ce dernier point est bien illustré par la partie descendante des TIV longs.

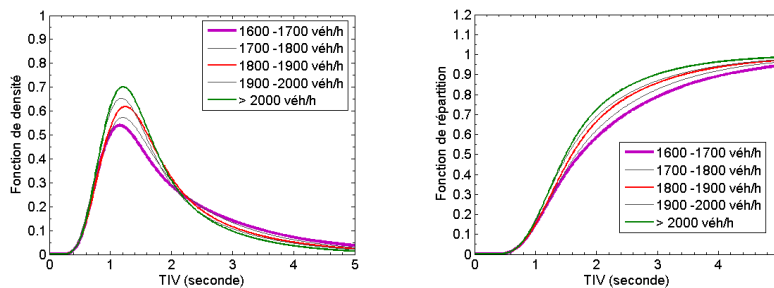


FIGURE 4.39 : La distribution des TIV selon le modèle Γ -GQM ajusté pour cinq niveaux de débit Q_6 élevé

4.5.5 L'existence de distributions bi-modales des TIV

Parmi tous les échantillons des TIV étudiés sur la VL de la RN118, les échantillons suivants ont un histogramme particulièrement différent des histogrammes *normaux* des TIV : les échantillons N°78, N°90 obtenus par la méthode de raffinement, les échantillons (400-500), (500-600) véh/h de la modalité jour-sec obtenus par la méthode de groupement. La Figure 4.40 illustre des cas bi-modaux des distributions de TIV.

Débit (véh/h)	μ_H	CV	α	β	θ	λ	P_χ^2	P_{K-S}
400-500	7.8925	1.1297	2.9259	5.2858	0.3401	0.1084	1.32×10^{-7}	0.0272
500-600	6.6772	1.1982	2.3056	4.4771	0.4457	0.1171	2.89×10^{-11}	0.0952
587 (N°78)	6.1266	1.3161	2.3492	4.5067	0.5316	0.1113	4.6×10^{-12}	0.0185
623 (N°90)	5.7794	1.4239	2.4754	4.6263	0.5681	0.1104	0.0144	0.1647

TABLE 4.13 : Paramètres du modèle Γ -GQM pour les 4 échantillons bi-modaux

Le Tableau 4.13 présente les caractéristiques de ces échantillons particuliers. En fonction du débit, la moyenne μ_H des TIV diminue tandis que le CV augmente. Les quatre échantillons ont tous des CV supérieurs à 1, les valeurs des paramètres α, β et λ du modèle Γ -GQM sont assez similaires. À cause du phénomène bi-modal, le modèle Γ -GQM ne peut pas être correctement ajusté aux données au sens du test du χ^2 . La Figure 4.41 montre l'effet bi-modal sur la fonction de répartition empirique.

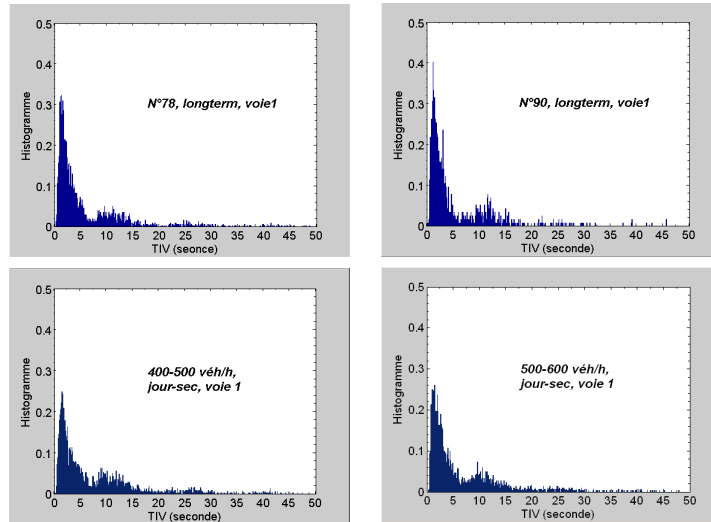


FIGURE 4.40 : L'ajustement du modèle Γ -GQM aux données empiriques pour les échantillons ayant un histogramme bi-modal

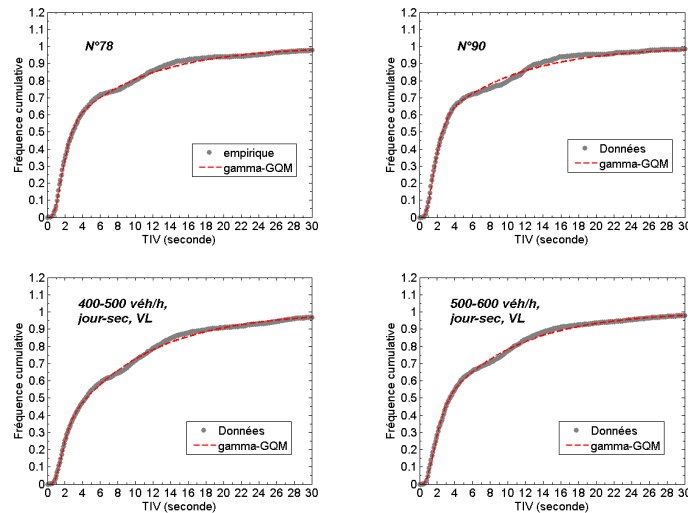


FIGURE 4.41 : Les échantillons ayant un histogramme bi-modal

4.5.6 Les échantillons dans la période du soir

Les modélisations des TIV dans les autres modalités de la période du soir, selon la condition météo et selon la voie ont aussi été réalisées. La Figure 4.42 montre brièvement le niveau d'ajustement du modèle Γ -GQM et du modèle DG en fonction du débit Q_6 dans toutes les conditions exogènes. On constate que :

- Les deux modèles Γ -GQM et DG ne sont rejetés dans aucun échantillon des modalités du **soir** sur la VL au niveau de signification de 5 % du test K-S.
- Les deux modèles fournissent aussi de bons ajustements pour tous les échantillons de la modalité **soir-pluie** sur la VR.
- Pour la modalité **soir-sec** sur la VR, le modèle DG est rejeté par le test K-S au niveau de signification de 5% et le modèle DG ne donne aucun résultat d'estimation pour certains échantillons de TIV ayant un débit Q_6 inférieur à 500 véh/h. Dans ce cas, le modèle Γ -GQM montre encore une bonne performance pour plusieurs échantillons.

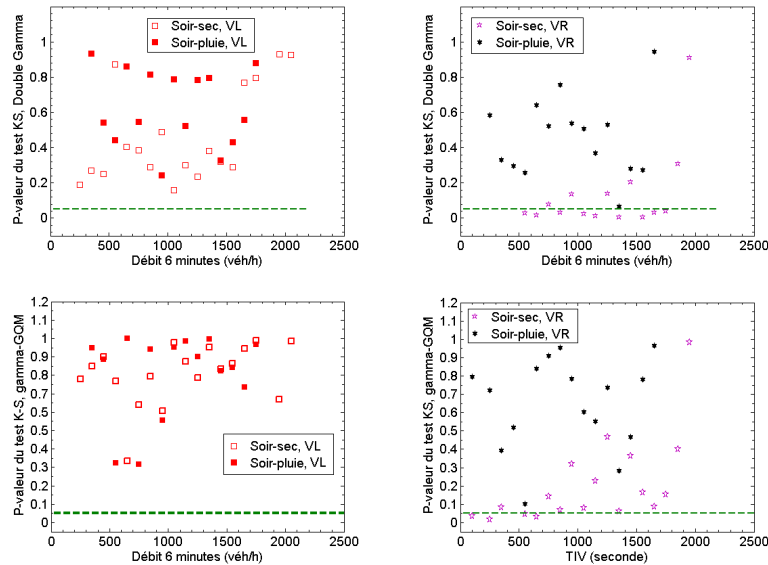


FIGURE 4.42 : L'ajustement des modèles Γ -GQM et DG le soir en fonction du débit Q_6 et selon la condition météorologique

Finalement, tous les résultats des paramètres du Site N°1 présentés dans les tableaux (B.2, B.3, B.4, B.5, B.6, B.7, B.8, B.9) de l'Annexe montrent la complexité des variations des paramètres notamment du modèle Γ -GQM selon la combinaison des facteurs de *gêne* et de *propulsion*. Les variations des paramètres du modèle DG semble plus continues, moins dispersées que celles des paramètres du modèle Γ -GQM.

4.6 Modélisation du comportement standard

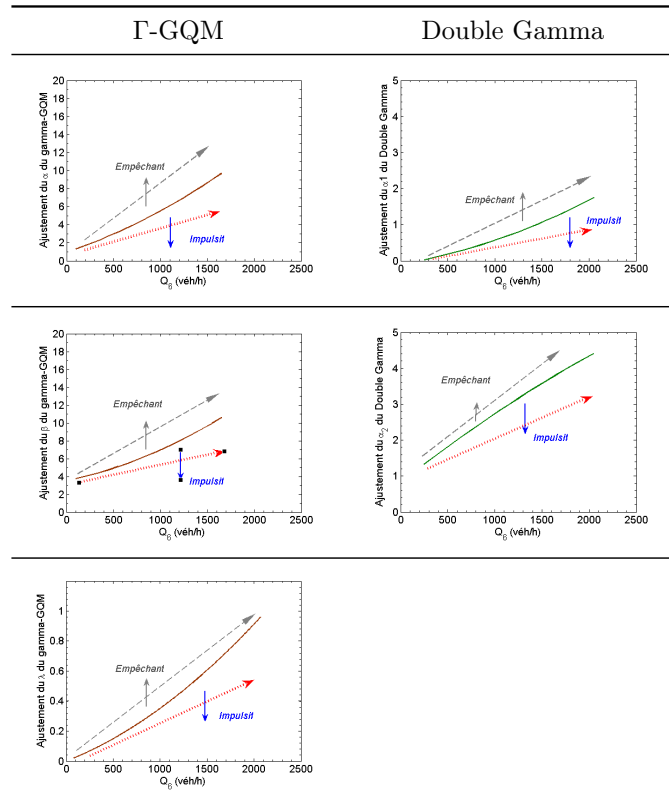
4.6.1 Schéma de principe

En partant des hypothèses de l'existence de facteurs exogènes et des évolutions des paramètres d'échelle des modèles, les schémas de variation des paramètres sont proposés dans le Tableau 4.14 pour les deux modèles étudiés, le modèle Γ -GQM et le modèle DG. Les schémas représentent donc le principe et les effets des facteurs exogènes sur le comportement standard. Les écarts entre les courbes semblent dépendre généralement de l'intensité de l'effet du facteur imposé.

L'effet des facteurs exogènes sur le paramètre β_1 du modèle DG n'est pas clair tandis que les effets des facteurs exogènes sur les paramètres θ des modèles Γ -GQM et DG sont observables. Cependant ces derniers sont difficiles à interpréter et modéliser. De toute façon, on suppose qu'il est possible de déduire des comportements de conduite sous des facteurs exogènes si l'on connaît *a priori* le comportement standard.

4.6.2 Régression des paramètres des modèles en fonction du débit Q_6

Les effets des facteurs exogènes sont quantifiés par des décalages des paramètres des modèles Γ -GQM et DG par rapport à ceux du comportement standard. Les niveaux des effets dépendent de l'intensité du facteur exogène étudié ce qui est hors du cadre de notre étude. Il semble que les facteurs affectent plus les paramètres d'échelle des modèles en trafic intermédiaire. En conséquence, seul le comportement standard est étudié par régression dans la section suivante.

TABLE 4.14 : Le schéma de principe de l'influence des effets exogènes \mathcal{F}_1

Suite aux analyses et aux interprétations proposées dans les sections précédentes, à partir de deux sources de données de la RN118 et de l'A6, et à partir de différentes méthodes d'échantillonnage, on suppose que les paramètres des modèles Γ -GQM et DG varient selon les principes suivants :

1. Il existe une façon de conduire des conducteurs générale, constante et déterminable sur la VL par temps **sec** le jour que l'on l'appelle le *comportement standard*. Ce comportement ne dépend que du type de réseau routier et des variables macroscopiques, particulièrement le débit Q_I .
2. Il y a deux types de facteurs exogènes *simples* (hors le trafic) influençant le *comportement standard* : le facteur *impulsif* (ou le facteur *de propulsion*) et le facteur *empêchant* (ou le facteur *de gêne*). Si le comportement standard peut être représenté par un modèle $\mathcal{P}(\Lambda)$ des TIV, avec $\Lambda = \{\lambda_i\}$ les paramètres d'échelle, les deux facteurs affectent en sens inverse les paramètres λ_i . Plus précisément, au même niveau du débit Q_I , le facteur *impulsif* réduit et le facteur *empêchant* augmente les paramètres Λ .

Les principes proposés semblent se manifester plus nettement en trafic intermédiaire. Pourtant, les facteurs exogènes *simples* veulent dire qu'ils ne sont pas facteurs *complexes* définis comme une combinaison des facteurs *simples* ; par exemple la condition de conduite par temps de **pluie** sur la VR est un facteur *complexe* qui se trouve hors de cette étude.

La méthode de régression polynomiale est utilisée. Les données d'entrée du processus de régression sont récapitulées pour les modèles dans le Tableau 4.15. Il est à noter que les données du débit Q_6 sont prises égales aux centres des intervalles de débit dans chaque échantillon. Les échantillons marqués « * » ne sont pas pris en compte dans la régression faute de l'interdépendance entre les paramètres ou de l'insuffisance d'ajustement. Dans ce cas-là, ils sont utilisés dans la phase de validation.

$Q_6/100$	Γ -GQM				DG				
	α	β	θ	λ	α_1	α_2	β_1	β_2	θ
0 – 2	1.6210	3.4028	0.0239	0.0511	0.0861	2.6044	0.3674	2.5843	0.6103
2 – 3*	5.2842	9.1940	0.2051	0.0727	0.0818	1.4296	1.4284	3.8826	0.6744
3 – 4	2.8094	5.0697	0.1940	0.0992	0.1284	1.6619	1.3862	3.6328	0.7076
4 – 5	2.9259	5.2858	0.3401	0.1084	0.1443	1.6801	2.1303	4.5192	0.6060
5 – 6	2.3056	4.4771	0.4457	0.1171	0.1594	1.7715	1.8694	4.0611	0.5038
6 – 7	3.0965	5.3591	0.3481	0.1669	0.2258	1.9288	1.9334	4.1856	0.5441
7 – 8	2.5441	4.7795	0.4412	0.1908	0.2438	1.8570	1.8685	4.1509	0.4793
8 – 9	6.0259	7.5271	0.1116	0.3008	0.3627	2.3543	1.3526	3.5676	0.4062
9 – 10	6.1040	6.8713	0.0586	0.3576	0.4345	2.5775	1.2408	3.3711	0.3238
10 – 11*	10.2902	10.1242	0.0382	0.3943	0.6419	2.9500	2.5277	4.9276	0.5570
11 – 12	5.6924	6.7503	0.0819	0.4663	0.8301	3.3623	2.8945	5.3592	0.5927
12 – 13	7.5667	8.2470	0.0992	0.5024	0.9006	3.2171	4.6153	7.1439	0.6595
13 – 14	8.4964	8.9388	0.1119	0.5436	0.9753	3.4047	4.3631	6.6917	0.5886
14 – 15	8.6051	10.2019	0.2317	0.5904	1.0918	3.6915	5.3712	7.9920	0.5277
15 – 16	8.3317	9.7588	0.2943	0.6093	1.0557	3.5234	5.0618	7.3687	0.4643
16 – 17	9.3353	10.3531	0.2723	0.6699	1.2138	3.9215	5.1101	7.2556	0.4282
17 – 18	8.2470	9.6999	0.3276	0.7682	1.4089	3.8734	6.8139	9.2188	0.5003
18 – 19	6.9555	8.9148	0.4716	0.8039	1.3693	3.7951	5.8230	8.2938	0.3820
19 – 20	7.6414	9.2790	0.4661	0.8212	1.3362	3.7709	5.5585	7.7534	0.3291
> 20	8.0099	9.8825	0.4854	0.9839	1.7741	4.5485	6.4364	8.8130	0.3244

TABLE 4.15 : Les entrées utilisées dans les régressions des paramètres des modèles Γ -GQM et DG pour le comportement standard

Dans le cadre de cette étude, seules les débits de 0 à 1650 véh/h en trafic fluide ou intermédiaire sont mis en jeu pour les modèles Γ -GQM et DG. Les résultats dans le Tableau 4.16 représente les courbes et droites de régression qui pourront être utilisées dans la pratique.

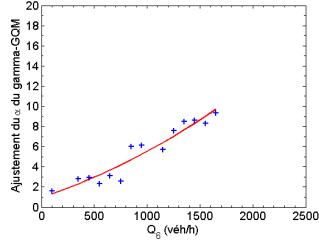
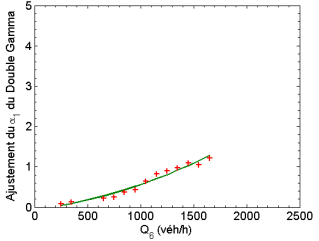
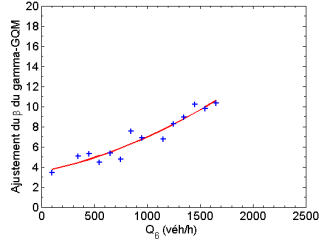
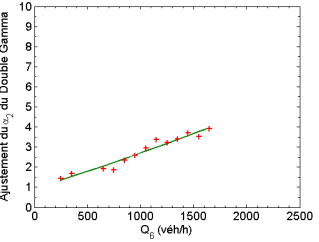
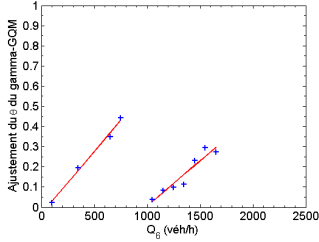
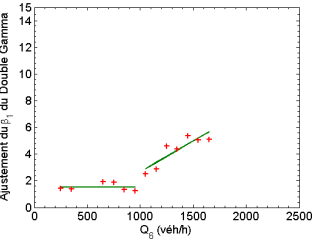
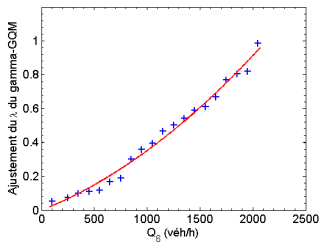
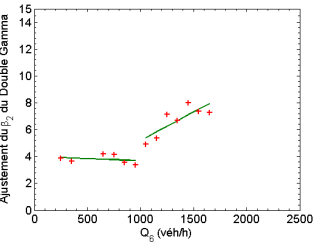
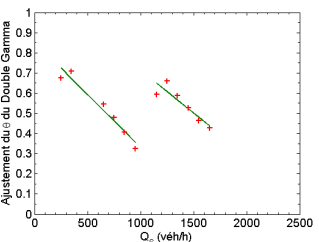
Γ -GQM	Double Gamma
	
$\alpha = 1.1268 \times 10^{-6} Q_6^2 + 0.0034 Q_6 + 1$	$\alpha_1 = 2.9108 \times 10^{-7} Q_6^2 + 3.3145 \times 10^{-4} Q_6 - 0.06$
	
$\beta = 1.2568 \times 10^{-6} Q_6^2 + 0.0022 Q_6 + 3.5$	$\alpha_2 = 1.0108 \times 10^{-7} Q_6^2 + 0.0017 Q_6 + 0.94$
	
$\theta^I = 6.1718 \times 10^{-4} Q_6 - 0.035$ $\theta^{II} = 4.4986 \times 10^{-4} Q_6 - 0.446$	$\beta_1^I = -1.0588 \times 10^{-5} Q_6 + 1.542$ $\beta_1^{II} = 0.0046 Q_6 - 1.1912$
	
$\lambda = 1.053 \times 10^{-7} Q_6^2 + 2.455 \times 10^{-4} Q_6$	$\beta_2^I = -3.1699 \times 10^{-4} Q_6 + 4$ $\beta_2^{II} = 0.0042 Q_6 + 0.963$
	
	$\theta^I = -5.2397 \times 10^{-4} Q_6 + 0.854$ $\theta^{II} = -4.1971 \times 10^{-4} Q_6 + 1.131$

TABLE 4.16 : Les régressions des paramètres des modèles Γ -GQM et DG pour le comportement standard, $R^2 > 0,9$

4.7 Conclusions du chapitre

1. Tout au long du chapitre, les modèles Γ -GQM et DG fournissent de très bons ajustements pour la majorité des échantillons de TIV quelque soit la condition prise en compte. Les échantillons horaires, les échantillons groupés, les échantillons raffinés ou les échantillons des états élémentaires sont tous bien calibrés par les modèles quelque soit le niveau de la demande macroscopique du trafic.
2. Le changement du contexte global (les facteurs exogènes et macroscopiques) induit un changement de la distribution des TIV illustré par l'évolution de paramètres des modèles Γ -GQM et DG. En conséquence, cela montre la possibilité d'établir un lien direct entre l'échelle macroscopique et l'échelle microscopique (en tenant en compte ou pas les conditions exogènes).
3. Les variations des paramètres des modèles en fonction des variables macroscopiques comme le débit, le taux d'occupation, la vitesse moyenne harmonique sont établies. Dans le modèle Γ -GQM, la relation entre α et β est confirmée; par ailleurs, l'interdépendance entre les paramètres α , β et θ est observée particulièrement en trafic dense; le paramètre λ est fortement corrélé au débit du trafic. Dans le modèle DG, les paramètres $\alpha_1, \alpha_2, \beta_1, \beta_2$ augmentent en fonction du débit. Dans les deux modèles, le paramètre θ reste difficile à interpréter totalement. Ce paramètre maintient certaines relations avec le pourcentage des véhicules contraints mais il semble qu'il soit considérablement influencé par les paramètres d'échelle α et λ .
4. La distribution des TIV est aussi étudiée le long du diagramme fondamental. La sensibilité des paramètres des modèles Γ -GQM et DG en fonction des états élémentaires est obtenue; elle est consistante avec celle obtenue dans le chapitre précédent. Le débit est confirmé comme le facteur macroscopique le plus important de la distribution des TIV.
5. L'analyse des facteurs exogènes conduit à distinguer deux types de facteurs : les facteurs *de gêne* et les facteurs *de propulsion*. Ces deux facteurs ont des effets différents sur les paramètres d'échelle α, λ du modèle Γ -GQM et α_1, α_2 du modèle DG du *comportement standard*.
6. Pourtant, les échantillons particuliers échappent parfois à la modélisation du modèle Γ -GQM.
7. Enfin, les paramètres des modèles sont mis en relation avec le débit Q_6 par la méthode de régression polynomiale. Ces résultats seront testés lors de la validation dans le chapitre suivant.

Chapitre 5

Niveaux mésoscopique et microscopique

Les entités dans la portée *mésoscopique* d'un véhicule sont visibles pour le conducteur du véhicule. Si les informations macroscopiques et exogènes ont un caractère général, les informations *mésoscopiques* sont plus concrètes et concernent les véhicules ou les groupes de véhicules aux alentours du véhicule en considération. Pourtant, les facteurs ne sont pas toujours quantifiables et certaines notions *mésoscopiques* n'ont pas encore une définition « officielle ». Par exemple, bien que la notion de *peloton* soit une notion essentielle du trafic à l'échelle *mésoscopique*, aucune définition quantitative de *peloton* n'est acceptée car les comportements de conduite observés à ce niveau dépendent fortement des données réelles. Ces dernières sont à leur tour très diversifiées en temps et en espace. Par ailleurs, peu d'études ont porté sur le TIV à cette échelle de trafic, les travaux de la distribution de TIV sont plus rares.

L'échelle *microscopique* concerne les informations du véhicule lui-même ou les informations partagées avec le véhicule en tête. L'étude du TIV à l'échelle microscopique sera intéressante si l'information de la vitesse des véhicules considérés est mise en jeu. Comme le comportement individuel est directement affecté par le véhicule de tête, certaines notions *mésoscopiques* n'impliquent que le véhicule précédent au lieu de tous les véhicules dans la portée *mésoscopique*. Par conséquent, sur ce point, la frontière entre les facteurs *mésoscopiques* et *microscopiques* est délicate. C'est aussi la raison pour laquelle, ce chapitre parlera en même temps des distributions des TIV aux niveaux méso et microscopique.

5.1 L'échelle mésoscopique

5.1.1 Le groupement des variables microscopiques

État de l'art

Athol [1] a étudié en 1965 le TIV pour mieux comprendre le comportement des conducteurs dans le trafic. L'approche d'Athol est mésoscopique, il distingue deux composantes du trafic : le *peloton* et le *groupe*. Sous ce point de vue, le *peloton* est défini comme un ensemble de véhicules consécutifs ayant un TIV inférieur à un seuil critique δ_H . Le *groupe* est au contraire l'ensemble de véhicules consécutifs ayant un TIV supérieur à δ_H .

Selon Athol, le seuil $\delta_H = 2,1 \text{ s}$ est le plus compatible avec les données dont il dispose car en adoptant ce seuil, une relation linéaire entre le pourcentage p_g des véhicules en *groupe* avec le débit Q a été trouvée. D'autant plus qu'en appliquant la relation trouvée au point $p_g = 0$ (autrement dit - ce pourcentage est d'ailleurs le complément à l'unité du

pourcentage de véhicules en peloton - 100 % de peloton), le débit Q obtenu est de 2500 véh/h ce qui est conforme à la capacité empirique, alors que tout autre seuil en supposant une relation entre p_g et Q conduirait à une valeur de débit en désaccord avec ces données empiriques.

Les statistiques ont montré la prédominance dans les pelotons des pelotons de taille 2 (peloton 2-véhicules) jusqu'au débit Q de 1700 véh/h, ce type de peloton représentant alors toujours plus de 50 % des pelotons. Toute autre taille supérieure n'a jamais été associée à plus de 20 % des pelotons. Pour les groupes, le groupe de 2-véhicules est aussi prédominant dans les données d'Athol. En général, Athol constate que plus la taille du *peloton* ou du *groupe* est grande, plus sa proportion relative est faible.

De plus, en utilisant un test statistique, Athol a constaté au niveau de signification de 5% qu'il y a pas de distance significative des moyennes et des variances des TIV selon la position dans le peloton n -véhicules. En conséquence, la moyenne des TIV en peloton sert à calculer la capacité dite *caractéristique* (ang. *characteristic volume*) de la route et la moyenne des TIV en groupe \bar{H}_g est supposée dépendre du débit Q . Athol a considéré que les TIV des groupes suivent une loi exponentielle \mathcal{E} de paramètre ρ , c. à. d $g(y) = e^{-\rho y}$. Le paramètre ρ varie différemment du paramètre λ d'un modèle \mathcal{E} appliqué à l'ensemble des TIV, comme le montre la Figure 5.1.

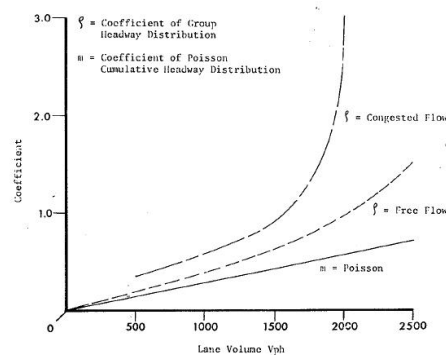


FIGURE 5.1 : Les paramètres d'intensité ρ et λ des modèles \mathcal{E} des TIV normaux et des TIV en groupe en fonction du débit Q (Source : Athol [1])

En résumé, l'approche mésoscopique d'Athol se base principalement sur la définition du seuil δ_H ce qui entraîne l'identification du *peloton* et du *groupe*. Athol suggère que la valeur convenable de δ_H est comprise entre 2 et 2,5 s.

Dans les modèles *complexes* (combiné, de mélange) des TIV, les notions de véhicules *contraints* et véhicules *libres* sont essentielles. Ces deux catégories de véhicules sont en effet caractérisées par des modèles probabilistes différents. Cependant, le seuil entre ces deux catégories est théorique ce qui incite donc à des études quantitatives afin d'éclaircir cette notion et afin de déterminer de façon quantitative ce seuil.

Tout d'abord, la détermination a été réalisée via l'estimation du paramètre θ dans les modèles complexes des TIV. En principe, ce paramètre représente la part des TIV contraints, tandis que le paramètre $\alpha_l = 1 - \theta$ correspond à la part des véhicules libres. Les auteurs proposent deux modèles simples pour calibrer α_l en fonction du débit q :

- La relation *linéaire* de Tanner (1962 et 1967) [130, 45] est de la forme $\theta = 1 - \Delta q$ où Δ est le TIV minimum. Dans la même approche, Sullivan et Troutbeck (1997) [71] ont proposé $\alpha_l = 0,8 - 0,0005 Q/n_v$ où Q le débit en véh/h et n_v le nombre de voies de la route.
- La relation *exponentielle* proposée par Brilon (1988) [131] est de la forme $\alpha_l = e^{Aq}$. Dans cette même approche, Akçelik *et al.* [72] ont proposé $\alpha_l = e^{-b\Delta q}$ où A est un

paramètre positif, Δ est le TIV minimum et b est le coefficient de peloton.

Si les approches ci-dessus proposent un seuil variable en fonction de l'état du trafic par l'intermédiaire des paramètres θ ou α_l , d'autres approches considèrent un seuil fixe de TIV δ_H , par exemple Aly en 1989 [65] a choisi $\delta_H = 6$ s, alors que Haging [77] a proposé une dépendance linéaire du seuil δ_H au débit Q du trafic.

En somme, l'approche mettant les paramètres θ et δ_H en relation avec le débit est faite surtout par des chercheurs qui étudient le modèle M3 de Cowan ou qui étudient la *queue* de la distribution des TIV en utilisant le modèle de mélange g -SPM comme Camiz [132], Hoogendoorn [133], etc.

Sous un point de vue différent, Vogel (2002) [134] a considéré qu'un véhicule libre est tel que sa vitesse ne dépend pas de la présence du véhicule *en tête*. En partant de cette définition encore qualitative, la méthode de détermination est de vérifier si la source d'influence constituée par le véhicule précédent restreint la vitesse du véhicule suivant.

Vogel a étudié le trafic fluide à une intersection urbaine et montré que le seuil δ_H du TIV n'était pas suffisant pour définir un véhicule libre. L'étude est basée sur la corrélation entre les vitesses des véhicules consécutifs dont le TIV est rangé dans des classes allant de moins d'une seconde à plus de 12 s. Vogel a proposé de prendre aussi en compte la DIV en groupant les véhicules de DIV de 10 m jusqu'à 170 m.

La méthode de Vogel s'appuie ensuite sur la régression linéaire par morceaux des valeurs de corrélation des vitesses en fonction du TIV et de la DIV en vue de déterminer les seuils critiques δ_H sur le TIV et δ_d sur la DIV, bien que cette relation soit curviligne. Tout d'abord, un seuil préalable est prédéterminé pour diviser les valeurs de corrélation en deux parties. Ensuite, chaque partie de valeurs de corrélation est régressée et les coefficients R^2 et r (pourcentage de variance expliquée et coefficient de corrélation) sont calculés. Enfin, le seuil critique est le TIV ou la DIV correspondant à l'intersection des deux droites de régression trouvées. Vogel a obtenu $\delta_H = 6$ s et $\delta_d = 75$ m. Le test de Student a permis de constater que les corrélations des vitesses sont plus dispersées en fonction de la DIV qu'en fonction du TIV. En outre, les résultats de régression selon le TIV sont bien plus significatifs que ceux selon la DIV.

Par ailleurs, les corrélations des vitesses entre les 2 branches principales et les 2 branches secondaires de l'intersection sont similaires ce qui incite à conclure que le seuil pourrait ne pas dépendre du niveau de trafic. En déterminant les seuils critiques selon les corrélations des vitesses, l'approche permet d'interpréter le changement des seuils en fonction de différents facteurs, par exemple : la limite de vitesse (quand la limite augmente, le seuil baisse), la réaction plus ou moins tardive de l'ensemble de conducteurs (plus elle est tôt, plus le seuil augmente), l'expérience du conducteur (les débutants ont des seuils plus petits que les conducteurs expérimentés), l'état psycho-physiologique du conducteur (fatigué ou attentif) (cf. Vogel [134]).

Finalement, Vogel utilise le seuil de 6 s pour distinguer entre « *conduire librement* » et « *sous quelque influence du véhicule devant* ». Cela aboutit à discriminer trois phases lorsqu'un véhicule rattrape un autre véhicule en tête plus lent :

1. Conduire librement à vitesse désirée jusqu'à ce que son TIV soit d'environ 6 s.
2. Approcher le *leader* et commencer à adapter sa vitesse à celle de celui-ci.
3. Conduire avec la même vitesse que le leader et garder un TIV relativement constant. Ce TIV constant va dépendre de facteurs comme la « *pression du temps* », la familiarité à l'environnement, etc.

Plus précisément, si l'on considère qu'un véhicule libre est celui dont la vitesse ne dépend pas de la vitesse du véhicule précédent, une méthode aussi utilisable est de tracer l'écart-type des vitesses relatives v_r selon des classes de TIV. Luttinen (1992) [97] a constaté

des écarts-types constants à partir du seuil δ_H de l'ordre de 9 s. Ce seuil veut dire que les conducteurs n'ajustent plus leur vitesse à celle du véhicule précédent lorsque le TIV est au delà de 9 s.

He *et al.* (2002) [115], ont étudié le TIV *préféré* (**ang.** *preferred time headway* (PTH)) qui est le TIV entre deux véhicules successifs roulant à des vitesses très proches, en pratique, lorsque la valeur absolue de la vitesse relative v_r est inférieure à 1,524 m/s (équivalent à 5 feet/s) (cf. Aycin [135]).

He *et al.* (2002) proposent une nouvelle méthode pour identifier l'état de poursuite en corrélant des vitesses relatives v_r et les TIV. Logiquement, lorsque les TIV sont courts, la corrélation est significative et la vitesse relative est faible en valeur absolue, alors qu'à partir d'une certaine valeur de TIV, la vitesse relative ne dépend plus du TIV et donc la pente de relation est approximativement horizontale. La nouvelle méthode consiste à trouver le seuil sur les TIV par changement de coordonnée tel qu'en deçà il y a corrélation et qu'au delà, la corrélation est nulle. Résultat : le seuil de TIV trouvé est de l'ordre de 6,1 s. En conséquence, les auteurs ont proposé 3 états de conduite comme suit :

1. TIV < 6 s : le véhicule poursuit le véhicule précédent,
2. TIV > 8 s : le véhicule est libre,
3. 6 s < TIV < 8 s : le véhicule est partiellement contraint.

Finalement, une autre approche qui permet de distinguer les composantes du trafic selon le comportement de conduite est de considérer que pour les TIV supérieurs au seuil δ_H , les TIV suivent une loi exponentielle \mathcal{E} . Cette supposition conduit à tester l'hypothèse nulle suivante :

$$H_0 : 1 - F_n(h) \sim e^{-\alpha h} \quad \text{pour } \forall h > \delta_H$$

Luttinen (1996) [11] a obtenu un seuil δ_H d'environ 8 s après avoir testé l'hypothèse H_0 . Le test utilisé est celui du χ^2 au niveau de signification de 5 %.

Application sur les données de la RN118

Comme le paramètre θ des modèles Γ -GQM et DG ne représente que la part des véhicules contraints, les approches de Vogel, de Luttinen et de He sont utilisées dans cette étude. En se basant sur les approches citées, l'application aux données de la RN118 a été réalisée. Les informations individuelles comme le TIV h_i , la DIV d_i , le TIV net g_i et la vitesse instantanée v_i sont divisées en groupes de valeurs croissantes. En conséquence, les classes correspondantes \mathcal{C}_h , \mathcal{C}_d , \mathcal{C}_g et \mathcal{C}_v comportent des véhicules individuels. Au sein de chaque classe, les coefficients de variation entre les variables microscopiques des véhicules consécutifs sont calculés et fournissent des comportements de conduite. Les valeurs moyennes et la variance des variables microscopiques dans les classes sont aussi calculées.

Dans ce chapitre, différentes périodes de données du Site N°1 de la RN118 sont étudiées. Les classes suivantes ont été constituées :

- \mathcal{C}_h de 0 à 30 s avec incrément de 0,5 ou 1 s par classe
- \mathcal{C}_g de 0 à 30 s avec incrément de 0,5 ou 1 s par classe .
- \mathcal{C}_d de 0 à 150 m avec incrément de 5 ou 10 m par classe .
- \mathcal{C}_v de 0 à 50 m/s avec incrément de 5 m/s par classe .

Les caractéristiques suivantes sont étudiées dans chaque classe : la moyenne \bar{L} et l'écart-type σ_L des longueurs des véhicules, la moyenne \bar{V} et l'écart-type σ_V des vitesses instantanées des véhicules, la moyenne \bar{V}_r et l'écart-type σ_r des vitesses relatives des véhicules, le coefficient de corrélation entre les longueurs des véhicules consécutifs ρ_{LL} , le coefficient de corrélation entre les longueurs et les vitesses des véhicules ρ_{LV} , le coefficient de corrélation

entre les longueurs et les vitesses relatives des véhicules ρ_{LV_r} , le coefficient de corrélation entre les vitesses des véhicules consécutifs ρ_{VV} et le coefficient de corrélation entre les vitesses et les vitesses relatives des véhicules ρ_{VV_r} . Où le coefficient de corrélation entre deux variables A de valeurs a_i et B de valeurs b_i est :

$$\rho_{AB} = \rho_{BA} = \frac{\sum_i (a_i - \bar{A})(b_i - \bar{B})}{\sqrt{\sum_i (a_i - \bar{A})^2 \sum_i (b_i - \bar{B})^2}} \quad (5.1)$$

avec \bar{A} et \bar{B} sont les moyennes des a_i et b_i respectivement.

5.1.2 Résultats

Les variations des coefficients étudiés sont assez larges et dépendent de la grandeur de l'incrément choisi. Par exemple, les variations des coefficients obtenues dans les classes \mathcal{C}_h avec un incrément de 0,5 s sont évidemment plus grandes que celles obtenues dans les classes \mathcal{C}_h avec un incrément de 1 s. Par ailleurs, ces grandes variations sont aussi la conséquence des différentes périodes de données enregistrées sur le Site N°1 de la RN118. En conséquence, la méthode quantitative de Vogel précitée est d'application difficile sur tout l'ensemble des données disponibles, d'autant plus que le groupement en classes entraîne des approximations au moins du même ordre que la valeur de l'incrément.

Bien que la dispersion soit non négligeable, il y a toujours des comportements généralement explicables à partir des résultats obtenus. Les interprétations sont présentées dans les sous sections suivantes :

Groupement en fonction de la DIV d_i

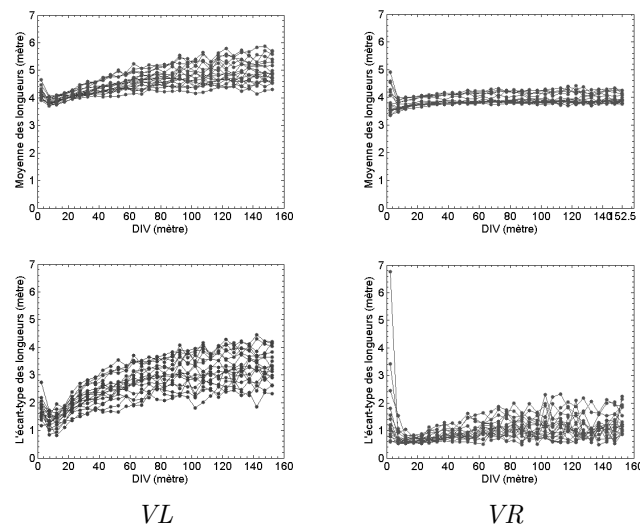


FIGURE 5.2 : Moyenne \bar{L} et écart-type σ_L des longueurs en fonction de la DIV

Sur la VL, plus la DIV est grande, plus la moyenne \bar{L} et l'écart-type σ_L des longueurs des véhicules sont élevés. Sur la VR, la moyenne \bar{L} est très stable et l'écart-type σ_L est beaucoup plus faible que sur la VL (cf. Figure 5.2). Ces résultats confirment que les véhicules plus longs circulent principalement sur la VL. De plus, sur la VL, les conducteurs ayant conscience de la longueur de leur véhicule ont tendance à laisser une distance proportionnelle à cette longueur. Les véhicules de longueur moins importante, ou bien les véhicules particuliers (VP) peuvent choisir une DIV courte ou longue, mais cela n'est pas

le cas pour les véhicules longs, ou bien les Poids lourds (PL). Ceci explique des plus grandes valeurs de σ_L aux grandes DIV. Il semble que la DIV devant les PL n'est pas choisie, elle résulte du plus ou moins grand éloignement du VP plus rapide devant.

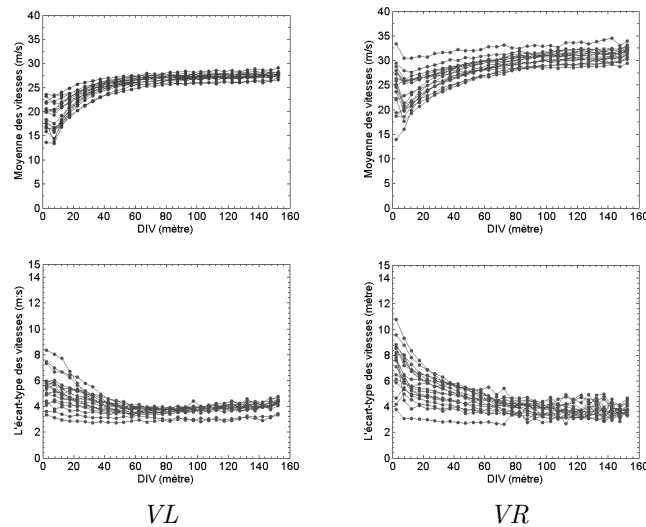


FIGURE 5.3 : Moyenne \bar{V} et écart-type σ_V des vitesses en fonction de la DIV

Chaque ligne de variation de la figure précédente représente une période de mesure en le même point sur le Site N°1 de la RN118, correspondant à un échantillon *primitif*. Bien que la moyenne \bar{V} soit plus faible dans les classes de DIV courte que dans les classes de DIV longue, l'écart-type σ_V est plus élevé dans les classes de DIV courte. Les variations des moyennes \bar{V} et des écarts-types σ_V des vitesses selon les périodes sont élevées parce que les durées des périodes sont différentes et les sous-périodes de jour, de soir, de la nuit sont éventuellement toute confondues.

Sur la VL, la moyenne et l'écart-type des vitesses deviennent stables pour les DIV supérieures à 70 m. La vitesse constante est de l'ordre de 25 m/s (soit 90 km/h) sur la VL, et de l'ordre 30 m/s (soit plus 108 km/h) sur la VR, ce qui correspond à la vitesse *limite*.

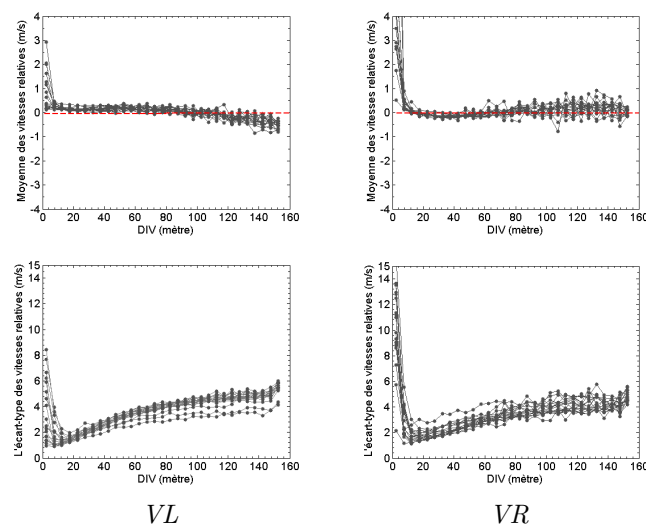


FIGURE 5.4 : Moyenne \bar{V}_r et écart-type σ_{V_r} des vitesses relatives en fonction de la DIV

Dans toutes les classes \mathcal{C}_d de DIV, les valeurs et la variation des moyennes \bar{V} et des

écarts-types σ_V sur la VR sont significativement plus élevées que celles correspondant sur la VL (cf. la Figure 5.3). En général, les véhicules sur la VR se déplacent plus vite que ceux sur la VL. Par ailleurs, les conducteurs sur la VR semblent plus mobiles et modifient leurs vitesses plus souvent pour s'adapter aux conditions du trafic.

L'étude de la vitesse relative permet de mieux comprendre le comportement des conducteurs par rapport au véhicule précédent au niveau de la vitesse. Sur la VL, la moyenne \bar{V}_r est toujours positive pour les DIV inférieures à 70 m, c.à.d la vitesse du véhicule précédent est plus faible. Dans les classes de DIV supérieure à 100 m, la moyenne \bar{V}_r est négative. Sur la VR, la moyenne \bar{V}_r est légèrement inférieure à zéro dans les classes de DIV de 20 à 60 m. Dans les classes de DIV supérieure à 100 m, la \bar{V}_r est positive. Dans les classes de DIV supérieure à 10 m, les moyennes \bar{V}_r sont entre -1 et 1 m/s ; l'écart-type σ_{V_r} augmente en fonction de la DIV.

Sur la VL, la Figure 5.5 montre tout d'abord que le coefficient ρ_{LL} est positif pour les DIV inférieures à 90 m ce qui indique qu'il est plus rare qu'un véhicule court (VP) suive de près un véhicule long (PL) que l'inverse - ceci peut s'expliquer par le fait que les PL sont distancés par les VP.

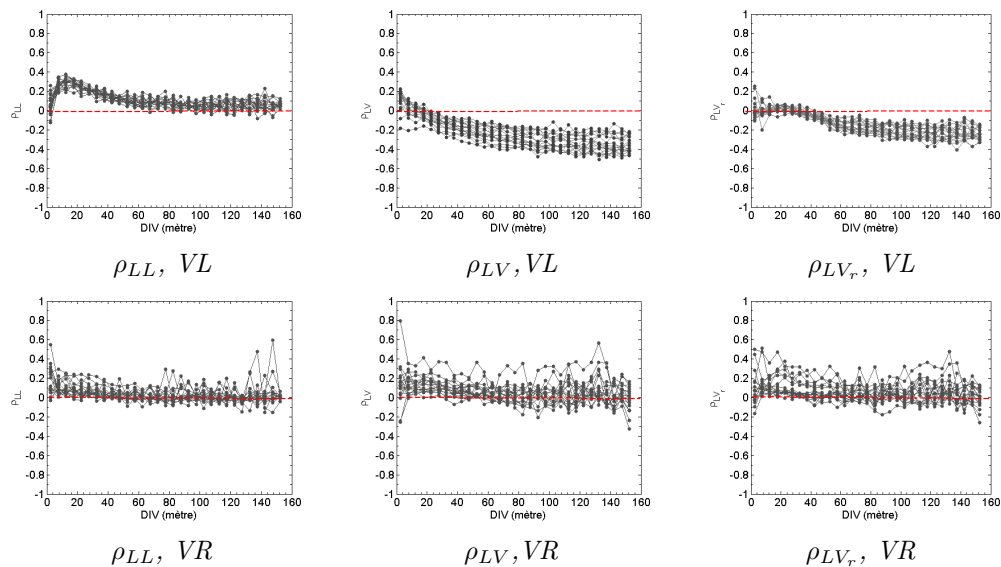


FIGURE 5.5 : Les coefficients ρ_{LL} , ρ_{LV} et ρ_{LV_r} en fonction de la DIV

Lorsque la DIV est grande, la corrélation entre les longueur est faible. Sur la VL, le coefficient ρ_{LV} diminue en fonction de la DIV et est majoritairement négatif. Ce résultat montre que plus le véhicule est long, plus la vitesse choisie est plus faible ce qui peut être dû à la limitation de vitesse des PL (de 90 km/h), plus faible que celle des VP.

Le coefficient ρ_{LV_r} est autour de zéro dans les classes de DIV inférieure à 30 m, et diminue ensuite en fonction de la DIV. Il semble que les vitesses relatives dans les classes de DIV courte soient proches de zéro et soient indépendantes de la longueur des véhicules. Dans les classes de DIV plus longue, il est confirmé que plus le véhicule est long, plus sa vitesse relative est faible.

Le comportement des conducteurs sur la VR est différent de celui sur la VL. Les remarques précédentes ne sont plus valables parce qu'il n'y a pas deux types bien différenciés de véhicules sur la VR. Il est à noter que la moyenne \bar{L} des longueurs des véhicules sur la VR est très stable.

Les coefficients ρ_{VV} et ρ_{VV_r} sont toujours positifs. À partir d'une DIV de 20 m quelque soit la voie, le coefficient ρ_{VV} diminue tandis le coefficient ρ_{VV_r} augmente. Le coefficient

ρ_{VV} représente la corrélation entre la vitesse d'un véhicule et la vitesse du véhicule précédent. Les valeurs positives et la variation de ρ_{VV} sont expliquées pour les DIV faibles par l'existence de pelotons où les véhicules suivants adaptent leur vitesse à celle du véhicule précédent ; lorsque la DIV augmente, cette adaptation s'annule. En général, la rapidité/lenteur du véhicule précédent semble stimuler/empêcher la volonté de rouler vite du véhicule suivant. Le niveau d'influence de la stimulation diminue en fonction de DIV.

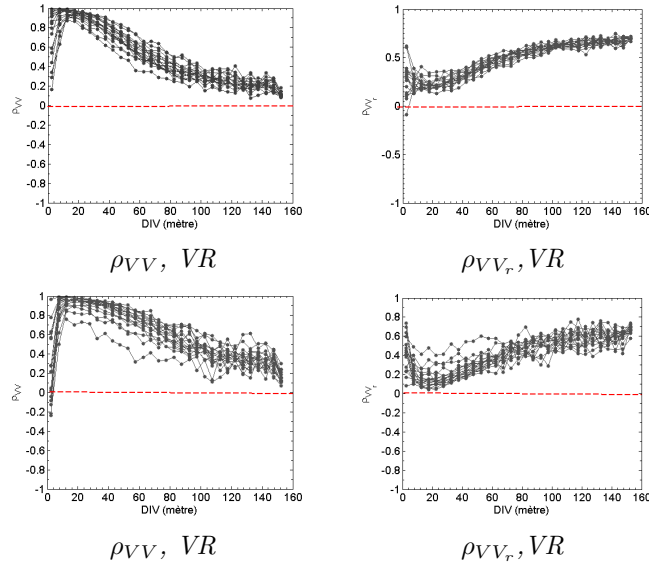


FIGURE 5.6 : Les coefficients ρ_{VV} et ρ_{VV_r} en fonction de la DIV

La variation du coefficient ρ_{VV_r} semble être expliquée par le coefficient ρ_{VV} si l'on suppose que les vitesses instantanées V_i des véhicules i sont i.i.d. En effet :

$$\rho_{VV_r} = \text{Corr}[V_i, (V_i - V_{i-1})] \sim 1 - \text{Corr}[V_i, V_{i-1}] \quad (5.2)$$

Groupement en fonction du TIV

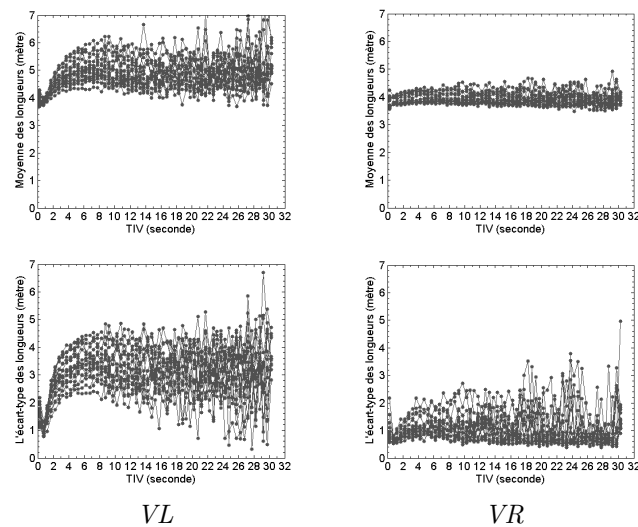


FIGURE 5.7 : Moyenne \bar{L} et écart-type σ_L des longueurs en fonction du TIV

En fonction du TIV, la moyenne \bar{L} et l'écart-type σ_L des longueurs des véhicules sont beaucoup plus stables sur la VR que sur la VL. La moyenne \bar{L} de 4 m indique que la

majorité des véhicules se déplaçant sur la VR sont des véhicules particuliers (VP). Sur la VL, la longueur ne dépend du TIV que lorsque le TIV est court. À partir d'un TIV de 6 à 7 s, il n'y a pas de tendance de la moyenne \bar{L} selon le TIV. La magnitude de variation dépend dans ces cas de la période étudiée.

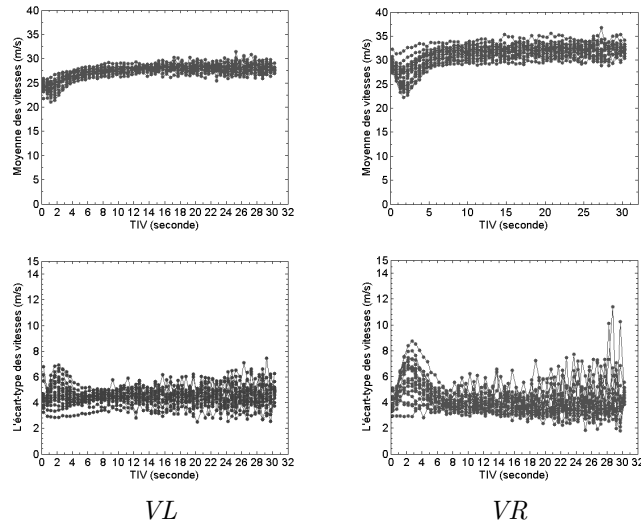


FIGURE 5.8 : Moyenne \bar{V} et écart-type σ_V des vitesses en fonction du TIV

En fonction du TIV, la moyenne \bar{V} et l'écart-type σ_V des vitesses sont plus faibles sur la VL que sur la VR. Dans les classes de TIV d'environ 2 à 3 s, les plus faibles valeurs des vitesses s'accompagnent de grande valeurs de l'écart-type σ_V . Il semble qu'à ce seuil, il y ait un certain nombre des freinages ou bien des véhicules zigzagant entre les voies. À partir d'un TIV supérieur à 7 s, la \bar{V} est assez constante, de l'ordre 25 m/s sur la VL et de l'ordre 30 m/s sur la VR, ce qui est cohérent avec les remarques de la section précédente.

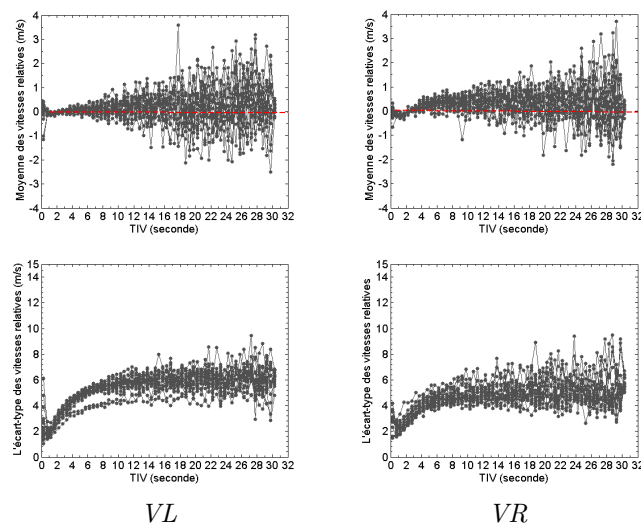


FIGURE 5.9 : Moyenne \bar{V}_r et écart-type σ_{V_r} des en fonction du TIV

La moyenne \bar{V}_r et l'écart-type σ_{V_r} des vitesses relatives sont faibles dans les classes de TIV court. Au contraire, la moyenne \bar{V}_r dans les classes de TIV élevé est très dispersée et dépend fortement de la période de mesure. À partir d'un certain seuil, disons 7 s, la variation des écarts-types σ_{V_r} semble ne plus dépendre du TIV. Cette observation est compatible avec les résultats de Luttinen [11] et de He *et al.* [115].

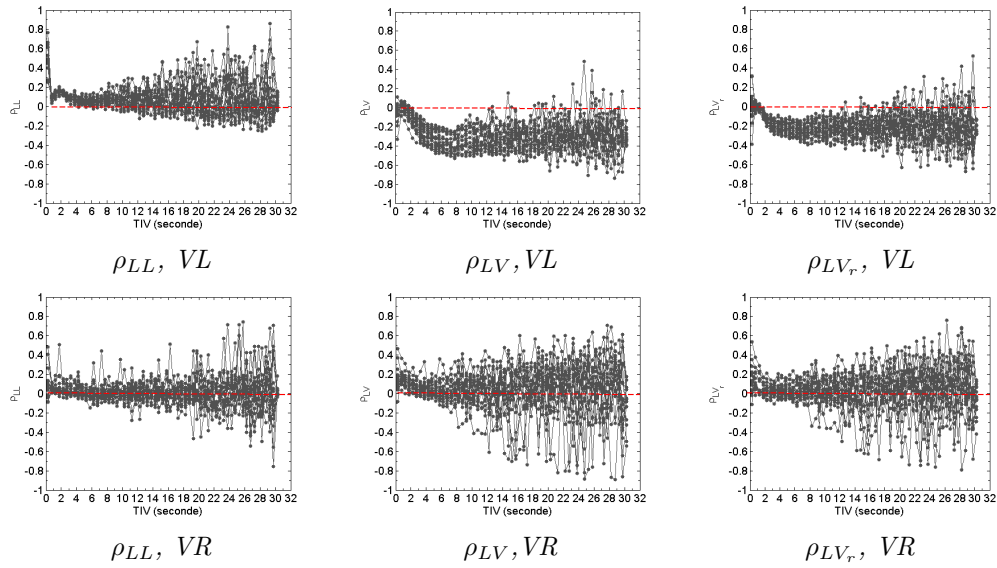


FIGURE 5.10 : Les coefficients ρ_{LL} , ρ_{LV} et ρ_{LV_r} en fonction du TIV

Comme dans le cas de la DIV, le coefficient ρ_{LL} sur la VL est principalement positif dans les classes de TIV court. Les coefficients ρ_{LV} et ρ_{LV_r} sont autour de zéro pour les TIV très courts (moins de 2 s) et diminuent jusqu'au TIV de 6 à 8 s. Pour les TIV supérieurs à 7 ou 8 s, les valeurs des coefficients fluctuent autour d'une valeur constante de l'ordre de -0,3 ou -0,2. Au contraire, ces mêmes coefficients de corrélation sur la VR fluctuent autour de la valeur zéro.

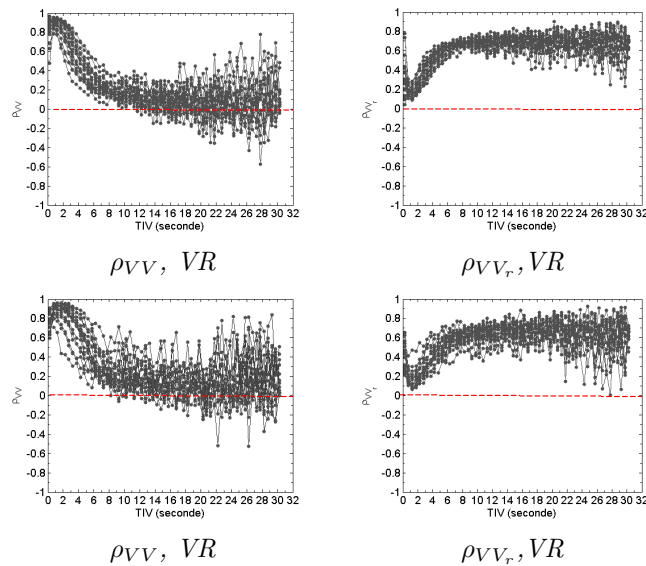


FIGURE 5.11 : Les coefficients ρ_{VV} et ρ_{VV_r} en fonction de la TIV

Les vitesses des véhicules successifs sont très corrélées dans les classes de TIV très courts (moins de 2 s). Le coefficient ρ_{VV} diminue en fonction du TIV et est positif dans la majorité des classes. Pour les classes de TIV supérieur à 7 s, les coefficients ρ_{VV} semblent osciller autour d'une valeur constante et faible tandis que les coefficients ρ_{VV_r} semblent ne plus dépendre du TIV.

Groupement en fonction de la vitesse

La Figure 5.12 confirme que la moyenne \bar{L} et l'écart-type σ_V des longueurs des véhicules sur la VL sont plus grandes que sur la VR. Il faut noter l'existence de classes de vitesse inférieure à 10 m/s et supérieure à 45 m/s dans lesquelles la variabilité est particulièrement élevée, probablement à cause d'un nombre insuffisant de données. Il apparaît qu'aux vitesses comprises entre 20 à 25 m/s correspondent des véhicules longs tandis que la plupart des VL se déplacent à une vitesse entre 25 à 40 m/s.

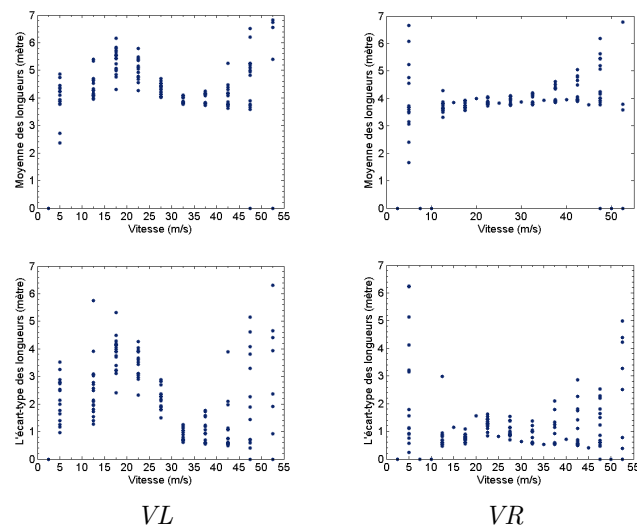


FIGURE 5.12 : Moyenne \bar{L} et écart-type σ_L en fonction de la vitesse

La moyenne \bar{H} des TIV augmente avec la vitesse dans les classes de vitesse de 10 à 45 m/s. En groupant les données selon la vitesse, la moyenne \bar{H} et l'écart-type σ_H des TIV sur la VR sont plus variables que ceux sur la VL selon la période de mesure.

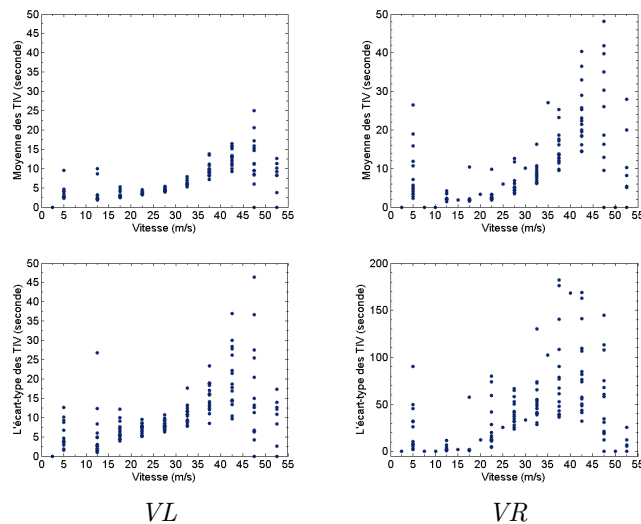


FIGURE 5.13 : Moyenne \bar{H} et écart-type σ_H des TIV en fonction de la vitesse

Dans les classes de vitesse inférieure à 30 m/s et supérieure à 10 m/s, la moyenne \bar{V}_r des vitesses relatives sur les deux voies est autour de zéro. À partir d'une vitesse supérieure à 30 m/s, la moyenne \bar{V}_r et l'écart-type σ_{V_r} des vitesses relatives augmentent en fonction de la vitesse du véhicule.

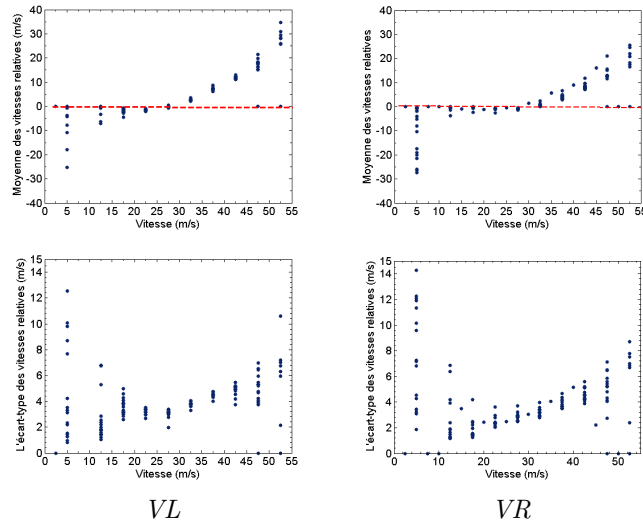


FIGURE 5.14 : Moyenne \bar{V}_r et écart-type σ_{V_r} en fonction de la vitesse

Finalement, on observe que le coefficient ρ_{HV_r} est proche de zéro dans les classes de vitesse de 30 m/s. Il est négatif dans les classes de vitesse inférieure à 30 m/s et positif dans le cas contraire.

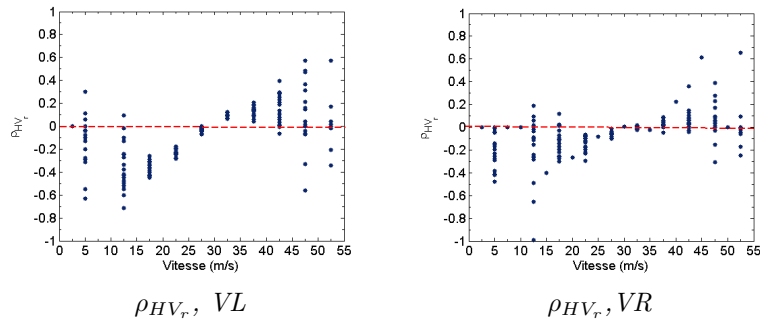


FIGURE 5.15 : Le coefficient ρ_{HV_r} en fonction de la vitesse

Discussion

En considérant tous les résultats obtenus dans la section précédente, on observe deux difficultés pour identifier les seuils permettant de définir le peloton. La première est que même sur un même lieu, le comportement des conducteurs change selon la période. Chaque ligne de points dans les figures précédentes représente une période et il est constaté que la variation d'une période à l'autre est non négligeable. La deuxième difficulté réside dans la largeur de chaque classe des variables groupées ce qui ne permet pas de calculer précisément un seuil à partir duquel le comportement des conducteurs connaît un changement significatif. En conséquence, sous un point de vue global et intuitif, la définition suivante est proposée :

Cette définition signifie qu'un peloton est un troupe des véhicules qui se déplacent ensemble avec des distances en temps et en espace suffisamment proches et une stabilité de vitesse au sein de l'ensemble.

Un véhicule est un suiveur dans un peloton si à la fois :

1. Son temps inter-véhiculaire **TIV** est inférieur ou égal à 7 s, \mathcal{B}_1 ,
2. Sa distance inter-véhiculaire **DIV** est inférieure ou égale à 100 m, \mathcal{B}_2 ,
3. Sa vitesse relative est inférieure ou égale en valeur absolue à 2 m/s, \mathcal{B}_3

$$i := \text{un suiveur} \iff \{h_i \leq 7, d_i \leq 100, |V_r| \leq 2\}$$

Par ailleurs, dans notre étude, un véhicule i est classé « long » et noté PL si sa longueur l_i supérieure ou égale à 7 m. Autrement dit :

$$i := \text{un PL} \iff \{l_i \leq 7\}$$

En conformément à ces nouvelles définitions, les véhicules qui ne sont pas *suiveurs* sont les véhicules *libres*. Un véhicule libre est soit un *leader* suivi par au moins un suiveur, soit un *baladeur* (ou bien un *véhicule isolé*) suivi par un autre véhicule libre. Les véhicules qui ne sont pas PL sont les véhicules courts, noté VP à la place de véhicule particulier.

Dans la section suivante, les trois approches mésoscopiques sont appliquées en étudiant la distribution des TIV :

1. La première est fondée sur la position des véhicules en peloton : les *leaders* et les *baladeurs* sont les véhicules P_ℓ en 1^{ème} position en peloton ; les véhicules P_2 en 2^{ème} position du peloton, les véhicules P_3 en 3^{ème} position du peloton et les véhicules P_* en $n^{\text{ème}}$ position du peloton, avec n est supérieur à 4.
2. La deuxième approche est fondée sur la configuration du trafic dont quatre catégories de poursuite sont considérés : catégorie (VP-VP) d'un VP suivant un VP, catégorie (VP-PL) d'un VP suivant un PL, catégorie (PL-VP) un PL suivant un VP et enfin catégorie (PL-PL) un PL suivant un PL.
3. La troisième approche est fondée sur la valeur du TIV T_p du véhicule précédent. Dans cette approche, les véhicules sont groupés selon les valeurs H_p des TIV des véhicules précédents.

5.1.3 Distribution des TIV et l'échelle mésoscopique

La position du véhicule en peloton

Athol a montré que la moyenne des TIV est stable quelque soit la position des véhicules en peloton. Cette section a pour but de représenter les distributions des TIV des véhicules en peloton selon leur position. Il est évident que si la définition du peloton ne met en jeu que le seuil δ_H sur le TIV, $h_i < \delta_H$, la distribution des TIV peut être obtenue directement par les modèles dans le chapitre de modélisation. En se basant sur la définition proposée, l'échantillonnage est réalisé par la méthode de groupement en distinguant trois périodes *jour-sec*, *jour-pluie*, *soir-sec*. Les modèles Γ -GQM et DG sont calibrés à partir des données de la RN118.

La Figure 5.16 représente les distributions des TIV des P_ℓ , P_1 et P_2 dans les trois périodes. On remarque que la distribution des TIV des P_ℓ est significativement différente de la distribution des TIV des suiveurs. Cependant, les distributions des suiveurs sont très proches les unes des autres.

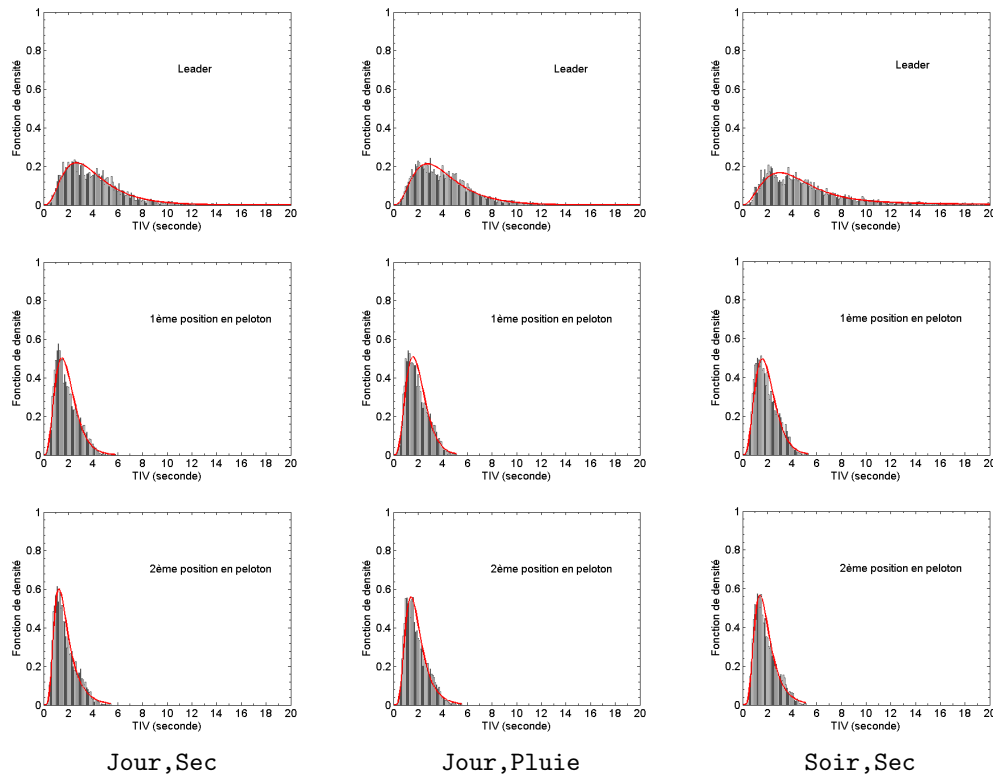


FIGURE 5.16 : Les distributions des TIV en peloton dans différentes périodes, empiriques et selon le modèle Γ -GQM

Avec la définition du peloton proposée, il y a des véhicules P_ℓ ayant un TIV court, et leur moyenne des TIV est inférieure à 5 s. On constate qu'en peloton, les distributions des TIV sont assez stables, autrement dit l'échelle mésoscopique est un bon niveau pour l'observation et la modélisation des TIV.

Pour expliquer la haute fréquence des TIV courts dans la distribution des véhicules libres P_ℓ , il semble qu'il y ait toujours création de véhicules libres dans le trafic. Cette création est faite par exemple dès lors qu'un véhicule a une vitesse réduite par rapport à celle du véhicule précédent. Dans ce cas, la condition \mathcal{B}_3 de la définition du « *suiveur* » n'est pas satisfaite.

Avec la définition du peloton proposée, on obtient évidemment l'équivalence des propriétés statistiques des TIV selon les différentes positions en peloton. La distribution des TIV montre que les comportements des conducteurs au sein d'un peloton sont généralement semblables et très différents de celui des véhicules libres P_ℓ . Le CV des TIV en peloton est entre 0,4 et 0,5 tandis que le CV des TIV des P_ℓ est supérieur à 0,5. Le CV sur la VR est particulièrement élevé dans la période du *soir*. La moyenne des TIV des P_ℓ est supérieure à 4 s tandis que celle des suiveurs est inférieure à 2 s. Cependant, la moyenne des TIV des véhicules P_2 est légèrement plus élevée que celle des autres suiveurs. Ce phénomène se produit probablement parce que les véhicules P_2 sont capable de prévoir davantage une possibilité de dépassement, ou bien parce que les pelotons de trois véhicules ou plus se forment par trafic dense avec moins de possibilité de dépassement.

La moyenne des TIV par temps de *pluie* est supérieure à celle par temps *sec* alors que le CV des TIV par temps de *pluie* est inférieur à celui par temps *sec*. La moyenne des TIV dans la période du *soir* est supérieure à celle dans la période du *jour* tandis que le CV des TIV dans la période du *soir* est inférieur à celui dans la période du *jour*. En comparant les impacts des deux conditions de météo (*pluie* ; *sec*), la *pluie* joue un rôle d'un facteur

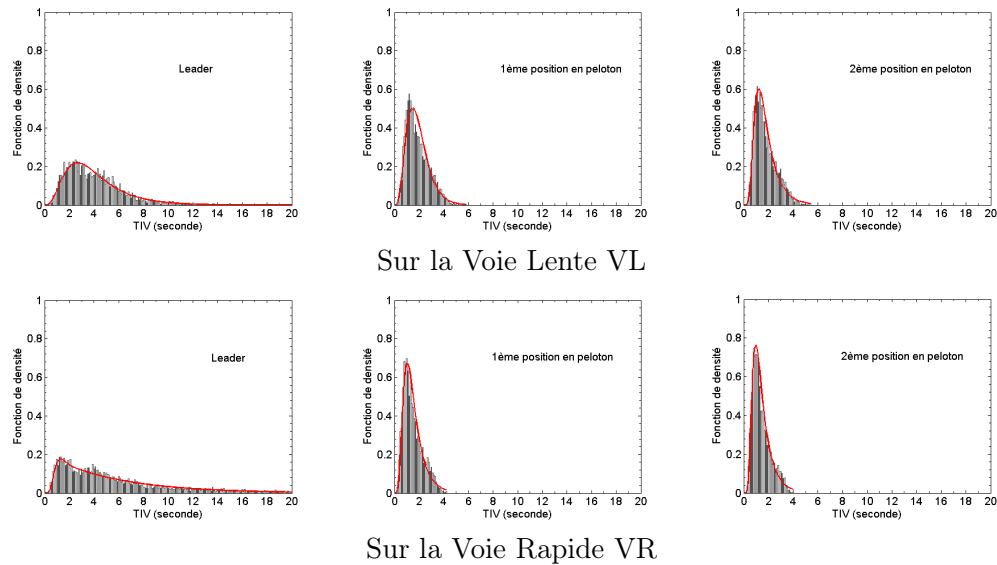


FIGURE 5.17 : Les distributions des TIV sur les deux voies (VL et VR), empiriques et selon le modèle Γ -GQM

de *gêne*. La moyenne des TIV augmente dans la période du *soir* probablement pour des raisons de sécurité - les différences de comportement selon la période de la journée sont un sujet de recherche qui n'est qu'effleuré dans cette thèse.

La moyenne des TIV sur la VR est également supérieure à celle sur la VL ceci étant probablement dû à une occupation inférieure de la VR. Au contraire du cas de la *pluie*, le CV des TIV sur la VR est supérieur à celui sur la VL indiquant un moins grand niveau de contrainte. Ce phénomène est obtenu quelque soit la période étudiée (*jour* ou *soir*). Dans ce cas, la VR semble jouer le rôle d'un facteur *de propulsion*. Les deux effets inverses de la VR et de la condition de *pluie* en tant que des facteurs exogènes affectant la distribution des TIV sont donc confirmés.

Le Tableau 5.1 représente l'ajustement de trois modèles Γ -GQM, DG et LNM $\ln \mathcal{N}_7$. Les résultats montrent que :

1. Le modèle LNM $\ln \mathcal{N}_7$ a une grande difficulté pour modéliser les TIV en peloton. Les symboles « * » signifient qu'aucune parmi les deux méthodes d'estimation MMVM et MMVL ne fournit un résultat numérique. Le calibrage utilisant le modèle LNM $\ln \mathcal{N}_7$ est rejeté pour tous les échantillons comportant des véhicules P_1, P_2 et P_3 .
2. Le modèle Γ -GQM n'est pas non plus un bon modèle pour les TIV en peloton. Le modèle Γ -GQM fournit de meilleurs ajustements pour les TIV des véhicules P_3 notamment des véhicules P_* .
3. Le modèle DG donne les meilleurs résultats pour les TIV en peloton. Dans 30 échantillons étudiés dans le Tableau 5.1, on ne peut pas rejeter 18 échantillons modélisés par le modèle DG. Les nombres d'échantillons correspondant pour les modèles Γ -GQM et LNM sont respectivement 6 et 3.

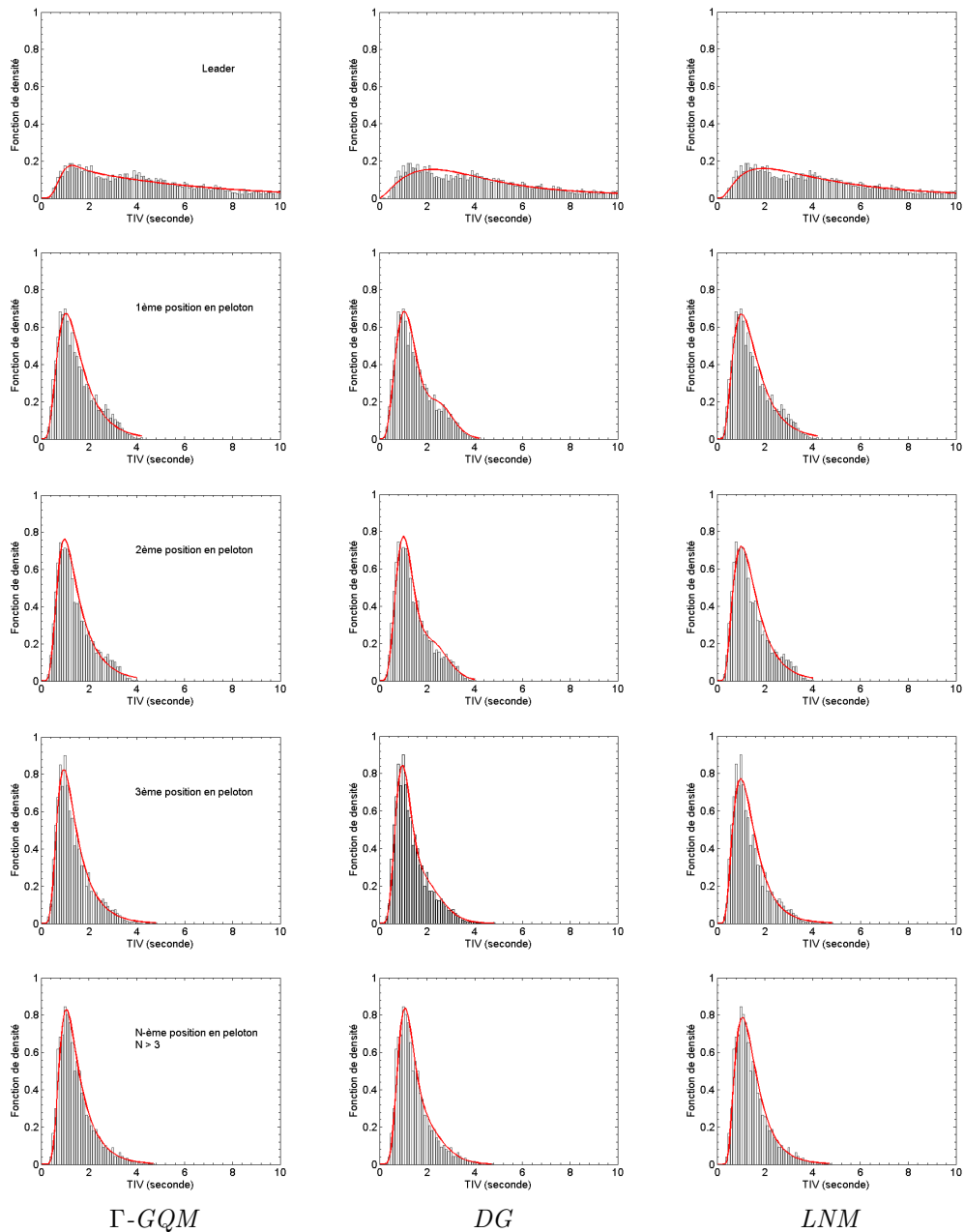


FIGURE 5.18 : Différents modèles calibrés et les distributions empiriques des TIV en peloton, VR, jour-sec

La configuration

Le Tableau 5.2 récapitule les résultats de la modélisation des TIV selon la configuration du trafic. Tout d'abord, on a observé très peu de cas où un PL suit un autre PL sur la VR. Il y a aussi un très faible nombre d'observations d'un VL suivant un PL sur la VR de la modalité **soir-sec**.

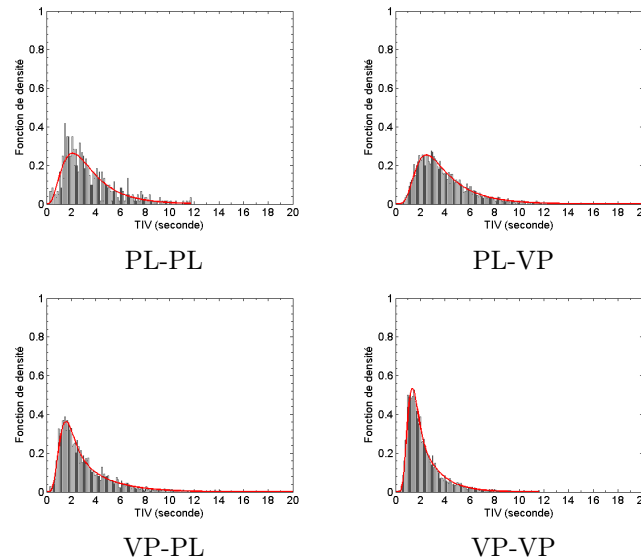


FIGURE 5.19 : Les distributions des TIV, empiriques et selon le modèle Γ -GQM, selon les configurations, modalité jour-pluie sur la VL

Au niveau des propriétés statistiques, on observe que la moyenne des TIV de la configuration (VP-VP) est la plus petite alors que la moyenne des TIV de la configuration (PL-VP) est la plus grande. Les véhicules VP grâce à leur taille courte et leur légèreté, sont les véhicules les plus dynamiques du trafic qui peuvent garder des TIV courts ou très courts et qui peuvent aussi distancer les véhicules PL derrière.

Sur la VL dans la période du jour, la moyenne des TIV de la configuration (PL-PL) est supérieure à celle de la configuration (VP-PL) quelque soit la condition de météo.

Sur la VL, la moyenne des TIV de la période de jour par temps de pluie est légèrement supérieure à celle de la période de jour par temps sec pour les trois configurations PL-PL, PL-VP et VP-PL. Sur la VL, la moyenne des TIV de la période de jour par temps de pluie est légèrement inférieure à celle de la période de jour par temps sec pour la configuration VP-VP, ceci pouvant peut-être être expliqué par un rabattement vers la VL de VP qui par temps sec circuleraient sur la VR. Pour toutes les configurations, la moyenne et le CV des TIV de la modalité **soir-sec** sont nettement supérieurs à ceux de la modalité **jour-sec**.

La moyenne et le CV des TIV sur la VR sont toujours supérieurs à ceux sur la VL quelque soit la période considérée. Dans la période de jour, le CV sur la VR est proche de ou plus grand que la valeur unité.

Le modèle DG fournit de meilleurs ajustements pour les distributions des TIV de la configuration PL-PL sur la VL. Le modèle Γ -GQM est meilleur que le modèle DG dans la plupart des cas étudiés. Cependant, le modèle Γ -GQM est rejeté dans les quatre configurations de la modalité **soir-sec**.

La Figure 5.19 représente la distribution des TIV selon les quatre catégories de poursuite utilisant le modèle Γ -GQM dans la période jour-pluie sur la VL. La Figure 5.20 montre l'avantage du modèle DG en modélisant les TIV autour du *mode* de la distribution des TIV. Ce modèle s'adapte aussi bien à la queue exponentielle de la distribution des TIV.

Position	P_ℓ	Véhicule P_N			
		$N = 1$	$N = 2$	$N = 3$, $N > 3$
Jour-Sec					
Voie Lente					
Moyenne	4.1013	1.9126	1.8054	1.7611	1.7428
CV	0.6013	0.4700	0.4817	0.4855	0.4852
P-value Γ -GQM	0.0051	0.0000	0.0001	0.0019	0.1161
P-value LNM	0.0000	*	*	*	0.0171
P-value DG	0.0290	0.0000	0.4033	0.0000	0.0000
Voie Rapide					
Moyenne	6.1489	1.5346	1.4683	1.3989	1.4847
CV	0.9508	0.5012	0.4908	0.4909	0.4617
P-value Γ -GQM	0.0001	0.0002	0.0044	0.0895	0.2655
P-value LNM	0.0137	0.0004	0.0031	0.0002	0.3727
P-value DG	0.0010	0.0880	0.1358	0.4030	0.2189
Jour-Pluie					
Voie Lente					
Moyenne	4.1683	1.9687	1.9178	1.8913	1.8853
CV	0.5805	0.4379	0.4560	0.4646	0.4520
P-value Γ -GQM	0.0114	0.0000	0.0001	0.0320	0.0799
P-value LNM	0.0000	*	*	*	0.0218
P-value DG	0.1264	0.4644	0.3780	0.0000	0.2902
Voie Rapide					
Moyenne	6.3435	1.6552	1.6429	1.6509	1.8543
CV	0.9727	0.4734	0.4766	0.4749	0.4613
P-value Γ -GQM	0.0000	0.0005	0.0114	0.0807	0.3903
P-value LNM	0.0978	*	0.0000	0.0000	*
P-value DG	0.0033	0.4109	0.0000	0.0001	0.7606
Soir-Sec					
Voie Lente					
Moyenne	5.9222	1.9840	1.9205	1.8817	1.8405
CV	0.9212	0.4453	0.4584	0.4556	0.4606
P-value Γ -GQM	0.0004	0.0001	0.0002	0.0000	0.0058
P-value LNM	0.0080	*	0.0014	0.0001	0.0258
P-value DG	0.0003	0.5515	0.6343	0.7843	0.5180
Voie Rapide					
Moyenne	9.6937	1.6014	1.5494	1.5232	1.6783
CV	2.1652	0.4798	0.4787	0.4933	0.4781
P-value Γ -GQM	0.0000	0.0000	0.0150	0.0008	0.0428
P-value LNM	-	*	*	0.0007	0.0631
P-value DG	-	0.3884	0.0000	0.4612	0.2888

TABLE 5.1 : La moyenne, le CV des TIV selon le rang dans le peloton et l'ajustement des modèles calibrés

Modèles	La configuration			
	PL-PL	PL-VP	VP-PL	VP-VP
Jour-Sec				
	Voie Lente			
Moyenne	3.4484	3.8847	2.9730	2.4542
CV	0.8607	0.6026	0.7220	0.7123
P-value Γ -GQM	0.0009	0.2866	0.5038	0.8162
P-value DG	0.0113	0.5194	0.2622	0.1404
	Voie Rapide			
Moyenne	*	5.1174	3.5884	2.9043
CV	*	1.4209	1.5847	1.2393
P-value Γ -GQM	*	0.1853	0.0306	0.0020
P-value DG	*	0.0887	0.0296	0.0002
Jour-Pluie				
	Voie Lente			
Moyenne	3.4694	4.0128	3.1420	2.3209
CV	0.6125	0.6137	0.9063	0.6504
P-value Γ -GQM	0.4464	0.7967	0.5587	0.5740
P-value DG	0.9125	0.4991	0.1413	0.1385
	Voie Rapide			
Moyenne	*	5.7351	4.0596	2.7035
CV	*	1.2272	1.3495	0.9873
P-value Γ -GQM	*	0.8157	0.7668	0.3390
P-value DG	*	0.7911	0.5361	0.0564
Soir-Sec				
	Voie Lente			
Moyenne	4.1858	5.3830	4.4565	3.7131
CV	0.8538	0.9489	1.0968	1.1414
P-value Γ -GQM	0.0119	0.0098	0.0000	0.0002
P-value DG	0.7965	0.0018	0.0001	0.0002
	Voie Rapide			
Moyenne	*	7.7633	*	4.8084
CV	*	1.6577	*	2.5829
P-value Γ -GQM	*	0.2118	*	0.0000
P-value DG	*	0.2131	*	0.0000

TABLE 5.2 : La moyenne, le CV des TIV et l'ajustement des modèles calibrés selon la configuration

Modèles	Le TIV précédent T_p^k					
	0 - 1 s	1-2 s	2-3 s	3-4 s	4-5 s	5-6 s
Jour-Sec						
	Voie Lente					
Moyenne	2.5445	2.7295	2.8172	3.0054	3.0151	3.1574
CV	0.7638	0.7615	0.7121	0.7235	0.7110	0.7897
P-value Γ -GQM	0.6194	0.6193	0.3532	0.9760	0.5233	0.5211
P-value DG	0.0586	0.0888	0.0605	0.3428	0.2614	0.0651
	Voie Rapide					
Moyenne	2.9022	2.8432	3.4176	4.1783	4.6852	4.6374
CV	1.3851	1.2751	1.2442	1.2691	1.2996	1.3043
P-value Γ -GQM	0.0001	0.0017	0.0098	0.0047	0.0063	0.0027
P-value DG	0.0000	0.0001	0.0005	0.0074	0.0027	0.0063
Jour-Pluie						
	Voie Lente					
Moyenne	2.6557	2.6530	3.0806	3.1401	3.2986	3.4851
CV	0.8028	0.7311	0.8493	1.0438	0.8257	0.9603
P-value Γ -GQM	0.7181	0.3993	0.0438	0.0514	0.1589	0.0426
P-value DG	0.2258	0.1972	0.0182	0.0002	0.0351	0.0043
	Voie Rapide					
Moyenne	3.3724	3.4440	3.6013	4.3270	4.5984	4.8499
CV	1.6208	1.3769	1.3598	1.4230	1.4846	1.3267
P-value Γ -GQM	0.0004	0.0020	0.0152	0.0974	0.0218	0.0547
P-value DG	0.0004	0.0001	0.0033	0.0501	0.0600	0.0255
Soir-Sec						
	Voie Lente					
Moyenne	2.8790	3.0162	3.4675	3.8794	4.2428	4.3390
CV	0.9161	0.9860	0.9797	1.0207	1.0413	0.9952
P-value Γ -GQM	0.0685	0.0012	0.0023	0.0066	0.0430	0.0663
P-value DG	0.0116	0.0003	0.0014	0.0017	0.0045	0.0092
	Voie Rapide					
Moyenne	3.3902	3.3243	4.2798	*	*	*
CV	2.1275	2.0693	2.1742	*	*	*
P-value Γ -GQM	0.0000	0.0002	0.0000	*	*	*
P-value DG	0.0000	0.0005	0.0000	*	*	*

TABLE 5.3 : Les distributions des TIV selon le TIV du véhicule précédent

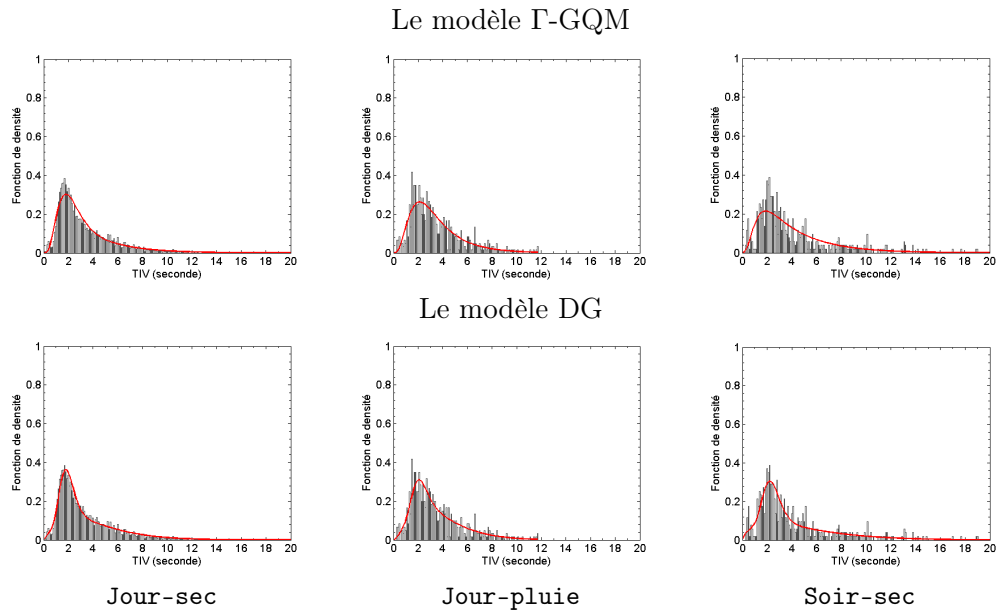


FIGURE 5.20 : La supériorité des ajustements du modèle DG sur ceux du modèle Γ -GQM pour la configuration (PL-PL) sur la VL

5.1.4 Distribution des TIV selon le TIV du véhicule précédent

Les TIV sont classés dans les classes \mathcal{C}_T des TIV h_p des véhicules précédents. L'incrément des classes choisi est d'une seconde. La classes k comporte donc les TIV h_i tels que $T_p^k := \{h_i \mid k - 1 < h_p \leq k\}$.

Pour les T_p^k avec $k \geq 9$ s, on observe que la distribution des TIV ne peut pas être calibrée par le modèle Γ -GQM ou par les autres modèles habituellement utilisés. Le Tableau 5.3 représente les six premières classes T_p^k , avec $k = \overline{1,6}$.

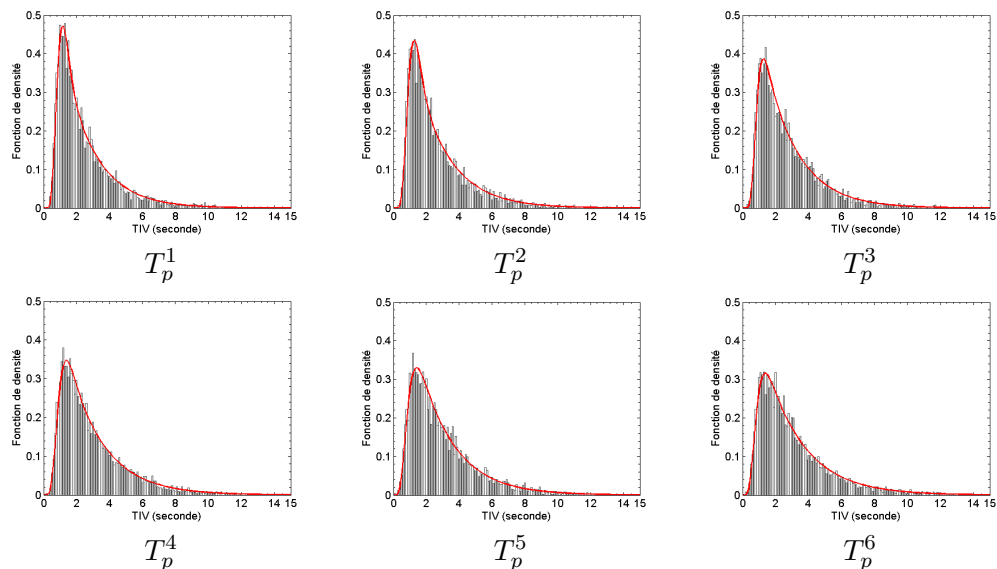


FIGURE 5.21 : L'ajustement du modèle Γ -GQM et les distributions des TIV sachant le TIV du véhicule précédent, jour-sec, VL

Tout d'abord, la moyenne des TIV augmente en fonction de la classe T_p^k pour k supérieur à 1. Dans la période (jour-sec), les CV sont assez équivalents dans les 6 classes. En général, les moyennes et les CV des TIV dans les périodes (jour-pluie) et (soir-sec)

sont supérieurs à ceux de la période (jour-sec). La moyenne et le CV sur la VR sont plus grands que ceux sur la VL.

Le modèle Γ -GQM et le modèle DG modélisent bien les distributions des six échantillons dans la période de jour sur la VL. Il n'y a qu'un seul échantillon bien calibré par le modèle Γ -GQM et deux échantillons bien calibrés par le modèle DG dans la période (jour-pluie) sur la VR. Dans les cas où les modèles ne sont pas rejetés au niveau de signification de 5 % au sens du test K-S, le modèle Γ -GQM est meilleur que le modèle DG.

5.2 L'échelle microscopique

La modélisation des TIV à l'échelle microscopique fournit aussi des connaissances supplémentaires sur le comportement des conducteurs. Dans cette section, deux questions sont posées :

1. Comment la vitesse instantanée du véhicule participe à la distribution des TIV ?
2. Sachant que le quotient de TIV h et de la vitesse instantanée v est un indicateur de sécurité routière, si on le considère comme une variable aléatoire choisie par les conducteurs, quelle est la distribution de cette variable ?

Ces questions font intervenir la variable vitesse instantanée du véhicule qui constitue avec le TIV deux des trois informations individuelles de base du trafic - la troisième information étant la longueur des véhicules.

5.2.1 Distribution des TIV et de la vitesse instantanée

Très peu d'études dans la littérature couplent les deux informations de TIV et de vitesse instantanée pour étudier l'interaction entre ces deux variables. Breiman (1962) [54] a supposé en trafic dense que la distribution des vitesses est indépendante de celle des TIV. Aly (1989) [65] a proposé des distributions de TIV différentes selon la vitesse de véhicule. Les deux principaux ensembles de véhicules discriminés, d'après Aly, sont les véhicules *en congestion* ayant une vitesse inférieure à 25 km/h et les véhicules *hors congestion* se déplaçant à une vitesse supérieure ou égale à 25 km/h.

Dans notre étude, les vitesses instantanées des véhicules sont groupées dans des échantillons de 0 à plus de 40 m/s avec un incrément de 2 m/s. En principe, il y donc 21 échantillons de vitesse instantanée, mais en réalité, le nombre d'échantillons obtenus dépend de la période observée. Par exemple, le nombre d'échantillons n_e du Site N°1 est récapitulé dans le Tableau 5.4 suivant :

Période	Voie Lente			Voie Rapide		
	v_{\min} m/s	v_{\max} m/s	n_e	v_{\min} m/s	v_{\max} m/s	n_e
jour-sec	(0-2)	> 40	21	(0-2)	> 40	21
jour-pluie	(6-8)	(36-38)	16	(10-12)	> 40	16
soir-sec	(2-4)	> 40	20	(0-2)	> 40	21
soir-pluie	(12-14)	(36-38)	13	(16-18)	(36-38)	11

TABLE 5.4 : Le nombre d'échantillons des TIV groupés selon la vitesse instantanée v

La Figure 5.13 dans la section précédente représente les moyennes et les écarts-types des TIV dans les différents échantillons primitifs sur les deux voies de la RN118. Ces résultats fournissent de premières connaissances sur les choix généraux des conducteurs. Cependant,

les durées de mesure des échantillons primitifs ne sont pas les mêmes. Dans cette section, les échantillons sont échantillonnés par la méthode de groupement qui distingue les modalités comme (**jour-sec**), (**jour-pluie**), etc., et ensuite groupe les TIV en fonction de la vitesse instantanée. La Figure 5.22 illustre les comparaisons des moyennes μ_H et des écart-types σ_H des TIV selon les différentes modalités.

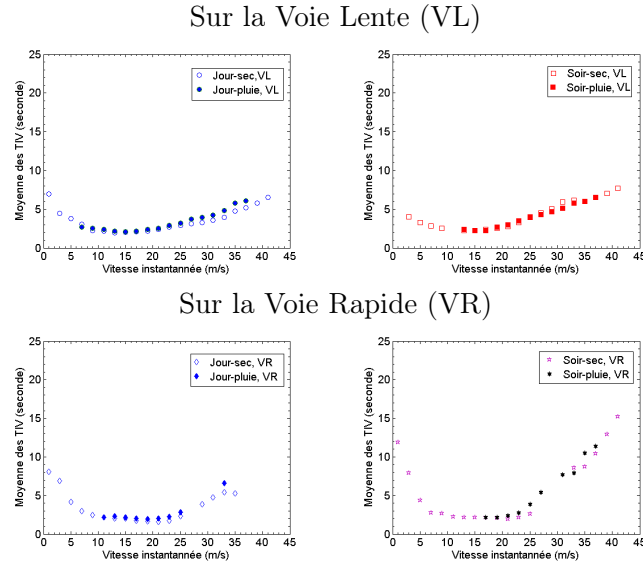


FIGURE 5.22 : Moyennes des TIV dans les échantillons groupés par vitesse instantanée v

La moyenne μ_H sur la VL par temps **sec** est généralement plus faible que celle sur la VL par temps de **pluie**. La moyenne μ_H sur la VL est minimale dans les échantillons de vitesse comprise entre 12 et 20 m/s. La moyenne μ_H sur la VR est minimale dans les échantillons de vitesse comprise entre 10 et 25 m/s.

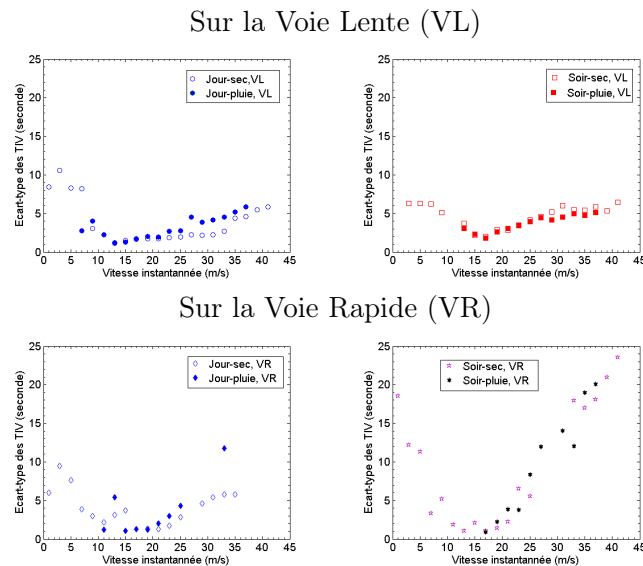


FIGURE 5.23 : Écart-types des TIV dans les échantillons groupés par vitesse instantanée v

L'écart-type des TIV commence par diminuer en fonction de la vitesse instantanée v jusqu'à l'échantillon de vitesse v entre 15 et 20 m/s avant d'augmenter (cf. Figure 5.23).

La moyenne μ_H et l'écart-type σ_H varient de la même façon et par conséquent le CV des TIV sur la VL est très stable et proche de la valeur unité pour les échantillons ayant une vitesse v supérieure à 20 m/s.

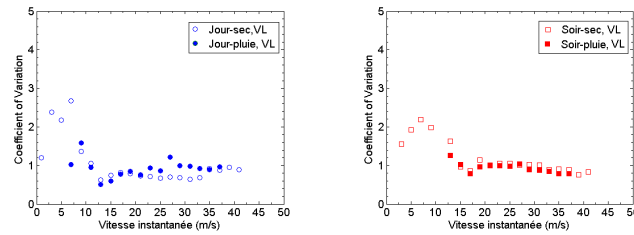


FIGURE 5.24 : CV des TIV dans les échantillons groupés par vitesse instantanée v

Le test du χ^2 est plus sévère que le test K-S et induit le rejet dans beaucoup d'échantillons. L'ajustement des modèles est le meilleur dans les modalités (jour-sec) et (jour-pluie) sur la VL et (soir-pluie) sur la VR. Dans ces trois modalités, les modèles Γ -GQM et DG fournissent de bons ajustements. L'ajustement sur la VR est beaucoup plus difficile.

Pour le modèle DG, le paramètre α_1 est inférieur ou égal à 1 tandis que le paramètre α_2 est supérieur à 1. Ces deux paramètres d'échelle sont assez constants dans l'intervalle de vitesse de 15 à 30 m/s avec α_1 est proche de 1 dans la modalité jour-sec. La Figure 5.25 montre que les paramètres α_1 et α_2 sont plus grands dans les périodes de jour que dans les périodes de soir.

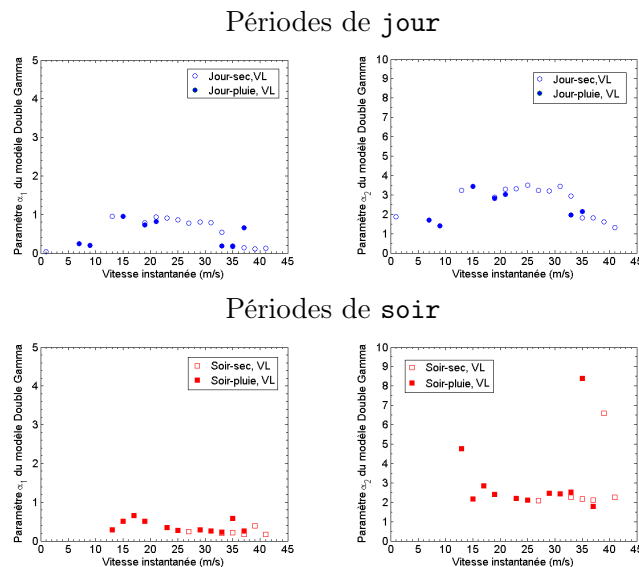
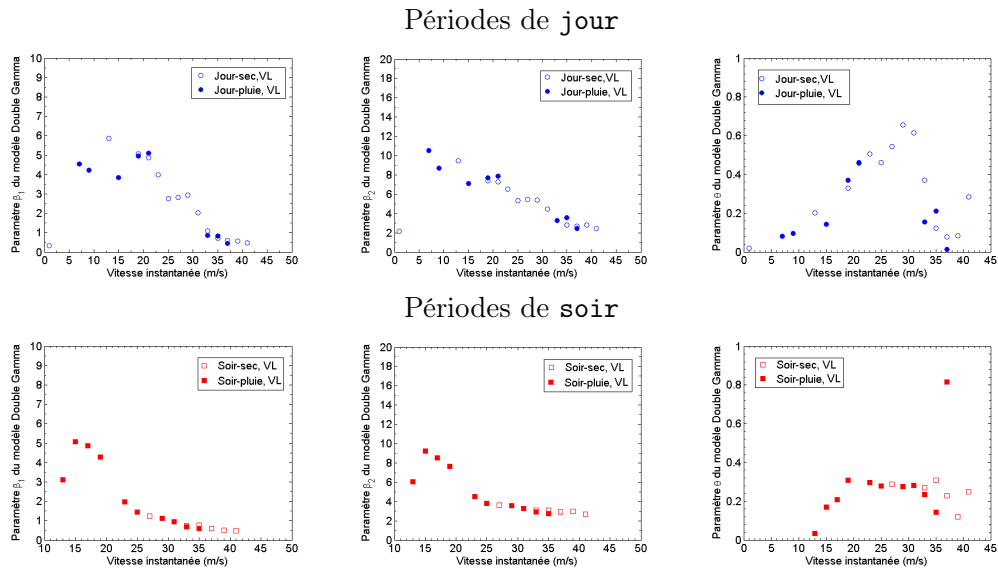


FIGURE 5.25 : Les paramètres α_1 et α_2 du modèle DG

Les paramètres β_1 et β_2 diminuent en fonction de la vitesse instantanée v . De l'échantillon de vitesse 15 m/s à celui de vitesse 30 m/s, la diminution peut être considérée comme linéaire (cf. Figure 5.26).

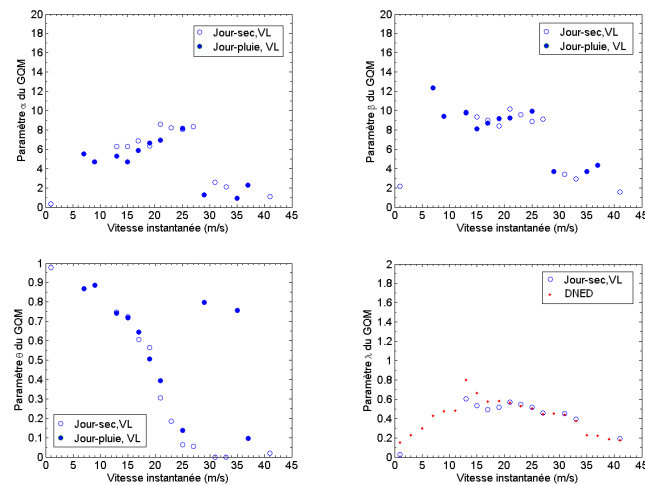
Le paramètre θ du modèle DG varie différemment selon la période étudiée. Entre les échantillons de vitesse de 15 à 30 m/s, θ augmente linéairement en fonction de la vitesse v pour la modalité jour-sec. Dans les périodes de soir, θ augmente à partir d'une valeur proche de zéro dans l'échantillon de vitesse de 13 m/s jusqu'à une valeur d'environ 0,3 dans l'échantillon de vitesse de 20 m/s ; θ est ensuite constant pour les vitesses entre 20 et 30 m/s (cf. Figure 5.26).

FIGURE 5.26 : Les paramètres β_1 et β_2 et θ du modèle DG

Pour le modèle Γ -GQM, les variations des paramètres α et β dans les périodes de jour changent d'allure lorsque la vitesse v atteint 30 m/s. Le paramètre α de la modalité jour-sec augmente, alors que le paramètre θ diminue, en fonction de la vitesse v pour les vitesses inférieures à 30 m/s. Les valeurs estimées de λ du modèle Γ -GQM sont proches de celles de λ du modèle \mathcal{E}_τ . Le paramètre λ est compris entre 0,4 et 0,6 dans les échantillons de vitesse v comprise entre 15 et 30 m/s.

Les variations des paramètres du modèle Γ -GQM en fonction de la vitesse instantanée v sont semblables à celles en fonction de la vitesse moyenne harmonique \mathcal{V}_6 (cf. Figure 4.6, Chapitre 4). On constate que le paramètre θ a une forte relation avec la vitesse des véhicules tant que celle-ci reste inférieure à 30 m/s.

Dans la modalité (soir-pluie) et pour les échantillons de vitesse v comprise entre 15 et 30 m/s, tous les paramètres α , β , θ et λ du modèle Γ -GQM diminuent en fonction de la vitesse v ; en outre le paramètre θ reste supérieur à 0,5 (cf. Figure 5.28).

FIGURE 5.27 : Les paramètres du modèle Γ -GQM dans la période du jour, VL

Discussion

L'intervalle de 15 à 30 m/s est équivalent à celui de 54 à 108 km/h. Il est à noter que la vitesse limite légale sur la RN118 est de 90 km/h sur la majorité des sections et de 110 km/h sur certaines sections. Dans l'intervalle indiqué, le comportement de conduite représenté par la distribution des TIV en fonction de la vitesse instantanée v peut être bien modélisé par le modèle Γ -GQM et le DG dans la période du jour.

Toujours dans cet intervalle de vitesse, les tendances des variations des paramètres estimés du modèle DG sont nettes notamment pour la modalité (jour-sec) sur la VL. Si l'on considère que la conduite dans cette modalité jour-sec sur la VL est un comportement *standard*, ce comportement est modélisable avec un très bon ajustement. L'impact des facteurs exogènes comme la pluie pourrait être déduit par décalage des valeurs des paramètres du comportement standard en fonction de l'intensité du facteur. Comme l'information de débit n'est pas utilisée ici aux niveaux microscopique et mésoscopique, le paramètre λ du modèle Γ -GQM ne représente plus directement la demande de trafic. Dans ce cas là, le modèle Γ -GQM n'a donc plus l'avantage par rapport au modèle DG de proposer une interprétation en terme de débit du phénomène. En conséquence, dans le cas présent, le modèle DG est préféré pour modéliser le comportement standard de conduite.

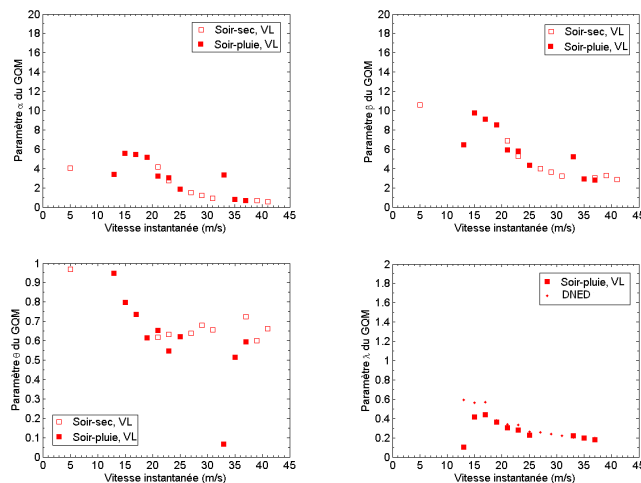


FIGURE 5.28 : Les paramètres du modèle Γ -GQM dans la période du soir, VL

5.2.2 La variable RHIS

La variable RHIS (Ratio of time Headway and Instantaneous Speed) associant le TIV et la vitesse a été l'objet d'une étude statistique initiale [109]. Cette variable joue le rôle d'un indicateur de sécurité routière. En effet, elle est tout d'abord une combinaison de deux informations individuelles essentielles du trafic. Par ailleurs, le rapport entre elles donne naissance à une grandeur dont l'unité est l'inverse de l'accélération. La RHIS est considéré comme une variable aléatoire qui se distribue différemment selon la voie et selon la condition météorologique.

On sait qu'un TIV faible, associé à une vitesse élevée - ce qui donne le rapport RHIS particulièrement faible - est dangereux [91], car si le premier véhicule venait à s'arrêter brutalement (comme arrêté par un mur de brique), le second véhicule, avec un RHIS trop faible, ne pourrait éviter la collision. Un calcul fondé sur la distance d'arrêt en fonction de la vitesse montre qu'il y a collision si la RHIS est inférieure à la moitié de l'inverse de la décélération maximale.

L'adjonction de la vitesse instantanée au dénominateur en gardant le TIV au numérateur n'empêche pas d'essayer d'utiliser les modèles de TIV étudiés dans les chapitres précédents. Le travail dans [109], est à notre connaissance, le premier sur cette variable. Pour cette raison, cette section est consacrée à calibrer la variable RHIS en ajustant les modèles Γ -GQM et DG aux 25 échantillons de l'*ensemble test* utilisés précédemment dans le Chapitre 2. Les modèles simples ne sont pas bons car le modèle Γ est très loin être compatible avec la distribution de RHIS [109], tandis que le modèle LNM $\ln \mathcal{N}_\tau$ ne donne aucun résultat numérique.

Le TIV est mesuré en secondes s et la vitesse instantanée est mesurée en mètre par seconde (m/s) ce qui donne l'unité de RHIS de s^2/m . Les données sont groupées dans les classes de largeur $0,01 s^2/m$ au lieu de $0,0072$ dans [109].

La Figure 5.29 représente l'ajustement du modèle Γ -GQM pour les échantillons de l'ensemble test.

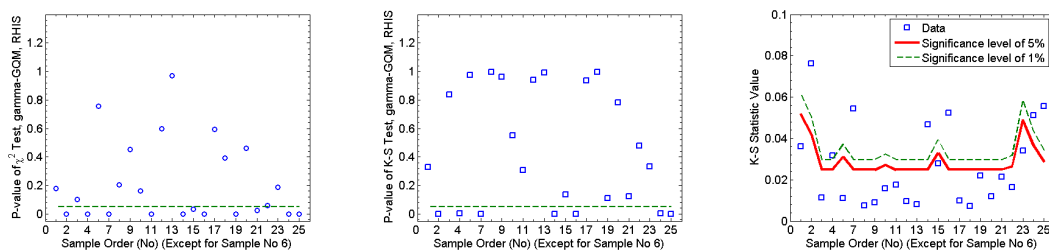


FIGURE 5.29 : P-valeur du test K-S et du test χ^2 , modèle Γ -GQM de la variable RHIS

Les résultats montrent que :

- Le modèle Γ -GQM fournit de très bons ajustements au sens du test du χ^2 et du test K–S pour la plupart d'échantillons examinés. Par contre, le modèle DG est rejeté dans la plupart des cas.
- Le calibrage est mauvais pour six échantillons N°2, N°7, N°14, N°16, N°24 et N°25. Il est constaté que le modèle Γ -GQM représente difficilement les distributions de la variable RHIS sur les données de l'autoroute A6 et sur la VR de la RN118.
- Le temps d'estimation des paramètres du modèle Γ -GQM de la variable RHIS est plus long que celui de la variable TIV. Le paramètres d'échelle α et λ donnent de la flexibilité au modèle Γ -GQM.

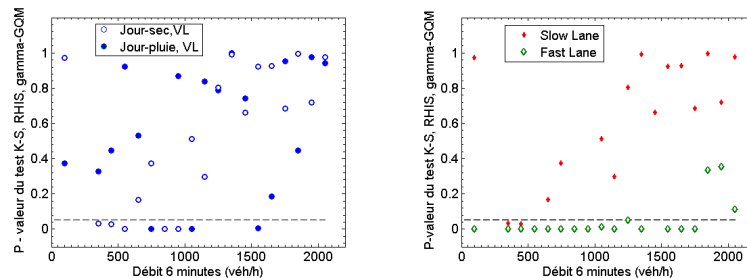


FIGURE 5.30 : Les P-valeurs de test K-S du modèle Γ -GQM pour la variable RHIS en fonction du débit Q_6

Sur la Figure 5.30 sont représentées les P-valeurs des calibrages du modèle Γ -GQM selon la voie, la période, la condition de météo et le débit Q_6 . Sur la VL, les ajustements sont généralement bons tant pour la modalité jour-sec que pour la modalité jour-pluie. Sur la VR, le modèle est rejeté à niveau de signification de 5 % dans 15 échantillons sur

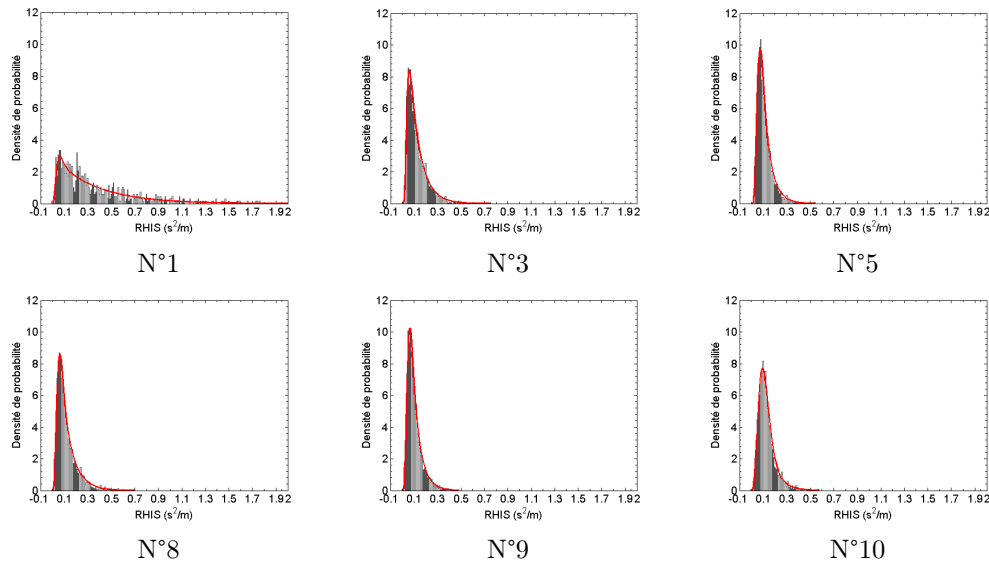


FIGURE 5.31 : Les densités de probabilité du modèle Γ -GQM pour la variable RHIS de l'ensemble test

18.

La Figure 5.32 illustre les variations des paramètres du modèle Γ -GQM pour la variable RHIS en fonction du débit Q_6 et selon la condition météo sur la VL. On obtient que :

- Les échelles de valeurs des paramètres α et λ servant à modéliser la variable RHIS sont complètement différentes de celles utilisées pour les TIV - ceci étant dû à l'introduction de la vitesse au dénominateur.
- Les tendances des paramètres en fonction du débit sont généralement semblables à celles constatées pour la variable TIV, particulièrement les paramètres β , λ . Le paramètre θ dans la période jour-sec varie aussi en deux phases.
- Le paramètre λ est toujours proportionnel au débit Q_6 quelque soit la condition de météo.

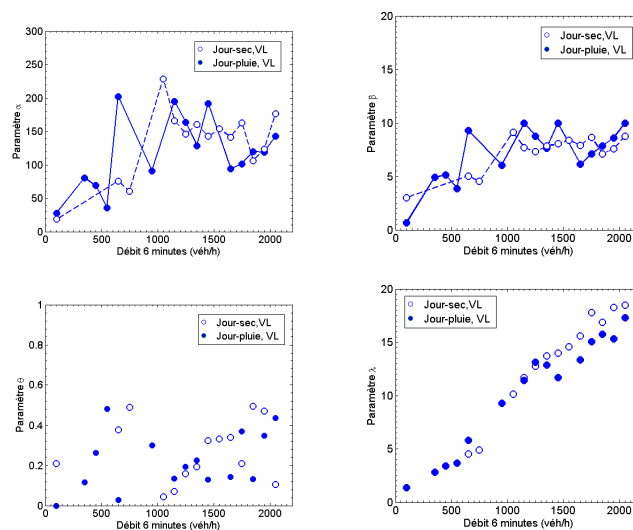


FIGURE 5.32 : Les paramètres du modèle Γ -GQM pour la variable RHIS en fonction du débit Q_6 selon les conditions de météo

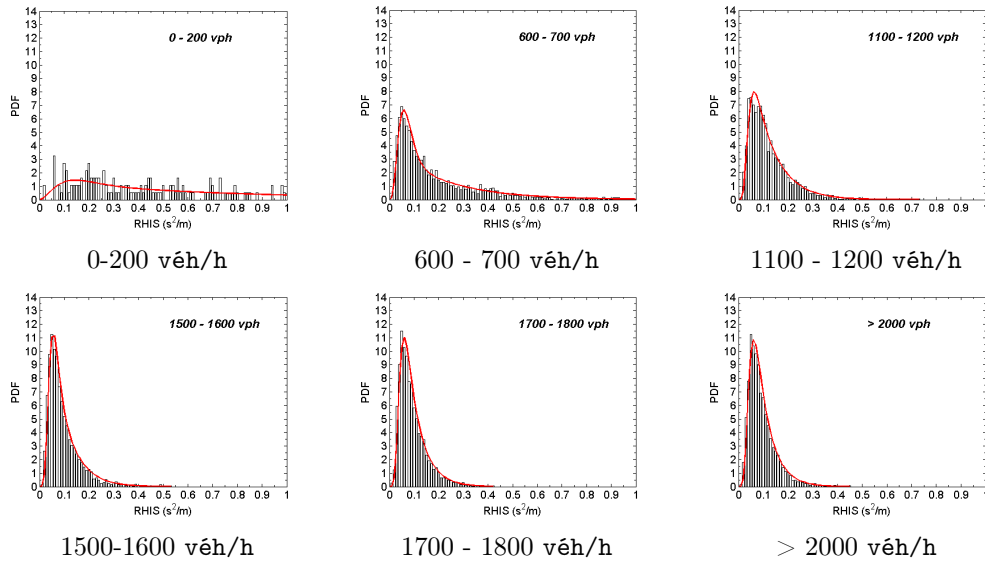


FIGURE 5.33 : La densité de probabilité du modèle Γ -GQM pour la variable RHIS, sur la VL de la RN118

5.3 Conclusion du chapitre

Tout d'abord, une analyse statistique est réalisée afin de trouver une définition convenable du peloton. La définition proposée est une combinaison des seuils de TIV, de distance intervéhiculaire (DIV) et de vitesse relative. À l'échelle mésoscopique, la modélisation des TIV en fonction de la position des véhicules en peloton, en fonction de la configuration de poursuite et en fonction des TIV des véhicules précédents sont réalisées en utilisant les modèles Γ -GQM, DG et éventuellement le modèle LNM. À l'échelle microscopique, la vitesse instantanée des véhicules est prise en compte. La distribution des TIV en fonction des vitesses instantanées groupées et la distribution de la variable RHIS - le ratio entre le TIV et la vitesse instantanée - sont modélisées. Le modèle Γ -GQM montre sa performance et sa flexibilité dans beaucoup de cas.

Chapitre 6

Validation

Dans l'ingénierie du trafic, ce sont les informations macroscopiques (débit, taux d'occupation) qui sont disponibles en pratique. Ce chapitre est consacré particulièrement à la validation à cette échelle. Il s'agit de valider un modèle et le calibrage réalisé sur des échantillons de TIV de différentes origines ayant des caractéristiques macroscopiques similaires. La validation des calibrages en présence de facteurs exogènes \mathcal{F}_1 et des autres facteurs $\mathcal{F}_3, \mathcal{F}_4$ correspondant à des niveaux d'agrégation plus bas n'est pas traitée ici. Dans le cadre de ce chapitre, nous examinons l'applicabilité du modèle Γ -GQM dans le cas du *comportement standard*.

6.1 Retour sur les buts de la modélisation probabiliste

Tant l'analyse statistique de la distribution des TIV que le calibrage des modèles probabilistes permettent d'atteindre trois objectifs :

1. Un premier objectif, *technique*, consiste à résumer le fichier de mesures en quelques paramètres : les moments, les principaux quantiles (pour ce qui concerne la distribution statistique), ou bien les paramètres des modèles probabilistes.
2. Le deuxième objectif consiste à éclaircir l'évolution de la distribution des TIV lorsque le *contexte* change. Ce *contexte* est caractérisé soit au niveau macroscopique (par le débit agrégé, la concentration, etc.), soit au niveau mésoscopique (par la configuration de poursuite, le peloton, etc.), soit au niveau microscopique (par la vitesse instantanée, etc.). Il est aussi caractérisé par les facteurs exogènes (condition météorologique, visibilité, état de la chaussée, le réseau, le rang de la voie, etc.). Les deux approches statistique et probabiliste sont *a priori* complémentaires, car les variations de la distribution en fonction des *contextes* entraînent la variation de certaines statistiques ainsi que celle de certains paramètres des modèles probabilistes. Cela amène des interprétations et donc une meilleure compréhension des phénomènes sous-jacents.
3. Au point de vue *opérationnel*, les distributions ou les modèles de TIV servent à établir des bases de *comparaison* et de *prévision* pour apprécier les TIV sur des sites proches par leur *contexte*.

Les deux premières remarques ont donné lieu à des développements dans les chapitres précédents tandis que la troisième remarque est déclinée dans ce chapitre pour les deux objectifs suivants :

- Lorsqu'il y a une seule caractéristique différenciant les *contextes* des deux sites (comme par exemple l'existence d'une limitation de vitesse sur l'un des sites), l'impact de cette caractéristique pourrait être *chiffrée* par l'évolution des statistiques de

la distribution ou par le changement de valeur des paramètres des modèles probabilistes, \mathcal{R}_1 .

- Lorsque les deux sites sont similaires et qu’il n’y a pas de mesure individuelle du trafic sur l’un des sites, les modèles probabilistes ou statistiques y remédient en générant des données fictives. La génération des données simulées se base donc sur la distribution statistique ou du modèle probabiliste du site de *référence*, \mathcal{R}_2 .

Ce dernier objectif de *prévision-comparaison* demande pourtant une validation des modèles tant statistique que probabiliste, ou encore un guide d’usage indiquant les marges d’erreurs. Les critères pour que deux échantillons soient jugés *similaires* par leur *contexte* doivent être précisés.

À l’échelle macroscopique, le débit Q est confirmé comme le premier facteur dont il faut tenir compte. Étant donné que le débit Q n’est autre que l’inverse de la moyenne des TIV, la connaissance du débit est nécessaire à la connaissance des TIV. Imprudemment peut-être ou à cause du manque de données, les études du TIV dans la littérature supposent explicitement ou implicitement que cette information de la moyenne des TIV est suffisante pour en déduire toute la distribution des TIV.

Cependant, on a montré dans les chapitres précédents le phénomène de *multi-distribution* de TIV et les effets des facteurs exogènes \mathcal{F}_1 sur la distribution des TIV du comportement standard au même niveau de débit Q . La moyenne des TIV semble ne pas être suffisante. En considérant aussi les remarques précédentes $\mathcal{R}_1, \mathcal{R}_2$, il vient deux premières questions (préliminaires à la phase de **Validation** de ce chapitre) :

1. Est-ce que les similitudes entre les statistiques générales - telles que, (en sus de la moyenne des TIV) le CV, le Mode, etc. - des deux échantillons suffisent pour que les deux échantillons soient *similaires* ?
2. Faut-il prendre en compte la similitude d’autres éléments de la distribution des TIV des deux échantillons (comme les *quantiles* de certains TIV « *importants* ») avant d’appliquer au second échantillon un modèle probabiliste calibré sur le premier ?

Une réponse « *Oui* » à cette dernière question légitime une validation restreinte à des points spécifiques de la distribution. L’approche issue de ce genre de questions est appelée l’approche de *statistique partielle*. Les TIV « *importants* » sont ceux qui sont souvent utilisés comme seuils en pratique comme les TIV inférieurs ou égaux à 1 s, à 2 s, à 4 s, à 7 s, etc.

En général, on considère qu’un modèle (par défaut un modèle probabiliste) est validé lorsque son application sans nouveau calibrage sur un nouveau jeu de données (le deuxième échantillon qui n’a pas servi au calibrage), n’entraîne pas une dégradation trop sensible des performances du modèle. On constate qu’une validation *partielle* est très pratique mais comporte des risques d’erreurs. En effet, les critères de similitude utilisés dans cette approche ne sont autre que des descriptions générales de la distribution des TIV. Même la deuxième question précitée n’a pour but que de contraindre la fonction de répartition modélisée des TIV à passer statistiquement près des valeurs empiriques des *quantiles* des TIV « *importants* ». Les schémas de la Figure 6.1 illustrent plus clairement cette approche de validation *partielle*.

6.2 Méthodes de validation

Le modèle probabiliste est utilisé pour reconstituer ou pour prédire les TIV et leur distribution. La question de validation est de savoir si un modèle calibré pour un échantillon est applicable à un autre échantillon relevant du même *contexte*, ou à un échantillon *similaire*. Cette section discute différentes approches de validation et les résultats des niveaux

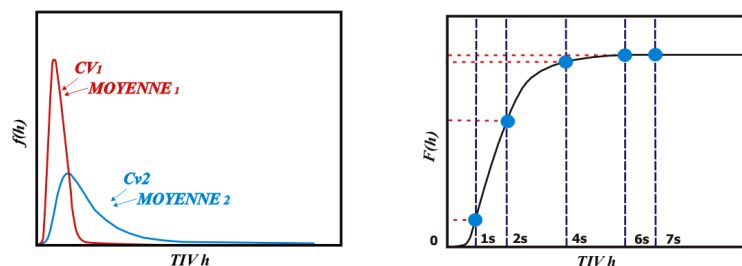


FIGURE 6.1 : La validation de statistiques partielles, type I et type II

de validation de ces approches.

6.2.1 Validation selon les relations établies sur les différents sites de la RN118 - Approche de régression

L'approche *de régression* s'appuie sur les relations établies entre les paramètres du modèle Γ -GQM et le débit Q_6 pour les données du Site N°1 de la RN118 (cf. Chapitre 4). Le Tableau 6.1 représente le niveau de validation en fonction des résultats de l'application du modèle sur les échantillons correspondants des autres sites de la RN118.

P-valeur du test K-S, le modèle Γ -GQM				
Q_6 véh/h	Site 2	Site 3	Site 4	Site 5
300-400	$2.098 \cdot 10^{-7}$	0.0744	$2.2420 \cdot 10^{-4}$	$3.0124 \cdot 10^{-4}$
400-500	$7.6426 \cdot 10^{-7}$	$1.6545 \cdot 10^{-12}$	$6.5768 \cdot 10^{-20}$	$1.7814 \cdot 10^{-18}$
500-600	$1.6218 \cdot 10^{-29}$	$3.6498 \cdot 10^{-30}$	$1.4839 \cdot 10^{-48}$	$1.6982 \cdot 10^{-27}$
600-700	$1.8908 \cdot 10^{-74}$	$1.1497 \cdot 10^{-79}$	$2.2985 \cdot 10^{-84}$	$6.9485 \cdot 10^{-94}$
700-800	$8.3412 \cdot 10^{-126}$	$6.2070 \cdot 10^{-161}$	$9.3018 \cdot 10^{-133}$	$1.7374 \cdot 10^{-156}$
800-900	$1.7293 \cdot 10^{-168}$	$9.8196 \cdot 10^{-265}$	$7.9859 \cdot 10^{-200}$	$1.4641 \cdot 10^{-221}$
900-1000				
1000-1100	$2.5089 \cdot 10^{-16}$	$1.7965 \cdot 10^{-7}$	$2.4477 \cdot 10^{-9}$	$8.8509 \cdot 10^{-9}$
1100-1200	$1.3054 \cdot 10^{-4}$	$9.111 \cdot 10^{-4}$	0.0079	$8.6479 \cdot 10^{-5}$
1200-1300	$2.5089 \cdot 10^{-16}$	0.0406	0.1837	0.0664
1300-1400	0.0876	$2.9347 \cdot 10^{-7}$	0.0303	$1.8878 \cdot 10^{-6}$
1400-1500	0.2385	$2.7636 \cdot 10^{-14}$	$2.6044 \cdot 10^{-11}$	$2.8595 \cdot 10^{-33}$
1500-1600	0.0220	$2.3985 \cdot 10^{-4}$	$8.2244 \cdot 10^{-9}$	$1.8695 \cdot 10^{-25}$
1600-1700	$1.7733 \cdot 10^{-15}$	–	$8.7256 \cdot 10^{-4}$	$1.0328 \cdot 10^{-7}$

TABLE 6.1 : Validation selon l'approche de régression sur les différents sites de la RN118

On constate que l'utilisation des relations établies donne de mauvais ajustements dans la majorité des échantillons utilisés par la validation. Les causes semblent être :

- En premier lieu, certains paramètres comme θ du modèle Γ -GQM n'ont pas une relation significative avec le débit Q_6 .
- Les paramètres α, β varient beaucoup selon les échantillons ; comme on a vu que le maximum de vraisemblance sur un échantillon était très sensible aux valeurs de ces paramètres, la déduction de leurs valeurs à partir des régressions pourrait alors être entachée d'erreur. Le paramètre λ est très sensible à des petites variations du débit Q_6 alors que les débits sont groupés en classes dont la largeur (100 véh/h) est sans doute trop grande.

- Les relations de régression liant les paramètres au débit peuvent être différentes selon les différents sites.

Les temps de calcul pourraient être réduits en utilisant une abaque des probabilités des TIV, comme celles établies par Grecco [60] et Khanasbis [66] - ces auteurs n'ayant toutefois pas procédé à la validation de leur modèle sur d'autres échantillons. Pour conclure, les résultats montrent que cette approche est fragile et difficile à mettre en place sur les autres jeux de données, que les calibrages ne sont pas transférables tels quels sur un autre site. Il faudra caractériser les sites et calibrer les modèles par type de site.

6.2.2 Validation de statistiques des échantillons - Approche de stockage

La validation de statistiques considère que les échantillons servant au calibrage et les échantillons sur lesquels ce calibrage est validé sont similaires si leurs propriétés statistiques sont simultanément équivalentes. L'approche de *stockage* signifie ici que les paramètres obtenus préalablement par le calibrage sont sauvegardés et constituent les éléments d'une base de référence. Les valeurs de ces paramètres sont utilisées pour un nouvel échantillon lorsque les caractéristiques statistiques de celui-ci sont proches de celles d'un des échantillons calibrés. Différents critères statistiques seront utilisés dans les sous-sections suivantes :

Validation de statistiques partielles (type I) sur les différents sites de la RN118

Cette validation correspond à la première question de l'approche de statistique partielle. Dans ce premier type de validation de statistiques partielles, le coefficient de variation CV, supposé connu, est utilisé conjointement avec la moyenne des TIV. En général, le CV fournit la forme de la distribution ce qui est une information des plus importantes.

P-valeur du test K-S, le modèle Γ -GQM						
Q_6 véh/h	300-400			700-800		
	Moyenne	CV	P_{KS}	Moyenne	CV	P_{KS}
Site 1	9.9282	1.0884	0.5265	4.8080	1.1217	0.3909
Site 2	11.9924	6.9724	$1.88 \cdot 10^{-4}$	4.8086	0.8339	$8.926 \cdot 10^{-26}$
Site 3	22.1155	7.5315	$2.76 \cdot 10^{-4}$	4.7763	0.7925	$1.3688 \cdot 10^{-42}$
Site 4	10.2924	0.9709	$1.26 \cdot 10^{-9}$	4.8440	0.8117	$7.7625 \cdot 10^{-27}$
Site 5	10.0929	1.0423	$1.89 \cdot 10^{-9}$	4.7409	0.7795	$1.6141 \cdot 10^{-36}$

TABLE 6.2 : Validation d'une statistiques partielles (type I) pour les échantillons des différents sites de la RN118, débits faibles

Les Tableaux 6.2 et 6.3 fournissent les résultats de validation. On constate que :

- La dégradation de l'ajustement lors de l'application du modèle sur différents sites est très importante si seule l'information sur la moyenne des TIV est prise en compte, d'autant plus qu'au même niveau de débit Q_6 en trafic fluide, la moyenne des TIV varie encore largement vu la largeur (relative) de la classe de débit.
- L'exigence de l'équivalence des CV en complément de celle des moyennes des TIV semble augmenter le niveau de validation (cf. Q_6 de (1400-1500) véh/h entre les sites N°1 et N°2; Q_6 de (1700-1800) véh/h entre les sites N°1 et N°2) Pourtant, ce n'est pas toujours le cas (cf. Q_6 de (300-400) véh/h entre les sites N°1 et N°4; Q_6 de (1400-1500) véh/h entre les sites N°1 et N°4)

Il semble qu'il y ait un effet du site (Site N°2 et Site N°4). Cependant, il est observé que l'information du CV est importante pour prédire la distribution des TIV.

P-valeur du test K-S, le modèle Γ -GQM						
Q_6 véh/h	1400-1500			1700-1800		
	Moyenne	CV	P_{KS}	Moyenne	CV	P_{KS}
Site 1	2.4869	0.6726	0.9805	2.0515	0.6196	0.9708
Site 2	2.5349	0.6600	0.0104	2.1077	0.6080	0.0490
Site 3	2.5098	0.5508	$4.587 \cdot 10^{-12}$	–	–	–
Site 4	2.5246	0.6387	$3.6957 \cdot 10^{-9}$	–	–	–
Site 5	2.5144	0.5956	$6.9564 \cdot 10^{-27}$	–	–	–

TABLE 6.3 : Validation d’une statistiques partielles (type I) pour les échantillons des différents site de la RN118, débits forts

Validation d’une statistique *totale*

Chaque échantillon de débit Q_6 du Site N°1 de la RN118 de la modalité (jour-sec) sur la VL contient beaucoup d’individus. Prenons par exemple la classe de débit Q_6 (400-500) véh/h, qui comporte 5 426 véhicules, la classe de débit Q_6 (1200-1300) véh/h, qui comporte 146 963 véhicules, etc. Pour des raisons de temps du calcul, on ne considère au maximum qu’un nombre M de véhicules dans chaque classe. Dans cette section, on examine la validation du modèle Γ -GQM dans des *sous échantillons* de chaque classe. Ces sous échantillons sont les résultats du glissement des données par un pas de Δ véhicules ; chacun contient encore M véhicules.

Soit un échantillon *groupé* de classe C_i regroupant $N > M$ véhicules ayant les TIV $H_0 = (h_1, h_2, \dots, h_M)$, les sous échantillons sont définis par $H_k = (h_{k\Delta+1}, \dots, h_{k\Delta+M})$, avec $k = 1, 2, \dots$. On a donc l’échantillon *de référence* H_0 qui avait servi au calibrage dans les chapitres précédents. Le choix de $M = 3000$ semble raisonnable, ce nombre est suffisant pour calibrer les quatre à six paramètres des modèles. Notons que dans la littérature, le nombre de TIV pris en compte dans chaque échantillon de calibrage était plus faible, citons par exemple $M_{\text{Buckley}} \leq 2071$ (cf. [57]), $M_{\text{Branston}} \leq 2238$ (cf. [58]), $M_{\text{Cowan}} \leq 1324$ (cf. [59]), $M_{\text{Luttinen}} \leq 600$ (cf. [11]), $M_{\text{Hoogendoorn}} \leq 1727$ (cf. [80]). Par ailleurs, on prend $\Delta = 100$ pour que la différence entre deux échantillons *glissants* H_k voisins soit non négligeable.

La validation par glissement consiste d’abord à appliquer les paramètres obtenus dans l’échantillon de référence H_0 sur les échantillons utilisés pour la validation H_k . Ensuite, pour évaluer la similitude entre des sous échantillons glissants H_k et l’échantillon de référence au niveau statistique, le test non paramétrique de K–S est utilisé. Ce test mesure la proximité des échantillons glissants H_k avec l’échantillon de référence H_0 .

Les résultats de validation sont présentés dans la Figure 6.2 dans laquelle la ligne rouge (en trait plein) représente la proximité statistique entre les échantillons glissants H_k et l’échantillon de référence H_0 . Cette proximité est mesurée par le test K–S non paramétrique et illustrée par la P-valeur. La ligne verte (en trait pointillé) représente les valeurs des P-valeurs du test K–S en appliquant les paramètres déjà estimés sur l’échantillon H_0 aux échantillons glissants H_k .

En sachant que la P-valeur de tout modèle (dont le modèle Γ -GQM) calibré sur un jeu de données diminuera normalement lorsque ce même modèle est appliqué sans nouveau calibrage sur un autre jeu de données, on constate que :

- Si l’échantillon H_k n’est pas similaire à l’échantillon H_0 au sens du test K–S non paramétrique, l’ajustement paramétrique est rejeté au niveau de signification de 5 %. En conséquence, l’*invalidation* statistique entraîne en général l’*invalidation* probabiliste.

- Un modèle bien calibré sur le premier jeu de données ne pourra obligatoirement pas donner, avec le même calibrage, de bons résultats sur le second jeu, si celui-ci est similaire au premier. La dégradation du niveau d’ajustement est observée par la position plus basse de la ligne verte (en trait pointillé) par rapport à la ligne rouge (en trait plein).
- Lorsque l’échantillon H_k est proche de H_0 au sens du test K–S non paramétrique, c.à.d lorsque sa P-valeur (la ligne rouge en trait plein) se trouve au dessus du seuil de 0,05, les paramètres du H_0 peuvent être appliqués au H_k dans la majorité de cas.
- Il existe plusieurs cas où une partie de la ligne rouge en trait plein est au dessous de la ligne verte en trait pointillé, ce qui signifie que le modèle Γ -GQM est plus proche de l’échantillon H_k au sens du test K–S que de l’échantillon de référence H_0 . Ces cas sont en fait tout à fait normaux car dans certaines conditions, la courbe de la fonction de répartition calibrée se situe entre les courbes correspondant respectivement aux TIV observés des échantillons H_0 et H_k , et elle est ici plus particulièrement plus proche de la dernière.
- En fonction du débit Q_6 (cf. Figure A.4 de l’Annexe), il semble que la validation soit acquise pour les débits intermédiaires compris entre 1000 et 1700 véh/h.

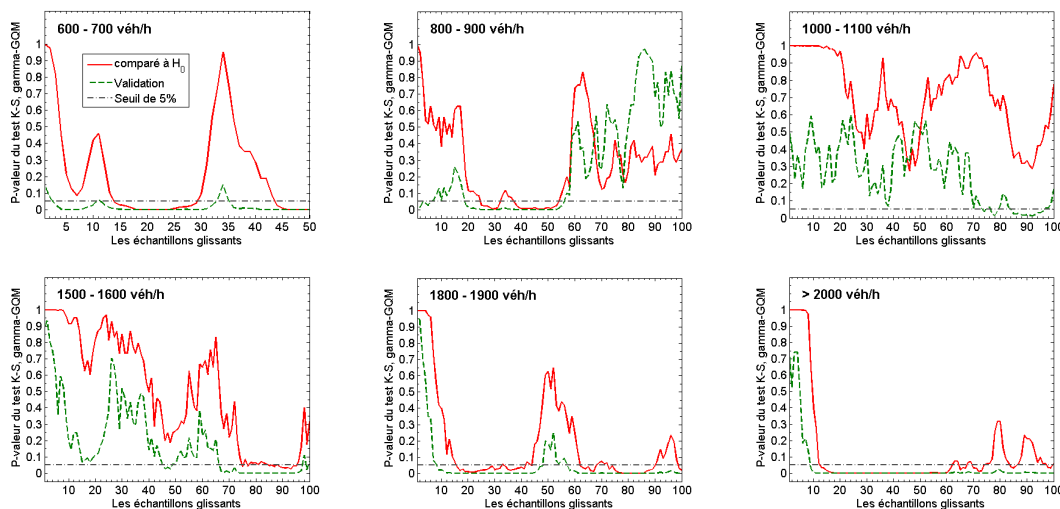


FIGURE 6.2 : Validation d’une statistique totale sur les échantillons glissants

6.2.3 Statistiques partielles (type II) sur les différents échantillons glissants du Site N°1 de la RN118

Cette validation correspond à la deuxième question de l’approche de statistique partielle. Certaines remarques de la Section 6.2.2 ne sont pas de surprises car la *validation statistique* entraîne en général la *validation probabiliste*. La validation d’une statistique *totale* dans la Section 6.2.2 assure au sens du test K–S la proximité statistiquement des courbes cumulées des deux échantillons H_k et H_0 . Pourtant, cette approche de validation n’a guère d’intérêt applicatif car dans certains cas, le rejet de la validation fondé sur la statistique totale est dû à une mauvaise modélisation de la fréquence de certains TIV qui n’ont pas d’intérêt en pratique.

Du point de vue des exploitants ou pour certaines recherches dans le domaine du trafic, des TIV « particuliers » (certains étant courts, d’autres longs) jouent des rôles plus importants que les autres valeurs de TIV. Par exemple, les chercheurs étudiant le trafic

au rond point ou en l'intersection sans feux s'intéressent particulièrement aux TIV longs, souvent supérieurs à 6 s tandis que les chercheurs en sécurité routière ne considèrent que les TIV courts, disons moins de 2 s. Dans ces cas-là, le critère de validation pertinent pondérera davantage la distance entre fréquences modélisées et empiriques de ces TIV particuliers.

Le critère de comparaison entre deux échantillons

Tout d'abord, on suppose que les deux échantillons H_i et H_j sont dans la même classe de débit Q_6 . Les hypothèses du test sont que :

$$\mathcal{H}_0 : p_I = p_J \quad (6.1)$$

$$\mathcal{H}_1 : p_I \neq p_J \quad (6.2)$$

p_I et p_J sont les vecteurs contenant les pourcentages à tester des deux échantillons H_i et H_j respectivement, I et J sont les index représentant les cas considérés en pratique. Suivant la littérature sur le TIV, les trois cas suivants sont examinés :

1. p_I, p_J sont les pourcentages des TIV inférieurs à 2 s, l'hypothèse \mathcal{H}_0 consiste à dire $p_2^I = p_2^J = p_2$
2. p_I, p_J sont les pourcentage des TIV supérieurs à 7 s, , l'hypothèse \mathcal{H}_0 consiste à dire $p_7^I = p_7^J = p_7$
3. p_I, p_J sont les vecteurs (de dimension 3) des pourcentages des TIV inférieurs à 2 s, de 2 à 7 s et supérieur à 7 s

Dans les deux premiers cas, le test de l'hypothèse nulle \mathcal{H}_0 se base sur le fait que si \mathcal{H}_0 est vraie, f_1 et f_2 sont des réalisations indépendantes de deux variables F_1 et F_2 suivant des lois normales :

$$F_i \sim \mathcal{N} \left(\hat{p}, \sqrt{\frac{\hat{p}(1-\hat{p})}{n_i}} \right) \quad (6.3)$$

Avec $i = 1, 2$, n_i sont les tailles des échantillons F_i et l'estimateur $\hat{p} = \frac{n_1 f_1 + n_2 f_2}{n_1 + n_2}$

Par conséquent, la variable $F_1 - F_2$ suit une loi normale centrée :

$$F_1 - F_2 \sim \mathcal{N} \left(0, \sqrt{\hat{p}(1-\hat{p})} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right) \quad (6.4)$$

Le troisième cas consiste à comparer deux pourcentages en même temps (le dernier pourcentage étant déduit des deux premiers) de deux échantillons dont un est l'échantillon de référence, le test du χ^2 est utilisé dans ce cas :

	Modalité 1	Modalité 2	Modalité 3	Total
	TIV < 2 s	2 s ≤ TIV < 7 s	% TIV ≥ 7 s	TIV
H_0	n_{11}	n_{12}	n_{13}	$n_{1.}$
H_k	n_{k1}	n_{k2}	n_{k3}	$n_{k.}$
Total	$n_{.1}$	$n_{.2}$	$n_{.3}$	n

TABLE 6.4 : Tableau de contingence du test du χ^2

Où n_{1j} et n_{kj} sont les nombres des TIV de la modalité j des échantillons 1 et k , n_1 , et n_k , étant les effectifs des échantillons 1 et k , n_j est le nombre total de véhicules de TIV j , on a :

$$n_{i.} = \sum_{j=1}^3 n_{ij} \quad (6.5)$$

$$n_{.j} = \sum_{i=1}^2 n_{ij} \quad (6.6)$$

$$n = \sum_i \sum_j n_{ij} \quad (6.7)$$

Dans ce cas, l'hypothèse \mathcal{H}_0 dit : « les deux échantillons proviennent de la même population » ; elle est testée contre l'hypothèse alternative \mathcal{H}_1 que : « les deux échantillons sont significativement différents ». On dit donc que les deux échantillons sont *similaires* si l'hypothèse \mathcal{H}_0 est vraie. On mesure la distance entre les deux échantillons comme suit :

$$d^2 = \sum_{i=1}^k \sum_{j=1}^r \frac{\left(n_{ij} - \frac{n_{i.} n_{.j}}{n} \right)^2}{\frac{n_{i.} n_{.j}}{n}} = \sum_{i=1}^k \sum_{j=1}^r \left(\frac{n_{ij}^2}{n_{i.} n_{.j}} - 1 \right) n \quad (6.8)$$

La mesure d^2 est la réalisation de la variable D^2 qui suit une loi du χ^2 de $(k-1) \times (r-1)$ degré de liberté si \mathcal{H}_0 est vraie.

En réalité, le test utilisant la loi normale dans le premier cas est strictement équivalent à celui de la loi χ^2 dans le deuxième cas quand $r = 2$. En effet, une χ^2 à un degré de liberté est le carré d'une loi normale, $\chi_1^2 \equiv \mathcal{N}(0, 1)$.

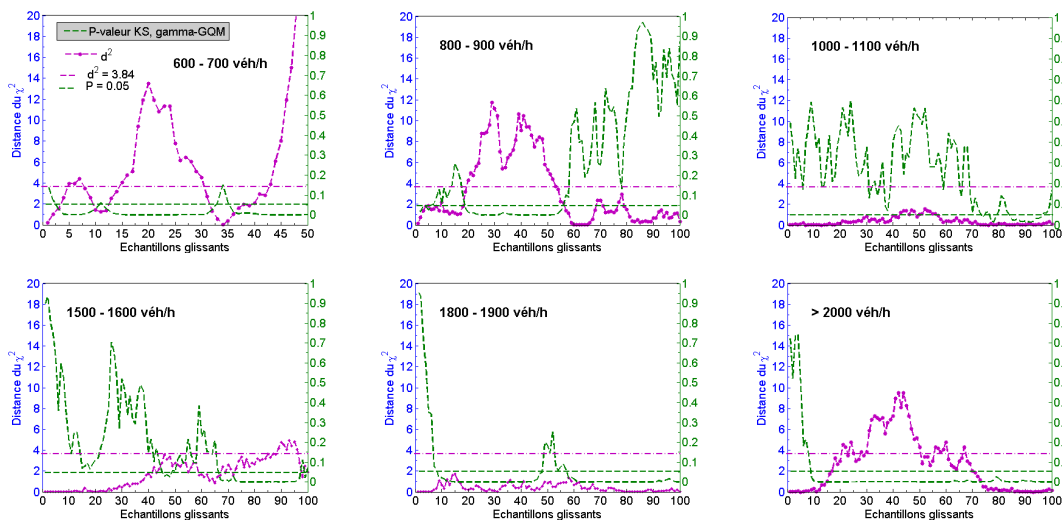


FIGURE 6.3 : Validation d'une statistique partielle de type II pour les échantillons glissants, critère de similitude p_2 des TIV < 2 s

6.2.4 Validation au sein du Site N°1

La Figure 6.3 montre (le carré de) la distance entre les échantillons H_k et l'échantillon de référence H_0 au sens de la similitude entre les pourcentages p_2 des TIV inférieurs à 2 s. On constate que :

- Dans les cas où les pourcentages p_2 diffèrent significativement entre H_k et H_0 , c. à. d. quand le carré de leur distance du χ^2 dépassent le seuil $d^2 = 0,384$, l'application du modèle Γ -GQM obtenu pour H_0 à l'échantillon H_k est rejetée.
- Pour les échantillons de débit Q_6 inférieur à 1800 véh/h, si la différence entre H_k et H_0 n'est pas significative, l'application d'un modèle bien calibré pour l'échantillon de référence H_0 sera dans la majorité des cas acceptée pour l'échantillon à valider H_k au niveau de signification de 5 %, ce qui est un élément de validation du modèle.

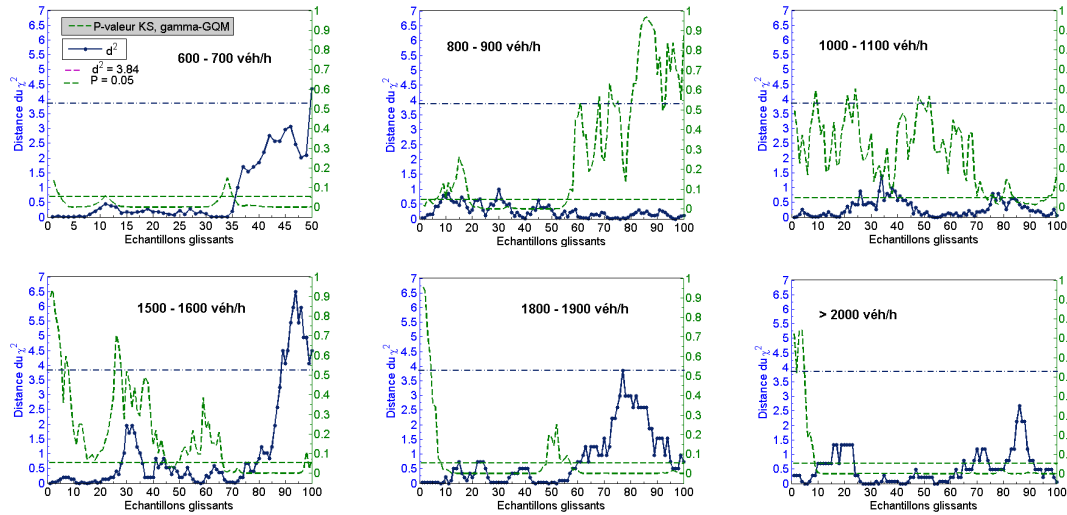


FIGURE 6.4 : Validation d'une statistique partielle de type II pour les échantillons glissants, critère de similitude p_7 des TIV ≥ 7 s

Il est à noter que pour les échantillons de haut débit, disons plus élevé que 1800 véh/h, le TIV est presque toujours inférieur à 10 s. La Figure 6.4 représente la distance entre l'échantillon à valider H_k et l'échantillon de référence H_0 au sens de la similitude entre les pourcentages p_7 des TIV supérieurs ou égaux à 7 s. On constate que dans la majorité des cas, la différence entre H_k et H_0 n'est pas significative. Dans chaque classe de débit, les pourcentages des TIV supérieurs ou égaux à 7 s sont assez stables. En conséquence, l'ajout de ce critère de validation n'est pas si « important ». Étant donné que le critère basé sur ce pourcentage p_7 joue un rôle faible, il est prévu que les résultats du troisième cas soient semblables à ceux du premier cas qui utilisent uniquement un critère de validation fondé sur le pourcentage p_2 . En effet, la Figure 6.5 montre des résultats similaires de la Figure 6.3.

6.3 Proposition d'une approche de validation et d'application

Face à la complexité des données réelles des TIV, l'application d'un modèle avec des valeurs exactes des paramètres semble difficile à valider et à évaluer. L'approche de régression et l'approche de stockage ne sont pas consistantes.

Vu que le modèle Γ -GQM utilisé fournit toujours de bons calibrages et que les courbes de la fonction de répartition sont ordonnées en fonction du débit Q_6 , une approche de validation est proposée. On observe qu'il existe un ensemble de distributions correspondant à chaque classe de débit Q_6 . La Figure 6.6 et la Figure 6.7 représentent toutes les distributions des échantillons glissants et de l'échantillon de référence superposées.

De plus, on remarque que :

- Les courbes de distributions se rapprochent particulièrement dans les classes de débit

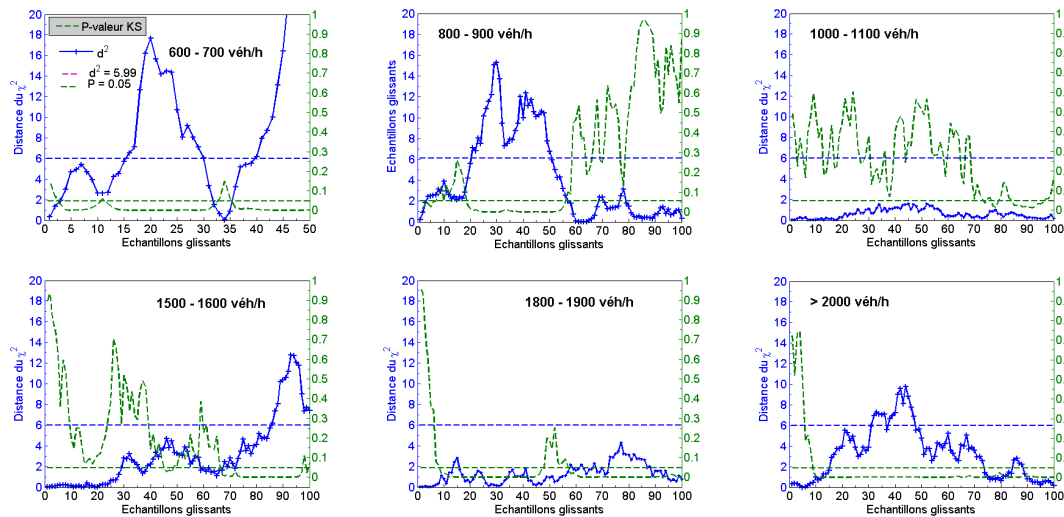


FIGURE 6.5 : Validation d'une statistique partielle de type II pour les échantillons glissants, critère de similitude $2 \leq TIV < 7$ s

compris entre 1000 et 2000 véh/h.

- Le phénomène de *multi-distribution* est plus net dans les classes de débit faible d'autant plus que deux familles de courbes semblent apparaître dans certaines classes de débit Q_6 .

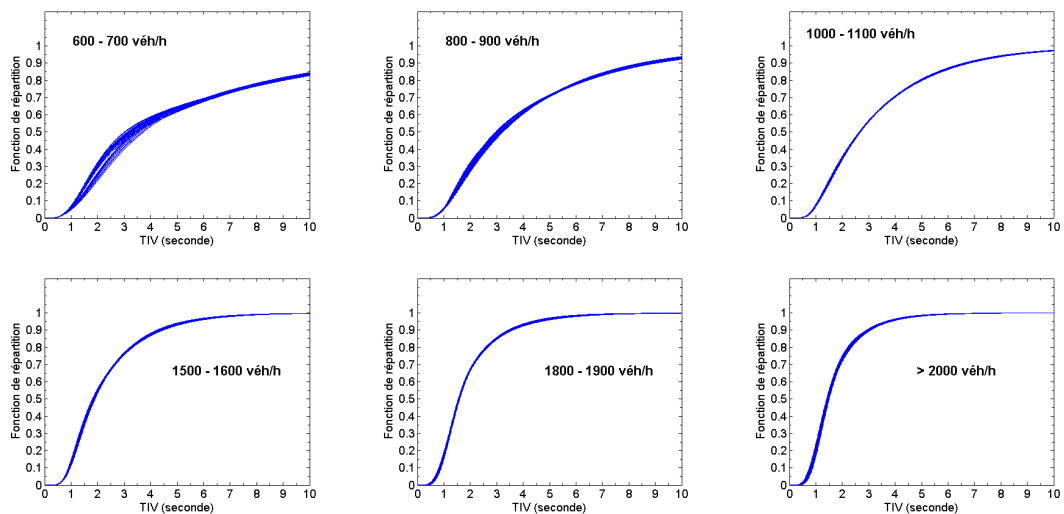


FIGURE 6.6 : Les ensembles de courbes de la fonction de répartition dans les classes de débit

L'approche proposée est appelée *approche de zonage* ou *enveloppe des distributions* qui prend en compte la proximité des distributions des TIV dans chaque classe d'échantillons. Comme la sensibilité des paramètres du modèle Γ -GQM est non négligeable, cette approche semble plus raisonnable pour valider les modèles de TIV. L'application de l'approche consiste dans des étapes suivantes :

1. Tout d'abord, une base de référence des paramètres estimés est établie. Cette base est obtenue par les calibrages des échantillons glissants de la base de données des TIV choisie *a priori*. Dans notre cas, cette base de référence des échantillons est le

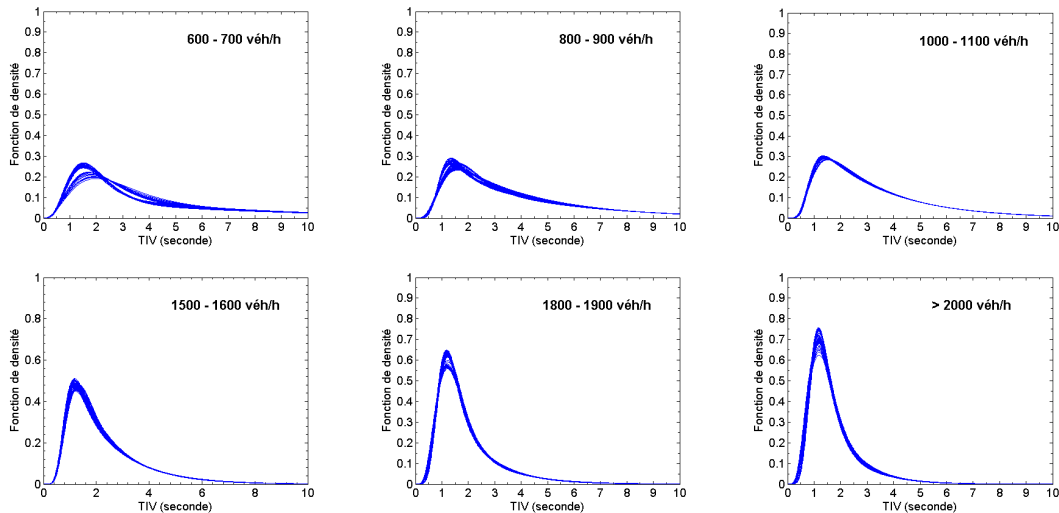


FIGURE 6.7 : Les ensembles de courbes de la fonction de densité dans les classes de débit

Site N°1 de la RN118.

$$\mathcal{B}_\Theta = \{\vec{\alpha}, \vec{\beta}, \vec{\theta}, \vec{\lambda}\}$$

Où $\vec{\alpha}, \vec{\beta}, \vec{\theta}, \vec{\lambda}$ sont les vecteurs des valeurs estimées des paramètres α, β, θ et λ . La taille des vecteurs est le nombre d'échantillons calibrés à l'entrée.

2. Les valeurs de la fonction de répartition F sont calculées aux points des TIV prédéfinis \vec{h} (disons de zéro à 20 s avec l'incrément de 0,01 s). Ensuite, les valeurs maximum et minimum obtenues sont utilisées pour déterminer une *zone de validation* \mathcal{Z} correspondant à une classe d'échantillons (par exemple la classe de débit (600-700) véh/h).

$$F(\vec{h}) = F(\vec{h}, \mathcal{B}_\Theta)$$

$$\text{Sup}_{\vec{h}} = \max_{\Theta} F(\vec{h})$$

$$\text{Inf}_{\vec{h}} = \min_{\Theta} F(\vec{h})$$

3. Pour évaluer le niveau de *transférabilité* du calibrage obtenu dans une classe \mathcal{C} de débit, on calibre aussi le modèle pour cette même classe \mathcal{C} mais à partir des autres sites de la RN118. Dans les cas où l'ajustement du modèle Γ -GQM n'est pas rejeté au niveau de signification de 5 % au sens du test K-S, les valeurs de la fonction de répartition F sont calculées aux points du vecteur h et on vérifie si elles se trouvent dans l'*enveloppe des distributions* $\mathcal{Z}_{\mathcal{C}}$ de la classe \mathcal{C} .
4. Finalement le nombre n des échantillons disponibles pour la validation satisfaisant le critère 3 ci-dessus divisé par le nombre total des échantillons considérés N mesure le niveau de validation :

$$\mathcal{A} = \frac{n}{N}$$

6.3.1 RN118 sur A6

Le Tableau 6.5 représente le niveau de *transférabilité* (ou de validation) sur les données de l'autoroute A6. Il est clair que l'application, à un autre réseau, d'un modèle ayant des paramètres obtenus sur un premier réseau n'est pas possible d'autant plus que le débit calculé sur l'A6 est un débit horaire Q_h , et non un débit 6 minutes Q_6 . Ce type de validation est particulièrement difficile pour les TIV courts comme le montre la Figure 6.8.

N°	Q	Classe RN118	A	N°	Q	Classe RN118	A
L1	1255	1200-1300	0	L7a	1330	1300-1400	0
L2	1272	1200-1300	0	L7b	1409	1400-1500	0
L3a	832	800-900	0	L7c	1405	1400-1500	0
L3b	814	800-900	0	L8	1402	1400-1500	0
L3c	810	800-900	0	L9a	782	700-800	0
L3d	862	800-900	0	L9b	772	700-800	0
L3e	838	800-900	0	L9c	793	700-800	0
L4	1259	1200-1300	0	L9d	834	800-900	0
L5	1196	1100-1200	0	L11	1436	1400-1500	0
L6a	802	800-900	0	L12a	735	700-800	0
L6b	802	800-900	0	L12b	790	700-800	0
L6c	828	800-900	0	L12c	801	800-900	0
L6d	807	800-900	0				

TABLE 6.5 : La validation sur les données de l'autoroute A6

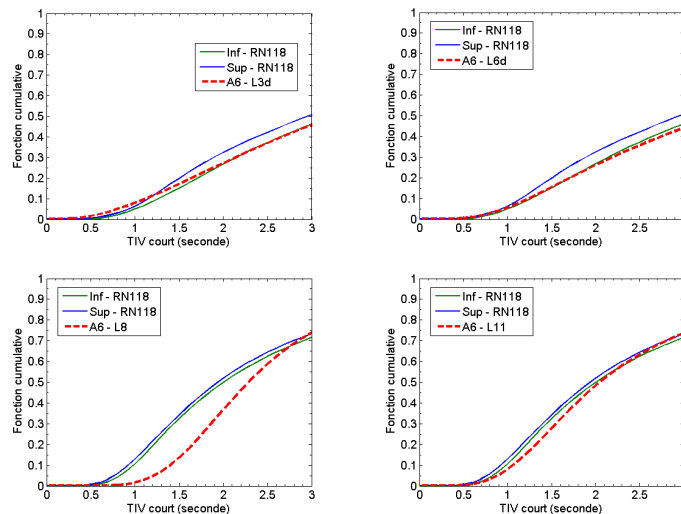


FIGURE 6.8 : Les ensembles de courbes de la fonction de répartition (A6 et RN118)

6.3.2 Validation sur différents sites de la RN118

Les cas marqués « - » signifient que le calibrage du modèle Γ -GQM ne s'est pas bien passé. Dans ce cas, on ne fait pas la validation du modèle pour la classe correspondante. Il est à noter que le modèle Γ -GQM ne donne pas toujours des estimations numériques satisfaisantes et qu'une seule erreur peut empêcher la sortie des résultats.

Les résultats de validation dans le Tableau 6.6 montrent que :

- L'effet de lieu existe car il y a des rejets systématiques comme dans le cas de validation pour les échantillons de l'autoroute A6.
- Comme la zone \mathcal{Z}_C est étroite pour les classes C de haut débit, une proximité statistique entre les échantillons H_k et H_0 peut être rejetée par cette méthode de validation. En conséquence, cette approche peut être améliorée en acceptant aussi une validation par exemple d'un échantillon H_k ayant sa courbe de la fonction de répartition proche au sens du test K-S des courbes $\text{Sup}_{\bar{h}}$ et $\text{Inf}_{\bar{h}}$.
- Le niveau de validation de cette approche de zonage dépend aussi de la base de

\mathcal{C} véh/h	Le niveau de validation \mathcal{A} (%)			
	Site N°2	Site N°3	Site N°4	Site N°5
(0-200)	–	–	–	–
(200-300)	–	–	–	–
(300-400)	–	–	–	–
(400-500)	–	–	–	–
(500-600)	–	–	–	–
(600-700)	–	–	–	–
(700-800)	16,67	0	35,62	45,45
(800-900)	–	0	1,33	0
(900-1000)	79	0	70,73	0
(1000-1100)	37	0	0	0
(1100-1200)	48,81	0	50,26	0
(1200-1300)	0	0	0	0
(1300-1400)	73,12	–	0	0
(1400-1500)	100	–	0	0
(1500-1600)	5	–	–	–
(1600-1700)	0	–	–	–
(1700-1800)	–	–	–	–
(1800-1900)	–	–	–	–
(1900-2000)	–	–	–	–
(> 2000)	–	–	–	–

TABLE 6.6 : Le niveau de validation dans différents sites de la RN118

référence \mathcal{B}_Θ . Déjà, des niveaux de validation élevés atteints dans certaines classes montrent que cette approche semble prometteuse pour la pratique.

6.4 Conclusion du chapitre

Différentes approches de validation sont abordées. Le processus de validation utilise soit des échantillons de la même classe (de débit), mais correspondant à différents sites de la RN118, soit des échantillons glissants dans la même classe d'échantillonnage du Site N°1 de la RN118. La validation conduit à distinguer la validation statistique issue de la comparaison des échantillons et la validation probabiliste de l'application des modèles. En général, une bonne validation statistique entraîne une bonne validation probabiliste et vice versa. L'effet du lieu est non négligeable voire fort. Enfin, l'approche par enveloppe des distributions proposée est une première approche qui semble prometteuse pour la pratique et l'utilisation des modèles Γ -GQM et DG.

Chapitre 7

Conclusion générale

L'objectif initial de la thèse était de modéliser le TIV par temps sec et par temps de pluie sur la route RN118. Ce but a été atteint et a ensuite été élargi en considérant d'autres facteurs aussi importants.

L'étude bibliographique indique que les modèles de TIV sont répartis en trois types : les modèles simples, les modèles combinés et les modèles de mélange. Une comparaison exhaustive entre les modèles probabilistes de TIV et entre les méthodes d'estimation montre que deux modèles de mélange, de performance équivalente, se détachent : le modèle Semi-Poisson (SPM) avec la loi Gamma comme la fonction de densité des véhicules contraints (Γ -SPM) et le modèle de file d'attente généralisée utilisant aussi la loi Gamma comme la fonction de densité des véhicules contraints (Γ -GQM). La facilité de calcul du modèle Γ -GQM conduit à employer ce modèle dans toute la modélisation des TIV.

Suite au classement et à la comparaison des modèles de TIV, de nouveaux modèles sont proposés, développés et testés. Six modèles sont bons ou très bons : le modèle HyperLNM, le modèle Hyper-GQM, le modèle GQM mixte, le modèle Double Gamma, le modèle Double Lognormal et le modèle Lognormal Gamma. En dehors des deux premiers de performance inférieure, les quatre derniers sont équivalents voire meilleurs que le modèle Γ -GQM. Les résultats des comparaisons des modèles et des méthodes d'estimation sont fiables car un ensemble test d'échantillons est proposé. Cet ensemble représente beaucoup d'états du trafic dans différents contextes de météo, de voie et de réseaux. De plus, les méthodes d'estimation pour les modèles complexes (les modèles combinés, de mélange et les nouveaux qui sont développés) sont exécutées selon le même processus d'estimation proposé, celui-ci étant exigeant au niveau de la stabilité des résultats sortants. Enfin, pour une raison pratique, le modèle Double Gamma est choisi à côté du modèle Γ -GQM dans toute la modélisation des TIV.

Avec ces deux modèles performants, la distribution des TIV est étudiée sous un point de vue hiérarchique. Trois dimensions d'étude du trafic sont considérées correspondant à trois échelles, les échelles macroscopique, mésoscopique et microscopique. Une quatrième dimension d'étude des TIV est constituée de facteurs exogènes au trafic proprement dit. Les facteurs à l'échelle macroscopique ainsi que les facteurs exogènes affectent l'ensemble des véhicules. Les facteurs macroscopiques sont tout particulièrement importants par leur impact au niveau quantitatif. Les facteurs exogènes quant à eux concernent la situation dans laquelle le trafic se déplace. L'effet des facteurs exogènes sur la distribution des TIV est considérée à chaque niveau des variables macroscopiques (à savoir le débit agrégé 6 minutes). Les résultats montrent que les évolutions des paramètres varient de façon interprétable en fonction du débit et au long du diagramme fondamental. De plus, selon leur effet, les facteurs exogènes sont distingués en deux types, à savoir le facteur « *empêchant* » et le facteur « *impulsif* ». Chacun de ces sous facteurs augmente ou réduit les paramètres

d'échelle des modèles Γ -GQM et DG.

L'échelle mésoscopique met en jeu les interactions des véhicules dans un groupe. En particulier, le véhicule précédent est considéré comme le principal facteur stimulant le changement de TIV, de vitesse, d'accélération comme le font les modèles de poursuite. Comme dans les modèles probabilistes construits précédemment, les véhicules sont généralement classés en véhicules libres et véhicules contraints dont les comportements de conduite sont différents ; une étude statistique préliminaire est réalisée afin de définir le véhicule suiveur et le véhicule libre de façon quantitative. Ensuite, les échantillons de TIV sont constitués selon les catégories mésoscopiques telles que la position d'un véhicule en peloton, la catégorie de poursuite (ou bien la configuration du trafic) et aussi selon la valeur des TIV des véhicules précédents. Cette approche fournit des connaissances intéressantes sur la distribution des TIV à l'échelle mésoscopique.

L'échelle microscopique est étudiée en utilisant conjointement d'autres informations sur les véhicules eux-mêmes. La distribution des TIV est modélisée en tenant en compte de la vitesse instantanée des véhicules. Par ailleurs, la variable microscopique RHIS formée par le rapport du TIV et de la vitesse instantanée est, elle aussi, bien calibrée par le modèle Γ -GQM.

Finalement, différentes approches de validation sont testées. L'utilisation directe des valeurs estimées des paramètres par l'approche de régression conduit toujours à des erreurs importantes, tandis que la classification des échantillons, de manière à ne considérer que des échantillons « *similaires* » selon les statistiques générales ne suffit pas à éviter une grande dégradation des ajustements. Pour la pratique, l'approche proposée par « *enveloppe des distributions* » semble prometteuse. Pourtant, l'effet du lieu a un impact important sur la validation utilisant cette approche.

Les perspectives à donner à ce travail de recherche consistent à étudier les autres bases de données en utilisant le modèle Γ -GQM et le modèle Double Gamma pour comparer et valider les modélisations. Les données des TIV de la nuit et des routes départementales sont intéressantes car les grandeurs des valeurs des TIV y sont différentes de celles de l'A6 et la RN118. Ceci demandera un processus de traitement considérablement différent. La Figure 7.1 illustre les valeurs des TIV sur une route départementale des Côtes d'Armor.

D'un point de vue technique, trois pistes principales sont à explorer :

1. La modélisation du TIV en utilisant le modèle GQM *mixte*, ce qui est un sujet ouvert car ce modèle comporte des paramètres physiques précis concernant les différents composants du TIV. On a bien vu que le modèle Γ -GQM à quatre paramètres a des paramètres très sensibles, tandis que le modèle DG à cinq paramètres a des paramètres beaucoup moins variables, pourtant les paramètres de ce dernier modèle apparaissent moins reliés aux phénomènes du trafic. Le modèle GQM *mixte* à six paramètres pourrait remédier aux inconvénients des deux modèles Γ -GQM et DG.
2. Une étude plus fine sur l'intensité des effets exogènes est intéressante. Par exemple, la condition de pluie peut être divisée en plusieurs niveaux plus homogènes de précipitation. Les effets plus *complexes* de différents facteurs pris simultanément peuvent aussi être étudiés en détail.
3. La modélisation du TIV à l'échelle mésoscopique (les véhicules en peloton) peut être ajustée par l'information macroscopique du flux de véhicules. Cela pourrait conduire à distinguer différents types de peloton au sens du comportement de conduite.

D'un point de vue pratique, les pistes suivantes sont à étudier :

1. La mise en application des modèles probabilistes de TIV par l'approche des *enveloppes de distribution*.

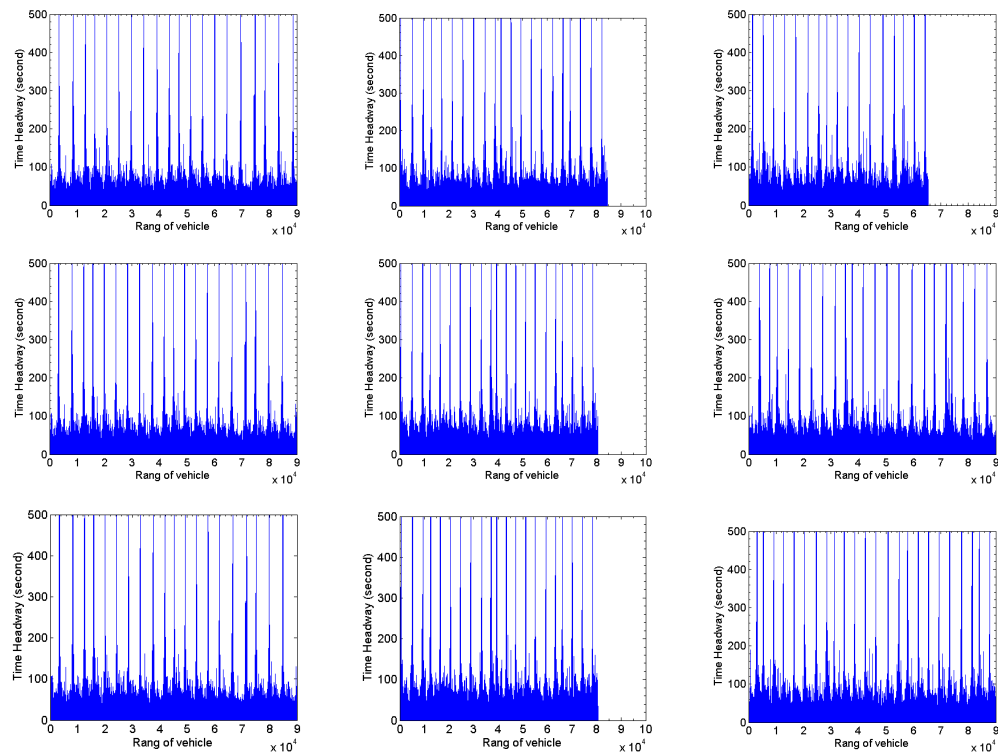


FIGURE 7.1 : Les valeurs des TIV sur une route départementale des Côtes d'Armor

2. L'utilisation des résultats à l'échelle mésoscopique pour mieux générer les véhicules entrants et pour vérifier la plausibilité des véhicules sortants d'un système de simulation.
3. L'utilisation des résultats à l'échelle microscopique pour mieux générer en un point donné simultanément les arrivées et les vitesses des véhicules entrants d'un système de simulation.

Bibliographie

- [1] P. ATHOL, “Headway groupings,” *Highway Research Record*, vol. 72, pp. 137–155, 1965.
- [2] W. F. ADAMS, “Road traffic considered as a random series,” *Journal of Institut Civil Engineering*, vol. 4, pp. 121–130, 1936.
- [3] M. BRACKSTONE and M. McDONALD, “Car-following : a historical review,” *Transportation Research Part F*, vol. 2, pp. 181–196, 1999.
- [4] L. A. PIPES, “An operational analysis of traffic dynamics,” *Journal of Applied Physics*, vol. 24, pp. 274–281, 1953.
- [5] E. KOMETANI and T. SASAKI, “Dynamic behavior of traffic with a nonlinear spacing-speed relationship,” *Proceedings of the Symposium on Theory of Traffic Flow*, pp. 105–119, 1959. Research Laboratories, General Motors.
- [6] P. G. GIPPS, “A behavioural car following model for computer simulation,” *Transportation Research Part B*, pp. 105–111, 1981.
- [7] R. E. CHANDLER, R. HERMAN, and E. W. MONTROLL, “Traffic dynamics : Studies in car following,” *Operations Research*, vol. 6, pp. 165–184, 1958.
- [8] W. HELLY, “Simulation of bottlenecks in single lane traffic flow,” *Proceeding of the Symposium on Theory of Traffic Flow*, pp. 207–238, 2010. Research Laboratories, General Motors, New York, Elsevier.
- [9] W. LEUTZBACH, *An Introduction to the Theory of Traffic Flow*. Springer-Verlag, 1988.
- [10] J. XING, “A parameter identification of a car following model,” *Proceeding of the Second World Congress on ATT*, pp. 1739–1745, 1995. Yakohama, November.
- [11] R. T. LUTTINEN, *Statistical Analysis of Vehicular Time Headways*. PhD thesis, University of Technology Lahti Center (Neopoli, Lahti, Finland), May 1996.
- [12] A. TORDEUX, S. LASSARRE, and M. ROUSSIGNOL, “An adaptive time gap car-following model,” *Transportation Research Part B*, vol. 44, pp. 1115–1131, 2010.
- [13] M. BANDO, K. HASEBE, N. A., and A. SHIBATA, “Dynamical model of traffic congestion and numerical simulation,” *Phys. Rev. E*, vol. 51, pp. 1035–1042, 1995.
- [14] H. LENZ, C. K. WAGNER, and R. SOLLACHER, “Multi-anticipative car-following model,” *The European Physical Journal B*, vol. 7, no. 2, pp. 331–335, 1999.
- [15] R. JUANG, Q. WU, and Z. ZHU, “Full velocity difference model for a car-following theory,” *Physical Review E*, vol. 64, no. 1, pp. 017101.1–017101.4, 2001.
- [16] G. F. NEWELL, “A simplified car-following theory : a lower order model,” *Transportation Research Part B*, vol. 36, no. 3, pp. 195–205, 2002.
- [17] M. TREIBER, A. HENNECKE, and D. HELBING, “Congested traffic states in empirical observations and microscopic simulations,” *Physical Review E*, vol. 62, no. 2, pp. 1805–1824, 2000.

- [18] R. M. MICHAELS, "Perceptual factors in car following," *Proceedings of the Second International Symposium on the Theory of Road Traffic Flow*, pp. 44–59, 1963. Paris, OCDE.
- [19] R. WIEDEMANN, "Simulation des straenverkehrsflusses," *Schriftenreihe des instuts für verkehrswesen*, 1974. Universität Karlsruhe, Heft 8.
- [20] W. LEUTZBACH and R. WIEDEMANN, "Development and applications of traffic simulation models at the karlsruhe institut fur verkehrswesen," *Traffic Engineering and Control*, vol. 27, no. 5, pp. 270–278, 1986.
- [21] C. KIKUCHI and P. CHAKROBORTY, "Car following model based on a fuzzy inference system," *Transportation Research Record*, vol. 1365, pp. 82–91, 1992.
- [22] M. C. DUNNE, R. W. ROTHERY, and R. B. POTTS, "A discrete markov model of vehicular traffic," *Transportaion Science*, vol. 2, no. 3, pp. 233–251, 1968.
- [23] X. CHEN, L. LI, and Y. ZHANG, "A markov model for headway/spacing distribution of road traffic," *IEEE Transactions on Intelligent Transportation Systems*, 2010. in press.
- [24] I. PRIGOGINE and F. C. ANDREWS, "A boltzmann-like approach for traffic flow," *Operations Research*, vol. 8, pp. 789–797, 1960.
- [25] I. PRIGOGINE and R. HERMANN, "Comment on "an analysis of the boltzmann-type statistical models for multi-lane traffic flow"," *Transportation Research*, vol. 4, pp. 113–116, 1970.
- [26] I. PRIGOGINE and R. HERMANN, *Kinetic Theory of Vehicular Traffic*. Elsevier, 1971.
- [27] P. NELSON, "A kinetic model of vehicular traffic and its associated bimodal equilibrium solutions," *Transport Theory and Statistical Physics*, vol. 24, pp. 383–409, 1995.
- [28] P. MUNJAL and J. PAHL, "An analysis of the boltzmanntype statistical models for multi-lane traffic flow," *Transportation Research*, vol. 3, pp. 151–163, 1969.
- [29] S. L. PAVERI-FONTANA, "On boltzmann-like treatments for traffic flow : A critical review of the basic model and an alternative proposal for dilute traffic analysis," *Transportation Research*, vol. 9, pp. 225–235, 1975.
- [30] W. F. PHILLIPS, "A kinetic model for traffic flow with continuum implications," *Transportation Planning and Technology*, vol. 5, pp. 131–138, 1979.
- [31] A. KLAR and R. WEGENER, "A hierarchy of models for multilane vehicular traffic," *SIAM J. Appl. Math.*, vol. 59, pp. 983–1011, 1999.
- [32] D. HELBING, "An interpretation of speed-flow concentration relationships using catastrophe theory," *Transportation Research A*, vol. 21, pp. 191–201, 1989.
- [33] D. HELBING, "Gas-kinetic derivation of navier-stokes-like traffic equations," *Physical Review E*, vol. 53, pp. 2366–2381, 1996.
- [34] P. NELSON and A. SOPASAKIS, "The chapman-enskog expansion : A novel approach to hierarchical extension of lighthill-whitham models," *Transportation and Traffic Theory : Proceedings of the 14th International Symposium on Transportation and Traffic Theory*, pp. 51–80, 1999. Pergamon A. Ceder, Edition.
- [35] G. F. NEWELL, "Approximation methods for queues with application to the fixed-cycle traffic light," *SIAM Review*, vol. 7, no. 2, pp. 223–240, 1965.
- [36] J. F. C. KINGMAN, "On queues in heavy traffic," *Journal of the Royal Statistical Society*, vol. 24, no. 2, pp. 383–392, 1962.

- [37] G. ZHANG, Y. WANG, H. WEI, and Y. CHEN, "Examining headway distribution models with urban freeway loop event data," *Transportation Research Record*, vol. 1999, pp. 141–149, 2007.
- [38] D. P. SULLIVAN and R. J. TROUTBECK, "The use of cowan's m3 headway distribution for modeling urban traffic flow," *Traffic Engineering and Control*, vol. 35, pp. 445–450, 1994.
- [39] R. J. TROUTBECK, "A review of the process to estimate the cowan m3 headway distribution parameters," *Traffic Engineering and Control*, vol. 38, pp. 600–603, 1997.
- [40] A. D. MAY, "Gap availability studies," *Highway Research Record*, no. 72, pp. 101–136, 1964.
- [41] J. P. KINZER, *Application of the Theory of Probability to Problems of Highway Traffic*. PhD thesis, Polytechnic Inst. of Brooklyn, 1933.
- [42] L. F. RADER, "Review of application of the theory of probability to problems of highway traffic," *Proc. Inst. of Traffic Engineering*, pp. 118–123, 1934.
- [43] F. GARWOOD, "An application of the theory of probability to the operation of vehicular-controlled traffic signals," *Journal Royal Stat. Soc.*, vol. 7, pp. 65–77, 1940.
- [44] GREENSHIELDS, "Traffic performance at urban street intersections," *Bureau of Highway Traffic, Tech. Rept.*, vol. 1, 1947.
- [45] J. C. TANNER, "The delay to pedestrians crossing a road," *Biometrika*, vol. 38, pp. 383–392, 1951.
- [46] D. L. GERLOUGH, "Use of poisson distribution in highway traffic," *Poisson and Traffic*, pp. 1–58, 1955. Eno Foundation.
- [47] A. SCHUHL, "Probability theory applied to vehicle distribution on two-lane highways," *In Poisson and Traffic*, 1955. The Eno Foundation for Highway Traffic Control, Saugatuck, Conn.
- [48] D. L. GERLOUGH, "Traffic inputs for simulation on a digital computer," *Highway Research Board Proc.*, vol. 38, pp. 480–492, 1959.
- [49] H. J. KELL, "Analyzing vehicular delay at intersections through simulation," *Highway Research Board*, vol. 356, pp. 28–39, 1962.
- [50] H. J. KELL, "A theory of traffic flow on urban streets," *Institute of Traffic Engineering, Western Section Proc.*, 1960. July.
- [51] R. M. LEWIS, "A proposed headway distribution for traffic simulation studies," *Traffic Engineering*, pp. 16–19, 1963.
- [52] R. M. LEWIS and H. L. MICHAEL, "Simulation of traffic flow to obtain volume warrants for intersection control," *Highway Research Record*, vol. 15, pp. 1–43, 1963.
- [53] F. A. HAIGHT, B. WHISLER, and W. W. MOSHER, "New statistical method for describing highway distributions of cars," *Highway Research Board Proceedings*, vol. 40, pp. 557–564, 1961.
- [54] L. BREIMAN, "The poisson tendency in traffic distribution," *Electronics Research Laboratory*, vol. 60, pp. 1–9, 1962.
- [55] I. GREENBERG, "The log-normal distribution of headways," *Australian Road Research*, vol. 2, pp. 14–20, 1966.
- [56] L. A. DAWSON, F. et CHIMINI, "The hyperlang probability distribution - a generalized traffic headway model," *Highway Research Board*, vol. 230, pp. 1–14, 1968.
- [57] D. J. BUCKLEY, "A semi-poisson model of traffic flow," *Transportation Science*, vol. 2, pp. 107–133, 1968.

- [58] D. BRANSTON, "Models of single lane time headway distributions," *Transportation Science*, vol. 10, no. 12, pp. 125–148, 1976.
- [59] R. J. COWAN, "Useful headway models," *Transportation Research*, vol. 9, pp. 371–375, 1975.
- [60] W. L. GRECCO and E. C. SWORD, "Prediction of parameters for schuhl's headway distribution," *Traffic Engineering*, vol. 38, no. 5, pp. 36–38, 1968.
- [61] J. E. TOLLE, "The log-normal headway distribution model," *Traffic Engineering and Control*, vol. 13, no. 1, pp. 22–24, 1971.
- [62] W. D. ASHTON, "Distribution for gaps in road traffic," *Journal of Institut Mathematics Application*, vol. 7, pp. 37–46, 1970.
- [63] P. WASIELEWSKI, "An integral equation for the semi-poisson headway distribution model," *Transportation Science*, vol. 8, pp. 237–247, 1974.
- [64] P. WASIELEWSKI, "Car-following headways on freeways interpreted by the semi-poisson headway distribution model," *Transportation Science*, vol. 13, pp. 36–55, 1979.
- [65] M. S. ALY, *Headway Distribution Model Interrelationships between Headway and Fundamental Traffic Flow Characteristics*. PhD thesis, Institute of Transportation Engineering and University of Karlsruhe, 1989.
- [66] S. KHASNABIS and C. L. HEIMBACK, "Headway distribution models for two-lane rural highways," *Transportation Research Record*, vol. 772, pp. 44–51, 1980.
- [67] T. CHISHAKI and Y. TAMURA, "Headway distribution model based on the distinction between leaders and followers," *Ninth International Symposium on Transportation and Traffic Theory*, pp. 43–63, 1984.
- [68] D. HEIDEMANN, "A theoretical model to calculate speed distribution as a function of density," *Proceedings of the 10th International Symposium on Transportation and Traffic Theory*, pp. 79–98, 1987. in M. H. Gartner and N.H.M. Wilson, eds., New York, USA.
- [69] D. HEIDEMANN, "A theoretical model to calculate time-headway distributions as a function of traffic density," *Proceedings of the 11th International Symposium on Transportation and Traffic Theory*, 1990. in M. Koshi, ed., Yokohama, Japan.
- [70] D. HEIDEMANN, "A theoretical model for distributions of speeds and time-headways on two-lane roads," *Proceedings of the 12th International Symposium on Transportation and Traffic Theory*, pp. 523–537, 1993. in C. F. Daganzo, ed., Berkeley, California, USA.
- [71] D. P. SULLIVAN and R. J. TROUTBECK, "An exponential relationship for the proportion of free vehicles on arterial roads," *Traffic Science*, vol. 31, pp. 21–33, 1997.
- [72] R. AKCELIK and E. CHUNG, "Calibration of the bunched exponential distribution of arrival headways," *Road Transport Research*, vol. 3, pp. 42–59, 1994.
- [73] R. T. LUTTINEN, "Properties of cowan's m3 headway distribution," *Transportation Research Record*, vol. 1678, pp. 189–196, 1999.
- [74] O. HAGRING, "Use of the cowan m3 distribution for modeling roundabout flow," *Transportation Engineering Control*, vol. 37, pp. 328–332, 1996.
- [75] O. HAGRING, "Estimation of critical gaps in two major streams," *Transport Research - B*, vol. 34, pp. 293–313, 2000.

- [76] O. HAGRING, "Estimation of parameters in distribution of headways in roundabouts," *Transportation Engineering*, vol. 128, pp. 403–411, 2002.
- [77] O. HAGRING, "Calibration of headway distribution," *Proc. Of the 13th Mini-Euro Conference Handling Uncertainty in the Analysis of Traffic and Transportation Systems and the 9th Meeting of the Euro Working group on Transportation, Intermodality, Sustainability and ITS*, vol. 0, pp. 323–328, 2002.
- [78] S. P. HOOGENDOORN and H. BOTMA, "The estimation of parameters in headway distributions," technical report, TRAIL Research School Delft, 1996.
- [79] S. P. HOOGENDOORN and H. BOTMA, "Modeling and estimation of headway distributions," *Transportation Research Record*, vol. 1591, pp. 14–22, 1997.
- [80] S. P. HOOGENDOORN and P. H. L. BOVY, "New estimation technique for vehicle-type-specific headway distributions," *Transportation Research Record*, vol. 1646, pp. 18–28, 1998.
- [81] K. GHOSH, A. MAJUMDAR, and D. CHOWDHURY, "Distribution of time-headways in a particle-hopping model of vehicular traffic," *Physical Review E*, vol. 58, pp. 4012–4015, 1997.
- [82] D. CHOWDHURY and et al., "Particle-hopping models of vehicular traffic : Distribution of distance headways and distance between jams," *Physica A*, vol. 246, pp. 471–486, 1997.
- [83] D. CHOWDHURY, A. PASUPATHY, and S. SINHA, "Distribution of time-and distance-headways in the nagel-schreckenberg model of vehicular traffic : effects of hindrances," *Eur. Phys. J. B*, vol. 5, pp. 781–786, 1998.
- [84] M. KRBALEK, P. SEBA, and P. WAGNER, "Headways in traffic flow : Remarks from a physical perspective," *Physical Review E*, vol. 64, p. 066119, 2001.
- [85] M. Krbalek, "Headway distribution of the asymmetric simple exclusion model," ., vol. 0, pp. 1–10, 2002.
- [86] M. KRBALEK and P. SEBA, "The statistical properties of the city transport in cuernavaca (mexico) and random matrix ensembles," *Journal of Physics A*, vol. 33, pp. 229–234, 2000.
- [87] M. KRBALEK and P. SEBA, "Headway statistics of public transport in mexican cities," *Journal of Physics A : Mathematical and General*, vol. 36, pp. 7–11, 2003.
- [88] R. PUEBOOBPAPHAN and D. PARK, "Time headway distribution of probe vehicles on single and multiple lane highway," *International Conference on Application of Advanced Technologies in Transportation, (Athens, Greece)*, 2008.
- [89] T. HELMUT, O. ERDINC, and R. S. KIRAN, "Approximated headway distributions of free-flowing traffic on ohio freeways for work zone traffic simulations," *Transportation Research Record*, vol. 1999, pp. 131–140, 2007.
- [90] D. GERLOUGH and M. J. HUBER, *Traffic Flow Theory*. TRB, Special Report 165, 1975.
- [91] M. ARON, M.-B. BIECHELER, and J.-F. PEYTAVIN, "Temps intervéhiculaires et vitesses : Quel enjeux de sécurité sur l'autoroute?," *Recherche Transports Sécurité*, vol. 71, pp. 3–20, Juillet-Septembre 1999. French.
- [92] D. H. HA, M. ARON, and S. COHEN, "Variations of the parameters of the time headway models according to macroscopic variables, fundamental diagram and exogenous effects," *Transportation Research Board (Washington D.C, US)*, 2011.

- [93] D. H. HA, M. ARON, and S. COHEN, "Time headway variable and probabilistic modeling," *Traffic Flow Theory Committee (Annecy, France)*, 2010. Transportation Research Part-C (in submission).
- [94] D. H. HA, M. ARON, and S. COHEN, "Comparison of time headway distribution in different traffic contexts," *World Conference of Transportation Research (Lisbon, Portugal)*, 2010.
- [95] D. H. HA, "Exogenous effects on the microscopic distributions," *Transportation Research Arena*, 2010. Competition Year2010.
- [96] D. H. HA, M. ARON, and S. COHEN, "Modélisation du temps intervéhiculaire sur l'autoroute de france," tech. rep., Actes N°134, JDD SPI-STIC 2009 INRETS, 2009.
- [97] R. T. LUTTINEN, "Statistical properties of vehicle time headways," *Transportation Research Record*, vol. 1365, pp. 92–98, 1992.
- [98] J. H. BANKS, "Another look at a priori relationships among traffic flow characteristics," *Transportation Research Record*, vol. 1510, pp. 1–10, 1989.
- [99] M. J. CASSIDY and B. COIFMAN, "Relation among average speed, speed, flow, and density and analogous relation between density and occupancy," *Transportation Research Record*, vol. 1591, pp. 1–6, 1989.
- [100] D. R. COX and P. A. W. LEWIS, *L'analyse statistique des séries d'événements*. Dunod, 1969.
- [101] L. BREIMAN and F. LAWRENCE, "Time scales, fluctuations and constant flow periods in uni-directional traffic," *Transportation Research*, vol. 7, 1973.
- [102] A. J. MILLER, "An empirical model for multilane road traffic," *Transportation Science*, vol. 4, 1970.
- [103] D. R. COX and A. STUART, "Some quick sign tests for trend in location and dispersion," *Biometrika*, vol. 42, pp. 80–95, 1955.
- [104] A. STUART and J. K. ORD, *Kendall's Advanced Theory of Statistics. Volume 2 : Classical Inference and Relationship*. Edward Arnold, 1991.
- [105] A. S. AL-GHAMDI, "Analysis of time headways on urban roads : Case study from riyadh," *Transportation Engineering*, vol. 127, pp. 289–294, 2001.
- [106] B. SILVERMAN, *Density Estimation for Statistics and Data Analysis*. Cfpman and Hall Ltd., 1986.
- [107] H. WOLD, *A study in the analysis of stationary time series*. Almqvist and Wiksell, 1954.
- [108] WALD and WOLFOWITZ, "On a test whether two samples are from the same population," *Annals of Mathematical Statistics*, vol. 2, pp. 147–162, 1940.
- [109] D. H. HA, "Temps intervéhiculaire sur route nationale - etude du intervéhiculaire divisé par la vitesse sur la route nationale n118," technical report, INRETS, 2007. French.
- [110] B. S. KERNER, *Introduction to Modern Traffic Flow Theory and Control : The Long Road to Three-Phase Traffic Theory*. Springer, 2009.
- [111] K. NAGEL and M. SCHRECKENBERG, "A cellular automaton model for freeway traffic," *J. Phys. I France*, vol. 2, pp. 2221–2229, 1992.
- [112] M. SCHRECKENBERG, A. SCHADSCHNEIDER, and K. NAGEL, 1995.
- [113] T. THEDEEN, "A note on the poisson tendency in traffic distribution," *The Annals of Mathematical Statistics*, vol. 35, no. 4, pp. 1823–1824, 1964.

- [114] A. D. MAY, *Traffic Flow Fundamentals*. Prentice-Hall, Englewood Cliffs, 1990.
- [115] M. HE, X. LIU, and J. RONG, "Driver's preferred time headway selection : Experimental findings," *Traffic and Transportation Studies, Proc. ICTTS 2002*, vol. 2, pp. 889–894, 2002.
- [116] S. MAMMAR and D. KOENIG, "Un scénario périurbain de route automatisée," *Recherche Transports Sécurité*, vol. 75, pp. 144–167, Avril-Juin 2002. French.
- [117] R. J. TROUTBECK, "Average delay at an unsignalized intersection with two major streams each having a dichotomized headway distribution," *Transportation Science*, vol. 20, pp. 272–286, 1986.
- [118] O. HAGRING, "A further generalization of tanners formula," *Transport Research - B*, vol. 32, pp. 423–429, 1998.
- [119] J. C. TANNER, "The capacity of an uncontrolled intersection," *Biometrika*, vol. 54, pp. 657–658, 1967.
- [120] J. D. GRIFFITHS, "Vehicle headways in urban areas," *Traffic Engineering Control*, vol. 32, no. 10, pp. 458–462, 1991.
- [121] A. D. MAY, "Traffic flow fundamentals," *Transportation Research Record*, vol. 1678, pp. 189–196, 1990.
- [122] A. C. COHEN and B. J. WHITTEN, "Estimation in the three - parameter lognormal distribution," *Journal of the American Statistical Association*, vol. 75, no. 370, pp. 399–404, 1980.
- [123] D. J. BARTHOLOMEW, 1959.
- [124] D. J. BUCKLEY, "A semi-poisson model of traffic flow," *Transportation Science*, vol. 2, pp. 107–133, 1968.
- [125] E. L. CROW and K. SHIMIZU, *Lognormal Distributions : Theory and Application*. Marcel Dekker, Inc, 1988.
- [126] D. GROSS and C. M. HARRIS, *Fundamentals of Queueing Theory*. John Wiley & Sons, 1985.
- [127] E. SEDDIKI, "Impact de la pluie sur le trafic autoroutier - etude aux niveaux micro-, méso- et macroscopique," *Recherche Transports Sécurité (RTS)*, vol. 41, pp. 13–24, 1993.
- [128] P. G. MICHAEL, F. C. LEEMING, and W. O. DWAYER, "Headway on urban streets : Observational data and an intervention to decrease tailgating," *Transportation Research - F*, vol. 3, pp. 55–64, 2000.
- [129] M. ARON, M.-B. BIECHELER, and J.-F. PEYTAVIN, "Observatoire des vitesses et des interdistances sur autoroutes," technical report, INRETS, 1997. French.
- [130] J. C. TANNER, "A theoretical analysis of delays at an uncontrolled intersection," *Biometrika*, vol. 49, pp. 163–170, 1962.
- [131] W. BRILON, *Intersections without traffic signals. Recent Developments in Calculation Methods for Unsignalized Intersections in West Germany*. Springer - Verlag, 1988.
- [132] S. CAMIZ, "A propos des pelotons : Identification et indicateurs de risque," rapport dera/gretia, DSCR/INRETS, 2000.
- [133] S. P. HOOGENDOORN, "Unified," *Transportation Research Part B*, vol. 1591, pp. 14–22, 2001.
- [134] K. VOGEL, "What characterizes a free vehicle in an urban area?," *Transportation Research - F*, vol. 5, pp. 313–327, 2002.

- [135] M. AYCIN and R. BENEKOHAL, "A linear acceleration car-following model development and validation," *Transportation Research Records*, vol. 1644, pp. 10–19, 1998.

Annexe A

Les Figures et Tableaux supplémentaires

k	$\lambda_j = \lambda_{-j}$	$h(\omega)$	Variance $f^2(\omega)$
0 (Daniell)	$\frac{\sin(j\epsilon)}{ j\epsilon }$ ($ j \leq n$)	$\frac{1}{2\epsilon}$ ($ \omega \leq \epsilon$)	$\frac{2}{\pi\epsilon}$
1 (Bartlett)	$\begin{cases} 1 - \frac{ j }{l} & (j \leq l) \\ 0 & (j > l) \end{cases}$	$\frac{1}{2\pi l} \left[\frac{\sin(l\omega/2)}{\sin(\omega/2)} \right]^2$	$\frac{2l}{3n}$
3 (Tukey)	$\begin{cases} \frac{1}{2} [1 + \cos(\pi j/l)] & (j \leq l) \\ 0 & (j > l) \end{cases}$	$\frac{1}{4\pi} \left[\frac{\sin(l + \omega/2)}{\sin(\omega/2)} + \frac{1}{2} \left(\frac{\sin[(l + 1/2)(\omega + \pi/l)]}{\sin(\omega + \pi/l)/2} + \frac{\sin[(l + 1/2)(\omega - \pi/l)]}{\sin(\omega - \pi/l)/2} \right) \right]$	$\frac{3l}{4n}$
4 (Parzen)	$\begin{cases} 1 - \frac{6j^2}{l^2(1 - j /l)} & (j \leq l/2) \\ 2(1 - j /l)^3 & (l/2 \leq j \leq l) \\ 0 & (j > l) \end{cases}$	$\frac{3}{8\pi l} \left[\frac{\sin(l\omega/4)}{\sin(\omega/4)} \right]^4$	$\frac{l}{2n}$

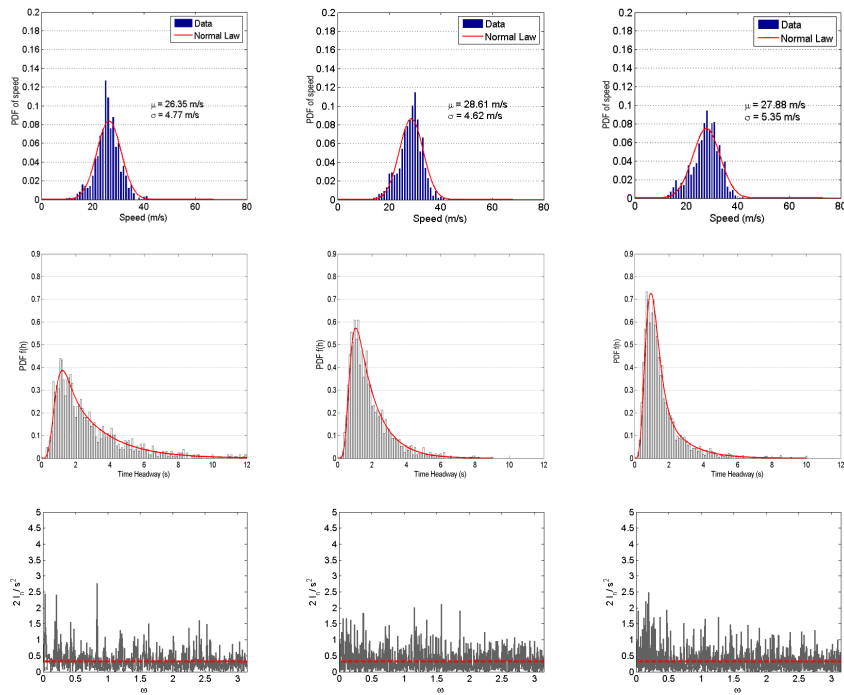
TABLE A.1 : Fonction de pondération spectrale

Site	PR	Virage	Limite de Vitesse	Signalisation	Millions de données
1	8591	R	110		1,5
2	–	VD	110		1,9
3	–	R	110		0,3
4	–	R	110		1,3
5	4356	R	110		1,2
6	3750	VD	90	Paneau CSA	1,7
7	3445	Avant/pendant VG	90		1,2
8	2169	Pendant/Après VG	–	Radar	0,7
9	1234	Avant/pendant VG	–		1,6

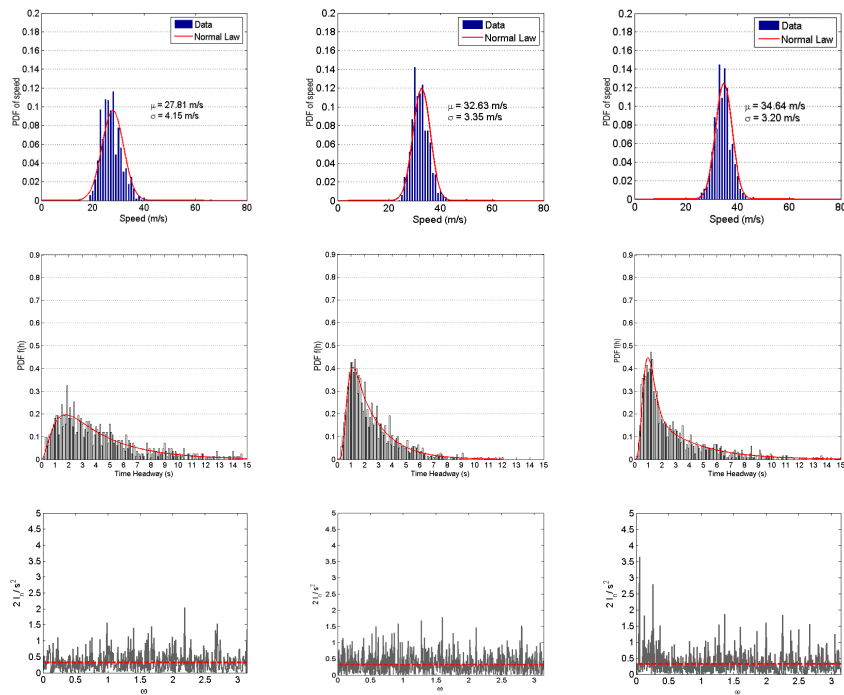
TABLE A.2 : L'information supplémentaire sur les sites de mesure de la RN118

Note :

- Les données ont été recueillies pendant 3-4 jours d'affilée hors week-end, sur 9 sites, du sens Province vers Paris, à partir de février 2005
- Le radar a été activé le 18 juillet 2005
- Le panneau CSA a été activé vers le 10 juillet 2005



Les échantillons L2, M2 et R2



Les échantillons L3a, M3a et R3a

FIGURE A.1 : Les échantillons sur les 3 voies dans la même période, Autoroute A6

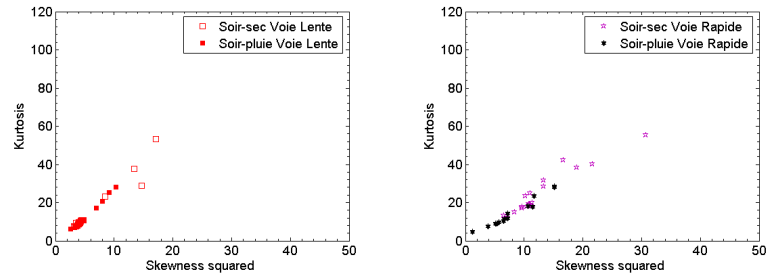


FIGURE A.2 : Relation $K-S^2$ des TIV du soir en fonction de Q_6 , RN118, Site N°1

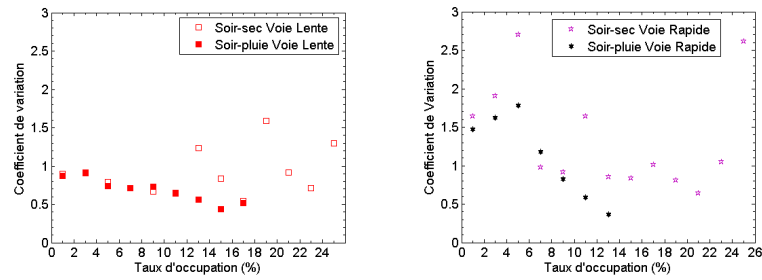


FIGURE A.3 : CV des TIV du soir en fonction de \mathcal{T}_6 , RN118, Site N°1

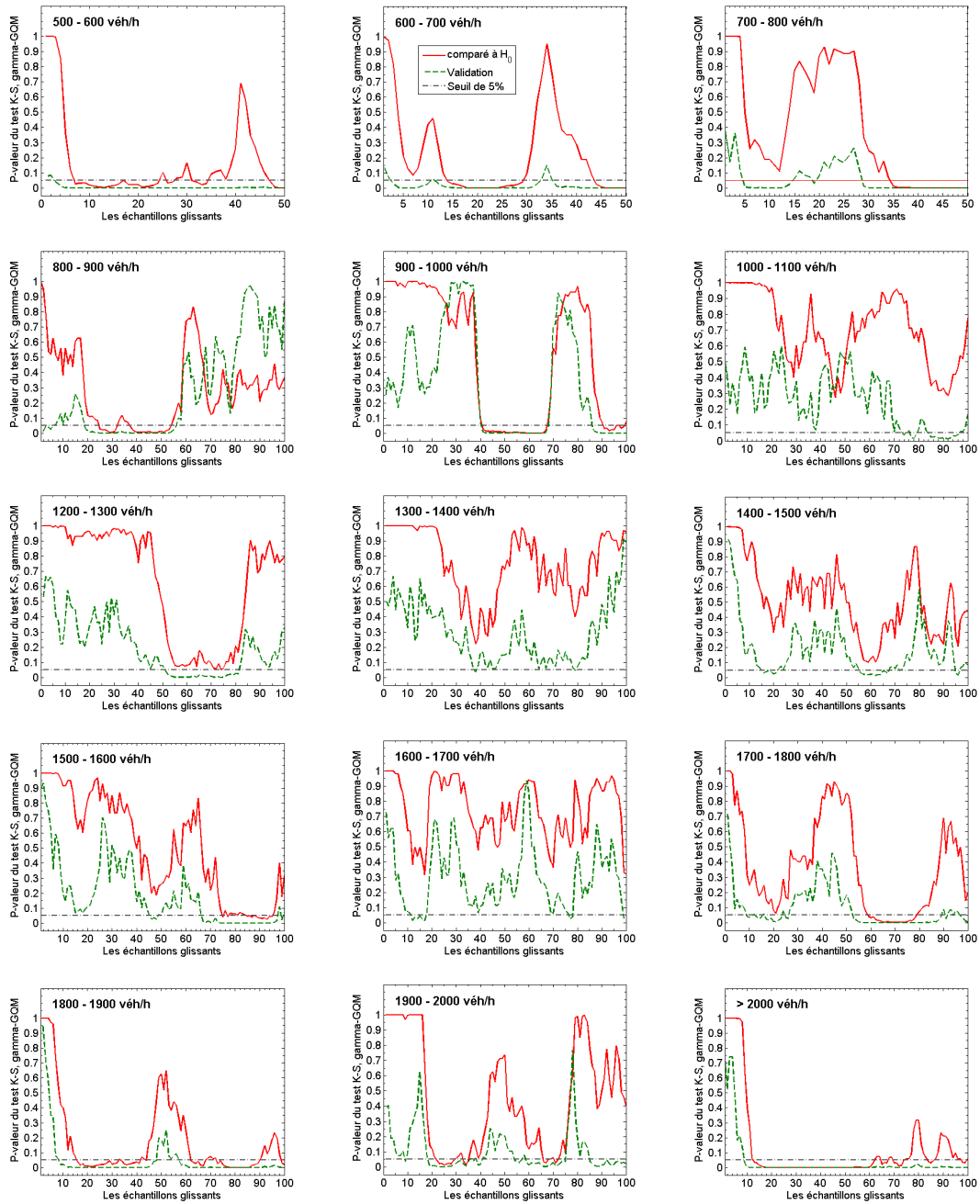


FIGURE A.4 : Validation d'une statistique totale sur les échantillons glissants, RN118, Site N°1

Annexe B

Les résultats numériques de l'A6 et de la RN118

N°	Profil	Q	Période	μ_H	σ_H	mode	CV	S_3	K_4	ρ_1	p
Voie Lente											
L1	Normal	1255	8h - 9h	2.8679	2.1509	1.15	0.75	2.0065	8.8503	0.0838	9.8805
L2	Normal	1272	8h - 9h	2.8272	2.1340	1.15	0.75	1.9017	8.3873	0.1029	10.3774
L3a	Normal	832	10h - 11h	4.3277	3.4652	1.95	0.8007	2.1096	10.6240	-0.0188	25.1504
L3b	Normal	814	11h - 12h	4.4185	3.4518	1.05	0.7812	1.8206	7.9305	-0.0002	22.6044
L3c	Normal	810	12h - 13h	4.4427	3.2923	1.75	0.7411	1.3393	4.8152	-0.0491	25.0617
L3d	Normal	862	13h - 14h	4.1759	3.0826	1.45	0.7382	1.3967	5.3127	-0.0662	23.5499
L3e	Normal	838	14h - 15h	4.2979	3.4778	2.35	0.8092	2.1891	12.0574	0.0116	21.4797
L4	Réd	1259	8h - 9h	2.8574	2.0536	1.25	0.7187	1.6483	5.9971	0.1090	11.8348
L5	Réd	1196	8h - 9h	3.0079	2.2625	1.25	0.7522	2.0830	9.5352	0.0770	11.8729
L6a	Réd	802	10h - 11h	4.4849	3.3444	1.75	0.7457	1.4962	5.6425	-0.0566	22.4439
L6b	Réd	802	11h - 12h	4.4890	3.3867	1.25	0.7544	1.5666	6.0378	0.0399	27.8055
L6c	Réd	828	12h - 13h	4.3242	3.1560	1.45	0.7299	1.4661	5.5192	0.0518	23.9130
L6d	Réd	807	13h - 14h	4.4569	3.2568	1.35	0.7307	1.4993	6.1290	-0.0307	21.0657
L7a	Réd Bou	1330	18h - 19h	4.4850	3.3256	1.75	0.7415	1.4811	5.6016	-0.0026	25.1128
L7b	Réd Bou	1409	19h - 20h	4.4673	3.3279	1.45	0.7450	1.5099	5.7033	0.0032	23.4209
L7c	Réd Bou	1405	20h - 21h	4.0889	3.0088	1.95	0.7359	1.5960	6.5790	-0.0129	21.5658
L8	Réd Bou	1402	11h - 12h	2.5675	1.2044	2.15	0.4691	1.7589	7.7147	-0.0304	8.8559
L9a	Normal	782	10h - 11h	4.5970	3.4985	2.05	0.7610	1.4006	4.9454	0.0071	26.0870
L9b	Normal	772	11h - 12h	4.6655	3.6316	2.05	0.7784	1.3970	4.9588	0.0032	29.4041
L9c	Normal	793	12h - 13h	4.5391	3.4968	1.45	0.7704	1.5122	5.4950	0.0457	24.33
L9d	Normal	834	13h - 14h	4.3094	3.2943	1.25	0.7644	1.9338	9.8669	-0.0086	21.2230
L11	Normal Bou	1436	14h - 15h	2.5048	1.5891	1.85	0.6344	1.9265	8.4091	-0.0095	8.4262
L12a	Réd	735	10h - 11h	4.8917	3.4949	1.85	0.7144	1.1960	4.1523	-0.0121	24.2177
L12b	Réd	790	11h - 12h	4.5519	3.4078	2.65	0.7486	1.4548	5.6046	0.0322	27.3418
L12c	Réd	801	12h - 13h	4.4933	3.3587	2.45	0.7475	1.3924	5.0522	0.0475	24.3446
Voie Médiane											
M1	Normal	1974	8h - 9h	1.8214	1.2401	1.05	0.6809	2.5278	14.0186	0.0824	1.5198
M2	Normal	1926	8h - 9h	1.8693	1.1604	1.05	0.6208	1.7326	7.1245	0.0568	1.2461
M3a	Normal	1411	10h - 11h	2.5475	1.8210	1.35	0.7148	1.7892	8.0790	0.0401	1.2757
M3b	Normal	1372	11h - 12h	2.6249	1.8965	1.15	0.7225	1.7486	7.4327	0.0268	1.2391
M3c	Normal	1433	12h - 13h	2.5107	1.7667	1.05	0.7037	1.8945	8.3677	0.0207	1.6050
M3d	Normal	1485	13h - 14h	2.4231	1.7999	1.15	0.7428	1.7574	7.2623	0.0611	0.8081
M3e	Normal	1581	14h - 15h	2.3693	1.7544	1.25	0.7405	2.2702	11.6689	0.0208	0.9881
M4	Réd	1802	8h - 9h	1.8309	1.2459	0.95	0.6805	2.2114	9.7412	0.0563	0.9434
M5	Réd	1978	8h - 9h	1.8192	1.2039	0.95	0.6617	1.8724	8.0163	0.0518	0.3033
M6a	Réd	1384	10h - 11h	2.6007	1.8584	1.05	0.7146	1.7919	7.8769	-0.0002	1.0838
M6b	Réd	1349	11h - 12h	2.6689	2.1097	1.15	0.7905	2.3967	12.4554	0.0769	1.3343
M6c	Réd	1397	12h - 13h	2.5943	1.9932	0.85	0.7683	2.2524	10.7926	-0.0179	0.6489
M6d	Réd	1343	13h - 14h	2.6820	2.1185	0.85	0.7899	2.2028	10.6131	0.0392	0.2234
M7a	Réd Bou	1442	18h - 19h	2.5981	1.8457	1.05	0.7104	1.7907	7.8809	-0.0024	1.0402
M7b	Réd Bou	1492	19h - 20h	2.6687	2.1489	1.15	0.8052	2.4233	12.2518	0.0514	1.2735
M7c	Réd Bou	1598	20h - 21h	2.6777	2.0479	0.85	0.7648	2.1602	10.3303	0.0305	0.5006
M8	Réd Bou	1457	11h - 12h	2.4679	1.1264	1.75	0.4564	1.9333	9.0010	-0.0201	0
M9a	Normal	1432	10h - 11h	2.5119	1.8506	1.55	0.7368	1.7262	6.5892	0.0313	1.8156
M9b	Normal	1384	11h - 12h	2.6005	1.8672	1.55	0.7180	2.2894	14.2086	0.0005	1.5896
M9c	Normal	1385	12h - 13h	2.5989	1.9102	0.85	0.7350	1.7270	6.6193	0.0469	1.2274
M9d	Normal	1452	13h - 14h	2.4800	1.8715	1.35	0.7546	2.3023	12.2502	0.0500	0.7576
M11	Normal Bou	1619	14h - 15h	2.2238	1.2482	1.25	0.5613	1.8943	9.3379	0.0186	0.9265
M12a	Réd	1319	10h - 11h	2.7268	1.9402	0.95	0.7115	1.7877	7.3440	0.0351	1.0614
M12b	Réd	1380	11h - 12h	2.6082	1.8844	0.95	0.7225	1.5237	5.6428	0.0765	0.8696
M12c	Réd	1397	12h - 13h	2.5763	1.9210	1.25	0.7456	2.0578	9.2812	0.0463	0.7874
Voie Rapide											
R1	Normal	2247	8h - 9h	1.5992	1.2373	0.80	0.7737	2.8429	14.5024	0.0923	0
R2	Normal	2251	8h - 9h	1.5977	1.1846	0.75	0.7415	2.6641	13.2349	0.1052	0
R3a	Normal	1205	10h - 11h	2.9878	3.0255	1.25	1.0126	2.7389	13.6941	0.0781	0
R3b	Normal	1158	11h - 12h	3.1089	3.3457	0.85	1.0762	2.7546	12.9277	0.0877	0
R3c	Normal	1253	12h - 13h	2.8748	3.0273	0.75	1.0531	3.2028	20.7203	0.0963	0
R3d	Normal	1243	13h - 14h	2.8930	3.1826	0.95	1.1001	3.1018	17.0170	0.0295	0
R3e	Normal	1280	14h - 15h	2.8019	3.0645	0.85	1.0937	2.7298	12.5659	0.1106	0.0781
R4	Réd	2032	8h - 9h	1.7716	1.4962	0.95	0.8445	3.3257	21.5800	0.0761	0
R5	Réd	1955	8h - 9h	1.8408	1.6859	0.95	0.9158	3.8625	32.9151	0.0940	0
R6a	Réd	933	10h - 11h	3.8534	4.3428	1.15	1.1270	3.0536	18.1176	0.0848	0
R6b	Réd	989	11h - 12h	3.6421	4.1556	0.85	1.1410	2.8301	13.8168	0.0236	0.2022
R6c	Réd	1024	12h - 13h	3.5175	4.1095	0.85	1.1683	3.1994	16.9206	0.0314	0
R6d	Réd	942	13h - 14h	3.8249	4.0374	1.35	1.0556	2.3731	10.2611	0.1202	0
R7a	Réd Bou	1459	18h - 19h	3.8043	4.3520	0.85	1.1440	3.0331	16.9353	0.0508	0.0685
R7b	Réd Bou	1530	19h - 20h	3.5679	4.0476	0.85	1.1345	2.9496	15.0062	0.0442	0.0654
R7c	Réd Bou	2390	20h - 21h	3.4222	3.8097	0.65	1.1132	2.7986	13.3262	0.0864	0
R8	Réd Bou	1444	11h - 12h	2.4862	1.5175	1.95	0.6104	6.9028	99.5440	0.1167	0
R9a	Normal	1176	10h - 11h	3.0580	3.0283	0.95	0.9903	2.2719	9.4252	0.1251	0
R9b	Normal	1151	11h - 12h	3.1075	3.1685	0.95	1.0196	2.5781	12.6828	0.0989	0.0866
R9c	Normal	1204	12h - 13h	2.9897	3.0999	0.85	1.0369	2.7301	14.0786	0.0798	0
R9d	Normal	1264	13h - 14h	2.8488	3.0512	0.75	1.0710	3.1017	18.3326	0.0740	0.0791
R11	Normal Bou	1757	14h - 15h	2.0479	1.3752	1.35	0.6715	7.6559	113.9996	0.0935	0.0569
R12a	Réd	953	10h - 11h	3.7763	4.3022	1.05	1.1393	3.1581	17.8927	-0.0060	0.1049
R12b	Réd	1030	11h - 12h	3.4978	3.8270	0.75	1.0941	2.7867	13.3548	0.1420	0
R12c	Réd	1126	12h - 13h	3.1965	3.2785	1.15	1.0257	2.4864	11.2125	0.0845	0

TABLE B.1 : Les propriétés statistiques du TIV de l'autoroute A6

$Q_6/100$	Statistiques										I-GQM										DG			
	véh/h	σ_H	CV	S_3	K_4	Mo	α	β	θ	λ	P_{K-S}	A_2	α_1	α_2	β_1	β_2	θ	P_{K-S}	A_2					
0 - 2	18.763	0.885	1.447	5.114	5.950	1.621	3.403	0.024	0.051	0.859	0.313	0.086	2.604	0.367	2.584	0.610	0.999	0.112						
2 - 3*	15.976	1.260	3.758	27.453	1.250	5.284	9.194	0.205	0.073	0.028	1.648	0.082	1.430	1.428	3.883	0.674	0.585	1.550						
3 - 4	10.806	1.088	4.104	46.509	2.050	2.809	5.070	0.194	0.099	0.527	0.520	0.128	1.662	1.386	3.633	0.708	0.153	1.548						
4 - 5	8.916	1.130	2.438	11.277	1.550	2.926	5.286	0.340	0.108	0.027	1.916	0.144	1.680	2.130	4.519	0.606	0.001	3.769						
5 - 6	8.001	1.198	3.103	19.159	1.550	2.306	4.477	0.446	0.117	0.095	2.804	0.159	1.772	1.869	4.061	0.504	0.018	3.885						
6 - 7	5.912	1.049	2.595	13.001	1.250	3.097	5.359	0.348	0.167	0.341	1.081	0.226	1.929	1.933	4.186	0.544	0.195	2.527						
7 - 8	5.393	1.122	4.096	30.214	1.650	2.544	4.780	0.441	0.191	0.391	1.893	0.244	1.857	1.869	4.151	0.480	0.115	3.594						
8 - 9	3.756	0.894	4.170	41.177	1.550	6.026	7.527	0.112	0.301	0.043	1.379	0.363	2.354	1.353	3.568	0.406	0.097	2.529						
9 - 10	2.952	0.785	2.704	17.568	1.350	6.104	6.871	0.059	0.358	0.358	1.209	0.435	2.578	1.241	3.371	0.324	0.061	2.951						
10 - 11*	2.572	0.751	2.304	12.736	1.350	10.290	10.124	0.038	0.394	0.346	0.669	0.642	2.950	2.528	4.928	0.557	0.115	2.211						
11 - 12	2.125	0.674	1.798	8.297	1.550	5.692	6.750	0.082	0.466	0.229	1.063	0.830	3.362	2.895	5.359	0.593	0.584	0.973						
12 - 13	1.977	0.686	1.852	8.114	1.250	7.567	8.247	0.099	0.502	0.576	0.873	0.901	3.217	4.615	7.144	0.660	0.348	1.420						
13 - 14	1.825	0.680	1.850	8.678	1.550	8.496	8.939	0.112	0.544	0.757	0.439	0.975	3.405	4.363	6.692	0.589	0.445	0.971						
14 - 15	1.673	0.673	1.935	8.820	1.350	8.605	10.202	0.232	0.590	0.981	0.199	1.092	3.692	5.371	7.992	0.528	0.562	0.799						
15 - 16	1.622	0.696	2.027	8.506	1.250	8.332	9.759	0.294	0.609	0.970	0.275	1.056	3.523	5.062	7.369	0.464	0.525	1.327						
16 - 17	1.468	0.669	1.834	7.143	1.150	9.335	10.353	0.272	0.670	0.865	0.306	1.214	3.922	5.110	7.256	0.428	0.144	1.159						
17 - 18	1.271	0.620	1.880	8.032	1.150	8.247	9.700	0.328	0.768	0.971	0.305	1.409	3.873	6.814	9.219	0.500	0.705	0.781						
18 - 19	1.190	0.614	2.241	10.608	1.350	6.956	8.915	0.472	0.804	0.989	0.199	1.369	3.795	5.823	8.294	0.382	0.697	0.764						
19 - 20	1.168	0.627	2.268	10.450	1.150	7.641	9.279	0.466	0.821	0.355	0.955	1.336	3.771	5.559	7.753	0.329	0.028	2.221						
> 20	0.981	0.558	2.210	10.952	1.350	8.010	9.883	0.485	0.984	0.854	0.286	1.774	4.549	6.436	8.813	0.324	0.636	0.628						

TABLE B.2 : Les résultats de calibrage sur le Site N°1, VL, jour-sec

$Q_6/100$	Statistiques											Γ-GQM					DG				
	véh/h	σ_H	CV	S_3	K_4	Mo	α	β	θ	λ	P_{K-S}	A_2	α_1	α_2	β_1	β_2	θ	P_{K-S}	A_2		
2 - 3*	12.213	0.853	1.358	4.344	2.550	0.916	1.626	0.013	0.079	0.924	0.340	0.267	8.983	0.183	1.848	0.179	0.818	0.288			
3 - 4	10.015	0.955	1.949	7.897	1.150	2.652	4.494	0.113	0.101	0.594	0.260	0.128	1.829	0.594	2.405	0.630	0.822	0.448			
4 - 5	8.549	1.083	3.920	34.433	1.850	2.621	5.224	0.256	0.126	0.699	0.344	0.159	1.792	1.213	3.546	0.598	0.854	0.549			
5 - 6	6.633	1.012	3.332	23.950	2.050	2.311	4.847	0.315	0.154	0.986	0.269	0.209	1.898	1.800	4.588	0.616	0.903	0.471			
6 - 7	4.842	0.859	2.570	14.314	4.550	5.769	7.766	0.065	0.218	0.751	0.381	0.289	2.365	1.058	3.404	0.486	0.550	0.784			
7 - 8	3.838	0.804	1.973	8.592	1.850	8.250	8.672	0.009	0.266	0.785	0.330	0.410	2.631	2.230	5.160	0.600	0.851	0.609			
8 - 9	3.173	0.752	1.934	8.544	1.350	7.072	8.470	0.045	0.316	0.811	0.333	0.501	2.985	1.785	4.410	0.502	0.3330	1.151			
9 - 10	2.965	0.778	2.486	14.673	1.550	5.981	8.134	0.139	0.351	0.716	0.358	0.530	2.806	2.279	5.013	0.521	0.193	1.766			
10 - 11*	2.904	0.839	8.983	217.014	1.350	8.656	9.484	0.047	0.403	0.644	0.778	0.610	2.841	2.339	4.975	0.528	0.158	Inf			
11 - 12	2.176	0.693	1.923	9.376	1.450	7.208	9.205	0.158	0.452	0.586	0.741	0.825	3.277	4.884	8.090	0.641	0.508	0.980			
12 - 13	1.971	0.681	1.994	9.288	1.150	10.975	11.368	0.047	0.513	0.710	0.209	0.898	3.538	3.593	6.393	0.515	0.302	1.454			
13 - 14	1.696	0.632	1.803	8.430	1.650	8.447	9.978	0.114	0.589	0.634	0.274	1.110	3.874	4.668	7.708	0.562	0.670	0.686			
14 - 15	1.590	0.634	2.427	14.570	1.250	11.363	12.299	0.058	0.661	0.796	0.352	1.126	3.916	4.133	7.214	0.440	0.247	1.525			
15 - 16	1.375	0.590	2.087	10.286	1.550	5.977	8.372	0.293	0.759	0.875	0.351	1.209	3.949	3.796	6.751	0.372	0.664	0.995			
16 - 17	1.224	0.557	2.154	10.577	1.650	5.105	8.149	0.518	0.800	0.748	0.599	1.265	4.065	4.077	7.308	0.286	0.490	1.180			
17 - 18	1.352	0.656	5.749	101.616	1.550	6.414	8.907	0.486	0.766	0.820	0.449	1.184	3.724	4.596	7.351	0.298	0.362	1.350			
18 - 19	1.048	0.536	2.148	10.318	1.250	6.860	9.920	0.540	0.900	0.736	0.266	1.654	4.926	5.533	8.781	0.266	0.666	0.531			
19 - 20	1.044	0.555	2.158	10.276	1.250	8.575	11.239	0.486	0.904	0.990	0.134	1.681	4.463	7.399	10.573	0.367	0.864	0.408			
> 20	1.064	0.596	2.545	12.662	1.050	19.268	19.455	0.222	1.001	0.991	0.190	1.551	4.251	6.312	8.803	0.291	0.755	0.549			

TABLE B.3 : Les résultats de calibrage sur le Site N°1, VL, jour-pluie

$Q_6/100$	Statistiques										Γ-GQM										DG			
	σ_H	CV	S_3	K_4	Mo	α	β	θ	λ	P_{K-S}	A_2	α_1	α_2	β_1	β_2	θ	P_{K-S}	A_2						
2-3*	14.514	1.176	3.839	28.875	2.050	2.480	5.902	0.249	0.075	0.780	0.490	0.096	1.558	1.711	4.952	0.711	0.188	1.388						
3-4	9.424	0.910	2.001	8.623	2.650	7.446	11.494	0.052	0.108	0.851	0.334	0.152	2.036	0.983	3.631	0.686	0.268	2.102						
4-5	7.052	0.861	1.966	8.705	2.250	3.486	6.340	0.103	0.141	0.900	0.157	0.212	2.191	1.376	4.240	0.704	0.252	1.203						
5-6	5.322	0.814	2.004	9.874	3.150	5.075	6.912	0.037	0.186	0.767	0.431	0.293	2.538	1.337	4.015	0.623	0.874	0.593						
6-7	5.205	0.934	3.666	37.609	1.550	4.139	6.702	0.176	0.208	0.335	1.212	0.276	2.189	1.396	3.846	0.545	0.405	2.383						
7-8	4.151	0.859	4.135	53.384	1.350	7.302	8.433	0.036	0.262	0.638	0.517	0.360	2.485	1.405	3.878	0.500	0.384	1.595						
8-9	3.248	0.760	2.004	9.217	1.350	8.623	9.985	0.036	0.309	0.793	0.250	0.504	2.918	2.140	4.955	0.564	0.288	1.708						
9-10	2.779	0.730	2.108	10.647	1.850	5.544	7.544	0.112	0.363	0.606	0.398	0.608	2.959	2.835	5.805	0.623	0.488	1.167						
10-11*	2.631	0.763	2.902	23.228	1.250	12.805	12.856	0.031	0.396	0.976	0.188	0.657	3.031	3.194	6.043	0.572	0.161	1.752						
11-12	2.268	0.722	2.009	9.135	1.550	7.017	8.752	0.153	0.447	0.874	0.207	0.730	3.185	3.075	5.752	0.511	0.300	1.294						
12-13	1.926	0.671	1.793	7.642	1.150	10.122	10.886	0.068	0.519	0.786	0.317	0.927	3.736	3.480	6.242	0.482	0.235	1.310						
13-14	1.806	0.678	2.085	10.330	1.150	7.053	9.060	0.229	0.558	0.953	0.201	0.954	3.485	4.141	6.949	0.499	0.383	1.145						
14-15	1.668	0.666	1.861	7.699	1.250	9.854	11.613	0.214	0.593	0.835	0.305	1.091	3.723	5.632	8.521	0.522	0.322	1.493						
15-16	1.496	0.638	1.920	8.056	1.050	8.044	10.140	0.277	0.667	0.864	0.316	1.193	4.019	4.752	7.550	0.424	0.290	1.187						
16-17	1.285	0.587	2.069	10.208	1.350	7.283	9.556	0.302	0.797	0.945	0.230	1.426	4.416	4.785	7.692	0.390	0.769	0.549						
17-18	1.239	0.596	2.182	10.791	1.550	6.633	8.973	0.423	0.795	0.988	0.140	1.385	4.112	5.134	7.925	0.375	0.797	0.559						
18-19	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-						
19-20	0.891	0.479	1.884	9.257	1.250	6.387	9.317	0.556	1.104	0.671	0.695	1.988	4.581	7.818	11.768	0.446	0.928	0.432						
> 20	0.892	0.513	1.775	7.238	1.350	7.542	9.741	0.530	1.056	0.987	0.173	1.936	4.699	6.819	9.483	0.334	0.926	0.208						

TABLE B.4 : Les résultats de calibrage sur le Site N°1, VL, soir-sec

$Q_6/100$	Statistiques										Γ-GQM										DG			
	σ_H	CV	S_3	K_4	M_0	α	β	θ	λ	P_{K-S}	A_2	α_1	α_2	β_1	β_2	θ	P_{K-S}	A_2						
3-4	8.701	0.880	2.032	9.156	3.250	5.595	10.280	0.073	0.115	0.950	0.244	0.1770	2.244	1.473	4.963	0.700	0.932	0.283						
4-5	6.948	0.861	1.928	7.833	2.750	2.795	6.526	0.180	0.143	0.885	0.226	0.210	2.310	1.190	4.283	0.605	0.542	0.953						
5-6	5.683	0.867	2.083	8.898	1.750	9.199	13.127	0.067	0.182	0.323	0.633	0.255	2.511	1.104	3.738	0.491	0.442	1.022						
6-7	4.759	0.847	2.631	17.083	2.050	3.394	7.057	0.254	0.211	1.000	0.142	0.330	2.596	1.956	5.342	0.563	0.861	0.558						
7-8	3.983	0.822	2.820	20.666	2.550	5.476	6.637	0.013	0.272	0.314	0.975	0.360	2.590	1.425	4.151	0.451	0.546	0.922						
8-9	3.209	0.748	1.903	7.663	1.550	10.560	12.673	0.030	0.314	0.942	0.175	0.507	3.103	2.003	5.026	0.494	0.814	0.587						
9-10	2.704	0.710	1.951	8.598	1.850	10.696	12.705	0.035	0.369	0.555	0.688	0.671	3.265	3.719	7.412	0.631	0.243	1.821						
10-11*	2.412	0.704	1.797	7.154	1.650	5.963	7.249	0.070	0.420	0.953	0.202	0.673	3.292	2.198	4.803	0.460	0.788	0.450						
11-12	2.282	0.724	3.208	28.186	1.650	5.102	7.859	0.258	0.461	0.984	0.227	0.720	3.140	3.023	6.126	0.482	0.525	1.225						
12-13	1.911	0.664	1.845	8.125	1.650	11.356	11.823	0.046	0.520	0.899	0.278	0.968	3.731	3.930	6.806	0.539	0.783	0.755						
13-14	1.843	0.686	3.027	25.166	1.750	5.189	8.022	0.375	0.548	0.996	0.137	0.903	3.454	3.794	6.919	0.431	0.796	0.735						
14-15	1.637	0.655	2.043	8.973	1.350	6.567	9.212	0.342	0.599	0.822	0.214	1.050	3.625	5.040	8.252	0.475	0.327	1.147						
15-16	1.512	0.649	2.207	10.928	1.350	6.402	8.702	0.368	0.652	0.840	0.353	1.122	3.615	5.105	8.008	0.461	0.431	1.094						
16-17	1.245	0.566	1.987	8.793	1.450	12.477	13.730	0.068	0.847	0.734	0.361	1.447	4.697	4.773	8.078	0.327	0.560	0.762						
17-18	1.257	0.593	1.616	6.194	0.950	20.000	18.979	0.089	0.779	0.967	0.171	1.638	4.771	6.509	9.092	0.476	0.880	0.386						
18-19	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-					
19-20	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-					
> 20	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-					

TABLE B.5 : Les résultats de calibrage sur le Site N°1, VL, soir-pluie

$Q_6/100$	Statistiques										Γ-GQM										DG				
	σ_H	CV	S_3	K_4	Mo	α	β	θ	λ	P_{K-S}	A_2	α_1	α_2	β_1	β_2	θ	P_{K-S}	A_2							
0-2	56.906	1.770	4.952	39.120	1.050	0.934	2.212	0.325	0.023	0.000	15.769	-	-	-	-	-	-	-	-						
2-3*	35.433	2.316	10.015	139.211	0.850	1.317	2.742	0.412	0.045	0.000	13.439	-	-	-	-	-	-	-	-						
3-4	16.198	1.581	10.082	204.599	1.050	2.506	3.754	0.302	0.080	0.002	4.264	-	-	-	-	-	-	-	-						
4-5	9.443	1.173	2.349	10.454	1.250	2.800	4.135	0.312	0.105	0.312	1.370	0.126	1.432	2.002	3.560	0.655	0.058	2.946	2.946						
5-6	7.568	1.182	2.307	9.585	1.050	2.559	3.886	0.384	0.126	0.049	1.840	0.153	1.541	1.677	3.152	0.550	0.097	2.850	2.850						
6-7	6.452	1.157	2.378	10.436	0.850	3.405	4.613	0.365	0.151	0.005	2.598	0.186	1.682	1.842	3.292	0.521	0.025	3.605	3.605						
7-8	5.579	1.152	2.466	11.104	1.150	4.039	5.245	0.392	0.172	0.179	1.702	0.221	1.681	2.810	4.309	0.544	0.045	3.242	3.242						
8-9	4.774	1.116	2.391	10.009	0.850	4.555	5.660	0.383	0.203	0.002	2.780	0.258	1.805	2.492	3.968	0.496	0.058	4.072	4.072						
9-10	4.611	1.195	3.502	24.072	1.050	3.963	5.002	0.439	0.216	0.069	2.049	0.268	1.625	3.171	4.566	0.523	0.010	4.056	4.056						
10-11*	4.029	1.156	3.018	15.414	0.850	4.631	5.730	0.463	0.239	0.113	2.365	0.308	1.796	3.348	4.772	0.467	0.039	4.324	4.324						
11-12	3.508	1.119	3.121	16.668	0.850	6.298	6.947	0.413	0.289	0.029	2.922	0.369	1.873	3.883	5.177	0.482	0.017	5.516	5.516						
12-13	3.214	1.111	3.168	17.801	0.950	5.446	6.306	0.486	0.296	0.072	2.016	0.388	1.920	3.859	5.147	0.431	0.048	3.994	3.994						
13-14	2.840	1.060	3.082	16.564	0.950	6.059	7.012	0.494	0.333	0.024	2.081	0.446	2.086	4.098	5.490	0.401	0.021	4.167	4.167						
14-15	2.541	1.015	3.338	20.835	0.850	5.760	6.739	0.517	0.362	0.472	0.936	0.518	2.139	4.960	6.332	0.430	0.076	2.528	2.528						
15-16	2.426	1.036	4.888	49.717	0.950	4.835	6.129	0.589	0.383	0.112	1.542	0.510	2.118	3.946	5.509	0.343	0.028	3.156	3.156						
16-17	2.096	0.961	3.892	26.160	1.050	4.439	5.914	0.666	0.393	0.188	1.702	0.534	2.107	4.182	5.891	0.305	0.076	2.901	2.901						
17-18	2.821	1.313	19.391	604.533	1.050	4.565	6.133	0.671	0.409	0.120	Inf	0.489	1.978	3.740	5.488	0.265	0.015	Inf	Inf						
18-19	1.550	0.794	3.353	21.610	1.050	7.358	8.592	0.523	0.609	0.515	1.095	0.924	2.869	5.768	7.507	0.360	0.164	2.777	2.777						
19-20	1.417	0.770	3.904	30.323	1.050	6.068	7.605	0.636	0.621	0.295	1.322	0.900	2.791	5.150	6.980	0.278	0.071	2.708	2.708						
> 20	1.167	0.672	3.479	27.607	1.150	7.972	9.565	0.578	0.784	0.612	0.562	1.267	3.556	6.286	8.271	0.283	0.149	1.587	1.587						

TABLE B.6 : Les résultats de calibrage sur le Site N°1, VR, jour-sec

$Q_6/100$	Statistiques										I-GQM						DG			
	σ_H	CV	S_3	K_4	Mo	α	β	θ	λ	P_{K-S}	A_2	α_1	α_2	β_1	β_2	θ	P_{K-S}	A_2		
0 - 2	83.200	1.484	2.927	13.163	0.750	0.228	1.312	0.332	0.013	0.165	1.344	-	-	-	-	-	-	-	-	
2 - 3*	17.215	1.089	2.162	10.029	1.050	5.941	7.926	0.152	0.059	0.910	0.208	2.038	3.977	0.072	1.408	0.217	0.985	0.269	0.893	
3 - 4	14.078	1.327	3.864	33.469	1.350	2.541	4.523	0.343	0.074	0.333	0.661	0.084	1.325	1.619	3.530	0.619	0.573	0.893	0.893	
4 - 5	9.324	1.158	2.364	10.383	0.850	4.092	5.878	0.267	0.111	0.028	1.987	3.541	47.483	0.127	1.021	0	0.000	20.377	20.377	
5 - 6	7.572	1.164	2.481	10.698	1.050	2.601	4.565	0.397	0.127	0.392	1.202	0.161	1.636	1.882	3.887	0.548	0.108	2.358	2.358	
6 - 7	6.764	1.217	2.908	14.923	1.050	2.815	4.448	0.418	0.147	0.007	3.335	0.171	1.580	1.697	3.387	0.491	0.023	4.696	4.696	
7 - 8	5.732	1.185	2.965	16.003	1.450	4.912	6.800	0.398	0.174	0.036	2.548	0.221	1.685	3.297	5.335	0.537	0.005	4.734	4.734	
8 - 9	5.149	1.202	3.107	16.647	1.050	4.195	5.833	0.447	0.191	0.010	2.596	0.235	1.675	2.863	4.674	0.483	0.019	4.568	4.568	
9 - 10	4.483	1.162	3.081	15.925	1.150	2.777	4.421	0.544	0.201	0.078	2.589	0.252	1.805	2.076	3.792	0.380	0.006	4.213	4.213	
10 - 11*	3.963	1.143	3.945	30.701	1.150	4.586	6.138	0.469	0.250	0.220	1.505	0.328	1.900	3.298	5.084	0.453	0.128	3.283	3.283	
11 - 12	3.405	1.088	3.514	24.228	1.050	3.672	5.286	0.549	0.266	0.226	1.295	0.353	1.888	3.261	5.110	0.414	0.192	2.393	2.393	
12 - 13	2.682	0.929	2.674	12.071	1.150	4.985	7.068	0.494	0.345	0.533	0.721	0.497	2.449	3.473	5.684	0.380	0.378	1.641	1.641	
13 - 14	2.562	0.957	3.630	24.290	1.150	4.055	6.097	0.590	0.350	0.599	1.009	0.475	2.282	3.183	5.322	0.321	0.233	1.917	1.917	
14 - 15	2.316	0.937	5.966	70.945	1.250	2.875	5.148	0.782	0.320	0.155	1.324	0.421	2.247	2.570	4.837	0.171	0.111	1.773	1.773	
15 - 16	1.565	0.666	5.319	66.661	1.550	3.873	7.280	0.758	0.515	0.543	0.374	0.749	3.016	3.552	6.983	0.186	0.314	0.801	0.801	
16 - 17	1.822	0.836	4.970	48.084	1.250	4.371	6.803	0.722	0.446	0.560	0.919	0.610	2.599	3.660	6.101	0.198	0.199	1.739	1.739	
17 - 18	1.349	0.651	3.088	18.439	1.250	4.585	7.260	0.670	0.614	0.315	0.902	0.906	3.288	3.869	6.578	0.194	0.103	1.860	1.860	
18 - 19	1.366	0.700	5.878	69.617	1.250	4.155	6.752	0.828	0.526	0.404	0.811	0.673	2.631	3.685	6.263	0.114	0.153	1.596	1.596	
19 - 20	1.114	0.602	3.028	19.291	1.250	6.378	8.880	0.633	0.801	0.893	0.280	1.278	3.866	5.186	7.839	0.224	0.586	0.659	0.659	
> 20	1.138	0.648	3.236	21.435	1.250	6.111	8.085	0.685	0.725	0.995	0.157	1.132	3.133	5.886	8.157	0.269	0.930	0.297	0.297	

TABLE B.7 : Les résultats de calibrage sur le Site N°1, VR, jour-pluie

$Q_6/100$	Statistiques										I-GQM						DG			
	σ_H	CV	S_3	K_4	Mo	α	β	θ	λ	P_{K-S}	A_2	α_1	α_2	β_1	β_2	θ	P_{K-S}	A_2		
0-2	47.779	1.347	3.650	28.606	1.150	1.254	2.729	0.217	0.024	0.035	3.444	-	-	-	-	-	-	-		
2-3	18.681	1.235	2.551	13.383	1.250	2.237	3.798	0.261	0.055	0.020	2.662	0.060	1.187	1.635	3.314	0.735	0.013	3.566		
3-4	13.270	1.260	3.188	23.610	0.850	2.207	3.736	0.323	0.077	0.084	2.006	3.745	19.225	0.083	0.874	0	0.000	41.580		
4-5*	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
5-6	8.155	1.236	3.311	24.871	1.050	2.963	4.512	0.368	0.125	0.046	2.291	0.148	1.487	1.959	3.643	0.578	0.027	3.595		
6-7	6.966	1.235	2.671	12.285	1.150	2.875	4.523	0.447	0.136	0.034	2.490	0.164	1.542	2.124	3.868	0.505	0.017	3.996		
7-8	5.764	1.204	4.076	42.262	1.050	2.995	4.624	0.454	0.168	0.144	2.160	0.216	1.724	2.237	3.989	0.484	0.080	3.552		
8-9	5.137	1.205	3.110	17.167	1.150	2.902	4.301	0.489	0.184	0.068	2.268	0.222	1.581	2.341	3.944	0.473	0.032	3.856		
9-10	4.419	1.157	3.285	19.403	1.250	3.458	5.160	0.520	0.206	0.322	1.056	0.269	1.737	3.129	5.076	0.455	0.135	2.458		
10-11	3.700	1.083	3.099	17.785	1.050	4.516	5.990	0.457	0.260	0.081	1.731	0.341	1.914	3.239	5.000	0.460	0.026	3.639		
11-12	3.366	1.073	3.360	19.854	1.150	4.072	5.685	0.512	0.280	0.229	2.280	0.365	1.987	3.019	4.842	0.399	0.012	4.368		
12-13	2.882	0.991	2.886	14.893	1.050	4.550	6.268	0.520	0.313	0.469	1.019	0.447	2.265	3.450	5.357	0.386	0.141	2.357		
13-14	2.722	1.021	3.343	18.217	1.150	3.752	5.435	0.613	0.318	0.063	2.581	0.410	2.127	2.831	4.617	0.291	0.004	4.552		
14-15	2.366	0.953	3.304	18.968	1.250	5.138	6.970	0.580	0.373	0.366	1.095	0.536	2.378	4.090	6.096	0.337	0.204	2.357		
15-16	2.064	0.881	3.229	17.838	1.050	4.311	6.110	0.602	0.430	0.167	2.241	0.557	2.499	3.013	4.924	0.249	0.006	4.395		
16-17	1.674	0.763	4.351	38.543	1.550	3.178	5.506	0.806	0.420	0.089	1.877	0.570	2.614	2.845	5.169	0.136	0.032	2.754		
17-18	1.454	0.699	4.651	40.058	1.550	3.685	6.401	0.846	0.450	0.153	0.893	0.611	2.633	3.391	6.092	0.112	0.038	1.651		
18-19	1.518	0.768	5.544	55.260	1.250	3.864	6.293	0.857	0.413	0.401	0.725	0.546	2.253	3.717	6.205	0.125	0.308	1.183		
19-20	1.056	0.576	3.640	31.649	1.250	5.285	8.056	0.780	0.714	0.985	0.245	1.115	3.263	5.238	8.213	0.195	0.910	0.319		
> 20	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		

TABLE B.8 : Les résultats de calibrage sur le Site N°1, VR, soir-sec

Q ₆ /100	Statistiques										Γ-QM										DG		
	σ _H	CV	S ₃	K ₄	M ₀	α	β	θ	λ	P _{K-S}	A ₂	α ₁	α ₂	β ₁	β ₂	θ	P _{K-S}	A ₂					
0-2	38.731	1.265	3.416	23.513	1.650	1.020	2.942	0.248	0.027	0.796	0.346	-	-	-	-	-	-	-	-				
2-3	16.948	1.132	1.957	7.459	1.750	0.975	2.616	0.330	0.055	0.721	0.417	0.071	1.604	0.707	2.333	0.609	0.586	0.431	0.431				
3-4	17.150	1.616	8.672	126.290	1.550	1.340	2.828	0.378	0.073	0.394	0.646	0.079	1.211	1.096	2.675	0.632	0.333	0.797	0.797				
4-5*	9.412	1.164	2.254	8.961	1.650	3.071	4.976	0.316	0.106	0.518	0.511	0.127	1.445	2.210	4.279	0.653	0.297	0.931	0.931				
5-6	7.231	1.122	2.689	14.261	1.050	6.605	8.086	0.220	0.150	0.103	1.158	0.172	1.683	1.403	3.111	0.558	0.259	1.149	1.149				
6-7	6.273	1.118	2.322	8.946	1.650	1.944	4.086	0.511	0.139	0.841	0.329	0.185	1.901	1.422	3.475	0.403	0.644	0.478	0.478				
7-8	5.164	1.095	2.393	9.757	1.450	2.981	5.013	0.468	0.175	0.909	0.291	0.234	1.779	2.597	4.827	0.497	0.525	0.742	0.742				
8-9	4.727	1.103	2.676	11.847	1.350	2.943	5.283	0.545	0.183	0.956	0.286	0.254	1.977	2.432	4.785	0.399	0.755	0.580	0.580				
9-10	4.043	1.068	2.545	10.482	1.450	2.994	5.143	0.573	0.207	0.787	0.468	0.296	2.118	2.476	4.653	0.361	0.539	0.780	0.780				
10-11	4.028	1.168	3.887	28.073	1.050	2.936	4.270	0.523	0.239	0.603	0.796	0.285	1.764	2.047	3.538	0.385	0.507	1.133	1.133				
11-12	3.253	1.062	3.391	17.757	1.250	2.984	5.187	0.678	0.243	0.553	0.608	0.327	2.074	2.545	4.747	0.267	0.368	0.958	0.958				
12-13	3.382	1.177	7.735	130.648	1.050	3.836	5.617	0.601	0.283	0.736	0.554	0.371	1.939	3.263	5.190	0.353	0.531	1.066	1.066				
13-14	2.364	0.879	3.262	18.098	1.450	3.906	6.549	0.643	0.352	0.284	1.023	0.513	2.593	3.205	5.834	0.268	0.066	1.900	1.900				
14-15	2.096	0.839	3.891	28.368	1.350	4.128	6.807	0.657	0.404	0.468	0.802	0.581	2.545	3.654	6.435	0.281	0.282	1.281	1.281				
15-16	1.846	0.789	2.565	11.519	1.150	5.896	7.706	0.458	0.524	0.780	0.562	0.739	2.908	3.591	5.717	0.320	0.274	0.915	0.915				
16-17	0.800	0.365	1.086	4.695	2.150	7.449	12.362	0.201	1.500	0.964	0.257	3.333	9.136	5.938	11.433	0.327	0.946	0.261	0.261				
17-18	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-				
18-19	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-				
19-20	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-				
> 20	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-				

TABLE B.9 : Les résultats de calibrage sur le Site N°1, VR, soir-pluie

Index

- approche
 - de statistique partielle, 224
 - de calibrage, 154
 - de classement, 154
 - de régression, 225
 - de zonage, 232
 - stockage, 226
- bootstrap, 147
- comportement
 - collectif, 23, 37
 - de l'ensemble, 33, 37, 151
 - individuel, 23, 37, 193
 - standard, 170, 188, 218, 223
- condition
 - météo, 166
 - pluie, 166
 - sec, 166
- conducteur
 - suiveur, 152
- contexte, 223
- critère
 - statistique, 38
- dépasseur, 152
- diagramme
 - fondamental, 156
- discontinuité, 182
- distribution
 - histogramme, 47
 - méthode du noyau, 47
 - multi-distribution, 224, 232
 - queue, 195
- échantillon, 38
 - bootstrap, 148
 - fins, 44
 - glissant, 227
 - groupé, 44
 - horaire, 175
 - image, 148
 - initial, 148
 - non-stationnaires, 38
 - particulier, 64
 - primitif, 38
 - raffiné, 42, 44
 - représentatif, 34
 - sous échantillon, 227
 - stationnaires, 38
 - stationnarité, 49
- échantillonnage, 38
 - analyse de tendance, 38
 - méthode de groupement, 38
 - méthode de raffinement, 41
- ensemble
 - de choix, 151
 - test, 113
- estimateur, 105
- état
 - élémentaire, 156, 160
 - gestionné, 154
 - de poursuite, 196
 - du trafic, 154
 - fluide, 154
 - intermédiaire, 154
 - physique, 152
 - psychologique, 152
- facteur, 151
 - complexe, 189
 - de gêne, 170
 - de propulsion, 171
 - empêchant, 189, 207
 - exogènes, 152, 166
 - impulsif, 189
 - mésoscopique, 152
 - macroscopique, 152
 - microscopique, 152
 - propulsif, 207
 - simple, 189
 - sous-facteurs, 152
- fonction
 - aléatoire, 48
 - de survie, 48
 - du noyau, 47
 - de déplacement, 76

- de vraisemblance, 107
- groupe, 193
- interdépendance, 165
- invalidation
 - probabiliste, 228
 - statistique, 227
- mode, 166, 177, 184, 209
- modèle
 - complexe, 194
 - de mélange, 100
 - exponentiel, 86
 - gamma, 88
 - mixte, 130
 - spécifique, 154
- modèle combiné, 94
- modèle
 - automate cellulaire, 81
 - complexe, 112
 - de Heidemann, 75
 - de poursuite, 23
 - macroscopique, 21
 - probabiliste, 85
 - simple, 32, 86
- peloton, 23, 40, 76, 153, 193
- période
 - d'agrégation, 39
 - temporelle, 39
- phase, 157, 167
 - changement de, 157, 167
 - d'accélération, 25
 - de conduite libre, 24
 - de décélération, 25
 - de poursuite, 24
 - de transition, 24
- poursuite
 - catégorie de, 152
 - fonction d'équilibre, 25
 - modèle GHR, 24
 - sensibilité, 24
 - stimuli, 24
- rééchantillonnage, 147
- régime, 154
 - régularité, 174
- statistique
 - descriptive, 166
- temps
 - d'ajustement, 115
 - d'application, 24
 - d'estimation, 115
 - d'exécution, 114
 - d'occupation, 29
 - de décision, 24
 - de perception, 24
 - de réaction, 24
- TIV
 - d'équilibre, 24
 - de relaxe, 129
 - gardé, 129
 - particulier, 228
 - préférée, 196
 - TIV embarqué, 28
 - TIV net, 28
 - TIV tracé, 28
 - zéro, 38, 124
- traceur, 32
- trafic
 - fluide, 160
- trajectoire, 21
- validation
 - niveau de, 233
 - zone de, 233
- variable
 - spatiale, 22
 - temporelle, 22
- véhicule
 - baladeur, 205
 - contraint, 23, 32, 152, 164, 194
 - en congestion, 214
 - en tête, 23, 152, 195
 - hors congestion, 214
 - isolé, 205
 - leader, 25, 31, 205
 - libre, 23, 32, 96, 152, 194, 205
 - particulier, 152
 - poids lourd, 152
 - retardé, 96
 - suiveur, 25, 31, 153, 200, 205
 - traceur, 88
- vitesse
 - désirée, 25
 - optimale, 25