

Sur la capacité opérationnelle des modèles d'affectation dynamique du trafic, et la convergence des algorithmes d'équilibrage

Hai Dang Mai

▶ **To cite this version:**

Hai Dang Mai. Sur la capacité opérationnelle des modèles d'affectation dynamique du trafic, et la convergence des algorithmes d'équilibrage. Sciences de l'ingénieur [physics]. Ecole des Ponts ParisTech, 2006. Français. NNT : . pastel-00002183

HAL Id: pastel-00002183
https://pastel.hal.science/pastel-00002183

Submitted on 1 Sep 2010

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.

Thèse de doctorat de l'Ecole Nationale des Ponts et Chaussées

Laboratoire Ville Mobilité Transports

Spécialité TRANSPORT

Présentée et soutenue publiquement
le 18 décembre 2006 par

Hai Dang MAI

Sur la capacité opérationnelle des modèles d'affectation dynamique du trafic, et la convergence des algorithmes d'équilibrage

Devant le jury composé de Messieurs

Marcel MIRAMOND, Professeur des universités, INSA de Lyon, président de jury et rapporteur.
Habib HAJ-SALEM, Directeur de recherche, HDR, INRETS-GRETIA, rapporteur.
Jean LATERRASSE, Professeur des universités, UMLV-LVMT, examinateur.
Fabien LEURENT, enseignant-chercheur, HDR, ENPC-LVMT, examinateur.
Vincent AGUILERA, enseignant-chercheur, Docteur, ENPC-LVMT, examinateur.
Patrick MALLEJACQ, Adjoint au Chef du CSTR, SETRA, examinateur.

Thèse financée par le Sétra

Sous la direction de :
Jean LATERRASSE et **Fabien LEURENT**

A mes parents

Remerciements

Mes remerciements vont d'abord et avant tout à mon directeur de thèse, le professeur Jean Laterrasse, ainsi qu'à mon co-directeur de thèse, Monsieur Fabien Leurent, sans lesquels je n'aurais pu mener à bien mes travaux.

J'ai bénéficié de leurs connaissances dans le domaine de la recherche et je leur suis reconnaissant pour m'avoir initié à leurs méthodes de recherche.

Je remercie également Monsieur Vincent Aguiléra qui m'a prodigué de précieux conseils au cours de mes recherches.

Je suis reconnaissant au gouvernement de la République Socialiste du Vietnam de m'avoir accordé une bourse qui m'a permis de continuer mes recherches dans ce beau pays qu'est la France.

Je remercie le SETRA d'avoir financé la réalisation de ma thèse.

Toutes les équipes du CSTR et des CETEs m'ont toujours apporté un soutien sans faille en me procurant toute la documentation scientifique dont j'avais besoin. Je n'oublierai jamais les sentiments d'amitié qu'elles m'ont témoignés pendant toute la durée de mes recherches.

Je remercie mes collègues du LVMT de m'avoir fait partager leurs connaissances et d'avoir bien voulu prendre de leur temps pour mener des discussions m'ayant permis d'avancer dans la rédaction de ma thèse.

Mes remerciements vont à tous mes amis qui m'ont soutenu pendant mon séjour en France.

Ma profonde reconnaissance enfin va à ma famille qui m'a accordé tout au long son soutien moral inébranlable.

Table des matières

Remerciements **5**

Introduction générale **21**

Contexte opérationnel, technique et scientifique 21

Le contexte opérationnel 21

Le contexte technique 22

Le contexte scientifique 23

Problématique et objectifs 24

Démarche et méthode 25

Contributions 26

Plan du mémoire 27

I Etat de développement de l'affectation dynamique 29

1 Etat de l'art de la simulation dynamique du trafic routier 31

1.1 Généralités sur la simulation dynamique du trafic routier . . . 32

1.1.1 Le problème d'affectation du trafic 33

1.1.2 Repères chronologiques 33

1.1.2.1 Modèles statiques 33

1.1.2.2 Modèles de simulation dynamique 34

1.1.3 Domaines d'application 35

1.2 Principes des modèles 36

1.2.1 Concernant l'écoulement 36

1.2.1.1 Modèles macroscopiques par e.d.p. 37

1.2.1.2 Modèles macroscopiques *analytiques* 38
1.2.1.3 Modèles microscopiques 38
1.2.1.4 Modèles mésoscopiques 39
1.2.1.5 Autres modèles d'écoulement 40
1.2.2 Concernant les déplacements 40
1.2.2.1 Notion de déplacement 40
1.2.2.2 Choix d'itinéraire 40
1.2.2.3 Choix de l'heure de départ 41
1.3 Exemples de modèles et outils logiciels 42
1.3.1 AIMSUN 42
1.3.1.1 Modèle de poursuite 43
1.3.1.2 Modèle de changement de voie 43
1.3.1.3 Gestion des intersections et voies d'accès . . . 44
1.3.1.4 Affectation du trafic 44
1.3.1.5 Capacité de traitement 45
1.3.2 METROPOLIS 46
1.3.3 STRADA 47
1.3.3.1 Modélisation de l'écoulement 48
1.3.3.2 Modélisation des intersection 49
1.3.3.3 Estimation du temps de parcours 50
1.3.3.4 Affectation du trafic 50
1.3.4 METACOR 51
1.3.4.1 Modélisation de l'écoulement 51
1.3.4.2 Modélisation de l'affectation 54
1.4 Modèles d'affectation dynamique fondés sur un modèle macroscopique analytique de l'écoulement 56
1.4.1 Modèle de Kuwahara et Akamatsu 57
1.4.2 Modèle de Bin Ran et al. 57
1.4.3 Modèle de Gentile et al. 59
1.4.4 LADTA 60

2 Etude de l'utilisation des modèles dynamiques en France 61
2.1 Objectif de l'étude 61

2.2 Méthode utilisée 61
2.2.1 Liste des destinataires de l'enquête 62
2.2.2 Période de réalisation de l'enquête 62
2.2.3 Liste des questions 62
2.3 Les acteurs du RST 63
2.3.1 Au CETE Nord-Picardie 63
2.3.2 Au CETE Méditerrannée 65
2.3.3 A l'INRETS 67
2.4 Commentaires et synthèse 70
2.4.1 Affectation dynamique/statique 70
2.4.2 Type de modèle utilisé 71
2.4.3 Echelle d'application 71
2.4.4 Constitution des données 72
2.4.5 Expériences d'application 72

II Prototypage du modèle LADTA 73

3 LADTA : un modèle analytique d'affectation dynamique du trafic 75
3.1 Introduction 75
3.2 Généralités sur le modèle LADTA 76
3.3 Variables du modèle 78
3.3.1 Objets et ensembles 78
3.3.2 Variables d'offre 79
3.3.3 Variables de demande 79
3.4 La formation des services 80
3.4.1 Principes physiques et économiques 80
3.4.2 Formulation mathématique 80
3.4.3 Algorithme de propagation sur un réseau dynamique . 81
3.5 Le choix des usagers 82
3.5.1 Principes économiques 82
3.5.2 Formulation mathématique 82
3.5.3 Un algorithme de répartition 84

3.6 Le chargement en volume 84
3.6.1 Principes physiques 85
3.6.2 Formulation mathématique du problème de chargement 85
3.6.3 Algorithme atomique de chargement 86
3.7 L'écoulement du flux 86
3.7.1 Principes physiques et économiques 86
3.7.2 Formulation mathématique 88
3.7.3 Calcul du temps de parcours 89
3.7.4 Algorithme de calcul du volume sortant 89
3.8 L'équilibre offre-demande 91
3.8.1 Formulation mathématique 91
3.8.2 Un algorithme de combinaisons convexes au niveau des arcs 91

4 Traitement d'un cas d'école **93**
4.1 Traitement manuel pour l'explication du fonctionnement du modèle 93
4.1.1 Traitement de l'exemple 94
4.1.1.1 Les hypothèses 94
4.1.1.2 Principe des calculs 95
4.1.1.3 Chargement de l'arc a_1 isolé 96
4.1.1.4 Etat d'équilibre 99
4.1.2 Commentaires sur la concurrence entre deux arcs . . . 104
4.1.2.1 Période de saturation 104
4.1.2.2 File d'attente 104
4.1.2.3 Temps global perdu 105
4.2 Traitement par le prototype 105
4.2.1 Codage du réseau 105
4.2.2 Vérification par étapes 106
4.2.2.1 Itération 1 106
4.2.2.2 Itération 2 107
4.2.3 Résultats à l'état de convergence 109
4.3 Série de tests 112

5 Elaboration du prototype **115**
5.1 Elaboration du prototype 115
5.1.1 Type de données 117
5.1.2 Fonctions utilitaires 117
5.1.2.1 Fonction de création d'un profil 117
5.1.2.2 Fonction d'évaluation 117
5.1.2.3 Fonction de réciproque 118
5.1.2.4 Fonction de différentiation 118
5.1.2.5 Fonction d'intégration 118
5.1.2.6 Fonction de cumulation 119
5.1.2.7 Fonction de prolongation 119
5.1.2.8 Fonction de coupure 119
5.1.2.9 Procédure d'enregistrement 119
5.1.2.10 Procédure d'édition graphique 119
5.2 Programmation de l'écoulement du flux 120
5.2.1 Fonction principale 120
5.2.2 Fonctions auxiliaires 121
5.3 Programmation de la formation des services 121
5.3.1 Fonction principale 121
5.3.2 Fonctions auxiliaires 122
5.4 Programmation du choix des usagers et de chargement en volume 123
5.4.1 Fonction principale 123
5.4.2 Fonctions auxiliaires 124
5.5 Programmation de l'équilibre offre-demande 125
5.5.1 Fonction principale 126
5.5.2 Fonctions auxiliaires 127
5.6 Le problème de fragmentation des profils 127
5.6.1 Position du problème 131
5.6.2 Pourquoi une fragmentation exponentielle des profils ? . 132
5.6.2.1 Sur la formation des chemins 133
5.6.2.2 Sur le choix des usagers 133
5.6.2.3 Sur le chargement en volume 133
5.6.2.4 Sur l'écoulement du flux 134

5.6.2.5 Sur l'équilibre offre-demande 134
5.6.3 Comment contrôler la fragmentation des profils ? 134
5.6.3.1 Fonction de simplification 134
5.6.3.2 Fonction de regroupement 135

III Formulations analytiques de l'équilibre dynamique et mesurage de la convergence 137

Introduction de partie 139
Contexte et problématique 139
Objectif de connaissance et questionnement 141
Plan de la partie 142

6 L'équilibre statique : formulation, obtention et mesurage 145
6.1 Hypothèses et notations 145
6.2 Formulation par arcs et par chemins 147
6.2.1 Conditions d'équilibre par arcs et destinations 147
6.2.2 Fonction de lagrangien et problème de point-selle . . . 148
6.2.3 Formulation extrémale primale par arcs et destinations 149
6.2.4 Formulation extrémale duale par arcs et destinations . 150
6.3 Formulation par chemins 153
6.3.1 Conditions d'équilibre par chemins 153
6.3.2 Fonction de lagrangien et problème de point-selle . . . 154
6.3.3 Formulation extrémale primale par chemins 155
6.3.4 Formulation extrémale duale par chemins 156
6.4 Problèmes-types et propriétés formelles 157
6.4.1 Expression générique pour l'affectation statique 157
6.4.2 Le problème de complémentarité non linéaire (NCP) . 158
6.4.3 Le problème d'inéquation variationnelle (VIP) 159
6.4.4 Le problème de point fixe (FPP) 160
6.4.5 Formulation extrémale primale (PEP) 160
6.4.6 Sur les formulations duales 161
6.5 Les correspondances entre formulations 162

6.5.1 Correspondance de dualité 162
6.5.2 La conservation du flux est identique par arcs et par chemins 162
6.5.3 Les inégalités de coût locales impliquent celles par chemins 163
6.5.4 Les écarts complémentaires sont identiques par arcs ou chemins 163
6.5.5 L'état optimum par arcs résout le problème par chemins 164
6.5.6 Un état optimum par chemins résout le problème par arcs 164
6.5.7 Commentaires 165
6.6 Algorithmes et calculabilité 167
6.6.1 Le problème auxiliaire : une affectation à coûts constants 167
6.6.2 Les méthodes de combinaison convexe 168
6.6.3 Algorithmes primaux traitant les chemins 169
6.6.4 Algorithmes duaux 170
6.6.5 Sur les algorithmes primaux-duaux 170
6.7 Critères pour mesurer un écart à l'équilibre 171
6.7.1 L'écart aux coûts minimum 171
6.7.2 L'écart primal-dual 173
6.7.3 Les écarts inter-itérations et leurs risques 174

7 Examen des modèles d'affectation dynamique 177
7.1 Nature de l'examen bibliographique, en lien avec le développement historique des modèles 177
7.2 Qualification de quelques modèles récents 179
7.3 Discussion 181

8 Formulations analytiques de l'équilibre dynamique et mesurage de la convergence 183
8.1 Relativité chronologique et correspondance temps-volume . . . 184
8.1.1 Notations et variables pour l'affectation dynamique . . 185
8.1.2 Structuration de l'affectation dynamique 187

8.1.3 Sur la démarche d'équilibrage dans le modèle LADTA . 188
8.1.4 La relativité chronologique du chargement en volume des arcs 189
8.1.5 La correspondance temps-volume est relative à une chronologie 190
8.1.6 Sur la concordance temps-volume 191
8.1.6.1 Définition de la concordance temps-volume . . 191
8.1.6.2 Propriétés de la concordance 192
8.1.6.3 Imposer la concordance 192
8.1.7 Sur la concordance "duale" volume-temps 193
8.1.8 La concordance des énergies par arcs et par chemins . . 194
8.1.9 Récapitulation 195
8.2 Sur la formulation de l'équilibre dynamique 196
8.2.1 Sur la formulation par chemins 197
8.2.1.1 Formulation relaxée de l'équilibre dynamique 197
8.2.1.2 Critère de convergence 198
8.2.1.3 Choisir une base de variables endogènes . . . 199
8.2.2 Formulations par arcs et destinations 202
8.2.2.1 Expression précise par arcs et destinations . . 202
8.2.2.2 Sur la relativité chronologique et les trois types de temps 204
8.2.2.3 Expression précise distinguant quatre types de temps 204
8.2.2.4 Expression générique 205
8.2.2.5 Propriétés de concordance 206
8.2.2.6 Un critère de convergence global et ses adaptations 207
8.2.2.7 Sur les approches algorithmiques 209
8.2.2.8 Formulation primale-duale avec liaison des potentiels 210
8.2.2.9 Formulation primale-duale avec relaxation des potentiels 211
8.2.3 Une cohérence temporelle pour des formulations par arcs 212

8.2.3.1 Enjeu et principe 212
8.2.3.2 Méthode d'inférence chronologique 213
8.2.3.3 Critère de convergence 214
8.3 Sur les algorithmes d'affectation dynamique 215
8.3.1 Combinaison convexe au niveau des arcs 216
8.3.1.1 Organigramme 216
8.3.1.2 Discussion 216
8.3.1.3 Critère global par chemins explicites 218
8.3.1.4 Critère global par chemins implicites 219
8.3.1.5 Critère global par arcs et destinations 219
8.3.2 Combinaison convexe par chemins avec séparation du chargement et de l'écoulement 220
8.3.2.1 Organigramme 220
8.3.2.2 Critère de convergence 220
8.3.2.3 Discussion 222
8.3.3 Combinaison convexe par chemins avec intégration du chargement et de l'écoulement 223
8.3.4 Quels autres algorithmes primaux sont possibles ? . . . 224
8.3.5 Algorithme hybride 224
8.3.5.1 Organigramme 225
8.3.5.2 Discussion 225
8.3.6 Algorithmes duaux 225
8.3.7 Algorithmes primaux-duaux 227
8.4 Critique du modèle LADTA : expression, formulation, critère et algorithme 228
8.4.1 Sur l'expression du modèle LADTA 228
8.4.2 Sur la formulation mathématique en problème de point fixe 229
8.4.3 Algorithmes atomiques pour traiter les chemins 229
8.4.3.1 Représenter des chemins par un faisceau d'atomes 230
8.4.3.2 Algorithme atomique de chargement en volume 231
8.4.3.3 Algorithme atomique pour actualiser un flot par chemins 231

8.4.4 Sur la complexité informatique 232
8.4.4.1 Qualification des fonctions 233
8.4.4.2 Evaluation des algorithmes 235

9 Expériences numériques **237**
9.1 Le plan d'expérience et le cas d'école 237
9.1.1 Le plan d'expérience 237
9.1.2 Position du cas d'école 238
9.1.3 Solution du cas d'école 239
9.2 Comportement des algorithmes et des critères 240
9.2.1 La similarité des équilibres 240
9.2.2 Réglage de la portée d'assimilation temporelle 242
9.2.3 Réglage de la suite des coefficients 244
9.2.4 Sur la combinaison convexe par arcs 247
9.2.5 Sur la combinaison convexe par chemins 247
9.2.6 Sur l'algorithme hybride 250
9.2.7 Synthèse 250
9.2.8 Sur la composition des critères de convergence 252

Conclusion de partie **257**
Récapitulation 257
Perspectives 258

Conclusion générale **261**
Récapitulation du travail 261
Confrontation des résultats aux objectifs de connaissance 262
Perspectives 263

Bibliographie **265**

A Logiciels de simulation dynamique du trafic routier **275**

B Liste des entretiens effectués **279**

C Tests du prototype **281**

D Estimation de matrices O-D dynamiques **283**

E Test de la performance du simulateur **319**

Table des figures

3.1 Quatre relations causales. 77

4.1 Réseau de deux arcs concurrents. 94
4.2 Profil temporel du volume de la demande. 95
4.3 Ecoulement sur l'arc 1 isolé. 97
4.4 Concurrence entre deux arcs et instant critique h_1. 98
4.5 Ecoulement sur l'arc a_1 à l'équilibre. 101
4.6 Partage du trafic à l'équilibre. 102
4.7 Temps de parcours à l'équilibre. 103
4.8 Le réseau $TNET$. 106
4.9 Ecoulement sur l'arc a_1 à la 2ème itération. 107
4.10 Concurrence entre arcs à la 2ème itération. 108
4.11 Partage du volume à la 2ème itération. 109
4.12 Résultat de la combinaison convexe sur l'arc a_1 à la deuxième itération. 110
4.13 Résultat de la combinaison convexe sur l'arc a_2 à la deuxième itération. 111
4.14 Volumes entrants des arcs à la 51ème itération. 112
4.15 Temps de parcours à la 51ème itération. 113

5.1 Exemple de profil temporel affine par morceaux. 116
5.2 Temps de traversée à la 71ème itération. 128
5.3 Temps de traversée à la 201ème itération. 128
5.4 Temps de traversée à la 5ème itération. 129
5.5 Temps de traversée à la 51ème itération. 129

5.6 Augmentation du nombre de points du profil de volume en entrée de l'arc a_1. 130

9.1 Réseau à quatre arcs. 239
9.2 Les méthodes de CC produisent un même équilibre. 241
9.3 L'algorithme hybride produit aussi l'équilibre. 241
9.4 Similarité des rythmes d'équilibrage, méthodes de CC. 242
9.5 Similarité des rythmes d'équilibrage, hybride et CC par chemins. 243
9.6 Influence de la portée d'assimilation. 244
9.7 Influence de la suite des coefficients sur le CIVA_A. 245
9.8 Influence de la suite des coefficients sur le CECM_R. 246
9.9 Influence de la suite des coefficients sur le CECM_A. 246
9.10 Evolution de CECM pour la combinaison convexe par arcs. . . 247
9.11 Evolution de CB pour la combinaison convexe par arcs. 248
9.12 Evolution de CIVA pour la combinaison convexe par arcs. . . 248
9.13 Evolution de CECM pour la combinaison convexe par chemins. 249
9.14 Evolution de CB pour la combinaison convexe par chemins. . . 249
9.15 Evolution de CIVA pour la combinaison convexe par chemins. 250
9.16 Evolution de CECM pour l'algorithme hybride. 251
9.17 Evolution de CBH_1 pour l'algorithme hybride. 251
9.18 Evolution de CBH_2 pour l'algorithme hybride. 252
9.19 Décroissance conjointe du CIVA et du CECM. 253
9.20 Totalisation des critères pour la CC par arcs. 254
9.21 Totalisation des critères pour la CC par chemins. 255
9.22 Totalisation des critères pour l'algorithme hybride. 255

C.1 Le réseau NET. 281
C.2 Le réseau $NET8$. 282
C.3 Le réseau MLV. 282

Introduction générale

Contexte opérationnel, technique et scientifique

Le contexte opérationnel

Sur un territoire, le développement social et économique d'une population passe par la diversification des biens et des services consommés, qu'ils soient marchands ou non, et par la spécialisation des productions individuelles. Chaque individu prend part à un jeu d'échanges, souvent aussi de production, pour des biens et services variés, en s'impliquant dans des activités variées qui l'amènent à fréquenter des lieux variés, et donc à se déplacer entre ces lieux. La demande de déplacement ainsi constituée s'exerce sur des réseaux de transport, dont le développement va de pair avec le développement général de la société et de son économie.

Le développement d'un réseau de transport, en structure sur le territoire, en capacité d'écoulement, et en qualité de service, requiert une planification : en effet, les infrastructures comme les véhicules d'exploitation sont des éléments matériels à durée de vie longue et au coût de revient très élevé, et les investissements correspondants sont à amortir à long terme. La planification consiste à concevoir les investissements matériels et l'organisation des services, à les sélectionner, et à les répartir dans le temps en fonction des besoins et des ressources financières.

Les infrastructures les plus importantes en capacité de transport et en niveau de service, sont conçues et dimensionnées afin d'écouler un certain volume de trafic. Quand ce volume croît, ou aux périodes de forte concentration comme les pointes journalières en milieu urbain au moment des migrations

domicile-travail et les pointes saisonnières en milieu interurbain au moment des migrations touristiques, des phénomènes de congestion se mettent en œuvre, qui dégradent la qualité de service (évaluée par la vitesse du flot) et même, en cas de saturation, réduisent la capacité d'écoulement. La congestion, et sa forme aiguë la saturation, font l'objet d'une gestion dynamique du trafic par les exploitants des réseaux : intensification des fréquences de desserte en transport public, intervention rapide sur incident en transport public comme en transport privé, mesures d'exploitation comme l'orientation du flot sur certains itinéraires, ou la réduction de la vitesse limite afin d'homogénéiser le flot ou de le retarder, et même mesures à caractère économique comme l'orientation modale ou la modulation tarifaire.

Le contexte technique

Par cette description rapide, nous avons exprimé les enjeux concrets et opérationnels de la gestion des réseaux de transport, pour le terme long de la planification et pour le terme court de l'exploitation. Ces enjeux sont considérables au plan économique, et complexes au plan de la connaissance et de l'analyse : le trafic est formé d'un grand nombre de déplacements, aux lieux variés d'origine et de destination et de passage, avec divers comportements de circulation qui distinguent notamment les poids lourds et les voitures sur le mode routier, et divers comportements économiques dans les choix d'itinéraire, d'horaire, de mode, de se déplacer etc.

C'est pourquoi la planification fait l'objet d'études au caractère analytique poussé, appelées des *études de prévision de trafic*, et réalisées au moyen d'un instrument technique dédié : le modèle d'affectation du trafic, où sont représentées : d'une part l'offre de transport en termes d'infrastructures et de services localisées dans un territoire d'étude, avec pour chaque service des caractères physiques et économiques en temps, en prix, en qualité ; d'autre part la demande de déplacement, conçue comme des flux d'usagers allant des zones d'origine vers les zones de destination (selon un découpage zonal du territoire d'étude) auxquels on associe des comportements économiques ; et enfin la rencontre entre l'offre et la demande, par la simulation des choix des

usagers et du fonctionnement technique du réseau, qui détermine les volumes et les prix / coûts généralisés des services sur le marché du transport.

Les modèles d'affectation et les études de prévision du trafic ont été développés depuis les années 1950, et sont désormais couramment répandus dans les pays développés, grâce à une ingénierie-conseil spécialisée. Mais les modèles utilisés sont en général statiques, ils réduisent le temps chronologique à une période fictive de référence durant laquelle les variables de trafic sont réputées constantes : cela n'empêche pas de modéliser des durées de parcours et certains aspects de la congestion, mais cela réduit la portée des simulations à la planification du réseau et à la conception de schémas de circulation. Il n'est pas possible de simuler des variations temporelles, ni dans les flux origine-destination (O-D) et leurs choix d'horaire de départ, ni dans les commandes d'exploitation comme la répartition de capacité entre les axes confluant en une jonction, ni la modulation temporelle des tarifs, ni les aspects dynamiques de la congestion, la formation, le développement puis la dissipation des saturations.

Le contexte scientifique

La recherche en ingénierie et modélisation du transport et du trafic, a développé les modèles statiques, leur contenu physique et économique, leurs algorithmes et leurs applications, depuis le milieu des années 1950 [BMW56], et fortement dans la période de 1970 à 1995. La fin de cette période correspond aux modèles d'affectation multiclasses et multicritères notamment le modèle prix-temps pour simuler la concurrence entre itinéraires différenciés en prix et qualité de service [Leu98]. Les développements plus récents des modèles d'affectation statique concernent surtout les modes de transport collectif.

Les limitations inhérentes à l'hypothèse statique ont été rapidement comprises : dès 1978, Merchant et Nemhauser [MN78] ont proposé un modèle d'affectation dynamique, restreint à une zone d'origine et avec un traitement sommaire, quasi-statique, de la congestion. Le modèle précurseur ayant servi à des applications opérationnelles est CONTRAM de Leonard *et al.*[LTB78],

conçu pour simuler des sous-réseaux routiers urbains, sans principe d'équilibre entre l'offre et la demande de déplacement. Dans CONTRAM, chaque déplacement est simulé depuis son origine jusqu'à sa destination, en regroupant les véhicules par "paquets" vers la destination, traités à des instants successifs discrétisés. C'est pourquoi le modèle CONTRAM est qualifié de "mésoscopique", plutôt que de "microscopique" qui correspondrait au traitement particulier de chaque véhicule. Dans les années 1990 ont été développés plusieurs simulateurs microscopiques : les plus diffusés commercialement sont AIMSUN, CORSIM, PARAMICS et VISSIM. Mais ils restent dénués de principe d'équilibre, et en fin de compte leur domaine d'application privilégié reste celui des sous-réseaux routiers en milieu urbain.

Depuis 1995 ont été conçus des modèles d'affectation dynamique plus proches de l'affectation statique, avec un traitement macroscopique et une formulation analytique : les applications rapportées concernent le plus souvent de simples cas d'école, hormis pour des modèles récents dont l'application semble encore réservée aux équipes de recherche qui les développent : les modèles de Gentile *et al.* [BGP05], METROPOLIS de De Palma *et al.* [DPMN96], DYNAMEQ de Florian *et al.* [MMN05], LADTA de Leurent [Leu03] et Leurent et Aguiléra (2005).

Problématique et objectifs

Notre thèse veut répondre à la question centrale suivante : quelle est la portée opérationnelle des modèles d'affectation dynamique du trafic ? Jusqu'où s'étend leur domaine de pertinence, quelle est leur capacité de simulation, et pourraient-ils se répandre en tant qu'instrument technique dans la panoplie de l'ingénieur-conseil en transport, et dans la pratique courante des études de trafic ?

Ces questions ont motivé l'inspiration et le financement de la thèse par le Service Technique des Routes et des Transports (Sétra), service technique central du ministère des transports en France, qui a récemment renouvelé son modèle d'affectation statique du trafic, et veut évaluer l'opportunité de disposer également d'un modèle d'affectation dynamique, que pourraient utiliser

les centres d'études techniques du même ministère.

Pour répondre à cette problématique de la portée opérationnelle, nous l'avons déclinée en quatre objectifs de connaissance, dont voici l'énoncé :

1. quel est le contenu "sémantique" d'un modèle d'affectation dynamique, quelles en sont les hypothèses d'ordre physique et économique ? De manière connexe, comment interpréter en pratique chaque hypothèse ?
2. évaluer la calculabilité d'un modèle : cette question comprend deux volets, d'une part l'efficacité / effectivité de la simulation à savoir produit-elle effectivement un équilibre dynamique du trafic ; d'autre part l'efficacité / efficience de la simulation, au prix de quel volume de calcul et donc de quel temps de réponse de l'ordinateur pour un utilisateur ?
3. quels sont les besoins du modèle en données d'entrée, les exigences de son alimentation en inputs ?
4. comment analyser et interpréter les données traitées par le modèle, non seulement en entrée mais aussi et surtout en sortie, puisque le modèle produit en résultat des flux par itinéraire et horaire de départ, par classe de trafic et segment de comportement, par relation origine-destination ou tronçon du réseau etc.

Tels sont les objectifs de connaissance fixés à notre thèse. Nous ne traiterons que brièvement, la question des types d'application à simuler, car des travaux à ce sujet sont déjà disponibles (ex. Peeta et Ziliaskopoulos [PZ01]). Pour certains de nos objectifs, nous ne nous contenterons pas d'examiner l'état de la recherche et l'état de la pratique : nous nous attacherons à développer le modèle d'affectation du LVMT, appelé LADTA pour **L**umped **A**nalytical **D**ynamic **T**raffic **A**ssignment.

Démarche et méthode

Afin de répondre aux objectifs de connaissance, nous avons mené d'une part une analyse bibliographique des modèles de simulation de trafic et des modèles dynamiques d'affectation ; et d'autre part une enquête sur leur usage, auprès de leurs utilisateurs dans le ministère français des transports. Surtout,

notre travail s'est inséré dans l'équipe Economie des Réseaux et Modélisation Offre-Demande (ERMOD) du LVMT[1], dont le modèle LADTA est conçu tout spécialement comme une extension directe de l'affectation statique : nous avons utilisé la formulation analytique et notamment son expression "générique" qui capture simplement les principes sémantiques, afin d'analyser le problème d'affectation, d'en comprendre le contenu sémantique et le traitement algorithmique. Nous avons également utilisé une version initiale des algorithmes, pour développer un programme de simulation informatique que nous appelons prototype, car il est codé en langage SCILAB ce qui privilégie l'accès interactif aux données dans l'environnement logiciel associé, au détriment de l'efficacité des opérations ; nous avons employé le prototype pour expérimenter les algorithmes sur des cas d'école.

Nous avons programmé et expérimenté en particulier :

Le traitement en temps continu des variables dynamiques : les profils horaires des flux et des temps par relation origine-destination, chemin et arc du réseau. Nous avons proposé et expérimenté des techniques "d'assimilation temporelle des morceaux" pour une fonction de profil, affine par morceaux.

Des algorithmes d'équilibrage : nous avons programmé une méthode de combinaison convexe des volumes par arcs proposée par Leurent [Leu03] et aussi une méthode de combinaison convexe des volumes par chemins, qui est différente de la précédente dans le cas dynamique, bien qu'identique dans le cas statique ; et aussi un nouvel algorithme qualifié d'*hybride*. Nous avons expérimenté le comportement de ces algorithmes en mesurant plusieurs critères de convergence à chaque itération du traitement.

Contributions

Sur un plan plus conceptuel, en association avec Fabien Leurent, nous avons apporté deux contributions. L'une concerne la conception de l'algo-

[1]Laboratoire Ville Mobiblité Transport, ENPC/INRETS/UMLV.

rithme hybride, qui traite conjointement des volumes et des temps au niveau des arcs du réseau. L'autre est une approche formalisée de la convergence d'un algorithme d'équlibrage relativement à une formulation mathématique donnée, autrement dit une théorie du mesurage de convergence, applicable à tout traitement algorithmique de l'équilibre dynamique, caractérisé principalement par une base de variables endogènes. Nous avons conçu des critères de convergence qui sont à la fois rigoureux et relativement économes dans leur évaluation.

Au total, notre travail relève de la recherche opérationnelle : avec l'analyse et le développement de formulation mathématique ; de la programmation informatique ; la conception d'un algorithme d'équilibrage et de critères de convergence ; et l'étude expérimentale d'algorithmes.

Plan du mémoire

Ce mémoire comprend neuf chapitres suivis d'une conclusion. Il est organisé en trois parties.

La première partie constate l'état de développement de l'affectation dynamique, d'abord au plan théorique des modèles de recherche, au chapitre 1, puis au plan pratique de l'usage en France, au chapitre 2.

La deuxième partie expose notre inculturation du modèle LADTA et notre contribution à son implémentation. Au chapitre 3 nous exposons les principes, le contenu et la structure de ce modèle. Au chapitre 4 nous traitons dans le détail un cas d'école, dont nous indiquons la solution de manière analytique, avant de la retrouver par l'exécution de l'algorithme d'équilibrage, dont nous suivons la progression sur une série d'itérations. Au chapitre 5, nous relatons l'élaboration de notre prototype, en termes de conception informatique et de structuration du code ; nous proposons également une analyse du problème de fragmentation exponentielle des profils temporels. Nous montrons que ce problème, découvert au cours des premières expérimentations, est intrinsinquement lié à la méthode des combinaisons convexes, et proposons un traitement y remédier, en procédant par agrégation.

La troisième partie développe les formulations analytiques de l'équilibre

dynamique, le mesurage de la convergence ; et nous traitons aussi des algorithmes et concevons l'algorithme hybride. Le chapitre 8 est la contribution théorique centrale de la thèse, avec la mise en évidence des complexités inhérentes à l'affectation dynamique relativement à l'affectation statique, l'investigation de formulations mathématiques alternatives, leur expression à haut niveau d'abstraction dite expression générique, l'étude des algorithmes et la conception de l'algorithme hybride, la théorie du mesurage de convergence. Enfin, le chapitre 9 donne des résultats expérimentaux, permettant d'éclairer certaines conclusions du chapitre 8.

Première partie

Etat de développement de l'affectation dynamique

Chapitre 1

Etat de l'art de la simulation dynamique du trafic routier

Introduction

Les techniques de simulation sont utilisées depuis de nombreuses années dans de multiples domaines, car elles permettent de reproduire un certain nombre de processus et de tester certaines hypothèses à des coûts moindres que les expérimentations en situation réelle. Dans le domaine du transport la simulation dynamique du trafic routier est considérée avec intérêt par de nombreux acteurs (bureaux d'études, laboratoires universitaires, exploitants), et ce sur une vaste gamme d'échelles spatiales et temporelles, depuis la simulation d'une intersection jusqu'à l'étude de réseaux nationaux ou transnationaux.

L'objectif de ce chapitre est de présenter un état de l'art de la simulation dynamique du trafic routier. Il comprend quatre sections.

La section 1.1 présente de façon générale l'objet de la simulation dynamique du trafic routier, dresse un bref historique du domaine et présente quelques applications dans le domaine des transports, ainsi que des outils opérationnels.

En section 1.2 sont rappelés les principes fondamentaux des modèles utilisés pour la simulation d'une part de l'écoulement du trafic, d'autre part des

déplacements. Concernant la simulation de l'écoulement du trafic, l'accent est mis sur les trois principales familles utilisés : macroscopique, microscopique, et mésoscopique. Concernant les déplacements, seront évoqués les paramètres fondamentaux caractérisant les choix des usagers, en particulier les choix de mode, d'itinéraire, et d'heure de départ.

En section 1.3, nous présentons une revue de modèles et outils logiciels existants.

Enfin, la section 1.4 s'intéresse aux modèles d'affectation dynamique fondés sur une représentation analytique de l'écoulement.

1.1 Généralités sur la simulation dynamique du trafic routier

L'objectif de la simulation dynamique est de reconstruire l'évolution du trafic, à la fois dans le temps et l'espace, et ce pour des échelles spatiales et temporelles qui varient grandement en fonction des études. Ainsi des simulations pourront être effectuées au niveau d'une région, d'une ville, d'un corridor autoroutier, ou, à une échelle plus fine, d'une rue ou de quelques carrefours. De même, l'échelle temporelle des études est diverse, depuis la représentation minute par minute de l'écoulement lors de l'heure de pointe sur un carrefour jusqu'à celle de l'écoulement sur une partie du réseau autoroutier transeuropéen en période estivale.

La vision classique de la modélisation du trafic, qui est celle de la planification des transports, comprend quatre étapes de modélisation pour représenter complètement les déplacements au sein d'une aire d'étude :

1. l'étape de génération de la demande ;
2. l'étape de distribution de la demande ;
3. l'étape de choix modal ;
4. l'étape d'affectation.

Pour traiter les problèmes d'exploitation, caractérisés par une échelle de temps plus courte que celle de la planification, il faut également tenir compte d'une cinquième étape, celle de l'écoulement.

1.1.1 Le problème d'affectation du trafic

L'affectation du trafic, en transport de voyageurs ou de marchandises, partage la demande de déplacements entre des origines et des destinations, sur les différents chemins d'un réseau. De façon générale, un modèle d'affectation attribue à chaque usager un itinéraire et un horaire de départ de façon à minimiser un coût généralisé de trajet qui agrège le prix, le temps, la pénibilité, une contrainte d'heure d'arrivée à destination, etc. Le coût généralisé est un prix d'échange du bien transport, entre en offreur de circulation (le réseau) et le demandeur (voyageur ou transporteur). L'offre est soumise à la congestion, car le temps de parcours se dégrade quand la circulation s'intensifie. Un demandeur réagit aux conditions de déplacement en renonçant au mouvement projeté, en différant son départ ou en changeant de chemin, et choisit son itinéraire en fonction d'une rationalité propre.

1.1.2 Repères chronologiques

Du fait du rôle important de la simulation du trafic, tant pour les problématique de la planification que de l'exploitation des réseaux, la recherche a constamment contribué, en particulier depuis les années 1950, à fournir aux praticiens des modèles et outils dont les évolutions et la commodité d'usage ont suivi le développement des moyens informatiques.

Au début, les modèles ont privilégié une approche qualifiée de *statique* dont les variables sont indépendantes du temps. Ensuite, en suivant à la fois le développement scientifique, notamment l'informatique, et les besoins des utilisateurs, les modèles ont pris en compte, dans une approche qualifiée de *dynamique* la dimension temporelle des variables caractéristiques du trafic.

1.1.2.1 Modèles statiques

En 1952, Wardrop [War52] définit deux principes équilibre pour l'approche statique.

Le premier principe correspond à un optimum individuel : chaque individu utilise l'itinéraire qui lui procure le coût généralisé le plus faible pour un

couple origine destination (O-D) donné. Le second principe est fondé sur l'optimum social : le partage des usagers entre les itinéraires concurrents se fait de telle sorte que le coût social généralisé (pour l'ensemble du système) soit minimum. Dans le cas général, ces deux équilibres ne sont pas équivalents.

En 1956, Beckmann [BMW56] a proposé une formulation mathématique pour l'équilibre statique entre offre et demande de déplacement sur un réseau de transport, cela pour chacun des deux principes de Wardrop, dans le cas d'une classe d'usagers et d'une demande élastique. De plus Beckmann a indiqué un algorithme pour calculer l'équilibre. Cette approche a été largement utilisée et développée, jusqu'au début des années 1990, avec des extensions à plusieurs classes de trafic, aux réseaux de transport collectif, au choix modal, à d'autres choix de transport (localisation, distribution spatiale entre relations O-D).

Les logiciels opérationnels de modélisation des déplacements, couramment utilisés par les praticiens des études de trafic, permettent de spécifier et de calculer divers modèles d'équilibre statique. Les logiciels statiques, tels que DAVISUM [AG, PTV02], EMME/2 [Spi84], MARS [Leu95], CIUDAD-SIM [LMQW03], TRANSCAD [Cor, Cal00], ne prennent pas en compte l'évolution de la congestion en fonction de l'heure de la journée. L'une des critiques formulées à l'encontre des modèles statiques est de supposer que la congestion reste constante durant la période étudiée. Dans ces modèles, le volume de demande, la durée de trajet, le prix ne varient pas au cours du temps.

1.1.2.2 Modèles de simulation dynamique

Les modèles dynamiques ont pris leur essor dans les années 1990, profitant des progrès importants des moyens informatiques. Dans les modèles dynamiques, la congestion, et par conséquent le choix d'itinéraire, varie au cours du temps. Dans certains modèles le choix de l'horaire de départ est également considéré.

De manière succincte, les modèles dynamiques peuvent être décrits comme suit : les usagers font face à un double choix, celui de l'heure de départ et celui de l'itinéraire. Le coût généralisé associé à un trajet est la somme pondérée

du temps de déplacement et d'une fonction des pénalités dues aux arrivées précoces ou tardives. Chaque individu choisit l'heure de départ et l'itinéraire qui minimisent cette fonction de coût généralisé. Un équilibre dynamique est atteint lorsque aucun usager ne peut modifier son itinéraire ou son heure de départ en vue de réduire strictement son coût généralisé.

Les modèles de simulation dynamique peuvent être classés en trois familles selon la manière de représenter l'écoulement des véhicules sur le réseau. : (i) modèles macroscopiques, (ii) modèles microscopiques, (iii) et modèles mésoscopiques. Une description plus précise de ces trois familles est donnée en section 1.2.

1.1.3 Domaines d'application

Les outils de modélisation et de simulation du trafic sont utilisés afin d'optimiser la gestion du trafic, depuis la planification jusqu'à l'exploitation des réseaux routiers. Parmi leurs nombreuses utilisations, on trouve :

- l'étude et la conception de nouvelles infrastructures ;
- l'évaluation de mesures d'exploitation ;
- la régulation du trafic en temps réel ;
- les études d'impact environnemental ;

Pour la planification, l'étude et la conception de nouvelles infrastructures, les résultats issus de simulations sont utilisés par les concepteurs pour décider de critères techniques (par exemple, le nombre de voies d'une autoroute) et/ou économiques (tarification).

Les outils de simulations sont également utilisés pour l'évaluation a priori de mesures d'exploitation. Il est ainsi possible d'estimer l'efficacité de dispositifs de régulation du trafic : régulation d'accès, feux de circulation, stations de péage, etc.

De plus, pour ce qui concerne l'aide à la décision et la régulation du trafic en temps réel, des applications possibles sont :

- Décision des stratégies de régulation d'accès.
- Estimation de l'état du réseau existant.
- Gestion du trafic (la congestion), des incidents et des parkings.

- Production d'information de guidage.
- Evaluation d'opération de régulation ou des études ITS.
- Définition des stratégies adaptatives optimales de régulation de carrefour à feux.
- Comparaison entre différentes stratégies de régulation de trafic, stratégies de guidage.

Enfin, les outils de simulation de trafic peuvent également être employés avec profit au-delà du champ strict de la gestion du trafic. Ainsi, par exemple, la simulation de trafic permet, par couplage avec d'autres outils de simulation, la réalisation d'études d'impact environnemental, comme :

- Mesure d'impact sur l'environnement de plans de circulation.
- Calcul d'émission de polluants.
- Estimation d'effet du système d'infrastructure sur la qualité de la vie.
- Estimation de nuisances sonores du trafic routier [Lec02].
- Calcul de la consommation d'énergie.

1.2 Principes des modèles

1.2.1 Concernant l'écoulement

A partir d'un état initial du réseau et de conditions aux limites (demande en entrée du réseau, contraintes en sortie, etc.), le modèle d'écoulement doit permettre de déterminer l'évolution des variables caractéristiques du trafic. L'écoulement du trafic peut être décomposé en deux parties : l'écoulement le long des axes et l'écoulement aux points d'échanges (carrefours, échangeurs). De plus, la littérature distingue classiquement trois familles de modèles pour l'écoulement : les modèles macroscopiques, mésoscopiques et microscopiques.

Notre présentation des modèles d'écoulement est fondée sur cette classification. Sont présentés dans un premier temps les modèles d'écoulement macroscopiques, en distinguant les modèles macroscopiques par équations aux dérivées partielles (e.d.p) en 1.2.1.1 et les modèles macroscopiques *analytiques* (1.2.1.2). Dans un deuxième temps sont esquissés les modèles microscopiques (1.2.1.3), et enfin les modèles mésoscopiques (1.2.1.4).

1.2.1.1 Modèles macroscopiques par e.d.p.

Dans les modèles macroscopiques par e.d.p., le trafic est représenté comme un fluide continu, caractérisé par des grandeurs moyennes qui sont les variables d'état du modèle. Ces variables, dépendant de l'espace[1] et du temps, sont :

- la densité ou concentration, exprimée en nombre de véhicules par unité de longueur ;
- le débit, exprimé en nombre de véhicules par unité de temps ;
- la vitesse, exprimée en unité de longueur par unité de temps.

Ces modèles utilisent des équations aux dérivées partielles liant les variables de débit, de vitesse, et de concentration, considérées dans deux dimensions : l'espace et le temps.

Trois équations reliant ces trois variables sont utilisées.

La première est l'équation de conservation qui exprime le fait que le nombre de véhicules présents dans une section à un instant $t + dt$ est égal au nombre présent à l'instant t, plus le nombre de véhicules entrés pendant dt, moins le nombre de véhicules sortis pendant dt.

La deuxième équation définit la vitesse du flot comme le rapport du débit sur la densité.

Enfin, la troisième équation varie suivant le modèle. Elle relie le débit à la concentration, elle est une épure de ce qui est observé expérimentalement.

Les modèles dits du premier ordre ([LW55]) sont caractérisés par l'utilisation d'une relation d'équilibre entre débit et densité, aussi appelée diagramme fondamental, et supposée vraie à tout instant et en tout point. Les transitions sont décrites comme des passages instantanés d'un état d'équilibre à un autre. Ces modèles ont l'avantage de la simplicité : ils permettent le calcul de solutions analytiques et sont dépourvus d'effets secondaires gênants. Ils présentent en revanche l'inconvénient de ne pas décrire les phases transitoires d'accélération et de décélération.

[1] Le plus souvent une seule dimension de l'espace est considérée, mais il existe des modèles macroscopiques utilisant une deuxième dimension d'espace, pour modéliser par exemple des changements de voie.

Les modèles du second ordre explicitent quant à eux l'accélération du flux, soit en exprimant une loi de poursuite entre véhicules ([Pay79], parmi d'autres), soit en modélisant des transitions entre phases dans une approche inspirée de la cinétique des gaz en théorie thermodynamique ([PH71]). Typiquement, les modèles du second ordre définissent la vitesse comme se relaxant vers la vitesse d'équilibre du diagramme fondamental. Les termes qui interviennent dans cette évolution sont des termes de retard et d'anticipation.

Ces modèles décrivent plus finement les transitions mais présentent des effets indésirables (débits négatifs) qui obligent à leur associer diverses formes de contraintes dans les calculs. Ils ne sont donc pas utilisables directement (pas de solution analytique) mais uniquement en simulation. La controverse est au demeurant vive entre partisans et détracteurs des différents types de modèles, et l'utilisateur peut avoir bien du mal à se faire sa propre opinion.

1.2.1.2 Modèles macroscopiques *analytiques*

Dans les modèles *analytiques* d'écoulement les variables sont le volume et le temps. Le problème d'écoulement de ces modèles est de déterminer le volume sortant puis le temps de traversée effectif en connaissant le volume entrant, les contraintes (capacité, loi FIFO, loi de conservation), le temps de traversée libre via l'analyse de l'évolution de la file d'attente avec stockage.

1.2.1.3 Modèles microscopiques

Dans les modèles microscopiques, l'état dynamique de chaque véhicule (position, vitesse, accélération) est représenté explicitement. Les caractéristiques de l'écoulement en section courante sont déduites d'un modèle appelé loi de poursuite, qui définit l'accélération de chaque véhicule suivant la situation de trafic rencontrée :

- en situation de *trafic fluide*, c'est-à-dire lorsque la distance inter-véhiculaire est assez grande pour négliger les interactions entre véhicules, l'accélération est calculée en fonction de l'écart entre la vitesse actuelle et une vitesse désirée, déduite en particulier des caractéristiques de l'infrastructure ;

– lorsque la distance inter-véhiculaire est inférieure à un seuil donné, le calcul de l'accélération prend en compte l'état dynamique du ou des véhicules situés immédiatement en aval (distance, vitesse, vitesse relative, etc).

Par ailleurs, toute modification de l'accélération est affectée d'un temps de réaction. Les distributions, dans la population des conducteurs, du temps de réaction, de la vitesse désirée, des bornes limitant l'accélération ainsi que d'autres coefficients d'ajustement constituent les paramètres essentiels des modèles microscopiques.

1.2.1.4 Modèles mésoscopiques

Les modèles mésoscopiques se présentent comme des intermédiaires entre la finesse des modèles microscopiques et la relative rusticité des modèles macroscopiques, en arguant du fait qu'ils représentent le trafic sous forme de "paquets" de véhicules dont la progression est décrite individuellement.

L'étude de ces modèles montre en fait que la séparation du trafic en paquets n'est en rien liée à l'écoulement, mais représente simplement une façon de décrire l'affectation du trafic, chaque paquet ayant une destination déterminée et suivant un itinéraire donné. L'écoulement en section courante est au contraire représenté de façon très rustique, soit par un temps de parcours constant indépendant de la densité du trafic, soit par une loi donnant, à l'entrée d'un paquet dans un tronçon du réseau, son temps de parcours en fonction de la densité. Dans ce dernier cas, l'écoulement est en fait géré par un modèle macroscopique élémentaire à l'échelle du tronçon.

L'approche mésoscopique de METROPOLIS [DPMN96] se traduit par le fait que la génération de véhicules utilise l'approche de désagrégation totale (par toutes les variables de comportement des usagers) pour la description de décisions individuelles. Par contre, l'écoulement est représenté de manière macroscopique par la loi vitesse-débit.

1.2.1.5 Autres modèles d'écoulement

Signalons enfin qu'il existe également d'autres modèles pour la modélisation de l'écoulement sur un arc. Citons par exemple le modèle de dispersion de peloton de Robertson [HRBR82], où encore le modèle *cell-transmission*, proposé par Daganzo [Dag94, Dag95]).

1.2.2 Concernant les déplacements

1.2.2.1 Notion de déplacement

Un *déplacement* est un trajet effectué par un usager sur un réseau de transport, de manière continue, pour se rendre d'un point d'origine à un point de destination. Un déplacement est décidé par un usager et concerne un certain motif (par exemple, domicile-travail). De plus, un déplacement peut être caractérisé par une heure de départ et/ou une heure d'arrivée.

Les usagers sont classés en plusieurs classes selon les critères qui peuvent être liés :

- au motif des déplacements ;
- au type de véhicule utilisé, qui peut par exemple conditionner les possibilités d'accès à certains arcs du réseau ;
- à des attributs économiques, comme la valeur du temps ;
- etc.

Pour chaque classe d'usagers, le volume de déplacements entre les lieux d'origine et de destination est représenté par une matrice origine-destination, ou matrice O-D.

1.2.2.2 Choix d'itinéraire

Il existe deux cas pour le problème du choix d'itinéraire : ce choix est exercé soit par les usagers, soit par l'exploitant du réseau selon des règles précises donc faciles à simuler dans un modèle. Nous ne considérons ici que la problématique du choix par les usagers.

Un réseau routier offre en général plusieurs itinéraires possibles pour une relation O-D. De plus, l'ensemble de ces itinéraires peut varier en fonction

de la classe de l'usager. Chaque itinéraire, compris comme un service de déplacement, est caractérisé par ses attributs de temps et de prix à partir des attributs des arcs. La formation de ce service est aussi fonction de l'heure de départ dans le cas des horaires variables.

Chaque classe d'usagers perçoit objectivement les caractères des services qui lui sont accessibles, les évalue de manière synthétique par un coût généralisé, et choisit le service optimal qui minimise son coût généralisé. Le volume de chaque classe sera réparti entre les services, ce qui correspond à l'affectation des déplacements sur le réseau.

A partir des principes de la formation des services et du choix des usagers présentés précédemment, deux grandes familles de modèles d'affectation ont été développées. La première, qualifiée de *déterministe*, est fondée sur la résolution de la formulation mathématique de l'équilibre usager au sens de Wardrop. Il convient de noter que cette approche fait l'hypothèse d'une connaissance parfaite, pour chaque usager, de l'état du réseau.

Afin de relaxer cette hypothèse, l'approche *stochastique* modélise une perception imparfaite des coûts généralisés, en ajoutant aux coûts généralisés déterministes un terme d'erreur aléatoire. Les choix des usagers sont alors effectués en fonction des coûts généralisés perçus.

Les modèles stochastiques d'équilibre d'usagers diffèrent selon la distribution de probabilité des termes d'erreur aléatoire. Le modèle le plus couramment utilisé est le modèle logit basé sur des termes d'erreur distribués de Gumbel. Le modèle de distribution gaussienne ou probit est aussi utilisé dans certaines études. Deux méthodes sont couramment utilisées pour calculer l'équilibre usager stochastique : soit par résolution analytique, lorsque celle-ci est possible, soit par la méthode de Monte-Carlo.

1.2.2.3 Choix de l'heure de départ

Le temps de parcours dépend toujours des conditions du trafic, notamment de la densité, et il varie selon les heures de la journée. Les usagers choisissent leur heure de départ selon soit leur motif de déplacement (par exemple, horaire cible d'arrivée pour un déplacement domicile-travail), soit

dans le but de minimiser leur coût généralisé de déplacement.

Lorsque la demande dépasse l'offre sur un arc du réseau, un phénomène de congestion apparaît, qui peut amener les usagers à reconsidérer leur choix initial d'itinéraire et/ou d'heure de départ.

En 2001, Leurent a étudié le problème du choix d'horaire [Leu01]. Il a mené une analyse physique puis mathématique pour formuler les principes du problème. Il détermine le coût généralisé moyen en fonction de l'horaire de départ. Puis, une solution permettant de calculer l'horaire de départ qui minimise le coût généralisé est proposée via une formule dérivée de la fonction du coût moyen. Cette formule détermine l'horaire de départ optimal à partir de l'horaire cible et des paramètres de distribution du temps de parcours, en comptant le péage marginal et les coûts marginaux d'avance, de retard, d'origine, de transport. Le coût généralisé minimal est finalement calculé à partir de l'horaire de départ optimal.

1.3 Exemples de modèles et outils logiciels

Cette section illustre chaque famille de modèle de simulation de trafic par la description détaillée d'un modèle et/ou outil logiciel appartenant à cette famille. L'ensemble des modèles/outils logiciels étudiés par l'auteur est donné en annexe A.

1.3.1 AIMSUN

AIMSUN, développé par TTS (Transport Simulation Systems), est un outil de simulation microscopique du trafic très répandu [FB93]. Ce simulateur est intégré dans la plate-forme multi-modèles appelée GETRAM (Generic Environment for TRaffic Analysis and Modeling) qui comprend :

- un éditeur graphique du réseau (TEDI),
- une banque de données,
- un module pour sauvegarder les résultats,
- un programme d'application permet d'interfacer avec d'autres modèles d'affectation et de simulation tels que EMME/2, TRANSYT, SYN-

CHRO, GIS.

- Une extension permet d'interfacer avec d'autres applications externes telles que le contrôle du trafic ou le système de gestion du trafic.

Cette architecture fait de AIMSUN un simulateur très ouvert et convivial, ayant la capacité de traiter un réseau de grande taille en milieu urbain. Concernant l'écoulement, comme tous les autres modèles microscopiques, AIMSUN simule le trafic par :

- un modèle de poursuite ;
- un modèle de changement de voie (pour *lane changing*) ;
- un modèle de simulation de trafic aux intersections et aux accès.

Pour l'affectation du trafic, AIMSUN utilise les modèles suivants :

- un modèle de formation des services (chemins) utilisant le PCC ;
- un modèle du choix des usagers utilisant le modèle LOGIT.

Nous allons maintenant décrire rapidement les modèles composants de AIMSUN.

1.3.1.1 Modèle de poursuite

Le modèle de poursuite de AIMSUN est une extension du modèle originel de Gipps [Gip81]. Les extensions approtées au modèle originel prennent en compte des données observées expérimentalement sur les comportements de conducteurs, et portent essentiellement sur :

- la méthode de calcul de la vitesse désirée ;
- l'influence de la situation de trafic présente dans les voies adjacentes sur le comportement de conducteur ;
- les effets de topologie (la pente de la route).

1.3.1.2 Modèle de changement de voie

Le modèle de changement de voie utilisé est également fondé sur le modèle de Gipps [Gip86]. Ce modèle simule le comportement de conducteur par un processus de décision d'après les principes suivants :

- A chaque instant, un véhicule s'interroge sur la nécessité de changer de voie. La décision dépend de plusieurs facteurs tels que la possibilité

de tourner, la distance vers le tournant suivant, la situation de trafic actuel, etc.

- Si la réponse est affirmative, deux nouvelles questions se posent. *Est-il est désirable de changer de voie ?* et *Est-il est possible de changer de voie ?* Pour répondre à cette dernière question, un modèle dit "Gap Acceptable" est utilisé.

1.3.1.3 Gestion des intersections et voies d'accès

AIMSUN considère les méthodes de contrôle suivantes afin de simuler le trafic aux intersections :

- par un plan de contrôle fixe ou adaptif ;
- par un plan de contrôle fourni par TRANSYT ;
- par l'interface avec le système de gestion du trafic.

1.3.1.4 Affectation du trafic

Pour l'affectation dynamique du trafic, le principe est de :

- mettre à jour temporellement les services (chemins) allant les origines vers les destinations selon la situation de trafic ;
- affecter les véhicules aux services par chaque période de temps ;
- recalculer les meilleurs services pour les véhicules correspondant à leurs positions vers leurs destinations.

Cette procédure exige donc un modèle de formation des services et un modèle de choix des usagers qui sont appliqués suivant la procédure suivante :

1. Calculer les PCCs d'après les coûts de traversée estimés initialement ;
2. Affecter le trafic aux chemins selon les informations des services/chemins, puis actualiser les coûts par chemin à partir des résultats de la simulation ;
3. Recalculer les PCCs utilisant les nouveaux coûts ;
4. Informer les usagers des nouvelles informations relatives aux services/-chemins ;
5. Retourner à l'étape 2.

Cette affectation est réalisée par période de temps. Pour la formation des services/chemins, le coût de traversée des chemins est calculé à partir de celui des arcs. Par arc et par classe d'usagers, le coût est déterminé grâce à une fonction de coût par défaut ou modifiable par utilisateur.

Pour le choix des usagers, un modèle de type Logit, appelé C-Logit, est utilisé. Il permet de calculer la probabilité P_r du choix d'un chemin r dans l'ensemble des chemins I_{rs} disponibles pour le relation O-D rs par la formule :

$$P_r = \frac{e^{\theta(V_r - CF_r)}}{\sum_{r' \in I_{rs}} e^{\theta(V_{r'} - CF_{r'})}}$$

où V_r est l'utilité perçue du chemin r, θ le facteur d'échelle du modèle Logit et CFr un *facteur d'analogie* (pour *commonality factor*) entre deux chemins. Il est donné par :

$$CF_r = \beta \cdot \sum_{r' \in I_{rs}} \left(\frac{L_{rr'}}{L_{r'}^{1/2} L_r^{1/2}} \right)^{\gamma}$$

avec L_r (resp. $L_{r'}$) la longueur du chemin r (resp. r') et $L_{rr'}$ la longueur de l'intersection des chemins r et r'.

1.3.1.5 Capacité de traitement

Nous avons identifié deux projets opérationnels pour lesquels les données caractéristiques du réseau traité sont disponibles. Ils sont décrits brièvement ci-après.

ISM (Intermodal Strategy Manager) : un projet de Réseau routier d'Hessian et de la Ville de Frankfurt, dont l'objectif est le développement et la démonstration d'un système de management de trafic intermodal basé sur un modèle de simulation pour évaluer les impacts des stratégies proposées. Les grandeurs essentielles de ce projet sont résumées dans le tableau suivant :

Longueur totale	812 miles
Nombre de sections (arcs)	17 375
Nombre de nœuds	2 276
Types de véhicules	bus, VP, PL
Nombre de zones (O-D)	50

Et un autre projet appliqué par AIMSUN est le grand réseau urbain de

Helsinki, qui est développé par Simulus. Les caractéristiques du réseau considéré sont les suivantes :

Longueur totale	1 138 miles
Nombre de sections (arcs)	3 188
Nombre de nœuds	1 193
Nombre de zones (O-D)	339

1.3.2 METROPOLIS

METROPOLIS [DPMN96] est un modèle mésoscopique qui détermine le temps de traversée en fonction du volume entrant instantané, de la concentration instantanée et les paramètres d'offre de l'arc (longueur et capacité). L'interaction des véhicules est limitée et décrite seulement par une fonction de congestion. Une file d'attente verticale est représentée par arc, et suit une loi FIFO.

L'approche mésoscopique dans METROPOLIS se traduit par le fait que la génération de véhicules utilise l'approche de désagrégation totale (par toutes les variables de comportement d'usagers) pour la description des décisions individuelles. Par contre, l'écoulement est représenté de manière macroscopique par la loi vitesse-débit.

Les modèles dynamiques représentent la congestion par période de temps. Chaque modèle utilise une unité de temps différente selon leur discrétisation des données. Dans le modèle considéré ici, les auteurs proposent deux modèles pour le temps : modèle *par pas de temps* et *par événement*.

Le modèle par pas de temps utilise un petit intervalle (en général, quelques secondes) qui autorise une représentation fine des phénomènes du trafic, notamment en saturation. Le modèle événementiel utilise une description de temps continue, et les variables utilisées sont affectées uniquement lors de la survenance d'un événement susceptible de modifier l'état du système. Chaque phénomène ou événement est daté. Ensuite une liste d'attente des événements est traitée, dans l'ordre de survenance des événements.

La demande est composée par la combinaison de modèles qui déterminent le choix modal, le choix de l'heure, et le choix d'itinéraire, en utilisant les formulations économiques de Vickrey pour calculer le coût généralisé, la proba-

bilité d'utilisation de voiture ou TC, et le temps de parcours. Les paramètres utilisés dans le modèle sont calibrés par une procédure à quatre étapes :

1. la collection des paramètres du réseau et matrice O-D ;
2. le calibrage des lois de congestion ;
3. le calibrage des paramètres dynamiques du comportement ;
4. le réajustement de la valeur des paramètres techniques.

METROPOLIS utilise la méthode de discrétisation par événement dans le but d'obtenir un coût informatique moins élevé par rapport au modèle par pas de temps.

Dans le problème d'écoulement, l'émission de véhicules sur le réseau est gérée par décisions individuelles (i.e. les variables de comportement d'usagers), en respectant la règle FIFO. L'écoulement sur chaque arc est simulé de manière macroscopique par la loi vitesse-débit, en analysant l'évolution de files d'attente verticales ou horizontales. Toutes les variables dépendent du temps, sauf la capacité en entrée et en sortie.

La programmation du modèle a été réalisée en C++. Il dispose d'une interface utilisateur pour le dialogue et l'analyse numérique et graphique des résultats. METROPOLIS possède une grande capacité de calcul pour les réseaux de grande taille, ainsi qu'une interface utilisateur conviviale.

1.3.3 STRADA

Le modèle STRADA est développé conjointement par le CERMICS (ENPC) et par le LICIT (ENTPE-INRETS). Son objectif est de simuler le trafic dans les conditions spécifiques de l'affectation dynamique [BLL96]. Ce modèle constitue la version discrétisée de la généralisation aux réseaux et aux conditions de l'affectation dynamique du modèle macroscopique du premier ordre de Lighthill-Whitham-Richards (modèle LWR), en utilisant le schéma de Godunov pour déterminer le débit. Cela permet de traiter les intersection complexes et les grands réseaux mixtes et de modéliser les divers éléments du réseau à une échelle adaptée aux contraintes de temps de calcul et/ou de capacité de recueil des données.

1.3.3.1 Modélisation de l'écoulement

Concernant l'écoulement, le modèle STRADA repose sur l'analyse réalisée par Lebacque du schéma de Godunov appliqué au modèle de LWR [BLL96]. Ce travail a donné une fonction dérivée partielle de la concentration $K(x,t)$ et de la fonction d'équilibre reliant le débit à la concentration et exprimant la relation fondamentale $Q(x,t) = Q_e(K,x)$ (relation de Daganzo, 1994). Cette fonction représente le principe du modèle d'écoulement de STRADA pour le cas inhomogène.

$$\frac{\partial K}{\partial t} + \frac{\partial}{\partial x} Q_e(K,x) = 0$$

où x est la position et t est le temps.

La solution de ce problème repose sur les concepts d'offre et de demande locales de trafic qui sont définis de la manière suivante :

– pour l'offre :

$$\Omega_e(K,x) = \begin{cases} Q_{max}(x^+) & \text{si} \quad K \leq K_{crit}(x^+) \\ Q_e(K,x^+) & \text{si} \quad K \geq K_{crit}(x^+) \end{cases}$$

– pour la demande :

$$\Delta_e(K,x) = \begin{cases} Q_e(K,x^-) & \text{si} \quad K \leq K_{crit}(x^-) \\ Q_{max}(K,x^-) & \text{si} \quad K \geq K_{crit}(x^-) \end{cases}$$

On notera les symboles x^+ et x^- qui indiquent que l'on prend les parties à droite et à gauche de la fonction fondamentale $Q(x,t) = Q_e(K,x)$ selon K_{crit}. Les définitions de l'offre et de la demande d'équilibre constituent le schéma de Godunov qui permet de calculer le débit en tout point x à tout instant t :

$$Q(x,t) = \min\{\Omega_e(x,t), \Delta_e(x,t)\}$$

Le cas homogène correspond au cas où la fonction d'équilibre $Q_e(K,x)$ ne dépend pas de la position x, c'est-à-dire le cas où les attributs du réseau et les comportement des usagers sont indépendants de la position, et ce cas est donc retenu pour la section courante. Par contre, le cas inhomogène est retenu pour les intersections complexes. En outre, ce modèle traite le problème discontinu

dans certains cas, lorsque le plan de capacité du réseau change de façon rapide.

Les extensions du modèle fondamental LWR, les conditions aux limites intrinsèques et la modélisation des intersections, ne sont pas suffisantes pour la modélisation de l'écoulement du trafic dans les conditions de l'affectation dynamique. Le modèle STRADA décrit aussi l'écoulement par catégories d'usagers, il s'agit de désagréger les débits et concentrations K et Q en débits et concentrations partiels, par classes d'usagers.

Concernant la discrétisation des tronçons, elle repose sur le schéma Godunov. Ce schéma conservatif du 1er ordre est, dans le cas homogène, convergent. Pour le cas inhomogène, chaque tronçon est divisé en segments dont les paramètres physiques sont résumés dans une fonction d'offre d'équilibre $\Omega_e(.,s)$ et une fonction de demande d'équilibre $\Delta_e(.,s)$.

L'état du tronçon est recalculé à chaque pas de temps, i.e le débit moyen au point x pendant l'intervalle de temps est le minimum de l'offre du segment aval et de la demande du segment amont à ce point pendant cet intervalle de temps :

$$Q(.,s) = \min\{\Omega_e(.,s), \Delta_e(.,s)\}$$

Pour assurer un bon comportement du modèle (et notamment la convergence de la discrétisation), l'inégalité suivante de type CFL (Courant-Friedrichs-Lewy) doit impérativement être satisfaite pour tout segment s :

$$\frac{V_{s,max}\Delta t}{l_s} \leq 1,$$

et ce rapport doit être aussi proche de 1 que possible afin de diminuer les effets de viscosité numérique.

1.3.3.2 Modélisation des intersection

Concernant la modélisation des intersections, les zones d'échange sont découpées en segment par points d'entrée et de sortie, et les carrefours complexes sont découpés en zones d'échange. Chaque zone a des points d'entrée et de sortie, le débit entre deux zones est estimé directement par le schéma

de Godunov présenté précédemment. Le trafic en zones d'échange qui sont distingués en deux types : zones d'échange englobantes (ZEE) et les zones partielles (ZEP), donc est désagrégé par flux entre les segments, qui seront simulés par pas de temps pour calculer l'état du trafic interne des zones. Deux équations de conservation relatives aux zones ZEE et ZEP sont enfin formulées pour calculer le nombre de véhicules dans ces zones, et ainsi déterminer l'offre et la demande globale pour chaque zone. Ces équations permettent de compléter le modèle de zones, et donc le modèle d'écoulement complet.

1.3.3.3 Estimation du temps de parcours

Concernant l'estimation du temps de parcours, ce modèle utilise le concept de temps de parcours instantané :

$$\begin{aligned} &V\frac{\partial T}{\partial x}+\frac{\partial T}{\partial t}=1 \\ &T\left(x_0,t\right)=0 \end{aligned}$$

En considérant la discrétisation :

$$\begin{aligned} &T_s^{t+1}=\beta_s^t T_{s-1}^t+\left(1-\beta_s^t\right)T_s^t+\Delta t \\ &T_0^t=0 \end{aligned}$$

avec $\beta_s^t=\frac{Q_s^t\Delta t}{N_s^t}$ et N_s^t le nombre de véhicules présents sur le segment s à l'instant t.

Mais la définition de ce temps n'est pas directement applicable : il devient infini à chaque interruption du trafic. Pour surmonter cette difficulté, le modèle estime le temps grâce à la dérivée partielle d'un indicateur R.

1.3.3.4 Affectation du trafic

Concernant l'affectation, STRADA utilise un modèle de type logit [LBLM96]. Au niveau d'une intersection donnée, on considère que les usagers devant se rendre à une destination et se trouvant à cette intersection ont le choix entre les différentes sorties de cette intersection. Pour chaque itinéraire possible passant par l'une des intersections en aval, on distingue d'une part le

temps de parcours de l'intersection traitée à l'intersection aval considérée, et d'autre part le temps de parcours minimum de l'intersection aval à la destination considérée. Le temps total de l'intersection traitée à la destination considérée est estimé. Enfin, la proportion des usagers présents sur l'intersection traitée et choisissant de passer par une intersection aval donnée pour atteindre leur destination est calculée par un modèle logit.

1.3.4 METACOR

Le modèle METACOR, développé depuis 1994 sous la direction de Habib Haj Salem du GRETIA/INRETS [SE94, SP98], est un modèle dynamique macroscopique de l'écoulement du trafic, fondé sur une approche macroscopique construite en analogie avec la théorie hydrodynamique. METACOR intègre deux types de modèles d'écoulement : l'un du 1er ordre, pour la partie urbaine et l'autre du 2ème ordre pour la partie interurbaine.

La topologie du réseau dans METACOR est modélisée par un graphe dont les éléments de base sont les tronçons autoroutiers, les intersections autoroutières, les tronçons urbains et les carrefours. Les entrées et les sorties du corridor sont modélisées comme étant respectivement les origines et les destinations.

Afin de résoudre numériquement les équations de METACOR, les auteurs procèdent à une discrétisation spatiale et temporelle. La discrétisation spatiale consiste à diviser un tronçon m en plusieurs segments i de même longueur L_m tout en respectant la contrainte de stabilité suivante : $L_m > v_{d,m}T$, avec $v_{d,m}$ la vitesse désirée sur le tronçon m, et T l'intervalle de temps discrétisé.

1.3.4.1 Modélisation de l'écoulement

Modèle urbain du 1er ordre Les variables du modèle urbain sont les suivantes :

- $\rho_{m,i}(k)$: densité sur le segment i du tronçon m à l'instant kT.
- $q_{m,i-1}(k)$ le débit sortant du segment i du tronçon m pendant l'intervalle de temps $[kT;(k+1)T]$.

- $q_{m,i}(k)$ le débit entrant sur le segment i du tronçon m pendant l'intervalle de temps $[kT;(k+1)T]$.
- $v_{m,i}(k)$ la vitesse moyenne des véhicules circulant sur le segment i,m à l'instant kT.
- $\rho_{max,i}(k)$ la densité maximale du segment i,m à l'instant kT.

Ces variables sont calculées sur toute la section et non pas par voie. Le débit de sortie du tronçon est calculé en fonction de l'état du trafic sur le carrefour. Le modèle calcule les débits sortant des autres segments en allant de l'aval vers l'amont.

Les équations du modèle urbain sont les suivantes :

- l'équation de conservation :

$$\rho_{m,i}(k+1) = \rho_{m,i}(k) + \frac{T}{L_m}\left[q_{m,i}(k) - q_{m,i-1}(k)\right]$$

- la relation débit-vitesse :

$$q_{m,i}(k) = \rho_{m,i}(k)\, v_{m,i}(k)$$

- le diagramme fondamental :

$$V_{m,i}(k) = V_{f,m}\left[1 - \frac{\rho_{m,i}(k)}{\rho_{max,m}}\right]$$

  Cette fonction exprime la vitesse comme une fonction monotone décroissante de la densité.
- Calcul du débit moyen : ce calcul prend en compte les réactions des conducteurs face aux changements qui affectent le trafic à l'aval, surtout en présence de congestion. Une valeur moyenne du débit traversant l'interface de deux segments adjacents est calculée à partir des valeurs de leur densités. La méthode utilisée est similaire à celle détaillée en 1.3.3 pour le modèle STRADA.

Modèle de carrefour : Les carrefours du réseau urbain sont subdivisés en mouvements et en zones d'échange. Par mouvement, on désigne les véhicules

allant d'une entrée vers une sortie ; par zones d'échange, on désigne la liaison entre la surface du carrefour et les voies de sortie. Cette modélisation permet de traiter les mouvements et les segments des tronçons de façon homogène.

Concernant le problème de tourne-à-gauche, le modèle utilise le concept de créneau acceptable de Lebacque. Un créneau est défini comme l'intervalle temporel séparant deux véhicules d'une voie prioritaire qu'un usager juge nécessaire pour pouvoir traverser la voie (ou s'y engager). Le modèle utilisé est celui du conducteur impatient de Mahmassani et Sheffi [MS81].

Modèle autoroutier, du second ordre : Les variables et les équations du modèle autoroutier sont similaires à celles du modèle urbain, en ajoutant le nombre de voies λ_m, et avec pour diagramme fondamental :

$$F\left(\rho_{m,i}\left(k\right)\right) = v_{d,m} \exp\left[-\frac{1}{a}\left(\frac{\rho_{m,i}\left(k\right)}{\rho_{cr,m}}\right)\right]$$

où $\rho_{cr,m}$ est la densité critique du tronçon m et a un paramètre réel.

L'équation dynamique de la vitesse utilisée dans le modèle autoroutier a été proposée par Papageorgiou [Pap88]. C'est une extension du développement en série de Taylor à l'ordre 1 de l'équation de Payne [PAY71] :

$$V\left(x, t+\tau\right) = F\left(\rho\left(x+\Delta x\right), t\right)$$

qui représente les réactions des conducteurs face aux changements qui affectent le trafic à l'aval. Cette extension permet de tenir compte des influences des rampes d'accès et du changements du nombre de voies sur la vitesse. Cette équation s'écrit :

$$\begin{array}{ll}
V_{m,i}\left(k+1\right) = V_{m,i}\left(k\right) + \frac{T}{\tau}\left[F\left(\rho_{m,i}\left(k\right)\right) - v_{m,i}k\right] & \textit{terme de relaxation} \\
+\frac{T}{L_m} v_{m,i}\left(k\right)\left[v_{m,i-1}\left(k\right) - v_{m,i}\left(k\right)\right] & \textit{terme de convection} \\
-\frac{v}{\tau}\frac{T}{L_m}\frac{\rho_{m,i+1}(k)-\rho_{m,i}(k)}{\rho_{m,i}(k)+\kappa} & \textit{terme d'anticipation} \\
-\frac{\delta T}{L_m}\frac{u_r(k)\cdot v_{m,i}(k)}{\rho_{m,i}(k)+\kappa} & \textit{terme "on ramps"} \\
+\frac{\Phi T}{L_m}\frac{\lambda_{m,i}-\lambda_{m,i+1}}{\lambda_{m,i}}\frac{\rho_{m,i}(k)}{\rho_{cr,m,i}} v_{m,i}^2\left(k\right) & \textit{terme "ramps drop"}
\end{array}$$

Modèle de l'intersection autoroutière : Les convergents, les divergents, le changement du nombre de voies dans un réseau autoroutier sont représentés par des nœuds avec les tronçons entrants et sortants. Dans un nœud n, le débit sortant de chaque sortie m peut être exprimé en fonction du débit entrant total Q_n dans ce nœud selon les pourcentages directionnels de sortie α_m, cela pour chaque instant kT : $q_m(k) = \alpha_m(k) Q_n(k)$.

1.3.4.2 Modélisation de l'affectation

Affectation sur un tronçon : Pour l'affectation sur un tronçon du réseau, le modèle utilisé est celui de D'Ans et Gazis, proposé en 1976, dont l'une des caractéristiques notable est de satisfaire la condition de conservation du flux entrant et sortant par relation O-D. A partir de la définition de la densité partielle par destination j du segment i du tronçon m comme :

$$\gamma_{m,i,j} = \frac{\rho_{m,i,j}}{\rho_{m,i}}, \forall j \in S_m$$

avec S_m l'ensemble des destinations possibles à partir du tronçon m.

Les équations de conservation du nombre de véhicules sont reformulées par destination de la manière suivante :

- dans le cas autoroutier :

$$\rho_{m,i}(k+1) = \rho_{m,i}(k) + \frac{T}{L_m \lambda_m} \left[\gamma_{m,i-1,j} q_{m,i-1}(k) - \gamma_{m,i,j} q_{m,i}(k)\right]$$

- dans le cas urbain :

$$\rho_{m,i}(k+1) = \rho_{m,i}(k) + \frac{T}{L_m} \left[\gamma_{m,i-1,j} q_{m,i-1}(k) - \gamma_{m,i,j} q_{m,i}(k)\right]$$

Affectation sur une intersection autoroutière : La modélisation de l'affectation dans un nœud résulte des trois équations suivantes :

- calcul du débit total entrant dans le nœud n vers la destination j :

$$q_{n,j} = \sum_{m \in I_n} Q_m \Gamma_{m,j}, \forall i \in S^n, S^n = \bigcup_{m \in O_n} S_m$$

avec :

- Q_m le débit sortant du tronçon entrant du nœud n,
- I_n l'ensemble des tronçons entrants du nœud n,
- $\Gamma_{m,j}$ le pourcentage des véhicules par destination j à la fin du tronçon m,
- O_n l'ensemble des tronçons sortants du nœud n.

– calcul du débit sortant :

$$q_m = \sum_{j \in S_m} q_{n,j} \beta_{n,j}^m$$

selon le taux d'affectation $\beta_{n,j}^m$ qui est fixé ou calculé selon une stratégie de guidage.

– calcul du pourcentage des véhicules sortants :

$$\gamma_{m,j} = \beta_{n,j}^m \frac{q_{n,j}}{q_m}$$

Ces équations affectent les véhicules entrants dans un nœud en fonction de leurs destinations et conformément aux taux d'affectation.

Affectation sur un carrefour : L'affectation sur un carrefour est plus complexe que sur un nœud autoroutier, car la capacité et la topologie du carrefour sont prises en compte. Le modèle affecte le trafic dans un carrefour selon $\alpha_{n,j}^{m,m'}$ le taux d'affectation des véhicules du mouvement (m, m') entrant dans le carrefour n vers la destination j. Les équations du modèle sont les suivantes :

– le débit total entrant est obtenu par :

$$q_{n,j,e}^{m'} = \sum_{m \in I_n} \alpha_{n,j}^{m,m'} \Gamma_{m,j} Q_m$$

– l'équation de conservation du nombre de véhicules présents dans la zone d'échange associée à la sortie m' s'écrit :

$$N_n^{m'}(k+1) = N_n^{m'}(k) + T\left[Q_{n,e}^{m'}(k) - Q_{n,s}^{m'}(k)\right]$$

avec $Q^{m'}_{n,e}(k)$ et $Q^{m'}_{n,s}(k)$ les débits totaux, respectivement entrant et sortant, de la zone d'échange. Rapportée à chaque destination cette équation devient :

$$N^{m'}_{n,j}(k+1) = N^{m'}_{n,j}(k) + T\left[Q^{m'}_{n,j,e}(k) - Q^{m'}_{n,j,s}(k)\right]$$

– le débit total sortant par m' est donné par :

$$q^{m'}_{n,j,s} = \gamma^{m'}_{n,j}(k)\, Q^{m'}_{n,s}$$

avec

$$\gamma^{m'}_{n,j}(k) = \frac{N^{m'}_{n,j}(k)}{N^{m'}_{n}(k)}$$

le pourcentage de véhicules présents dans la zone d'échange associée à la sortie m' vers la destination j.

Ces équations décrivent l'entrée des véhicules dans le carrefour par mouvement et par destination, et les affectent selon leurs destinations conformément aux taux d'affectation.

1.4 Modèles d'affectation dynamique fondés sur un modèle macroscopique analytique de l'écoulement

Un certain nombre de modèles d'affectation dynamique ont pour caractérisque particulière d'utiliser un modèle d'écoulement simplifié, par comparaison avec les modèles d'écoulement évoqués ci-avant. Nous qualifions ces modèles d'écoulement d'*analytiques*, et présentons ci-après des modèles d'affectation dynamiques fondés sur ce type de modèles d'écoulement.

1.4.1 Modèle de Kuwahara et Akamatsu

En 1997, Kuwahara et Akamatsu [KA97] ont proposé un modèle macroscopique et dynamique d'affectation réactive du trafic avec file d'attente. Les variables du modèle tels que le volume et le temps sont représentées en fonction du temps, c'est-à-dire par des profils temporels. Deux contraintes physiques sont repectées : l'une est la loi conservation du flux, et l'autre est la discipline "non-dépassement". Les deux contraintes sont formulées mathématiquement pour caractériser l'écoulement du trafic.

Bien que le modèle n'explicite pas très clairement les quatre problèmes d'affectation, on peut les comprendre suivant : la formation des services est réalisée par la méthode de PCC. Le temps de traversée de PCC est déterminé de manière réactive, d'après le temps de traversée des arcs.

Le choix des usagers et le chargement en volume sont combinés. A partir de l'état de concurrence des PCCs, le volume arrivé à un nœud, qui est déterminé par le problème d'écoulement du flux, est réparti sur les arcs à l'aval du nœud selon quatre principes de répartition.

Le problème de l'écoulement du flux fournit le volume sortant et le temps de traversée d'après le volume entrant, le temps libre et la capacité en sortie, en utilisant un modèle de file d'attente verticale.

Dans le modèle, le problème d'affectation est résolu par un algorithme comprenant six étapes. L'algorithme évalue le volume entrant sur tous les arcs du réseau sur un intervalle de temps. Cet intervalle est soumis à la contrainte suivante : il ne doit pas excéder le temps de parcours libre de tout arc du réseau.

1.4.2 Modèle de Bin Ran et al.

Bin Ran et al. [SRHC00] ont intégré un modèle de type dispersion de peloton à leur modèle analytique d'affectation dynamique. L'objectif de cette intégration est d'améliorer la représentation de propagation du flux, en respectant la loi FIFO.

Pour l'écoulement, le volume sortant et le temps de traversée effectif sont obtenus par l'analyse d'évolution de la file d'attente et la congestion. D'autre

part, le modèle considère à la fois sur la contrainte de la capacité en sortie de l'arc traité et de la capacité en entrée de l'arc à l'aval pour déterminer le volume sortant. La capacité en entrée est divisée par le pourcentage d'entrée pour chaque arc initial.

Dans le modèle de dispersion de peloton, il y a deux méthodes proposées pour la simulation : le méthode *linkwise* et la méthode *stepwise*. La méthode *linkwise* simule le trafic sur chaque arc, en traitant les arcs successivement et indépendamment pour toute la durée prise en compte. Cette méthode ne peut donc pas représenter l'évolution dans l'espace des files d'attente. Par contre, la méthode *stepwise* simule le trafic sur l'ensemble du réseau par pas de temps (comme dans une simulation microscopique). Les phénomènes de congestion peuvent donc être représentés.

L'implémentation de ces deux méthodes aboutit à des modèles analytiques d'affectation différents : dans le premier cas *link-based*, c'est-à-dire privilégiant les variables relatives aux arcs, et dans le second cas *path-based*, privilégiant les variables relatives aux chemins.

Le modèle affecte le trafic sur un réseau par classes d'usagers, une classe appartenant à l'un des trois types suivants :

Classe 1 *fixed route*

Classe 2 *stochastic dynamic user optimal route*

Classe 3 *dynamic user optimal*.

Seules les classes de type 2 et 3 apparaissent dans l'expression du choix des usagers, car les itinéraires des classes de type 1 sont fixés.

L'algorithme des moyennes successives est utilisé pour combiner le choix des usagers et la dispersion de peloton, au cours de deux itérations. La première, qualifiée d'*itération externe*, a pour objectif de tester la convergence des volumes déterminés par le modèle d'écoulement, et la seconde, qualifiée d'*itération interne*, teste la convergence des volumes déterminés par l'affectation du trafic.

1.4.3 Modèle de Gentile et al.

Un modèle d'affectation dynamique du trafic journalier a été développé par des auteurs de l'université de Rome [BGP05]. Le modèle caractérise un équilibre offre-demande par un cercle de dépendance des sous-modèles qui établissent les caractéristiques de la capacité de l'arc, ensuite du chemin, puis les choix d'itinéraire optimal, et finalement les flux sur chaque chemin et chaque arc en fonction de l'horaire journalière. Les variables du modèle sont des profiles temporels affines par morceaux qui représentent le temps et le volume.

Le problème d'écoulement est traité à la fois les trois parties de l'arc dont l'entrée, la sortie, et section courant (le long de l'arc). Pour tous les deux sections d'entrée et de sortie, le modèle utilise un seul algorithme qui détermine l'heure de sortie et le volume sortant de la section traitée. La méthode de calcul est l'analyse d'évolution de la file d'attente verticale, dont les hypothèses sont le volume entrant, la capacité constante en section, l'horaire de départ. Pour la section courant, un nouveau modèle pour l'écoulement nommé *Onde Cinématique Moyenne* est appliqué. La nouvelle méthode représente la congestion à cause de l'interaction interne de l'arc entre les pelotons, qui sont divisés par morceaux du profil de volume entrant. L'interaction de pelotons de véhicules sera apparue lorsqu'il y a une contrainte de la capacité en sortie. Grâce à la théorie de l'onde de choc, et la méthode de calcul de la vitesse des pelotons successifs, la vitesse de pelotons et de l'onde de choc est déterminée. Donc, le temps de traversée est évalué selon ces vitesses et le long de l'arc. Enfin, le volume sortant est calculé à partir de volume entrant et le temps de traversée calculé.

A partir des résultats d'une application numérique, les auteurs ont fait une comparaison entre nouveau modèle OCM et autres modèles existant. En générale, le modèle OCM est le plus efficace, dont les résultats sont exacts et réalistes d'une part, et le coût informatique est le moins cher d'autre part.

1.4.4 LADTA

Le modèle LADTA de Leurent [Leu03] est situé dans la famille des modèles analytiques d'affectation avec stockage vertical. Ce modèle simule les interactions entre l'offre et la demande de manière dynamique et macroscopique sur un réseau routier. Ses variables sont des profils temporels de volume, de temps et de prix, qui portent sur un itinéraire ou une relation origine-destination ou un arc, cela par classe d'usagers.

La mise en oeuvre opérationnelle du modèle LADTA constituant l'un des objectifs de cette thèse, une présentation beaucoup plus détaillée en est donnée au chapitre 3.

Chapitre 2

Etude de l'utilisation des modèles dynamiques en France

2.1 Objectif de l'étude

L'un des objectifs de cette thèse est la mise en œuvre d'un modèle d'affectation dynamique du trafic routier, avec pour objectif opérationnel une application numérique sur le site de la Vallée du Rhône. La revue des modèles utilisés au sein du réseau technique est le point de départ logique d'un tel travail. L'objectif de cette étude est de prendre connaissance de l'état de la pratique dans la modélisation dynamique du réseau routier i.e. :

- des modèles utilisés ;
- du type de réseau d'étude et de sa taille ;
- des méthodes d'estimation des matrices O-D ;
- des résultats obtenus.

2.2 Méthode utilisée

Cette étude a été réalisée par une enquête à questionnaire direct.

2.2.1 Liste des destinataires de l'enquête

La liste des destinataires a été établie à l'aide des renseignements fournis par l'équipe CSTR au SETRA. Elle figure en annexe B.

La plupart des personnels contactés travaillent au sein du Réseau Scientifique et Technique du Ministète de l'Equipement (RST), principalement en CETE et au CERTU. Les entretiens ont été réalisés par téléphone et par courrier électronique. Classiquement le premier contact était téléphonique et suivi de l'envoi du questionnaire par courrier électronique.

2.2.2 Période de réalisation de l'enquête

L'étude a été réalisée durant le dernier trimestre de l'année 2005.

2.2.3 Liste des questions

La fiche de questionnaire comportait cinq sections, numérotées de **S1** à **S5**, détaillées ci-après.

S1. Logiciels utilisés

- Nom du logiciel
- Fournisseur
- Informations complémentaires

S2. Applications traitées et en cours de traitement

- Site(s) étudié(s)
- Période de simulation (HPM/HPS)
- Nombre de relations O-D
- Nombre de carrefours/intersections
- Nombre d'arcs
- Disponibilité des documents d'étude

S3. Constitution des données

- Constitution de la matrice O-D
- Constitution des temps de parcours
- Constitution des itinéraires (fixes ou calculés)

S4. Commentaires sur les logiciels/les applications

S5. Contacts principaux pour des échanges techniques sur l'affectation dynamique

2.3 Les acteurs du RST

Nous n'avons eu que deux réponses positives concernant l'affectation dynamique. Les autres personnes interrogées, soit n'étaient pas joignables, soit n'utilisaient pas l'affectation dynamique dans le cadre de leur travail.

2.3.1 Au CETE Nord-Picardie

S1. Logiciel utilisé

DynaSim, développé par la société DYNALOGIC.

S2. Applications réalisées/en cours de traitement

Site : Axe central de circulation à Tergnier dans l'Aisne (ancienne route nationale, dénommée boulevard Gambetta).

Période : heure de pointe du soir (16h30 - 17h30).

Nombre de relations O-D : environ 160.

Nombre d'intersections : 4 carrefours à feux, 12 intersection secondaires en T/en croix (STOP ou Cédez-le-Passage).

Nombre d'arcs : DynaSim repose sur la modélisation des trajectoires des véhicules décomposées en arcs (cette décomposition en arcs est assez libre ; le modélisateur doit juste s'assurer qu'au niveau de chaque intersection, il y a l'extrémité d'un arc)

Disponibilité de documents d'études : il existe un rapport d'études consultable.

S3. Constitution des données

Constitution de la matrice O-D : la matrice O-D a été construite à partir :

- d'une enquête cordon O-D VL/PL par relevé des plaques minéralogiques sur la période [16h-18h]. L'enquête cordon comportait des postes par sens au niveau de chaque branche des carrefours à feux ainsi bien entendu que des postes en entrée et en sortie de l'axe étudié (soit 18 postes deux sens).
- des comptages directionnels réalisés le jour de l'enquête cordon au niveau de chacune des intersections secondaires (comptage sur une demi-heure, par sens de circulation, pour VL/PL).

Une première matrice a été construite à partir de l'enquête cordon, puis complétée à partir des comptages directionnels. Cette matrice a ensuite été éclatée sur l'ensemble des relations O-D.

Mesure des temps de parcours : le temps de parcours a été mesuré le jour de l'enquête par un véhicule, ainsi que sur d'autres journées. Le modèle a été calé sur la congestion au niveau des feux. Les longueurs de queue ont été recensées au moment des phases de rouge.

Constitution des itinéraires : les itinéraires ont été calculés. Il n'y a pas plusieurs itinéraires possibles par O-D, car le réseau modélisé était assez simple.

S4. Commentaires sur le logiciel/l'application

- DynaSim est un bon outil de communication car le rendu sous forme de simulation plaît beaucoup aux services techniques et aux élus. Ce type de simulation ne demande pas beaucoup de données en entrée.
- DynaSim demande beaucoup de temps pour le calage du modèle. Il faut constamment redessiner certains arcs (trajectoires des véhicules), revoir la signalisation ou insérer une signalisation fictive pour obtenir des comportements de véhicules crédibles. La manipulation du logiciel sur cette phase de calage s'est révélée ardue et fastidieuse.
- Les utilisateurs ne sont pas persuadé que le modèle ait parfaitement représenté la situation actuelle de congestion.

2.3.2 Au CETE Méditerrannée

S1. Logiciel utilisé

Aimsun NG, développé par l'Université Polytechnique de Catalogne et commercialisé par la société T.S.S.

S2. Applications traitées et en cours de traitement

Site : le réseau d'étude se compose de plusieurs parties :
- un secteur autoroutier entre le péage de Bandol et celui de Cuers ;
- autoroute A570, ainsi que le réseau primaire autour de l'autoroute ;
- les grands axes de traversée de Toulon.

Période : heure de pointe du soir

Nombre de relations O-D : 150 centroïdes

Nombre d'intersections : 677 nœuds , dont 150 carrefours à feux

Nombre d'arcs : 2929 sections qui représente 300 kml (il y a des sections à plusieurs voies).

Disponibilité de documents d'études : il existe un rapport d'étude intitulé *Prise en main du logiciel* AIMSUN *NG*.

S3. Constitution des données

La codification détaillée du réseau est à la base de la micro-simulation, qui se veut pouvoir retraduire les conditions réelles de circulation. Toutes les caractéristiques et les attributs du réseau sont définis pour une affectation microscopique.

Constitution de la matrice O-D : un système de matrice O-D évoluant selon la période de la journée pour chacun des types de véhicules sera constitué par l'utilisation un macro-modèle statique EMME/2.

Temps de parcours : le temps de parcours sera constitué par l'observation.

Paramètres du modèle d'écoulement : les lois de poursuite classiques de la simulation microscopique sont appliquées, impliquant donc des paramètres à calibrer.

Pour le calage, plusieurs calages/validations du modèle ont été proposés, mais de manière très générale, aucune méthode de calage n'a été présentée.

S4. Commentaires sur le logiciel/l'application

L'équipe contactée au CETE Méditerrannée vient de prendre en main le logiciel AIMSUN pour tester des stratégies de gestion de trafic. C'est un micro-simulateur performant pour affecter les déplacements sur le réseau de manière microscopique grâce à des lois paramétrées par défaut tels que : loi de poursuite des véhicules, stratégie de changement de voie, affectation prédictive ou réactive sur l'itinéraire avec simulation en pourcentages de flux aux intersections ou par matrice O-D, injection stochastique des véhicules sur le réseau à partir des flux de génération avec le temps d'arrivée suivant

une loi exponentielle, uniforme, normale ou constante, loi d'insertion dans une intersection avec la possibilité de modéliser les rond-points.

De plus, cet outil permet de prendre en compte les actions : la priorité des transports en commun, la régulation d'accès, les feux adaptifs, la gestion des incidents, les panneaux à messages variables, etc.

AIMSUN, un logiciel commercial, a été utilisé pour plusieurs études de trafic en Europe. Pour les études de CETE Méditerranée, une proposition d'une méthodologie d'utilisation combinée de EMME/2 et AIMSUN a été présentée.

2.3.3 A l'INRETS

S1. Logiciel utilisé

METACOR, développé par l'axe *Modélisation et simulation de système de transport complexe* du GRETIA/INRETS, dont les applications ont été réalisées dans le cadre de plusieurs projets européens : EUROCOR (site FR, NL), DACCORD (site FR, NL), QOUVADIS (sites Glasgow (UK), Aalborg (DK)). La plateforme multi-modèles MAGISTER développée par l'axe *Modélisation et simulation de système de transport complexe* du GRETIA/INRETS, a été utilisée dans le cadre du projet européen EURAMP en 2006 [HS06].

S2. Applications réalisées/en cours de traitement

Site : le réseau autoroutier de l'Ile de France dans le cadre du PDU en Ile de France 2001 : Installation du contrôle d'accès généralisé (300 accès classiques, 50 convergents autoroutiers).

- 200 km d'autoroute
- 70 accès classiques
- 50 contrôlables, dont 20 convergents autoroutiers

Période : 6h - 21h :30, pas de temps de simulation 5 sec.

Nombre de tronçons : 250, dont 622 segments.

La dimension du problème : 1 800 000 (très grande dimension).

Disponibilité de documents d'études : il existe une présentation d'étude de trafic du réseau autoroutier IDF.

Site : la Défense (1998) dans le cadre d'une analyse et évaluation de l'impact de l'aménagement du boulevard Circulaire de la Défense : passage en urbain à horizon 2005.

Période : inconnue.

Nombre de tronçons : 285, dont 60 km de longueur.

Nombre de nœuds : 118 carrefours (30 à feux), 88 nœuds.

Nombre d'origines : 41.

Nombre de destination : 44.

Site : IDF (BP+A6W) pour une étude de l'impact du contrôle d'accès isolé et coordonné et du contrôle d'accès autoroute/autoroute. Le réseau simulé comprend :

- BPI (sud) : 9 kms avec 7 " on-ramps "
- A6a : 10 kms avec 3 " on-ramps "
- A6b : 10 kms avec 5 " on-ramps "
- A6w : 20 kms avec 6 " on-ramps "
- MTM : A6a/BPI

Site : la partie sud intérieure du corridor périphérique de Paris (porte d'Italie et porte de Sèvres incluant le réseau de surface (les Boulevards des Maréchaux).

Période : 6h-12h.

Nombre de tronçons : 122.

Nombre de nIJuds : 38 (23 carrefours, 15 nIJuds autoroutiers).

Nombre d'origines : 24.

Nombre de destination : 23.

S3. Constitution des données

Constitution de la topologie du réseau : la topologie du réseau mixte est modélisée par un graphe orienté, dont les éléments sont identifié automatiquement par l'interface. Les paramètres géométriques saisis ou par défaut sont chargés automatiquement dans des boîtes de dialogue.

Constitution de la matrice O-D : l'estimation de la matrice O-D à partir des comptages utilisant les estimateurs tels que DAVIS (cas de la Défense), MCONTRAM (cas de la Ville D'Aalborg). Pour le cas du réseau autoroutier de l'IdF, les demandes aux origines et les pourcentages directionnels aux divergents sont estimés par une moyenne historique des 4 jours sélectionnés.

Constitution des paramètres : les paramètres utilisés dans les équations du modèle sont déterminés par le calibrage et validés. Par exemple, dans le cas du réseau autoroutier de l'IdF, le calibrage a été réalisé quantitativement et qualitativement. Le calibrage quantitatif a été lancé par axe (le réseau est décomposé par axe), dont deux jours de calibrage et deux jours de validation. Les paramètres trouvés par le calibrage par axe ont été intégrés pour le réseau en entier. La congestion est reconstituée en utilisant la technique de "essai/erreur". Et le système SYTADIN a été utilisé comme référence.

S4. Commentaires sur le logiciel/l'application

Le logiciel METACOR et la plate forme MAGISTER sont dotés d'une interface graphique conviviale, développée par la société PHOENIX ISI. Elle permet à l'utilisateur de saisir, de configurer et de dessiner rapidement et facilement le réseau, et les paramètres du modèle. Une visualisation dynamique

permet à l'utilisateur de visualiser les états du trafic en cours de simulation. Avec plusieurs applications européens de moyenne et grande dimension (réseau autoroutier IdF), METACOR a permis le développement et l'évaluation hors ligne de plusieurs stratégies de régulation (contrôle d'accès autoroute sur autoroute, régulation intégrée sur un corridor) et des stratégies d'information des usagers via les PMVs. La plate forme MAGISTER permet des simulations suivant différents modèles de représentation du trafic (LWR : 1er ordre classique ; LWRACB : 1er ordre à accélération bornée ; METACOR ; ARZ : 2ème ordre). Ainsi dans les années à venir, les modèles théoriques seront implantés et rendus disponibles dans MAGISTER pour répondre aux besoins des recherches et des études en simulation (Modèle microscopique, Modèle hybride (macro+micro)).

Concernant les applications, MAGISTER a été utilisé pour l'évaluation de la régulation isolés et coordonnée sur l'autoroute au sud de Paris (A6W). C'est un outil en cours de développement et enrichissement. C'est un équipement de recherche et d'application destiné aux chercheurs et aux RST, aux exploitants etc.

2.4 Commentaires et synthèse

Cette section vise, à partir des résultats obtenus lors de l'enquête réalisée précédemment, à présenter l'état d'application en pratique des modèles/logiciels d'affectation dynamique au sein du RST.

2.4.1 Affectation dynamique/statique

Dans les bureaux d'études publics français, tels que les CETEs, le CETUR, le SETRA, etc., la plupart des études du trafic routier sont des études statiques réalisées à l'aide de logiciels connus :

- TRANSCAD pour le SETRA, le CETE de Metz, le CETE d'Aix ;
- TRIPS pour la DREIF, le CETE de Lyon ;
- EMME/2 pour le CETE de Lille, le CETE d'Aix ;
- DAVISUM pour les CETEs de Nantes, du Sud-ouest, de Lyon, de Metz.

Pour l'application en pratique des modèles dynamiques, il y a peu d'études réalisées :

- le CETE Méditerrannée commence à prendre en main le logiciel AIMSUN ;
- le CETE Nord a utilisé le logiciel DYNASIM pour une petite application dans le cadre d'une gestion locale de trafic ;
- l'outil METROPOLIS a été utilisé pour une application sur le réseau francilien ;
- l'outil METACOR a été utilisé pour quelques applications en IdF, notamment les corridors du réseau autoroutier francilien, et divers sites à l'étranger ;

2.4.2 Type de modèle utilisé

La plupart des logiciels de simulation dynamique du trafic utilisés au sein du RST sont des modèles microscopiques : AIMSUN, DYNASIM, INTEGRATION.

Un seul modèle mésoscopique, METROPOLIS, a été appliqué au site Ile de France.

Les modèles dynamiques d'affectation macroscopique ne sont pas encore utilisés en pratique au sein du RST.

Seuls des modèles issus de la recherche, avec des petites applications au niveau du cas d'école ont été mis en opération. Citons en particulier CIUDADSIM [LMQW03], STRADA [BLL96].

Au sein du GRETIA/INRETS, le modèle macroscopique METACOR [SE94, SP98] a traité des cas en IdF et à l'étranger pour les corridors (réseau mixte de moyen taille). Un développement d'une plateforme multi-modèles MAGISTER est en cours de réalisation.

2.4.3 Echelle d'application

Les études de trafic ont été réalisées par l'utilisation des modèles dynamiques en France jusqu'à aujourd'hui seulement pour des réseaux locaux de

petite taille. Diverses applications en pratique ont été mises en œuvre au niveau urbain, mais aucune au niveau interurbain.

Nous pensons que l'une des raisons de cet état de fait est que les outils mis en oeuvre sont microscopiques.

2.4.4 Constitution des données

Les données concernant le réseau telles que le temps de parcours et les paramètres d'affectation sont constituées par l'observation. La matrice O-D est constituée par des comptages et des outils statiques, mais la méthode d'estimation de la matrice O-D dynamique n'est pas encore présentée clairement.

2.4.5 Expériences d'application

Les personnes qui utilisent les logiciels dynamiquesont commenté les applications réalisées ou en cours de traitement. Ces critiques sur les modèles utilisés ont un rôle important pour notre application dans le cadre de la thèse.

Cela permet de démontrer la nécessité de mise en pratique de notre outil/modèle, qui est un modèle macroscopique pour des études interurbaines.

Deuxième partie

Prototypage du modèle LADTA

Chapitre 3

LADTA : un modèle analytique d'affectation dynamique du trafic

3.1 Introduction

Un modèle d'affectation dynamique du trafic routier a été développé par Leurent [Leu03, Leu04]. Ce modèle fait partie de la famille de modèles analytiques d'affectation avec stockage vertical, et simule les interactions entre l'offre et la demande de manière dynamique et macroscopique sur un réseau routier.

L'objectif de LADTA est de modéliser les déplacements sur un réseau de manière à la fois réaliste et aussi simple que possible. Le réalisme du modèle porte sur la représentation du réseau et des services, sur la description des usagers avec leur comportement, sur la simulation de l'écoulement et des phénomènes de congestion, et ainsi sur la représentation des files d'attente et les priorités entre flux.

Du fait de cet objectif prioritaire de réalisme, le modèle est complexe. Cette complexité a été limitée de deux manières : premièrement, en adoptant une représentation macroscopique des flux, comme dans le cadre d'une simulation statique, deuxièmement en négligeant les aspects dynamiques internes à un arc autres que le phénomène de file d'attente.

La construction du modèle a été menée selon les principes de l'analyse

physique et économique, ainsi que mathématique en vue d'intégrer les autres problèmes indépendants que sont : (i) la formation des services, (ii) le choix des usagers, (iii) le chargement en volume, et (iv) l'écoulement du flux.

Chacun de ces problèmes est modélisé par une ou plusieurs équations, et résolu par un algorithme efficace. Une version restreinte du modèle a été programmée en 2003 par l'auteur de cette thèse dans le cadre d'un stage de DEA Transport, sous la forme d'un prototype qui a été validé par deux exemples simples. Ce prototype, codé en SciLab, ne permet de traiter que des réseaux de petite taille. Afin de disposer d'un outil performant, une nouvelle version qui permet de traiter des réseaux de grande taille a été programmée en C++ par Vincent Aguiléra en 2005.

Ce chapitre comprend huit sections. La section 3.2 rappelle rapidement les principes, l'architecture et caractéristique d'équilibre offre-demande du modèle Ladta.

Une version restreinte du modèle est ensuite présentée avec dans un premier temps la définition des variables utilisées (section 3.3), puis la présentation des quatre problèmes fondamentaux que sont : la formation des services (section 3.4), le choix des usagers (en 3.5), le chargement en volume (en 3.6) et l'écoulement du flux (en 3.7). La question du calcul de l'équilibre offre-demande est ensuite abordée (en 3.8).

3.2 Généralités sur le modèle Ladta

Dans le modèle Ladta l'offre de transport est constituée d'un ensemble de service de déplacement, eux-même décrits comme des ensembles de couples itinéraire-horaire de départ. La demande est représentée par un ensemble d'usagers, qui utilisent les services de transport selon le principe de l'optimum individuel : chacun choisit le service qui maximise son utilité privée.

Ce modèle prédit les interactions entre offre et demande en fonction du temps dans chacun des quatre problèmes suivants :

la formation des services : un service de déplacement est caractérisé par un itinéraire sur le réseau et un horaire de départ. Le problème de la

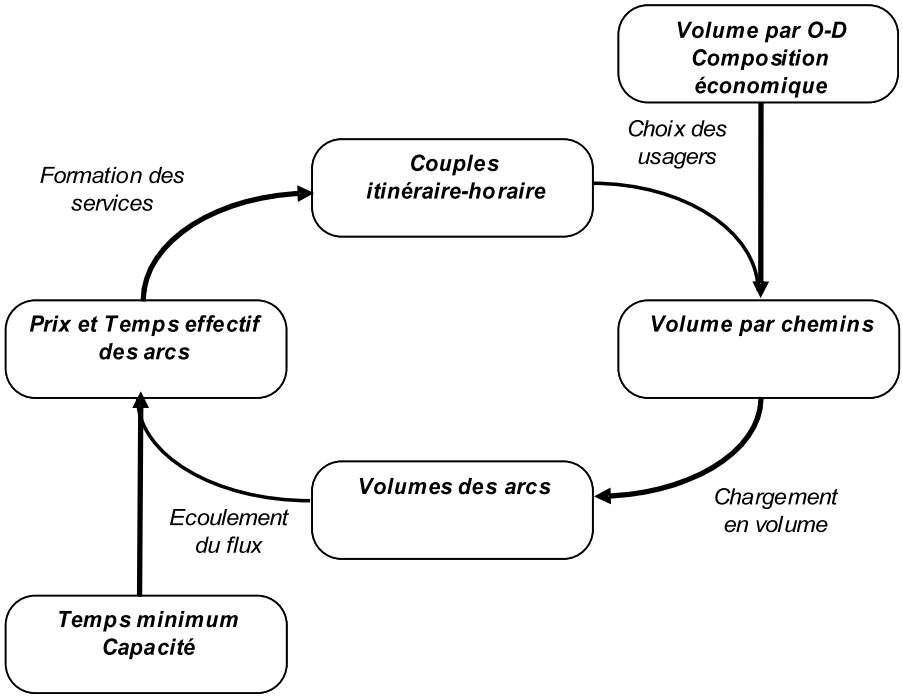


FIG. 3.1 – Quatre relations causales.

formation des services concerne la formation des services optimaux et donc de leurs temps de traversée et de leurs prix, d'après les temps et les prix des arcs du réseau.

le choix des usagers : chaque usager choisit le service optimal dont le coût généralisé est minimal. Par classe d'usagers, le volume de demande est distribué entre les services optimaux en fonction de l'horaire de départ.

le chargement en volume : le volume de chaque arc dépend des volumes des chemins qui empruntent l'arc et du temps de traversée, donc des volumes des services optimaux.

l'écoulement du flux : le volume sortant d'un arc dépend du volume entrant, de la capacité en sortie de l'arc, et du temps de traversée minimum. Donc, le temps minimum se transforme en temps effectif.

La figure 3.1 représente les quatre problèmes, leur situations respectives et leurs complémentarités. Elle fait apparaître un cercle de dépendances, caractéristique d'un problème d'équilibre offre-demande. Dans le cadre de cette

thèse, nous ne considèrons qu'une version restreinte du modèle LADTA qui calcule l'équilibre offre-demande avec une seule classe d'usagers, sans péage, sans élasticité de la demande, et sans d'horaire de départ. Dans les parties qui suivent, nous précisons successivement chacun des problèmes de la version restreinte, avec leurs analyses physiques et économique, en explicitant les équations mathématiques correspondantes. Puis, un algorithme efficace est proposé pour chaque problème. Enfin, un algorithme d'équilibrage utilisant la méthode des combinaisons convexes est appliqué pour calculer l'équilibre offre-demande représenté par le cercle de dépendances ci-dessus. Les analyses, les équations et les algorithmes du modèle complet ont été modifiés afin d'être mis en correspondance avec sa version restreinte.

3.3 Variables du modèle

Avant de présenter chaque problème du modèle, nous définissons d'abord les objets, les ensembles, les variables d'offre et de demande qui seront utilisés pour l'affectation du trafic dans les parties qui suivent.

3.3.1 Objets et ensembles

h le temps d'horloge, à valeurs dans un ensemble H, qui repère l'horaire de passage des usagers sur une position donnée du réseau. On note h_0 un instant de référence dans H.

N ensemble de nœuds n.

A l'ensemble des arcs a.

I l'ensemble des relations O-D i.

R_i l'ensemble des itinéraires r par relation O-D i.

T_i l'ensemble des tranches d'horaire d'entrée composé par la concurrence entre les temps des chemins par relation O-D i.

3.3.2 Variables d'offre

On caractérise l'offre par des variables de temps de traversée, d'horaire de passage, et de volume entrant ou sortant, cela par horaire d'entrée, au niveau des arcs ou des itinéraires. Pour un horaire d'entrée h, au niveau d'un arc a :

$t_a(h)$	le temps efficace de traversée en fonction de l'heure d'entrée h en arc a.
$t_{a0}(h)$	le temps minimum de traversée, correspondant à l'absence de toute attente.
$H_a(h) = h + t_a(h)$	l'horaire de sortie effectif de l'arc.
$H_{a0}(h) = h + t_{a0}(h)$	l'horaire d'arrivée minimal en sortie.
$\overline{H}_a(\hbar) = H_a^{[-1]}(\hbar)$	l'horaire de départ effectif en fonction de l'horaire de sortie.
$\overline{H}_{a0}(\hbar)$	l'horaire maximal d'entrée, qui est la réciproque de l'horaire d'arrivée minimal en sortie $H_{a0}(h)$.
$K_a(h)$	la capacité en sortie cumulée jusqu'à h.
$X_a^+(h)$	le volume entrant cumulé jusqu'à h.
$X_a^-(h)$	le volume sortant cumulé jusqu'à h.

Au niveau d'un itinéraire r, on adapte les notations en remplaçant l'indice a par l'indice r.

3.3.3 Variables de demande

Dans le cas de la version restreinte du modèle complet, la seule variable considérée pour une relation O-D i donnée est le volume de demande Q_i.

3.4 La formation des services

Le problème de formation des chemins[1] a pour objectifs de :

- caractériser les chemins, en formant leurs temps de traversée à partir des temps des arcs ;
- sélectionner les chemins efficaces qui présentent un temps minimum selon les horaires d'entrée des usagers, pour une relation origine-destination.

L'algorithme utilisé est un algorithme de propagation sur un réseau dynamique [Leu06]. Donc, les inputs du problème sont les profils temporels de temps de traversée effectif des arcs qui composent les chemins entre origines et destinations. Les outputs du problème sont les chemins efficaces avec leur topologie et leur profil temporel de temps de traversée.

3.4.1 Principes physiques et économiques

La caractérisation d'un chemin r repose sur un premier principe d'agrégation : le temps de traversée $t_r(h)$ dépend des $t_a(h')$ pour $a \in r$ et $h \geq h'$.

Le deuxième principe concerne la propriété d'efficacité : un chemin r est efficace pour une relation O-D et pour une tranche d'horaire d'entrée τ si son temps de traversée effectif est minimal sur τ.

On se limite ici aux chemins sans multiplicité, i.e. qui n'utilisent pas plus d'une fois un arc donné.

3.4.2 Formulation mathématique

Pour former le temps d'un chemin r, on calcule récursivement les temps des arcs a qui composent r, grâce aux sous-chemins $r > a$:

$$t_r(h) = \sum_{a \in r} t_a\left(h + t_{r>a}(h)\right) \tag{3.1}$$

[1] Dans le modèle complet, ce problème est nommé "formation des services" car on forme des services de transport, c'est-à-dire des couples itinéraire-horaire de départ. Pour cette version restreinte, on considère seulement les itinéraires avec leur temps de traversée.

En fonction de l'horaire de sortie $\hbar$, cette formule devient :

$$t_r^- (h) = \sum_{a \in r} t_a^- \left(\hbar - t_{a<r}^- (\hbar) \right) \tag{3.2}$$

Pour exprimer la propriété d'efficacité des chemins, les chemins optimaux pour une relation origine-destination sont définis de la façon suivante :

$$t_r^* (\tau) = \min \left\{ t_r (\tau), r \in R_i, \tau \in T_i \right\} \tag{3.3}$$

i.e. le temps de traversée minimum pour la tranche d'horaire d'entrée τ .

$$R_i^* = \left\{ r \in R_i, \tau \in T_i, t_r (\tau) = t_r^* (\tau) \right\} \tag{3.4}$$

i.e. l'ensemble des chemins optimaux pour la relation O-D i.

T_i est l'ensemble de tranches d'horaire d'entrée qui est composé par la concurrence de temps de traversée entre les chemins d'une relation i.

De plus la cohérence physique implique une condition de non-dépassement :

$$h \leq h^{'} \Rightarrow h + t_a (h) \leq h^{'} + t_a \left(h^{'} \right) \tag{3.5}$$

3.4.3 Algorithme de propagation sur un réseau dynamique

L'algorithme de propagation est établi pour une destination d. Toutes les origines de couple $i \approx (O, d)$ seront traitées simultanément.

Un chemin r ayant pour nœud final d est représenté par un atome $\alpha_r = \left[n, a, \overline{H}_r, t_r^-, \theta \right]$, avec :

- $n = n_r^+$ le nœud initial du chemin r ;
- $a = a_r^+$ l'arc initial du chemin r, donc $r = a \cup r^{'}$ et $r^{'} > a$ est un sous-chemin de l'arc a ;
- et $\overline{H}_r$ est le profil temporel d'horaire d'entrée en fonction d'horaire sortie de chemin r ;
- α_r est un pointeur sur l'atome de r. Un atome virtuel α_ϕ est introduit pour la référence.

Le détail de cet algorithme est donné en page 83.

3.5 Le choix des usagers

L'objectif du problème du choix des usagers est de distribuer la demande entre les chemins efficaces. Les données du problème sont :

- les itinéraires/chemins efficaces entre origines et destinations du réseau, leurs profils temporels de temps de traversée effectif ;
- les profils temporels de volume de la demande par relation origine-destination.

En sortie, on souhaite obtenir le profil de volume entrant sur chaque itinéraire efficace. Seul le temps de traversée est considéré comme un critère de choix entre les chemins.

3.5.1 Principes économiques

Les principes de comportement économique des usagers sont les suivants :

- Chaque usager d'une relation i est informé des chemins disponibles, qui forment un ensemble R_i.
- Il est parfaitement informé sur le temps de traversée correspondant à son heure d'entrée dans chaque chemin R_i.
- Il choisit le chemin qui minimise son temps de traversée.

Autrement dit, par horaire d'entrée, la demande se concentre sur le chemin dont le temps de traversée est minimal.

3.5.2 Formulation mathématique

Le choix du chemin optimal, ou plus court chemin (PCC), est caractérisé par les équations :

$$dX_r^+(\tau) \geq 0, \forall \tau \in T_i \tag{3.6}$$

$$\sum_{r \in R_i} dX_r^+(\tau) = dQ_i(\tau), \forall \tau \in T_i \tag{3.7}$$

Entrées :
d la destination à traiter
L liste d'attente
Sorties :
M ensemble d'atomes actifs
W_n chemins efficaces pour chaque nœud $n \in N$
début
$M \leftarrow \emptyset$
pour chaque $n \in N$ **faire**
$W_n \leftarrow \emptyset$
fin
$L \leftarrow \{\alpha_\phi\}$
$W_d \leftarrow \{\alpha_\phi\}$
tant que $L \neq \emptyset$ **faire**
Soit $\alpha'_r = \left[n', a', \overline{H}'_r, t'_{r^-}, \theta'\right]$ un atome de L de nombre d'arcs minimal, et θ un pointeur sur α'_r.
$L \leftarrow L \setminus \{\alpha'_r\}$
si α'_r *est efficace par rapport à* $W_{n'}$ **alors**
$M \leftarrow M \cup \{\alpha'_r\}$
$W_{n'} \leftarrow W_{n'} \cup \{\alpha'_r\}$
pour chaque *arc* b *de nœud final* n' **faire**
$r \leftarrow b \cup r'$
$\overline{H}_r = \overline{H}_b \circ \overline{H}'_r$
$t_{r^-}(\hbar) \leftarrow t'_{r^-}(\hbar) + t^-_b\left(\overline{H}'_r(\hbar)\right)$
$\alpha_r = \left[n^+_b, b, \overline{H}_r, t^-_r, \theta\right]$
si α_r *est efficace par rapport à* $W - n$ **alors**
$L \leftarrow L \cup \{\alpha_r\}$
fin
$W_{n'} \leftarrow W_{n'} \setminus \{\alpha \in W_{n'}, \alpha \text{ dominé par } \alpha'_r\}$
fin
fin
fin
pour chaque *nœud* n *qui est une origine vers* d **faire**
pour chaque $\alpha \in W_n$ **faire**
$H_r \leftarrow \overline{H}_r^{(-1)}$
$t_r(h) = H_r(h) - h$
fin
fin
fin

Algorithme 3.1 : Algorithme de propagation

$$t_r(\tau) \geq t_r^*(\tau), \forall r \in R_i \tag{3.8}$$

$$(t_r(\tau) - t_r^*(\tau))\, dX_r^+(\tau) = 0, \forall r \in R_i \tag{3.9}$$

qui signifient que le volume de demande se concentre sur les chemins dont le temps est minimum, et ce pour chaque d'horaire d'entrée.

3.5.3 Un algorithme de répartition

Il s'agit d'énumérer les tranches d'horaire d'entrée qui sont composées par la concurrence entre les temps de traversée des chemins par relation i.

Par relation $i = (o, d)$, on connaît l'ensemble W_o des chemins efficaces depuis o, issus de la formation des chemins.

On remet $X_r^+(h) \leftarrow 0, \forall (h \in H, r \in W_o)$

On énumère les tranches d'horaire d'entrée, en identifiant le chemin r dont le temps est minimum pour chaque , on obtient un ensemble de couples $[\tau, r]$.

Pour chaque $[\tau, r]$ on charge totalement la demande sur le chemin r, et on obtient alors

$$dX_r^+(\tau) = dQ_i(\tau)$$

3.6 Le chargement en volume

L'objectif du problème de chargement en volume est d'établir les profils des volumes entrants pour chaque arc sachant les volumes entrants des chemins.

Les données du problème sont :

- les profils des volumes entrants des chemins r passant par a, i.e. les $\{X_r^+(h), r \wedge a\}$
- Les temps de traversée des portions de ces itinéraires qui sont en amont de a, i.e. les $t_{r>a}(h)$.

3.6.1 Principes physiques

Pour chaque chemin r passant a, les déplacements via r qui entrent dans a jusqu'à h sont exactement ceux entrés dans r à un horaire h' qui vérifie la condition de concordance exprimée par l'équation 3.10 :

$$h' + t_{r>a}(h') \leq h \tag{3.10}$$

Cette condition de concordance est la clef du problème de chargement.

3.6.2 Formulation mathématique du problème de chargement

L'horaire d'entrée en a sachant l'entrée dans r à h' est calculée grâce à l'équation :

$$H_{r>a}(h') = h' + t_{r>a}(h') \tag{3.11}$$

car c'est est exactement l'horaire de sortie du sous-chemin $r > a$ en amont de a.

Sa fonction réciproque calcule l'horaire d'entrée en r sachant l'entrée en a à h :

$$\overline{H}_{r>a}(h) = sup\{h' \in H, H_{u,r>a}(h') \leq h\} \tag{3.12}$$

La condition de concordance est reformulée :

$$h' + t_{r>a}(h') \leq f \Leftrightarrow H_{r>a}(h') \leq h \Leftrightarrow h' \leq \overline{H}_{r>a}(h) \tag{3.13}$$

Alors, la partie $X_r^+\left[\overline{H}_{r>a}(h)\right]$ contribue à $X_a^+(h)$.

En considérant tous les chemins passant par a d'une relation O-D i, la formulation mathématique du problème de chargement est représentée par l'équation de propagation suivante :

$$X_a^+(h) - X_a^+(h_0) = \sum_{i \in I} \sum_{r \in R_i \wedge a} X_r^+\left[\overline{H}_{r>a}(h)\right] - X_r^+\left[\overline{H}_{r>a}(h_0)\right] \tag{3.14}$$

qui calcule le volume entrant de l'arc a pendant $[h_0; h]$.

3.6.3 Algorithme atomique de chargement

L'algorithme atomique de chargement est établi pour une destination d. Toutes les origines de I sont traitées ensemble.

Pour chaque chemin/atome, on considère en fait $\tilde{X}_r^+$ qui cumule X_r^+ et tous les $X_{r'}^+$ des chemins $r' = r'' \cup r$ qui prolongent r.

X_r^+, le volume chargé sur le chemin r, calculé par une distribution de la demande par O-D, résulte du problème de choix des usagers. Pour chaque chemin/atome r, on considère $\tilde{X}_r^+$, qui cumule X_r^+ et tous les $X_{r'}^+$ des chemins r' dont r est un sous-chemin.

Pour chaque arc a, une variable X_a^+ contient le volume entrant en a qui s'oriente vers la destination d.

Le détail de l'algorithme de chargement figure en page 87.

3.7 L'écoulement du flux

Au niveau d'un arc considéré isolément, connaissant les profils temporels de volume entrant, de la capacité en sortie et du temps de traversée minimal, l'objectif du problème est d'établir le profil temporel du temps de traversée effectif de l'arc.

3.7.1 Principes physiques et économiques

Dans le cas d'une seule classe d'usagers, l'écoulement met en œuvre deux principes physiques de limitation du volume sortant :

- le volume sortant est limité par le volume entrant, avec le décalage temporel correspondant au temps de traversée minimal ;
- le volume sortant est limité par la capacité en sortie.

Et un principe économique de minimisation du temps :

- le volume s'écoule aussi vite que possible, la capacité est utilisée pleinement quand un stock (file d'attente) est présent en sortie.

Entrées :
M : issu de la formation des services
$X^+_{r\in R}$: issus du choix des usagers
Sorties :
$X^+_{a\in A}$
début
  pour chaque $a \in A$ **faire**
    $X^+_a \leftarrow 0$
  fin
  pour chaque $r \in M$ **faire**
    $\tilde{X}^+_r \leftarrow 0$
  fin
  pour chaque $i = (o, d)$ **faire**
    pour chaque $o \in O$ **faire**
      pour chaque $r \in W_o$ **faire**
        $\tilde{X}^+_r \leftarrow X^+_r$
      fin
    fin
  fin
  tant que $M \neq \emptyset$ **faire**
    Soit $\alpha_r = [n, a, Hb_r, t_{r^-}, \theta]$ un atome de M de nombre d'arcs maximal.
    $X^+_a(h) \leftarrow X^+_a(h) + \tilde{X}^+_r(h)$
    $X^+_{r'}(h) \leftarrow X^+_{r'}(h) + \tilde{X}^+_r\left(\overline{H}_a(h)\right)$
    avec r' le chemin en aval de r, pointé par θ.
  fin
fin

Algorithme 3.2 : Algorithme atomique de chargement

3.7.2 Formulation mathématique

Le but est d'établir les profils $X_a^-(\hbar)$ et $t_a(h)$ en fonction des profils $X_a^+(h)$, $K_a(h)$, et $t_{a0}(h)$, en suivant les principes énoncés en section 3.7.1.

La limitation du volume sortant par le volume entrant se formule :

$$X_a^-(h + t_{a0}(h)) - X_{a0}^- \leq X_{a0}^+(h) - X_{a0}^+ + N_{a0} \tag{3.15}$$

avec $X_{a0}^+ = X_a^+(h_0)$, $X_{a0}^- = X_a^-(h_0)$, et N_{a0} le volume présent sur l'arc a à l'instant de référence h_0. On pose $\Delta_{a0} = X_{a0}^- - X_{a0}^+ + N_{a0}$.

On définit la fonction d'horaire minimum de sortie $H_{a0}(h) = h + t_{a0}(h)$.

La fonction réciproque de H_{a0} est $\overline{H}(\hbar) = \sup\{h, H_{a0}(h) \leq \hbar\}$.

On peut donc reformuler l'équation 3.15 :

$$X_a^-(h) \leq X_a^+\left(\overline{H}_{a0}(h)\right) + \Delta_{a0} \tag{3.16}$$

La limitation du volume sortant par la capacité d'écoulement se formule :

$$X_a^-(h) - X_a^-(h-s) \leq K_a(h) - K_a(h-s), \forall s \in S_h = [0; h - \inf H] \tag{3.17}$$

En combinant les équations 3.16 et 3.17 au point $h' = h - s$, il vient :

$$X_a^-(h) \leq \inf\left\{X_a^+\left(\overline{H}_{a0}(h-s)\right) + K_a(h) - K_a(h-s) + \Delta_{a0}, s \in S_h\right\} \tag{3.18}$$

En considérant le principe économique de minimisation du temps, il vient :

$$X_a^-(h) = \inf\left\{X_a^+\left(\overline{H}_{a0}(h-s)\right) + K_a(h) - K_a(h-s) + \Delta_{a0}, s \in S_h\right\} \tag{3.19}$$

Le temps de traversée $t_a(h)$ pour l'horaire d'entrée h vérifie $X_a^+(h) + \Delta_{a0} = X_a^-(h + t_a(h))$, et donc :

$$t_a(h) = X_a^{-[-1]}\left(X_a^+(h) + \Delta_{a0}\right) - h \tag{3.20}$$

3.7.3 Calcul du temps de parcours

L'algorithme de calcul du temps de parcours est l'algorithme 3.3.

Entrées :
X_a^+ : profil de volume entrant.
K_a : profil de capacité.
t_{a0} : profil de temps de parcours libre.
Sorties :
t_a : profil de temps de parcours.
début
  Calculer $H_{a0}(h) = h + t_{a0}(h)$
  Calculer le profil réciproque $\overline{H}_{a0}$ de H_{a0}
  Calculer $X_a^-(h)$ selon l'équation 3.19
  Calculer le profil réciproque $X_a^{-[-1]}$ de X_a^-
  En déduire $t_a(h) = X_a^{-[-1]}(X_a^+(h) + \Delta_{a0}) - h$
fin

Algorithme 3.3 : Algorithme de calcul du temps de parcours

3.7.4 Algorithme de calcul du volume sortant

Afin de caractériser le profil du volume sortant d'après l'équation 3.19, nous calculons d'abord le volume arrivé en sortie grâce à l'équation suivante :

$$X_a^{arr}(h) = X_a^+\left(\overline{H}_{a0}(h)\right) + \Delta_{a0} \tag{3.21}$$

En connaissant le volume arrivé en sortant, nous proposons un algorithme qui permet d'évaluer la file d'attente verticale d'après le principe économique d'utilisation pleine de la capacité en sortie (voir algorithme 3.4 en page 90).

Dans cet algorithme, on note $\dot{Y} = \frac{dX_a^{arr}}{dt}$ et $\dot{X} = \frac{dX_a^-}{dt}$. Les notations i_Y, i_K et i_X sont les indices identifiant respectivement un morceau du profil temporel de volume arrivé, de capacité, et de volume sortant. Les instants h_1 et h_2 sont utilisés pour calculer l'état du trafic dans l'arc, autrement dit, pour calculer le volume du stock en sortie $S_1(h_1)$, $S_2(h_2)$, en déduire le volume sortant selon l'état de stock noté par la variable booléenne *Stock*.

Entrées :
X_a^+ : profil de volume entrant.
K_a : profil de capacité.
t_{a0} : profil de temps de parcours libre.
Sorties :
X_a^- : profil de volume en sortie.
début
  $h_1 \leftarrow \inf H$; $S_1 \leftarrow 0$; $i_X \leftarrow 1$; $i_Y \leftarrow 1$; $i_K \leftarrow 1$; $h_x(1) \leftarrow h_1$
  si $\left(\dot{Y}[1] > \dot{K}[1]\right)$ **alors**
    $Stock \leftarrow Vrai$; $\dot{X}[1] \leftarrow \dot{K}[1]$
  sinon
    $Stock \leftarrow Faux$; $\dot{X}[1] \leftarrow \dot{Y}[1]$
  fin
  répéter
    $h_2 \leftarrow \min\{h_K[i_K+1]; h_Y[i_Y+1]\}$
    $S_2 \leftarrow S_1 + (h_2 - h_1) \cdot \left(\dot{Y}[i_Y] - \dot{K}[i_K]\right)$
    si $(Stock = Vrai) \wedge (S_2 < 0)$ **alors**
      $i_X \leftarrow i_X + 1$
      $h_X[i_X] \leftarrow h_1 + S_1 / \left(\dot{K}[i_K] - \dot{Y}[i_Y]\right)$
      $h_1 \leftarrow h_X[i_X]$; $\dot{X}[i_X] \leftarrow \dot{Y}[i_Y]$
      $S_1 \leftarrow 0$; $Stock \leftarrow Faux$
    sinon
      si $h_K[i_K+1] = h_2$ **alors** $i_K \leftarrow i_K + 1$
      si $h_Y[i_Y+1] = h_2$ **alors** $i_Y \leftarrow i_Y + 1$
      $Stock \leftarrow (S_2 > 0) \vee \left(\dot{Y}[i_Y] > \dot{K}[i_K]\right)$
      si $(Stock = Vrai)$ **alors**
        si $\dot{X}[i_X] \neq \dot{K}[iK]$ **alors**
          $i_X \leftarrow i_X + 1$; $h_X[i_X] \leftarrow h_2$; $\dot{X}[i_X] \leftarrow \dot{K}[i_K]$
        fin
      sinon si $\dot{X}[i_X] \neq \dot{Y}[iY]$ **alors**
        $i_X \leftarrow i_X + 1$; $h_X[i_X] \leftarrow h_2$; $\dot{X}[i_X] \leftarrow \dot{Y}[i_Y]$
      fin
      si $(S_2 < 0)$ **alors** $S_2 \leftarrow 0$; $h_1 \leftarrow h_2$; $S_1 \leftarrow S_2$
    fin
  jusqu'à $\dot{Y}[i_Y] > \dot{K}[i_K]$ *ou* $Stock = Faux$;
fin

Algorithme 3.4 : Algorithme de calcul du volume en sortie

3.8 L'équilibre offre-demande

L'équilibre offre-demande est atteint quand les volumes et les conditions de déplacement se correspondent mutuellement, dans les quatre problèmes ci-dessus. Pour obtenir l'équilibre offre-demande, on résout conjointement les quatre problèmes par un algorithme dédié. Les principaux inputs sont :

- les paramètres de la demande, en particulier les profils temporels de volume de demande.
- les paramètres de l'offre, comprenant les profils temporels de capacité en sortie des arcs, les profils temporels de temps de traversée minimal.

Les outputs sont les profils temporels du volume entrant sur les arcs, qui vérifient une condition de convergence.

3.8.1 Formulation mathématique

- $t_R = F_S\left(t_A\right)$ pour la formation des chemins.
- $X_R^+ = F_U\left(t_R\right)$ pour les choix des usagers.
- $X_A^+ = F_V\left(X_R^+\right)$ pour le chargement en volume.
- $t_A = F_F\left(X_A^+\right)$ pour l'écoulement du flux.

En remplaçant les variables X_R^+ et t_A par leurs expressions en fonction de X_A^+, on obtient une formulation réduite équivalente qui est un problème de point fixe :

$$X_A^+ \in F_V \circ F_U \circ F_S \circ F_F\left(X_A^+\right) \tag{3.22}$$

3.8.2 Un algorithme de combinaisons convexes au niveau des arcs

Voir algorithme 3.5, page 92.

début

$k \leftarrow 0$; $X_A^{+[0]} \leftarrow 0$

répéter

$Y_A^{+[k]} \leftarrow F_V \circ F_U \circ F_S \circ F_F \left(X_A^{+[k]} \right)$

$X_A^{+[k+1]} \leftarrow (1 - w_k) \cdot X_A^{+[k]} + w_k \cdot Y_A^{+[k]}$ avec w_k une suite décroissante de $[0; 1]$.

$k \leftarrow k + 1$

jusqu'à $X_A^{+[k]}$ *vérifie un critère de convergence* ;

fin

Algorithme 3.5 : Algorithme des combinaisons convexes au niveau des arcs

Chapitre 4

Traitement d'un cas d'école

Introduction

Le traitement d'un cas d'école présente par l'exemple le fonctionnement du modèle LADTA. Un exemple très simple permet d'effectuer les calculs manuellement et de confronter ces résultats à ceux fournis par le prototype.

Ce chapitre se compose de trois sections :

- la section 4.1 explique le fonctionnement du modèle sur un réseau simple, composé de deux arcs concurrents. Elle analyse en particulier l'état d'équilibre de ce réseau ;
- la section 4.2 expose les détails des calculs réalisés par le prototype sur ce même exemple ;
- la section 4.3 présente une série de tests ;

4.1 Traitement manuel pour l'explication du fonctionnement du modèle

Dans ce chapitre, nous traitons d'un d'équilibre offre-demande pour un exemple très simple de deux arcs concurrents avec une demande variable selon les horaires d'une journée. Les étapes de ce test permettent premièrement d'expliquer comment le modèle théorique simule les déplacements sur un réseau routier, puis d'avoir un instrument de renseignement pour le travail

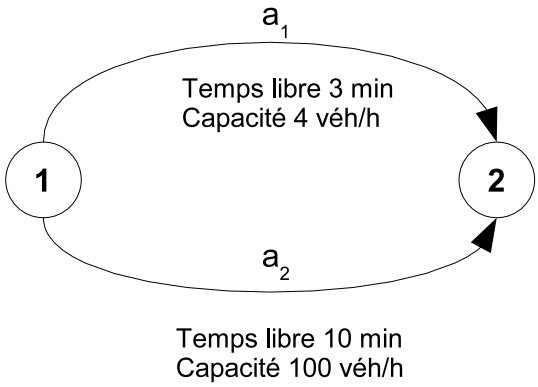


FIG. 4.1 – Réseau de deux arcs concurrents.

de programmation du prototype, et enfin d'examiner les résultats obtenus par le prototype.

4.1.1 Traitement de l'exemple

4.1.1.1 Les hypothèses

Tout d'abord, on suppose que les usagers connaissent l'état du réseau. Il s'agit de traiter la concurrence entre les deux arcs représenté figure 4.1.

Pour l'arc a_1, la capacité en sortie est de 4 véh/h et le temps libre est de 3 min.

Pour l'arc a_2, la capacité est de 100 véh/h, et le temps libre de 10 min.

La demande est représentée par un profil temporel de trois morceaux du volume cumulé de 0h à 23h, dont le total est de 67 véhicules (cf. figure 4.2).

Ces hypothèses sont proposées afin d'une part de mettre en évidence la concurrence entre deux les deux arcs a_1 et a_2 et d'autre part de faciliter les calculs. La capacité de l'arc a_1 est limitée à 4 véh/h, avec un temps libre de seulement 3 min. Pour l'arc a_2 le temps libre est toujours de 10 min mais la capacité est 100 véh/h. La capacité de l'arc a_2 est donc très supérieure à la demande maximale, qui est de 5 véh/h.

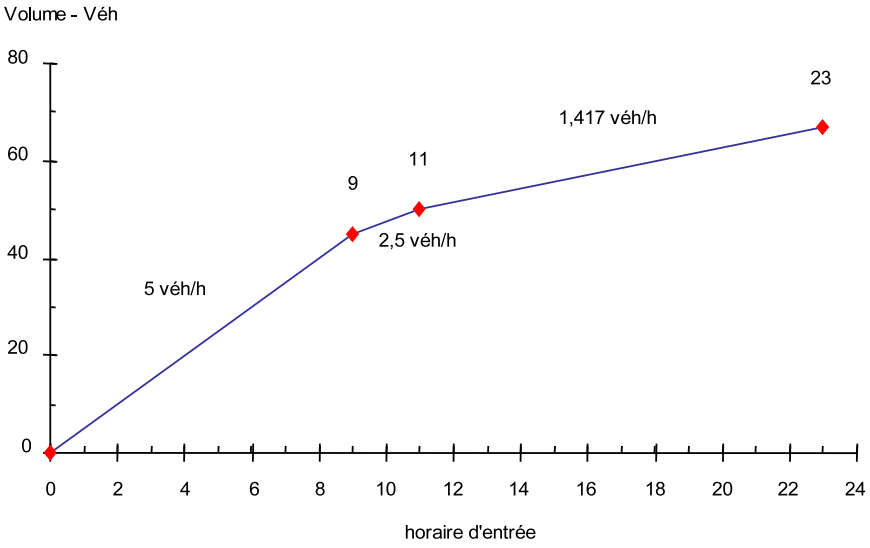


FIG. 4.2 – Profil temporel du volume de la demande.

4.1.1.2 Principe des calculs

1. D'abord, la demande est concentrée sur l'arc a_1.
2. Puis, elle reporte vers l'arc a_2, pour certaines tranches de temps. On nomme les bornes de ces intervalles des *instants critiques*.

A partir des hypothèses et des principes présentés précédemment, nous procédons à une analyse de l'évolution du trafic dans le réseau de l'exemple via les étapes suivantes :

- A l'instant $h_0 = 0$, le temps de parcours libre de l'arc a_1 est inférieur à celui de l'arc a_2. Les usagers choisissent donc l'arc a_1. Le volume de la demande reste concentré sur l'arc a_1 jusqu'à h_1 où le temps de parcours effectif de l'arc a_1 atteint 10 min, soit le temps de parcours libre de l'arc a_2. C'est la congestion sur a_1 qui explique l'augmentation du temps de parcours. En effet, la demande (5 véh/h entre h_0 à h_1) est supérieure à la capacité de l'arc a_1 (4 véh/h).
- De h_1 à h_2 la demande se partage entre les deux arcs jusqu'à l'instant

h_2, pour lequel la demande est inférieure à la capacité de l'arc a_1. Les temps effectif des deux arcs sont sur cette période de 10 min, c'est à dire que le volume chargé sur l'arc a_1 est toujours égal à la capacité de l'arc a_1, et que le reste de la demande est chargé sur l'arc a_2.

– Enfin, à partir de h_2 la demande est concentrée sur l'arc a_1. L'arc a_1 reste cependant congestionné jusqu'à l'instant h_3 où le volume cumulé de la demande est égal à la capacité cumulée de l'arc a_1. Le temps de parcours sur l'arc a_2 est de 10 min pour tous les instants car sa capacité est supérieure à la demande.

Grâce à cette analyse, nous avons calculé les instants critiques h_1, h_2, h_3. Les étapes de calcul et les résultats sont présentés successivement ci-après.

4.1.1.3 Chargement de l'arc a_1 isolé

Calculons séparément l'écoulement sur a_1, en supposant la demande concentrée sur a_1. Celà permet de déterminer h_1, car dans la période $[h_0; h_1]$ la demande est concentrée sur l'arc a_1.

L'évolution de trafic sur l'arc a_1 est représentée graphiquement dans les figures 4.3 et 4.4.

De 0h à 9h, le volume chargé sur l'arc a_1 est de 5 véh/h, mais la capacité est de 4 véh/h. Il y a donc un 1 véh/h en attente sur a_1. Le stock à l'instant $h = 9$h est de 9 véhicules. Cela entraîne une augmentation du temps de traversée, qui passe de 3 min pour $h = 0$h à 138 min pour $h = 9$h. Ainsi, le véhicule entrant sur a_1 à 9h va sortir de a_1 à 11h18min.

Le calcul du temps de traversée effectif à l'instant $h = 9$ h est le suivant :

– Volume entrant cumulé à 9h :

$$5\text{véh/h} \times 9\text{h} = 45\text{véh}$$

– Heure de sortie du 45ème véhicule :

$$45\text{véh} \div 4\text{véh/h} + 3\text{min} = 11\text{h}18\text{min}$$

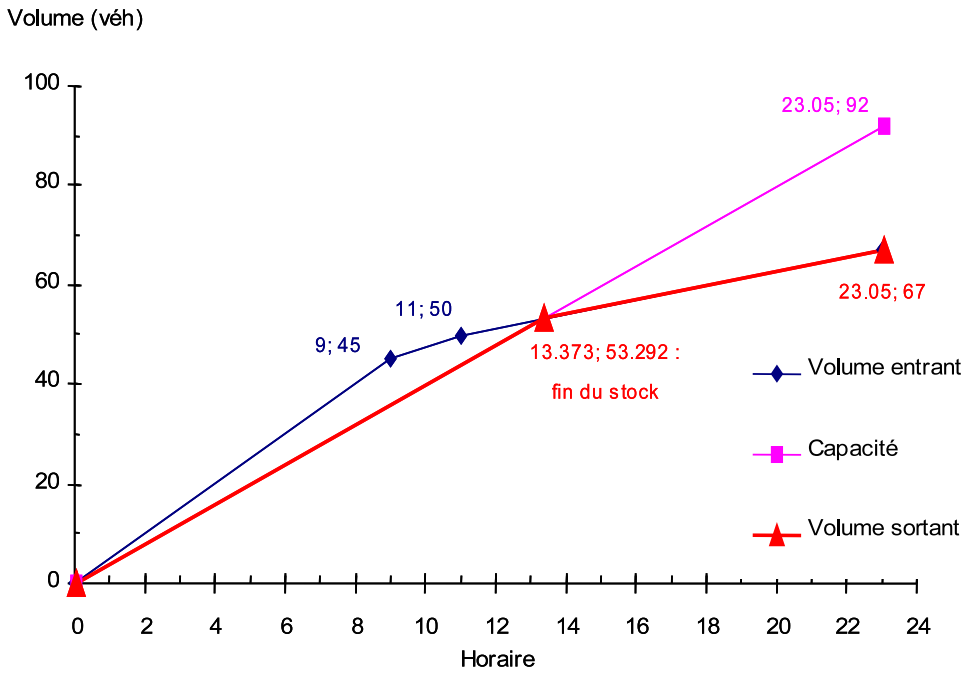


FIG. 4.3 – Ecoulement sur l'arc 1 isolé.

– Temps de traversée pour le véhicule entrant à 9h :

$$11\text{h}18\text{min} - 9\text{h} = 2\text{h}18\text{min}$$

– De 9h à 11h, le volume chargé sur l'arc a_1 est de $2,5$ véh/h. Sa capacité est de 4 véh/h. Le stock diminue donc à la vitesse de $1,5$ véh/h (déstockage). Celà entraîne une diminution de temps de traversée, qui passe de 2h18min pour $h = 9\text{h}$ à 1h33min pour $h = 11\text{h}$. Le stock restant sur a_1 est de 6 véh à l'instant $h = 11\text{h}$. Le calcul du temps de traversée effectif est alors :
  – Volume entrant cumulé à 11h :

$$45\text{véh} + 2,5\text{véh/h} \times 2\text{h} = 50\text{véh}$$

  – Heure de sortie du 50ème véhicule :

$$50\text{véh} \div 4\text{véh/h} + 3\text{min} = 12\text{h}33\text{min}$$

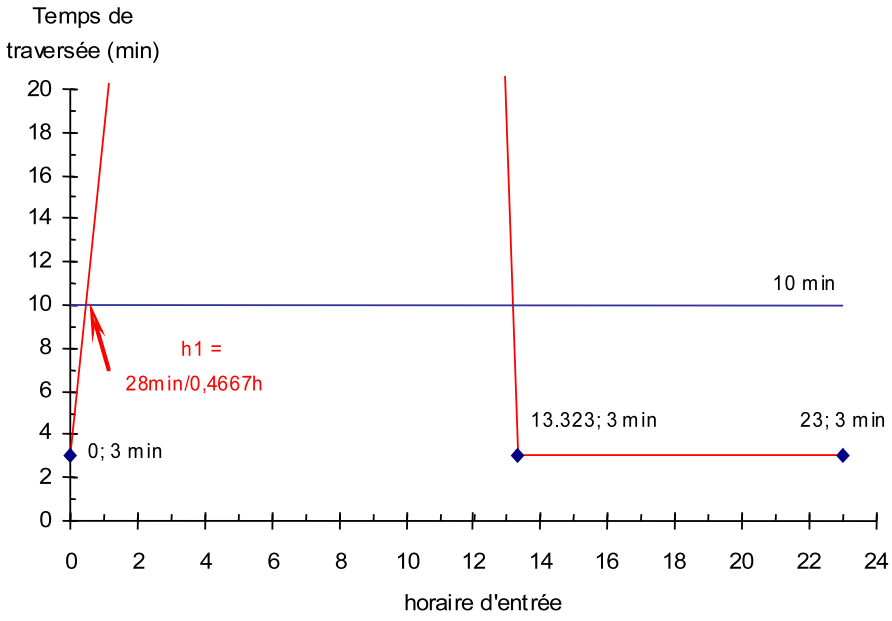


FIG. 4.4 – Concurrence entre deux arcs et instant critique h_1.

– Temps de traversée pour le véhicule entrant à 11h :

$$12\text{h}33\text{min} - 11\text{h} = 1\text{h}33\text{min}$$

– De 11h à 23h, le volume chargé sur l'arc a_1 est de $1,417$véh/h. Le stock diminue donc à la vitesse de $2,583$véh/h jusqu'à l'épuisement à l'instant $h = 13\text{h}19\text{min}22\text{s}$. Cet instant est calculé suivant :

$$6\text{véh} \div 2,583\text{véh/h} + 11\text{h} = 13\text{h}19\text{min}22\text{s}$$

Le temps de traversée redevient alors égal au temps libre, c'est à dire 3min. A partir de cet instant, le trafic est fluide et la congestion est finie.

A partir du profil temporel de temps de traversée déterminé par le chargement de la demande sur l'arc a_1 isolé, on peut calculer l'instant $h_1 = 0,4667\text{h}$ pour lequel le temps de traversée de a_1 atteint 10min. h_1 correspond à l'instant à partir duquel l'arc a_2 devient plus performant que l'arc a_1. Le calcul de l'instant h_1 est le suivant :

– Volume entrant à l'instant h_1 :

$$h_1 \times 5\text{véh/h}$$

– Ce volume entrant est totalement sorti à l'instant :

$$h_1 + 10\text{min}$$

– La capacité en sortie :

$$4\text{véh/h}$$

– Le volume sortant à l'instant $h_1 + 10\text{min}$:

$$((h_1 + 10\text{min}) - 3\text{min}) \times 4\text{véh/h}$$

– On a donc :

$$h_1 \times 5\text{véh/h} = ((h_1 + 10\text{min}) - 3\text{min}) \times 4\text{véh/h} = (7 \div 60) \times 4$$

– Soit :

$$h_1 = 0,4667\text{h}$$

4.1.1.4 Etat d'équilibre

L'évolution du trafic sur l'arc a_1 à l'état d'équilibre est représentée par la figure 4.5. Pour l'arc a_2, nous ne présentons que le chargement en volume dans la figure 4.6, car il n'y a pas de congestion dans cet arc. Cette évolution d'équilibre se déroule en trois temps :

– Nous avons déterminé l'instant critique $h_1 = 0,4667\text{h}$ à partir duquel l'arc a_2 est utilisé. La période de stockage sur a_1 est de h_0 à h_1. A

l'instant h_1, le stock sur a_1 est déterminé par :

$$(5\text{véh/h} - 4\text{véh/h}) \times (h_1 - h_0) = 0,46675\text{véh}$$

La congestion sur a_1 entraîne une augmentation du temps de traversée, qui devient égal à 10min.

- A partir de l'instant h_1, le volume de la demande se partage entre les deux arcs, et ce jusqu'à l'instant $h_2 = 9\text{h}$ pour lequel le volume de la demande devient inférieur à la capacité de l'arc a_1. Ce partage de la demande assure l'égalité de temps de traversée entre les deux arcs. Pendant cette période, le stock sur l'arc a_1 est donc toujours de $0,4667$véh. Le temps de traversée de deux arcs est toujours égal à 10 minutes.

  Les deux instants critiques h_1 et h_2 déterminent un état d'équilibre offre-demande. On peut distribuer la demande entre deux arcs concurrents (cf. figure 4.6), et définir également leur temps effectif (cf. figure 4.7) :
  - Pour l'arc a_1, le volume chargé est :
    - de 0h à $h_1 = 0,467\text{h}$: 5véh/h (toute la demande).
    - de h_1 à $h_2 = 9\text{h}$: 4véh/h (la capacité de l'arc a_1).
    - de h_2 jusqu'à la fin de période (23h) : toute la demande.
  - Pour l'arc a_2, le volume chargé de h_1 à h_2 est de 1véh/h.
- A partir de l'instant h_2, le stock sur l'arc a_1 diminue (déstockage). Le temps de traversée de l'arc a_1 décroît pour devenir égal au temps libre de l'arc a_1, soit 3min. La demande se concentre alors à nouveau uniquement sur l'arc a_1. Le trafic sur l'arc a_2 est nul.
- Enfin, déterminons l'instant h_3 correspondant à l'épuisement du stock sur a_1 :
  - A l'instant h_2, la capacité est de 4véh/h, le volume chargé de $2,5$véh/h, donc la capacité utilisée pour le déstockage est de $1,5$ véh/h.
  - le stock est de $0,4667$véh.

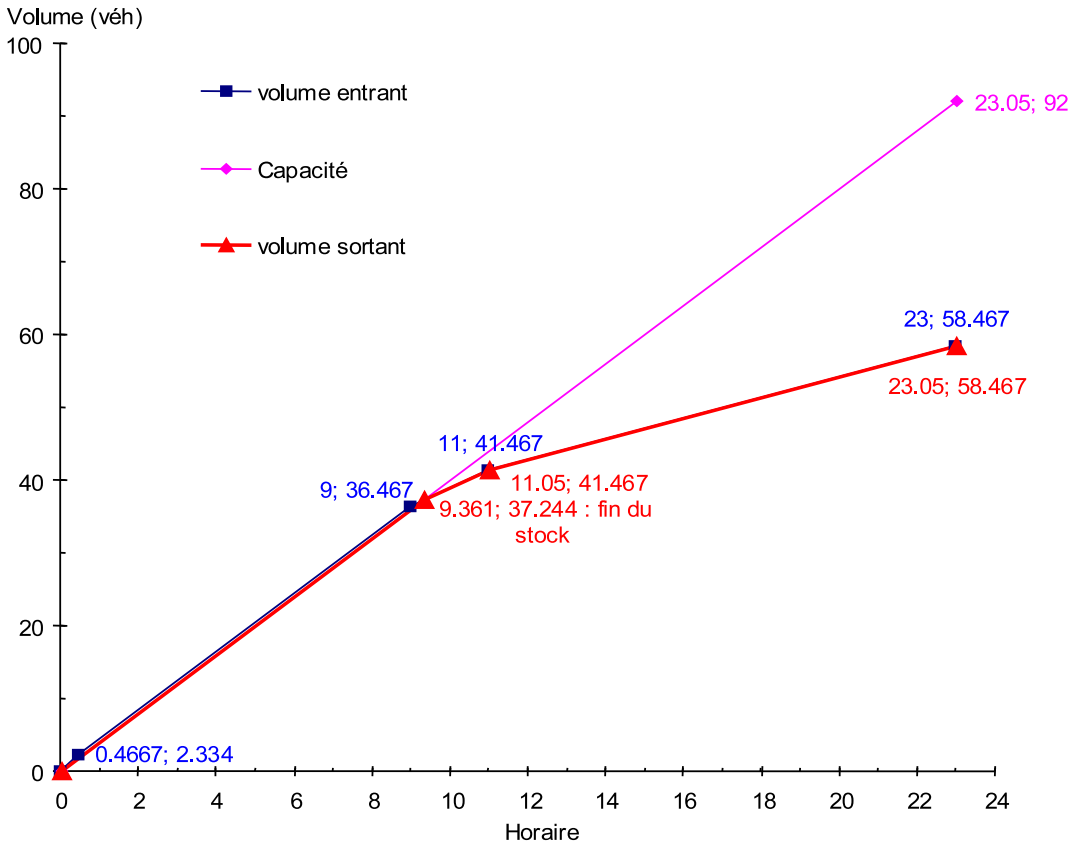


FIG. 4.5 – Ecoulement sur l'arc a_1 à l'équilibre.

– Le temps nécessaire pour déstockage est donc de :

$$0,4667 \div 1,5 = 0,311\text{h}$$

– h_3 vaut donc :

$$h_3 = 9,311\text{h}$$

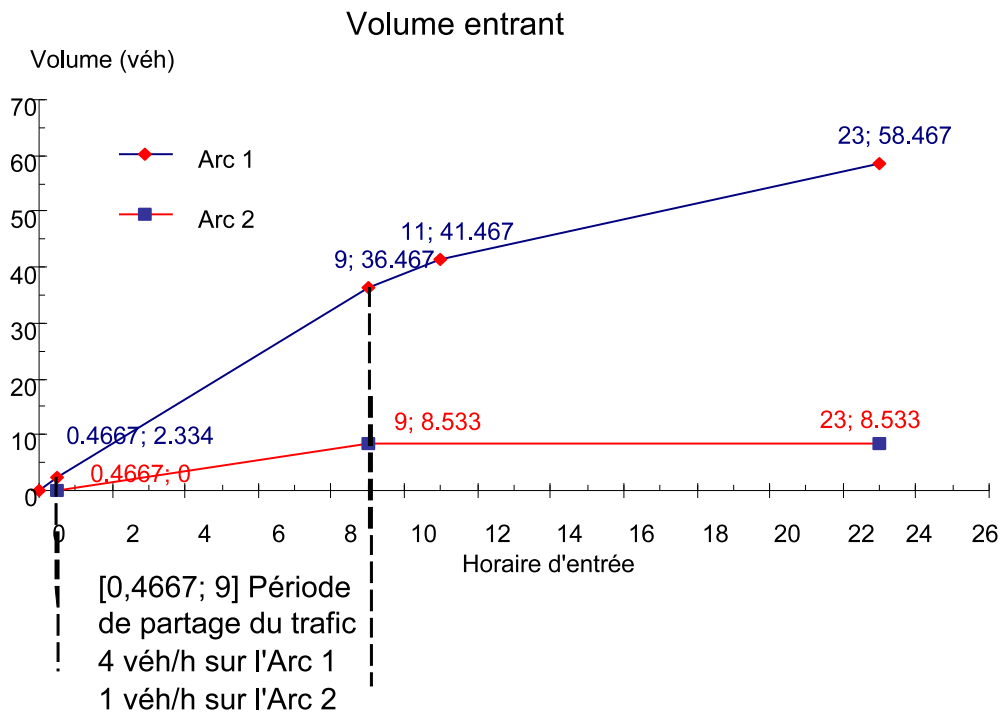


FIG. 4.6 – Partage du trafic à l'équilibre.

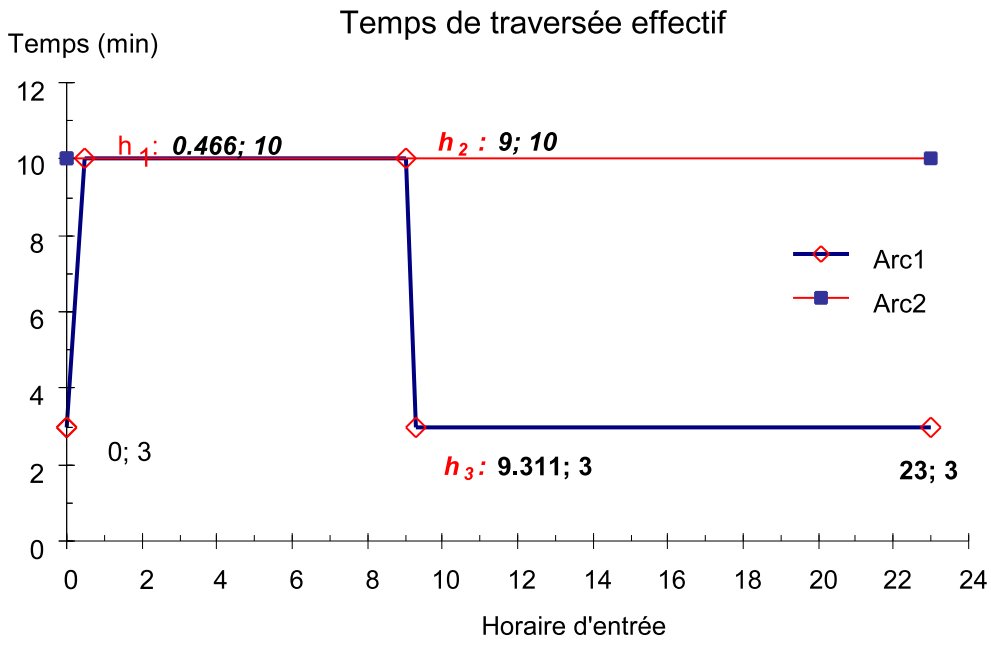


FIG. 4.7 – Temps de parcours à l'équilibre.

4.1.2 Commentaires sur la concurrence entre deux arcs

Dans cette partie, on compare deux cas : l'arc a_1 isolé et les deux arcs a_1 et a_2 en concurrence. Cette comparaison est réalisée en termes de période de congestion, de files d'attentes et de temps perdu global.

4.1.2.1 Période de saturation

Arc a_1 isolé : Stockage sur l'intervalle $[0,05\text{h}; 9,05\text{h}]$

Déstockage sur l'intervalle $[9,05\text{h}; 13,373\text{h}]$

Concurrence des deux arcs : il n'y a pas de saturation sur l'arc a_2. Sur l'arc a_1, la période de saturation sur l'intervalle $[0,05\text{h}; 9,361\text{h}]$ est divisée en trois périodes :

- stockage sur l'intervalle $[0,05\text{h}; 0,517\text{h}]$;
- stock constant sur l'intervalle $[0,517\text{h}; 9,05\text{h}]$;
- déstockage sur l'intervalle $[9,05\text{h}; 9,361\text{h}]$.

4.1.2.2 File d'attente

Arc a_1 isolé : la file d'attente débute à $0,05$h. Son évolution est la suivante :

- de $0,05$ h à $9,05$h, elle passe de 0 à 9 véhicules ;
- de $9,05$h à $11,05$h, elle diminue de 9 véhicules jusqu'à 6 véhicules ;
- de $11,05$h à $13,373$, elle diminue de 6 véhicules jusqu'à 0.

Concurrence des deux arcs : il n'y pas de file d'attente sur l'arc a_2. La file d'attente sur l'arc a_1 évolue comme suit :

- de $0,05$h à $0,517$h, elle passe de 0 à $0,467$ véhicules ;
- de $0,517$h à $9,05$h, elle reste constante, égale à $0,467$ véhicule ;
- de $9,05$h à $9,361$h, elle diminue de $0,467$ véhicules jusqu'à 0.

4.1.2.3 Temps global perdu

Arc a_1 isolé : A partir des volumes de la file d'attente et des périodes correspondantes de saturation, on peut calculer le temps perdu global, soit :

$$\frac{9 \times (9,05 - 0,05) + (9+6) \times (11,05 - 9,05) + 6 \times (13,373 - 11,05)}{2} = 62,469\text{h}$$

Concurrence des deux arcs : A partir des volumes de la file d'attente et des périodes correspondantes de saturation, on peut calculer le temps perdu global, soit :

$$\frac{0,467 \times (0,517 - 0,05) + 0,467 * (9,361 - 9,05)}{2} + 0,467 \times (9,05 - 0,5167) = 4,164\text{h}$$

4.2 Traitement par le prototype

Pour la simulation par le prototype, le réseau de deux arcs concurrents doit être recodé en trois arcs, dont un est fictif. Cette modification de topologie du réseau assure la distinction de deux arcs qui ont le même nœud initial et nœud final, car dans le prototype chaque arc du réseau est identifié par le nœud initial et final.

4.2.1 Codage du réseau

La structure du réseau nommé $TNET$ est représenté figure 4.8. Ce réseau est composé de 3 arcs :

- a_{12}, Temps de parcours libre : 3min, Capacité : 4véh/h
- a_{13}, Temps de parcours libre : 10min, Capacité : 100véh/h
- a_{32}, Temps de parcours libre : 0min, Capacité : 100véh/h

L'arc a_{32} est un arc fictif.

Un seul couple O-D est posé sur le réseau avec la demande représentée par un profil temporel de volume de déplacement (cf. figure 4.2).

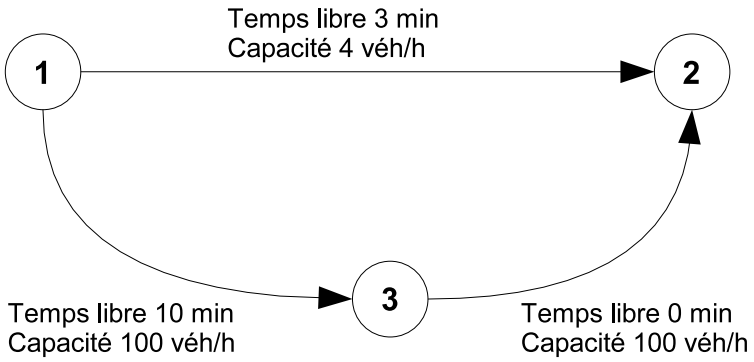


FIG. 4.8 – Le réseau $TNET$.

4.2.2 Vérification par étapes

La vérification étape par étape a été réalisé pour les six premières itérations. Les résultats des deux premières itérations sont présentés dans les figures 4.9, 4.10, 4.11 et 4.12.

4.2.2.1 Itération 1

$k = 0$, $w_0 = 1$, $X_A^{(0)} = 0$.

Calcul de l'état auxilaire $Y_A^{(0)}$:

Ecoulement : Pas encore de chargement en volume sur les 2 arcs

Formation des services : temps effectif = temps libre

Choix des usagers et chargement en volume : la demande concentre totalement sur l'arc a_1.

Combinaison : la combinaison du volume est égale à l'état auxiliaire.

$$X_A^{(0)} = Y_A^{(0)}$$

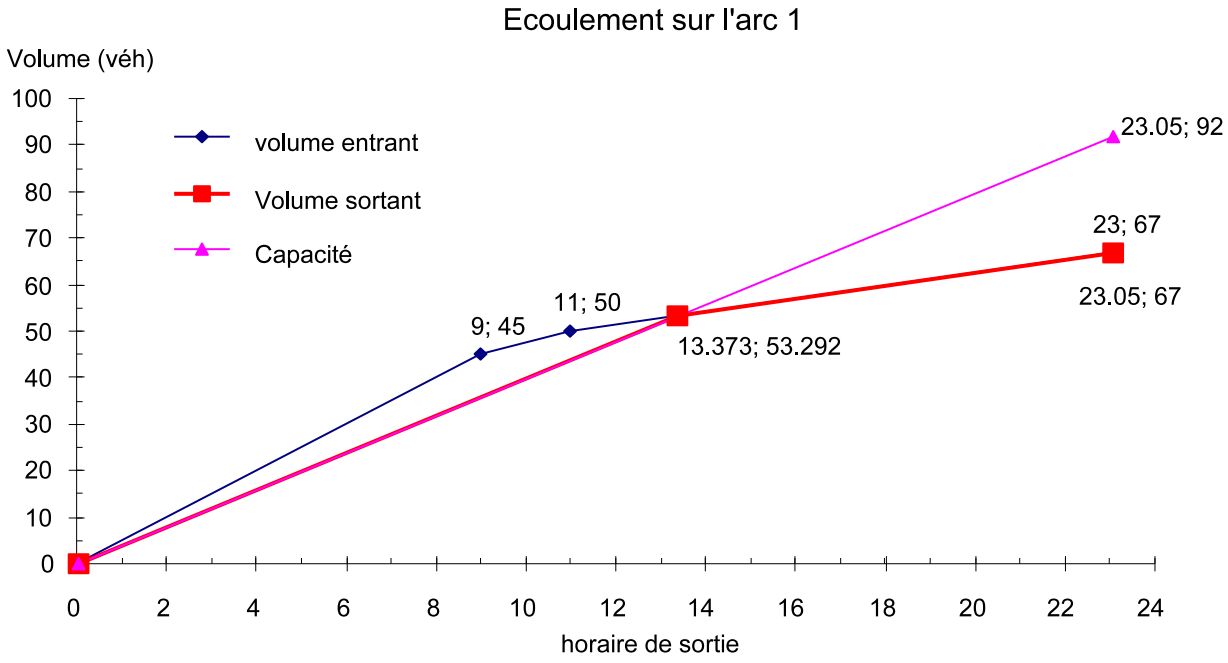


FIG. 4.9 – Ecoulement sur l'arc a_1 à la 2ème itération.

4.2.2.2 Itération 2

$k = 1$, $w_1 = 1/2$, $X_A^{(0)} = Y_A^{(0)}$.

Calcul de l'état auxilaire $Y_A^{(1)}$:

Ecoulement : La figure 4.9 représente l'écoulement sur l'arc a_1.

Formation des services : la figure 4.10 présente la concurrence entre les chemins $1-2$ et $1-3-2$ du réseau $TNET$, qui permet de déterminer les intervalles efficaces de chaque chemin.

Choix des usagers et chargement en volume : la figure 4.11 présente le partage du volume entre deux chemins concurrents, sachant les intervalles efficaces déterminés précédemment. Les volumes chargés sur les arcs

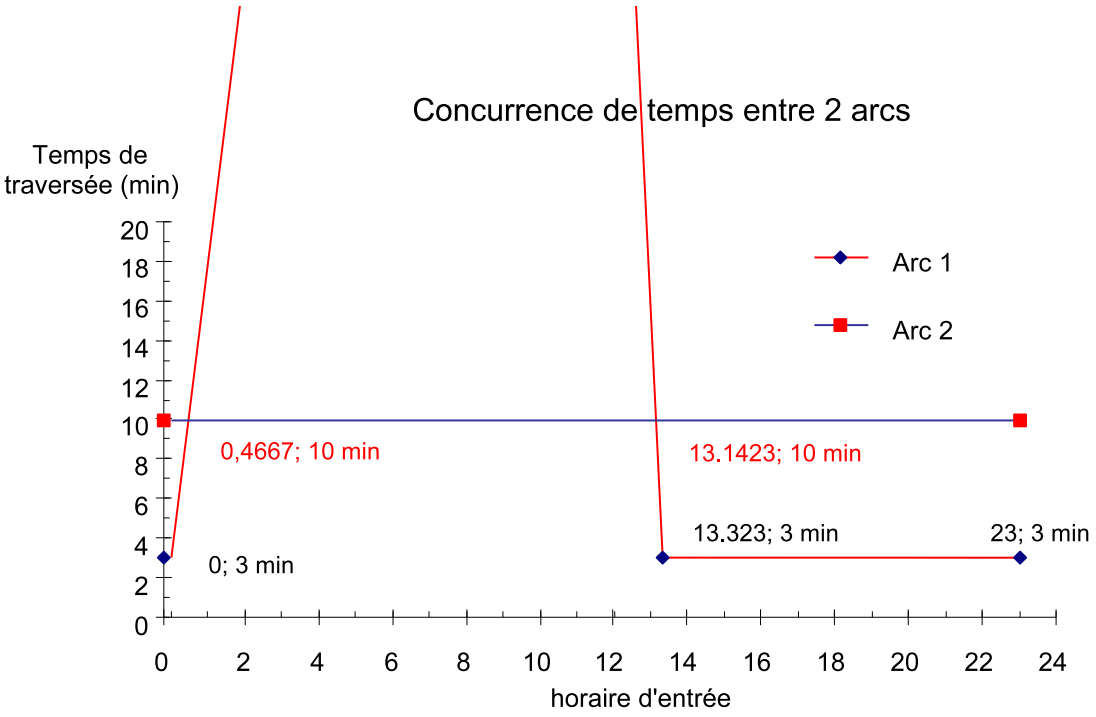


FIG. 4.10 – Concurrence entre arcs à la 2ème itération.

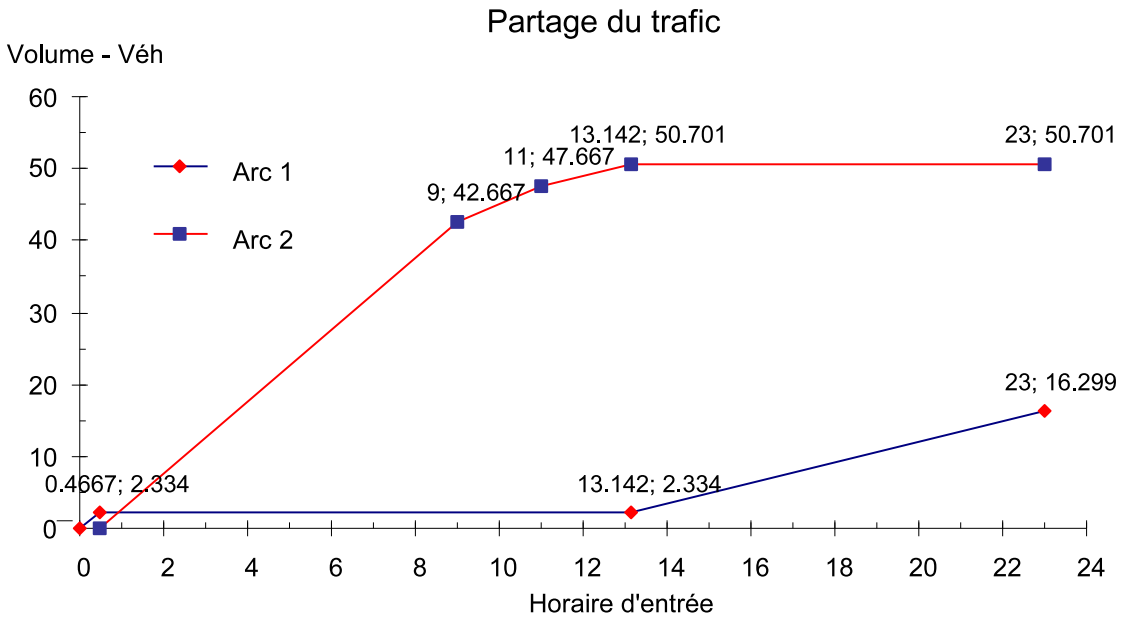


FIG. 4.11 – Partage du volume à la 2ème itération.

sont également calculés.

Combinaison : Combinaison de l'état courant et auxiliaire ($w_k = 0.5$) : voir la figure 4.12 pour a_1 la figure 4.13 pour a_2.

Sur l'arc a_1,

$$X_{a_1}^{(1)} = \frac{1}{2} X_{a_1}^{(0)} + \frac{1}{2} Y_{a_1}^{(0)}$$

Sur l'arc a_1,

$$X_{a_2}^{(1)} = \frac{1}{2} X_{a_2}^{(0)} + \frac{1}{2} Y_{a_2}^{(0)}$$

4.2.3 Résultats à l'état de convergence

La simulation du trafic au réseau $TNET$ a été réalisée par trois suites de coefficients de combinaison pour 250 itérations. La 51ème itération a donné un niveau de convergence acceptable. De plus, l'affectation par la suite $w_k =$

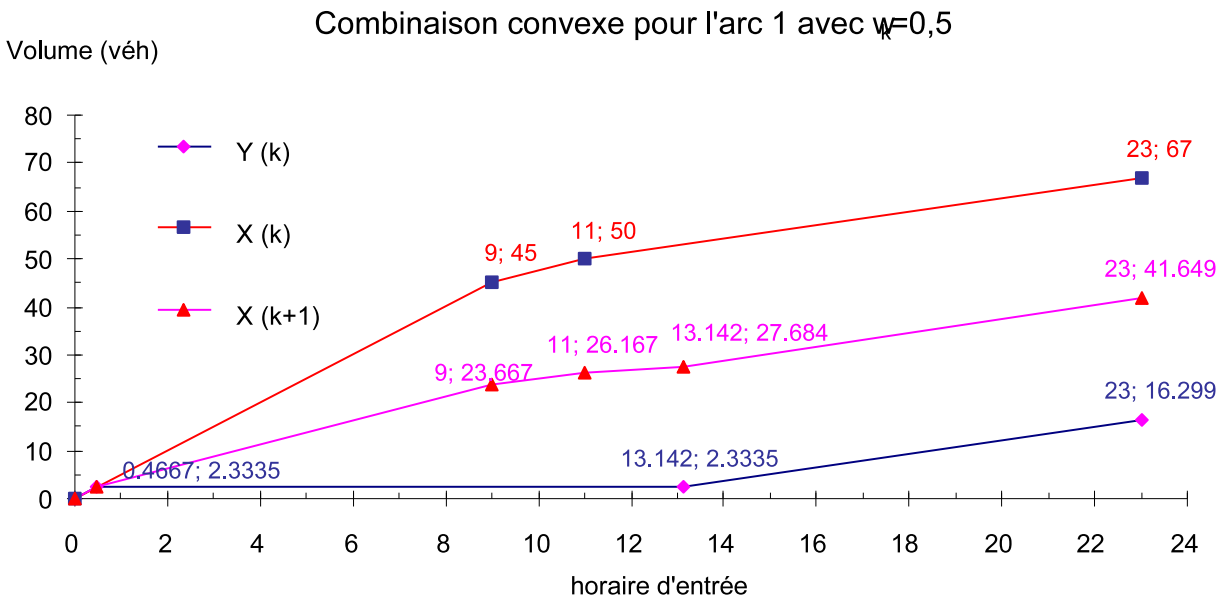


FIG. 4.12 – Résultat de la combinaison convexe sur l'arc a_1 à la deuxième itération.

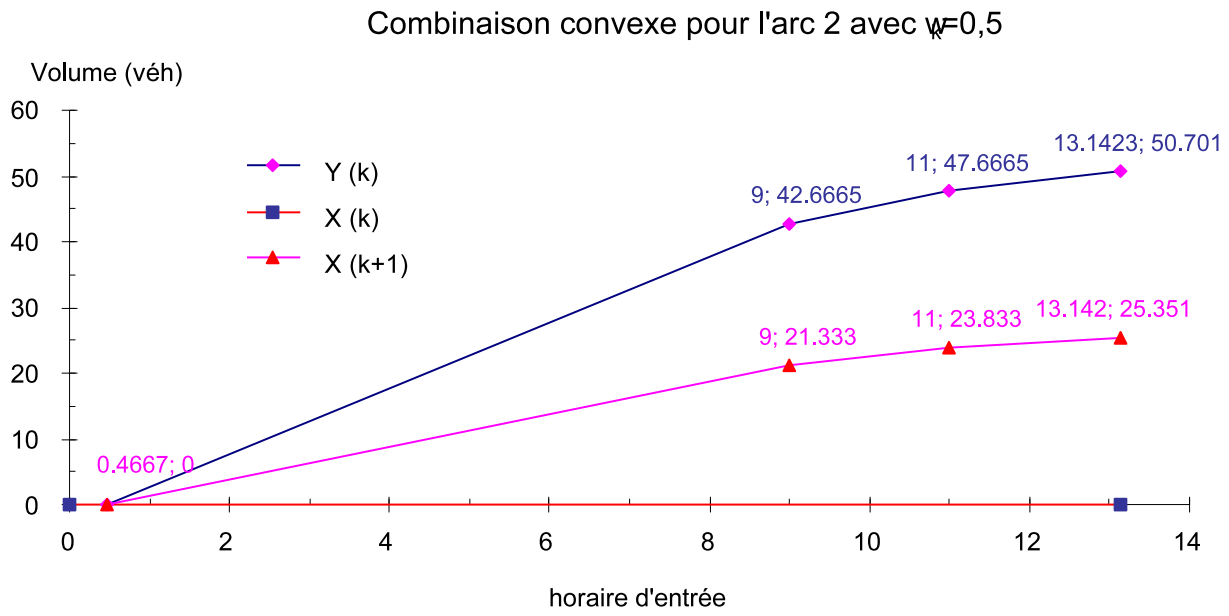


FIG. 4.13 – Résultat de la combinaison convexe sur l'arc a_2 à la deuxième itération.

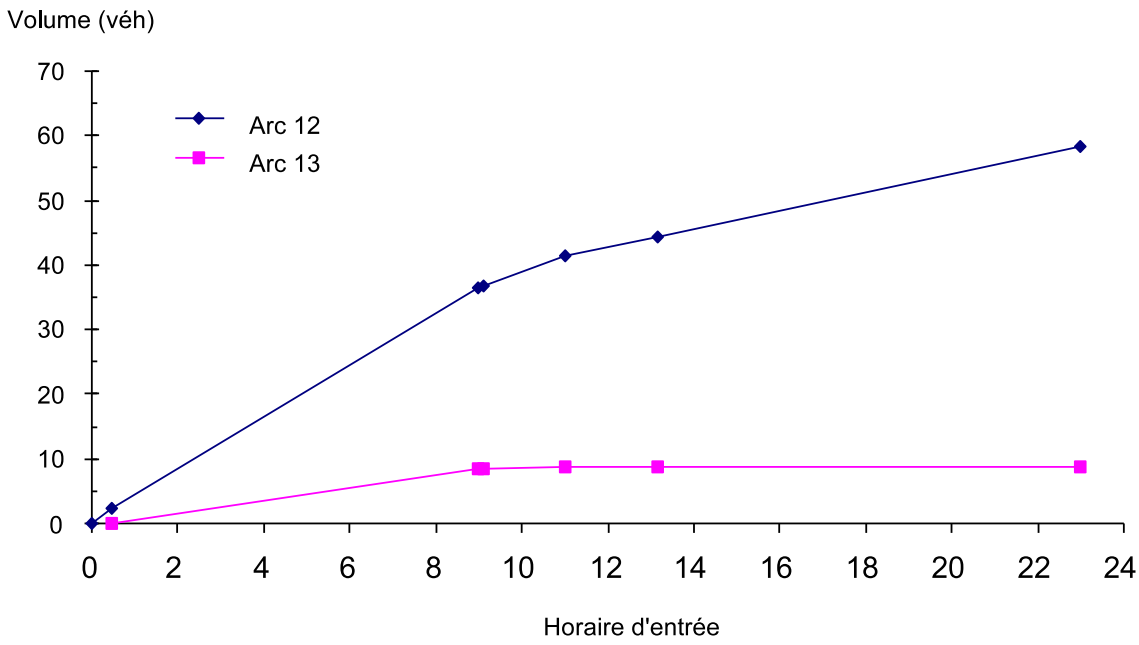


FIG. 4.14 – Volumes entrants des arcs à la 51ème itération.

$\frac{1}{1+k}$ a fourni une meilleure convergence. Les résultat obtenus sont :

- Les profils de volume entrant sur les arcs à l'état d'équilibre offre-demande. Ils sont présentés dans la figure 4.14.
- Les profils de temps de traversée de deux chemins concurrents $1-2$ et $1-3-2$ sont présentés dans la figure 4.15.

4.3 Série de tests

Une série de tests est proposé afin de :

- s'assurer du fonctionnement correct du prototype pour des cas plus complexes ;
- débusquer les phénomènes spécifiques qui peuvent apparaître pendant le processus de calcul.

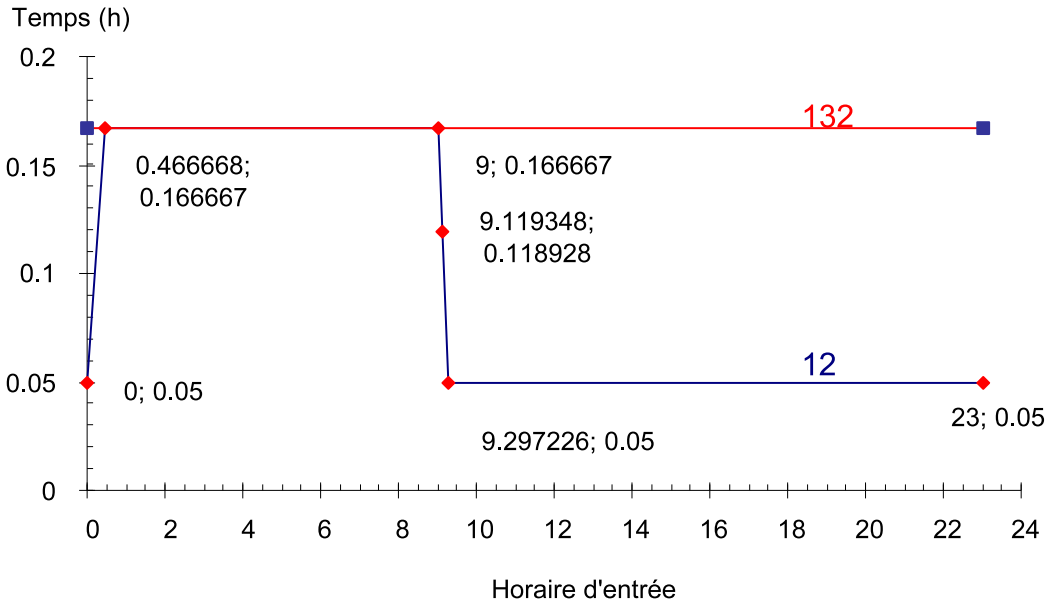


FIG. 4.15 – Temps de parcours à la 51ème itération.

– mener une première analyse de la convergence d'un algorithme d'équilibrage du modèle LADTA. Cette question sera discutée dans un chapitre spécifique au plus tard (cf. chapitre 9).

La série de test consiste en quatres réseaux d'exemple : $TNET$ (cf. figure 4.8), NET (cf. figure C.1), $NET8$ (cf. figure C.2), et réseau baïonnette de la ville Marne-la-Vallée MLV (cf. figure C.3). Les tests sont synthétisés dans le tableau 4.1.

Tous les tests cités dans ce tableau sont calculés par l'algorithme des combinaisons convexe au niveau des arcs et au niveau des chemins. Nous avons donc réalisé au toal 44 tests.

Réseau	Tps base	N. itér.	$w_k = \frac{1}{1+k}$	$w_k = \frac{1}{5+3k}$	$w_k = \frac{1}{5+k/3}$
$TNET$	0	250	×	×	×
NET	1	100	×		
	2	100	×		
	3	100	×	×	×
$NET8$	0	50	×	×	×
	3	100	×		
	5	100	×	×	×
MLV	0	20	×		
	12	100	×	×	×
	18	100	×	×	×

TAB. 4.1 – Série de tests.

Chapitre 5

Elaboration du prototype

Introduction

Le modèle théorique de Leurent doit être mis en opération afin de démontrer que le modèle est calculable, et que de plus l'algorithme du modèle est convergent. Pour ces besoins, le modèle théorique a été codé en SciLab, un langage informatique interprété, pour une version restreinte.

Ce travail de programmation a abouti à un prototype dont les principes d'élaboration et de programmation sont présentés dans la section 5.1. La programmation des fonctions d'écoulement, de formation des services, de choix des usagers, puis de de l'équilibre offre-demande, font l'objet des sections 5.2 à 5.5.

La section 5.6 discute du problème de fragmentation des profils au cours du calcul, mis en évidence par l'expérience lors des premières utilisation du prototype. Il s'agit d'une difficulté réelle, intrinsinquement liée aux principes de calcul choisis, et essentiellement liée au caractère dynamique du modèle. La discussion est conclue par la présentation des méthodes proposées pour tenter de remédier à ce problème.

5.1 Elaboration du prototype

Le prototype a été écrit en langage SciLab, en sept étapes :

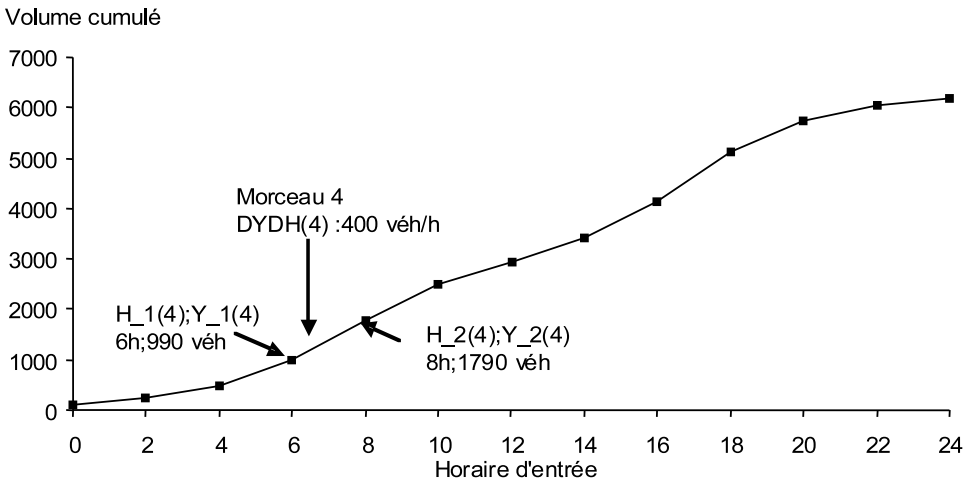


FIG. 5.1 – Exemple de profil temporel affine par morceaux.

- première étape : définition du type de données *profil temporel* modélisant des fonctions du temps, linéaires par morceaux. Puis définition des fonctions de traitement de profils temporels, ainsi que des procédures d'entrée des données et de sortie des résultats.
- 2ème, 3ème, 4ème, 5ème étapes : les fonctions et les sous fonctions, qui traitent successivement le problème de l'écoulement du flux, de la formation des services, du choix des usagers, du chargement en volume, sont mises en organigramme et codées.
- 6ème étape : Les critères de convergence sont codés afin de tester le niveau de convergence à chaque itération de l'algorithme d'équilibrage.
- 7ème étape : La fonction principale, traitant le cercle de dépendances des quatre relations causales qui caractérise le problème d'équilibre offre-demande, est établie.

Dans le cadre d'un prototype du modèle théorique, nous ne programmons qu'une boîte à outils d'affectation dynamique pour une seule classe d'usagers, avec l'horaire de départ fixé, sans péage et sans élasticité de la demande.

5.1.1 Type de données

Les variables principales sont des profils temporels. Par simplicité, on a choisi de considérer uniquement des profils affines par morceaux (cf. figure 5.1).

Afin de coder un profil temporel, le type de données `Prof` a été défini. Il contient des champs pour stocker le nom d'un profil, le nombre de morceaux, les paramètres des morceaux (les sommets initial et final, la pente), etc. Ensuite, une fonction de création d'un profil temporel `MakeProf` est programmée afin de créer une donnée de type `Prof`.

```
Prof ← tlist(
        [NameP; "h0"; "hmax"; "hmin"; "Nb"; "NamesM";
        "H_1"; "H_2"; "Y_1"; "Y_2"; "DYDH"],
        h0, hmax, hmin, Nb, 1:Nb, H_1, H_2, Y_1, Y_2, DYDH)
```

5.1.2 Fonctions utilitaires

Afin de réaliser les calculs sur les profils temporels, ainsi que les traitements des inputs et des outputs du prototype, nous avons d'abord programmé un ensemble de fonctions utilitaires, dont la liste est présentée rapidement ci-dessous :

5.1.2.1 Fonction de création d'un profil

`Prof` ← **MakeProf**(`typef`, `NameP`, `h0`, `hmax`, `Nb`, `H_1`, `Y_1`, `Y_2`, `ymax`)

Cette fonction retourne un profil temporel à partir de trois types des données en input : par un fichier existant qui contient les données d'un profil, par la fenêtre d'interface du SciLab, les données calculées en cours de fonctionnement du prototype.

5.1.2.2 Fonction d'évaluation

`y` ← **EvalProf**(`P`, `h`)

Cette fonction calcule la valeur y du profil temporel P à l'instant h. Si h est un vecteur de dimension n, la valeur retournée y est également un vecteur de dimension n, dont chaque composante y[i] est la valeur de P en x[i].

En cas de discontinuité de P, c'est-à-dire s'il existe i tel que x[i] = x[i+1] et y[i] ≠ y[i+1], on suppose que P est continue à gauche.

5.1.2.3 Fonction de réciproque

h ← **RecProf**(P , y)

Cette fonction calcule la réciproque du profil temporel P : l'instant h correspondant à la valeur y. Cette fonction est utilisée uniquement pour calculer le temps de traversée effectif dans le problème d'écoulement.

Dans le cas où l'une au moins des pente DYDH est nulle sur les profils du volume entrant ou sortant des arcs, c'est-à-dire si P n'est pas inversible, la valeur de la réciproque est choisie d'après les principes d'une analyse physique et économique de l'écoulement du flux sur l'arc.

5.1.2.4 Fonction de différentiation

dydh ← **DiffProf**(P , h)

Cette fonction retourne la valeur de la pente dydh du profil P à l'instant h. Elle sera donc utilisée quand on veut d'une part passer d'un profil de volume cumulé à un profil de débit, d'autre part prendre le débit à un instant h.

5.1.2.5 Fonction d'intégration

Y ← **IntProf**(P , h1 , h2)

Cette fonction calcule l'intégrale du profil P sur l'intervalle [h1;h2]+. Elle est utilisée par exemple lorsque l'on veut passer d'un profil de débit à un profil de volume cumulé, ou afin d'évaluer la différence entre deux profils.

5.1.2.6 Fonction de cumulation

`Prof` ← **ProfCumul**(`P`,`h0`)

Cette fonction crée un profil de volume cumulé à partir d'un profil de volume en intensité `P`, en utilisant la fonction d'intégration.

5.1.2.7 Fonction de prolongation

`Prof` ← **Prolon_Prof**(`P`,`min_h`,`max_h`,`cas_plg`)

Cette fonction permet de prolonger un profil. Il y a trois cas de prolongation (paramètre `cas_plg`) qui dépendent du type de profil (profil des temps, profil d'horaire, ou profil de volume) et de l'objectif de prolongation.

5.1.2.8 Fonction de coupure

`Prof` ← **Coupe_Prof_V**(`P`)

Cette fonction permet de couper les morceaux initiaux et finaux d'un profil de volume dont la pente `DYDH` est égale à 0.

5.1.2.9 Procédure d'enregistrement

`void` ← **SaveProfL**(`P`,`savefile`)

`void` ← **SaveProfL**(`Dir_Data`,`ListP`,`FName`,`Note`,`L_ARC`,`savefile`)

Cette fonction permet d'enregistrer les paramètres d'un profil ou d'une liste de profils sous deux types de fichier :

- fichier $*$.`pro` ou $*$.`lpr` qui contiennent les paramètres codés de profil ;
- fichier $*$.`csv`, qui peut être réutilisé dans un outil bureautique.

5.1.2.10 Procédure d'édition graphique

`void` ← **PlotProf**(
`P`, `Thn`, `colorP`, `hormin`, `vermin`,
`hormax`, `vermax`, `NumW`, `title`, `hor`, `ver`)

Cette fonction permet d'éditer un profil dans une fenêtre graphique de SciLab.

5.2 Programmation de l'écoulement du flux

La fonction d'écoulement du flux `FF` est programmée pour l'ensemble `A` des arcs. Elle est implémentée suivant les principes exposés au 3.7.

5.2.1 Fonction principale

(`X_A_out`, `t_A`, `Sto_A`) ← **FF**(`X_A_in`, `K_A`, `t0_A`, ...)

Cette fonction calcule, pour l'ensemble des arcs :
- les profils de volumes sortants `X_A_out` ;
- les profils de temps de traversée effectif `t_A` ;
- les profils de volumes de stock `Sto_A`.

Ses paramètres principaux sont :
- les profils de volume entrant `X_A_in` ;
- les profils de capacités en sortie `K_A` ;
- les profils de temps de traversée minimum `t0_A`.

Les étapes du traitement sont les suivantes :

1. chargement des paramètres.
2. les profils d'horaires de passage tels que le profil d'horaire de départ selon les horaires sortis, le profil d'horaire d'arrivée en sortie selon les horaires entrés, en sachant le temps de traversée, sont successivement établis par la fonction auxiliaire `Cal_H` (voir 5.2.2).
3. le profil de volume d'arrivée en sortie de l'arc est établi, connaissant les profils de volume entrant, d'horaires d'arrivée et d'horaires de départ.
4. le profil de volume sortant est établi par une analyse de stockage en sortie de l'arc, sous contrainte de la capacité en sortie, d'après l'organigramme de l'algorithme de calcul du volume sortant (cf. figure 3.1).

5. le profil de temps de traversée effectif est établi par le calcul d'horaire de départ du volume entrant et d'horaire de sortie du volume sortant, en recourant à la fonction réciproque, en suivant l'algorithme de calcul du temps de parcours (algorithme 3.3, page 89).

5.2.2 Fonctions auxiliaires

La fonction `Cal_Xout` calcule le profil de volume sortant et du profil de stock. Cette fonction implémente l'algorithme 3.4, page 90, discuté en section 3.7.4.

```
(X_A_out, Sto_A) ← Cal_Xout(
        X_A_arr, K_A, H_max, NameP_Xout, NameP_Stock)
```

La fonction `Cal_H` calcule les profils de l'horaire de passage.

```
(H, Hb) ← Cal_H(t, Names_H)
```

5.3 Programmation de la formation des services

La fonction de formation des services `FS_DIJ` est programmée pour une destination et toutes les origines, grâce à l'algorithme de propagation sur un réseau dynamique (voir section 3.4 et l'algorithme 3.1).

5.3.1 Fonction principale

```
(M_s, W_n_s) ← FS_DIJ(N, s, t_A, ...)
```

Cette fonction forme et choisit les chemins efficaces, en calculant récursivement leur temps de traversée.

Ses paramètres principaux sont :

- `N` : le réseau traité ;
- `s` : la destination traitée ;
- `t_A` : les temps de traversée effectif des arcs ;
- ... : des paramètres auxiliaires.

Elle retourne :

- `M_s` : liste des atomes efficaces pour la destination `s`. Cette liste est utilisée en argument de la fonction de chargement en volume `FUV`.
- `W_N_s` : liste des atomes efficaces pour la destination `s`, triée par nœud. Cette liste est utlisée en argument de la fonction de distribution de la demande `Distribution_R`.

Les étapes de programmation sont les suivantes :

1. chargement des paramètres.
2. les profils d'horaire de passage de chaque arc sont établis par la sous-fonction `Cal_H`.
3. l'ensemble des chemins/atomes efficaces du réseau et l'ensemble de chemins/atomes efficaces pour chaque nœud du réseau sont établis par l'algorithme de propagation, en utilisant les fonctions auxiliaires suivantes :
   - `SFE`, qui juge de l'efficacité d'un chemin par rapport aux autres dans un ensemble de chemins, entre un nœud n et une destination s.
   - `Cal_int`, qui calcule les intervalles efficaces de chaque chemin entre un nœud n et une destination s.
   - `Composition_T`, qui compose le profil de temps d'un chemin r d'après le profil de temps du sous chemin r' et le profil de temps de l'arc initial a.

5.3.2 Fonctions auxiliaires

La fonction `SFE` compare l'efficacité de l'atome `alpha` à celles de l'ensemble des atomes `W_n` qui ont le même nœud initial.

`check_sfe` ← **SFE**(`alpha , W_n`)

La fonction `cal_int` calcule les intervalles efficaces de chaque atome dans l'ensemble des atomes `W_n`.

`Ieff_R` ← **cal_int**(`W_n`)

La fonction auxiliaire `Composition_T` de composition de deux profils $\bar{t}_a$ et $\bar{t}_{r'}$ permet de calculer le temps de traversée $\bar{t}_r(h)$ à partir de l'expression :

$$\bar{t}_r(h) = \bar{t}_{r'}(h) + \bar{t}_a(h - \bar{t}_r(h))$$

`(tr_ns)` ← **Composition_T**`(tr_ms , ta)`

5.4 Programmation du choix des usagers et de chargement en volume

Les deux problèmes sont programmés pour une destination donnée et toutes les origines correspondantes. Dans le cas du prototype, nous combinons la programmation du choix des usagers à celle du chargement en volume.

Pour le problème des choix des usagers nous codons une fonction de distribution de demande pour tous les itinéraires d'une relation O-D traitée : `Distribution_r`. Cette fonction partage la demande entre les itinéraires efficaces. Les inputs sont le profil de demande et l'ensemble des itinéraires efficaces, représentés par les atomes d'une relation O-D considérée. Les outputs sont les profils de volume chargés par chaque itinéraire.

Le problème de chargement en volume est traité en recourant à l'algorithme atomique de chargement depuis une destination fixée. Afin de calculer le volume chargé sur l'arc et sur le chemin considérés dans chaque boucle de propagation, nous établissons deux sous-fonctions de composition `Composition_Xa` et `Composition_Xr`.

5.4.1 Fonction principale

```
(Y_A_s , Y_R_s , Ieff_R_s) ← FUV(
        N , s , Q_s , M_s , W_n , t_A , ...)
```

Les paramètres de la fonction de chargement en volume **FUV** sont :

- `N` : la structure du réseau ;
- `s` : la destination traitée ;

- `Q_s` : les profils de demande pour chaque relation O-D de destination `s` ;
- `M_s` : l'ensemble de chemins/atomes efficaces ;
- `W_n` : les ensembles de chemins/atomes efficaces par nœud ;
- `t_A` : les profils de temps de traversée effectif des arcs ;
- ... : des paramètres auxiliaires.

Les outputs sont :

- `Y_A_s` : les profils de volume entrant des arcs, pour l'état auxiliaire `Y` ;
- `Y_R_s` : les profils de volume entrant des chemins de destination `s`, pour l'état auxiliaire `Y` ;
- `Ieff_R_s` : les intervalles efficaces pour chaque chemin de destination `s`.

Les étapes de traitement du chargement en volume sont les suivantes :

1. chargement des paramètres ;
2. les profils de volume entrant des arcs et des chemins efficaces sont initialisés à 0 ;
3. les profils de volume entrant des chemins par O-D sont calculés par la fonction auxiliaire `Distribution_R` ;
4. les profils de volume entrant des arcs sont établis par l'algorithme atomique de chargement.

Les étapes de traitement de la fonction auxiliaire `Distribution_R` sont les suivantes :

1. chargement des paramètres ;
2. les intervalles efficaces de chaque chemin dans `W_o_s` sont déterminés par la fonction auxiliaire `cal_int` ;
3. pour chaque chemin `r` dans `W_o_s`, la fonction auxiliaire `distribution_r` est utilisée afin de distribuer la demande sur `r`.

5.4.2 Fonctions auxiliaires

Fonction `distribution_r` de distribution de la demande pour un chemin `r` de `W_o_s`.

`Y_r` ← **distribution_r**(`Ieff_r , Q_o_s , r`)

Fonction `Distribution_R` de distribution de la demande pour l'ensemble `R` des chemins efficaces de `W_o_s`.

`Y_R_o_s` ← **Distribution_R**(`Ieff_R_o_s , W_o , Q_o_s`)

`Xa` ← **Composition_Xa**(`Xa , Xr`)

Fonction de composition du volume d'entrée d'un arc `a`.

`Xr` ← **Composition_Xr**(`Xr , Xra , tab , Ha`)

Fonction de composition du volume d'entrée d'un chemin `r`.

5.5 Programmation de l'équilibre offre-demande

La fonction `MSA` de calcul de l'équilibre offre-demande est appliquée à toutes les relations O-D par l'algorithme d'équilibrage MSA utilisant la méthode des combinaisons convexes au niveau des arcs.

Pour chaque itération de l'algorithme, un état auxiliaire est déterminé par l'appel successif des fonctions d'écoulement, de formation des services, de choix des usagers et de chargement en volume.

L'état courant à l'itération `k`+1 est déterminé par la fonction auxiliaire `Combine_Xa_Ya`, qui retourne la combinaison convexe de l'état auxiliaire `Ya` avec l'état courant `Xa` à l'itération `k`, en utilisant une suite de cœfficients `wk`.

La convergence est testée par divers critères discutés en partie 3 de cette thèse.

Les inputs sont la structure du réseau (variable `Network`), les profils de demande des relations O-D, les profils de capacité des arcs, les profils de temps de traversée minimum des arcs, et le nombre d'itérations.

Les outputs sont les profils de volume entrant sur les arcs, les profils de temps de traversées effectifs des arcs, les profils d'horaire de départ et de sortie.

Les étapes du traitement sont les suivantes :

1. la variable `Network`, les profils de demande de toutes relations O-D, les profils de la capacité en sortie, les valeurs de volume stock à l'instant de référence de tous les arcs du réseau sont chargés. Ces inputs seront fournis aux fonctions de quatre problèmes.
2. les profils de volume entrant des arcs et des chemins efficaces seront initialisés à 0.
3. Pour chaque itération, l'état auxiliaire est calculé selon les étapes suivantes :
   (a) Pour chaque arc du réseau, la fonction d'écoulement est appelée. Puis, pour chaque arc, le profil de temps de travaersée effectif est calculé, à partir des profils de volumes cumulés en entrée et en sortie.
   (b) Pour chaque destination, la fonction de formation des services est appelée. Elle détermine l'ensemble de chemins/atomes efficaces et les ensembles de chemins/atomes efficaces pour chaque couple O-D.
   (c) Pour chaque destination, la fonction de chargement est appelée.

On obtient ainsi les profils de volume entrant des arcs pour chaque destination.

Terminaison d'une itération. En fin d'itération, l'état courant est combiné avec l'état auxiliaire calculé en cours d'itération, puis le niveau de convergence est évalué. Différents critères de convergence ont été définis. Leur étude fait l'objet de la troisième partie de cette thèse.

5.5.1 Fonction principale

`void ← MSA_CCA(N, t0_A, K_A, Q, ...)`

La fonction `MSA_CCA` calcule l'équilibre offre-demande par la méthode des combinaisons convexes.

5.5.2 Fonctions auxiliaires

Les fonctions auxiliaires ci-dessous sont utilisées principalement lors de la combinaison convexe et du calcul des critères de convergence.

`Xa` ← **Combine_Xa_Ya**(`Xa , Ya , wk`)

Fonction de combinaison.

Cri_Int ← **Cri_Int**(`L_Xa , L_Xa1`)

Calcul du niveau de convergence d'après le critère intuitif.

`Moy_II` ← **Dif_P**(`P1 , P2`)

Calcul de la différence carrée moyenne de ces deux profils.

Cri_dual ← **Cri_dual**(`L_tA , L_Xa , L_Ya`)

Calcul du niveau de convergence d'après le critère d'écart de dualité.

`Dif` ← **Dif_V**(`P1 , P2 , P3`)

Calcul de l'écart de dualité entre l'état courant et l'état auxiliaire.

5.6 Le problème de fragmentation des profils

L'étude, à l'aide du prototype, du cas d'école décrit au chapitre 4 nous a permis de mettre en évidence un phénomène important : une fragmentation croissante des profils au cours des calculs, notamment autour des instants critiques.

Observons tout d'abord ce phénomène, en suivant, au cours du calcul, l'évolution du profil de temps de parcours sur les arcs a_1 et a_2. Les figures 5.4 à 5.3 représentent les profils de temps de parcours aux itérations 6, 51, 71, et 201, avec une attention particulière sur les points situés entre les instants critiques h_2 et h_3.

A l'itération 201, plus de 20 points sont apparus sur l'intervalle $[h_2; h_3]$, correspondant à la phase de déstockage sur l'arc a_1. Le graphique ne permet pas d'illustrer un point important : tous les morceaux correspondant à ces nouveaux points possèdent une pente légèrement différente de la pente moyenne entre les instants critiques h_2 et h_3.

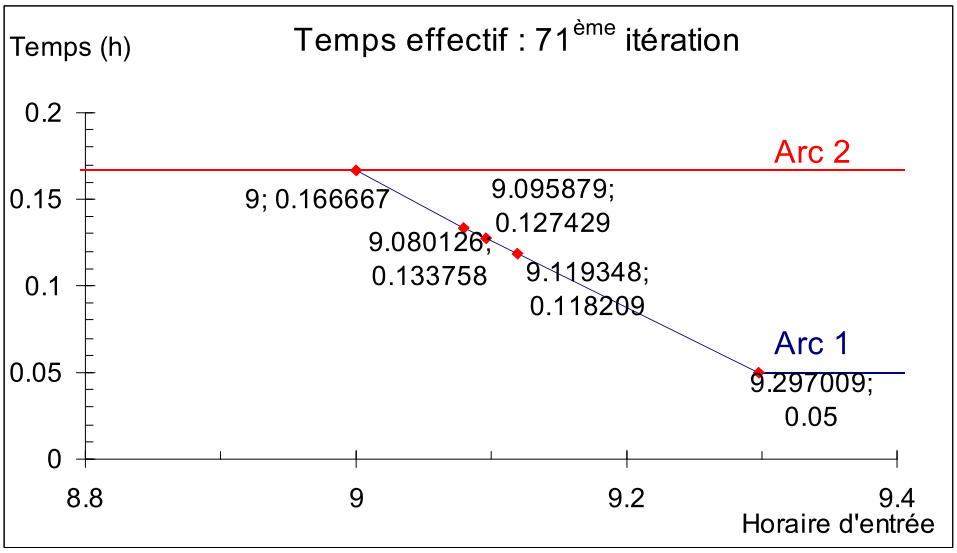


FIG. 5.2 – Temps de traversée à la 71ème itération.

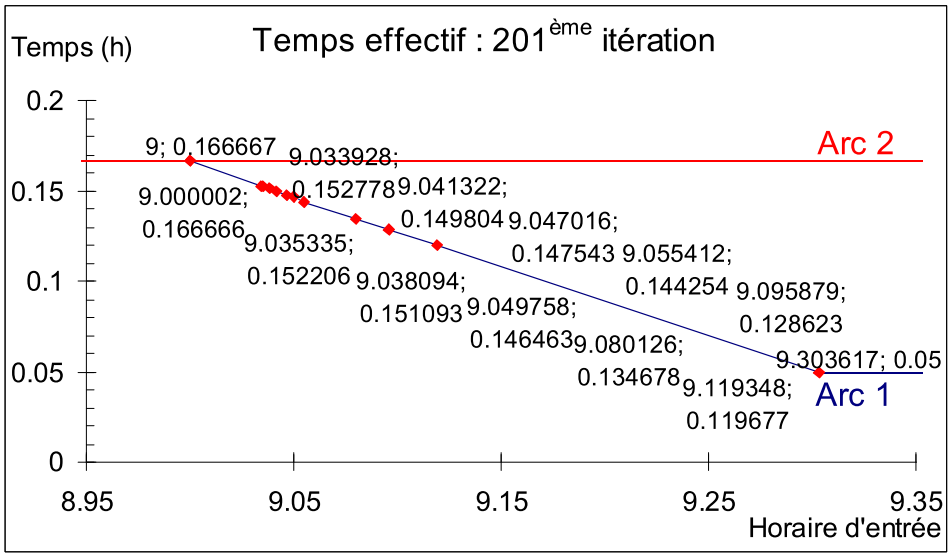


FIG. 5.3 – Temps de traversée à la 201ème itération.

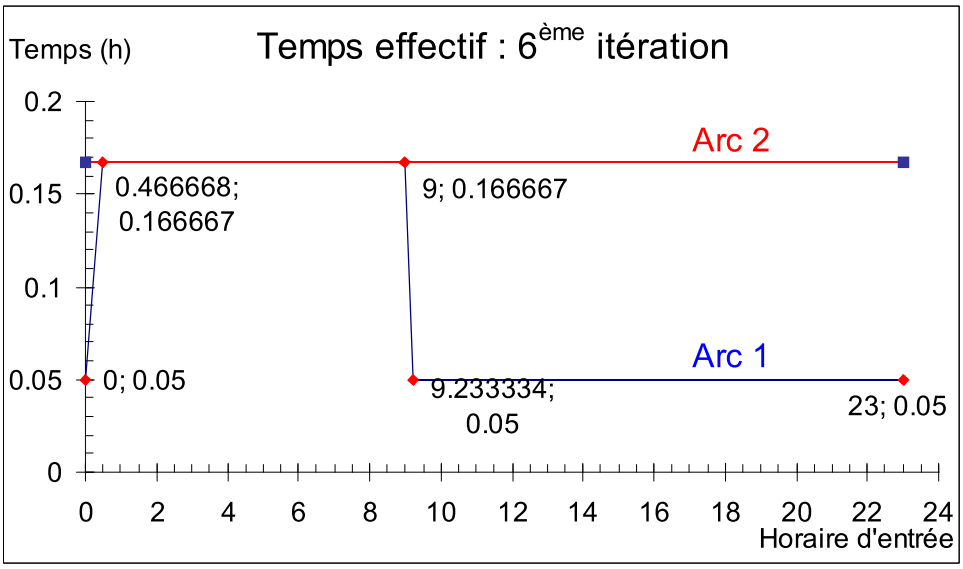


FIG. 5.4 – Temps de traversée à la 5ème itération.

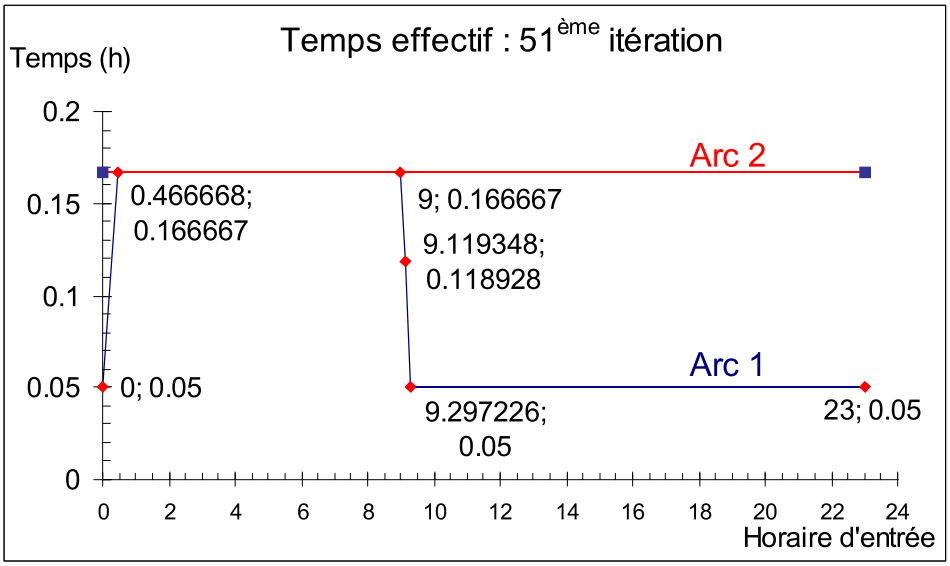


FIG. 5.5 – Temps de traversée à la 51ème itération.

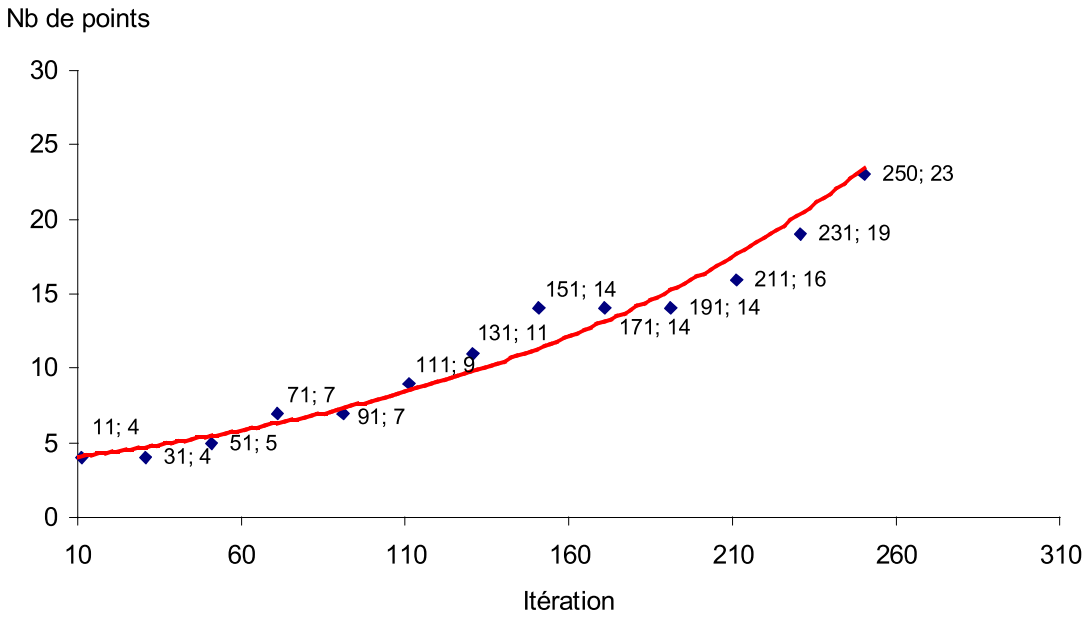


FIG. 5.6 – Augmentation du nombre de points du profil de volume en entrée de l'arc a_1.

De plus, la taille des données augmente de façon exponentielle au cours des calculs, comme l'illustre la figure 5.6. La courbe de tendance, en rouge, correspond à une régression exponentielle des points expérimentaux, en bleu.

Le caractère exponentiel de la fragmentation des profils au cours du calcul impose une attention toute particulière. En effet, celà pourrait gravement compromettre le caractère opérationnel du modèle LADTA, en limitant de façon drastique le nombre d'itérations calculables en pratique, du fait de l'augmentation conjointe de la taille des données et, corrélativement, du temps de calcul à chaque itération. Sur de grands réseaux, cette limite pourrait même être atteinte bien avant l'équilibre.

Il importe donc : d'une part, de comprendre et d'analyser quelles sont les étapes de calcul qui entraînent une augmentation de la dimension des profils de volume ; d'autre part, de proposer des moyens pour contrôler ce phénomène.

5.6.1 Position du problème

Les variables du modèle sont des fonctions affines par morceaux, appelés profils. Essentiellement, ces profils sont codés comme des tableaux de coordonnées (`x`,`y`). Nous dirons dans la suite qu'un profil est de dimension `n` s'il contient `n` couples de coordonnées (`x`,`y`).

Il convient alors de noter que si `P1` et `P2` sont des profils de dimensions respectives m et n, le résultat de toute combinaison linéaire de `P1` et `P2` est un profil de dimension comprise entre $\max\{m, n\}$ et $m + n$. Le cas le plus favorable est celui pour lequel les supports des profils coïncident, c'est-à-dire, en supposant $m \leq n$, que l'ensemble des `P1.x[i]` est inclus dans l'ensemble des `P2.x[j]`. Le cas le plus défavorable est atteint lorsque les supports de `P1` et `P2` sont disjoints. Cette remarque vaut également pour la plupart des fonctions prenant en argument deux profils et retournant un profil.

Ainsi, au cours du calcul de l'équilibre par la méthode des moyennes successives, la taille des données ne peut que croître (au mieux, rester constante) à chaque itération. En particulier, les profils de volume en entrée de chaque arc à l'itération $k+1$ ont nécessairement une taille supérieure ou égale à celle des profils de volume en entrée de chaque arc à l'itération k, car ils résultent d'une combinaison linéaire entre les profils de volume de l'état courant et les profils de volume de l'état auxiliaire.

Dans le pire des cas, cette inflation peut être exponentielle, avec un doublement, éventuellement à chaque itération, de la taille des données nécessaires pour stocker les profils de volume en entrée de chaque arc. Au mieux, cette taille reste constante d'une itération à l'autre.

Deux questions sont donc posées :

1. Pourquoi une fragmentation exponentielle des profils ?

2. Quels peuvent-être les moyens de contrôler la fragmentation des profils sans nuire à l'exactitude des calculs ?

5.6.2 Pourquoi une fragmentation exponentielle des profils ?

Une piste de réponse à la première question est celle des inévitables erreurs numériques. Il se peut en effet que, du fait de la précision finie des calculs effectués par ordinateur, deux points d'un profil qui seraient égaux à l'issue d'un calcul exact soient distincts à l'issue d'un calcul par ordinateur.

Cependant, si des problèmes d'approximation numériques sont pas à négliger, il existe une autre explication, liée de façon plus fondamentale au principe même du modèle LADTA.

Rappelons que :

- à l'équilibre, entre les instants h_1 et h_2, la demande sur l'arc a_1 doit être égale à la capacité (de 4 véh/h) de façon à maintenir un temps de parcours constant égal à 10 min.
- par construction du modèle LADTA et de l'algorithme d'équilibrage, le volume en entrée de a_1 est, sur l'intervalle $[h_1; h_2]$, le résultat d'une suite de combinaisons de type

$$X_1^{[k+1]} = (1 - w_k)X_1^{[k]} + w_k Y_1^{[k]}$$

  où $Y_1^{[k]}$ est, du fait du principe d'affectation tout ou rien :
  - de pente nulle lorsque $t_1^{[k]}$ excède le temps de parcours libre t_2 de l'arc a_2, égal à 10 min ;
  - de pente 5 véh/h lorsque $t_1^{[k]}$ est inférieur à 10 min.

En moyenne, au bout d'un nombre assez grand d'itérations, l'affectation de la demande se produit donc, pour chaque h de $[h_1; h_2]$, quatre fois sur cinq sur l'arc a_1 et une fois sur cinq sur l'arc a_2, ce pour assurer le débit à l'équilibre de 4 véh/h sur l'intervalle $[h_1; h_2]$.

Considérons maintenant un intervalle $[h'_1; h'_2]$ inclus dans $[h_1; h_2]$ tel que, à l'itération k, $t_1^{[k]}(h'_1) < t_2$ et $t_1^{[k]}(h'_2) > t_2$. Alors, il existe un instant h' compris strictement entre h'_1 et h'_2 tel que, à l'itération $k+1$, la demande sera affectée sur l'arc a_1 sur l'intervalle $[h'_1; h']$ et sur l'arc a_2 sur l'intervalle $[h'; h'_2]$.

S'il se produit alors que, pour $k' > k$, $t_1^{[k']}(h'_1) < t_2$, $t_1^{[k']}(h') > t_2$ et $t_1^{[k']}(h'_2) < t_2$, il se trouve que l'intervalle de départ $[h'_1; h'_2]$ a été subdivisé en deux intervalles pour lesquels ce même raisonnement peut être appliqué à nouveau, éventuellement au signe près.

C'est ce qui explique le caractère exponentiel de la fragmentation des profils au cours des calculs constaté par l'expérience.

Nous complétons cette analyse par une inspection plus détaillée des mécanismes de fragmentation des profils temporels à chaque étape de l'algorithme.

5.6.2.1 Sur la formation des chemins

Dans cet algorithme, l'étape de propagation calcule les profils $\overline{H}_r = \overline{H}_b \circ \overline{H}_{r'}$ et $t_r^-(\hbar) = t_{r'}^-(\hbar) + t_b^-\left(\overline{H}_{r'}(\hbar)\right)$. Ce calcul récursif pourrait doubler le nombre de morceaux pour les profils résultats. Mais la plupart de ces nouveaux morceaux auraient la même pente.

5.6.2.2 Sur le choix des usagers

Dans le choix des usagers, la demande se partage selon la concurrence entre les chemins par tranche d'horaire d'entrée. Le nombre de morceaux sur les profils entrants des chemins dépend du nombre de tranches de temps énumérées. Donc, plus nombreuses tranches de temps, plus nombreux morceaux sur le profil du volume entrant de chemins. Cet état de fait est de plus en plus ardu à mesure que le nombre d'itérations augmente et ceci en raison de la concurrence entre les chemins. Mais les nouveaux morceaux issus de ce problème n'ont pas le même pente comme la formation des chemins.

5.6.2.3 Sur le chargement en volume

Nous rencontrons un problème identique de multiplication des morceaux dans l'algorithme atomique de chargement causée par l'étape de propagation qui compose le volume cumulé pour les chemins et le volume chargé pour les arcs. Ces calculs peuvent aussi multiplier le nombre de morceaux sur les profils en résultat. Une partie de ces nouveaux morceaux peuvent avoir une pente identique.

5.6.2.4 Sur l'écoulement du flux

La multiplication des morceaux dans le problème d'écoulement est dû à :

- d'une part, l'algorithme de calcul du volume sortant, qui analyse l'état du stock, par instant de stockage h_1 et de déstockage h_2, selon le volume arrivé et la capacité en sortie, pourrait faire multiplier le nombre de morceaux selon celui de volume arrivé et de capacité. De plus, le nombre de morceaux du volume arrivé pourrait également doubler par profil de temps minimum et profil de volume entrant de l'arc.
- d'autre part, le calcul du temps effectif $t_a(h)$ qui peut doubler le nombre de morceaux du profil de temps effectif par celui de profil du volume entrant et sortant.

5.6.2.5 Sur l'équilibre offre-demande

L'algorithme des combinaisons convexes peut multiplier le nombre de morceaux par le nombre d'interactions entre les situations respectives d'un problème à l'autre dans le cercle de dépendance. Autrement dit, l'output d'un problème devient l'input du problème suivant, ce qui risque de faire exploser le nombre de morceaux qui croit également avec le nombre d'itération.

5.6.3 Comment contrôler la fragmentation des profils ?

Il est nécessaire dans un premier temps de s'assurer que toute portion continue d'un profil ne contient pas des morceaux successifs de même pente. Pour ce faire, toute fonction qui retourne un profil le fait par l'intermédiaire de la fonction `SimpliP`, décrite en 5.6.3.1.

Cependant, à l'évidence, cette simplification ne permet pas à elle seule de résoudre le problème de fragmentation des profils. La technique proposée à cette fin fait l'objet du 5.6.3.2.

5.6.3.1 Fonction de simplification

`Prof` ← **SimpliP**(`P`,`tol`)

Cette fonction permet de simplifier un profil par l'assimilation des morceaux contigus ayant le même pente DYDH à une valeur de tolérance près (paramètre tol). Elle est utilisée à la fin de toutes les fonctions qui établissent les profils. La façon d'assimiler les morceaux est la suivante :

morceau examiné n :

- point initial :[H_1(n);Y_1(n)]
- point final :[H_2(n);Y_2(n)]
- pente :DYDH(n)

morceau suivant :

- point initial :[H_1(n+1);Y_1(n+1)]
- point final :[H_2(n+1);Y_2(n+1)]
- pente :DYDH(n+1)

si le paramètre de tolérance est validé, alors regrouper n avec n+1.

Morceau d'assimilation est le suivant :

- point initial : [H_1(n);Y_1(n)]
- point final : [H_2(n+1);Y_2(n+1)]
- pente : à calculer

5.6.3.2 Fonction de regroupement

L'idée principale de la fonction de regroupement est d'effectuer un filtrage "passe-bas" du support d'un profil. Tout morceau d'un profil dont l'intervalle support est de longueur inférieure à une valeur seuil donnée intmin est regroupé avec ses voisins. Cette procédure assure que, pour une simulation de durée totale H, le nombre maximal de morceaux d'un profil ne peut excéder $\frac{H}{intmin}$, ce qui permet de déterminer la borne supérieure de la taille mémoire maximale à la simulation.

Cette façon de procéder n'est pas sans risque et peut avoir un impact non négligeable sur la convergence. Les conséquences de la procédure de regroupement sur la convergence de l'algorithme d'équilibrage est discutée au chapitre 6.

Prof ← **Group_M**(P, intmin)

Cette fonction permet de regrouper un morceau dont la tranche de temps [H_2 − H_1] est inférieure à un temps de base donné avec le morceau voisin. Cette fonction est utilisée afin de limiter l'émiettement que subissent les profils temporels au cours des calculs. Cette fonction ne traite que les profils de volume entrant des arcs, et qui sont des fonctions strictement croissantes. La façon de regrouper des morceaux est la suivante :

Pour le morceau n :

- Point initial : [H_1(n); Y_1(n)]
- Point final : [H_2(n); Y_2(n)]
- Pente : DYDH(n)

Morceau précédent :

- Point initial : [H_1(n−1); Y_1(n−1)]
- Point final : [H_2(n−1); Y_2(n−1)]
- Pente : DYDH(n−1)

Morceau suivant :

- Point initial : [H_1(n+1); Y_1(n+1)]
- Point final : [H_2(n+1); Y_2(n+1)]
- Pente : DYDH(n+1)

si [H_2(n)−H_1(n)] < intmin (temps de base), alors regrouper n avec soit n+1, soit n−1 selon la différence de la pente entre les morceaux.

Morceau de regroupement est le suivant :

- soit : point initial [H_1(n−1) ;Y_1(n−1)] ; point final [H_2(n);Y_2(n)] ; pente à calculer.
- soit : point initial [H_1(n) ;Y_1(n)] ; point final [H_2(n+1);Y_2(n+1)] ; pente à calculer.

Cela est réalisé successivement pour chaque morceau d'un profil.

Troisième partie

Formulations analytiques de l'équilibre dynamique et mesurage de la convergence

Introduction de partie

La troisième partie de cette thèse est consacrée à l'étude de critères de convergence pour caractériser l'équilibre offre-demande du modèle LADTA. Cette étude nous a conduit dans un premier temps à analyser, de façon à la fois large et rigoureuse, la notion d'équilibre offre-demande dans le cas de l'affectation dynamique. En retour, cette première phase nous a permis d'exhiber des notions nouvelles — relativité chronologique, correspondance temps-volume — qui fondent une grille d'analyse, que nous croyons novatrice et bien fondée, permettant de caractériser quelques bonnes propriétés d'un critère de convergence. La suite de cette introduction expose de façon plus détaillée la démarche adoptée.

Contexte et problématique

Dans un problème d'affectation du trafic aux itinéraires sur un réseau, l'équilibre du trafic traduit l'adaptation à long terme entre l'offre de services et la demande de déplacement : à l'équilibre, aucun usager n'aurait intérêt à changer de service de manière isolé. Ce principe, appelé principe de l'optimum individuel ou encore premier principe de Wardrop, permet d'interpréter simplement les résultats d'une affectation : sur une relation O-D et pour une classe d'usagers, pour une période d'horaires de départ, les itinéraires utilisés ont un coût identique et minimum parmi les chemins accessibles.

Ce principe de comportement individuel peut être couplé avec un modèle de formation du trafic et un modèle de congestion. Alors, les conditions d'équilibre contrôlent la répartition de la congestion sur les itinéraires. Il est important de simuler ce mécanisme pour éclairer la planification d'un réseau

de transport, afin de reproduire la répartition des usagers entre les services, et donc le niveau d'utilisation des routes et de leurs capacités.

Dans la pratique des études, l'affectation statique à l'équilibre du trafic est un modèle bien maîtrisé : on dispose de garanties théoriques sur l'existence et l'unicité (partielle) d'un état d'équilibre, et sur la convergence des algorithmes d'équilibrage (i.e. des algorithmes pour le calcul de l'équilibre). Rappelons qu'un tel algorithme procède par itérations successives afin d'approcher progressivement l'équilibre. Le principal algorithme utilisé, la méthode des combinaisons convexes, est robuste, et l'utilisateur du modèle se contente de commander un nombre déterminé d'itérations (typiquement 50 ou 100), et de mesurer à l'issue un indicateur de variation des débits des arcs entre deux itérations successives.

La situation est moins satisfaisante pour l'affectation dynamique, qu'elle soit simulée par un modèle microscopique ou macroscopique. Pour un modèle par simulation microscopique, même s'il n'y a pas de notion d'équilibre à proprement parler, la question se pose d'une convergence en moyenne pour un échantillon de simulations aux paramètres exogènes identiques : jusqu'à présent cette question n'a pas été traitée théoriquement ; les praticiens rencontrent des difficultés fortes[1], et ceux qui partagent leur expérience restent très modestes dans leurs préconisations. Dans les simulations macroscopiques avec priorité à l'écoulement du trafic, les enjeux de convergence sont plutôt relatif au schéma d'intégration des équations aux dérivées partielles : là encore on manque d'investigation sur les interactions potentielles avec un modèle de comportement des usagers. Enfin, pour la simulation macroscopique avec priorité au comportement des usagers comme dans le modèle LADTA, il semble que les chercheurs comme les utilisateurs se contentent de mesurer la convergence par des critères de type empirique, intuitif : qui ont été adaptés rapidement de l'affectation statique. C'est pourquoi nous avons entrepris une investigation systématique de la convergence dans un modèle d'affectation macroscopique du trafic, investigation dont le moment est venu de préciser les objectifs.

[1] témoignages de Cofiroute et de la Dirif pour des simulations sur des réseaux routiers franciliens

Objectif de connaissance et questionnement

Notre objectif de connaissance s'exprime par les questions suivantes :

- Comment, dans une simulation, mesurer rigoureusement l'écart entre un état courant (en termes de volumes et de coûts) et un état d'équilibre ?
- Quel algorithme convient-il d'employer pour progresser vers un état d'équilibre, et tout au moins mesurer l'écart ?
- Et plus en profondeur encore, comment caractériser un équilibre, en fonction de quelles variables endogènes basiques, selon quelle formule analytique ?

Nous posons ces questions pour tout modèle macroscopique d'affectation dynamique du trafic pour l'équilibre offre-demande, et en particulier pour le modèle LADTA.

Notre investigation porte en priorité sur les critères de convergence, dont nous voulons discerner les éventuelles propriétés : le pouvoir de résolution, i.e. qu'une certaine valeur du critère révèle effectivement un état d'équilibre ; l'interprétation en termes physiques et économiques ; la calculabilité i.e. la possibilité et la commodité de mesurer le critère.

Cette investigation nous a amenés à examiner la formulation mathématique de l'équilibre, et à envisager diverses reformulations, variées en fonction des variables endogènes basiques et des contraintes imposées ou relaxées. Elle nous a aussi conduits à examiner les algorithmes d'équilibrage, leurs variables et leurs traitements, notamment afin d'évaluer la calculabilité d'un critère.

Démarche

Nous avons commencé par une revue bibliographique de la littérature scientifique consacrée aux modèles macroscopiques d'affectation du trafic : revue peu fructueuse concernant notre question prioritaire, car les articles publiés traitent avant tout de formulations mathématiques et peu d'applications pratiques rapportées avec suffisamment de détail, et la question de la convergence n'a pas encore été traitée spécifiquement pour des problèmes

d'affectation de grande dimension.

Nous avons alors revisité le modèle d'affectation et la formulation de l'équilibre, afin de construire des critères d'écart et de convergence dont on puisse garantir le pouvoir de résolution. Pour cela nous avons procédé en deux temps : d'abord un retour critique sur l'affectation statique ; puis un examen approfondi, une véritable inquisition, de l'affectation dynamique. Nous avons ainsi identifié et caractérisé des complexités spécifiques d'une affectation dynamique : d'une part, la relativité chronologique des volumes des arcs ; d'autre part, la correspondance entre les temps et les flux, correspondance qui ne débouche que dans certains cas sur une concordance. L'équilibre nécessite une telle concordance, qui n'est toutefois pas une condition suffisante.

Ayant reconnu ces aspects, nous avons proposé des critères effectifs, en indiquant la possibilité de les mesurer pour un algorithme d'équilibrage particulier, la méthode des combinaisons convexes : qui en fait se dédouble dans le cas dynamique, entre un traitement au niveau des arcs et un traitement au niveau des chemins, aux résultats distincts. Enfin, nous avons proposé diverses formulations pour l'équilibre dynamique du trafic, qui sont susceptibles de favoriser la mesure des critères. Au total, nous avons fourni les concepts nécessaires pour formuler l'équilibre de diverses manières, pour faciliter la spécification d'une formulation, pour concevoir des traitements algorithmiques, et pour mesurer l'écart à un équilibre dynamique. Nous avons proposé des formulations économes, des algorithmes et des critères de convergence rigoureux.

Plan de la partie

Le corps de cette troisème partie est composé de quatre chapitres.

Au chapitre 6 nous examinons l'affectation statique : les diverses formulations de l'équilibre, leurs correspondances. Nous examinons les principaux critères employés pour mesurer un écart à l'équilibre et nous en discutons les propriétés, en particulier leur calculabilité dans un algorithme de combinaison convexe.

Le court chapitre 7 procède à un examen de la question de la conver-

gence des modèles d'affectation dynamique. Il donne brièvement l'état de la recherche concernant les formulations, les algorithmes et leur comportement de convergence vers l'équilibre, en matière d'affectation dynamique du trafic.

Le chapitre 8 est central, en ce qu'il expose nos contributions essentielles à la formulation analytique de l'équilibre dynamique d'une part, et à la définition d'algorithmes et de critères de convergence rigoureux d'autre part.

Enfin, le chapitre 9 expose une expérience numérique menée dans un double objectif : d'une part, pour comparer deux algorithmes de combinaison convexe portant sur les volumes respectivement par arcs ou par chemins, ainsi qu'un nouvel algorithme qualifié d'*hybride*, et d'autre part pour comparer plusieurs critères de convergence.

Une conclusion de partie fait la synthèse de ces contributions.

Chapitre 6

L'équilibre statique : formulation, obtention et mesurage

Introduction

Nous allons passer en revue les formulations, leurs correspondances, les algorithmes et les critères d'écart à l'équilibre dans le cas de l'affectation statique du trafic, afin de repérer les principaux critères et d'en caractériser les propriétés.

6.1 Hypothèses et notations

Définissons d'abord les éléments d'un modèle d'affectation statique, en nous restreignant à :

- du côté de la demande, une seule classe d'usagers, dont le choix d'itinéraire suit un comportement homogène et déterministe, avec un coût généralisé réduit au temps de parcours, et sans élasticité du volume au niveau de service.
- Du côté de l'offre, des itinéraires sans restriction d'accès aux arcs autre que la condition de non-multiplicité (pour l'occurrence d'un même arc dans un chemin), et des fonctions de congestion qui ne dépendent que du débit local de l'arc.

Voici les notations utilisées :

$G = (N, A)$ un réseau de transport, avec N l'ensemble des nœuds n et A l'ensemble des arcs $a = (n_a^+, n_a^+)$ à sommets dans N.

A_n^+ (resp. A_n^-) l'ensemble des arcs sortants (resp. entrants) du nœud n.

T_a le temps de traversée de l'arc a.

x_a le débit (flux) de l'arc a.

t_a la fonction temps-débit de l'arc a.

S l'ensemble des sommets de destination s.

O_s l'ensemble des sommets d'origine o pour la destination s.

$I = \bigcup_{s \in S} O_s$ l'ensemble des relations O-D $i = (o, s)$.

R_i l'ensemble des chemins sans multiplicité r sur la relation O-D i.

q_i le flux sur l'O-D i.

x_r le débit (flux) du chemin r.

T_r le temps de parcours du chemin r.

σ_{ar} indicateur binaire égal à 1 si r traverse a, noté $a \in r$, et 0 sinon.

$x_a^s = \sum_{o \in O_s} \sum_{r \in R_{os}} x_r$ débit (flux) de l'arc a pour la destination s.

u_n^s coût (éventuellement réduit au temps) depuis n jusqu'à la destination s.

6.2 Formulation par arcs et par chemins

6.2.1 Conditions d'équilibre par arcs et destinations

On peut caractériser l'équilibre du trafic par le système d'équations suivant :

$$x_a^s \geq 0, \forall s \in S, a \in A \tag{6.1a}$$

$$\sum_{a \in A_n^+} x_a^s = q_{ns} + \sum_{a \in A_n^-} x_a^s, \forall s \in S, n \in N \tag{6.1b}$$

$$x_a = \sum_{s \in S} x_a^s, \forall a \in A \tag{6.1c}$$

$$T_a = t_a(x_a) \tag{6.1d}$$

$$\nu_a = T_a, \forall a \in A \tag{6.1e}$$

$$u_n^s \leq \nu_a + u_m^s, \forall s \in S, a = (n, m) \in A \tag{6.1f}$$

$$x_a^s \left(\nu_a + u_m^s - u_n^s\right) = 0, \forall s \in S, a = (n, m) \in A \tag{6.1g}$$

Ce système exprime :

- que les débits par arcs et destinations sont non-négatifs (6.1a) et respectent la conservation du flux en chaque nœud pour chaque destination (6.1b) ;
- que le flux total d'un arc se décompose par destinations (6.1c) et induit le temps de parcours (6.1d) ;
- que le coût par arc est réduit au temps de parcours (6.1e) et induit le coût vers chaque destination (6.1f et 6.1g).

La condition (6.1g) est précisément une *condition des écarts complémentaires* entre le flux et les coûts par destination, et jointe à (6.1a) et (6.1f) elle impose que u_n^s soit le coût minimal de n à s sur la base des coûts des arcs ν_a.

6.2.2 Fonction de lagrangien et problème de point-selle

Le système de conditions (6.1) se compose de relations linéaires d'égalité ou d'inégalité, à l'exception des fonctions de congestion (6.1d) qui pourraient prendre une forme plus générale $t_a(x_A)$ ou même $t_a\left(x_A^S\right)$ en notant $x_A = (x_a)_{a\in A}$ et $x_A^S = (x_a^s)_{s\in S, a\in A}$.

Dans le cas d'une dépendance simple $t_a(x_a)$, les conditions (6.1) sont équivalentes aux conditions de point-selle pour la fonction de lagrangien suivante :

$$\mathcal{L}_A\left(x_A^S, x_A, u_I, \nu_A\right) = \sum_{a\in A}\int_0^{x_a} t_a(z)dz - \sum_{a\in A}\nu_a\cdot\left(x_a - \sum_{s\in S}x_a^s\right)$$

$$+\sum_{s\in S, n\in N}u_n^s\cdot\left(q_{ns} + \sum_{a\in A_n^-}x_a^s + \sum_{a\in A_n^+}x_a^s\right) \tag{6.2a}$$

Ce lagrangien admet pour dérivées partielles :

$$\frac{\partial\mathcal{L}_A}{\partial x_a^s} = \nu_a + u_m^s - u_n^s, \forall s\in S, a = (n,m)\in A \tag{6.2b}$$

$$\frac{\partial\mathcal{L}_A}{\partial x_a} = t_a\left(x_a\right) - \nu_a, \forall a\in A \tag{6.2c}$$

$$\frac{\partial\mathcal{L}_A}{\partial u_n^s} = q_{ns} + \sum_{a\in A_n^-}x_a^s - \sum_{a\in A_n^+}x_a^s, \forall s\in S, n\in N \tag{6.2d}$$

$$\frac{\partial\mathcal{L}_A}{\partial \nu_a} = \sum_{s\in S}x_a^s - x_a, \forall a\in A \tag{6.2e}$$

Le lagrangien est à minimiser en fonction de x_A^S et x_A sous les contraintes $x_a^s \geq 0$, et à maximiser en fonction de u_I et ν_A, d'où les conditions de point-selle suivantes :

$$\frac{\partial\mathcal{L}_A}{\partial x_a^s} \geq 0, \forall s\in S, a\in A \tag{6.2f}$$

$$x_a^s\cdot\frac{\partial\mathcal{L}_A}{\partial x_a^s} = 0, \forall s\in S, a\in A \tag{6.2g}$$

$$\frac{\partial \mathcal{L}_A}{\partial x_a} = 0, \forall a \in A \tag{6.2h}$$

$$\frac{\partial \mathcal{L}_A}{\partial u_n^s} = 0, \forall s \in S, n \in N \tag{6.2i}$$

$$\frac{\partial \mathcal{L}_A}{\partial \nu_a} = 0, \forall a \in A \tag{6.2j}$$

Effectivement (6.2f) coïncide avec (6.1f), (6.2g) avec (6.1g), (6.2h) avec (6.1d et 6.1e), (6.2i) avec (6.1b) et (6.2j) avec (6.1c) : les conditions de point-selle du lagrangien équivalent à la définition de l'équilibre.

6.2.3 Formulation extrémale primale par arcs et destinations

Quand le problème primal-dual par arcs et destinations (6.2) admet un lagrangien $\mathcal{L}_A$, un équilibre satisfait les conditions de point-selle du lagrangien, et on peut considérer une fonction objectif primale :

$$J_A\left(x_A^S, x_A\right) = \max_{U_I, \nu_A} \mathcal{L}_A\left(x_A^S, x_A, u_I, \nu_A\right)$$

dont la minimisation en fonction de x_A^S et x_A équivaut à la solution du problème de point-selle, et donc à l'équilibre du trafic formulé par arcs et destinations (6.1).

On montre que la fonction objectif primale admet l'expression :

$$J_A\left(x_A^S, x_A\right) = \sum_{a \in A} \int_0^{x_a} t_a\left(z\right) dz \tag{6.3a}$$

Le programme primal de minimiser J_A reste soumis aux contraintes suivantes :

$$x_a^s \geq 0, \forall s \in S, a \in A \tag{6.3b}$$

$$\sum_{a \in A_n^+} x_a^s = q_{ns} + \sum_{a \in A_n^-} x_a^s, \forall s \in S, n \in N \tag{6.3c}$$

$$x_a = \sum_{s \in S} x_a^s, \forall a \in A \tag{6.3d}$$

En effet, la maximisation du lagrangien en fonction de u_I et ν_A implique (6.3c) et (6.3d) respectivement ; et donc l'annulation des deux derniers termes du lagrangien $\mathcal{L}_A$.

6.2.4 Formulation extrémale duale par arcs et destinations

Symétriquement à l'établissement du programme primal, à partir du problème de point-selle du lagrangien on peut établir un programme dual, pour la fonction objectif

$$G_A(u_I, \nu_A) = \min_{x_A^S, x_A} \mathcal{L}_A\left(x_A^S, x_A, u_I, \nu_A\right)$$

dont la maximisation en fonction de u_I et ν_A équivaut à la solution du problème de point-selle, et donc à l'équilibre du trafic formulé par arcs et destinations (6.1).

On montre que la fonction objectif duale admet l'expression

$$G_A(u_I, \nu_A) = \sum_{a \in A} \int_{t_{a0}}^{\nu_a} t_a^{-1}(\theta)\, d\theta - \sum_{s \in S, n \in N} q_{ns} \cdot u_n^s \tag{6.4a}$$

Le programme dual reste soumis aux contraintes

$$u_n^s \leq u_m^s + \nu_a, \forall s \in S, a = (n, m) \in A \tag{6.4b}$$

$$\nu_a \geq t_{a0} = t_a(0), \forall a \in A \tag{6.4c}$$

En effet, la minimisation du lagrangien en fonction de x_A^S et x_A implique (6.2f) qui coïncide avec (6.4b), et (6.2h) qui implique (6.4c). La condition (6.2g) implique que seuls les arcs qui sont efficaces pour la destination s compte tenu des potentiels u_m^s et u_n^s et du coût ν_a , i.e. ceux tels que $u_n^s \leq$

$u_m^s + \nu_a$, peuvent porter un flux x_a^s positif. Ainsi les deux termes extraits du lagrangien,

$$\sum_{a \in A, s \in S} \nu_A \cdot x_a^s$$

et

$$\sum_{s \in S, n \in N} u_n^s \cdot \left(\sum_{a \in A_n^-} x_a^s - \sum_{a \in A_n^+} x_a^s \right) = \sum_{s \in S, a \approx (n,m) \in A} x_a^s \cdot (u_m^s - u_n^s)$$

se contrebalancent, et la fonction duale se réduit à

$$\mathcal{L}_A \left(x_A^S, x_A, u_I, \mu_A \right) = \sum_{a \in A} \left(\int_0^{x_a} t_a(z)\, dz - \nu_A \cdot x_a \right) + \sum_{s \in S, n \in N} u_n^s \cdot q_{ns}$$

Si la fonction de congestion t_a est inversible de réciproque t_a^{-1}, (6.2h) équivaut à $x_a = t_a^{-1}(\nu_a)$, et on peut remplacer

$$\begin{aligned} \int_0^{x_a} t_a(z)\, dz - \nu_A \cdot x_a &= \left[z \cdot t_a(z) \right]_0^{x_a} - \int_0^{x_a} z \cdot \dot{t}_a(z)\, dz - \nu_a \cdot x_a \\ &= - \int_0^{x_a} z \cdot \dot{t}_a(z)\, dz \\ &= - \int_{t_a}^{\nu_a} t_a^{-1}(\theta)\, d\theta \end{aligned}$$

en intégrant par parties, par le changement de variable $\theta = t_a(z)$.

C'est pourquoi la fonction duale suit l'expression (6.4a).

Le programme dual en fonction de u_I et ν_A se simplifie encore en un programme dual restreint en fonction uniquement de ν_A, car la minimisation selon u_I, jointe à la minimisation selon x_A^S et x_A, entraîne que

$$\sum_{a \in A_n^+} x_a^s = q_{ns} + \sum_{a \in A_n^-} x_a^s, \forall s \in S, n \in N$$

$$u_n^s \leq \nu_a + u_m^s, \forall s \in S, a \approx (n,m) \in A$$

$$x_a^s \cdot (\nu_a + u_m^s - u_n^s) = 0, \forall s \in S, a \approx (n,m) \in A$$

L'inégalité locale, reportée sur tous les arcs a qui composent un chemin r de o à s, implique que

$$\nu_r = \sum_{a \in r} \nu_a \geq \sum_{a \in r} u^s \left(n_a^+ \right) - u^s \left(n_a^- \right) = u^s \left(n_r^+ \right) - u^s \left(n_r^- \right) = u_o^s$$

La conservation du flux entraîne que x_A^s est un flot de destination s. Si $q_o s > 0$ alors il y a un chemin r de o à s de flux $x_r > 0$, et le long d'un tel chemin $x_a^s > 0$ pour tout $a \in r$. La condition d'écart complémentaire implique que

$$\nu_a = u^s\left(n_a^+\right) - u^s\left(n_a^-\right), \forall a \in r$$

$$\nu_r = \sum_{a \in r} \nu_a = u_o^s$$

Donc u_o^s est égal au coût d'un chemin chargé, et par conséquent les chemins chargés ont un coût identique et minimal parmi les chemins accessibles. On le note $U_{os}(\nu_A)$, car il dépend uniquement de ν_A.

Ainsi $q_{os} \cdot u_o^s = q_{os} \cdot U_{os}(\nu_A)$, ce qui est vrai aussi pour $q_{os} = 0$. Donc

$$\tilde{G}_A(\nu_A) = \sum_{a \in A} \int_{t_{a_0}}^{\nu_a} t_a^{-1}(\theta)\, d\theta - \sum_{s \in S, n \in N} q_{ns} \cdot U_{ns}(\nu_A) \tag{6.4d}$$

De plus, la conservation des flux implique que

$$\sum_{n \in N} q_{ns} \cdot u_n^s = \sum_{n \in N} u_n^s \cdot \left(\sum_{a \in A_n^+} x_a^s - \sum_{a \in A_n^-} x_a^s \right)$$

$$= \sum_{a \approx (n,m) \in A} x_a^s \cdot (u_n^s - u_m^s) = \sum_{a \in A} x_a^s \cdot \nu_a$$

et donc

$$\sum_{s \in S, n \in N} q_{ns} \cdot U_{ns}(\nu_A) = \sum_{a \in A} \nu_A \cdot X_a(\nu_A) \tag{6.4e}$$

avec $X_a(\nu_A) = \sum_{s \in S} x_a^s$ le débit total sur l'arc a issu de l'affectation des q_I aux plus courts chemins sur la base des coûts ν_A.

6.3 Formulation par chemins

6.3.1 Conditions d'équilibre par chemins

L'équilibre du trafic se caractérise également de la manière suivante :

$$x_r \geq 0, \forall i \in I, r \in R_i \tag{6.5a}$$

$$\sum_{r \in i} x_r = q_i, \forall i \in I \tag{6.5b}$$

$$x_a = \sum_{r \in a} x_r, \forall a \in A \tag{6.5c}$$

$$T_a = t_a\left(x_a\right), \forall a \in A \tag{6.5d}$$

$$\nu_a = T_a, \forall a \in A \tag{6.5e}$$

$$\nu_r = \sum_{a \in r} \nu_a, \forall r \in R = \bigcup_{i \in I} R_i \tag{6.5f}$$

$$\nu_r \geq u_i, \forall i \in I, r \in R_i \tag{6.5g}$$

$$x_r\left(\nu_r - u_i\right) = 0, \forall i \in I, r \in R_i \tag{6.5h}$$

Ce système exprime que les flux des chemins sont non-négatifs (6.5a), qu'ils portent les flux O-D (6.5b), qu'ils composent les débits des arcs (6.5c) qui eux-mêmes induisent les temps de parcours au niveau des arcs (6.5d). Ces temps représentent les coûts par arcs (6.5e) et ceux-ci induisent les coûts des chemins (6.5f). La condition (6.5h) est une condition des écarts complémentaires, qui jointe à (6.5g) implique que u_i est le coût minimal sur la relation O-D i, minimal parmi les chemins de cette relation. (6.5h) stipule que seuls les chemins de coût minimal peuvent porter un flux positif.

6.3.2 Fonction de lagrangien et problème de point-selle

Le système de conditions (6.5) se compose de relations linéaires d'égalité ou d'inégalité, à l'exception là encore des fonctions de congestion (6.5d). Pour une dépendance simple, (6.5) équivaut aux conditions de point-selle pour la fonction suivante de lagrangien :

$$\mathcal{L}_R\left(x_R, x_A, u_I, \nu_A\right) = \sum_{a\in A}\int_0^{x_a} t_a\left(z\right)dz - \sum_{a\in A}\nu_a\cdot\left(x_a - \sum_{r\in a}x_r\right) + \sum_{i\in I}u_i\cdot\left(q_i - \sum_{r\in I}x_r\right) \tag{6.6a}$$

Ce lagrangien admet pour dérivées partielles :

$$\frac{\partial\mathcal{L}_R}{\partial x_r} = \nu_r - u_{i(r)}, \forall i\in I, r\in R_i \tag{6.6b}$$

$$\frac{\partial\mathcal{L}_R}{\partial x_a} = t_a\left(x_a\right) - \nu_a, \forall a\in A \tag{6.6c}$$

$$\frac{\partial\mathcal{L}_R}{\partial u_i} = q_i - \sum_{r\in I}x_r, \forall i\in I \tag{6.6d}$$

$$\frac{\partial\mathcal{L}_R}{\partial \nu_a} = \sum_{r\in a}x_r - x_a, \forall a\in A \tag{6.6e}$$

Le lagrangien est à minimiser en fonction de x_R et x_A sous les contraintes $x_r \geq 0$, et à maximiser en fonction de u_I et ν_A, d'où les conditions de point-selle suivantes :

$$\frac{\partial\mathcal{L}_R}{\partial x_r} \geq 0, \forall i\in I, r\in R_i \tag{6.6f}$$

$$x_r\cdot\frac{\partial\mathcal{L}_R}{\partial x_r} = 0, \forall i\in I, r\in R_i \tag{6.6g}$$

$$\frac{\partial\mathcal{L}_R}{\partial x_a} = 0, \forall a\in A \tag{6.6h}$$

$$\frac{\partial\mathcal{L}_R}{\partial u_i} = 0, \forall i\in I \tag{6.6i}$$

$$\frac{\partial \mathcal{L}_R}{\partial \nu_a} = 0, \forall a \in A \tag{6.6j}$$

Effectivement, (6.6f) coïncide avec (6.5g), (6.6g) avec (6.5h), (6.6h) avec (6.5d et 6.5e), (6.6i) avec (6.5b) et (6.6j) avec (6.5c).

6.3.3 Formulation extrémale primale par chemins

Quand le problème primal-dual par chemins (6.5) admet un lagrangien $\mathcal{L}_R$, un équilibre satisfait les conditions de point-selle du lagrangien, et à la suite de Beckmann *et al.* [BMW56] on peut considérer une fonction objectif primale

$$J_R(x_R, x_A) = \max_{u_I, \nu_A} \mathcal{L}_R(x_R, x_A, u_I, \nu_A)$$

dont la minimisation en fonction de x_R et x_A équivaut à la solution du problème de point-selle, et donc à l'équilibre du trafic formulé par chemins (6.5).

On montre que

$$J_R(x_R, x_A) = \sum_{a \in A} \int_0^{x_a} t_a(z)\, dz \tag{6.7a}$$

Le programme primal de minimiser J_R reste soumis aux contraintes :

$$x_r \geq 0, \forall i \in I, r \in R_i \tag{6.7b}$$

$$\sum_{r \in i} x_r = q_i, \forall i \in I \tag{6.7c}$$

$$x_a = \sum_{r \in a} x_r, \forall a \in A \tag{6.7d}$$

En effet, la maximisation du lagrangien en fonction de u_I et ν_A implique (6.7c) et (6.7d) respectivement, et par conséquent l'annulation des deux derniers termes du lagrangien $\mathcal{L}_R$.

6.3.4 Formulation extrémale duale par chemins

Symétriquement à l'établissement du programme primal, à partir du problème de point-selle du lagrangien du système (6.6) on peut établir un programme dual, pour la fonction objectif

$$G_R\left(u_I,\nu_A\right)=\min_{x_R,x_A}\mathcal{L}_R\left(x_R,x_A,u_I,\nu_A\right)$$

dont la maximisation en fonction de u_I et ν_A équivaut à la solution du problème de point-selle, et donc à l'équilibre du trafic formulé par chemins du système d'équations (6.5).

On montre que la fonction duale par chemins admet l'expression suivante

$$G_R\left(u_I,\nu_A\right)=\sum_{a\in A}\int_{t_{a0}}^{\nu_a}t_a^{-1}\left(\theta\right)d\theta-\sum_{i\in I}q_i\cdot u_i \tag{6.8a}$$

En notant $\nu_r=\sum_{a\in r}\nu_a$, le programme dual reste soumis aux contraintes

$$\nu_r\geq u_i,\forall i\in I,r\in R_i \tag{6.8b}$$

$$\nu_a\geq t_{a0}=t_a\left(0\right),\forall a\in A \tag{6.8c}$$

En effet, la minimisation du lagrangien en fonction de x_R et x_A implique (6.6h) donc (6.5d et 6.5e), et (6.6f) donc (6.8b), et (6.6g) donc (6.5h). Aussi seuls sont chargés les chemins de coût ν_r minimal sur leur relation O-D.

Dans le lagrangien optimisé en x_R et x_A, les deux termes

$$\sum_{a\in A,r\in a}\nu_a\cdot x_r\text{ et }\sum_{i\in I}u_i\cdot\sum_{r\in R_i}x_r$$

se contrebalancent, et la fonction duale se réduit à

$$G_R\left(u_I,\nu_A\right)=\mathcal{L}_R\left(\widehat{x}_A^S,\widehat{x}_A,u_I,\nu_A\right)=\sum_{a\in A}\left[\int_0^{x_a}t_a\left(z\right)dz-\nu_a\cdot\widehat{x}_a\right]+\sum_{i\in I}u_i\cdot q_i$$

Le programme dual se simplifie en un programme dual restreint en fonction uniquement de ν_A, en remplaçant u_I par $U_I\left(\nu_A\right)$ comme pour la formu-

lation par arcs et destinations.

6.4 Problèmes-types et propriétés formelles

Nous allons passer en revue les formulations mathématiques en problème type d'optimisation et de recherche opérationnelle, qui caractérisent l'affectation statique et permettent d'établir des propriétés formelles concernant l'existence et l'unicité d'une solution. Pour présenter les formulations nous utilisons une "expression générique" du modèle, afin de nous abstraire de la lourdeur des détails mathématiques.

6.4.1 Expression générique pour l'affectation statique

Nous exprimons le modèle d'affectation statique par un système de quatre conditions :

$$t_R = F_S\left(t_A\right) \tag{6.9a}$$

$$x_R \in F_U\left(t_R\right) \tag{6.9b}$$

$$x_A = F_V\left(x_R\right) \tag{6.9c}$$

$$t_A = F_F\left(x_A\right) \tag{6.9d}$$

Ce système exprime que les temps des chemins résultent des temps des arcs par une fonction de *formation des services* f_S qui est linéaire (6.9a) ; que les flux des chemins dépendent de leurs temps de traversée, par une fonction de *choix des usagers* f_U (6.9b) ; que les flux des arcs découlent de ceux des chemins, par une fonction de *chargement en volume* f_V qui est linéaire (6.9c) ; que les temps des arcs résultent de leur flux, par une fonction de congestion appelée aussi une *fonction d'écoulement du trafic* et notée f_F (pour Flowing) (6.9d).

On peut formuler le modèle également par arcs et destinations, de la manière suivante :

$$u_N^S = \tilde{f}_S\left(t_A\right) \tag{6.10a}$$

$$x_A^S \in \tilde{f}_U\left(u_N^S, t_A\right) \tag{6.10b}$$

$$x_A = \tilde{f}_V\left(x_A^S\right) \tag{6.10c}$$

$$t_A = \tilde{f}_F\left(x_A\right) \tag{6.10d}$$

Dans la suite de la section, compte tenu de l'équivalence entre la formulation par arcs et destinations et la formulation par chemins, nous nous concentrons sur l'expression générique par chemins.

Donnons encore trois définitions fondamentales :

E_R ensemble faisable des flux de chemins

$$E_r = \left\{ x_R = \left[x_r\right]_{r \in R}, x_r \geq 0, \sum_{r \in i} x_r = q_i, \forall i \in I \right\}$$

V_R fonction de coût des chemins, réduite ici au temps

$$V_R\left(x_R\right) = f_S \circ f_F \circ f_V\left(x_R\right)$$

F_R fonction de coût réduite des chemins, en notant $\mu_r = \min_{\rho \in i(r)} V_\rho$

$$F_R\left(x_R\right) = V_R\left(x_R\right) - \mu_R\left(V_R\left(x_R\right)\right)$$

6.4.2 Le problème de complémentarité non linéaire (NCP)

L'affectation statique se formule comme le problème de complémentarité non linéaire suivant :

$$\textit{Trouver } x_R \in E_R \textit{ tel que } \begin{cases} x_R & \geq & 0 \\ F_R(x_R) & \geq & 0 \\ x_R \cdot F_R(x_R) & = & 0 \end{cases}$$

On reconnaît là une version condensée de la formulation primale-duale de l'équilibre, avec notamment la condition des écarts complémentaires.

6.4.3 Le problème d'inéquation variationnelle (VIP)

L'affectation statique se formule aussi comme le problème d'inéquation variationnelle suivant :

$$\textit{Trouver } {x_R}^* \in E_R \textit{ tel que } V_R({x_R}^*) \cdot (x_R - {x_R}^*) \geq 0, \forall x_R \in E_R$$

Pour $x_R, {x_R}^* \in E_R$, le produit $\mu_R \cdot x_R = u_I \cdot q_I = \mu_R \cdot {x_R}^*$, donc

$$V_R(x_R^*) \cdot (x_R - x_R^*) = F_R(x_R^*) \cdot (x_R - x_R^*)$$

Le problème (VIP) équivaut au problème (NCP) car :

- Une solution x_R^* du problème (NCP) vérifie $F_R(x_R^*) \cdot x_R^* = 0$, et donc $V_R(x_R^*) \cdot (x_R - x_R^*) = F_R(x_R^*) \cdot x_R$. Comme $F_R(x_R^*) \geq 0$ et $x_R \geq 0$ (car x_R est dans E_R), le produit scalaire est non négatif, et donc x_R^* résoud aussi l'inéquation variationnelle.
- Une solution x_R^* du problème (VIP) appartient à E_R, donc $x_R^* \geq 0$. Par construction, $F_R(x_R^*) \geq 0$. Il reste à vérifier la condition des écarts complémentaires. Considérons un chemin r tel que $x_r^* > 0$ et supposons $F_r(x_R^*) > 0$. Alors il existe $\rho \in R_{i(r)}$ tel que $V_r > V_\rho = \mu_i$. Donc, pour le vecteur x_R formé à partir de x_R^* en retirant un petit $\epsilon > 0$ de x_r et en l'ajoutant à x_ρ, on a

  $$F_R(x_R^*) \cdot x_R = F_R(x_R^*) \cdot x_R^* + \epsilon \cdot (V_\rho - V_r) < F_R(x_R^*) \cdot x_R^*$$

  et donc

  $$V_R(x_R^*) \cdot (x_R - x_R^*) < 0$$

  Contradiction.

Par conséquent, $x_r^* \cdot F_r(x_R^*) = 0, \forall r \in R$ et x_R^* résoud le problème (NCP).

On dispose de théorèmes généraux établissant des propriétés formelles pour les inéquations variationnelles : si V_R est continue, et si E_R est compact (ou V_R est coercive ou fortement monotone), alors il existe une solution. Celle-ci est unique si V_R est strictement monotone.

6.4.4 Le problème de point fixe (FPP)

L'affectation statique se formule également comme le problème de point fixe pour l'application multivoque

$$W_R(x_R) = \arg\min \{V_R(x_R) \cdot (y_R - x_R), y_R \in E_R\}$$

Le problème FPP se formule :

Trouver $x_R \in E_R$ *tel que* $x_R \in W_R(x_R)$

Par construction de W_R, le problème FPP relatif à W_R équivaut au problème (VIP) relatif à V_R, car x_R résout le problème FPP si et seulement si la fonction $y_R \mapsto V_R(x_R) \cdot (y_R - x_R)$ admet 0 comme valeur minimale, ce qui est équivalent au fait que x_R résout le problème (VIP).

Pour le problème de point fixe, un théorème (de Brouwer) établit l'existence d'une solution quand E_R est compact et W_R est univoque et continue. Un théorème établit l'unicité d'une solution quand de plus W_R est strictement contractante.

6.4.5 Formulation extrémale primale (PEP)

Quand la fonction de congestion est simple, l'affectation statique se formule aussi comme le programme d'optimisation primal suivant :

Trouver $x_R^* \in E_R$ *solution de* $\min J(x_R)$, *pour* $x_R \in E_R$

Pour assurer l'existence d'une solution, il suffit que J soit continue, et que E_R soit compact ou J fortement convexe. S'il existe une solution, si J est

convexe alors l'ensemble des solutions est convexe. De plus, si la convexité est stricte alors la solution est unique.

La formulation extrémale implique la formulation en (VIP), qui exprime les conditions d'optimalité du premier ordre sur une solution du programme extrémal. En effet $V_R = \nabla J$.

6.4.6 Sur les formulations duales

La composition de la fonction de coûts par chemins,

$$V_R = f_S \circ f_F \circ f_V (x_R)$$

court-circuite les variables duales, et toute formulation basée sur V_R (ou W_R, ou J) est intrinsèquement primale, en fonction des variables endogènes basiques x_R.

Une formulation duale a pour variables endogènes basiques, le champ ν_A des coûts par arcs : dans le modèle basique d'équilibre, on se restreint aux temps des arcs t_A. Les autres variables endogènes s'en déduisent par les relations suivantes :

$$t_R = F_S (\nu_A) \tag{6.11a}$$

$$X_R \in F_U (t_R) \tag{6.11b}$$

$$X_A = F_V (X_R) \tag{6.11c}$$

$$\nu_A \approx t_A = F_F (x_A) \tag{6.11d}$$

La formulation primale-duale de l'équilibre n'est pas complètement symétrique entre les variables primales et duales. Le modèle de congestion fait dépendre ν_A de x_A, et en cela la formulation privilégie les variables primales. Dans le cas simple où la congestion sur chaque arc est une fonction monotone $t_a = t_a (x_a)$, on inverse cette fonction, et on remplace (6.11d) par :

$$x_A = f_F^{-1}(\nu_A) \tag{6.11e}$$

Alors l'approche duale consiste essentiellement à accorder $F_V(X_R(\nu_A))$ et $f_F^{-1}(\nu_A)$, i.e. les deux manières d'obtenir les flux des arcs.

Dans le programme extrémal dual restreint, la fonction objectif $\tilde{G}$ à maximiser a pour (sous-)gradient, en notant $X_R(\nu_A) \in f_U \circ f_S(\nu_A)$,

$$\nabla\tilde{G}(\nu_A) = f_F^{-1}(\nu_A) - F_V(X_R(\nu_A))$$

Ainsi les conditions d'optimalité (annulation du gradient) consistent à accorder les deux représentations des flux des arcs.

La condition privilégiée pour l'approche duale est quand l'application f_U est univoque, car alors $X_R(\nu_A) = f_U \circ f_S(\nu_A)$ i.e. le sous-gradient est un gradient.

6.5 Les correspondances entre formulations

6.5.1 Correspondance de dualité

Les programmes extrémaux respectivement primal et dual, sont en dualité l'un de l'autre. Ils sont mutuellement équivalents, et chacun d'eux équivaut au problème de point-selle du lagrangien. Cela vaut d'une part pour la formulation par arcs et destinations et d'autre part pour la formulation par chemins.

6.5.2 La conservation du flux est identique par arcs et par chemins

Regroupons par destination $s \in S$ et arc $a \in A$ les flux des chemins r, sous la notation $x_a^s = \sum_{r \in s \cap a} x_r$. Nous pouvons examiner le bilan des flux des chemins qui desservent s en chaque nœud $n \in N$, qui est la fonction suivante

du flot x_R :

$$\sum_{a \in A_n^+} x_a^s - \sum_{a \in A_n^-} x_a^s$$

Un chemin r qui passe par le nœud n, ou bien commence en n auquel cas il n'appartient qu'à la première somme, ou bien commence strictement en amont de n auquel cas il contribue de manière équilibrée aux deux sommes. Donc, la différence des deux sommes se réduit à la somme des flux des chemins qui desservent s et commencent en n, et donc à q_{ns} en raison de (6.5b).

Ainsi les conditions (6.5b) et (6.1b) sont équivalentes.

6.5.3 Les inégalités de coût locales impliquent celles par chemins

Dans la formulation primale-duale par arcs et destinations (1), considérons l'inégalité (6.1f) entre le coût local ν_a de l'arc $a = (n, m)$ et les potentiels u_n^s, u_m^s pour la destination s :

$$\nu_a \geq u_n^s - u_m^s$$

En sommant le long des arcs qui composent un chemin r, il vient

$$\nu_r = \sum_{a \in r} \nu_a \geq \sum_{a \in r} u^s\left(n_a^+\right) - u^s\left(n_a^-\right) = u^s\left(n_r^+\right) - u^s\left(n_r^-\right) = u_o^s$$

Ainsi les potentiels u_n^s de la formulation par arcs et destinations, satisfont les conditions associées aux potentiels u_i dans la formulation par chemins.

6.5.4 Les écarts complémentaires sont identiques par arcs ou chemins

Dans la formulation primale-duale par chemins, considérons la fonction d'écart complémentaire par relation O-D $i = (n, s)$:

$$Z_i^R = \sum_{r \in I} x_r \cdot (\nu_r - u_i)$$

Elle vérifie que

$$\begin{aligned}\textstyle\sum_{r\in i} x_r \cdot (\nu_r - u_i) &= \textstyle\sum_{r\in i} x_r \cdot (\sum_{a\in r} \nu_a) - (\sum_{r\in i} x_r) \cdot u_i \\ &= \textstyle\sum_{a\in A} \nu_a \cdot x_a^i - q_i \cdot u_i\end{aligned}$$

en notant $x_a^i = \sum_{r\in a\cap i} x_r$

En totalisant sur les relations i qui desservent la destination s, puis en utilisant la conservation locale du flux de la destination, il vient

$$Z_s^R = \sum_{a\in A} \nu_a.x_a^s - \sum_{n\in N} q_{ns}.u_{n/R}^s = \sum_{a\in A} x_a^s.\left(\nu_a + u_{/N}^s\left(n_a^-\right) - u_{/N}^s\left(n_a^+\right)\right)$$

Cette expression correspond à la fonction d'écart complémentaire par destination dans la formulation primale-duale par arcs et destinations

$$Z_s^A = \sum_{a\in A} x_a^s.\left(\nu_a + u_{/A}^s\left(n_a^-\right) - u_{/A}^s\left(n_a^+\right)\right)$$

quand on peut identifier $u_{I/R}$ et $u_{A/A}^S$.

6.5.5 L'état optimum par arcs résout le problème par chemins

En un état optimal du problème par arcs et destinations (1), les variables duales u_n^s satisfont les conditions sur les variables duales du problème par chemins (5) : tant les inégalités de coût, que la condition des écarts complémentaires. Donc en choisissant un flot par chemins qui correspond aux variables primales x_a^s de l'état optimal par arcs, ce flot par chemins associé aux variables duales est solution optimale du problème par chemins.

6.5.6 Un état optimum par chemins résout le problème par arcs

La démonstration en sens réciproque est moins simple qu'en sens direct, car les variables duales par chemins en un état quelconque (éventuellement non optimal) ne vérifient pas nécessairement les contraintes locales de coût

$\nu_a \geq u_n^s - u_m^s$ pour $a \approx (n, m)$.

Définissons donc des potentiels auxiliaires $\hat{u}_n^s$ de la manière suivante :

$$\hat{u}_n^s = \min \{\nu_r, r \in R_{ns}\}$$

Les inégalités de coût des chemins impliquent que $\hat{u}_n^s \geq u_n^s$.

De plus, ces potentiels auxiliaires vérifient les contraintes locales de coût des arcs :

$$\nu_a \geq \hat{u}_n^s - \hat{u}_m^s$$

par définition des plus courts chemins qui induisent les $\hat{u}$.

Par construction du champ $\hat{u}$, pour une relation O-D $i = (n, s)$ et tout chemin r de n à s, on a

$$\nu_r \geq \hat{u}_n^s \geq u_n^s$$

Comme les x_r sont non négatifs, il vient

$$0 \leq x_r \cdot (\nu_r - \hat{u}_n^s) \leq x_r \cdot (\nu_r - u_n^s)$$

donc la condition des écarts complémentaires sur l'état optimal par chemins $\left(x_R, u_N^S\right)$ entraîne celle pour l'état $\left(x_R, \hat{u}_N^S\right)$ qui est donc optimal pour le problème par chemins. Cet état est également optimal pour le problème par arcs, puisqu'il vérifie la conservation locale du flux par destination (cf. 6.5.2) ; les contraintes locales sur les coûts(cf. 6.5.3) ; la condition des écarts complémentaires (cf. 6.5.4).

6.5.7 Commentaires

Pour le modèle basique d'équilibre statique, les formulations primale-duale par arcs et chemins sont équivalentes. L'équivalence persiste pour certaines extensions :

- la prise en compte de plusieurs classes de trafic, éventuellement avec des restrictions d'accès à certains arcs pour certaines classes.
- des fonctions de congestion complexes, qui relient le temps d'un arc à plusieurs flux par arcs.

- l'explicitation de contraintes de capacité sur certains flux par arcs.
- des fonctions de coût qui incorporent non seulement le temps subi, mais aussi d'autres termes liés à la traversée des arcs : frais monétaires, inconfort...
- l'élasticité du volume de demande au coût généralisé, par classe d'usagers et par relation O-D.
- le choix d'itinéraire sur des réseaux à lignes de transport public.
- la sophistication du choix : pour un réseau multimodal, pour le choix conjoint du mode et de l'itinéraire ; ou pour le choix conjoint de destination et d'itinéraire ; ou même d'origine et destination et itinéraire.

Mais il existe certaines extensions qui portent spécifiquement au niveau des chemins, et ne peuvent être formulées seulement au niveau des arcs :

- des fonctions de coût d'itinéraire, non linéaires en fonction des coûts des arcs ; ou dont certains termes dépendent de la traversée de plusieurs arcs non nécessairement consécutifs, notamment des péages par couple de gares d'accès et sortie d'un sous-réseau.
- certaines segmentations de la demande, non seulement par relation origine-destination et classe de trafic, mais encore selon certains arbitrages et notamment l'arbitrage entre prix et temps : modèles multicritères.
- une répartition du flux O-D entre itinéraires, plus sophistiquée que l'affectation aux options de coût minimum : notamment quand on incorpore une composante aléatoire au coût d'itinéraire. Ainsi des modèles de choix discret de type logit ou probit etc.

Les extensions "de seconde espèce" rompent la linéarité de la relation entre les arcs et les chemins concernant la formation des coûts, et alors il n'y a pas de formulation par arcs et destinations. De plus, quand on étend le modèle basique, le rôle des variables de coût évolue : on ne peut considérer les coûts comme des variables duales que si le modèle jouit d'une formulation primale-duale ou extrémale. Le principal obstacle à cela n'est pas le raffinement du modèle de demande, mais plutôt la considération de fonctions de congestion à plusieurs variables endogènes de flux par arcs ou par classes de trafic.

6.6 Algorithmes et calculabilité

Cette section a pour objet de relier les algorithmes d'équilibrage aux formulations mathématiques, en vue de préparer l'examen des critères de convergence en termes d'une part de variables endogènes, d'autre part de calculabilité.

Le principal algorithme utilisé en pratique est la méthode des combinaisons convexes, qui consiste à résoudre une suite de problèmes auxiliaires (6.6.1) et qui a de bonnes propriétés théoriques (6.6.2). D'autres algorithmes primaux traitent explicitement les itinéraires : ils sont beaucoup plus efficaces en nombre d'itinéraires, mais leur itération typique coûte plus cher (6.6.3). Il existe des algorithmes duaux, qui manquent d'efficacité pour le modèle basique (6.6.4). Enfin des algorithmes primaux-duaux n'ont été développés, que pour des variables duales associées à des contraintes de capacité (6.6.5).

6.6.1 Le problème auxiliaire : une affectation à coûts constants

Pour faciliter la minimisation de la fonction objectif primale, ou de manière équivalente la résolution de l'inéquation variationnelle (problème VIP), on procède par étapes, chaque étape consistant à résoudre un problème auxiliaire relatif à un état courant du flux. Dans ce problème, on fixe les coûts à leur valeur en l'état courant.

Ainsi, en fonction des flux par chemins, à l'état courant x_R le problème auxiliaire est

$$\min \tilde{J}(y_R) = J(x_R) + \nabla J(x_R) \cdot (y_R - x_R) \tag{6.12}$$

soumis aux contraintes (6.7b) à (6.7d).

Ce problème auxiliaire est un programme d'optimisation linéaire sous contraintes, facile à résoudre car il se décompose selon les relations O-D, et pour chaque O-D c'est une affectation du flux q_i aux plus courts chemins sur la base des temps t_A induits par l'état x_R.

La méthode du problème auxiliaire a été théorisée par Guy Cohen [Coh80,

Coh88]. Elle s'applique aussi à un problème formulé en inéquation variationnelle, en remplaçant ∇J par V_R.

6.6.2 Les méthodes de combinaison convexe

Dans une méthode de combinaison convexe, on fait évoluer par des itérations successives un vecteur de flux afin de lui faire vérifier l'équilibre. Le vecteur x_A vérifie à tout moment les contraintes de flux, et dans une itération on le combine à un flux faisable auxiliaire y_A selon certaines proportions $(1-\xi_k, \xi_k)$ à l'itération k.

Les étapes essentielles d'une méthode de combinaison convexe sont présentées par l'algorithme 6.1.

début
  initialisation :
  $k \leftarrow 0$
  $x_A \leftarrow 0$
  répéter
    $t_A \leftarrow [t_a(x_a)]_{a \in A}$
    Chargement
    sur la base des temps t_A, affecter aux plus courts chemins la matrice O-D q_I, ce qui produit un flot auxiliaire y_A.
    Combinaison
    $x_A \leftarrow (1-\xi_k)\, x_A + \xi_k y_A$
    $k \leftarrow k+1$
  jusqu'à x_A *vérifie une condition de convergence* ;
fin

Algorithme 6.1 : Algorithme des combinaisons convexes.

On fixe $\xi_0 = 1$ pour obtenir un flot faisable. Aux itérations suivantes, dans la variante *de Frank-Wolfe*, on choisit $\xi_k \in [0;1]$ tel que l'état $(1-\xi_k)\, x_A + \xi_k y_A$ minimise la fonction objectif primale sur le segment $[x_A; y_A]$.

Dans la variante dite *des moyennes successives*, on utilise $\xi_k = \frac{1}{k+1}$ ou une fonction proche, l'essentiel étant de vérifier que $\sum_k \xi_k = \infty$ et $\sum_k \xi_k^2 < \infty$.

L'algorithme est spécifié en fonction d'un vecteur de flux par arcs x_A. Le traitement d'un flot par chemins x_R est entièrement équivalent, et produit les mêmes coûts t_A et le même vecteur x_A.

Il est facile d'évaluer le critère d'écart complémentaire au point x_A, grâce au vecteur auxiliaire y_A qui par définition minimise $V_R(x_R) \cdot (y_R - x_R)$:

$$\begin{aligned} \textstyle\sum_{r\in R} x_r \cdot \left(t_r - u_{i(r)}\right) &= \textstyle\sum_{r\in R} x_t \cdot t_r - \sum_{i\in I} \left(\sum_{r\in I} x_r\right) \cdot u_i \\ &= \textstyle\sum_{a\in A} x_a \cdot t_a - \sum_{i\in I} q_i \cdot u_i \\ &= \textstyle\sum_{a\in A} x_a \cdot t_a - \sum_{a\in A} y_a \cdot t_a \\ &= t_A (x_A - y_A) \end{aligned} \tag{6.13}$$

Au total, la méthode des combinaisons convexes est un algorithme simple dont la convergence vers l'équilibre est prouvée (cf. la revue très large de Patriksson [Pat94]). Cependant elle charge à certaines itérations des chemins qui ne sont pas optimaux in fine, et ensuite elle les décharge lentement mais jamais totalement, ce qui ralentit la convergence. De plus la direction de descente $y_A - x_A$ est de moins en moins productive à mesure que l'on approche de l'équilibre.

6.6.3 Algorithmes primaux traitant les chemins

La méthode des combinaisons convexes identifie des chemins à l'étape de chargement. Mais comme cette information ne sert qu'à calculer le vecteur auxiliaire, on ne la conserve pas pour la suite. D'autres algorithmes ont été proposés qui utilisent spécifiquement les chemins. Les deux plus efficaces traitent à chaque itération la succession des relations O-D :

- dans l'algorithme d'équilibrage par transvasement, on recherche un plus court chemin sur la relation, et on transvase du flux du chemin le plus chargé vers ce chemin de coût minimum, jusqu'à égaliser les coûts ou vider le chemin plus cher.
- dans l'algorithme de décomposition simpliciale désagrégée (DSD de Larsson et Patriksson [LP94]), on résout par la méthode de Newton un problème auxiliaire dont la fonction objectif est une approximation quadratique de la fonction objectif primale, en fonction des flux des

chemins de la relation.

Ces algorithmes convergent en peu d'itérations (excellente précision en 5 itérations), mais chaque itération présente une complexité élevée. Les critères de convergence de la méthode des combinaisons convexes sont disponibles, ainsi que des critères en fonction des chemins.

6.6.4 Algorithmes duaux

Le principal algorithme dual pour l'affectation statique, est une adaptation par Fukushima [Fuk84] de la méthode de sous-gradient de Polyak [Pol67] appliquée au programme dual restreint en fonction des ν_A. A chaque itération, le vecteur ν_A est modifié selon une direction de montée $t_A^{-1}(\nu_A) - X_A(\nu_A)$, le vecteur $X_A(\nu_A)$ étant le vecteur de flux par arcs issu d'une affectation de la matrice O-D aux plus courts chemins sur la base des coûts ν_A.

Les pas de montée sont des ξ_k pré-déterminés, devant vérifier les mêmes conditions que pour la méthode de combinaison convexe.

L'algorithme dual est spécifié au niveau des arcs. La spécification au niveau des chemins est équivalente puisque la relation entre les chemins et les arcs est linéaire à la fois du côté des flux et du côté des coûts.

Un autre algorithme dual a été proposé par Cascetta et Cantarella [CC95], à savoir une méthode des combinaisons convexes portant sur les coûts des arcs. Elle s'applique de manière privilégiée aux modèles stochastiques, pour lesquels la formulation en problème de point fixe est efficace (la fonction étant alors une application contractante).

6.6.5 Sur les algorithmes primaux-duaux

Comme la faisabilité primale (conservation du flux) est désirable et facile à obtenir, il existe peu d'algorithmes primaux-duaux qui la relaxent. La principale circonstance concerne les contraintes de capacité sur les flux des arcs, parfois appelées des *contraintes associées* : les multiplicateurs associés sont des variables duales.

Le schéma usuel d'algorithme primal-dual pour l'affectation statique est une méthode de pénalisation. Il consiste à résoudre une suite de problèmes

non contraints, dans chacun desquels l'effet des contraintes relaxées est incorporé à la fonction de coût au moyen des variables duales, qui sont fixées pour ce problème. A l'issue du problème, appelé *itération maîtresse*, on actualise les variables duales, et donc les fonctions de coût du prochain problème.

Notamment la méthode de lagrangien augmenté a été appliquée pour traiter des contraintes associées.

6.7 Critères pour mesurer un écart à l'équilibre

Pour l'affectation statique, on dispose de deux principaux critères de convergence rigoureux : l'écart aux coûts minimum, présenté en 6.7.1 et l'écart primal-dual, objet du 6.7.2. Les praticiens utilisent également des critères d'écart inter-itération en débit ou en temps, au niveau des arcs ou des chemins, mais nous en déconseillons l'emploi pour les algorithmes de combinaison convexe, pour les raisons exposées au 6.7.3.

6.7.1 L'écart aux coûts minimum

Dans les formulations en inéquation variationnelle et en problème de point fixe, nous avons considéré l'application

$$y_R \mapsto V_R(x_R) \cdot (y_R - x_R)$$

en signalant que si x_R minimise la fonction, alors la valeur optimale est nulle et x_R résout l'inéquation variationnelle. À x_R fixé, le problème auxiliaire

$$\min V_R(x_R) \cdot (y_R - x_R)$$

admet une solution pour une valeur optimale négative ou nulle, l'annulation équivalant à la solution de l'inéquation variationnelle.

Donc, en la solution y_R du problème auxiliaire associé à x_R, le critère

$$V_R(x_R) \cdot (y_R - x_R) \tag{6.14}$$

est positif et s'annule seulement à l'équilibre : il convient comme critère de convergence.

Ainsi ce critère a un bon pouvoir de résolution. Il jouit également d'une bonne interprétation économique : comme le problème auxiliaire se décompose par relation O-D,

$$\sum_{r\in R} t_r \cdot (x_r - y_r) = \sum_{r\in R} x_r \cdot t_r - \sum_{i\in I} q_i \cdot u_i \text{ avec } q_i = \sum_{r\in i} y_r \text{ et } u_i = \min_{\rho\in i} t_\rho$$

$$= \sum_{r\in R} x_r \cdot \left(t_r - \min_{\rho\in i} t_\rho\right)$$

Cette égalité justifie l'appellation d'*écarts aux coûts minimum.*

De plus, le critère est facile à calculer, car d'après (6.13)

$$V_R\left(x_R\right) \cdot \left(x_R - y_R\right) = t_A\left(x_A\right) \cdot \left(x_A - y_A\right)$$

Le critèe possède également une interprétation mathématique grâce à la théorie de la dualité : au point $\nu_A = t_A\left(x_A\right)$, le flot chargé aux plus courts chemins sur la base des temps ν_A est le $X_A\left(\nu_A\right)$ de l'équation (6.4e). C'est aussi le flot y_A solution du programme auxiliaire en x_A, donc la formule (6.4d) devient ici

$$\tilde{G}\left(t_A\left(x_A\right)\right) = \sum_{a\in A} \int_0^{x_a} t_a\left(z\right) dz + t_a \cdot \left(y_a - x_a\right)$$

C'est pourquoi l'écart $J\left(x_A\right) - \tilde{G}\left(t_A\left(x_A\right)\right) = \sum_{a\in A} t_a \cdot \left(x_a - y_a\right)$ est aussi appelé l'*écart de dualité.*

Pour mieux apprécier l'ordre de grandeur du critère, on l'évalue relativement au coût minimum global, sous la forme

$$\frac{t_A \cdot \left(x_A - y_A\right)}{t_A \cdot x_A}$$

Ce critère relativisé mesure, en moyenne sur le flux O-D global, l'écart relatif entre les coûts des chemins chargés et le coût minimum.

6.7.2 L'écart primal-dual

Quand on résout la formulation extrémale primale par une suite de problèmes auxiliaires, chacun d'eux fournit une borne inférieure β_k sur la valeur optimale du programme originel, J^* :

$$\beta_k = J\left(x_A^{(k)}\right) + t_A\left(x_A\right) \cdot \left(y_A^{(k)} - x_A^{(k)}\right) \tag{6.15}$$

En effet, par définition d'une solution optimale y_R du problème auxiliaire en x_R , on vérifie qu'en un état d'équilibre x_R^*,

$$\tilde{J}\left(y_R\right) \leq \tilde{J}\left(x_R^*\right)$$

Or, si J est convexe (notamment quand les t_a sont croissantes), on sait que

$$J^* = J\left(x_R^*\right) \geq J\left(x_R\right) + \nabla J\left(x_R\right) \cdot \left(x_R^* - x_R\right)$$

et en combinant les deux inégalités et la définition de $\tilde{J}\left(x_R\right)$, il vient que

$$J^* \geq \tilde{J}\left(y_R\right)$$

Donc $\tilde{J}\left(y_R\right) = J\left(x_R\right) + \nabla J\left(x_R\right) \cdot \left(y_R - x_R\right)$ constitue une borne inférieur sur J^*.

En définissant

$$\alpha_k = \max\left\{\alpha_{k-1}, \beta_k\right\} \tag{6.16}$$

chaque α_k est aussi une borne inférieure sur J^*, plus stricte que β_k. C'est pourquoi l'écart $J\left(x_A^{(k)}\right) - \alpha_k$ majore l'écart $J\left(x_A^{(k)}\right) - J^*$ de plus près que l'écart aux coûts minimum. Nous l'appelons l'*écart primal-dual* afin de l'identifier en particulier. C'est aussi un écart de dualité.

Ce critère a un bon pouvoir de résolution. Il est facile à calculer, il jouit d'une interprétation mathématique claire, mais son interprétation physico-économique est moins intuitive que l'écart aux coûts minimum.

6.7.3 Les écarts inter-itérations et leurs risques

Dans les études opérationnelles, et aussi dans les communications scientifiques traitant des modèles d'affectation statique sophistiqués, on emploie surtout la méthode des combinaisons convexes, et on mesure la convergence par un critère d'écart entre les variables endogènes considérées à deux itérations successives, l'actuelle et la précédente. Ainsi l'écart inter-itération en débit des arcs se formule

$$\sqrt{\frac{1}{|A|}\sum_{a\in A}\left|x_a^{(k+1)}-x_a^{(k)}\right|^2} \tag{6.17}$$

avec éventuellement des pondérations par la longueur et la capacité des arcs. Un tel critère indique le niveau de variation d'une itération à la suivante. Il s'annule à l'équilibre.

Mais il est difficile de relativiser sa valeur afin de décider d'arrêter les itérations ! En effet, notamment, dans une méthode des combinaisons convexes, chaque terme vérifie, par construction, que

$$\left|x_a^{(k+1)}-x_a^{(k)}\right| = \xi_k \left|y_a^{(k)}-x_a^{(k)}\right| \tag{6.18}$$

Puisque la suite des coefficients ξ_k tend vers 0 au cours de l'algorithme, une valeur faible du critère pourrait ne rien indiquer d'autre que cette décroissance du coefficient, sans que la distance $\left|y_a^{(k)}-x_a^{(k)}\right|$ ne diminue pour autant.

C'est pourquoi nous en déconseillons l'usage.

D'autres critères du même type concernent les temps

$$\sqrt{\frac{1}{|A|}\sum_{a\in A}\left|t_a^{(k+1)}-t_a^{(k)}\right|^2}$$

ou les coûts subis

$$\frac{1}{|A|}\sum_{a\in A} x_a^{(k+1)}\left|t_a^{(k+1)}-t_a^{(k)}\right|$$

éventuellement par chemins, en remplaçant A par R dans les variables de

temps et flux.

Ils présentent les mêmes qualités et risques que le critère inter-itération en débit des arcs de l'équation (6.17).

Chapitre 7

Examen des modèles d'affectation dynamique

Nous esquisserons un historique du développement des modèles pour délimiter le champ de notre revue en 7.1, puis nous évoquerons brièvement une série de modèles récents en 7.2. Nous pourrons alors discuter des expressions, formulations, algorithmes et critères de convergence en 7.3.

7.1 Nature de l'examen bibliographique, en lien avec le développement historique des modèles

Les prémices de l'affectation dynamique remontent aux années 1960, avec le problème de plus court chemin dynamique traité à la Ford-Bellman par Jocksh [Jok66] puis à la Dijkstra par Dreyfus [Dre69]. Le premier modèle temporalisé intégrant la congestion remonte à Merchant et Nemhauser [MN78], qui ont considéré une seule destination et une congestion "quasi-statique " sous la forme d'une relation temps-débit sans notion de file d'attente. Dans les années 1980 plusieurs extensions dynamiques ont été proposées à partir de l'affectation statique : traitement quasi-statique pour le modèle Saturn de Van Liet [VV82] et le modèle de Janson [Jan91]. Ces modèles quasi-statiques

ont été consolidés et capitalisés dans la monographie de Ran et Boyce [RB96], avec une formalisation déjà assez mature, mais une formulation insuffisamment rigoureuse : la formulation extrêmale prônée alors, a été depuis prouvée défectueuse.

En parallèle, les simulateurs microscopiques ont été développés progressivement. Depuis les lois de poursuite posées par Pipes et Reuschhell dans les années 1950 [Pip53], en passant par la simulation microscopique de jonctions dans les années 1970 et 1980, pour aboutir dans les années 1990 à des simulateurs capables de traiter sur micro-ordinateur, des réseaux de dimension moyenne (quelques dizaines de jonctions, quelques centaines d'arcs) : notamment CORSIM et VISSIM [Fel94].

Le modèle CONTRAM de Leonard *et al.* [LTB78] a été un précurseur dès les années 1980, avec un traitement "mésoscopique" par paquets de véhicules.

En parallèle également, a été développé la modélisation dynamique macroscopique de l'écoulement : principes des modèles par équations aux dérivées partielles du premier ordre par Lighthill et Whitham [LW55] et Roberts (1956), modèles du second ordre d'inspiration cinétique (Prigogine et Herman, [PH71]) ou mécanique, comme Payne [Pay79]. Des développements pour le contrôle dynamique du trafic ont eu lieu dans les années 1980, notamment par Papageorgiou et Hadj-Salem [PBHS90].

Le rapprochement entre les deux branches de modélisation dynamique macroscopique, d'une part l'affectation et d'autre part l'écoulement, a été déclenché par le modèle de Newell [New93]. Celui-ci a traité l'écoulement de manière simplifiée en termes de volumes cumulés, plutôt que de débit. Ce traitement était déjà familier aux ingénieurs du trafic pour analyser les phénomènes d'attente à une jonction par le modèle de goulot (file d'attente verticale) de May [May64].

Le modèle de goulot, qui traite la congestion de manière plus réaliste que la relation temps-débit, a été intégré dans l'affectation dynamique par Kuwahara et Akamatsu pour l'affectation depuis une origine [KA93], puis pour l'affectation à l'optimum individuel sur la base des temps instantanés et immédiats [KA97], ensuite pour l'optimum collectif et pour l'optimum individuel sur la base des temps propagés [KA01]. Leur formulation est un

problème de point fixe en fonction des débits par arcs et destinations. Elle est résolue par combinaison convexe.

Nous faisons commencer notre revue à ces modèles de Kuwahara et Akamatsu, et nous la bornons à des modèles d'affectation dynamique avec file d'attente. Pour plus d'information concernant les modèles antérieurs, nous renvoyons aux revues de Ben-Akiva *et al.*[BADPK91] et de Peeta et Ziliaskopoulos [PZ01].

7.2 Qualification de quelques modèles récents

Wie, Tobin et Carey [WTC02] ont exprimé un modèle d'équilibre à congestion quasi-statique, choix d'itinéraire et d'horaire et avec une élasticité de la demande : ils l'ont formulé comme un problème de point fixe en fonction des débits par arcs et destinations, et résolu par la méthode des combinaisons convexes.

Le modèle de goulot a été substitué par Ran *et al.* [RSHC00] au modèle quasi-statique de congestion locale. Il en résulte un modèle d'équilibre formulé comme une inéquation variationnelle en fonction des débits des chemins, résolu par un algorithme de combinaison convexe, sans intégration du chargement en volume et de l'écoulement du trafic.

Tong et Wong [TW00] ont exprimé un modèle d'équilibre avec file d'attente, formulé comme une inéquation variationnelle en fonction des débits des chemins, et traité par un algorithme de combinaison convexe, en intégrant le chargement des arcs et l'écoulement du trafic. Leur modèle a été adapté par Varia et Dhingra [VS04] à une répartition logit du volume O-D entre les itinéraires concurrents.

Le modèle METROPOLIS de De Palma *et al.* [DPMN96] incorpore aussi un choix logit entre itinéraires, et y ajoute un choix d'horaire traité également par un modèle logit. Ce modèle de demande permet une formulation duale en fonction des temps des arcs, et l'algorithme est une méthode des combinaisons convexes sur les variables de temps. Dans METROPOLIS, le chargement et l'écoulement sont traités de manière intégrée par une simulation mésoscopique. C'est la même approche que Tong et Wong, et on peut la rapprocher

du modèle DynaMeq proposé plus récemment par Florian *et al.* [MMN05], où le chargement et l'écoulement sont traités de manière intégrée par simulation microscopique décentralisée par arcs. En 2005 ce modèle paraissait le simulateur commercial le plus efficace pour l'affectation dynamique.

Leurent a proposé le modèle Ladta [Leu03] : équilibre avec files d'attentes, choix d'itinéraire et éventuellement choix d'horaire, avec éventuellement plusieurs classes d'usagers, distribution des arbitrages entre prix et temps, élasticité du volume de demande au coût généralisé. Le modèle est formulé comme un problème de point fixe en fonction des volumes cumulés par arcs. De plus l'auteur a donné une expression générique qui facilite l'analyse et la modélisation : notamment la formulation de la formation des services et du chargement en volume ; et des algorithmes efficaces pour ces deux problèmes, le choix des usagers et le modèle de goulot. Le modèle d'écoulement a été étendu aux files d'attente horizontales avec plusieurs classes de trafic [Leu04] ; il a servi à formuler les coûts marginaux sociaux de congestion [Leu05].

Parallèlement, Bellei *et al.* [BGMP05] ont développé un traitement similaire, pour un modèle d'équilibre avec files d'attente et répartition logit du volume O-D entre itinéraires. Gentile *et al.* [GMP05] ont aussi considéré les files d'attente horizontales pour une classe de véhicules, avec une relation temps-débit pour les régimes hors saturation. Dernièrement ces auteurs ont traité la répartition modale entre modes individuels et modes collectifs.

Lin et Heydecker [LH05] ont exprimé l'équilibre dynamique avec un choix d'itinéraire et d'horaire selon un modèle logit, et un modèle de congestion par file d'attente et relation temps-débit en régime non saturé. Pour une destination, ils ont donné deux formulations extrémales de l'équilibre en fonction des débits des arcs, tout en proposant une résolution par combinaison convexe. Cependant ces auteurs n'ont pas indiqué comment traiter conjointement plusieurs destinations.

Friesz et Mookherjee [FM06] ont formulé un modèle d'équilibre avec congestion par file d'attente et relation temps-débit, comme une inéquation variationnelle en fonction des flux par chemins. Ils ont proposé un algorithme de point fixe pour la résoudre.

7.3 Discussion

De cette revue nous retenons les éléments suivants :

- Que la modélisation du choix des usagers concerne le choix d'itinéraire, souvent aussi le choix d'horaire, parfois le choix de se déplacer (élasticité du volume de demande au niveau de service).
- Que la modélisation macroscopique de la congestion incorpore désormais les files d'attente, parfois complétées par les relations temps-débit pour les régimes non saturés. Les modèles quasi-statiques ont été abandonnés depuis 2002.
- Que les formulations ont pour variables endogènes basiques, souvent les flux par arcs et destinations, assez souvent les flux par chemins, parfois les volumes cumulés par arcs, exceptionnellement les temps des arcs.
- Que désormais les problèmes types usuels sont l'inéquation variationnelle et le problème de point fixe. Les formulations extrémales initiales étaient défectueuses ; celles de Lin et Heydecker [LH05] paraît prometteuse mais nécessite d'être complétée pour des destinations multiples.
- Que les algorithmes d'équilibrage se réduisent à la méthode des combinaisons convexes, qui porte le plus souvent sur les flux par arcs et destinations, assez souvent sur les flux par chemins, parfois sur les volumes cumulés par arcs, exceptionnellement sur les temps par arcs.
- Que dans les communications scientifiques, le principal critère d'écart à l'équilibre est l'écart aux coûts minimum ; plus rarement, l'écart inter-itération en débits des arcs ; exceptionnellement, l'écart inter-itération en temps des arcs.
- Que seuls les modèles METROPOLIS et DYNAMEQ ont été appliqués à des réseaux de grande dimension. Les autres applications rapportées sont de simples cas d'école à quelques nœuds et arcs, une ou quelques relations O-D, ce qui ne permet pas d'apprécier concrètement le comportement de convergence des algorithmes d'équilibrage.

Chapitre 8

Formulations analytiques de l'équilibre dynamique et mesurage de la convergence

Introduction

Après l'examen circonstancié de l'affectation statique du chapitre 6, et la revue de certains modèles dynamiques du chapitre 7, nous allons examiner plus en profondeur l'affectation dynamique afin d'en révéler les complexités spécifiques.

La section 8.1 revient à la formulation de l'équilibre dans le modèle LADTA, exprimée à un niveau d'abstraction suffisant pour représenter largement l'affectation dynamique à l'équilibre offre-demande. Nous mettons en évidence les aspects spécifiques du cas dynamique : la relativité chronologique des volumes des arcs ; la correspondance des temps et des volumes, ainsi que les cas de concordance. Nous tirons des conséquences concernant la calculabilité de certains critères ; ou plus précisément, de certains termes qui entrent dans la composition des critères.

Dans la section 8.2, nous investiguons systématiquement diverses formulations pour l'affectation dynamique. Nous reprenons l'expression générique par chemins pour discuter l'aptitude de chaque jeu de variables endogènes,

volumes ou temps par arcs ou par chemins, à servir de variables endogènes basiques et donc à engendrer les autres variables endogènes par des relations imposées, tandis que les relations relaxées sont incorporées au critère de convergence. Puis nous donnons deux expressions précises et une expression générique par arcs et destinations, et nous examinons là aussi les critères de convergence et les approches algorithmiques : en particulier des approches primales-duales paraissent prometteuses. Ensuite nous revenons aux formulations par arcs pour montrer que le volume par arcs reste lié à un volume par chemins sous-jacents (cohérence temporelle), ce que nous exploitons pour obtenir des critères de convergence à la fois rigoureux et relativement économes.

La section 8.3 traite des algorithmes d'équilibrage pour l'affectation dynamique. Nous évoquons brièvement les algorithmes duaux et primaux-duaux, ainsi que les algorithmes primaux autres que les méthodes de combinaison convexe, pour en constater l'état très faible de développement. Surtout, nous discutons plusieurs méthodes de combinaison convexe : au niveau respectivement des arcs ou des chemins, avec séparation ou intégration du chargement en volume et de l'écoulement du trafic. Nous indiquons des critères de convergence rigoureux, et nous proposons un algorithme hybride qui traite conjointement des volumes et des temps par arcs.

Nous critiquons le modèle LADTA à la section 8.4, considérant successivement son expression et sa formulation. Pour faciliter des traitements algorithmiques qui explicitent les chemins, nous représentons les chemins par des faisceaux d'atomes, et nous développons des algorithmes "atomiques" pour le chargement des arcs en volume et pour l'actualisation d'un volume par chemins.

8.1 Relativité chronologique et correspondance temps-volume

Nous rappelons d'abord, en 8.1.1, les notations, puis, en 8.1.2, la structuration du problème. Nous discutons ensuite la démarche d'équilibrage du

modèle LADTA (8.1.3). Puis nous mettons en évidence que le chargement du volume par chemins aux arcs du réseau, est relatif à une chronologie (8.1.4). Par conséquent, la correspondance entre les temps et les flux dépend aussi de la chronologie (8.1.5) ; une éventuelle concordance entre les flux et les temps des chemins nécessite un algorithme de chargement spécialement adapté : celui de Leurent [Leu03] ne convient pas (8.1.6). Comme une telle concordance est requise pour des critères d'écart au coût minimum, nous la recherchons dans plusieurs circonstances : concordance "duale" entre les temps et les flux, qui induit une concordance "primale" pour un flot auxiliaire (8.1.7). De manière connexe, nous définissons une propriété de concordance des énergies et nous établissons des conditions suffisantes (8.1.8). Nous terminons en récapitulant les effets détectés et les résultats établis (8.1.9).

8.1.1 Notations et variables pour l'affectation dynamique

Rappelons d'abord les éléments du modèle d'affectation dynamique LADTA de [Leu03], dans une version basique qui considère :

- du côté de la demande, une seule classe d'usagers, dont le choix d'itinéraire suit un comportement homogène et déterministe, avec un coût généralisé réduit au temps de parcours, et sans élasticité du volume au niveau de service.
- Du côté de l'offre, des itinéraires sans restriction d'accès aux arcs autre que la condition de non-multiplicité (pour l'occurrence d'un même arc dans un chemin), et des fonctions de congestion par délai minimal de propagation et file d'attente verticale en sortie d'arc.

Voici les notations correspondantes :

$G = (N, A)$	réseau de transport, avec N l'ensemble des nœuds n et A l'ensemble des arcs $a = (n, m)$ à sommets dans N.
A_n^+ (resp. A_n^-)	l'ensemble des arcs sortants (resp. entrants) du nœud n.

h	horaire dans une période H.
$X_a^+(h)$	volume entré dans l'arc a, cumulé jusqu'à h.
$X_a^-(\hbar)$	volume sorti de l'arc a cumulé jusqu'à l'instant de sortie $\hbar$.
$t_a(h)$	temps effectif de traversée de l'arc a en y entrant à h.
$H_a(h) = h + t_a(h)$	fonction de transfert chronologique, qui donne l'horaire de sortie de a sachant l'entrée à h.
$\overline{H}_a(\hbar)$	fonction réciproque de H_a, associe à un instant de sortie $\hbar$ l'horaire d'entrée correspondant.
S	ensemble des sommets de destination s.
O_s	ensemble des sommets d'origine o pour la destination s.
$I = \bigcup_{s \in S} O_s$	ensemble des relations O-D $i = (o, s)$.
R_i	ensemble des chemins sans multiplicité r sur la relation O-D i.
$Q_i(h)$	volume O-D de la relation i cumulé jusqu'à l'instant de départ h.
$X_r^+(h)$	volume entré dans le chemin r cumulé jusqu'à h.
$X_r^-(\hbar)$	volume sorti du chemin r cumulé jusqu'à l'instant de sortie $\hbar$.
$t_r(h)$	temps effectif de traversée du chemin r en y entrant à h.

$H_r(h) = h +$ horaire de sortie de r sachant l'entrée à h. Fonction de transfert chronologique d'un chemin. $t_r(h)$

$\overline{H}_r(\hbar)$ fonction réciproque de H_r, associe à un instant de sortie $\hbar$ l'horaire d'entrée correspondant.

$r > a$ pour un chemin r et un arc a qu'il traverse, $r > a$ est la partie de r strictement en amont de a.

$r < a$ pour un chemin r et un arc a qu'il traverse, $r < a$ est la partie de r strictement en aval de a.

Les variables endogènes du modèle sont les volumes entrants et sortants par chemin (X_r^+ et X_r^-) et par arc (X_a^+ et X_a^-), que nous appelopns des variables primales. Ainsi que les temps de traversée par chemin (t_r) et par arc (t_a) que nous appelons des variables duales puisque nous assimilons la durée de déplacement au coût généralisé subi par l'usager.

Les volumes et les temps sont couplés par divers effets :

- le couplage physique de la congestion : au niveau local d'un arc, le volume induit le temps de traversée.
- le couplage économique des comportements : chaque usager choisit son chemin en fonction du coût, donc du temps.
- la propagation en réseau.

Dans le cas dynamique, la propagation en réseau, combinée à l'aspect chronologique, induit une complexité supplémentaire par rapport au cas statique.

8.1.2 Structuration de l'affectation dynamique

Dans [Leu03], l'affectation dynamique a été décomposée en quatre problèmes.

Le problème de *Formation des Services* F_S

$$t_R = F_S(t_A) \tag{8.1a}$$

de *Choix des Usagers* F_U

$$X_R^+ \in F_U\left(t_R\right) \tag{8.1b}$$

de *Chargement en Volume* F_V

$$X_A^+ = F_V\left(X_R^+, t_R\right) \tag{8.1c}$$

et d'*Ecoulement du Trafic* F_F (pour *Flowing*)

$$t_A \in F_F\left(X_A^+\right) \tag{8.1d}$$

Un état $\left(X_R^+, X_A^+, t_A, t_R\right)$ qui satisfait le système d'équations (8.1) est un équilibre dynamique du trafic.

A ce niveau d'abstraction, la formulation est générique pour des modèles variés d'affectation dynamique macroscopique : notamment Akamatsu et Kuwahara [KA97, KA01], Tong et Wong [TW00] ; Friesz *et al.* [FBST01] etc. Au plan physique, elle englobe la plupart des modèles d'écoulement. Au plan économique, elle englobe l'optimum de l'usager comme l'optimum du système ; diverses sensibilités au temps, car on pourrait négliger dans F_U l'influence de t_R et considérer des temps exogènes.

8.1.3 Sur la démarche d'équilibrage dans le modèle LADTA

L'approche originelle de l'équilibrage dans le modèle LADTA, conçue indépendamment par Leurent [Leu03] et Gentile [GMP05], consiste à formuler l'équilibre comme un problème de point fixe en fonction des X_a^+ qui sont les variables endogènes basiques, et à approcher l'équilibre par une méthode des combinaisons convexes sur ces variables.

La formulation en problème de point fixe est obtenue en remplaçant les autres variables endogènes par leur expression en fonction de X_A^+ :

$$t_A \in F_F\left(X_A^+\right) \tag{8.2a}$$

$$t_R \in F_S \circ F_F \left(X_A^+ \right) \tag{8.2b}$$

$$X_R^+ \in F_U \circ F_S \circ F_F \left(X_A^+ \right) \tag{8.2c}$$

$$X_A^+ \in F_V \left(F_U \left(\bullet \right), \bullet \right) \circ F_S \circ F_F \left(X_A^+ \right) \tag{8.2d}$$

La dernière condition est justement une condition de point fixe, pour l'application

$$F_V \left(F_U \left(\bullet \right), \bullet \right) \circ F_S \circ F_F$$

L'algorithme d'équilibrage est très économe en information, donc en espace de mémoire : une combinaison convexe X_A^+ est la simple transposition dynamique de la combinaison convexe sur x_A dans le cas statique.

Le critère de convergence proposé jusqu'à présent est un écart inter-itération sur les volumes des arcs, de la forme

$$\left\| X_A^{+(k+1)} - X_A^{+(k)} \right\|^2 = \frac{1}{|A|} \sum_{a \in A} \int \left| X_A^{+(k+1)} \left(h \right) - X_A^{+(k)} \left(h \right) \right|^2 dh$$

Il présente le risque inhérent à la méthode des combinaisons convexes, de tendre vers 0 simplement parce que les coefficients de combinaison ξ_k tendent vers 0. Il serait préférable de disposer d'un critère de convergence plus rigoureux, ce qui motive notre examen des formulations et des algorithmes des sections 8.1.5 et 8.1.6.

8.1.4 La relativité chronologique du chargement en volume des arcs

Dans l'affectation statique, les flux des arcs et des chemins sont liés par une condition simple de totalisation :

$$x_a = \sum_{r \in a} x_r$$

Dans l'affectation dynamique, la composition du volume des arcs à partir des volumes des chemins fait l'objet du problème de chargement en volume, $X_A^+ = F_V\left(X_R^+, t_R\right)$.

Précisément, voici l'équation de chargement de [Leu03] :

$$X_a^+\Big|_\eta^h = \sum_{r \in a} X_r^+ \circ \overline{H}_{r>a}\Big|_\eta^h \tag{8.3}$$

Elle exprime que le volume entrant dans l'arc a entre η et h est la somme des volumes des chemins r qui traversent l'arc, et que le volume du chemin r qui entre dans a entre η et h est entré dans r pendant l'intervalle $\left[\underline{h}; \overline{h}\right]$ tel que $\underline{h} + t_{r>a}\left(\underline{h}\right) = \eta$ et $\overline{h} + t_{r>a}\left(\overline{h}\right) = h$, soit

$$\underline{h} = \overline{H}_{r>a}\left(\eta\right)$$

$$\overline{h} = \overline{H}_{r>a}\left(h\right)$$

Ainsi les fonctions t_R interviennent dans le chargement en volume des arcs par l'intermédiaire des fonctions de transfert temporel H_R et leurs réciproques $\overline{H}_R$. Nous appelons cette influence, *la relativité chronologique du chargement*, et nous appelons les H_R *les fonctions de chronologie des chemins* associées aux temps t_R.

8.1.5 La correspondance temps-volume est relative à une chronologie

Dans l'affectation statique, la correspondance entre temps et flux au niveau d'un arc est très simple, avec une fonction de congestion t_a :

$$t_a = \mathrm{t}_a\left(x_a\right)$$

Au niveau d'un chemin la correspondance reste simple :

$$t_r = \sum_{a \in r} t_a \text{ avec } x_a = \sum_{r \in a} x_r$$

Dans l'affectation dynamique, la correspondance entre temps et volume au niveau d'un arc est relativement simple, par application d'un modèle d'écoulement local. Les $[t_a(h), h \in H]$ résultent des $[X_a^+(h), h \in H]$ qui sont confrontés aux conditions d'écoulement local (délai minimal $t_{a0}(h)$, débit de capacité $K_a(h)$). A ce stade nous évitons de considérer les interactions entre les volumes d'arcs voisins). Mais au niveau des chemins, la fonction de correspondance temps-volume a la forme

$$\begin{aligned} t_R &= F_S(t_A) \\ &\in F_S \circ F_F\left(X_A^+\right) \\ &\in F_S \circ F_F \circ F_V\left(X_R^+, \tilde{t}_R\right) \end{aligned} \tag{8.4}$$

Elle dépend d'une chronologie $\tilde{H}_R = \left[h + \tilde{t}_r(h), h \in H\right]_{r \in R}$ qui détermine le chargement.

8.1.6 Sur la concordance temps-volume

8.1.6.1 Définition de la concordance temps-volume

La concordance temps-volume est la condition d'identité entre t_R et $\tilde{t}_R$ dans l'application temps-volume des chemins (8.4). C'est une condition de point fixe pour t_R relativement à X_R^+ :

$$t_R \in F_S \circ F_F \circ F_V\left(X_R^+, t_R\right) \tag{8.5}$$

Cette concordance est nécessaire pour l'équilibre dynamique du trafic. Elle n'en est pas une condition suffisante, car elle indique seulement que X_R^+ et t_R sont en couplage physique de congestion. Elle n'indique rien concernant un éventuel couplage économique de X_R^+ et t_R du fait du choix des usagers.

La conjonction des deux couplages physique et économique, (8.5) et $X_R^+ \in F_U(t_R)$, définit un équilibre dynamique du trafic.

8.1.6.2 Propriétés de la concordance

Quand la concordance est vérifiée, pour mesurer l'écart à l'équilibre il suffit de mesurer l'écart entre X_R^+ et $F_U(t_R)$. Pour cela, l'écart aux coûts minimaux convient :

$$\sum_{r\in R}\int_{h\in H}\left[t_r(h)-u_{i(r)}(h)\right]dX_r^+(h) \tag{8.6}$$

avec $u_i(h)=\min_{\rho\in i}t_\rho(h)$

De plus, pour un couple par chemins $\left(X_R^+,t_R\right)$ en concordance temps-volume et le couple par arcs $\left(X_A^+,t_A\right)$ qui lui correspond, on vérifie la propriété de *concordance des énergies*

$$\sum_{r\in R}\int_{h\in H}t_r(h)\,dX_r^+(h)=\sum_{a\in A}\int_{h\in H}t_a(h)\,dX_a^+(h)$$

En effet, X_A^+ découle de X_R^+ selon la chronologie associée à t_R, et t_R est formé à partir de t_A. Ces conditions suffisent à établir la concordance des énergies par arcs et par chemins, indépendamment de la liaison entre t_A et X_A^+ via le modèle d'écoulement.

8.1.6.3 Imposer la concordance

Pour obtenir la concordance temps-volume, l'approche "naturelle" suivie notamment par Tong et Wong [TW00] est d'affecter progressivement, instant après instant, les volumes X_R^+ sur le réseau, en déterminant simultanément X_A^+ et t_A. Autrement dit, c'est un traitement intégré du chargement en volume et de l'écoulement du trafic, comme dans une simulation microscopique. Tong et Wong ont proposé une simulation mésoscopique qui fonctionne comme une simulation microscopique mais par "paquets" de véhicules, pour faire le lien avec le niveau macroscopique.

Un tel algorithme présente certaines complexités :

– en temps de calcul, l'ordre de complexité est le produit du nombre de paquets et du nombre d'instants de présence pour chaque paquet. Le nombre de paquets est proportionnel au nombre de chemins et d'ins-

tants de départ. Le nombre moyen d'instants de présence dépend du pas horaire choisi et de la durée moyenne de déplacement.

- en espace de mémoire informatique, la complexité dépend du nombre d'arc des chemins, et du nombre de paquets présents en moyenne. Chaque paquet renvoie à son chemin et à une position le long du chemin, donc c'est la somme des deux nombres qui intervient et non le produit.

De plus la discrétisation par paquets et par périodes élémentaires induit des approximations, dont l'effet global est une imprécision d'autant plus forte que le problème d'affectation est de plus grande dimension.

Ces coûts de complexité et d'imprécision sont à mettre en balance avec le grand avantage de la concordance temps-volume : la disposition du critère d'écart aux coûts minimum pour mesurer la convergence.

8.1.7 Sur la concordance "duale" volume-temps

En partant non pas des volumes, mais des temps, la chaîne de causalité est : t_A, puis

$$t_R = F_S\left(t_A\right) \tag{8.7a}$$

$$Y_R^+ \in F_U \circ F_S\left(t_A\right) \tag{8.7b}$$

$$Y_A^+ = F_V\left(Y_R^+, t_R\right) \tag{8.7c}$$

$$\tilde{t}_A \in F_F\left(Y_A^+\right) \tag{8.7d}$$

Il reste à mettre en relation t_A et $\tilde{t}_A$. mais nous laissons là ce sujet, qui revient à concevoir un algorithme d'équilibrage dual pour l'affectation dynamique, afin de nous concentrer sur une propriété des volumes et temps ainsi constitués. Comme t_R est formé à partir de t_A, et Y_A^+ découle de Y_R^+ selon la chronologie associée à t_R, les conditions d'application nécessaires à

la concordance des énergies par arcs et par chemins sont réunies :

$$\sum_{r\in R}\int_{h\in H} t_r(h)\,dY_r^+(h) = \sum_{a\in A}\int_{h\in H} t_a(h)\,dY_a^+(h)$$

Quand Y_R^+ résulte d'une affectation aux chemins de temps minimum, cette égalité permet d'évaluer le coût minimal global sur la base des temps t_R :

$$\sum_{i\in I}\int_{h\in H} u_i(h)\,dQ_i(h) \text{ avec } u_i(h) = \min_{r\in I} t_r(h)$$

En effet, par construction de Y_R^+,

$$u_i(h)\,dQ_i(h) = \sum_{r\in I} t_r(h)\,dY_r^+(h)$$

et donc

$$\begin{array}{rcl} \sum_{i\in I}\int_{h\in H} u_i(h)\,dQ_i(h) & = & \sum_{r\in R}\int_{h\in H} t_r(h)\,dY_r^+(h) \\ & = & \sum_{a\in A}\int_{h\in H} t_a(h)\,dY_a^+(h) \end{array} \quad (8.8)$$

8.1.8 La concordance des énergies par arcs et par chemins

Dans l'affectation statique, on vérifie l'égalité des puissances entre les arcs et les chemins :

$$\sum_{r\in R} t_r \cdot x_r = \sum_{a\in A} t_a \cdot x_a$$

sous la condition $t_r = \sum_{a\in r} t_a$ et $x_a = \sum_{r\in a} x_r$.

La transposition dynamique de cette concordance nécessite les deux conditions suivantes : d'une part, que t_R soit formé à partir de t_A (par la formation des services) ; et d'autre part, que X_A^+ découle de X_R^+ selon la chronologie associée à t_R.

Sous ces conditions :

$$t_r(h) = \sum_{a\in r} t_a(H_{r>a}(h))$$

$$\int_{\underline{h}}^{\overline{h}} t_r(h)\, dX_r^+(h) = \sum_{a\in r} \int_{\underline{h}}^{\overline{h}} t_a\left(H_{r>a}(h)\right) dX_r^+(h)$$

$$= \sum_{a\in r} \int_{\overline{H}_{r>a}(\underline{h})}^{\overline{H}_{r>a}(\overline{h})} t_a(h)\, dX_r^+\left(\overline{H}_{r>a}(h)\right)$$

Donc :

$$\sum_{r\in R} \int_{h\in H} t_r(h)\, dX_r^+(h) = \sum_{a\in A} \int_{h\in H} t_a(h) \left[\sum_{r\in a} dX_r^+\left(\overline{H}_{r>a}(h)\right)\right]$$

$$= \sum_{a\in A} \int_{h\in H} t_a(h)\, dX_a^+(h) \tag{8.9}$$

en raison du chargement en volume $X_A^+ = F_V\left(X_R^+, t_R\right)$, et à la condition que l'intégration dans le temps couvre l'ensemble des instants où circule du flux, en chaque endroit (arc ou chemin).

8.1.9 Récapitulation

Nous avons mis en évidence :

- La relativité chronologique du chargement des arcs en volume.
- Que la correspondance temps-volume, au niveau des chemins, est relative à une chronologie.
- Que la correspondance temps-volume débouche toujours sur une concordance dans l'affectation statique, mais pas dans l'affectation dynamique.
- Que la concordance temps-volume i.e. $t_R = F_S \circ F_F \circ F_V\left(X_R^+, t_R\right)$ est une condition nécessaire de l'équilibre dynamique.
- Que la concordance peut être imposée en opérant conjointement le chargement en volume et l'écoulement du trafic : approche de Tong et Wong [TW00].
- Qu'entre des temps (T_R, T_A) déduits les uns des autres, et des volumes $\left(X_A^+, X_R^+\right)$ déduits les uns des autres selon la chronologie associée à T_R, les énergies par arcs et par chemins coïncident, concordent.

– Que cette concordance des énergies est vérifiée notamment en cas de concordance temps-volume, et aussi pour des volumes Y_R^+, Y_A^+ respectivement affectés et chargés uniquement à partir de temps T_R, T_A déduits les uns des autres.

8.2 Sur la formulation de l'équilibre dynamique

Nous avons constaté que la formulation du modèle LADTA comme un problème de point fixe en fonction des volumes des arcs, permet un traitement algorithmique économe, mais dénué de concordance temps-volume et dont la convergence n'est pas fermement établie. Tandis que la formulation par chemins avec intégration du chargement et de l'écoulement, due à Tong et Wong [TW00], paraît plus robuste puisque la concordance temps-volume est assurée et que le critère de convergence est rigoureux, mais nécessite un traitement beaucoup plus coûteux.

Ici notre objectif est de concevoir des formulations pour l'équilibre dynamique, qui soient robustes, exemptes de risque et relativement peu coûteuses.

Pour ce faire, nous allons investiguer systématiquement des formulations, en nous concentrant sur l'expression d'un modèle, sur la sélection de variables endogènes basiques, sur la formulation d'un critère de convergence. Ainsi notre approche de la formulation se situe en amont de la formulation caractéristique en problème-type de recherche opérationnelle, comme de l'analyse des propriétés formelles des fonctions impliquées : nous privilégions l'expression de l'équilibre en fonction de certaines variables endogènes, et pour une telle expression nous privilégions autant que possible un niveau générique afin de nous abstraire des particularités formelles.

Notre plan d'investigation se compose de trois parties. En section 8.2.1 nous reprenons la formulation par chemins, en envisageant diverses relaxations des conditions d'équilibre et en examinant l'aptitude des divers jeux de variables endogènes à servir de base endogène. Nous donnons des critères rigoureux de convergence, qui nécessitent d'expliciter les chemins.

En 8.2.2 nous exprimons une formulation par arcs et par destinations pour l'affectation dynamique, d'abord de manière précise, puis de manière

générique. En reprenant les éléments d'analyse du chapitre précédent, nous montrons que la relativité chronologique concerne d'une part la condition de conservation du flux, et d'autre part la fonction de choix des usagers. La concordance des énergies est une propriété triviale de la formulation par arcs et destinations ; qui jouit aussi de propriétés plus utiles concernant l'écart de complémentarité. Nous proposons des critères de convergence et des approches algorithmiques, et nous présentons deux formulations prometteuses : l'une primale-duale en fonction des volumes par arcs et destinations et aussi des coûts par arcs, l'autre primale-duale également en fonction aussi des potentiels par destination.

Enfin, en 8.2.3, nous établissons une méthode pour assurer la cohérence temporelle dans une formulation en fonction des volumes des arcs, à partir des propriétés de la formulation par arcs et destinations.

8.2.1 Sur la formulation par chemins

8.2.1.1 Formulation relaxée de l'équilibre dynamique

En 8.1.2 nous avons donné une expression générique pour l'équilibre dynamique du trafic, dans le système d'équations (8.1). Dans cette expression, nous pouvons considérer que les variables endogènes $\left(X_R^+, X_A^+, t_A, t_R\right)$ sont toutes basiques, à déterminer conjointement. Cela revient à relaxer les relations fonctionnelles directes entre certaines variables, notamment la formation des services et l'écoulement du trafic. Pour exprimer cela, nous remplaçons chaque symbole d'égalité ou d'appartenance dans (8.1) par le symbole $\leftrightarrow$ qui signifie *rapprochement à opérer* :

$$t_R \leftrightarrow F_S\left(t_A\right) \tag{8.10a}$$

$$X_R^+ \leftrightarrow F_U\left(t_R\right) \tag{8.10b}$$

$$X_A^+ \leftrightarrow F_V\left(X_R^+, t_R\right) \tag{8.10c}$$

$$t_A \leftrightarrow F_F\left(X_A^+\right) \tag{8.10d}$$

8.2.1.2 Critère de convergence

Pour mesurer la proximité à l'équilibre, il suffit d'évaluer comment sont remplies les conditions d'équilibre, en mesurant un écart global que l'on peut formuler comme une somme d'écarts sur chaque condition :

$$\begin{aligned} &d_S\left[t_R, F_S\left(t_A\right)\right] + d_U\left[X_R^+, F_U\left(t_R\right)\right] \\ &+d_V\left[X_A^+, F_V\left(X_R^+, t_R\right)\right] + d_F\left[t_A, F_F\left(X_A^+\right)\right] \end{aligned} \tag{8.11}$$

où la notation d_I indique une distance, et l'indice I dans $\{S, U, V, F\}$ marque la condition d'équilibre concernée.

Le critère global (8.11) est rigoureux. Certains termes s'annulent quand on impose des conditions entre les variables endogènes. Cependant, l'expression générale fait apparaître les volumes par chemins et les temps par chemins, et a priori l'évaluation du critère constitue un traitement informatique coûteux.

Plus difficile est la question de comment approcher un équilibre, par quelle adaptation progressive des variables endogènes. La stratégie usuelle est de choisir une partie des variables endogènes, partie alors qualifiée de variables basiques, pour d'une part en déduire les autres variables endogènes au travers de contraintes *de définition* ; et d'autre part, faire évoluer au cours d'itérations successives les valeurs des variables basiques.

Afin de déduire les variables endogènes les unes des autres, on peut ainsi imposer l'une ou l'autre des conditions suivantes, de manière isolée :

$$
\begin{array}{ll}
t_R = F_S\left(t_A\right) & \text{donc } d_S\left[t_R, F_S\left(t_A\right)\right] = 0 \\
X_R^+ \in F_U\left(t_R\right) & \text{donc } d_U\left[X_R^+, F_U\left(t_R\right)\right] = 0 \\
X_A^+ = F_V\left(X_R^+, t_R\right) & \text{donc } d_V\left[X_A^+, F_V\left(X_R^+, t_R\right)\right] = 0 \\
t_A \in F_F\left(X_A^+\right) & \text{donc } d_F\left[t_A, F_F\left(X_A^+\right)\right] = 0
\end{array}
$$

Mais il n'y a qu'en un état d'équilibre que l'on arrive à imposer conjointement toutes ces conditions ; dans une approche progressive vers l'équilibre, on choisit certaines conditions qui sont compatibles mutuellement, et donc une base de variables endogènes. A contrario d'un tel mouvement de réduction, certaines démarches de résolution mettent en oeuvre des variables auxiliaires qui augmentent le jeu de variables endogènes. Nous nous contentons ci-après de considérer les variables $\left(X_R^+, X_A^+, t_A, t_R\right)$ et d'éventuelles simplifications, telles que des variables par arcs et destinations pour simplifier la description par chemins.

8.2.1.3 Choisir une base de variables endogènes

Concernant les *volumes par chemins*, nous avons constaté que les variables endogènes X_R^+ peuvent servir de base pour l'affectation dynamique, en en déduisant les autres variables endogènes par une fonction intégrée de chargement et écoulement, et par la fonction de formation des services :

$$
\begin{array}{rcl}
\left(X_A^+, t_A\right) & \in & F_{FV}\left(X_R^+\right) \\
t_R & = & F_S\left(t_A\right) \\
X_R^+ & \in & F_U\left(t_R\right)
\end{array}
\tag{8.12}
$$

Nous avons constaté le coût d'expliciter les volumes des chemins, et plus encore de traiter conjointement le chargement et l'écoulement.

Une simplification potentielle est de traiter plutôt les volumes par arcs et destinations, notés X_A^{S+}. Nous y revenons au 8.2.2. Une autre simplification

est de séparer le chargement et l'écoulement. Nous y revenons au 8.3 avec un algorithme de combinaison convexe au niveau des chemins, qui incorpore cette séparation.

Concernant les *volumes par arcs* X_A^+, ces variables endogènes peuvent servir de base pour l'affectation dynamique en en déduisant les autres variables de la manière suivante :

$$\begin{array}{rcl} t_A & \in & F_F\left(X_A^+\right) \\ t_R & = & F_S\left(t_A\right) \\ X_R^+ & \in & F_U\left(t_R\right) \\ X_A^+ & = & F_V\left(X_R^+, t_R\right) \end{array} \tag{8.13}$$

Une telle expression est très économe si l'on évite d'expliciter les volumes des chemins. La configuration la plus favorable est un modèle de choix des usagers tel que l'application F_U est univoque, auquel cas X_A^+ permet d'engendrer X_R^+. En particulier, les modèles logit de choix discret d'itinéraire et d'horaire de départ, de Bellei, Gentile *et al.* [BGP05] possèdent cette propriété. Dans le modèle LADTA originel de [Leu03], la base endogène X_A^+ fait la synthèse de variables auxiliaires Y_A^+ , qui elles-mêmes découlent de volumes par chemins Y_R^+ solution d'un problème auxiliaire d'affectation pour les temps issus de X_A^+. C'est le traitement algorithmique qui évite d'expliciter les volumes des chemins [1].

Mais nous avons identifié un risque inhérent à ce traitement : que le volume par arcs X_A^+ obtenu par les combinaisons convexes successives ne soit plus lié à un volume par chemins $\tilde{X}_R^+$ et une chronologie associée à des temps $\tilde{t}_R$, en ce sens qu'il n'existe pas de couple $\left(\tilde{X}_R^+, \tilde{t}_R\right)$ solution de

$$X_A^+ = F_V\left(\tilde{X}_R^+, \tilde{t}_R\right) \tag{8.14}$$

et ce, indépendamment de tout lien entre $\tilde{X}_R^+$ et $\tilde{t}_R$.

Comme la combinaison convexe préserve au moins les volumes cumulés

[1] plus précisément : qui permet de n'expliciter que les Y_R^+, sans les conserver pour la suite, et sachant que ce sont des volumes par chemins particulièrement simples puisque relatifs à un certain sous-ensemble de chemins, le sous-ensemble des chemins optimaux pour les coûts fixés.

sur l'ensemble de la période d'affectation, ce risque tient plus aux temps qu'aux volumes, et nous l'appelons le risque *d'incohérence temporelle* pour les volumes des arcs.

Concernant les *coûts ou temps par arcs* ν_A , ces variables endogènes peuvent servir de base pour l'affectation dynamique en en déduisant les autres variables de la manière suivante :

$$\begin{array}{rcl} t_R & = & F_S\left(\nu_A\right) \\ X_R^+ & \in & F_U\left(t_R\right) \\ X_A^+ & = & F_V\left(X_R^+, t_R\right) \\ \nu_A & \in & F_F\left(X_A^+\right) \end{array} \tag{8.15}$$

Le risque principal concerne la faisabilité primale, i.e. l'obtention d'un vecteur X_R^+ ou X_A^{S+} qui vérifie effectivement les contraintes. Lorsque le modèle de comportement des usagers, F_U , associe une solution unique à un t_R donné, la faisabilité primale est assurée. C'est la configuration privilégiée pour une approche duale, de manière très similaire à la base endogène X_A^+.

Concernant les *temps ou coûts par chemins* ν_R , ces variables endogènes peuvent difficilement servir de base endogène pour l'affectation dynamique. Cela nécessiterait trop de contraintes d'ordre topologique entre les divers chemins selon leurs interactions et imbrications variées. Une solution de simplicité est de s'en tenir aux variables ν_A.

Une autre solution est de considérer aussi des variables duales de potentiel, $u_N^S = \left[u_N^S\right]_{n\in N, s\in S}$, qui représentent le coût entre un sommet courant n et une destination s. On exprimerait alors le coût d'un chemin, ou plutôt d'un cheminement, comme le coût de l'arc initial plus le coût depuis le sommet final de cet arc jusqu'à la destination. Alors on pourrait déduire les autres variables de la base $\left(\nu_A, u_N^S\right)$ de la manière suivante :

$$\begin{array}{rcl} X_R^+ & \in & \tilde{F}_U\left(u_N^S, \nu_A\right) \\ X_A^+ & = & F_V\left(X_R^+, u_N^S, \nu_A\right) \\ t_A & \in & F_F\left(X_A^+\right) \end{array} \tag{8.16}$$

En résumé, les meilleurs jeux de variables endogènes basiques sont : d'une

part, des jeux simples $\left\{X_R^+\right\}$, $\left\{X_A^{S+}\right\}$, $\left\{X_A^+\right\}$, $\{\nu_A\}$; et d'autre part des couples $\left\{X_A^{S+}, \nu_A\right\}$ et $\left\{u_N^S, \nu_A\right\}$, qui conduisent à des formulations par arcs et par destinations, respectivement primale-duale ou complètement duale.

8.2.2 Formulations par arcs et destinations

Nous allons commencer par exprimer l'équilibre dynamique par arcs et par destinations de manière précise (8.2.2.1). Puis nous montrons que la relativité chronologique concerne d'une part la condition de conservation du flux, et d'autre part la fonction de choix des usagers (8.2.2.2) : cela nous amène à distinguer trois variables de coûts ou temps par arcs, à savoir les écarts de chronologie Θ_A, les durées de traversée t_A, et les coûts ou temps économiques ν_A (8.2.2.3). Nous donnons alors une expression générique (8.2.2.4). Ensuite nous examinons les propriétés de la formulation par arcs et destinations concernant la concordance des énergies, et l'écart aux coûts minimum (8.2.2.5). Après cela, nous proposons des critères de convergence (8.2.2.6) et des approches algorithmiques (8.2.2.7), et nous présentons deux formulations prometteuses : l'une primale-duale en fonction des volumes par arcs et destinations et aussi des coûts par arcs (8.2.2.8), l'autre primale-duale également en fonction aussi des potentiels par destination (8.2.2.9).

8.2.2.1 Expression précise par arcs et destinations

Nous caractérisons l'équilibre dynamique du trafic par le système de conditions suivant, en fonction de variables par arcs et par destinations :

$$\begin{gathered} dX_a^{s+}(h) \geq 0, \\ \forall s \in S, a \in A, h \in H \end{gathered} \tag{8.17a}$$

$$\begin{gathered} dX_a^{s+}(h) = dX_a^{s-}\left(h + t_a(h)\right), \\ \forall s \in S, a \in A, h \in H \end{gathered} \tag{8.17b}$$

$$\begin{gathered} \sum_{a \in A_n^+} dX_a^{s+}(h) = dQ_{ns}(h) - \sum_{a \in A_n^-} dX_a^{s-}(h), \\ \forall s \in S, n \in N, h \in H \end{gathered} \tag{8.17c}$$

$$dX_a^+ = \sum_{s \in S} dX_a^{s+}, \forall a \in A \tag{8.17d}$$

$$t_a = F_{F_a}\left(X_a^+\right), \forall a \in A \tag{8.17e}$$

$$\nu_a = T_a, \forall a \in A \tag{8.17f}$$

$$\begin{gathered} u_n^s(h) \leq \nu_a(h) + u_m^s(h + t_a(h)), \\ \forall s \in S, a \approx (n,m) \in A, h \in H \end{gathered} \tag{8.17g}$$

$$\begin{gathered} [\nu_a(h) + u_m^s(h + t_a(h)) - u_n^s(h)] \cdot dX_a^{s+}(h) = 0, \\ \forall s \in S, a \approx (n,m) \in A, h \in H \end{gathered} \tag{8.17h}$$

Il est convenu qu'un élément différentiel dX signifie la variation élémentaire de la variable X pendant une durée élémentaire dh.

Le système d'équations (8.17) exprime que les flux par arcs et destinations sont non-négatifs (8.17a), se propagent le long d'un arc selon un délai de propagation (8.17b) et respectent la conservation du flux en chaque nœud pour chaque destination (8.17c) ; que le flux total d'un arc se décompose par destinations (8.17d) et induit le temps de traversée (8.17e) ; que le coût par arc est réduit au temps de parcours (8.17f) et induit le coût vers chaque destination (8.17g et 8.17h). La condition (8.17h) est précisément une *condition des écarts complémentaires* entre le flux et les coûts par destination, et jointe à (8.17a) et (8.17g) elle impose que u_n^s soit le coût minimal de n à s sur la base des coûts des arcs ν_a.

La condition des écarts complémentaires, jointe à la conservation locale du flux, implique qu'un usager présent en n est également un usager partant de n à h. L'attente au nœud n n'est pas une option de déplacement, et l'usager s'engage immédiatement sur l'arc le plus avantageux à l'instant de sa présence, sans compter sur d'éventuelles améliorations ultérieures. Ainsi la traversée d'un nœud obéit à une logique de *First In, First Out* (FIFO), que l'on impose généralement aussi au niveau de chaque arc dans le modèle d'écoulement (8.17e).

8.2.2.2 Sur la relativité chronologique et les trois types de temps

Dans le système (8.17), la durée de traversée $t_a(h)$ intervient notamment dans la propagation du flux au niveau local d'un arc (8.17b), dans la conservation du flux au niveau local d'un nœud (8.17c), dans les conditions de préférence des usagers (8.17g, 8.17h). Nous retrouvons l'effet de relativité chronologique, qui ici concerne le niveau local des arcs et des nœuds et non le niveau des chemins.

Afin de considérer clairement cet effet de relativité chronologique, nous distinguons trois types de temps. Premièrement, le *temps chronologique*, autrement dit l'horaire. Deuxièmement, le *temps physique*, la durée, qui concerne le temps de traversée et le modèle d'écoulement. Troisièmement, le *temps économique* qui détermine le choix des usagers.

8.2.2.3 Expression précise distinguant quatre types de temps

Convenons de noter :

$\theta_A(h)$ un écart de chronologie entre les instants d'entrée h et de sortie $h + \theta_A(h)$ pour un usager donné.

$t_a(h)$ une durée de traversée de l'arc a en y entrant à h.

$\nu_a(h)$ un temps économique, dans un modèle de choix des usagers où le coût se réduit au temps.

$\tau_a(h)$ un délai de propagation entre les potentiels des sommets de l'arc a. Cette distinction est surtout formelle.

On peut alors reformuler l'équilibre dynamique du trafic de la manière suivante :

$$\begin{gathered} dX_a^{s+}(h) \geq 0, \\ \forall s \in S, a \in A, h \in H \end{gathered} \tag{8.18a}$$

$$\begin{gathered} dX_a^{s+}(h) = dX_a^{s-}(h + \theta_a(h)), \\ \forall s \in S, a \in A, h \in H \end{gathered} \tag{8.18b}$$

$$\sum_{a\in A_n^+} dX_a^{s+}(h) = dQ_{ns}(h) + \sum_{a\in A_n^-} dX_a^{s-}(h), \quad \forall s\in S, n\in N, h\in H \tag{8.18c}$$

$$dX_a^+ = \sum_{s\in S} dX_a^{s+}, \forall a\in A \tag{8.18d}$$

$$t_a = F_{F_a}\left(X_a^+\right), \forall a\in A \tag{8.18e}$$

$$\nu_a = t_a, \forall a\in A \tag{8.18f}$$

$$u_n^s(h) \le \nu_a(h) + u_m^s(h+\tau_a(h)), \quad \forall s\in S, a=(m,n)\in A, h\in H \tag{8.18g}$$

$$[\nu_a(h) + u_m^s(h+\tau_a(h)) - u_n^s(h)]\cdot dX_a^{s+}(h) = 0, \quad \forall s\in S, a=(m,n)\in A, h\in H \tag{8.18h}$$

$$\theta_a = t_a, \forall a\in A \tag{8.18i}$$

$$\tau_a = \theta_a, \forall a\in A \tag{8.18j}$$

Cette reformulation distingue les quatre types de temps, et se prête à analyser leurs effets respectifs et à relaxer les *correspondances temporelles* (8.18f, 8.18i et 8.18j). A priori il convient de fixer τ_a à θ_a dans le modèle de choix des usagers, car il s'agit bien de chronologie.

8.2.2.4 Expression générique

Nous synthétisons l'expression précise (8.18) par l'expression générique suivante :

$$\left(X_A^{S+}, \theta_A\right) \in D_X\left(Q_N^S\right) \tag{8.19a}$$

$$u_N^S \in D_U(\nu_A, \theta_A) \tag{8.19b}$$

$$X_A^{S+} \in \tilde{F}_{VU}\left(\nu_A, u_N^S, \theta_A\right) \tag{8.19c}$$

$$t_A \in \tilde{F}_F\left(X_A^{S+}\right) \tag{8.19d}$$

$$\nu_A = t_A \tag{8.19e}$$

$$\theta_A = t_A \tag{8.19f}$$

La condition (8.19a) synthétise (8.18a) à (8.18c) pour exprimer la faisabilité des volumes par arcs et destinations relativement à l'écart chronologique θ_A. La condition (8.19b) de faisabilité duale reprend (8.18g). La condition (8.19c) exprime le choix des usagers et le chargement en volume de manière intégrée, elle reprend (8.18h). La condition (8.19d) exprime le modèle d'écoulement, tout comme (8.18e). Enfin les correspondances temporelles (8.19f) reprennent (8.18f) et (8.18i).

8.2.2.5 Propriétés de concordance

Dans la formulation par arcs et destinations, on vérifie trivialement la concordance des énergies :

$$\sum_{s\in S}\sum_{a\in A}\int_{h\in H}\nu_a(h)\,dX_a^{s+}(h) = \sum_{a\in A}\int_{h\in H}\nu_a(h)\left[\sum_{s\in S}dX_a^{s+}(h)\right]$$

$$= \sum_{a\in A}\int_{h\in H}\nu_a(h)\,dX_a^{+}(h) \;\text{ avec } dX_a^{+} = \sum_{s\in S}dX_a^{s+} \tag{8.20}$$

En effet tous les termes concernent le niveau local d'un arc. Cette raison induit aussi la concordance temps-volume, qui est réduite ici à

$$\theta_A = t_A \in \tilde{F}_F\left(X_A^{S+}\right) \tag{8.21}$$

Pour saisir la signification de cette forme réduite, il faut bien interpréter θ_A comme un écart de chronologie. Cet écart met en relation les volumes entrants

X_A^{S+} et aussi les volumes sortants X_A^{S-}. Donc la condition de concordance est une contrainte aussi forte dans la formulation par arcs et destinations que dans la formulation par chemins.

Une autre propriété concerne l'écart complémentaire $Z = \sum_{s\in S, a\in A} Z_a^s$, de terme élémentaire

$$Z_a^s = \int_{h\in H} \left[\nu_a(h) + u_m^s(h+\tau_a(h)) - u_n^s(h)\right] dX_a^{s+}(h) \text{ , pour } a \approx (m,n)$$

On peut transformer, en changeant de variable d'intégration $\hbar = h + \tau_a(h)$,

$$\int_{h\in H} u_m^s(h+\tau_a(h))\, dX_a^{s+}(h) = \int_{\hbar\in H} u_m^s(\hbar)\, dX_{a/\tau}^{s-}(\hbar)$$

En sommant sur les arcs, il vient :

$$\sum_{a\in A} \int_{h\in H} u_m^s(h+\tau_a(h))\, dX_a^{s+}(h) = \sum_{n\in N, a\in A_n^-} \int_{h\in H} u_n^s(h)\, dX_a^{s-}(h)$$

et donc :

$$\sum_{a\in A} Z_a^s = \sum_{a\in A} \int_{h\in H} \nu_a(h)\, dX_a^{s+}(h) - \sum_{n\in N} \int_{h\in H} u_n^s(h)\, dQ_{ns}(h)$$

$$+ \sum_{n\in N} \int_{h\in H} u_n^s(h) \cdot \sum_{a\in A_n^-} \left(dX_{a/\tau}^{s-}(h) - dX_{a/\theta}^{s-}(h)\right) \tag{8.22}$$

en notant $X_{a/\theta}^{s-}$ le volume sortant qui intervient dans la conservation locale du flux.

Quand $\tau_a = \theta_a$, le dernier terme s'annule, et l'écart complémentaire s'exprime simplement comme la différence entre le total des coûts par arcs et le total des potentiels, i.e. un écart aux coûts minimum.

8.2.2.6 Un critère de convergence global et ses adaptations

Comme avec la formulation par chemins, pour mesurer la proximité à l'équilibre, il suffit d'évaluer comment sont remplies les conditions d'équilibre, en mesurant un écart global que l'on peut formuler comme une somme

d'écarts sur chaque condition :

$$\begin{aligned} & d_X\left[X_A^{S+}, \theta_A, Q_N^S\right] \\ + \; & d_U\left[u_N^S, \nu_A, \theta_A\right] \\ + \; & d_{VU}\left[X_A^{S+}, \tilde{F}_{VU}\left(\nu_A, u_N^S, \theta_A\right)\right] \\ + \; & d_F\left[t_A, \tilde{F}_F\left(X_A^{S+}\right)\right] \\ + \; & d\left[\theta_A, t_A\right] \\ + \; & d\left[\nu_A, t_A\right] \end{aligned} \tag{8.23}$$

La considération de relations entre les variables endogènes $\left\{X_A^{S+}, u_N^S, \nu_A, t_A, \theta_A\right\}$ et notamment les correspondances temporelles entre les trois dernières, permet de simplifier le critère global de convergence, puisqu'alors les termes d'écart associés s'annulent.

Pour donner un exemple, supposons que

1. $t_A \in \tilde{F}_F\left(X_A^{S+}\right)$, donc $d_F\left[t_A, \tilde{F}_F\left(X_A^{S+}\right)\right] = 0$
2. $\nu_A = t_A$, donc $d\left[\nu_A, t_A\right] = 0$
3. θ_A suit sa définition, donc $d_X\left[X_A^{S+}, \theta_A, Q_N^S\right] = 0$
4. $\tau_A = \theta_A$ pour les potentiels, donc

$$\tilde{d}_{VU} = \sum_{a \in A} \int_{h \in H} \nu_a(h)\, dX_a^+(h) - \sum_{s \in S, n \in N} \int_{h \in H} u_N^S(h)\, dQ_N^S(h)$$

5. u_N^S est tiré de ν_A et θ_A par une recherche des plus courts chemins dynamiques, donc $d_U\left[u_N^S, \nu_A, \theta_A\right] = 0$

Alors le critère global de convergence se réduit à

$$\tilde{d}_{VU} + d_F\left[\theta_A, \tilde{F}_F\left(X_A^{S+}\right)\right]$$

et il est rigoureux pour mesurer l'écart à l'équilibre dans la formulation en fonction de la base endogène $\left\{X_A^{S+}, \theta_A\right\}$.

8.2.2.7 Sur les approches algorithmiques

Nous allons investiguer les approches algorithmiques à un stade en amont, en termes de sélection d'une base endogène et de relaxation de conditions d'équilibre.

Le jeux de variables endogènes est a priori $\left\{X_A^{S+}, u_N^S, \nu_A, t_A, \theta_A\right\}$: X_A^{S+} est le seul jeu de variables proprement primales, θ_A est quasi-primale car elle concerne la propagation des volumes bien plus que des coûts, tandis que ν_A, t_A et u_N^S sont des variables duales.

Dans une approche *primale* avec une base endogène réduite à X_A^{S+}, les coûts découlent essentiellement de $\tilde{F}_F$, et on relaxe les conditions (8.19a), (8.19c). Donc le critère d'écart présente la forme suivante, en supposant $\nu_A = t_A \in \tilde{F}_F\left(X_A^{S+}\right)$ et que u_N^S est déduit de ν_A et de θ_A :

$$d_X\left[X_A^{S+}, \theta_A, Q_N^S\right] + d_{VU}\left[X_A^{S+}, \tilde{F}_{VU}\left(\nu_A, u_N^S, \theta_A\right)\right]$$

La partie d_X dans le critère de convergence est nécessaire, bien qu'omise à tort dans certains modèles (e.g. [WTC02]).

Dans une approche *duale* de base endogène $\{\nu_A\}$, il est difficile d'obtenir un vecteur primal faisable pour (8.19a) et (8.19c) donc on relaxe au moins l'une de ces deux conditions. Si $\tilde{F}_U$ est une application univoque, il convient d'imposer (8.19c) et de relaxer (8.19a), et alors le critère de convergence est réduit à un écart relatif à (8.19a) plus un écart entre ν_A et $t_A \in \tilde{F}_F\left(X_A^{S+}(\nu_A)\right)$.

Ces formulations respectivement primale et duale étaient connues avant notre investigation. Celle-ci en précise les conditions de validité et les critères de convergence, notamment en ce qui concerne la faisabilité primale (8.19a). Cet aspect nous paraît crucial. Dans une formulation par arcs et destinations de l'équilibre dynamique, la relativité chronologique rend délicate l'obtention de la faisabilité primale, qui devient un objectif en soi. C'est pourquoi nous préconisons de privilégier des approches primales-duales dont la base endogène incorpore d'une part X_A^{S+}, d'autre part un ou plusieurs jeux de variables parmi $\left\{u_N^S, \nu_A, t_A, \theta_A\right\}$.

La considération explicite de θ_A vise à imposer (8.19a), en transformant le problème de faisabilité primale en un critère d'écart entre θ_A et t_A (ou ν_A). En procédant ainsi, la variable θ_A est peu flexible et mieux vaut travailler également un jeu ν_A ou t_A. Ci-après nous examinons cette possibilité, en distinguant selon que les potentiels u_N^S sont déduits de ν_A donc contraints, ou sont autonomes.

Il subsiste évidemment bien d'autres possibilités de formulation. Au moins une formulation existe pour les relations (8.19a) à (8.19f) en supposant que les conditions sélectionnées dans le sous-ensemble sont imposées, et les autres relaxées.

8.2.2.8 Formulation primale-duale avec liaison des potentiels

Les variables endogènes basiques sont $\left\{X_A^{S+}, \nu_A\right\}$. Les contraintes intégrées sont : d'une part, la conservation du flux compte tenu des volumes par relation origine-destination, ce qui induit de fait les écarts chronologiques θ_A ; d'autre part, la formation des potentiels par destination, qui est proche de la formation des services. D'où la formulation :

$$\left(X_A^{S+}, \theta_A\right) \in D_X\left(Q_N^S\right) \tag{8.24a}$$

$$u_N^S = \tilde{F}_S\left(\nu_A, \theta_A\right) \tag{8.24b}$$

Les contraintes relaxées sont : d'une part, le choix des usagers et le chargement en volume, traités séquentiellement et notés $\tilde{F}_{VU}$; d'autre part, la formation du temps local par le modèle d'écoulement, en fonction du volume total par arcs X_A^+ :

$$X_A^+ = \sum_{s \in S} X_A^{s+} \tag{8.24c}$$

$$X_A^{S+} \leftrightarrow \tilde{F}_{VU}\left(\nu_A, u_N^S, \theta_A\right) \tag{8.24d}$$

$$\nu_A \leftrightarrow F_F\left(X_A^+\right) \tag{8.24e}$$

Un critère de convergence a la forme

$$d_{VU}\left[X_A^{S+}, \tilde{F}_{VU}\left(\nu_A, u_N^S, \theta_A\right)\right] + d_F\left[\nu_A, F_F\left(\sum_{s\in S} X_A^{s+}\right)\right] + d\left[\theta_A, \nu_A\right] \tag{8.24f}$$

Par rapport à la formulation par chemins, l'intérêt est double : une économie de coût informatique, et aussi la possibilité d'adapter localement le flot par destination X_A^{s+} au coût vers la destination.

8.2.2.9 Formulation primale-duale avec relaxation des potentiels

Les variables endogènes basiques sont $\left\{X_A^{S+}, \nu_A, u_N^S\right\}$. Les contraintes intégrées concernent la conservation du flux compte tenu des volumes par relation origine-destination, ce qui induit de fait les écarts chronologiques θ_A. D'où la formulation :

$$\left(X_A^{S+}, \theta_A\right) \in D_X\left(Q_N^S\right) \tag{8.25a}$$

Les contraintes relaxées incorporent de plus la formation des potentiels par destination :

$$u_N^S \leftrightarrow \tilde{F}_S\left(\nu_A, \theta_A\right) \tag{8.25b}$$

$$X_A^+ = \sum_{s\in S} X_A^{s+} \tag{8.25c}$$

$$X_A^{S+} \leftrightarrow \tilde{F}_{VU}\left(u_N^S, \nu_A\right) \tag{8.25d}$$

$$\nu_A \leftrightarrow F_F\left(X_A^+\right) \tag{8.25e}$$

Un critère de convergence a la forme

$$d_S\left[u_N^S, \tilde{F}_S\left(\nu_A, \theta_A\right)\right] + d_{VU}\left[X_A^{S+}, \tilde{F}_{VU}\left(\nu_A, u_N^S, \theta_A\right)\right]$$

$$+d_F\left[\nu_A, F_F\left(\sum_{s\in S} X_A^{s+}\right)\right] + d\left[\theta_A, \nu_A\right] \tag{8.25f}$$

Par rapport à la formulation avec liaison, l'intérêt de la relaxation des potentiels est de répartir la réaction à la correspondance temps-volume entre les potentiels et les temps. A priori, il incombe aux variables de potentiel u_N^S de porter la chronologie, tandis que les variables locales ν_A peuvent réagir plus rapidement aux volumes chargés localement. Une autre potentialité de cette formulation, est de particulariser le traitement des destinations au travers des potentiels u_N^S.

8.2.3 Une cohérence temporelle pour des formulations par arcs

Les formulations duales par arcs et destinations, avec une base endogène ν_A , peuvent être considérées comme des formulations duales uniquement par arcs puisque la base endogène ne dépend pas des destinations.

Ici nous revenons plutôt aux formulations primales par arcs, avec une base endogène X_A^+. Nous avons signalé au 8.2.1.3 le risque d'incohérence temporelle. Maintenant, grâce à l'analyse des formulations par arcs et destinations, nous allons établir une méthode pour imposer une cohérence temporelle.

8.2.3.1 Enjeu et principe

L'enjeu de la cohérence temporelle, pour des volumes par arcs X_A^+, est qu'il existe des volumes par chemins X_R^+ et des temps par chemins t_R tels que $X_A^+ = F_V\left(X_R^+, t_R\right)$.

Cette condition ne nécessite pas de liaison entre X_R^+ et t_R. Mais on peut supposer que les temps par chemins t_R résultent de temps par arcs t_A, afin d'assurer une propagation cohérente sur le réseau. Alors on peut exprimer la cohérence temporelle au niveau des arcs, en travaillant avec des temps θ_A et des volumes par arcs et destinations X_A^{S+}, de la manière suivante :

$$X_A^+ = \sum_{s\in S} X_A^{s+} \text{ et } \left(X_A^{S+}, \theta_A\right) \in D_X\left(Q_N^S\right) \tag{8.26}$$

Pour imposer la cohérence temporelle, le principe est que X_A^+ se décompose en des volumes par arcs et destinations X_A^{S+}, qui eux-mêmes sont des volumes faisables pour un écart de chronologie θ_A. Les conditions (8.26) se ramènent à des conditions sur les volumes entrants X_A^+ et X_A^{S+} et les volumes sortants X_A^- et X_A^{S-} qui leur sont associés d'après les écarts chronologiques θ_A :

$$X_A^+ = \sum_{s \in S} X_A^{s+} \tag{8.27a}$$

$$X_A^- = \sum_{s \in S} X_A^{s-} \tag{8.27b}$$

$$X_a^{s-}\left(h + \theta_a\left(h\right)\right) = X_a^{s+}\left(h\right) \tag{8.27c}$$

Par définition des volumes cumulés, les fonctions X_a^{s+} et X_a^{s-} sont monotones croissantes, de même que les X_a^+ et X_a^-. En sommant (8.27c) sur les destinations, il vient :

$$X_a^-\left(h + \theta_a\left(h\right)\right) = X_a^+\left(h\right) \tag{8.27d}$$

et pour déterminer un écart chronologique θ_a en fonction de X_a^+ et X_a^-, il suffit que

$$X_a^-\left(h\right) \leq X_a^+\left(h\right) \tag{8.27e}$$

8.2.3.2 Méthode d'inférence chronologique

Pour maintenir la cohérence temporelle de variables X_A^+ tout au long d'un traitement algorithmique, il suffit de maintenir des volumes sortants associés X_A^- tels que X_A^+ et X_A^- soient tous deux des volumes cumulés décomposables par destination et vérifiant (8.27e). En effet, dans ces conditions nous pouvons inférer l'écart chronologique θ_A par la méthode suivante, en déterminant successivement :

– la fonction réciproque du volume sortant

$$X_a^{-[-1]}\left(z\right) = \inf\left\{h, X_a^-\left(h\right) > z\right\}$$

– La fonction d'horaire de sortie par arc

$$H_a(h) = X_a^{-[-1]} \circ X_a^+(h)$$

– La fonction d'écart chronologique

$$\theta_a(h) = H_a(h) - h$$

Ce traitement constitue la *méthode d'inférence chronologique* (KIM).

Comme la fonction X_a^- est croissante, sa réciproque $X_a^{-[-1]}$ également. Or, par définition, $X_a^{-[-1]} \circ X_a^-(h) \geq h$. Donc, si $X_a^+(h) \geq X_a^-(h)$, alors

$$H_a(h) = X_a^{-[-1]} \circ X_a^+(h) \geq h$$

et donc l'écart chronologique est non-négatif.

Dans le cas particulier important où le traitement algorithmique est une méthode des combinaisons convexes, les variables X_A^+ et X_A^- évoluent selon la même combinaison :

$$X_A^{+(k)} = \sum_{s \in S} X_A^{s+(k)} \text{ et } Y_A^{+(k)} = \sum_{s \in S} Y_A^{s+(k)}$$

$$X_A^{-(k)} = \sum_{s \in S} X_A^{s-(k)} \text{ et } Y_A^{-(k)} = \sum_{s \in S} Y_A^{s-(k)}$$

$$X_A^{+(k+1)} = (1-\xi_k) X_A^{+(k)} + \xi_k Y_A^{+(k)} \text{ et } X_A^{-(k+1)} = (1-\xi_k) X_A^{-(k)} + \xi_k Y_A^{-(k)}$$

Comme $X_A^{+(k)} \geq X_A^{-(k)}$ et $Y_A^{+(k)} \geq Y_A^{-(k)}$, de même $X_A^{+(k+1)} \geq X_A^{-(k+1)}$ et la méthode d'inférence chronologique est valide, elle produit un écart topologique $\theta_A^{(k+1)}$.

8.2.3.3 Critère de convergence

Par construction, les volumes par arcs et destinations sous-jacents aux X_A^+ sont faisables pour l'écart chronologique θ_A, donc la condition (8.19a) est vérifiée et n'engendre pas d'écart à l'équilibre. Si de plus $t_A \in F_F\left(X_A^+\right)$ alors (8.19e) est aussi vérifié et n'engendre pas d'écart. Si de plus u_N^S résultent

de ν_A et θ_A selon $u_N^S = \tilde{F}_S(\nu_A, \theta_A)$ alors $d_U\left[u_N^S, \nu_A, \theta_A\right]$ s'annule.

Le critère global de convergence (éq. 8.23) se réduit ainsi à

$$d_{VU}\left[X_A^{S+}, \tilde{F}_{VU}\left(\nu_A, \hat{u}_N^S\right)\right] + d\left[\theta_A, t_A\right] + d\left[\nu_A, t_A\right] \tag{8.28}$$

En imposant une liaison entre ν_A et θ_A, l'écart $d\left[\nu_A, t_A\right]$ s'annule et le critère de convergence devient

$$d_{VU}\left[X_A^{S+}, \tilde{F}_{VU}\left(\nu_A, \hat{u}_N^S\right)\right] + d_F\left[\theta_A, F_F\left(X_a^+\right)\right] \tag{8.29}$$

Une dernière simplification est possible, en utilisant la solution Y_A^{S+} du problème auxiliaire d'affectation sur la base des coûts $\hat{\nu}_A(\theta_A)$ induits par θ_A : comme X_A^+ et Y_A^+ sont cohérents pour la chronologie associée à θ_A, l'écart complémentaire vaut $\hat{\nu}_A(\theta_A) \cdot \left(X_A^+ - Y_A^+\right)$, il mesure

$$d_{VU}\left[X_A^{S+}, \tilde{F}_{VU}\left(\nu_A, \hat{u}_N^S, \theta_A\right)\right]$$

Alors le critère global se réduit à

$$\hat{\nu}_A(\theta_A) \cdot \left(X_A^+ - Y_A^+\right) + d_F\left[\theta_A, F_F\left(X_A^+\right)\right] \tag{8.30}$$

C'est une fonction uniquement des variables par arcs, et donc ce critère jouit à la fois d'un bon pouvoir de résolution et d'une bonne calculabilité.

8.3 Sur les algorithmes d'affectation dynamique

Un grand nombre d'algorithmes ont été proposés pour l'affectation statique du trafic. La revue de Patriksson [Pat94] couvre plus d'un millier d'articles consacrés à la question et publiés entre 1954 et 1993. Pour l'affectation dynamique, nous allons passer en revue brièvement les principaux algorithmes d'équilibrage proposés jusqu'ici, et recommander des développements.

Dans la section précédente, nous avons examiné la partie amont du traitement algorithmique : la sélection de variables endogènes basiques, la sélection

des contraintes à imposer pour en déduire les autres variables endogènes, et la position d'un critère de convergence. Nous avons notamment identifié des approches respectivement primale, duale et primale-duale.

Dans cette section, l'objectif est de développer la partie aval du traitement algorithmique, en spécifiant des algorithmes et en en discutant les propriétés, notamment concernant les critères de convergence.

Nous traiterons surtout des approches primales par combinaison convexe. Tout d'abord au niveau des arcs comme dans le modèle LADTA en 8.3.1, puis au niveau des chemins : avec séparation du chargement et de l'écoulement en 8.3.2 ou avec intégration de ces deux fonctions, comme dans le modèle de Tong et Wong, en 8.3.3. Nous discuterons ensuite l'absence d'autres algorithmes primaux contrairement à l'affectation statique en 8.3.4. Puis nous proposerons en 8.3.5 un nouvel algorithme qualifié d'*hybride*, qui procède par combinaison convexe sur les volumes des arcs et adaptation des coûts des arcs en liaison avec un écart chronologique cohérent avec le volume. Cet algorithme conserve la facilité de traitement du modèle LADTA tout en comportant un critère de convergence rigoureux et facile à évaluer. Enfin, nous évoquerons brièvement les approches duales en 8.3.6 et primales-duales en 8.3.7.

8.3.1 Combinaison convexe au niveau des arcs

8.3.1.1 Organigramme

En notant k le compteur d'itérations, ϵ un seuil de tolérance et ξ_k les coefficients de combinaison (avec $\xi_0 = 1$), l'algorithme de combinaison convexe sur les volumes des arcs procède comme décrit par l'algorithme 8.1.

8.3.1.2 Discussion

Résumons en une ligne d'équation le principe d'une méthode de combinaison convexe au niveau des arcs :

$$X_A^{+(k+1)} = (1 - \xi_k)\, X_A^{+(k)} + \xi_k F_V \left(F_U \left(\bullet\right), \bullet\right) \circ F_S \circ F_F \left(X_A^{+(k)}\right) \qquad (8.31)$$

début

initialisation :

$k \leftarrow 0$

$t_A \leftarrow [t_{a0}(h)]_{h \in H, a \in A}$

$W_A \leftarrow 0$

répéter

$X_A^+ \leftarrow W_A$

Formation des services

$t_R \leftarrow F_S(t_A)$

Choix des usagers

$Y_R^+ \leftarrow F_U(t_R)$

Chargement en volume

$Y_A^+ \leftarrow F_V\left(Y_R^+, t_R\right)$

Combinaison convexe

$W_A \leftarrow (1 - \xi_k) X_A^+ + \xi_k Y_A^+$

Ecoulement

$t_A \leftarrow F_F\left(X_A^+\right)$

$k \leftarrow k + 1$

jusqu'à $\left|W_A - X_A^+\right| < \epsilon$;

fin

Algorithme 8.1 : Algorithme des combinaisons convexes sur les volumes des arcs.

Cette équation indique l'enchaînement des traitements opérés sur un état courant du volume par arcs, $X_A^{+(k)}$: écoulement local du trafic (F_F) puis formation des services (F_S), et ensuite le choix des usagers (F_U) et le chargement en volume (F_V), pour finir par la combinaison convexe stricto sensu.

Cet algorithme présente une grande simplicité et il est très économe.

Au 8.2.1.3 nous nous étions interrogés sur un éventuel risque d'incohérence temporelle concernant les volumes des arcs dans cette approche. Mais nous avons précédemment démontré l'absence d'un tel risque, et montré comment maintenir un champ temporel θ_A d'écart chronologique associé au volume courant.

Une autre question porte sur la disposition d'un critère de convergence rigoureux, compte tenu du risque d'atténuation artificielle associé à l'écart inter-itération en volume des arcs dans une méthode des moyennes successives. A cette question nous répondons ci-après positivement, par plusieurs méthodes dont chacune est suffisante : d'une part avec un critère global par chemins, qui nécessite d'expliciter les chemins ; d'autre part, avec un critère global par chemins, aménagé de manière à éviter d'expliciter les chemins ; et enfin, avec un critère global par arcs et destinations.

8.3.1.3 Critère global par chemins explicites

Compte tenu des relations entre variables endogènes, le critère global de la formulation par chemins (8.11) devient

$$d_U \left[X_R^+, F_U \left(t_R \right) \right] + d_V \left[X_A^+, F_V \left(X_R^+, t_R \right) \right] \tag{8.32}$$

Pour l'évaluer, il convient d'expliciter X_R^+ et de le charger en volume sur les arcs. De plus l'écart aux coûts minimum (la partie d_U) ne peut être exprimée en fonction de X_A^+ car a priori ce n'est pas un volume faisable pour la chronologie associée à t_A et t_R.

8.3.1.4 Critère global par chemins implicites

La deuxième approche est basée également sur le critère global du système (8.11) de la formulation par chemins, mais suit une stratégie de contournementé : on utilise le champ temporel θ_A d'écart chronologique associé à X_A^+ afin d'évaluer chaque terme de (8.32) par rapport à θ_A. Et pour tenir compte de la substitution de θ_A à t_A, on incorpore au critère un écart spécifique $d[\theta_A, t_A]$.

Précisément, on évalue $d_U\left[X_R^+, F_U(\theta_R)\right]$ par la formule $\hat{\nu}_A(\theta_A)\cdot\left(X_A^+ - Z_A^+\right)$ avec Z_A^+ la solution du problème auxiliaire d'affectation sur la base des coûts $\hat{\nu}_A(\theta_A)$ induits par θ_A.

Donc un critère rigoureux est

$$\hat{\nu}_A(\theta_A)\cdot\left(X_A^+ - Z_A^+\right) + d[\theta_A, t_A] \tag{8.33}$$

8.3.1.5 Critère global par arcs et destinations

L'algorithme de combinaison convexe au niveau des arcs peut être rattaché également à la formulation par arcs et destinations de l'équilibre : dans le critère global (eq. 8.23), le terme d_X s'annule ; le terme d_U est redondant avec l'écart entre θ_A et t_A ; les termes d_F et $d[\nu_A, t_A]$ s'annulent. Donc le critère est réduit à

$$d_{VU}\left[X_A^{S+}, \tilde{F}_{VU}\left(\nu_A, u_N^S, \theta_A\right)\right] + d[\theta_A, t_A] \tag{8.34}$$

Le terme d_{VU} peut être évalué par $\hat{\nu}_A(\theta_A)\cdot\left(X_A^+ - Z_A^+\right)$ comme dans la deuxième approche, et ainsi la deuxième et la troisième approche sont équivalentes. Pour incorporer ce critère à l'organigramme de l'algorithme, voir l'organigramme de l'algorithme hybride en page 226.

8.3.2 Combinaison convexe par chemins avec séparation du chargement et de l'écoulement

8.3.2.1 Organigramme

Avec les mêmes notations k et ξ_k que précédemment, l'algorithme de combinaison convexe sur les volumes des chemins procède comme décrit par l'algorithme 8.2.

début
  initialisation :
  $k \leftarrow 0$
  $t_A \leftarrow [t_{a0}(h)]_{h \in H, a \in A}$
  $X_R^+ \leftarrow 0$
  répéter
    Formation des services
    $t_R \leftarrow F_S(t_A)$
    Choix des usagers
    $Y_R^+ \leftarrow F_U(t_R)$
    Combinaison convexe
    $X_R^+ \leftarrow (1 - \xi_k) X_R^+ + \xi_k Y_R^+$
    Chargement en volume
    $X_A^+ \leftarrow F_V\left(X_R^+, t_R\right)$
    $Y_A^+ \leftarrow F_V\left(Y_R^+, t_R\right)$
    Ecoulement
    $t_A \leftarrow F_F\left(X_A^+\right)$
    $k \leftarrow k + 1$
  jusqu'à *l'écart aux coûts minimum pour* X_R^+ *est inférieur à un seuil fixé* ;
fin

Algorithme 8.2 : Algorithme des combinaisons convexes sur les volumes des chemins.

8.3.2.2 Critère de convergence

Le critère global de la formulation par chemins (8.11) s'exprime ici

$$d_S\left[{t_R}^{(k)}, F_S\left(t_A^{(k)}\right)\right] + d_U\left[{X_R^+}^{(k)}, F_U\left({t_R}^{(k)}\right)\right]$$

$$+d_V\left[X_A^{+(k)}, F_V\left(X_R^{+(k)}, {t_R}^{(k)}\right)\right] + d_F\left[{t_A}^{(k)}, F_F\left(X_A^{+(k)}\right)\right]$$

Compte tenu des relations entre les variables endogènes, il se réduit à

$$d_U\left[X_R^{+(k)}, F_U\left({t_R}^{(k)}\right)\right] + d_V\left[X_A^{+(k)}, F_V\left(X_R^{+(k)}, {t_R}^{(k)}\right)\right]$$

Comme $X_A^{+(k)} = F_V\left(X_R^{+(k)}, {t_R}^{(k-1)}\right)$, nous pouvons remplacer la seconde partie du critère par un écart entre ${t_R}^{(k-1)}$ et ${t_R}^{(k)}$, et donc encore par un écart entre ${t_A}^{(k-1)}$ et ${t_A}^{(k)}$, ce qui conduit à un critère global

$$d_U\left[X_R^{+(k)}, F_U\left({t_R}^{(k)}\right)\right] + d\left[{t_A}^{(k-1)}, {t_A}^{(k)}\right] \tag{8.35}$$

Ce critère est composé d'un critère inter-itération sur les temps des arcs, et d'un écart aux coûts minimum.

L'écart aux coûts minimum, pour un modèle basique d'équilibre dynamique, se formule

$$Z = \sum_{r\in R}\int_{h\in H}\left[t_r(h) - u_{i(r)}(h)\right] dX_r^+(h) \text{ avec } u_{i(r)}(h) = \min_{\rho\in i} t_\rho(h)$$

Il vaut

$$Z = \sum_{a\in A}\int_{h\in H} t_a(h)\left(dX_a^+(h) - dY_a^+(h)\right)$$

si la concordance des énergies s'applique.

A l'itération k, on évalue ce critère sans considérer $\tilde{X}_A^{+(k)} = F_V\left(X_R^{+(k)}, {t_R}^{(k)}\right)$ afin d'éviter de charger sur les arcs $X_R^{+(k)}$ selon la chronologie associée à ${t_R}^{(k)}$ (car les $X_A^{+(k)}$ déterminés à l'étape k de l'algorithme correspondent à $F_V\left(X_R^{+(k)}, {t_R}^{(k-1)}\right)$.

On considère uniquement $X_A^{+(k+1)}$ et $Y_A^{+(k)}$, grâce à la relation suivante, qui est valide car ces volumes par arcs découlent de volumes par chemins chargés selon une chronologie identique :

$$X_A^{+(k+1)} = (1-\xi_k)\,\tilde{X}_A^{+(k)} + \xi_k Y_A^{+(k)}$$

donc

$$X_A^{+(k+1)} - Y_A^{+(k)} = (1 - \xi_k)\left(\tilde{X}_A^{+(k)} - Y_A^{+(k)}\right)$$

et donc

$$Z = \frac{1}{1 - \xi_k} \sum_{a \in A} \int_{h \in H} \left[dX_a^{+(k+1)}(h) - dY_a^{+(k)}(h) \right] \tag{8.36}$$

A priori, les volumes par arcs obtenus par cette méthode, notés $X^+_{A/R}$, diffèrent de ceux obtenus par la combinaison convexe au niveau des arcs, notés $X^+_{A/A}$, même si les coefficients de combinaison sont identiques.

8.3.2.3 Discussion

Cet algorithme s'exprime de la manière suivante :

$$\begin{array}{rcl} {t_R}^{(k)} & = & F_S \circ F_F \circ F_V \left(X_R^{+(k)}, {t_R}^{(k-1)} \right) \\ X_R^{+(k+1)} & = & (1 - \xi_k)\, X_R^{+(k)} + \xi_k F_U \left({t_R}^{(k)} \right) \end{array} \tag{8.37}$$

Ici on affecte l'état courant des volumes par chemins, $X_R^{+(k)}$, aux arcs du réseau par le chargement en volume selon la chronologie associée aux temps par chemins antérieurs (F_V), pour former les temps locaux par le modèle d'écoulement (F_F), puis former les temps des services (F_S), et réaffecter la demande (F_U). La chronologie du chargement évolue en fonction de la chronologie antérieure et du volume courant.

Nous avons déjà signalé la lourdeur d'expliciter X_R^+, de le charger sur les arcs et de l'actualiser à chaque itération. Ce coût balance les avantages de la méthode : convergence plus robuste vers l'équilibre, et mesurable plus rigoureusement grâce au critère d'écart aux coûts minimum.

Cet algorithme de combinaison convexe par chemins est intermédiaire entre d'une part les algorithmes de combinaison convexe par arcs de Leurent [Leu03], ou Gentile *et al.* [GMP05] et d'autre part l'algorithme de Tong et Wong [TW00] présenté ci-après.

En théorie, la combinaison convexe au niveau des arcs traite implicitement les volumes par chemins $\tilde{X}_R^+$ sous-jacents aux volumes par arcs $\tilde{X}_A^+$ selon la

chronologie associée aux délais entrée-sortie θ_A du volume par arcs. Donc

$$\tilde{X}_A^+ = F_V\left(\tilde{X}_R^+, F_S\left(\theta_A\right)\right) \text{ avec } \theta_A \in \tilde{X}_A^{-[-1]}$$

Tandis que dans la combinaison convexe au niveau des chemins, la relation est

$$X_A^+ = F_V\left(X_R^+, F_S\left(t_A\right)\right) \text{ avec } t_A \in F_F\left(X_A^+\right)$$

8.3.3 Combinaison convexe par chemins avec intégration du chargement et de l'écoulement

Cet algorithme s'exprime de la manière suivante, en notant $F_{FV} : X_R^+ \mapsto t_A$ le traitement conjoint du chargement en volume et de l'écoulement local [2] :

$$X_R^{+(k+1)} = \left(1 - \xi_k\right) X_R^{+(k)} + \xi_k F_U \circ F_S \circ F_{FV}\left(X_R^{+(k)}\right) \tag{8.38}$$

Cette méthode due à Tong et Wong [TW00] a pour avantages : une convergence robuste vers l'équilibre, qui est mesurable rigoureusement par le critère d'écart aux coûts minimum. Mais le coût informatique est très élevé, et la précision du traitement conjoint F_{FV} n'est pas évidente a priori, ce qui pourrait freiner la convergence vers un équilibre.

L'intégration paraît avantageuse aussi pour éviter tout écart entre t_R et $F_S \circ F_F \circ F_V\left(X_R^+\right)$. Et donc pour éviter d'éventuelles "résonances" dans l'algorithme d'équilibrage entre de tels écarts, et les ajustements induits par le comportement économique des usagers (dans la fonction F_U). Cependant l'avantage est obtenu à un fort surcoût de complexité.

De plus, le chargement en volume et la congestion peuvent rendre efficaces à une itération des chemins qui n'avaient pas encore été utilisés, ce qui rend partiellement obsolète et inopérante la liaison entre t_R et X_R^+.

[2] ce traitement intégré produit conjointement t_A, t_R et X_A^+ tels que $X_A^+ = F_V\left(X_R^+, t_R\right)$ et $t_A = F_F\left(X_A^+\right)$ et $t_R = F_S\left(t_A\right)$

8.3.4 Quels autres algorithmes primaux sont possibles ?

Par définition, un algorithme primal maintient au cours des itérations la conservation du flux. Pour l'affectation statique, des algorithmes variés ont été proposés avec un traitement au niveau des chemins : en particulier l'égalisation par transvasement et la décomposition simpliciale désagrégée, mentionnées au chapitre 6.

Dans ces algorithmes :

- on s'appuie sur une relation simple entre les arcs et les chemins, tant pour les flux que pour les temps.
- on tente de modifier un flux x_r, en portant exactement la même variation Δx_r sur tous les arcs a du chemin, afin de modifier chaque temps t_a de Δt_a et donc de changer le temps du chemin t_r de $\Delta t_r = \sum_{a \in r} \Delta t_a$.
- cette approche est opérée par relation O-D, pour au moins deux chemins concurrents entre lesquels on souhaite échanger du flux.
- de plus, on traite explicitement tous les effets induits sur les autres chemins et relations O-D, grâce à la fonction objectif du programme extrémal primal.

Dans l'affectation dynamique, une variation ΔX_r^+ provoque des ΔX_a^+ le long des arcs a de r en fonction de la chronologie du chargement en volume : le changement sur X_a^+ se répercute directement sur t_a, au prix d'un traitement de tous les instants, puis sur t_r, au prix de traiter tous les instants autant de fois qu'il y a d'arcs dans le chemin. Le coût est donc beaucoup plus élevé que dans l'affectation statique, et surtout, la principale complexité tient à la chronologie du chargement : si t_r évolue alors cette chronologie évolue elle aussi - sauf à choisir de dissocier l'évolution de la chronologie et celle du volume. C'est l'approche primale-duale que nous esquissons plus loin.

8.3.5 Algorithme hybride

Dans l'analyse des formulations, nous avons démontré en 8.3.3 la possibilité de traiter un volume par arcs X_A^+ et un écart chronologique θ_A de manière cohérente et simple : et qu'en fixant les temps de parcours à θ_A, la solution du problème auxiliaire d'affectation procure le moyen d'évaluer facilement un

critère rigoureux de convergence. Nous proposons d'utiliser ces propriétés au sein d'un algorithme d'équilibrage que nous qualifions d'hybride, car il traite conjointement X_A^+, par combinaison convexe, et un champ temporel τ_A qui peut dépendre à la fois des durées de traversée induites par l'écoulement, t_A, et de l'écart chronologique, θ_A.

8.3.5.1 Organigramme

Avec les mêmes notations k et ξ_k que précédemment, l'algorithme hybride procède comme décrit par l'algorithme 8.3.

8.3.5.2 Discussion

Intuitivement, Y_A^+ réagit aux délais courants τ_A qui dépendent de θ_A et t_A liés à X_A^+, et la combinaison convexe de X_A^+ et Y_A^+ doit rapprocher de l'équilibre, tout en préservant la faisabilité primale.

Quand τ_A est réduit à t_A, la méthode équivaut à la combinaison convexe par arcs, et la différence tient seulement au calcul d'un critère rigoureux de convergence. Mais la méthode hybride donne de la flexibilité pour composer τ_A à partir de t_A et θ_A, ce qui pourrait procurer d'autres avantages.

8.3.6 Algorithmes duaux

Au chapitre 6 nous avons rappelé les algorithmes duaux pour l'affectation statique : la méthode de sous-gradient proposée par Fukushima [Fuk84], ainsi que la méthode de combinaison convexe sur les temps proposée par Cascetta et Cantarella [CC95].

Jusqu'ici seule la combinaison convexe sur les temps a été développée pour l'affectation dynamique, par De Palma *et al.* pour le modèle METROPOLIS, dans lequel le volume O-D est réparti entre les chemins conformément à un modèle de choix discret logit. L'algorithme paraît capable de faire décroître le critère de convergence retenu, à savoir un écart inter-itération sur les temps des arcs. Cependant un tel mesurage de convergence est soumis au risque de décroissance artificielle. De plus il semble qu'après quelques di-

début

initialisation :

$k \leftarrow 0$

$t_A \leftarrow [t_{a0}(h)]_{h \in H, a \in A}$

$\theta_A \leftarrow t_A$

$\tau_A \leftarrow t_A$

$W_A^+ \leftarrow 0$

$W_A^- \leftarrow 0$

répéter

$X_A^+ \leftarrow W_A^+$

$X_A^- \leftarrow W_A^-$

Formation des services

$\tau_R \leftarrow F_S(\tau_A)$

Choix des usagers

$Y_R^+ \leftarrow F_U(t_R)$

Chargement en volume

$Y_A^+ \leftarrow F_V\left(Y_R^+, \tau_R\right)$ $Y_A^- \leftarrow Y_A^+ \circ (Id_A + \tau_A)^{[-1]}$

Combinaison convexe

$W_A^+ \leftarrow (1 - \xi_k) X_A^+ + \xi_k Y_A^+$

$W_A^- \leftarrow (1 - \xi_k) X_A^- + \xi_k Y_A^-$

Ecart au coût minimum

$\theta_R \leftarrow F_S(\theta_A)$

$Z_R^+ \leftarrow F_U(\theta_R)$

$Z_A^+ \leftarrow F_V\left(Z_R^+, \theta_R\right)$

$CECM \leftarrow \theta_A \cdot \left(X_A^+ - Z_A^+\right)$

Inférence chronologique

Déterminer θ_A d'après X_A^+ et X_A^-

Ecoulement

$t_A \leftarrow F_F\left(X_A^+\right)$

Evolution du délai

Déterminer τ_A d'après t_A, θ_A, et éventuellement les valeurs antérieures

$k \leftarrow k + 1$

jusqu'à $CECM + d[\tau_A, t_A] + d[\tau_A, \theta_A] < \epsilon$;

fin

Algorithme 8.3 : Algorithme hybride.

zaines d'itérations le critère oscille autour d'un seuil, au lieu de continuer à décroître.

Il conviendrait de concevoir et d'expérimenter un algorithme de sous-gradient pour l'affectation dynamique : cela permettrait au moins de clarifier les relations marginales entre les variables, et les dépendances entre les états instantanés de deux instants distincts.

8.3.7 Algorithmes primaux-duaux

Jusqu'ici les algorithmes primaux-duaux n'ont pas été retenus pour l'affectation statique, en raison des correspondances directes entre les flux et les temps, les arcs et les chemins. Nous avons montré comment le programme dual par arcs et relations O-D se ramène quasi-spontanément au programme dual uniquement par arcs, grâce à la décomposition naturelle du problème d'affectation par relations O-D, quand on relaxe le couplage de congestion au niveau de chaque arc.

Comme les correspondances entre volumes et temps, et entre arcs et chemins, sont nettement plus complexes dans l'affectation dynamique, a priori il y a plus de potentiel d'application pour une approche primale-duale.

Il conviendrait de concevoir et d'expérimenter au moins les trois approches suivantes :

- l'une fondée sur le couple de variables endogènes $\left(X_R^+, \nu_A\right)$.
- l'autre fondée sur $\left(X_R^+, u_I\right)$ avec u_i un potentiel de coût vers la destination.
- une approche primale-duale par arcs et destination, fondée sur $\left(X_A^{S+}, u_N^S\right)$ avec $X_A^{S+} = \left[X_A^{s+}\right]_{s \in S}$ et X_a^{s+} le profil temporel du volume cumulé de l'arc a vers la destination s.

La dernière approche mentionnée (en $\left(X_A^{S+}, u_N^S\right)$) a de fait été développée par Lin et Heydecker [LH05] puisque leur modèle dynamique est traité par destination s, dans l'objectif de produire des volumes X_A^{s+} consistants avec un champ u_N^S. Cependant ces auteurs n'ont donné aucune indication sur la manière de coordonner les destinations, et donc sur la partie duale de l'approche primale-duale.

8.4 Critique du modèle LADTA : expression, formulation, critère et algorithme

Ayant identifié les caractères spécifiques de l'affectation dynamique, nous pouvons à présent critiquer le modèle LADTA dans son expression (8.4.1), sa formulation et son critère de convergence (8.4.2). Notre critique vise à évaluer chaque aspect tel que dans la proposition originelle, et à proposer des compléments ou des alternatives, de manière constructive. En particulier nous développons des algorithmes "atomiques" pour traiter un flot par chemins à un coût informatique réduit (8.4.3), avant de discuter la complexité informatique des algorithmes respectivement par arcs, ou par arcs et destinations ou par chemins (8.4.4).

8.4.1 Sur l'expression du modèle LADTA

Le modèle LADTA est exprimé à deux niveaux d'abstraction : un niveau générique rappelé au 8.1, et un niveau précis où l'on explicite l'expression mathématique interne des quatre fonctions.

L'expression symbolique des variables endogènes, X_R^+, X_A^+, t_A, t_R, sert aux deux niveaux. L'expression générique par chemins suffit pour spécifier la structure du modèle, pour spécifier un algorithme d'équilibrage, pour spécifier un critère rigoureux de convergence, et pour investiguer les caractères spécifiques de l'affectation dynamique.

Au 8.2.2 nous avons donné une expression de l'équilibre par arcs et chemins, et au 8.2.3 nous avons établi des propriétés pour la formulation par arcs : même si l'expression par arcs et destinations se prête moins bien au développement du contenu sémantique du modèle, elle nous a permis d'établir des critères de convergence rigoureux et économes.

8.4.2 Sur la formulation mathématique en problème de point fixe

La formulation mathématique consiste à mettre l'expression du modèle sous la forme d'un problème type de recherche opérationnelle, ce qui fournit un cadre d'analyse formelle et de résolution algorithmique : la formulation doit servir à établir les propriétés de continuité et de sensibilité pour les fonctions essentielles du modèle, et aussi à exprimer des algorithmes.

Pour un modèle d'affectation, la formulation sert à tracer les grandes lignes de l'analyse formelle de l'équilibre : l'existence d'un équilibre peut être garantie par la continuité de chaque fonction ; la stabilité de l'équilibre, par une sensibilité modérée de chaque fonction.

L'expression générique du modèle LADTA a permis de le formuler en problème type de recherche opérationnelle : un problème de point fixe en fonction des variables endogènes basiques X_A^+.

Cette formulation appelle les commentaires suivants :

- modularité du traitement, avec un algorithme-composant pour opérer chacune des quatre fonctions essentielles.
- les algorithmes-composants sont efficaces.
- simplicité de l'algorithme d'équilibrage, dont la consommation en mémoire informatique est minimale.
- le critère de convergence originel, un écart inter-itération sur le volume des arcs, n'a qu'un pouvoir de résolution modéré. Les critères de convergence rigoureux établis au 8.3 induisent un doublement du coût de traitement informatique.
- Il n'y a pas de risque d'incohérence temporelle pour les volumes des arcs , pour lesquelles il existe des volumes par chemins sous-jacents.

Au total, les choix de formulation du modèle LADTA privilégient la simplicité des calculs.

8.4.3 Algorithmes atomiques pour traiter les chemins

Pour traiter explicitement les chemins, en volume et en temps, dans un algorithme de combinaison convexe, il est nécessaire :

– d'expliciter les volumes des chemins, les variables X_R^+.
– de les actualiser à chaque itération.
– à chaque itération, de propager X_R^+ sur le réseau selon la chronologie associée à t_R, afin d'obtenir $X_A^+ = F_V\left(X_R^+, t_R\right)$.
– alors la tâche de propagation est dupliquée, puisqu'on propage déjà l'état auxiliaire $Y_A^+ = F_V\left(Y_R^+, t_R\right)$.
– avec X_A^+, Y_A^+ et les temps des arcs t_A sous-jacents aux t_R, on obtient le critère d'écart aux coûts minimaux.

Le surcoût informatique est élevé :

– duplication du chargement en volume.
– surtout, le traitement de X_R^+ : explicitation, actulisation et propagation. Celle-ci est la tâche la plus coûteuse car elle nécessite d'évaluer le temps de chemins qui peuvent ne pas être efficaces compte tenu des t_A en vigueur à l'itération courante : le chargement en volume de X_R^+ est plus complexe que celui de Y_R^+.

Afin de limiter le surcoût, nous proposons une approche informatique pour réduire autant que possible le volume des calculs : une représentation informatique par faisceau d'atomes (8.4.3.1), utilisée dans des algorithmes " atomiques " pour charger en volume les arcs (8.4.3.2) et pour actualiser le volume (8.4.3.3).

8.4.3.1 Représenter des chemins par un faisceau d'atomes

Pour optimiser la formation des services et le chargement en volume, leurs algorithmes respectifs mettent en oeuvre un objet informatique original : un *faisceau d'atomes* dans le cadre de la théorie du réseau de services [Leu06].

Un faisceau représente un ensemble de chemins vers une même destination, et chaque chemin r est représenté comme une séquence α , un atome α représentant la traversée d'un arc a par les chemins r dont la partie en aval de a est commune.

Chaque atome renvoie à l'atome suivant dans la séquence vers la destination ; ainsi la description d'un chemin se réduit à son atome initial et quelques variables gérées au niveau de cet atome.

Jusqu'à présent le faisceau d'atomes a été utilisé pour opérer la formation des services, $t_R = F_S(t_A)$, le choix des usagers pour l'état auxiliaire $Y_R^+ = F_U(t_R)$, et le chargement aux arcs du volume auxiliaire, $Y_A^+ = F_V\left(Y_R^+, t_R\right)$. Nous proposons de l'utiliser également pour gérer le volume courant X_R^+.

8.4.3.2 Algorithme atomique de chargement en volume

Chaque atome α doit contenir :
- l'identité de son arc.
- un pointeur sur son successeur β , noté σ_α.
- deux profils temporels Z_α^+ et Z_α^- des volumes respectivement entrant dans et sortant de l'atome, pour l'ensemble des chemins transitant par l'atome.
- le profil temporel X_α^+ du volume entrant dans le chemin qui correspond strictement à l'atome et à ses descendants dans le faisceau : c'est le chemin intrinsèque de l'atome.

On associe aussi au faisceau une liste de pointeurs sur les atomes, par ordre décroissant de hauteur (nombre d'arcs du chemin intrinsèque).

L'algorithme de chargement en volume pour X_R^+ consiste à traiter successivement les destinations. On parcourt le faisceau d'une destination dans l'ordre décroissant des hauteurs d'atomes, afin de consolider progressivement les flux :
- avant le parcours, chaque X_α^+ contient le X_r^+ du chemin intrinsèque de l'atome α. On initialise $Z_\alpha^+ \leftarrow X_\alpha^+$.
- dans le parcours, on traite chaque atome :
  - en propageant son volume jusqu'à la sortie de l'arc : $Z_\alpha^- \leftarrow Z_\alpha^+ \circ \overline{H}_a$
  - en transmettant ce volume sortant à l'entrée de l'atome suivant β désigné par le pointeur σ_α : $Z_\beta^+ \leftarrow Z_\beta^+ + Z_\alpha^-$

8.4.3.3 Algorithme atomique pour actualiser un flot par chemins

L'actualisation du flot courant X_R par combinaison avec un flot auxiliaire Y_R est une opération plus délicate : il s'agit de combiner deux faisceaux en un troisième, en identifiant les parties aval communes des chemins. Pour ce faire,

nous proposons l'algorithme atomique suivant afin de construire le faisceau de W_R à partir des faisceaux de X_R et Y_R (en fait, il y a un faisceau par destination pour chaque flot) :

- à chaque atome de X_R, Y_R est associée une liste de ses atomes immédiatement prédécesseurs dans son faisceau.
- à chaque atome ω de W_R sont associés deux pointeurs, vers d'éventuels correspondants l'un dans X_R et l'autre dans Y_R.
- on initialise le faisceau de W_R par un atome fictif de hauteur nulle qui modélise la destination, et dont les correspondants sont les atomes fictifs de X_R et Y_R.
- on construit progressivement le faisceau de W_R en traitant ses atomes par ordre croissant de hauteur (approche breadth-first). Pur un atome ω , on parcourt conjointement les deux listes L_X et L_Y des prédécesseurs immédiats de ses correspondants ω_X et ω_Y , afin de créer les atomes antécédents de ω en détectant les atomes "d'intersection de L_X et L_Y " - i.e. dont l'arc est identique. Ce parcours est efficace si les deux listes L_X et L_Y sont conservées dans le même ordre des arcs. On conserve cet ordre pour la liste d'antécédents L_W associée à ω. Pour chaque ω' antécédent de ω, on définit le pointeur sur ω, l'éventuel correspondant dans X_R issu de L_X, l'éventuel correspondant dans Y_R issu de L_Y, et le volume $X_{\omega'}$ est formé par combinaison des volumes des correspondants dans X_R et Y_R.

Après avoir composé le faisceau combiné associé à W_R, il ne reste plus qu'à actualiser X_R^+ en remplaçant son faisceau par celui de W_R.

8.4.4 Sur la complexité informatique

Terminons ces considérations algorithmiques par une évaluation sommaire de la complexité informatique des traitements évoqués. Nous procédons à l'évaluation en deux étapes : d'abord nous qualifions les "blocs de traitement" que sont les fonctions de formation des services, de chargement en volume, d'actualisation d'un volume : nous évaluons les blocs les uns relativement aux autres et nous dégageons deux facteurs de complexité ; ensuite, nous évaluons

la complexité des algorithmes en fonction des blocs qui les composent.

8.4.4.1 Qualification des fonctions

Nous distinguons les fonctions suivantes :

- La formation des services efficaces, notée FSE.
- Le chargement en volume de services efficaces, noté CSE, qui correspond aux Y_R^+ solution d'un problème auxiliaire.
- Le chargement en volume de services quelconques, noté CSQ, qui correspond aux X_R^+ courants.
- L'actualisation d'un volume par chemins, notée AVC, également pour des variables X_R^+.

Nous allons qualifier chacune de ces fonctions en supposant que les variables de volume cumulé et de temps sont des fonctions affines par morceaux : un tel morceau de profil horaire, constitue l'élément de base dans le traitement informatique des variables.

Pour la formation des services efficaces, nous supposons que la recherche des plus courts chemins dynamiques est opérée par l'algorithme atomique à la Dijkstra de Leurent [Leu06], qui conserve uniquement des morceaux efficaces, avec un atome par morceau. Il présente une complexité minimale pour sa tâche. Chaque morceau est traité d'abord à la création de l'atome correspondant, avec un calcul de ses coûts et une insertion en liste d'attente, puis en tant que morceau courant sortant de la liste d'attente, pour éventuellement créer des successeurs. La liste d'attente est rangée en ordre chronologique des horaires initiaux des morceaux, ce qui facilite l'insertion et le retrait de l'atome d'horaire minimal. Au total, le nombre d'opérations est lié au nombre N_A d'atomes, avec un facteur de proportionnalité lié au nombre de successeurs pour chaque atome, noté K en moyenne, et à la profondeur de la liste d'attente qui est en $\log(N_A)$ pour un codage par heap.

$$O\left(FSE\right) = N_A\left(\log(N_A) + K\right)$$

Pour le chargement des services efficaces, traité par un algorithme spécifique de Leurent [Leu06], aux morceaux des profils temporels des t_R s'ajoutent

ceux des volumes Q_i. Les atomes pour X_R^+ sont donc plus nombreux que pour t_R. Cependant il n'y a pas d'opérations de tri, les traitements consistent à cumuler et propager des volumes ce qui procède surtout par des additions et soustractions, à comparer aux additions pour former les coûts des services. Grossièrement, nous évaluons que

$$O\left(CSE\right) = O\left(FSE\right)$$

Le chargement en volume de services quelconques est une tâche plus lourde. Il faut à la fois évaluer les temps des services composés (pour la propagation), et calculer le morceau en flux et en support horaire. De plus le nombre d'atomes est plus élevé, d'un facteur désigné par λ. En effet les services quelconques sont plus nombreux que les services efficaces d'une itération, puisqu'ils consolident des services efficaces déterminés au cours d'itérations successives. Grossièrement, nous évaluons que

$$O\left(CSQ\right) = \lambda\left(O\left(FSE\right) + O\left(CSE\right)\right)$$

Une valeur de λ comprise entre 3 et 5 paraît raisonnable vu que les algorithmes sophistiqués d'affectation statique convergent en 5 itérations, donc n'engendrent pas plus de cinq itinéraires efficaces par relation O-D.

L'actualisation d'un volume par chemins, si on la traite par un algorithme atomique, nécessite de construire des faisceaux et de parcourir deux faisceaux existants : par morceau on opère un calcul pour réduire le volume si le morceau n'a qu'un antécédent, ou pour composer le volume en combinant des antécédents issus des deux faisceaux de base. Grossièrement, nous évaluons que

$$O\left(AVC\right) = O\left(CSQ\right)$$

en raison du nombre équivalent d'atomes et des équivalences entre les opérations élémentaires par atome.

Au total, en traitant les chemins par faisceaux d'atomes et en appliquant des algorithmes atomiques, une évaluation plausible est

$$O(AVC) = O(CSQ) = 2\lambda O(FSE) = 2\lambda O(CSE)$$

Nous notons U l'unité de compte $O(FSE)$.

Un traitement individuel des chemins, sans utiliser les faisceaux d'atomes et la récursivité inhérente, engendrerait un nombre de morceaux supérieur. Nous notons μ le facteur multiplicatif, dont nous évaluons très forfaitairement l'ordre de grandeur à $3 - 10$, selon la profondeur du réseau.

Ainsi, entre un chargement des services opérés par faisceau sur des services efficaces, et un chargement opéré sans faisceau sur des services quelconques, le rapport des coûts informatiques est de 1 à $2\lambda\mu$, soit de 1 à $20 - 50$.

8.4.4.2 Evaluation des algorithmes

La méthode des combinaisons convexes par arcs, dans sa version originelle avec un critère de convergence réduit à l'écart inter-itération en volume des arcs, appelle à chaque itération les fonctions FSE et CSE pour calculer Y_R^+ puis Y_A^+, soit $2U$.

Dans la version avec critère de convergence rigoureux, il faut rajouter une fonction FSE et une fonction CSE pour calculer aussi Z_R^+ puis Z_A^+ , soit $2U$ de plus qui portent le coût à $4U$.

La méthode des combinaisons convexes par chemins avec séparation du chargement et de l'écoulement, appelle à chaque itération une FSE, un CSE, un CSQ et un AVC. Son coût s'élève à $(2 + 4\lambda)U$ si les chemins sont traités par faisceaux et algorithmes atomiques, ou à $(2 + 4\lambda)\mu U$ si les chemins sont traités individuellement.

Ainsi le rapport de complexité entre les algorithmes de combinaison convexe respectivement par arcs et par chemins, est de 1 à :

- $(1 + 2\lambda) \approx 10$ sans critère rigoureux par arcs et avec traitement atomique des chemins ;
- $(1 + 2\lambda)\mu \approx 50$ sans critère rigoureux par arcs et sans traitement atomique des chemins ;

- $(\frac{1}{2}+\lambda) \approx 5$ avec critère rigoureux par arcs et avec traitement atomique des chemins ;
- $(\frac{1}{2}+\lambda)\mu \approx 25$avec critère rigoureux par arcs et sans traitement atomique des chemins.

Chapitre 9

Expériences numériques

Introduction

Afin d'expérimenter et comparer les deux méthodes de combinaison convexe au niveau respectivement des ars et des chemins, ainsi que les critères de convergence associés, nous avons d'une part programmé les algorithmes et les critères, et d'autre part conçu un cas d'école. Les algorithmes et les critères ont été programmés et incorporés au prototype du modèle LADTA, dans l'environnement logiciel SCILAB.

Notre objectif de connaissance porte d'une part sur le comportement des algorithmes en nombre d'itérations pour atteindre un certain niveau de convergence ; et d'autre part, sur le comportement des critères d'écart qui interviennent dans un critère global de convergence.

Nous présentons d'abord le plan d'expérience et le cas d'école en 9.1. Puis nous indiquons le comportement des algorithmes et des critères en 9.2, avant de discuter la pondération relative des critères d'écart qui interviennent dans un critère global de convergence 9.2.8.

9.1 Le plan d'expérience et le cas d'école

9.1.1 Le plan d'expérience

Notre plan d'expérience est le suivant :

1. appliquer la combinaison convexe sur les volumes des arcs, en mesurant le critère d'écart inter-itération en volume des arcs, noté CIVA_A, le critère d'écart inter-itération en temps des arcs noté CITA_A, le critère d'écart aux coûts minimum noté CECM_A, ainsi que le critère d'écart temporel entre θ_A et t_A, noté CB_A.
2. comparer les évolutions des critères CIVA_A, CITA_A, CECM_A et CB_Aau cours des itérations.
3. appliquer la combinaison convexe sur les volumes des chemins, en mesurant le critère d'écart inter-itération en volume des arcs, noté CIVA_R, le critère d'écart inter-itération en temps des arcs, noté CITA_R, et surtout le critère d'écart aux coûts minimum, noté CECM_R.
4. comparer les évolutions des critères CIVA_R, CITA_R et CECM_R au cours des itérations.
5. comparer les évolutions des critères CIVA_A et CIVA_R, de CECM_A et CECM_R au cours des itérations.
6. appliquer l'algorithme hybride, en mesurant le critère d'écart entre τ_A et t_A, noté CBH_1, le critère d'écart entre τ_A et t_A, noté CBH_2, et le critère d'écart aux coûts minimum, noté CECM_H.
7. Comparer les évolutions des critères CBH_1, CBH_2 et CECM_H au cours des itérations.
8. tester plusieurs valeurs du paramètre γ dans l'algorithme hybride.
9. Et mesurer l'écart entre les volumes des arcs issus des trois algorithmes.

Ayant procédé aux simulations correspondantes, nous pourrons alors comparer les trois algorithmes.

De plus, nous testerons plusieurs suites de coefficients pour les combinaisons convexes, ainsi que plusieurs portées d'assimilation temporelle des morceaux de profil horaire.

9.1.2 Position du cas d'école

Le cas d'école concerne un réseau à une seule relation O-D, avec quatre arcs $\{b, d, a, e\}$ disposés en deux maillons parallèles $\{b, d\}$ puis $\{a, e\}$, et qui

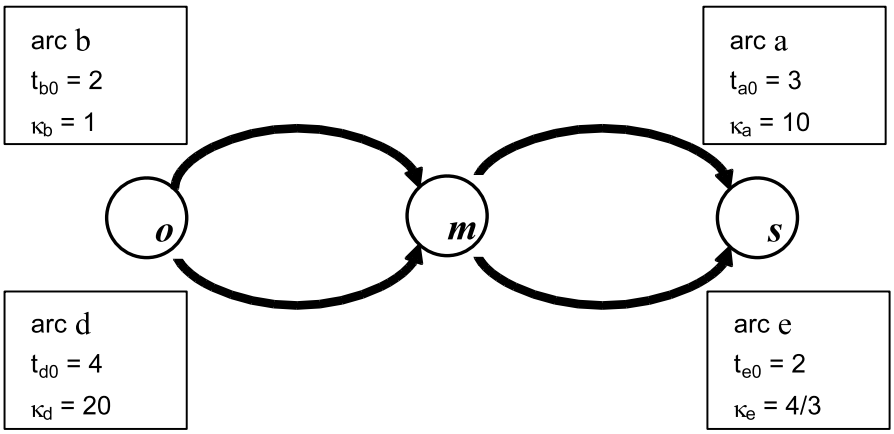


FIG. 9.1 – Réseau à quatre arcs.

forment quatre itinéraires depuis l'origine vers la destination : ba, be, da, de (cf. fig. 9.1).

Le volume O-D cumulé a pour profil temporel $h \mapsto A \min\{\max\{h, 0\}, 1\}$, ce que nous notons par $Q_i = 4.\,[0;1]$ car le débit est de 4 sur l'intervalle $[0;1]$ et 0 ailleurs.

Les arcs a ont des délais minimaux t_{a0} et des débits capacitifs κ_a invariables, indiqués sur la figure, et fixés de manière que le volume puisse saturer les arcs b et e, tandis que les arcs a et d ont un temps constant. Ainsi plusieurs itinéraires concurrents peuvent être saturés, dont l'un, be, peut même comporter deux files d'attente consécutives.

9.1.3 Solution du cas d'école

De l'origine o à la jonction m, l'arc b est meilleur que l'arc d tant que son temps de parcours reste inférieur à $t_{d0} = 4$, donc depuis l'instant initial h_0^* jusqu'à l'horaire de départ critique $h_1^* = 2/3$h. Ensuite le flux O-D se partage entre les deux arcs de manière à égaliser les temps, et donc à les stabiliser à la valeur de 4h puisque l'arc d n'est pas saturé. Ainsi jusqu'à 1h le temps de b reste constant, et donc le flux entrant dans b est égal à sa capacité en sortie, soit 1 véh/h.

Entre m et la destination s, tant que seul l'arc b est utilisé le flux est limité à la capacité en sortie de b, soit 1véh/h. Comme e est plus court et que ce flux n'atteint pas sa capacité, seul e est utilisé tant que b est seul utilisé. Quand d est à son tour utilisé, le flux entrant dans m passe à 4véh/h puisque les 3véh/h passant par d rejoignent les 1véh/h sortant de b. e reste utilisé seul tant que son temps reste inférieur à celui de a soit 3h. L'instant où le flux via m passe de 1 à 4 véh/h est $h_1^* + t_{d0} = 14/3$h, et après cet instant le volume qui passe représente $4.\,[14/3; 5]$ soit 4/3véh.

Le débit sature la capacité de e et l'écoulement de ce volume entrant en 1/3h nécessite 1h en sortie de e, à partir de l'horaire $14/3 + 2 = 20/3$h, et jusqu'à l'horaire $20/3 + 3 = 23/3$h. Pendant cette période la durée de traversée de e reste inférieure à 8/3h donc e reste meilleur que l'arc a, qui n'est pas utilisé à l'équilibre.

En résumé, à l'équilibre $X_{be}^+ = 4.\,[0; 2/3] + 1.\,[2/3; 1]$ et $X_{de}^+ = 3.\,[2/3; 1]$, tandis que $X_{da}^+ = X_{ba}^+ = 0$.

9.2 Comportement des algorithmes et des critères

Les simulations menées ont manifesté la similarité des équilibres et des rythmes d'équilibrage en 9.2.1 ; l'influence non négligeable de la portée d'assimilation dans les itérations initiales en 9.2.2 ; l'importance de bien choisir la suite des coefficients de combinaison en 9.2.3. Nous résumerons les constats en 9.2.7 après avoir examiné en particulier la combinaison convexe par arcs (9.2.4), celle par chemins (9.2.5), ainsi que l'algorithme hybride (9.2.6).

9.2.1 La similarité des équilibres

Les volumes ${X_{A/A}^+}^{(k)}$ et ${X_{A/R}^+}^{(k)}$ obtenus au fil des itérations k par les combinaisons convexes au niveau resepectivement des arcs (A) et des chemins (R), se rapprochent progressivement l'un de l'autre (cf. figure 9.2), et de même les volumes $X_{A/H}^{+[k]}$ issus de l'algorithme hybride se rapprochent des $X_{A/A}^{+[k]}$ (figure 9.3. Comme dans chaque cas les critères d'écart décroissent

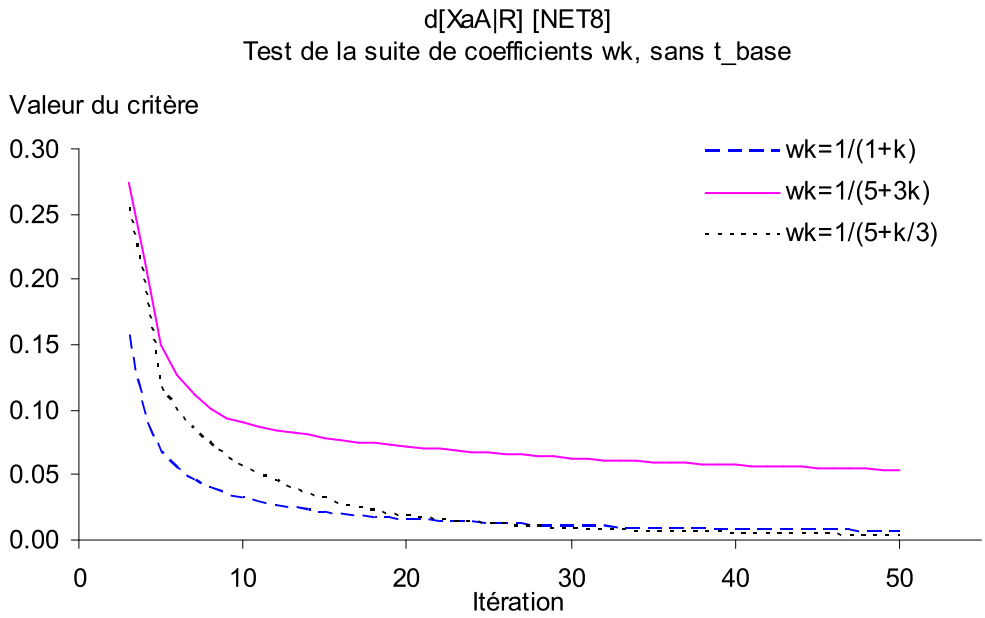


FIG. 9.2 – Les méthodes de CC produisent un même équilibre.

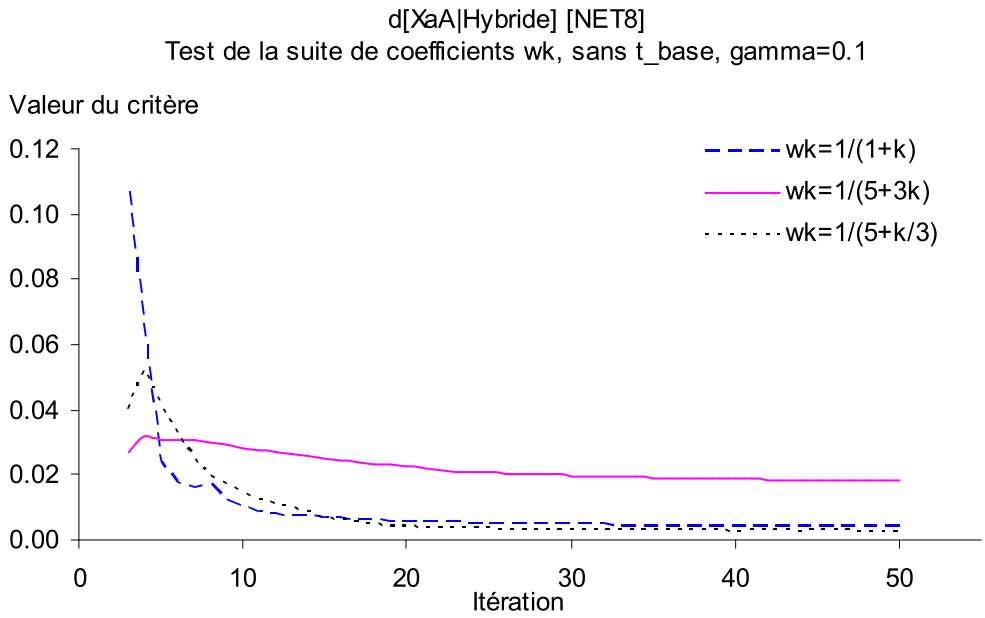


FIG. 9.3 – L'algorithme hybride produit aussi l'équilibre.

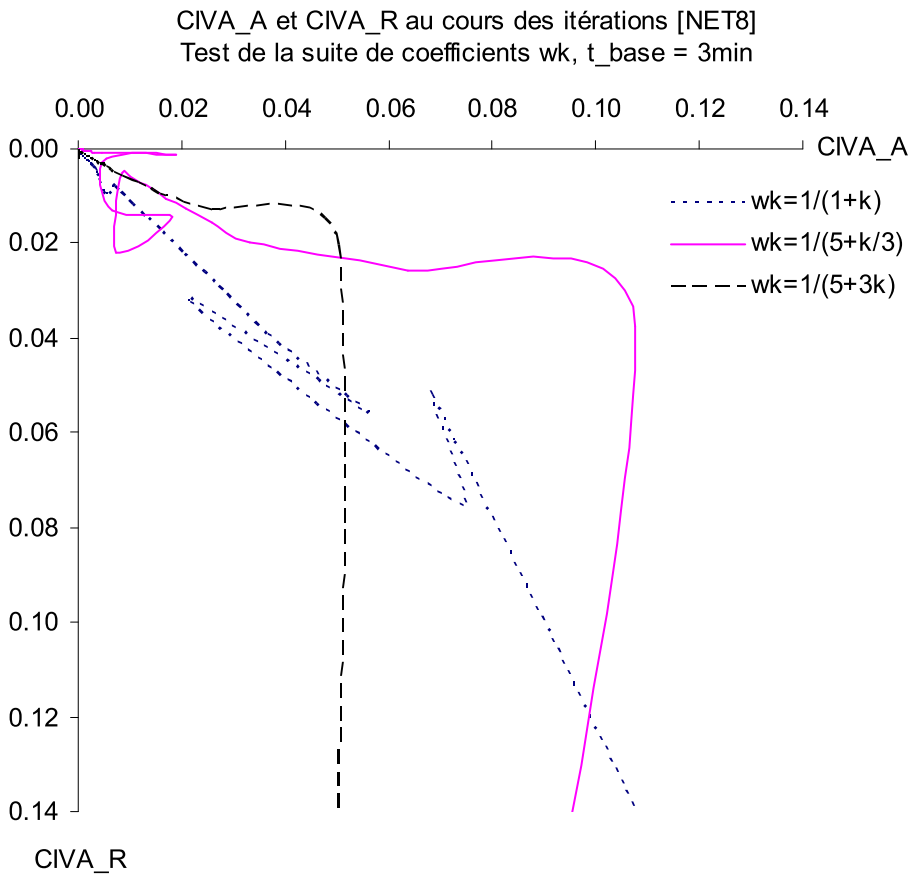


FIG. 9.4 – Similarité des rythmes d'équilibrage, méthodes de CC.

vers 0, nous en concluons que les trois algorithmes convergent vers le même équilibre.

De plus, les rythmes de progression sont similaires pour les deux méthodes, puisque les critères d'écart CIVA_A et CIVA_R ont des valeurs comparables (valeurs proches de la ligne diagonale sur la figure 9.4).

Ce constat s'étend à l'algorithme hybride (figure 9.5).

9.2.2 Réglage de la portée d'assimilation temporelle

Pour réduire les calculs, plutôt que recourir à une discrétisation statique selon l'horaire, nous pouvons simplifier les profils horaires en limitant la fragmentation des morceaux, par assimilation des morceaux de taille faible et de

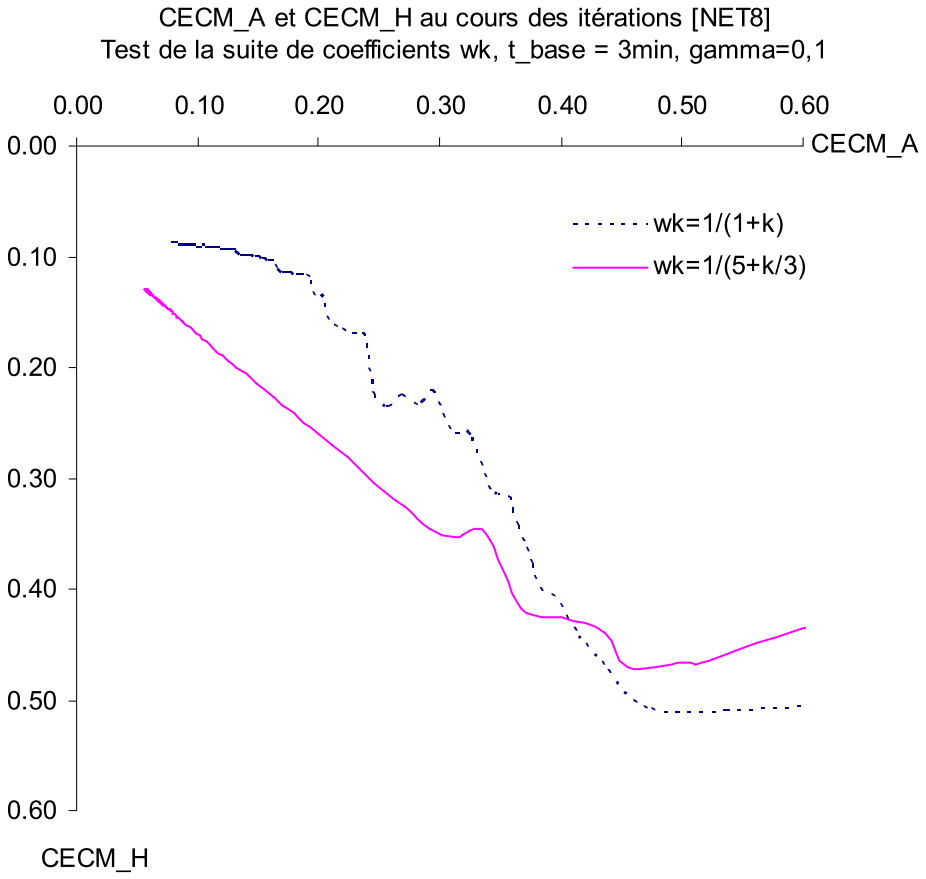


FIG. 9.5 – Similarité des rythmes d'équilibrage, hybride et CC par chemins.

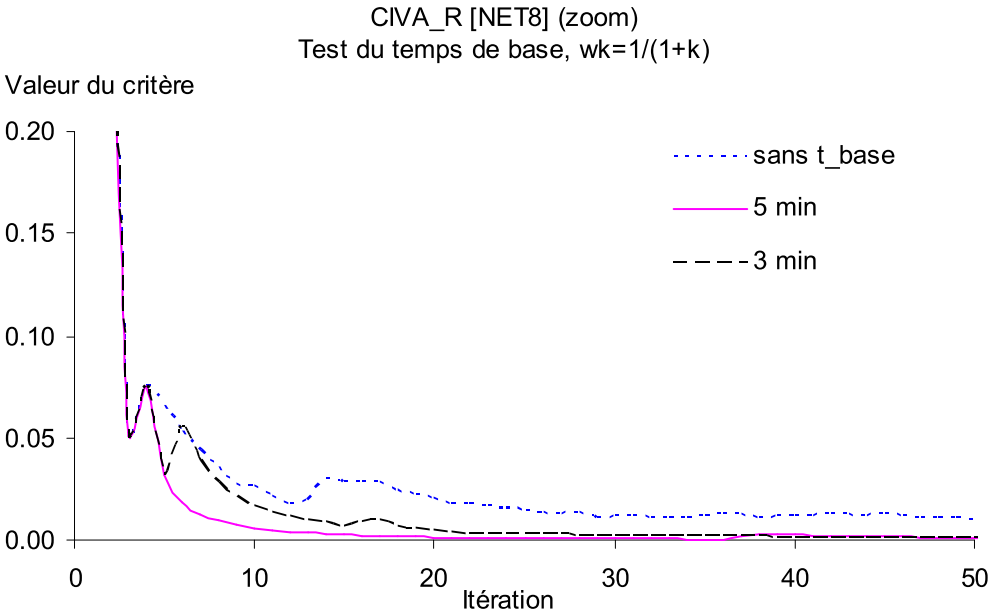


FIG. 9.6 – Influence de la portée d'assimilation.

pente voisine. Cette procédure est paramétrée principalement par la "portée d'assimilation temporelle", qui est la limite supérieure des "tailles faibles" pour les morceaux.

Nous avons testé trois valeurs pour ce paramètre : 0, 3 ou 5 minutes. L'influence est imperceptible sur l'écart aux coûts minimum ; mais plus forte sur l'écart inter-itération en volume (voir figure 9.6).

De manière a priori surprenant, l'assimilation améliore ce critère : une explication possible est que l'assimilation stabilise les points intermédiaires des profils temporels. Cela incite à accepter la procédure d'assimilation : une valeur de 5 minutes convient aussi bien qu'une valeur de 3 minutes.

9.2.3 Réglage de la suite des coefficients

Nous avons testé trois suites de coefficients ξ_k pour les combinaisons convexes :

– $\xi_k = \frac{1}{1+k}$, suite usuelle dans la littérature scientifique.

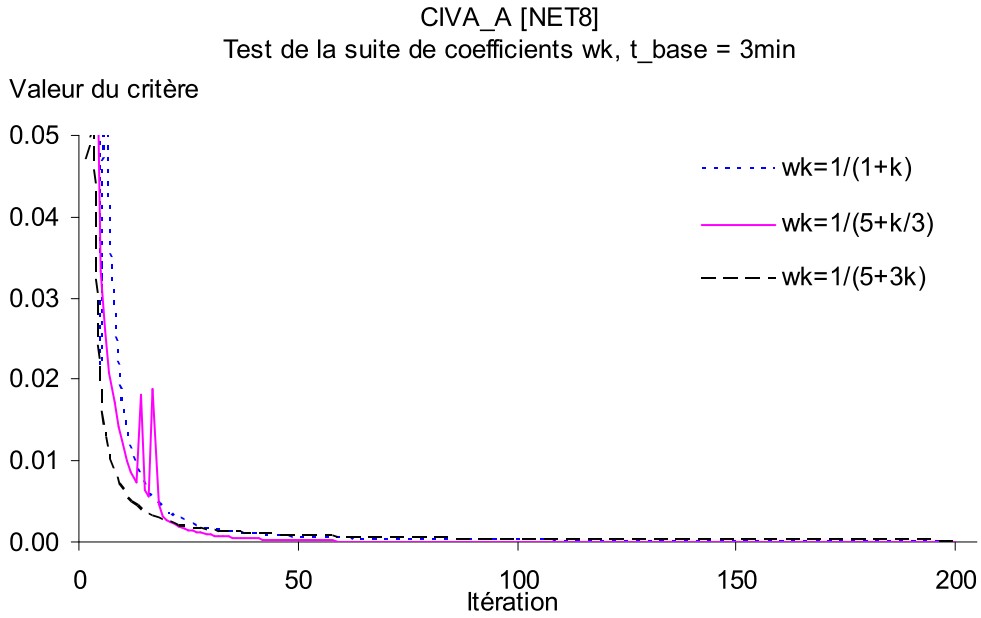


FIG. 9.7 – Influence de la suite des coefficients sur le CIVA_A.

- $\xi_k = \frac{1}{5+k/3}$ qui débute à un niveau plus faible mais ensuite prend des valeurs triples : éprouvée plus efficace par les auteurs dans des affectations statiques.
- $\xi_k = \frac{1}{5+3k}$ qui débute à un niveau plus faible encore et ensuite décroît plus rapidement.

L'influence de la suite choisie est manifeste sur chaque critère d'écart. Pour le CIVA, la suite fortement décroissante réduit plus vite l'écart, très significativement dans les dix premières itérations, moins nettement dans les suivantes où les trois suites donnent une même enveloppe inférieure (figure 9.7).

Pour le CECM_R, les décroissances initiales sont différenciées entre les suites mais l'évolution à long terme assez semblable (figure 9.8), tandis que pour le CECM_A la suite la plus décroissante peine à faire décroître le critère (figure 9.9).

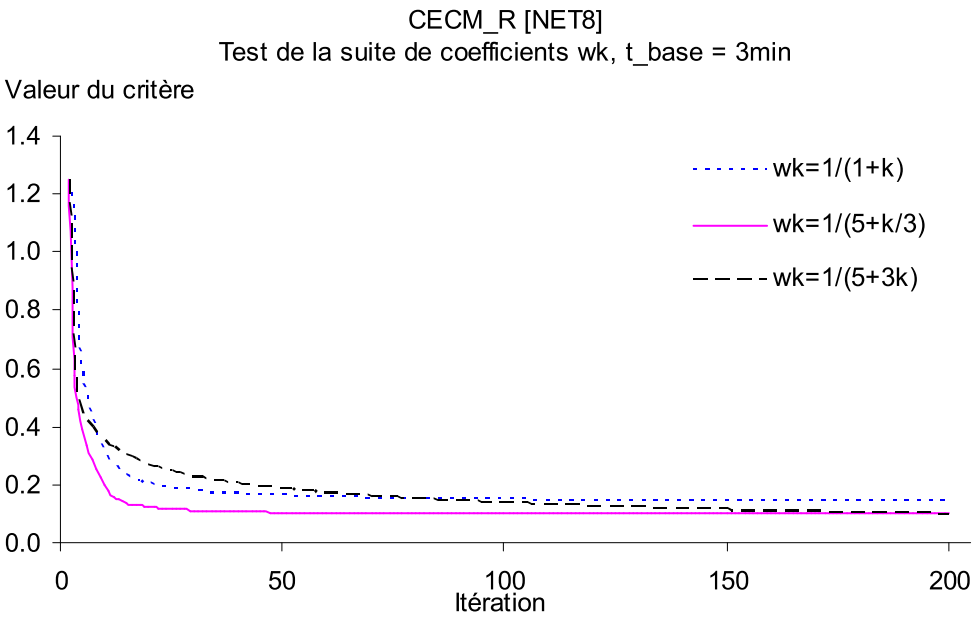


FIG. 9.8 – Influence de la suite des coefficients sur le CECM$_R$.

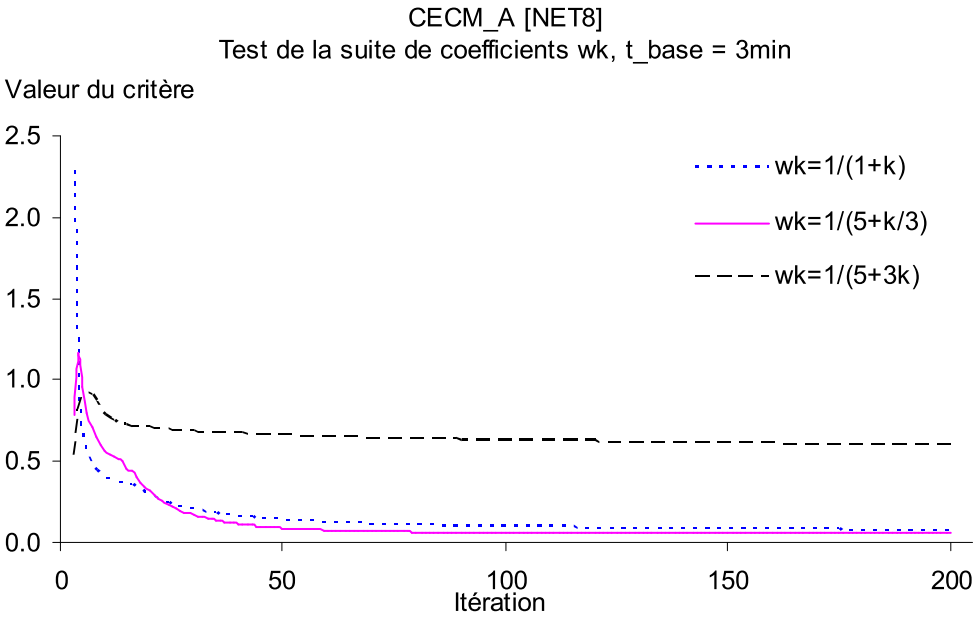


FIG. 9.9 – Influence de la suite des coefficients sur le CECM$_A$.

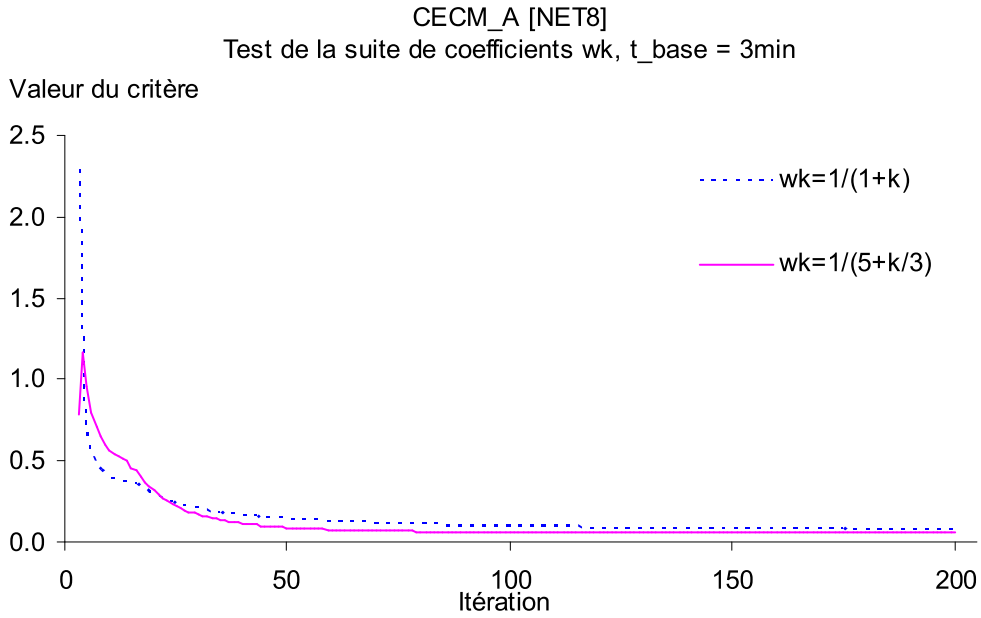


FIG. 9.10 – Evolution de CECM pour la combinaison convexe par arcs.

9.2.4 Sur la combinaison convexe par arcs

La méthode de combinaison convexe par arcs converge de manière robuste sur le cas d'école, avec une décroissance régulière des critères d'écart : d'une part le CECM_A (figure 9.10), d'autre par le CB_A (figure 9.11).

9.2.5 Sur la combinaison convexe par chemins

De même la méthode de combinaison convexe par chemins converge de manière robuste sur le cas d'école. Le critère CECM_R décroît assez lentement (figure 9.13), tandis que le critère CB_R décroît beaucoup plus vite (figure 9.14). Ceci suggère que pour cette application, le critère CECM suffit empiriquement pour mesurer la convergence.

Quand au critère CIVA, il décroît vite (figure 9.15).

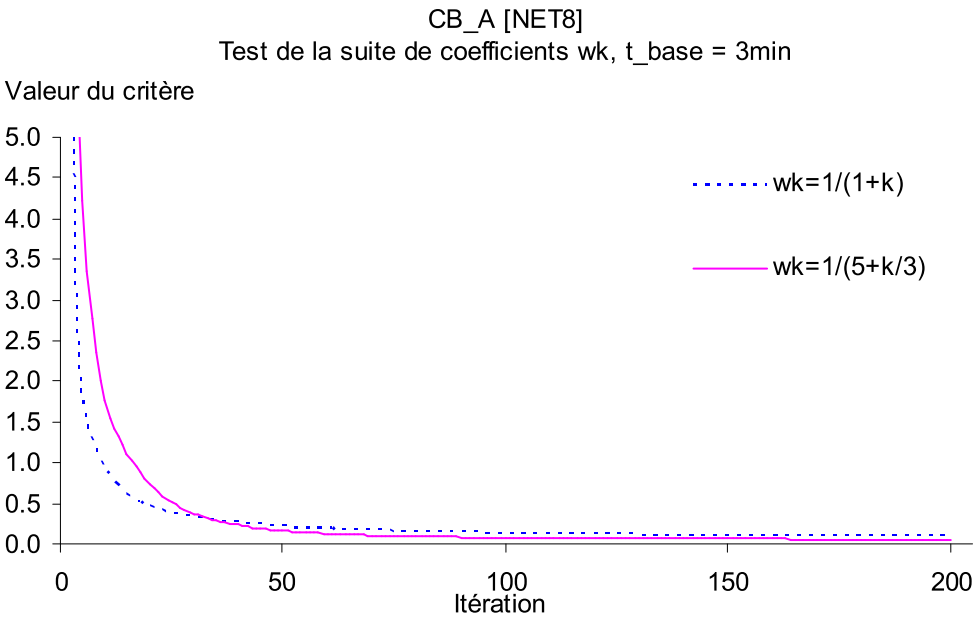


FIG. 9.11 – Evolution de CB pour la combinaison convexe par arcs.

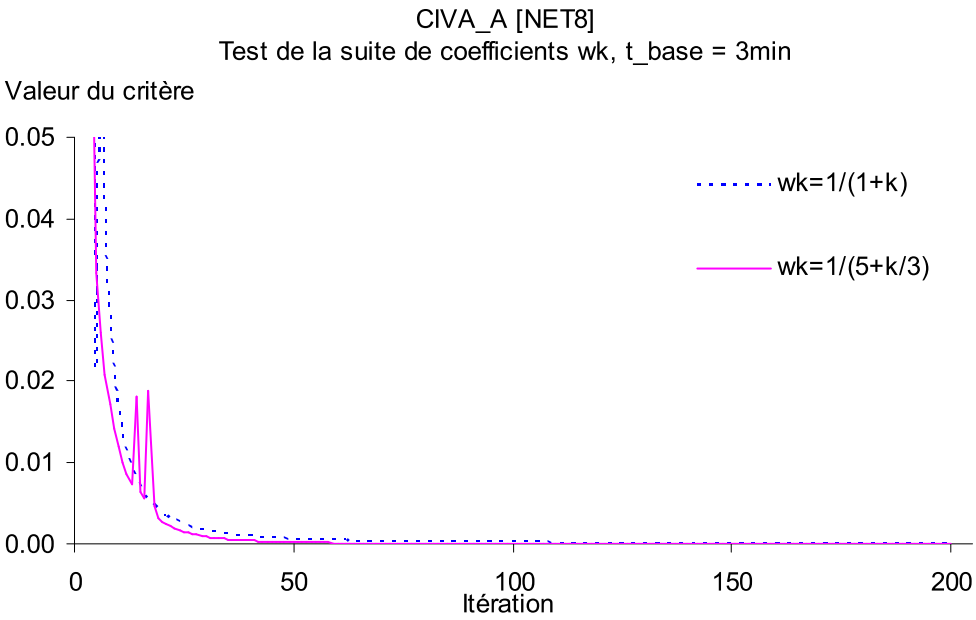


FIG. 9.12 – Evolution de CIVA pour la combinaison convexe par arcs.

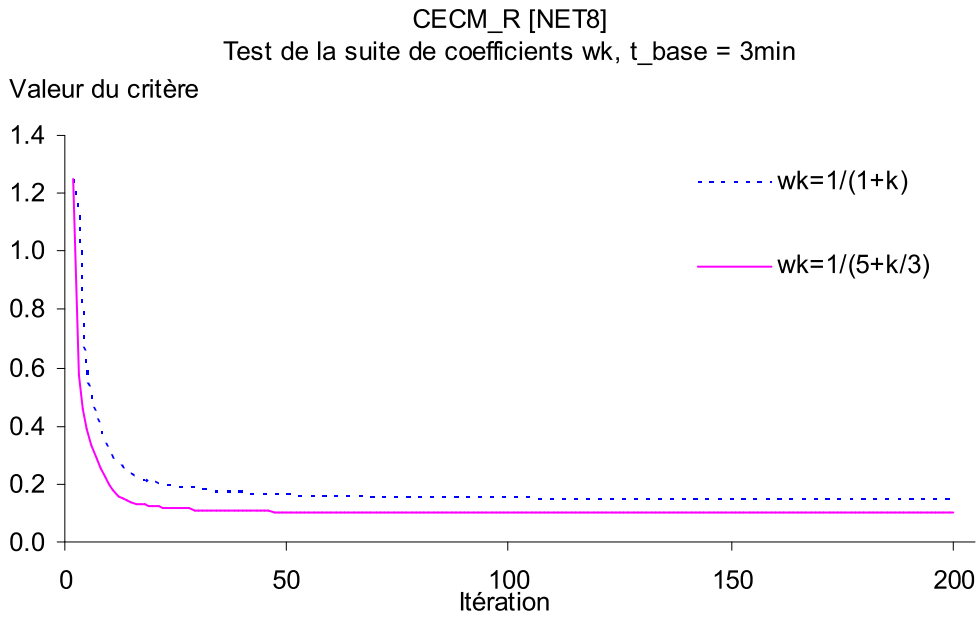


FIG. 9.13 – Evolution de CECM pour la combinaison convexe par chemins.

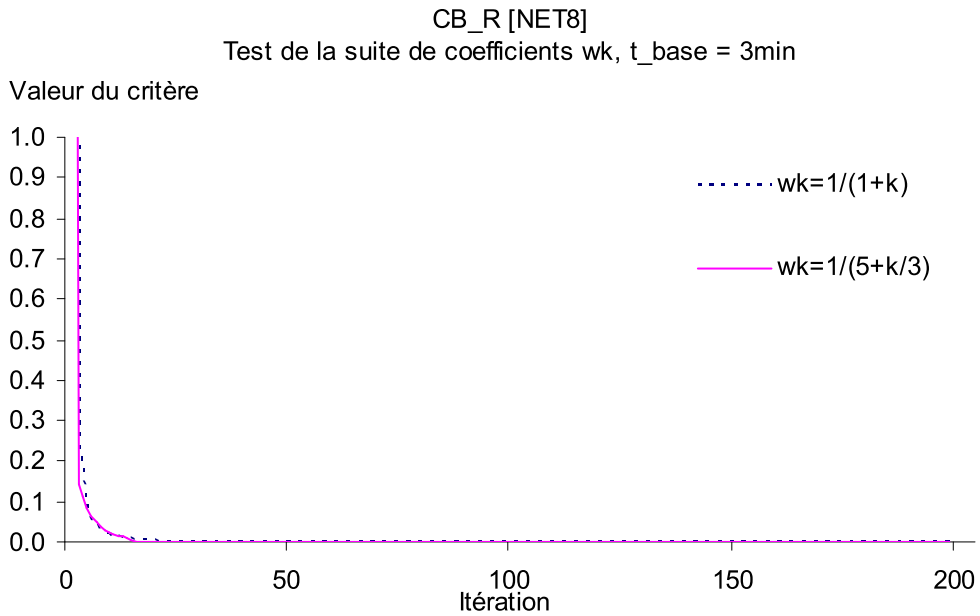


FIG. 9.14 – Evolution de CB pour la combinaison convexe par chemins.

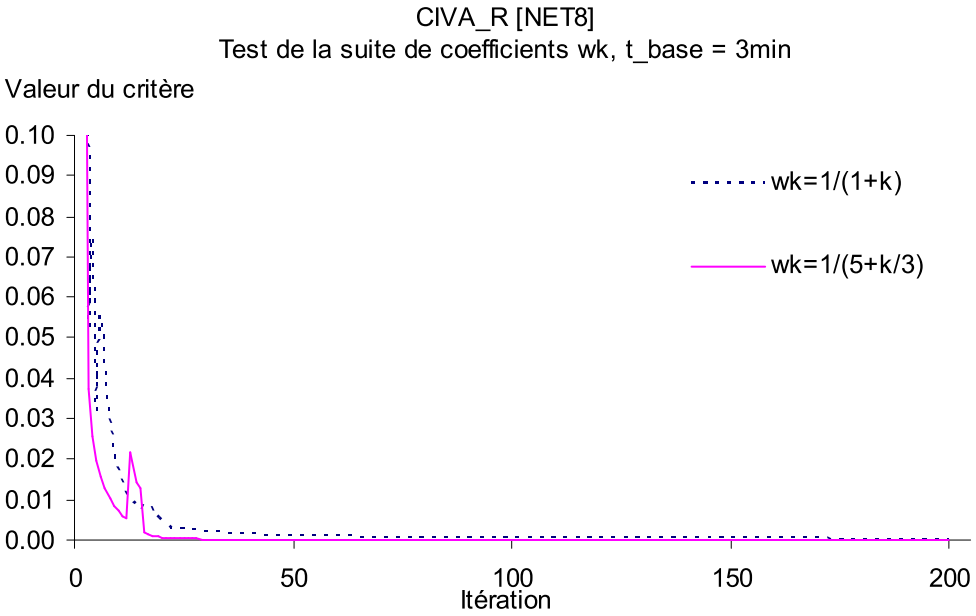


FIG. 9.15 – Evolution de CIVA pour la combinaison convexe par chemins.

9.2.6 Sur l'algorithme hybride

Le paramètre γ doit être suffisamment réduit pour procurer la convergence : les valeurs 0.1 et 0.05 conviennent pour le cas d'école, mais pas la valeur 0.2.

Pour la valeur 0.05, le critère CECM_H décroît régulièrement (figure 9.16). Le critère CBH_1 également, et sensiblement au même rythme (figure 9.17). CBH_2 aussi, plus rapidement encore (figure 9.18).

9.2.7 Synthèse

Nous avons constaté la convergence robuste des trois algorithmes sur le cas d'école, avec pour l'algorithme hybride la nécessité de fixer assez bas le paramètre γ. La convergence est robuste mais lente, plusieurs dizaines d'itérations sont nécessaires pour stabiliser les volumes des arcs.

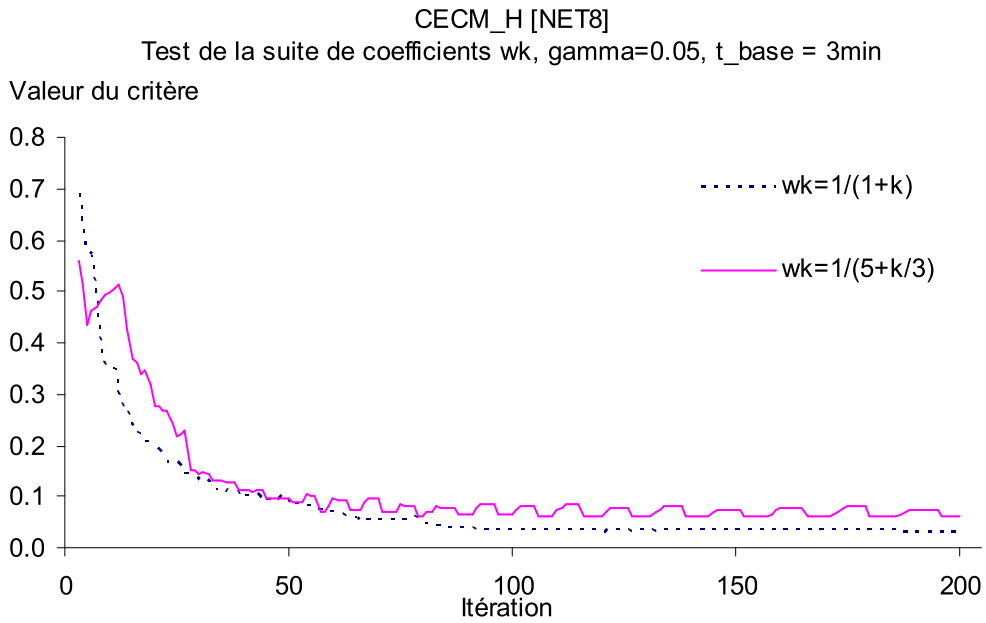


FIG. 9.16 – Evolution de CECM pour l'algorithme hybride.

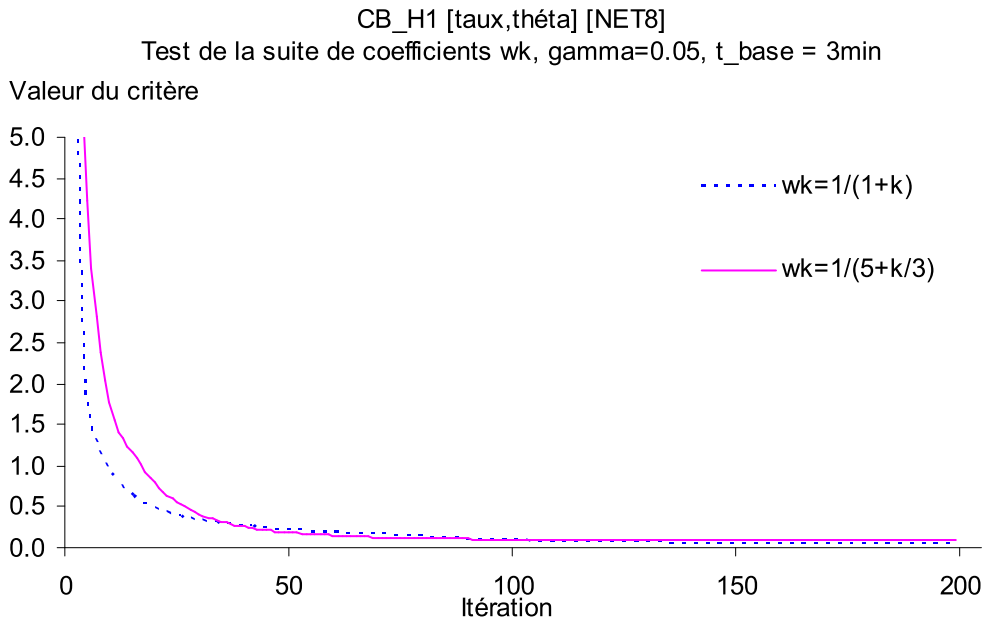


FIG. 9.17 – Evolution de CBH_1 pour l'algorithme hybride.

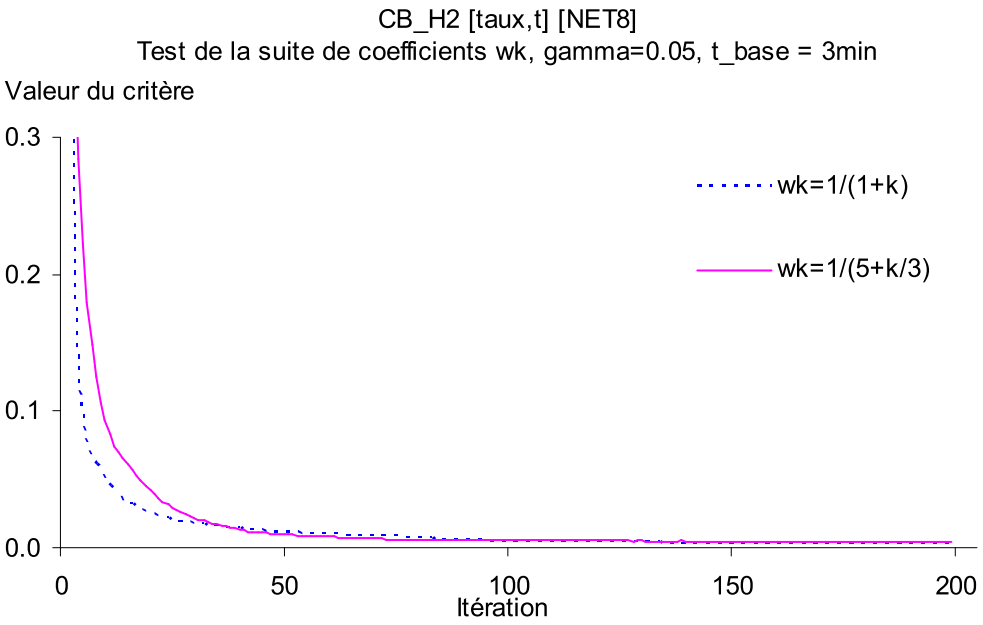


FIG. 9.18 – Evolution de CBH_2 pour l'algorithme hybride.

9.2.8 Sur la composition des critères de convergence

Nous avons constaté la nécessité théorique (8.2, 8.3) et l'utilité empirique (9.2) d'un critère global de convergence composé de plusieurs critères d'écart, au nombre de deux pour les méthodes de combinaison convexe. Dans l'analyse théorique nous avons exprimé le critère global de manière abstraite, générique en fonction des critères d'écart, en notant l'opérateur de combinaison par un signe de somme. Dans une application concrète, pour disposer d'un critère global synthétique, il convient de spécifier davantage la combinaison à opérer.

La combinaison la plus simple paraît une somme pondérée, avec des pondérations qui harmonisent les unités de compte, et aussi équilibrent les contributions de chaque critère d'écart. L'unité de compte harmonisée est a priori l'unité de coût total (ou de temps passé total), car le critère d'écart aux coûts minimum est exprimé dans cette unité et se prête mal à une transformation.

Afin d'exprimer en unité de coût l'écart entre deux variables de temps

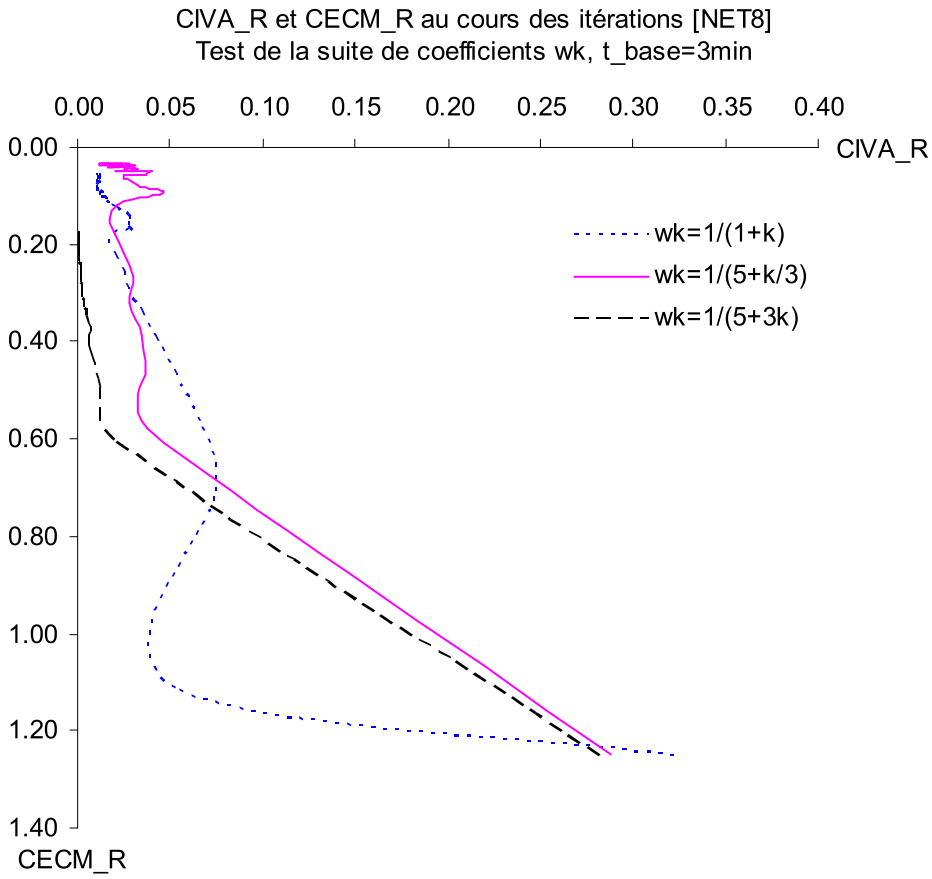


FIG. 9.19 – Décroissance conjointe du CIVA et du CECM.

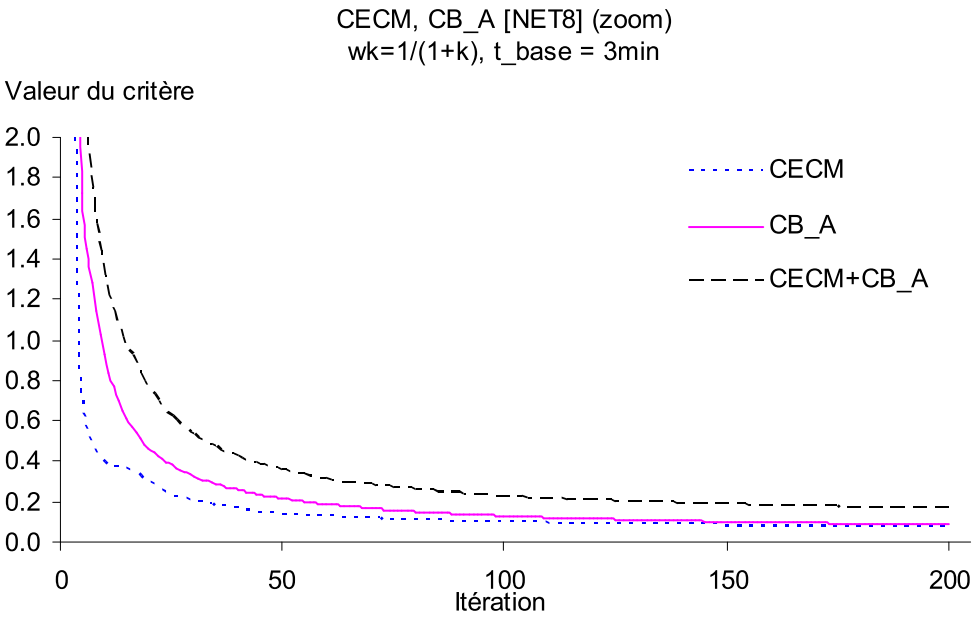


FIG. 9.20 – Totalisation des critères pour la CC par arcs.

par arcs, notées ici t_A et θ_A, nous proposons de mesurer

$$\text{CITA}_{bis} = \sum_{a \in A} \int_{h \in H} |t_a(h) - \theta_a(h)| \, dX_a^+(h)$$

qui est dans l'unité désirée, non-négatif, et ne s'annule qu'en l'absence de volume là où les temps diffèrent. Compte tenu des organigrammes des méthodes de combinaison convexe, ce critère ne convient qu'après l'itération initiale, quand t_A représente effectivement la congestion induite par un volume X_A^+ faisable.

La pondération entre les deux critères, CECM et CITA_{bis}, est à fixer dans chaque application à un réseau donné, en fonction de leurs valeurs respectives au cours d'une affectation choisie comme référence pour le calibrage.

La figure 9.20 illustre la superposition des critères d'écart qui composent l'écart global, pour la méthode des combinaisons convexes par arcs. La figure 9.21 concerne la méthode des combinaisons convexes par chemins, tandis que la figure 9.22 porte sur l'algorithme hybride.

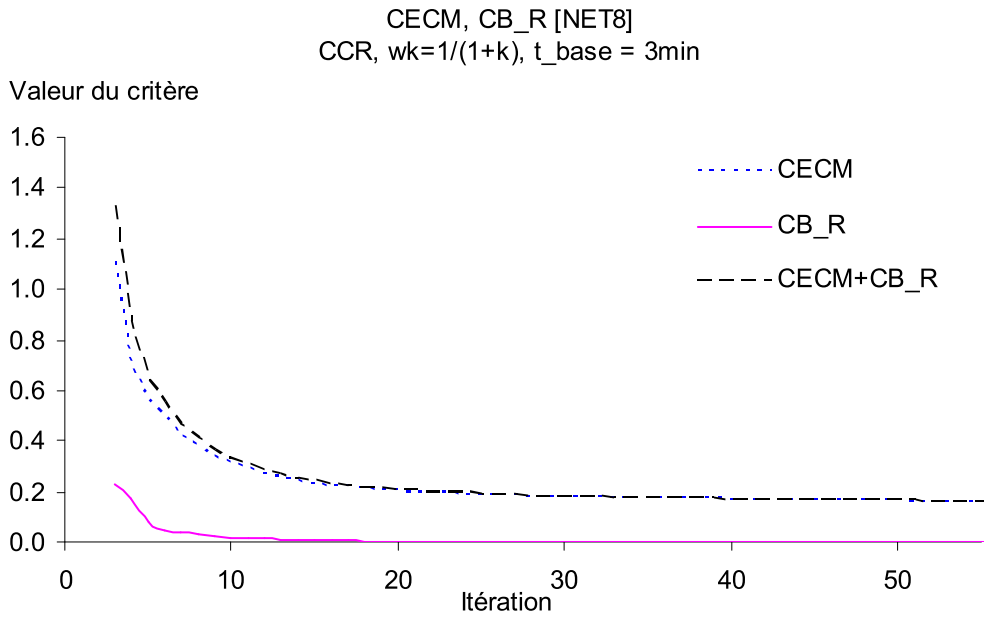


FIG. 9.21 – Totalisation des critères pour la CC par chemins.

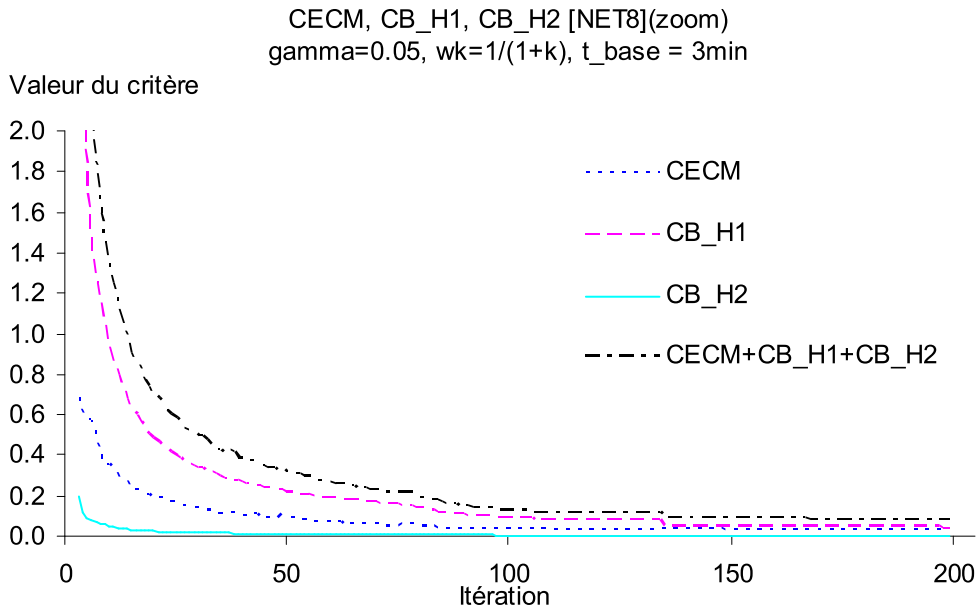


FIG. 9.22 – Totalisation des critères pour l'algorithme hybride.

Conclusion de partie

Récapitulation

Nous avons investigué l'équilibre dynamique du trafic en termes d'expression et de formulation analytique ; de mesurage d'équilibre et de convergence ; d'algorithmes de résolution.

Pour ce faire, nous avons d'abord résumé les formulations analytiques, les algorithmes d'équilibrage et les critères de convergence relatifs à l'affectation statique : celle-ci présente des relations simples entre le niveau local des arcs, et le niveau plus large des chemins, tant pour la formation des temps et des coûts, que pour la formation des flux. Dans l'affectation dynamique, les relations entre les arcs et les chemins sont plus complexes : le chargement d'un volume par chemin sur les arcs qui composent le chemin, met en IJuvre une chronologie de propagation. Or celle-ci interagit avec les effets de congestion, donc avec la formation du trafic sur le réseau, donc avec le choix d'itinéraire par les usagers selon leur comportement économique. Nous avons identifié, dénommé et caractérisé ces complexités spécifiques, et mis en évidence les surcoûts par rapport à l'affectation statique.

Nous avons alors recherché des formulations et des approches algorithmiques qui combinent la robustesse et l'économie. Au plan des formulations analytiques, nous avons examiné systématiquement les potentialités des variables de volume ou de temps, par arcs ou par chemins, pour servir de variables endogènes basiques dans le problème d'affectation dynamique exprimé par chemins. Puis nous avons exprimé l'équilibre en variables par arcs et destinations, et distingué les divers effets temporels en associant à chacun

d'eux des variables de temps par arcs dédiées. Nous avons examiné là aussi les critères de convergence et les approches algorithmiques. En particulier des approches primales-duales paraissent prometteuses. Ensuite nous sommes revenus aux formulations par arcs pour montrer que le volume par arcs reste lié à un volume par chemins sous-jacents (cohérence temporelle), ce que nous avons exploité pour obtenir des critères de convergence à la fois rigoureux et relativement économes.

Au plan des algorithmes d'équilibrage, nous avons évoqué brièvement des méthodes qui restent à développer : duales, primales-duales, primales autres que par combinaison convexe. Pour les méthodes primales de combinaison convexe, au niveau des arcs ou des chemins, nous avons établi des critères de convergence rigoureux et discuté les coûts informatiques. De plus, nous avons proposé un algorithme hybride qui traite conjointement des volumes et des temps par arcs.

Perspectives

Jusqu'à présent, seules certaines approches algorithmiques simplifiées ont permis de simuler l'affectation dynamique à l'équilibre pour des problèmes de grande dimension.

Nous pouvons citer à ce sujet, d'une part, l'approche LADTA que nous poursuivons et à laquelle s'apparente le modèle de Gentile *et al* [GMP05]. D'autre part, le modèle METROPOLIS de De Palma *et al.* [DPMN96] ; et d'autre part encore, le modèle DYNAMEQ de Florian *et al.* [MMN05].

D'autres approches voisines en formulation mais différentes dans le traitement algorithmique, ont été proposées qui paraissent plus coûteuses : notamment le traitement explicite des volumes par chemins, avec séparation ou intégration du chargement en volume et de l'écoulement du trafic. Cependant elles n'ont pas été appliquées à des problèmes de grande dimension, probablement en raison de leur coût élevé.

C'est pourquoi nous recommandons de développer des approches intermédiaires, susceptibles de combiner robustesse et économie de traitement : tout particulièrement, une formulation primale-duale, avec pour variables en-

dogènes basiques : des volumes par arcs et destinations, des temps par arcs, et, éventuellement, des potentiels par destination.

Conclusion générale

Récapitulation du travail

Au cours de cette thèse, nous avons constaté l'état de la recherche et l'état de la pratique en matière d'affectation dynamique du trafic : les modèles d'équilibre dynamique commencent à abonder dans la littérature scientifique mais sont encore très peu appliqués à des problèmes de grande dimension, qui sont réalisées par des équipes de recherche et non des bureaux d'ingénierie. Après cette approche externe, nous avons investigué l'affectation dynamique de l'intérieur, en reprenant, implémentant, expérimentant et finalement développant, le modèle LADTA porté par l'équipe de recherche ERMOD du LVMT.

Nous avons exposé les principes, et montré sur un cas d'école la réelle complexité d'abondance des traitements et des résultats : la notion d'instant critique s'est révélée précieuse pour analyser l'évolution de la congestion et repérer les intervalles temporels d'efficacité économique pour un itinéraire sur une relation origine-destination. Nous avons proposé une technique d'assimilation temporelle des morceaux de profil horaire, dont nous avons montré qu'elle ne compromet pas la convergence vers l'équilibre, et qu'elle améliore significativement le temps de calcul nécessaire pour atteindre un niveau donné de convergence.

Enfin, nous avons contribué à la théorie de l'affectation dynamique en développant une approche formalisée du mesurage de la convergence, et en concevant un algorithme "hybride" pour calculer l'équilibre dynamique du trafic.

Confrontation des résultats aux objectifs de connaissance

Face aux objectifs de connaissance posés dans l'introduction, le moment est venu de répondre en nous appuyant sur les résultats obtenus. Nous reprenons les quatre objectifs dans l'ordre de l'introduction :

1. un modèle d'affectation dynamique met en relation des hypothèses physiques relatives à la structure du réseau et à l'organisation de son exploitation, à la circulation et la propagation du flux, à l'écoulement local du flux, et à la formation du flux par agrégation des déplacements individuels ; et des hypothèses économiques relatives aux demandes des usagers entre les zones d'origine et les zones de destination par période d'horaire de départ, pour des usagers qui composent diverses classes de trafic (comportement de circulation) et divers segments de préférences (comportement économique), et peuvent prendre diverses décisions, notamment le choix d'itinéraire, éventuellement le choix d'horaire et d'autres choix encore. Ces caractères physiques et économiques sont traités selon le lieu et le temps, de manière géographique et chronologique. Ainsi la dimension chronologique constitue bien le caractère additionnel par rapport à l'affectation statique ; et de fait elle confère une complexité spécifique à l'affectation dynamique, par l'effet de "relativité chronologique".

2. sur l'effectivité d'une simulation, l'existence d'un équilibre solution de la formulation analytique est une propriété formelle connue (sous des hypothèses faibles de continuité des fonctions) : en revanche le mesurage de la convergence n'avait pas fait l'objet d'une analyse spécifique, que nous avons menée en donnant des critères de convergence rigoureux et relativement parcimonieux. Concernant l'efficience de la simulation, nous avons obtenu des résultats probants en nmobre d'itérations concernant des cas d'école ; ces résultats ont été confirmés sur une application de grande dimension traitée par Vincent Aguiléra, et rapportée en annexe.

3. Concernant les besoins du modèle en données d'entrée, nous avons indiqué les données nécessaires en présentant les applications. Au plan opérationnel, la principale complexité additionnelle par rapport à l'affectation statique nous paraît tenir à la constitution de matrices origine-destination temporalisées. Nous rapportons en annexe notre investigation à ce sujet. En effet, un traitement dynamique suppose une certaine finesse dans la description des profils temporels : primo dans le déroulement d'une période d'affectation disons la journée, ce qui introduit un premier facteur de complexité de l'ordre de 5-20 ; secundo pour représenter tous les jours d'une année il convient de distinguer des jours-types, ce qui introduit un second facteur de complexité de l'ordre de 5 à 20 : au total, l'utilisateur du modèle doit traiter non pas une matrice en moyenne annuelle mais un grand nombre, de l'ordre de 25-400, de matrices par période horaire.

4. concernant l'analyse des données en sortie du modèle, nous avons laissé cette question pour d'autres travaux, sauf en ce qui concerne les critères de convergence dont l'évaluation est cruciale ; nous avons discuté leur interprétation.

Perspectives

Compte tenu de nos résultats, il nous semble que la calculabilité est une question résolue au moins pour le modèle LADTA ; désormais l'interrogation principale concernant l'opérationnalité, porte sur l'ergonomie du modèle pour son utilisateur, sur l'accès interactif aux données et à des synthèses des données dont les formes restent largement à concevoir.

D'autres sujets de recherche sont en perspective :

– Sophistiquer, raffiner la représentation physique et économique. Leurent a formulé un modèle de choix d'itinéraire et d'horaire, avec des usagers segmentés par classes de circulation et par segment de préférence concernant l'horaire de départ et les arbitrages entre le prix et le temps de parcours [Leu03]. Ces aspects économiques, et des aspects physiques

relatifs à l'écoulement, restent à incorporer au simulateur et à expérimenter.

- Développer l'analyse formelle de l'équilibre dynamique ; la propriété d'existence est démontrée, mais l'unicité et la stabilité restent à investiguer en indiquant des conditions nécessaires et des conditions suffisantes, et en caractérisant les formes fonctionnelles pour lesquelles les "conditions classiques" s'appliquent (coercivité de l'inéquation variationnelle, contraction de l'application de point fixe etc).
- Développer les approches algorithmiques et les algorithmes d'équilibrage. L'affectation dynamique offre bien davantage de variables endogènes que l'affectation statique, et donc bien plus de "points d'entrée" pour une approche algorithmique, or jusqu'ici seules les bases de variables endogènes reprises de l'affectation statique ont été utilisées pour l'affectation dynamique, sauf notre algorithme hybride qui ainsi constitue une innovation pionnière. De plus, un algorithme d'équilibrage ne se réduit pas à la sélection de la base de variables endogènes : les méthodes utilisées jusqu'à présent, par combinaison convexe des volumes (ou des temps dans METROPOLIS), sont directement issues de l'affectation statique, mais les meilleurs algorithmes d'affectation statique n'ont pu être adaptés au cas dynamique.

Il y a là un large champ de développement.

Bibliographie

[Adm] Federal Highway Administration. Corsim/corflo, distribué par center for microcomputers in transportation (mctrans). consulté en 2004.

[AG] PTV Planung Transport Verkehr AG. Davisum. consulté en 2004.

[BADPK91] M. Ben-Akiva, A. De Palma, and I. Kaysi. Dynamic network models and driver information systems. *Transportation research A*, 25(5), 1991.

[BD00] L. Bloomberg and J. Dale. Comparison of vissim and corsim traffic simulation models on a congested network. *Transportation Research Record*, (1727), 2000.

[BGMP05] Giuseppe Bellei, Guido Gentile, Lorenzo Meschini, and Natale Papola. Macroscopic arc performance models with capacity constraints for within-day dynamic traffic assignment. *European Journal of Operational Research*, 2005. article in press.

[BGP05] G. Bellei, G. Gentile, and N. Papola. A within-day dynamic traffic assignment model for urban road networks. *Transportation Research, Part B*, 39(1), 2005.

[BLL96] C Buisson, J.P. Lebacque, and J.B. Lesort. Strada, a discretized macroscopic model of vehicular traffic flow in complex networks based on the godunov scheme. In *Proc. of the CESA 96 Symposium on Modelling, Analysis and Simulation*, Lille, France, 1996.

[BMW56] M. Beckmann, C.B. McGuire, and C.B. Winsten. Studies in the economics of transportation. *Yale University Press*, 1956.

[Cal00] Caliper Corporation, 1172 BEACON Stress, NEWTON MA, USA. *TransCAD User's Guide*, 2000.

[CC95] E. Cascetta and G.E. Cantarella. Dynamic processes and equilibrium in transportation networks : towards a unifying theory. *Transportation Science*, 29(4), 1995.

[CD96] G.D.B. Cameron and C.I.D. Duncan. Paramics : Parallel microscopic simulation of road traffic. *J. Supercomputing*, 1(10), 1996.

[Coh80] Guy Cohen. Auxiliary problem principle and decomposition of optimization problems. *Journal of Optimization Theory and Applications*, 32(3), 1980.

[Coh88] Guy Cohen. Auxiliary problem principle extended to variational inequalities. *Journal of Optimization Theory and Applications*, 59(2), 1988.

[Cor] Caliper Corporation. Transcad. `http://www.caliper.com`. Consulté en 2004.

[Dag94] C.F. Daganzo. The cell transmission model : A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research B*, 28B(4), 1994.

[Dag95] C.F. Daganzo. The cell transmission model, part ii : Network traffic. *Transportation Research B*, 29B(2), 1995.

[DPMN96] A. De Palma, F. Marchal, and Y. Nesterov. Metropolis : a modular system for dynamic traffic simulation. In *Proc. of the 76th TRB annual conference.*, TRB, Washington DC, 1996.

[Dre69] S.E. Dreyfus. An appraisal of some shortest-path algorithms. *Operations Research*, 17, 1969.

[FB93] J. Ferrer and J. Barcelo. Aimsun2 : Advanced interactive microscopic simulator for urban and non-urban networks, internal

report. Technical report, Departamento de Estad istica e Investigacion Operativa, Facultad de Informatica, Universitat Polit tecnica de Catalynya, 1993.

[FBST01] T.L. Friesz, D. Bernstein, Z. Suo, and R.L. Tobin. Dynamic network user equilibrium with state-dependent time lags. *Network and Spatial Economics*, 1, 2001.

[Fel94] M. Fellendorf. Vissim : Ein instrument zur beurteilung verkehrsabhängiger steuerung. In *Tagungsband zum Kolloquium Verkehrsabhängige Steuerung am Knotenpunkt*, 1994.

[FM06] T.L. Friesz and R. Mookherjee. Solving the dynamic network user equilibrium problem with state-dependent time shifts. *Transportation Research B*, 40(3), 2006.

[Fuk84] M. Fukushima. On the dual approach to the traffic assignment problem. *Transportation Research B*, 18(3), 1984.

[Gip81] P.G. Gipps. A behavioural car-following model for computer simulation. *Transportation Research, Part B*, 15(2), 1981.

[Gip86] P.G. Gipps. Model for the structure of lane-changing decisions. *Transportation Research, Part B*, 20(5), 1986.

[GMP05] Guido Gentile, Lorenzo Meschini, and Natale Papola. Macroscopic arc performance models with capacity constraints for within-day dynamic traffic assignment. *Transportation Research B*, 39, 2005.

[Hid05] Peter Hidas. A functional evaluation of the aimsun, paramics and vissim microsimulation models. *Road and Transport Research*, 14(4), 2005.

[HRBR82] P.B. Hunt, D.I. Robertson, R.D. Bretherton, and M. C. Royle. The scoot on-line traffic signal optimization technique. *Traffic Engineering and Control*, 23, 1982.

[HS06] Habib Haj Salem. Network-wide ramp metering, delivrable d3.3, euramp project, Juin 2006.

[Jan91] B.N. Janson. Dynamic traffic assignment for urban road networks. *Transportation Research, part B*, (25), 1991.

[Jok66] H. Joksch. The shortest route problem with constraints. *Journal of Mathematical Analysis and Applications*, 14, 1966.

[KA] Inc. KLD Associates. Watsim. `http://www.kldassociates.com`. date de consultation en 2004.

[KA93] Masao Kuwahara and Takashi Akamatsu. Dynamic equilibrium assignment with queues for a one-to-many od pattern. *TRANSPORTATION AND TRAFFIC THEORY*, 1993.

[KA97] M. Kuwahara and T. Akamatsu. Decomposition of the reactive dynamic assignments with queues for a many-to-many origin-destination pattern. *Transportation Research, Part B : Methodological*, 31B(1), 1997.

[KA01] M. Kuwahara and T. Akamatsu. Dynamic user optimal assignment with physical queues for a many-to-many od pattern. *Transportation Research Part B*, 35(5), 2001.

[LBLM96] J.P. Lebacque, C. Buisson, J.B. Lesort, and H. Mongeot. The strada model for dynamic assignment. In *Proc. of the Third World Congress on Intelligent Transport Systems*, 1996.

[Lec02] L. Leclercq. *Modélisation dynamique du trafic et applications à lŠestimation du bruit routier*. PhD thesis, INSA, 2002.

[Leu95] Fabien Leurent. Une boîte-à-outils pour opérer diverses affectations statiques du trafic routier. Technical Report 198, INRETS, Arcueil, France, 1995.

[Leu98] Fabien Leurent. Sensitivity and error analysis of the dual criteria traffic assignment model. *Transportation Research. Part B*, 32(3), 1998.

[Leu01] Fabien Leurent. *Modèles désagrégés du trafic*. Les collections de l'INRETS, 2001.

[Leu03] Fabien Leurent. On network assignment and supply-demand equilibrium : an analysis framework and a simple dynamic model. In *Proc. of the 2003 European Transport Conference*, 2003.

[Leu04] Fabien Leurent. The multiclass flowing problem in a dynamic traffic assignment model. In *Proc. of the 2004 European Transport Conference*, 2004.

[Leu05] Fabien Leurent. The network cost of congestion : analysis and computation of marginal social cost disaggregated by o-d pair and departure time. In *Proc. of the 2005 European Transport Conference*, 2005.

[Leu06] Fabien Leurent. *Structures de réseau et modèles de cheminement*. Tec et Doc. Editions Lavoisier, Cachan, France, 2006.

[LH05] Yongtaek Lim Lim and Benjamin Heydecker. Dynamic departure time and stochastic user equilibrium assignment. *Transportation Research B*, 39, 2005.

[LMQW03] P.A. Lotito, E. Mancinelli, J.-P. Quadrat, and L. Wynter. The traffic assignment toolboxes of scilab. Technical Report 0281, INRIA, 2003.

[LP94] T. Larsson and M. Patriksson. An augmented lagrangean scheme for capacitated traffic assignment problems. In *Proceedings of the 2nd meeting of the EURO Working Group on Urban Traffic and Transortation*, 1994.

[LTB78] D. R. Leonard, J. B. Tough, and P. C. Baguley. Contram : A traffic assignment model for predicting flows and queues during peak periods. *Transport and Road Research Laboratory (Great Britain), TRRL Report*, (841), 1978.

[LW55] M.J. Lighthill and G.B. Whitham. On kinematic waves : II. a theory of traffic flow on long crowded roads. *In Proceedings of the Royal Society*, A229, 1955.

[May64] A.D. May. Experimentation with manual and automatic ramp control. *Highway Research Record*, (59), 1964.

[MMN05] Florian M., Mahut M., and Tremblay N. A simulation-based dynamic traffic assignment model : Dynameq. Technical report, CRT, University of Montreal and INRO Montreal, QC, Canada, 2005.

[MN78] D.K Merchant and G.L Nemhauser. A model and an algorithm for the dynamic traffic assignment problems. *Transportation Science*, 12(3), 1978.

[MS81] Hani Mahmassani and Yosef Sheffi. Using gap sequences to estimate gap acceptance functions. *Transportation Research, Part B.*, 15(3), 1981.

[MTI] University of Maryland Maryland Tranportation Initiative. Dynasmart X et dynasmart P. consulté en 2004.

[New93] G.F. Newell. A simplified theory of kinematic waves in highway traffic, part I : general theory. part II : queuing at freeway bottlenecks. part III : multi-destination flows. *Transportation Research Part B : Methodological*, 27, 1993.

[NR01] Kai Nagel and Marcus Rickert. Parallel implementation of the transims micro-simulation. *Parallel Computing*, 27(12) :1611–1639, 2001.

[NSP+97] K. Nagel, P. Stretz, M. Pieck, S. Leckey, R. Donnelly, and C.L. Barrett. Transims traffic flow characteristics. Technical Report LA-UR-97-3530, Los Alamos National Laboratory report, 1997.

[Pap88] Markos Papageorgiou. Modelling and real-time control of traffic flow on the southern part of boulevard périphérique. Technical report, INRETS, Arcueil, France, 1988. Rapport de Contrat INRETS.

[Pat94] M. Patriksson. The traffic assignment problem, models and methods. *VSP*, 1994.

[PAY71] HAROLD J PAYNE. Models of freeway traffic and control. *MATHEMATICAL MODELS OF PUBLIC SYSTEMS*, 1971.

[Pay79] H.J. Payne. Freflo : A macroscopic simulation model for freeway traffic. *Transportation Research Record*, 722, 1979.

[PBHS90] M. Papageorgiou, J.M. Blosseville, and H. Haj-Salem. Modelling and real-time control of traffic flow on the southern part of boulevard peripherique in paris : Part ii : Coordinated on-ramp metering. *Transportation Research Part A*, 24(5), 1990.

[PH71] I. Prigogine and R. Herman. *Kinetic Theory of Vehicular Traffic*. Elsevier, 1971.

[Pip53] Louis A. Pipes. An operational analysis of traffic dynamics. *Journal of Applied Physics*, 24(3), 1953.

[PNDK02] Srinivas S. Pulugurtha, Shashi S. Nambisan, Mukund Dangeti, and Mohamed Kaseko. Simulating and analyzing incidents using corsim and vissim traffic simulation software. *Proceedings of the International Conference on Applications of Advanced Technologies in Transportation Engineering*, 2002.

[Pol67] BT Polyak. A general method for solving extremal problems. *DAN SSSR*, 174(1), 1967.

[PTV02] PTV Planung Transport Verkehr AG, StumpfstraSSe 1, D-76131 Karlsruhe, Allemagne. *Manuel de lŠutilisateur DAVISUM, Version 7.5*, 2002.

[PW99] Panos D. Prevedouros and Yuhao Wang. Simulation of large freeway and arterial network with corsim, integration, and watsim. *Transportation Research Record*, 1678, 1999.

[PZ01] S. Peeta and A. Ziliaskopoulos. Foundations of dynamic traffic assignment : The past, the present and the futur. *Network and Spatial Economics 1 (3/4)*, 2001.

[RB96] B. Ran and D. Boyce. *Modeling Dynamic Transportation Networks*. Springer-Verlag, 1996.

[RSHC00] B. Ran, M.S. Shin, R.R. He, and K. Choi. Introducing platoon dispersion into an analytical dynamic assignment process. *Transportation Research Records*, 1733, 2000.

[SDD95] Mark Smith, Gordon Duncan, and Stephen Druitt. Paramics : Microscopic traffic simulation for congestion management. *IEE Colloquium (Digest)*, (39), 1995.

[SE94] H. Haj Salem and N. Elloumi. METACOR : A macroscopic simulation model of corridor traffic, and its calibration and validation at the corridor périphérique in paris. In *Congrès Mondial, ATT Conférence, Paris*, 1994.

[SP98] H. Haj Salem and M. Papageorgiou. METACOR : A dynamic macroscopic simulation tool for corridor traffic. In *CESA'98-IEEE, Nabeul, Tunisia*, Avril 1998.

[Spi84] H. Spiess. *Contributions à la théorie et aux outils de planification de réseaux de transport urbain.* PhD thesis, Département d'informatique et de recherche opérationnelle, Centre de recherche sur les transports, Université de Montréal, 1984.

[SRHC00] Michael S. Shin, Bin Ran, Rachel R. He, and Keechoo Choi. Introducing platoon dispersion into an analytical dynamic assignment process. *Transportation Research Records*, 2000.

[TBM] Transport, Road Research Laboratory (Great Britain), and Mott MacDonald. Contram : Continuous traffic assignment model. date de consultation en 2004.

[TW00] C.O. Tong and S.C. Wong. A predictive dynamic traffic assignment model in congested capacity-constrained road networks. *Transportation Research B*, 34, 2000.

[VS04] H.R. Varia and Dhingra S.L. Dynamic user equilibrium traffic assignment on congested multidestination network. *Journal of Transportation Engineering*, 130(2), 2004.

[VV82] D. Van Vliet. Saturn : a modern assignment model. *Traffic Engineering and Control*, 23, 1982.

[War52] J.G. Wardrop. Some theoretical aspects of road traffic research. *Proceedings of the Institute of Civil Engineers, Part II 1*, 1952.

[WP98] Yuhao Wang and Panos D. Prevedouros. Comparison of integration, tsis/corsim, and watsim in replicating volumes and speeds on three small networks. *Transportation Research Record*, 1644, 1998.

[WTC02] Byung-Wook Wie, Roger L. Tobin, and Malachy Carey. The existence, uniqueness and computation of an arc-based dynamic network user equilibrium formulation. *Transportation Research Part B : Methodological*, 36(10), 2002.

Annexe A

Logiciels de simulation dynamique du trafic routier

Logiciel	Caractéristiques	Auteur/distributeur	Système d'exploitation
Corsim [Adm, WP98, PW99, PNDK02, BD00]	Microscopique, stochastique, loi de poursuite, réseaux mixtes (urbains et autoroutiers).	Développé par FHWA. Vendu par PC Trans et MC Trans. Utilisé aux Etats-Unis	Windows 9x, NT, XP, depuis 30 ans, version 5.1 en 2003
Corflo	Macroscopique, stochastique, réseaux urbains et interurbain	Développé par FHWA. Vendu par PC Trans et MC Trans. Utilisé aux Etats-Unis.	Windows
Dynasmart [MTI]	Développé par l'équipe du professeur Mahmassani, université du Texas, Austin, financé par la FHWA. Vendu par PC Trans et MC Trans. Utilisé aux Etats-Unis.	Windows, dernière version 0.91 en 2001	

Logiciel	Caractéristiques	Auteur/distributeur	Système d'exploitation
Transims [NSP+97, NR01]	Microscopiques, régions métropoles.	Développé par Clean Air Act et l'ISTEA aux Etats-Unis. Utilisé aux Etats-Unis, Albuquerque et Portland.	
Watsim [KA, WP98, PW99]	Microscopique, autoroutes et réseaux de surface.	Développé par KLD et Associates, utilisé aux Etats-Unis.	Windows, Linux. Dernière version en 2001.
Contram [LTB78, TBM]			
Paramics [SDD95, CD96]	Microscopique, stochastique, réseau autoroutier et urbain de grande échelle spatiale.	Développé en Ecosse. Distribué par la société SIAS en Grande-Bretagne, par Quadstone Ltd. aux Etats-Unis. Utilisé également au Japon et en Allemagne.	
Vissim [BD00, Hid05]	Microscopique, réseau urbain et interurbain, notamment intersections multimodales.	Développé par la société PTV en Allemagne. Distribué en France par Isis.	
AIMSUN	Microscopique, réseau urbain et interurbain, intersections à feux.	Développé par l'université Polytechnique de Catalogne et DHV aux Pays-Bas. Utilisé en France.	

Logiciel	Caractéristiques	Auteur/distributeur	Système d'exploitation
Dynasim	Microscopique, réseau urbain, interurbain, autorourier.	Développé par la société ynalogic en France.	
Sitra	Microscopique, réseau urbain.	Développé par J. H. Gabard, issu d'un travail de recherche avec l'Obera-CSCD. Distribué en France par la société Sodit.	
Integration	Mésoscopique, voies rapides urbaines.	Développé par Michel van Aerde et al. au Canada. Distribué par la société Goudappel Coffeng.	

Annexe B

Liste des entretiens effectués

1. Isabelle CABANNE, CERTU/MOB/PMD
2. Christophe DAMAS, CERTU/SYS/GTT
3. Yoann LA CORTE, DREIF/DIT/Groupe Etudes
4. André de PALMA
5. Danye ABOKI, CETE Méditerr. /DISTOA/TRI
6. Jean-Baptiste FEREY, CETE Lyon/DVT/GDU Bron
7. Pascal GABET, CETE Nord Picardie/DVT/Depl
8. Muriel MARIOTTO, CETE Méditerr/DISTOA/TRU
9. Jean-Baptiste LESORT, ENTPE/LICIT

Annexe C

Tests du prototype

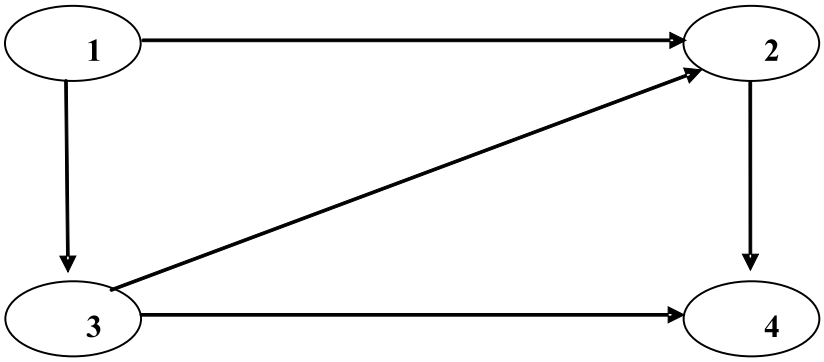


FIG. C.1 – Le réseau NET.

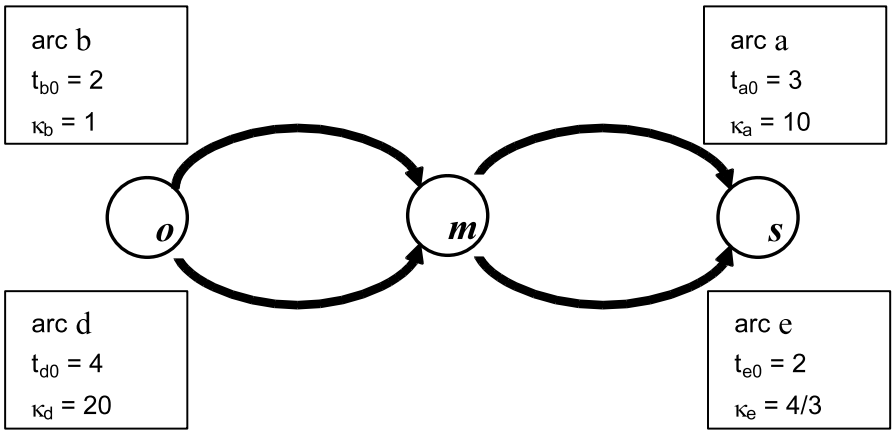


FIG. C.2 – Le réseau $NET8$.

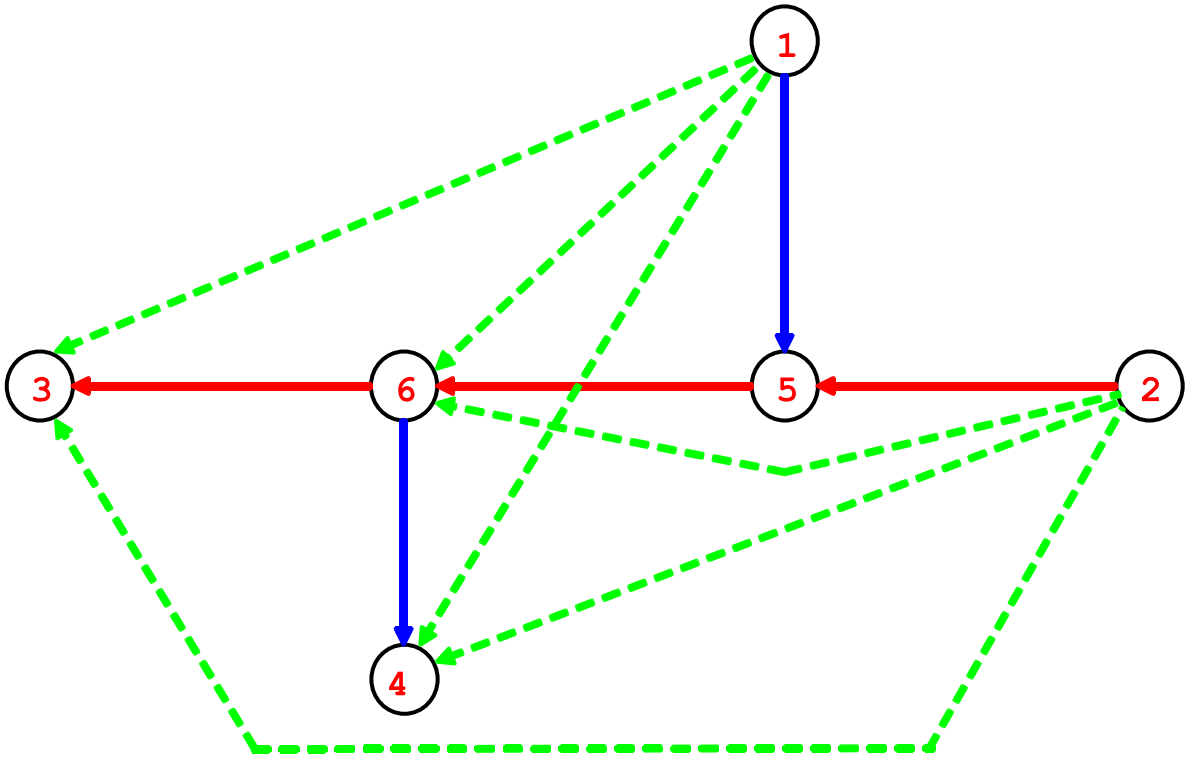


FIG. C.3 – Le réseau MLV.

Annexe D

Estimation de matrices O-D dynamiques

Annexe E

Test de la performance du simulateur

D.1 Introduction 286

D.2 Revue sur l'estimation de matrice statique 286

D.2.1 Inférence par entropie 288
D.2.2 Estimation d'un modèle de distribution 289
D.2.3 Exemple d'estimation d'une matrice O-D 291
D.2.4 Les diverses sources d'information possibles 293
D.2.4.1 Les comptages et les observations des plaques 294
D.2.4.2 Observation de couples origine-destination 294
D.2.4.3 Matrice a priori 295

D.3 Passage du cas statique au cas dynamique 295

D.3.1 La temporalisation de la propagation 295
D.3.2 Points importants du problème de temporalisation 296

D.4 Revue des méthodes dynamiques existantes 297

D.4.1 Approche de Bierlaire [BC02] 298
D.4.1.1 Les équations du modèle 298
D.4.1.2 Algorithmes appliqués 300
D.4.1.3 Commentaires du cas d'école 302
D.4.1.4 Comparaison entre deux méthodes : filtre de Kalman et LSQR 303
D.4.2 Approche de Van der Zijpp [VZ96] 303
D.4.2.1 Les équations d'estimation 303
D.4.2.2 Méthodes utilisées 305
D.4.2.3 Estimation avec multi-sources d'observation 306
D.4.3 Approche de Zhang et Mustard [ZM02] 307
D.4.4.1 Formulation du problème 307
D.4.4.1 Algorithme du modèle 308
D.4.5 Synthèse sur l'estimation dynamique 309

D.5 Estimateur de la matrice OD dynamique 311

D.5.1 Algorithme « Filtre de Kalman » 311
D.5.1.1 Hypothèses 311
D.5.1.2 Algorithme d'estimation 312
D.5.1.3 Explications supplémentaires 313
D.5.1.4 Algorithme de calcul de la matrice d'affectation 314
D.5.2 Programmation d'un estimateur 315
8.5.3 Test d'exemple numérique 315

Bibliographie 317

D.1 Introduction

La modélisation du trafic sur une aire d'étude comprend : (i) une représentation de l'offre de transport, (ii) une représentation de la demande de déplacement, (iii) une affectation de la demande sur le réseau. La modélisation de la demande est une étape centrale qui permet de fournir au modèle d'affectation la matrice origine-destination de déplacement. L'estimation est une méthode de constitution de la matrice OD de manière indirecte qui est plus économique que les mesures directes (par collage de papillons, relevé de numéros minéralogiques, et photographies aériennes, etc…). Pour le cas statique, il existe plusieurs méthodes indirectes pour l'estimation d'une matrice OD, par exemple : l'approche gravitaire, l'approche d'affectation par équirépartition, ou encore l'approche par entropie. Dans le cas dynamique, notons que la matrice OD est temporalisée, les contraintes pour l'estimation sont donc plus nombreuses.

Cette annexe aborde les méthodes existantes pour l'estimation d'une matrice OD dynamique. On recherche une méthode pour estimer une matrice O-D temporelle par des sources d'observation telles que des comptages sur les arcs, des matrices entre gares de péage autoroutier, des enquêtes origine-destination en divers points du réseau, une matrice statique, une matrice dynamique antérieure.

D'abord, une revue sur les méthodes d'estimation de matrice statique est présentée dans la section D.2 avec quelques tests simples. Puis, nous proposons une extension rapide au cas dynamique dans la section D.3. Pour l'état de l'art du problème d'estimation dynamique, nous avons analysé bibliographiquement quelques méthodes récentes dans la section D.4. Une de ces méthodes a été choisie, puis implémentée, en développant les algorithmes pour une version simple de l'estimateur. Ces derniers travaux sont présentés dans la section D.5.

D.2 Revue sur l'estimation de matrice statique

Nous considérons ici le problème d'estimation de la matrice O-D avec des informations de trafic, notamment les comptages sur les arcs du réseau. La production d'une matrice OD est un problème inverse de l'affectation de trafic, qui nous pose également un problème « sur-détermination ». En d'autre terme, le nombre des solutions du problème d'estimation est potentiellement infini pour une affectation donnée du trafic. Afin de le démontrer, un exemple simple est présenté suivant (cf. figure D.1) :

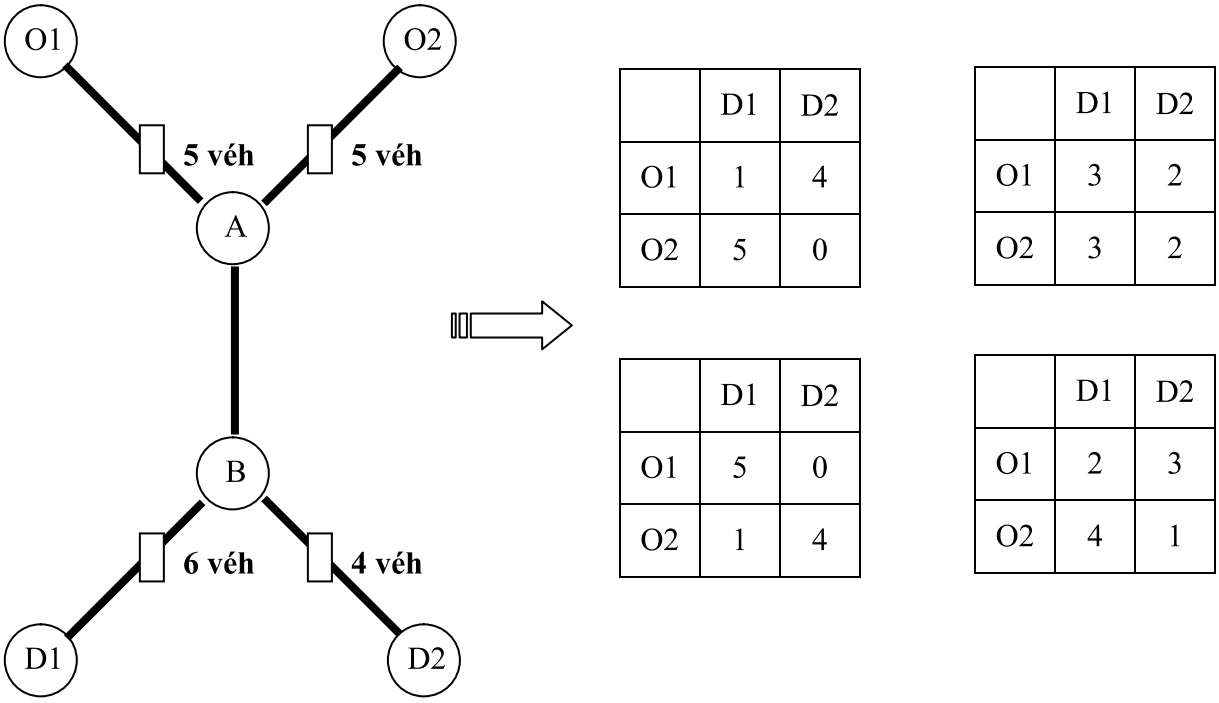


Figure D.1 Exemple d'illustration du problème « sur-détermination »

Afin de résoudre ce problème « sur-détermination », il est nécessaire d'avoir des stratégies de choix dans un ensemble de solutions possibles selon un principe de sélection. Le principe de sélection, c'est-à-dire la fonction objectif, mesure une ou des distances d'évaluation d'une solution par rapport des informations de référence.

Dans la littérature, un certain nombre de méthodes sont proposées, la plupart d'entre elles ont été développées d'une manière indépendante. Il y a deux méthodes considérées :

1. L'inférence d'une matrice OD par maximum d'entropie est utilisée dans la plupart des logiciels de planification. Cette méthode essaie d'inférer la matrice la plus vraisemblable qui soit compatible avec les informations disponibles (ex : comptages, les marges en émission et réception, matrice a priori…). La vraisemblance est fondée sur la notion d'entropie en théorie de l'information.

2. L'stimation d'un modèle économique de distribution spatiale, le plus souvent avec un modèle gravitaire. Dans ce groupe on suppose qu'un modèle de distribution est capable d'expliquer les déplacements sur un réseau. Le but de cette méthode d'estimation est de fixer les valeurs de paramètres dans un problème statistique bien posé : cadre probabiliste clair, davantage d'observations que de paramètres, et identifiabilité.

De plus, l'approche d'affectation par équirépartition tente de trouver une matrice qui soit à la fois compatible avec les données de comptage et le premier principe d'équilibre de WARDROP (estimation basée sur l'affectation à l'équilibre). Dans ce groupe, un ajout de condition portant sur les temps de parcoures est posé dans le processus d'estimation afin de bien connaître le choix d'itinéraire, puis la distribution des flux O-D.

D.2.1 Inférence par entropie

L'objectif de l'inférence d'une matrice OD est de :

- choisir une valeur vraisemblable compte tenu des informations disponibles
- Parmi plusieurs valeurs vraisemblables possibles : un principe de sélection, le maximum d'entropie.

Le modèle entropique est présenté rapidement ci-dessous :

- Objectif = exprimer des conditions d'optimalité pour soutenir la discussion économique
- Fonction objectif :

$$\max J(Q) \equiv -\sum_i Q_i \ln\left(\frac{Q_i}{\tilde{Q}_i}\right)$$

$$\text{Sous les contraintes } \begin{cases} Q_i \geq 0 \\ \sum_{i\in O} Q_i = E_o \leftarrow \alpha_o \\ \sum_{i\in D} Q_i = R_d \leftarrow \beta_d \\ \sum_i P_{a|i} Q_i = V_a \leftarrow \lambda_a \end{cases}$$

Par une analyse mathématique du modèle entropique, le résultat justifie l'algorithme multi-proportionnel :

- Début : $k \leftarrow 0,\ Q_i^{[0]} \leftarrow \tilde{Q}_i$

- Traiter les contraintes externes :

- $\forall a, o$ résoudre en λ_a : $\sum_i P_{a|i} Q_i \exp\left[-\lambda_a P_{a|i}\right] = V_a$ et

- $\forall i$ actualiser $Q_i \leftarrow Q_i \exp\left[-\lambda_a P_{a|i}\right]$¨

- Traiter les destinations :

- $\forall d$ calculer $f_d \leftarrow R_d \Big/ \sum_{i \in d} Q_i^{[2k+1]}$ et

- $\forall i \in d$ actualiser $Q_i^{[2k+2]} \leftarrow Q_i^{[2k+1]} f_d$

- **Test de convergence** : Si pas atteint, poser $k \leftarrow k+1$

Notations

i : relation O-D

Q_i : volume vrai par relation i.

$\widetilde{Q}_i$: volume O-D a priori (connaissance initiale).

o : zone d'origine, avec une émission E_o.

d : zone de destination, avec une réception R_d.

a : contrainte externe, avec un volume total V_a (ex. un arc observé).

$P_{a|i}$: proportion du volume Q_i dans le comptage a.

Dans l'application, concernant les contraintes, on distingue trois cas :

- simplement contraint : avec seulement les marges des colonnes de la matrice O-D
- doublement contraint : avec les marges en linge et en colonne
- prendre en compte les autres contraintes (ex. par classe de distance O-D)

D.2.2 Estimation d'un modèle de distribution

Cette famille estime les paramètres d'un modèle de distribution spatiale. On commence par donner une statistique à chaque comptage. Le principe est de relier les paramètres du modèle de distribution, aux flux sur les arcs, grâce à la contrainte de propagation.

Les équations d'observation du flux sur un arc a

$$V_a = X_a + \varepsilon_a$$

où V_a : volume observé, c'est une variable aléatoire. Si le résultat de l'observation équivaut au tirage dans une loi de Poisson de paramètre λ_v pendant durée τ dont la longueur est Δ_τ, alors l'espérance $E[V_a] = \lambda_a \Delta_\tau$ et la variance $Var[V_a] = \lambda_a \Delta_\tau$

ε_a : erreur d'observation, qui dépend de la méthode d'observation et qui fait l'objet de la qualification statistique.

$\lambda_a \Delta_\tau = E[V_a] = X_a + E[\varepsilon_a]$, a priori $E[\varepsilon_a]=0$, donc $\lambda_a \Delta_\tau = X_a$

$V[V_a] = V[\varepsilon_a] = \lambda_a \Delta_\tau$

X_a : Volume modélisé.

Celui-ci est obtenu grâce à la contrainte de propagation :

$$X_a = \sum_i P_{a|i} Q_i$$

Avec i : indice de relation O-D

$Q_i(\theta)$: volume modélisé par le modèle de distribution

θ : paramètre du modèle.

$P_{a|i}$: la part de marché de l'arc a sur l'O-D i, est connue soit par enquête, soit par un modèle d'affectation.

Si X_a est modélisé en véh/h noté x_a, l'équation d'observation donc devient :

$$V_a = \Delta_\tau . x_a + \varepsilon_a$$

Cette estimation est faite en minimisant un critère d'écart entre les flux par arcs respectivement modélisés et observés, de forme :

$$\min\left[\sum_a \frac{1}{Var(\varepsilon_a)}\left[V_a - X_a(Q_i)\right]^2\right] \qquad \text{(d.1)}$$

Il s'agit d'un problème d'estimation qui est bien posé dès que le nombre d'observations (non redondantes) dépasse le nombre de paramètres.

L'équation (d.1) nous fournit la première solution avec une hypothèse de Poisson. Ensuite, nous passons à la deuxième solution pour un modèle explicite quand on a des hypothèses plus précises sur les observations, dit solution par le maximum de vraisemblance. Le paramètre θ du modèle de distribution $Q_i(\theta)$ est obtenu en maximisant une fonction de vraisemblance (Log-vraisemblance), de forme :

$$\max \sum_a V_a \left(\ln\left(\frac{\lambda_a \Delta_\tau}{V_a}\right) - 1 \right) - \lambda_a \Delta_\tau$$

avec $\lambda_a \Delta_\tau = \sum_i P_{a|i} Q_i(\theta)$

D.2.3 Exemple d'estimation d'une matrice O-D

Les hypothèses

- La structure du réseau et les positions de comptage sont illustrées dans la figure D.2 :

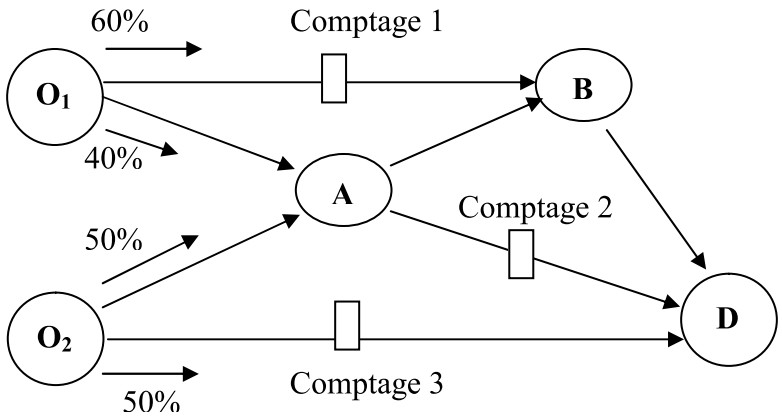


Figure D.2 Réseau d'exemple pour l'estimation d'une matrice O-D

- Les comptages

Arc	Flux Observé	σ (ε_a)	Pa|i
O_1 B : comptage 1	150	50	*60%* pour O_1D
O_2 D : comptage 2	200	60	40% pour O_1D ; 30% pour O_2D
A D : comptage 3	100	40	*50%* pour O_2D

Les résultats d'estimation

Les solutions de la matrice O-D : Q_{O1D} et Q_{O2D}. Les résultats d'estimation présentés ci-dessous est calculés grâce à l'outil « SOLVEUR » du logiciel EXCEL.

Méthode directe

A partir de comptage 1, on calcule directement le flux Q_{O1D} = (150/60%)= 250

A partir de comptage 3, le flux Q_{O2D} = (100/50%)= 200

Méthode sophistiquée

Q_{O1D} = 277,72849 et Q_{O2D} = 243,12323

Arc	Informations	

$$\frac{1}{Var(\varepsilon_a)}[V_a - X_a(Q_i)]^2$$

	Va	$X_a = \sum_i P_{a\|i} Q_i$	$\sigma(\varepsilon_a)$	
AD	200	184,0283671	40	0,159
O1B	150	166,6370963	50	0,111
O2D	100	121,561616	60	0,129
			Somme	**0,399** (min)

Estimation d'un modèle gravitaire par minimum d'écart entre les flux des arcs modélisés et observés

Equation du modèle gravitaire $\begin{cases} Q_{ij} = E_i R_j f(G_{ij}) \\ où \;\; f(G_{ij}) = G_{ij}^{-\theta} \end{cases}$

Emissions E_{O1} =	400	G_{01D} =	10
Emissions E_{O2} =	300	G_{02D} =	10
Réception R_D =	500	θ =	2,83635

Calculer Q_{ij} :

Q_{O1D}	291,52799
Q_{O2D}	218,64599

Puis on calcule les flux sur les arcs observés et la somme des écarts entre les observations et les estimations
Enfin, on estime le paramètre theta en minimisant l'écart par outil "solveur"

Arc	Informations			$\frac{1}{Var(\varepsilon_a)}[V_a - X_a(Q_i)]^2$
	Va	$X_a = \sum_i P_{a\|i} Q_i$	$\sigma(\varepsilon_a)$	
AD	200	182,2049914	40	0,197914
O1B	150	174,9167917	50	0,248339
O2D	100	109,3229948	60	0,024144
			Somme	**0,470397 (min)**

Estimation d'un modèle gravitaire par maximum de vraisemblance

Equation du modèle gravitaire $\begin{cases} Q_{ij} = E_i R_j f(G_{ij}) \\ où \;\; f(G_{ij}) = G_{ij}^{-\theta} \end{cases}$

Emissions E_{O1} =	400	G_{01D} =	10
Emissions E_{O2} =	300	G_{02D} =	10
Réception R_D =	500	θ =	2,851937463

Calculer Q_{ij} :

Q_{O1D}	281,25
Q_{O2D}	210,9375

On maximise la fonction de vraisemblance LV suivante pour obtenir le paramètre théta du

modèle gravitaire

Arc	Informations			LV
	Va	$X_a = \sum_i P_{a\|i} Q_i$	Var(ε_a)	
AD	200	175,7812505	40	-401,596658
O1B	150	168,7500005	50	-301,082545
O2D	100	105,4687503	60	-200,144299
			Somme	**-902,823502** (max)

Afin de représenter une très bonne détermination pour le paramètre Théta, une graphique dans la figure D.3 a été utilisée.

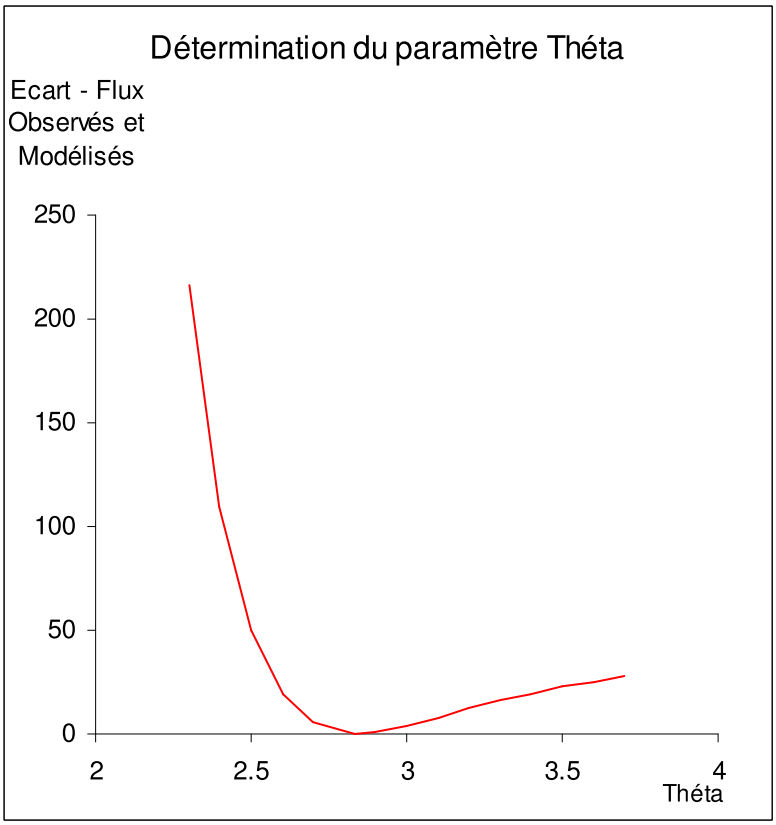


Figure D.3 illustration de la détermination du paramètre d'un modèle de distribution

D.2.4 Les diverses sources d'information possibles

Il y a plusieurs sources d'information possibles qui permettent de fournir des observations de trafic nécessaire à l'estimation des matrices O-D. Ces sources seront bien utilisées dans les estimateurs comme des inputs ou des contraintes afin d'obtenir des matrices estimées les plus proches de vraies matrices.

D.2.4.1 Les comptages et les observations des plaques

- Comptage par arc : les comptages peuvent être réalisés sur arc isolé dans le réseau. Si on n'intercepte pas les déplacements, on mesure uniquement le volume total par tranche de temps, sans connaître la structure par couples origine-destination. Les flux observés sur les arcs isolés sont les données les plus courantes.

- Comptage sur ligne-écran (coupure) : une série de comptages est effectués sur les coupures de manière à couvrir l'ensemble des flux. Les coupures seront décomposées selon de grands secteurs géographiques afin d'obtenir les informations globales de trafic permettant de mieux connaître la distribution de la demande sur les grands axes.

- Emission/réception de zone : cette information est connue par des enquêtes (ex. enquête ménage). Les émissions et les réceptions de zones seront utilisées successivement comme les marges en ligne et en colonne de matrice O-D. Observation sur la coupure et émission/réception de zone ont le même nature d'information, mais autre source.

- Porte d'observation des plaques : ces portes d'observation souvent équipées en entrée et en sortie du réseau permettent d'identifier les véhicules. De ces informations traitées, on peut connaître trajectoire des véhicules qui sera utilisée comme une importante information dans quelques méthodes d'estimation par multi-sources d'information.

D.2.4.2 Observation de couples origine-destination

- Enquête par interview, au bord de la route, sur les O-D. De ces informations, on obtient non seulement le volume de flux, mais la structure par couple O-D (la proportion des flux par O-D dans un comptage).

- Comptage par entrée – sortie (ex. réseau autoroute avec gare de péage) : ces comptages donnent les volumes sur l'entrée et la sortie du réseau, ou d'un sous-réseau. Dans ce cas, le volume V_{es} par chaque entrée /sortie est mesuré très précisément. Nous avons deux cas :

a. 1ère cas : si les origines et destinations sont les gares d'entrée-sortie, alors la matrice est bien mesurée.

b. 2ème cas : si on a le volume par chaque entrée/sortie avec des zones de demande. On suppose que chaque entrée/sortie est un arc fictif du réseau (entrée – nœud initial et sortie – nœud final), on obtient :

$X_{es} = \sum_i P_{es|i} Q_i$, avec $P_{es|i}$ est obtenu par un modèle d'affectation.

D.2.4.3 Matrice a priori

Cette matrice est une référence importante dans la plupart de méthode d'estimation. En général, cette matrice est déjà estimée antérieurement, puis elle est utilisée afin d'initialiser le processus d'estimation dans tous algorithmes proposés.

Elle est peut aussi être une matrice grossière, mesurée avec une précision faible, par exemple : une matrice OD tirée d'une enquête « ménages ».

D.3 Passage du cas statique au cas dynamique

D.3.1 La temporalisation de la propagation

En regard du cas statique, nous avons utilisé l'équation de propagation au niveau du volume par OD, ce qui exige le calcul de la proportion $P_{a|i}$. Cette proportion peut être déterminée facilement grâce à un modèle statique ou à un enquête au bord de la route.

Dans le cas dynamique, nous faisons un passage de la 2ème famille du cas statique au cas dynamique, en prenant en compte la temporalisation : période de départ au niveau de chaque O-D, et la proportion du flux O-D écoulé sur l'arc considéré selon l'heure de départ. Tout cela nous donne la difficulté à calculer des proportions à chaque instant $P_{a|i,\bar{H}_{ai}(h)}$. Pour éviter ces calculs, l'équation de propagation sera reformulée au niveau des chemins au lieu des couples O-D. Notons que les contraintes de propagation sont plus complexes car elles sont temporalisées. Par chaque arc observé, chaque variable devient une fonction du temps, l'équation de propagation du modèle de la 2ème famille est reformulée :

- la propagation dans le modèle :

$$\Delta X^+_{a\tau} = X^+_a(h2) - X^+_a(h1) = \sum_i \sum_{r \in i \wedge a} (X^+_r(\bar{H}_{r>a}(h2) - X^+_r(\bar{H}_{r>a}(h1)) \qquad \text{(d.2)}$$

- observation dynamique :

$$\Delta V^+_{a\tau} = \Delta X^+_{a\tau} + \varepsilon_{a\tau} \qquad \text{(d.3)}$$

où τ: indice de la durée d'observation, $\Delta_\tau = h2 - h1$, on suppose que chaque arc est observé pour T durées successives, $\tau = 1..N$.

$\Delta V^+_{a\tau}$: volume entrant observé pendant τ

$X^+_a(h)$: volume entrant cumulé modélisé à l'instant h sur l'arc a.

$\overline{H}_{r>a}(h)$: heure de départ pour chemin $r>a$ en amont de l'arc a.

$X_r^+(\overline{H}_{r>a}(h))$: volume entrant cumulé sur chemin r à l'instant $\overline{H}_{r>a}(h)$.

$\varepsilon_{a\tau}$: erreur d'observation pendant τ.

Pour simplifier le problème, nous supposons que les temps de parcours sont connus à l'avance.

La solution du problème est les paramètres estimés d'un modèle de distribution qui modélise $X_r^+(\overline{H}_{r>a}(h))$, en minimisant un critère d'écart entre les flux par arcs respectivement modélisés et observés. Ce critère est reformulé suivant :

$$\min\left[\sum_a\sum_\tau \frac{1}{Var(\varepsilon_{a\tau})}\left[\Delta V_{a\tau}^+ - \Delta Z_{a\tau}^+\right]^2\right] \qquad \text{(d.4)}$$

sous le contrainte $X_r^+(\overline{H}_{r>a}(h2)) \geq X_r^+(\overline{H}_{r>a}(h1)) \geq 0$, $h2>h1 \ \forall h$.

Pour 2ème solution, la fonction de log-vraisemblance est reformulée :

$$\max \sum_a\sum_\tau \Delta V_{a\tau}^+\left(\ln(\frac{\lambda_a\Delta_\tau}{\Delta V_{a\tau}^+})-1\right)-\lambda_a\Delta_\tau \qquad \text{(d.5)}$$

sous le contrainte $X_r^+(\overline{H}_{r>a}(h2)) \geq X_r^+(\overline{H}_{r>a}(h1)) \geq 0$, $h2>h1 \ \forall h$.

avec $\lambda_a\Delta_\tau$ est déterminé par équation de propagation (d.2)

Le volume par OD sera obtenu par l'équation suivante :

$$Q_i(h) = \sum_{r\in i} X_r^+(h) \qquad \text{(d.6)}$$

D.3.2 Points importants du problème de temporalisation

La temporalisation des variables nous donne des difficultés dans l'estimation des matrices dynamiques :

- Les variables et les contraintes plus nombreuses.
- Coefficients $P_{a|i,\overline{H}_{ai}(h)}$ à calculer par un modèle d'affectation dynamique.
- Problème d'interaction temporelle entre les observations des différentes heures pour une route donnée.

Ces questions délicates sont à « poser » à nos références bibliographiques sur les méthodes d'estimation dynamique. De ces raisons, nous allons faire une étude sur des méthodes d'estimation dynamique existantes proposées par différentes auteurs dans la partie suivante.

D.4 Revue des méthodes dynamiques existantes

Les matrices temporelles de flux O-D est un des inputs nécessaires pour la simulation dynamique du trafic. L'Estimation et la prédiction dynamique de ces matrices O-D nous pose un travail difficile à réaliser en pratique avec des diverses sources de données toujours incomplètes et insuffisances. Plusieurs méthodes dynamiques en théorie ont été proposées depuis des années 80 : Okutani et Stephanades en 1984 [OS84] ; Kremer et Keller en 1987 [KK87]; Nihan et Davis en 1987 [ND87] ; Bell en 1991 [Bell91]. Et puis, des méthodes opérationnelles ont été développées et mises en pratique dans les années 90 et nos jours, grâce aux recherches réalisées par divers auteurs : Ashok et Ben-Akiva en 1993 [ABA93] ; Bierlaire et Crittin 2002 [BC02] ; Van der Zijpp en 1996 [VZ96]; Chang et Wu en 1994 [CW94]; Cascetta et al., 1993 [CIM93] ; Zhang et Mustard en 2002 [ZM02] ; Wu et Davis en 2004 [WD04] ; Bell et al. en 2001 [YMB01]; etc.) . Ces méthodes utilisent de manière très efficace des sources de données possibles fournies par des observations et des données historiques afin d'obtenir une matrice la plus proche à réaliste.

En manière d'information fournie, on peut distingue l'estimation dynamique des matrices O-D en deux groupes : (a) pour « réseau fermé » et (b) pour « réseau ouvert ». Dans le cas « réseau fermé », les informations sont valables sur tous les entrées et les sorties du réseau et pour touts les points de temps. On peut citer des approches récentes au cas « réseau fermé » telles que [Bell91], [CW94], [CT96] et [VZ96]. Toutes ces approches essaient d'estimer la proportion du volume pour chaque O-D par chaque tranche de temps, en supposant que les informations sont complètes en entrée et en sortie du réseau. Les méthodes pour « réseau ouvert » telles que [ABA93], [BC02], [CIM93] et [WD04] s'appuient sur la réflexion du volume de demande des couples O-D à l'instant de départ sur les flux des arcs à l'instant observé par un paramètre d'affectation.

Dans le cadre de la thèse, effectivement, nous ne feront une analyse bibliographique que sur quelques méthodes récentes. De ces méthodes, nous souhaitons d'obtenir une base solide pour des applications numériques de la thèse.

D.4.1 Approche de Bierlaire [BC02]

En 1993, Ashok et Ben Akiva [ABA93]ont proposé un organigramme pour l'estimation dynamique des matrices O-D. Le problème est formulé par l'approche filtre de Kalman , où le vecteur d'état consiste en les déviations du flux O-D par rapport aux résultats a priori basés sur les données historiques. Grâce à l'ensemble de comptages obtenus sur les arcs à la fin de chaque intervalle, les déviations du flux O-D prévues à cet intervalle sont modifiées. Ce processus utilise les informations affectées telles que le temps de traversée et la proportion du choix d'itinéraire. De plus, à partir des matrices estimées pour l'intervalle actuel, le modèle peut facilement prévoir les matrices correspondantes aux intervalles de départ au futur, et mettre à jour ces matrices pour les intervalles de départ précédents. En 2002, Bierlaire a développé cette méthode par une combinaison des formulations proposées par Cascetta et al. [CIM93], et celles proposées par Ashok et Ben-Akiva, en comparant avec algorithme LSQR proposé par Paige et Saunders en 1982 [PS82] pour la solution du problème « least square ». La comparaison entre deux algorithmes a été réalisée à la fois théoriquement et numériquement. De cette comparaison, il nous indique quels avantages et désavantages obtenus par utilisation de deux méthodes.

D.4.1.1 Les équations du modèle

Le modèle est composé par trois équations :

a. Equation de transition qui représente la relation entre la déviation du flux O-D par rapport à sa valeur historique à l'instant h et celle à l'instant précédant, grâce à la formulation autorégressive.

$$\partial\Delta Q_\tau = \sum_{p=\tau-q'}^{\tau-1} f_\tau^p \partial\Delta Q_p + \varepsilon_{Q\tau} \qquad \text{(d.4.1)}$$

où $\partial\Delta Q_\tau = \Delta Q_\tau - \Delta\tilde{Q}_\tau$: vecteur de déviation pendant τ

f_p^h : matrice $(n_{OD} \times n_{OD})$ qui représente la contribution de $\partial\Delta Q_p$ à $\partial\Delta Q_\tau$. Elle peut être calculée par modèle de régression linéaire.

q' : nombre d'intervalles de temps précédents influant sur $\partial\Delta Q_\tau$.

$\varepsilon_{Q\tau}$: vecteur des variables aléatoires capturant les erreurs, avec :

- $E\left[\varepsilon_{Q\tau}\right] = 0$,

- $E\left[\varepsilon_{Q\tau}\,\varepsilon'_{Ql}\right]=U_\tau\delta_{\tau l}$ où U_τ est une matrice variance-covariance, $\delta_{\tau l}=1$ si τ=l et = 0 sinon.

b. Equation de propagation du volume par OD aux les arcs du réseau.

$$\Delta X_\tau=\sum_{p=\tau-p'}^{\tau}a_\tau^p\Delta Q_p \qquad \text{(d.4.2)}$$

c. Equation d'observation qui capture la relation entre les variables d'état (les flux O-D) et les observations (les comptages sur l'arc).

$$\Delta V_\tau=\Delta X_\tau+\varepsilon_{V\tau} \qquad \text{(d.4.3)}$$

Puis, on remplace ΔX_τ dans (d.4.3) par l'équation (d.4.2), on obtient :

$$\Delta V_\tau=\sum_{p=\tau-p'}^{\tau}a_\tau^p\Delta Q_p+\varepsilon_{V\tau}$$

Pour représenter l'équation d'observation par la déviation, on utilise :

$\partial\Delta V_\tau=\Delta V_\tau-\sum a_\tau^p\Delta\tilde{Q}_p$ avec ΔV_τ est un vecteur contenant les comptages pendant τ.

Finalement, on a :

$$\partial\Delta V_\tau=\sum_{p=\tau-p'}^{\tau}a_\tau^p\partial\Delta Q_p+\varepsilon_{V\tau} \qquad \text{(d.4.4)}$$

a_h^p : matrice de (*nombre arcs observés* x n_{OD}), dit matrice « affectation » qui capture topologique du réseau, le choix d'itinéraire et temps de traversée et qui permet de projeter les flux O-D partant pendant T_p au flux de l'arc observé pendant T_k.

p' : nombre maximum d'intervalles nécessaire pour traverser entre couple O-D.

$\varepsilon_{Q\tau}$: vecteur des variables aléatoires capturant les erreurs d'observation pendant τ. Avec :

- $E\left[\varepsilon_{V\tau}\right]=0$,

- $E\left[\varepsilon_{V\tau}\,\varepsilon'_{Vl}\right]=R_\tau\delta_{\tau l}$ où R_τ est une matrice variance-covariance, $\delta_{\tau l}=1$ si τ=l et = 0 sinon.

Bierlaire a présenté la méthode « least square » pour l'estimation/prédiction dynamique des matrices O-D en temps réel par la formulation suivante, en proposant les données observées pour une période de 1 à *h*, et N est la fin du période estimée :

$$\min_{\Delta Q} \sum_{\tau=1}^{h} \left\| \Omega_\tau^{-1} C_\tau^N Q - \Omega_\tau^{-1} \upsilon_\tau \right\|_2^2 + \sum_{\tau=h+1}^{N} \left\| \Omega_\tau^{-1} C_\tau^N Q \right\|_2^2 \qquad \text{(d.4.5)}$$

$$\text{où} \qquad Q = \begin{pmatrix} \partial \Delta Q_1 \\ \vdots \\ \partial \Delta Q_{N-1} \\ \partial \Delta Q_N \end{pmatrix}; \qquad \upsilon_\tau = \begin{pmatrix} 0_{n_{OD} \times 1} \\ \partial \Delta V_\tau \end{pmatrix};$$

$$\Theta_\tau = \Omega_\tau \Omega_\tau^T = \begin{pmatrix} U_\tau & 0 \\ 0 & R_\tau \end{pmatrix} = \begin{pmatrix} P_\tau P_\tau^T & 0 \\ 0 & S_\tau S_\tau^T \end{pmatrix}$$

$$\text{si } q' > p' \qquad C_\tau^N = \begin{pmatrix} 0 & \cdots & 0 & -f_\tau^{\tau-q'} & \cdots & \cdots & \cdots & -f_\tau^{\tau-1} & I & 0 & \cdots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & a_\tau^{\tau-p'} & \cdots & a_\tau^{\tau-1} & a_\tau^{\tau} & 0 & \cdots & 0 \end{pmatrix} = \begin{pmatrix} C_\tau^1 \\ C_\tau^2 \end{pmatrix}$$

$$\text{si } q' \le p' \qquad C_\tau^N = \begin{pmatrix} 0 & \cdots & \cdots & \cdots & 0 & -f_\tau^{\tau-q'} & \cdots & -f_\tau^{\tau-1} & I & 0 & \cdots & 0 \\ 0 & \cdots & 0 & a_\tau^{\tau-p'} & \cdots & \cdots & \cdots & a_\tau^{\tau-1} & a_\tau^{\tau} & 0 & \cdots & 0 \end{pmatrix} = \begin{pmatrix} C_\tau^1 \\ C_\tau^2 \end{pmatrix}$$

D.4.1.2 Algorithmes appliqués

Bierlaire a proposé deux solutions, en appliquant deux algorithmes « filtre de Kalman » et LSQR, pour le problème des moindres carrés dans l'estimation des matrices en temps réel. L'algorithme utilisant un filtre de Kalman proposé par Bierlaire est exactement correspondant avec algorithme de Ashok et Ben-Akiva en 1993.

a. Kalman Filter

L'algorithme « filtre de Kalman » résout la formulation (d.4.5) de façon itérative. Supposons que le problème est résolu pendant l'intervalle précédant $\tau - 1$, avec les résultats $Q_{\tau-1}$ et variance-covariance $H_{\tau-1}$. La mise en date pour l'intervalle k de ces résultats est faite en deux étapes : (1) 1ère étape incorpore l'équation de transition pour obtenir $\hat{Q}_\tau$ et $\hat{H}_\tau$, (2) 2ème étape incorpore l'équation d'observation pour obtenir Q_τ et H_τ. Les solutions du problème sont les suivantes :

$$\hat{Q}_\tau = \begin{pmatrix} I \\ F_{\tau-1} \end{pmatrix} Q_{\tau-1}$$

Avec $$F_{\tau-1} = \begin{pmatrix} 0 & \ldots & 0 & f_\tau^{\tau-q'} & \ldots & f_\tau^{\tau-1} \end{pmatrix}$$

$$\hat{H}_\tau = \left(\begin{array}{c|c} H_{\tau-1} & H_{\tau-1} F_{\tau-1}^T \\ \hline F_{\tau-1} H_{\tau-1} & F_{\tau-1} H_{\tau-1} F_{\tau-1}^T + U_\tau \end{array} \right)$$

Puis, l'équation d'observation est incorporée, fondé sur Bertsekas [Bert95] :

$$\begin{aligned} H_\tau &= \hat{H}_\tau + (C_\tau^2)^T R_\tau^{-1} C_\tau^2 \\ Q_\tau &= \hat{Q}_\tau + H_\tau^{-1} (C_\tau^2)^T (R_\tau^{-1} \partial \Delta V_\tau - R_\tau^{-1} C_\tau^2 \hat{Q}_\tau) \end{aligned} \qquad \text{(d.4.6)}$$

A partir de formule Sherman-Morrison-Woodbury [GVL96], on obtient :

$$H_\tau^{-1} = (I - K_\tau C_\tau^2) \hat{H}_\tau^{-1} \qquad \text{(d.4.7)}$$

où $$K_\tau = \hat{H}_\tau^{-1} (C_\tau^2)^T (R_\tau + C_\tau^2 \hat{H}_\tau^{-1} (C_\tau^2)^T)^{-1} \qquad \text{(d.4.8)}$$

et enfin $$Q_\tau = \hat{Q}_\tau + K_\tau (\partial \Delta V_\tau - C_\tau^2 \hat{Q}_\tau) \qquad \text{(d.4.9)}$$

b. LSQR

LSQR est une méthode itérative pour résoudre le problème des moindres carrés

$$\min_x \left\| Ax - b \right\|_2^2 \qquad \text{(d.4.10)}$$

L'algorithme LSQR a été détaillé par Paige et Saunders. Il n'est pas construit pour l'application en temps réel, heureusement on peut faire la modification suivante pour l'appliquer au problème d'estimation dynamique. Le problème des moindres carrés est reformulé avec $y = x - \bar{x}$ et $\bar{x}$ est une estimation précédente.

$$\min_x \left\| Ay - (b - A\bar{x}) \right\|_2^2 \qquad \text{(d.4.11)}$$

puis $$x^* = LSQR(A, b, \bar{x}) = \bar{x} + \arg\min_y \left\| Ay - (b - A\bar{x}) \right\|_2^2 \qquad \text{(d.4.12)}$$

L'algorithme LSQR dans l'aspect dynamique est le suivant :

Initialisation quand il n'y a pas de données de comptage, matrices O-D historique deviennent les meilleures estimations. Donc, initialiser le problème avec $k=0$, poser $Q_0 \leftarrow 0$ $\forall h=1, \ldots, N$

Boucle $k=1, \ldots, N$ par chaque itération, incorporer les comptages, et puis mettre à jour les matrices estimées et prévues en accordant avec :

$$Q_\tau = LSQR\left(\sum_{h=1}^{\tau}\Omega_h^{-1}C_h, \sum_{h=1}^{\tau}\Omega_h^{-1}z_h, Q_{\tau-1}\right) \qquad \text{(d.4.13)}$$

D.4.1.3 Commentaires du cas d'école

Le cas d'école dans cet article traite un réseau simple de 5 arcs avec trois relations O-D. Les flux OD de ce réseau seront estimés par les deux algorithmes présentés précédemment. Les hypothèses du problème sont les suivants :

- Couple OD : (1-5) ; (1-6) ; (2-6)

- Nombre d'intervalles : 15, avec T minutes pour chaque intervalle,

- Temps de traversée pour tous les arcs est de T minute,

- Capacité des arcs est illimité,

- Trois vraies matrices ont été données, puis les matrices historiques seront obtenues par la perturbation aléatoire des vraies matrices,

- Les matrices d'affectation et d'autorégressive sont déterminées à l'avance

- Les flux chargés sur tous les 5 arcs sont calculés par une affectation des vraies matrices au réseau par équation de propagation utilisant les matrices d'affectation, puis les comptages fictifs seront obtenus par la perturbation aléatoire des flux modélisés.

- Les matrices variance-covariance sont données à l'avance.

Nous donnons des commentaires sur cette application numérique :

- Les données d'application ne sont pas claires (les matrices historiques, les comptages sur les arcs obtenus par la perturbation aléatoire)

- La structure du réseau ne fait pas le choix d'itinéraire, c'est-à-dire pas de chemins concurrents (un seul arc 34 sur la coupure entre zone d'émission et zone de réception).

- Il y a trop de capteurs (comptages sur les arcs observés = 5) par rapport aux flux O-D estimés (flux OD = 3),

- La capacité est illimitée pour tous les arcs, c'est-à-dire la congestion modélisée n'est pas prise en compte.

Pour représentation de la méthode de Bierlaire plus claire et plus efficace, nous réalisons une nouvelle application avec les objectifs suivants :

- un exemple numérique complètement documenté et juste.

- traitement de deux cas : l'un avec beaucoup de capteurs et l'autre avec peu de capteurs.

- le choix d'itinéraire et la congestion sont pris en compte.

L'exemple sera traité pour le réseau du cas statique avec les données temporalisés qui permet de monter le choix d'itinéraire et la congestion.

D.4.1.4 Comparaison entre deux méthodes : filtre de Kalman et LSQR

En général, il nous semble que l'algorithme « filtre de Kalman » résout seulement, à chaque intervalle k, les matrices Ck, Ωk et le vecteur zk, mais LSQR résout ces matrices pour tous les intervalles précédents. C'est pourquoi l'algorithme « filtre de Kalman » a été proposé largement dans la littérature du problème d'estimation en temps réel. Bierlaire a fait une comparaison théorique (la complexité du problème), puis deux cas numérique (1) réseau de la centre Artery (Boston) et (2) Réseau Ivrine. De ces comparaisons, à la fin il a nous démontré que la performance de LSQR est plus forte que « filtre de Kalman », notamment le nombre d'itérations de LSQR est très moins que « filtre de Kalman » pour la convergence (cf. [BC02] pour les détails).

D.4.2 Approche de Van der Zijpp [VZ96]

En 1996, Van der Zijpp a proposé une méthode d'estimation dynamique des matrices O-D à partir d'une combinaison des comptages de trafic et des informations d'identification automatique des véhicules. Cette méthode estime la probabilité $p_{ij}(k)$ d'un véhicule qui entre par entrée *i* pendant la période *k* et qui sortira par sortie *j* du réseau.

D.4.2.1 Les équations d'estimation

Notons, la probabilité p_τ peut varier lentement en fonction du temps, le modèle « marche aléatoire » peut être donc appliqué :

$$p_{\tau+1} = p_\tau + \varepsilon_{p\tau} \qquad \text{(d.4.14)}$$

où p_τ : un vecteur de probabilité $\in$ [0 1]

$\varepsilon_{p\tau}$: un vecteur de variable aléatoire capturant les erreurs. Avec

- $E\left[\varepsilon_{p\tau}\right]=0$

- $E\left[\varepsilon_{p\tau}\,\varepsilon_{pl}^{'}\right]=S_{\tau}\delta_{\tau l}$

Et la distribution du flux O-D avec la période de départ t est fournie par la distribution multinomial suivante (cf. [ND89]) :

$$P\left[\Delta Q_{\tau}\left|\Delta\tilde{E}_{\tau},p_{\tau}\right.\right]=\prod_{i=1}^{m}\Delta E_{i\tau}!\prod_{j=1}^{n}\frac{p_{ij\tau}^{\Delta Q_{ij\tau}}}{\Delta Q_{ij\tau}} \qquad \text{(d.4.15)}$$

où ΔQ_{τ} : vecteur des flux O-D pendant l'intervalle τ

$\Delta\tilde{E}_{\tau}$: vecteur des vrais flux d'entrée pendant l'intervalle τ

m, n : nombre d'entrées et de sorties

Les éléments du vecteur des flux idéalisés des arcs ΔX_{τ} sont la somme de tous flux O-D avec le point de départ τ passent une position donnée. Cela est formulé par équation suivante :

$$\Delta X_{\tau}=U'\Delta Q_{\tau},\ \text{avec}\ U_{\varphi(i,j),a}=\delta_{ija}\ ;\ \varphi(i,j)=(n-1)i+j \qquad \text{(d.4.16)}$$

où $a=1\ldots l$: nombre des arcs observés.

$\delta_{ija}=1$ si route *i-j* utilise l'arc *a*, =0 sinon.

Combiner les équations (d.4.15) et (d.4.16), en supposant :

$$E\left[\Delta Q_{\tau}\left|\Delta E_{\tau}\right.\right]=H_{\tau}'p_{\tau}$$

$$H_{\varphi(i,j),a\tau}=\delta_{ija}\Delta E_{i\tau}$$

On obtient l'équation d'observation suivante :

$$\Delta V_{\tau}=H_{\tau}'p_{\tau}+\varepsilon_{a\tau} \qquad \text{(d.4.17)}$$

$\varepsilon_{a\tau}$: vecteur présentant les erreurs d'observation dans ΔV_{τ} et ΔE_{τ}, avec la variance-covariance R_{τ} de $\varepsilon_{a\tau}$:

- $E\left[\varepsilon_{a\tau}\right]=0$

- $E\left[\varepsilon_{a\tau}\,\varepsilon_{a\tau}'\right]=R_{\tau}$

(cf. [VZ96] pour les détails de R_τ)

D.4.2.2 Méthodes utilisées

Van der Zijpp a utilise les trois méthodes basées sur DCLS (Discounted Constrained Least Squares), l'approche de Bayesian, et l'approche «filtre de Kalman» pour résoudre le problème d'estimation présenté précédemment.

a. DCLS

Cette méthode estime la probabilité $\overline{p}_\tau \equiv \arg\min J(p,\tau)$, avec

$$J(p,\tau) = \sum_{k=1}^{\tau} \lambda^{\tau-k} \left\| \Delta V_k - H_k' p \right\|^2 \qquad \text{(d.4.18)}$$

où $\lambda \in [0\ 1]$ est le poids portant sur les anciennes observations.

Cette méthode est facile à implémenter, mais ses désavantages sont impossibles d'exploiter les détails spécifiés dans les équations précédentes.

b. Approche de Bayesian

Les informations contenues dans les nouvelles observations peut être incorporées dans la distribution de probabilité subjective utilisant règle Bayes :

$$prob\left[p_\tau \left| \Delta V_1 \ldots \Delta V_\tau \right. \right] = \frac{prob\left[\Delta V_\tau \left| p_\tau, \Delta V_1 \ldots \Delta V_{\tau-1} \right. \right] \cdot prob\left[p_\tau \left| \Delta V_1 \ldots \Delta V_{\tau-1} \right. \right]}{prob\left[\Delta V_\tau \left| \Delta V_1 \ldots \Delta V_{\tau-1} \right. \right]} \qquad \text{(d.4.19)}$$

Le dénominateur est une normalisation constante, car il est invariant par rapport à p_τ. Le numérateur comprends le produit de la fonction vraisemblance (gauche) et la distribution a priori (droite). Notons que la fonction vraisemblance, distribution a priori, distribution a posteriori sont « normal multivarié », on obtient donc :

$$\begin{aligned} \overline{p}_\tau^+ &= \overline{p}_\tau^- + K_\tau \left[\Delta V_\tau - H_\tau' \overline{p}_\tau^- \right] \\ K_\tau &= \Sigma_{p\tau}^- H_\tau \left[H_\tau' \Sigma_{p\tau}^- H_\tau + R_\tau \right]^{-1} \\ \Sigma_{p\tau}^+ &= \Sigma_{p\tau}^- - \Sigma_{p\tau}^- H_\tau \left[H_\tau' \Sigma_{p\tau}^- H_\tau + R_\tau \right]^{-1} H_\tau' \Sigma_{p\tau}^- \end{aligned} \qquad \text{(d.4.20)}$$

où $\overline{p}_\tau^-$ et $\Sigma_{p\tau}^-$ sont la moyenne et covariance a priori de p_τ

$\overline{p}_\tau^+$ et $\Sigma_{p\tau}^+$ sont la moyenne et covariance a posteriori de p_τ

Les équations (d.4.19) sont proches équations de la méthode « filtre de Kalman ». La différence entre deux méthodes est d'interpréter $\overline{p}_t$. Cela amène quelques différences en pratique (cf. [VZ96]).

D.4.2.3 Estimation avec multi-sources d'observation

Van der Zijpp a proposé une nouvelle méthode d'estimation dynamique des matrices O-D à partir d'une combinaison de comptages de trafic et d'observation de trajectoire des véhicules. Notons ΔG_τ est un vecteur qui contient les observations de trajectoire $\{\Delta G_{xy\tau} : x < y\}$. $\Delta G_{xy\tau}$ présente le nombre de véhicules constatés au site *x* pendant l'intervalle τ et au site *y*, pas au site en amont de *x* et au site en aval de *y*.

La distribution de la contribution d'observation de trajectoire suit la distribution multinomial suivante :

$$prob\left[\Delta G_\tau \middle| \Delta\tilde{E}_\tau, p_\tau\right] = \prod_{i=1}^{m} \Delta\tilde{E}_i ! \prod_{j=1}^{n} \prod_{r<s \vee xy=0} \frac{\pi_{xy}^{ij} p_{ij\tau}^{\Delta G_{xy\tau}^{ij}}}{\Delta G_{xy\tau}^{ij}} \qquad \text{(d.4.21)}$$

Où π_{xy}^{ij} est la probabilité de contribution de flux O-D $\Delta G_{ij\tau}$ à $\Delta G_{xy\tau}$

$\Delta G_{xy\tau}^{ij}$ est la contribution de flux O-D $\Delta G_{ij\tau}$ à $\Delta G_{xy\tau}$, et la probabilité d'un véhicule qui sera contribué à $\Delta G_{xy\tau}^{ij}$ est égale à $p_{ij\tau}.\pi_{xy}^{ij}$.

$$\begin{aligned} E\left[\Delta G_\tau \middle| x_\tau\right] &= R'_\tau y_\tau \\ R_{\varphi(i,j),\theta(x,y),\tau} &= \Delta E_{i\tau}.\pi_{xy}^{ij} \end{aligned} \qquad \text{(d.4.22)}$$

Et enfin, l'équation d'observation est la suivante :

$$\Delta G_\tau = R'_\tau p_\tau + \varepsilon_{G\tau} \qquad \text{(d.4.23)}$$

(cf. [VZ96] pour les détails des paramètres)

Pour le cas de combinaison des sources d'observation (par ex : comptage de trafic et observation de trajectoire), les deux équations (d.4.17) et (d.4.23) peuvent être reformulées dans une seule équation d'observation combinée suivante :

$$\begin{bmatrix} \Delta V_\tau \\ \Delta G_\tau \end{bmatrix} = \begin{bmatrix} H'_\tau \\ G'_\tau \end{bmatrix} p_\tau + \begin{bmatrix} \varepsilon_{a\tau} \\ \varepsilon_{G\tau} \end{bmatrix} \qquad \text{(d.4.24)}$$

Grâce à des tests numériques, Van der Zijpp montre les bénéfices de la nouvelle méthode et la réduction des erreurs d'estimation si on combine les observations de trajectoire avec les comptages de trafic. De plus, la dépendance entre les observations doit être considérée en pratique.

D.4.3 Approche de Zhang et Mustard [ZM02]

Dans cette approche, le problème d'estimation est formulé comme un problème de programmation mathématique basé sur le principe maximum d'entropie pour le cas dynamique. La programmation a une structure hiérarchique : au haut niveau, la matrice est estimée avec les proportions de choix d'itinéraire données, et au bas niveau les proportions de choix d'itinéraire sont déterminées par un modèle d'affectation dynamique avec la matrice déterminée au haut niveau (le simulateur dynamique CONTRAM est utilisé au bas niveau).

D.4.4.1 Formulation du problème

Le problème d'estimation de cette méthode est formulé par équation suivante :

$$\min Z = W_Q.Q + W_V.V \qquad \text{(d.4.25)}$$

avec le contrainte $\Delta V_{ap} = \sum_{\tau=1}^{p} \sum_{ij} \Delta Q_{ij\tau} P_{ij\tau}^{ap}, \quad \forall a, p$

ou $\Delta Q_{ij\tau}$: le flux O-D entre *i* et *j* pendant l'intervalle τ, $\Delta Q = \left\{\Delta Q_{ij\tau}\right\}$

$\Delta \tilde{Q}_{ij\tau}$: le flux O-D entre i et j a priori pendant l'intervalle τ, $\Delta \tilde{Q} = \left\{\Delta \tilde{Q}_{ij\tau}\right\}$

ΔX_{ap} : le flux sur l'arc *a* obtenu par l'affectation dynamique utilisant la matrice estimée pendant l'intervalle *p*, $\Delta X = \left\{\Delta X_{ap}\right\}$

ΔV_{ap} : le flux observé sur l'arc *a* pendant l'intervalle *p*, $\Delta V = \left\{\Delta V_{ap}\right\}$

$P_{ij\tau}^{ap}$: la proportion des flux O-D (i,j) départ à l'intervalle τ passant l'arc *a* à l'intervalle *p*.

Dans l'équation (d.4.25) :

- Q et V sont respectivement les différences entre la matrice O-D ΔQ et $\Delta \tilde{Q}$, entre les flux des arcs ΔX et ΔV.

- W_Q et W_V sont respectivement les poids pour les matrices O-D ΔQ et $\Delta \tilde{Q}$ et les flux des arcs ΔX et ΔV.

Cette méthode utilise fonction négative d'entropie pour résoudre ce problème, équation (d.4.25) est donc reformulée comme :

$$\min Z = W_Q \sum_{\tau} \left[\sum_{ij} \Delta Q_{ij\tau} \left(\ln \Delta Q_{ij\tau} / \Delta \tilde{Q}_{ij\tau} - 1 \right) \right] + W_V \sum_{p} \left[\sum_{a} \Delta X_{ap} \left(\ln \Delta X_{ap} / \Delta V_{ap} - 1 \right) \right]$$

avec le contrainte $\Delta V_{ap} = \sum_{\tau=1}^{p} \sum_{ij} \Delta Q_{ij\tau} P_{ij\tau}^{ap}, \quad \forall a, p$

L'algorithme heuristique est développé pour résoudre ce problème. Il a deux sub-problèmes : (1) Estimation des matrices temporelles sachant les proportions de choix d'itinéraire, (2) Affectation dynamique utilisant les matrices de la demande estimées.

L'algorithme itératif suppose, à chaque itération *k*, une solution courante $\left[\Delta Q^{[k]}, \Delta X^{[k]} \right]$, où $\Delta X^{[k]}$ issus d'une affectation $\Delta Q^{[k]}$ au réseau. Le traitement du problème d'estimation sachant $\{ P_{ij\tau}^{ap}(\Delta Q^{[k]}) \}$ fournit une solution auxiliaire de la matrice ΔQ^*. L'affectation de trafic sur le réseau avec ΔQ^* fournit les flux auxiliaires des arcs ΔX^*. Et puis, on a une solution auxiliaire $\left[\Delta Q^*, \Delta X^* \right]$.

Ensuit, on utilise la méthode MSA pour calculer une nouvelle solution :

$$\Delta Q^{[k+1]} = \Delta Q^{[k]} + \lambda_k \left(\Delta Q^* - \Delta Q^{[k]} \right) \qquad \text{(d.4.26)}$$

avec $\lambda_k = 1/k$

et une nouvelle solution pour les flux des arcs $\Delta X^{[k+1]}$ sera obtenue par l'affectation de trafic sur le réseau utilisant nouvelle matrice $\Delta Q^{[k+1]}$.

D.4.4.1 Algorithme du modèle

Les étapes de l'algorithme proposé sont les suivantes :

Etape 1 : Initialisation. Opérer CONTRAM pour affecter la matrice a priori $\left\{\Delta\tilde{Q}_{ij\tau}\right\}$ sur le réseau. Archiver les résultats des flux des arcs et les proportions de choix d'itinéraire, $\left\{\Delta X_{ap}^{[0]}\right\}$ et $\left\{P_{ij\tau}^{ap[0]}\right\}$. Poser $k \leftarrow 0$.

Etape 2 : Résoudre le problème d'estimation des matrices et archiver la matrice auxiliaire $\left\{\Delta Q_{ij\tau}^{*}\right\}$, en sachant les proportions de choix d'itinéraire $\left\{P_{ij\tau}^{ap[k]}\right\}$ et $\left\{\Delta X_{ap}^{[k]}\right\}$ issues d'une opération de CONTRAM.

Etape 3 : Affecter la matrice $\left\{\Delta Q_{ij\tau}^{*}\right\}$ sur le réseau utilisant COMTRAM et Archiver les flux auxiliaires des arcs $\left\{\Delta X_{ap}^{*}\right\}$.

Etape 4 : Utiliser la méthode MSA pour calculer une nouvelle solution $\left\{\Delta Q_{ij\tau}^{[k+1]}\right\}$ par l'équation (27)

Etape 5 : Affecter nouvelle matrice $\left\{\Delta Q_{ij\tau}^{[k+1]}\right\}$ sur le réseau utilisant CONTRAM. Archiver nouvelle solution pour les flux des arcs et les proportions de choix d'itinéraire $\Delta X_{ap}^{[k+1]}$ et $\left\{P_{ij\tau}^{ap[k+1]}\right\}$

Etape 6 : Poser $k \leftarrow k+1$ et retourner à la étape 2 jusqu'à les changements des flux dans la matrice O-D et des flux des arcs sont acceptable pour obtenir une convergence.

Cette a été testée par une application sur le grand réseau KENT pour une journée (1209 arcs, 643 nœuds, 288 tranche de temps pour 24 heures, 67 origines, 67 destinations, 2143 relation O-D, 208 arcs observés). Les résultats numériques nous montent qu'elle est très efficace sur les premières itérations, mais la convergence est difficile à obtenir à cause d'utilisation de la méthode MSA. La solution pour ce problème est de calculer approximativement le « step-length » λ optimal pour réduire le nombre d'itération.

D.4.5 Synthèse sur l'estimation dynamique

Pour avoir une vision globale sur les méthodes récentes d'estimation dynamique des matrices O-D à partir des observations actuelles et des estimations a priori, nous synthétisons les méthodes considérées dans le tableau suivant :

Caractéristiques	Bierlaire (2002)	Approche de Van der Zijpp (1996)	Zhang et Mustard (2002)
Equations mathématiques pour l'estimation	Combinaison des formulations originales « least square » proposé par Cascetta et al (1993) et de la déviation des variables d'état proposée par Ashok et Ben-Akiva (1993) dans les équations de transition et d'observation.	Equation du modèle « random walk » , en utilisant la distribution multinomial pour calculer la probabilité du choix d'itinéraire. Cela fournit une équation d'observation qui permet d'estimer la matrice à partir des comptages.	L'équation d'inférence par maximum d'entropie
Algorithme	- Application la méthode « least square » à l'algorithme « Kalman Filter » et LSQR (Paige et Saunders, 1982)	- Application de trois méthodes DCLS, Bayesian, et Kalman Filter aux équations mathématiques pour estimer les matrices dynamique à partir des comptages et des observations de trajectoire.	Algorithme heuristique et MSA sont utilisé
Sources d'observations	- Comptages sur les quelques arcs dans quelques périodes. - Matrice a priori	- Comptages sur les arcs - Observations de trajectoire -Matrice a priori	- Comptages sur les arcs - matrice a priori
Estimation /inférence	Estimation « on-line »	Estimation « on-line »	Inférence par maximum d'entropie
Mesure des erreurs			
Outils à l'aide	- Développé en Matlab - Utilisation de DynaMIT pour la référence		COMTRAM est utilisé pour affecter les matrices a priori afin d'obtenir les flux chargés sur les arcs et les proportions de choix d'itinéraires

Applications numériques	- Central Artery, Boston (221 arcs et 183 nœuds, 10 relations O-D, 35 arc observés) pour une durée de 1 heure, avec tranche de temps de 15 min. - Irvine, Ca (618 acrs, 296 nœuds, 627 relation O-D) pour une durée 1 heure, avec tranche de temps de 15 min.	Réseau Amsterdam avec les données fournies par système MARE (1994). Un sub-réseau est collecté pour le test numérique avec 5 entrées, 5 sorties, et 11 km de longueur. La durée d'estimation est de 0 à 20 heures avec tranche de temps de 5 min.	Réseau KENT (1209 arcs, 643 nœuds, 288 tranche de temps pour 24 heures,2143 relation O-D, 208 arcs observés)

D.5 Estimateur de la matrice OD dynamique

Grâce à une étude bibliographique sur l'estimation de la matrice OD dynamique, nous avons choisi une des méthodes récentes discutées dans la section D.4 pour estimer une matrice OD dynamique à partir des comptages sur les arcs. Il s'agit de Bierlaire et Crittin en 2002 [BC02], dont l'algorithme que nous allons développer dans la section D.5.3. Une première version de l'estimateur a été programmée, en utilisant le progiciel SCILAB. Nous avons également réalisé un premier test d'exemple simple afin de valider le fonctionnement de l'estimateur.

D.5.1 Algorithme « Filtre de Kalman »

Nous présentons ici un algorithme d'estimation d'une matrice OD dynamique par méthode de filtre de Kalman proposée par Bierlaire, en développant un nouvel algorithme qui permet de déterminer la matrice d'affectation qui est un des paramètres importants que Bierlaire n'a pas encore précisé dans son article.

D.5.1.1 Hypothèses

Afin d'estimer une matrice OD dynamique, on a besoins des données en inputs qui sont citées successivement comme les suivantes :

- la topologie du réseau,

- les profils de temps de traversée des arcs, noté $t_a(h)$. Ces profils sont obtenus grâce aux observations sur le réseau tels que la vitesse mesurée par les capteurs équipés sur les arcs, l'heure d'entrée et de sortie aux gares de péage, la vitesse libre pour les arcs non-observés, etc.,

- matrice OD a priori (matrice historique), noté $\Delta\widetilde{Q}$,

- matrice autorégressive f_p^h (cf. section D.4.1.1),

- matrice d'affectation a_h^p (cf. section D.4.1.1),

- les paramètres chronologiques tels que K le nombre d'intervalle d'estimation, p' nombre d'intervalles précédents influant sur l'intervalle traité, qui détermine la taille de la matrice autorégressive, q' nombre d'intervalles maximum pour traverser entre couple O-D, qui détermine la matrice d'affectation, l'intervalle initial k_0.

- Matrice variance-covariance présentant les erreurs d'estimation de la matrice estimée Q (écart actuel-historique par flux OD) pour chaque intervalle U_k,

- matrice d'amorçage de la journée $Q_{(k_0-1)}$ et matrice d'amorçage variance-corvariance $H_{(k_0-1)}$ (cf. section D.4.1.2),

- les vecteurs qui contiennent les volumes des arcs observés pour chaque intervalle de temps $\{\Delta V_k : \forall k \in K\}$, dont la taille est égale au nombre d'arcs observés A.

- Matrice variance-covariance présentant les erreurs d'observation V (écart actuel-historique par flux des arcs) pour chaque intervalle R_k.

D.5.1.2 Algorithme d'estimation

Nous précisons maintenant un algorithme d'estimation d'une matrice OD dynamique d'après méthode de filtre de Kalman :

1. Initialisation : poser $k \leftarrow k_0$; $Q \leftarrow Q_{(k_0-1)}$; $H \leftarrow H_{(k_0-1)}$

2. Détermination des paramètres :

$$F_{k-1} = \begin{pmatrix} 0 & \ldots & 0 & f_k^{k-q'} & \ldots & f_k^{k-1} \end{pmatrix}$$

$$C_k^2 = \begin{pmatrix} 0 & \ldots & 0 & a_k^{k-q'} & \ldots & a_k^{k} \end{pmatrix}$$

$$\partial\Delta V_k = \Delta V_k - \sum_{p=k-p'}^{k} a_k^p \Delta\tilde{Q}_p$$

3. Etape autorégressive :

Calculer : $$\hat{Q}_k = \begin{pmatrix} I \\ F_{k-1} \end{pmatrix} Q_{k-1}$$

et $$\hat{H}_k = \left(\begin{array}{c|c} H_{k-1} & H_{k-1}F_{k-1}^T \\ \hline F_{k-1}H_{k-1} & F_{k-1}H_{k-1}F_{k-1}^T + U_k \end{array} \right)$$

4. Etape d'ajustement :

Calculer :
$$H_k = \hat{H}_k + (C_k^2)^T R_k^{-1} C_k^2$$
$$Q_k = \hat{Q}_k + H_k^{-1}(C_k^2)^T (R_k^{-1}\partial\Delta V_k - R_k^{-1}C_k^2\hat{Q}_k)$$

5. Terminaison : si $k = K$, alors terminer.

Sinon poser $k = k+1$, retourner à l'étape 2.

D.5.1.3 Explications supplémentaires

Afin de mettre en évidence l'algorithme « filtre de Kalman » ci-dessus et aussi les variables, les paramètres utilisés, nous précisons quelques points supplémentaires concernant l'évolution de la taille des matrices et des vecteurs au cours de calcul. Par chaque étape de l'algorithme, les évolutions de la taille des matrices sont les suivantes :

- l'étape 2 calcule la matrice autorégressive extensive, la matrice d'affectation extensive, et l'écart entre volume des arcs observés actuel et historique. Dans cette étape, les tailles des matrices sont les suivantes :

F_{k-1} : $\left(n_{OD} \times (k-1).n_{OD}\right)$, dont la partie "zero" : $\left(n_{OD} \times (k-1-q'_k).n_{OD}\right)$,

C_k^2 : $(n_{Obs} \times k.n_{OD})$, dont la partie "zero" : $\left(n_{OD} \times (k-p'_k).n_{OD}\right)$,

$\partial\Delta V_k$: $(A \times 1)$.

- Etape auxiliaire utilise l'équation autorégressive pour estimer l'écart entre volume actuel et historique $\hat{Q}_k$, puis la matrice variance-covariance $\hat{H}_k$ jusqu'à l'intervalle *k*. Dans cette étape, les tailles des matrices sont les suivantes :

I : $(k-1).n_{OD} \times (k-1).n_{OD}$ (Matrice d'identité),

$Q_{(k-1)}$: $(k-1).n_{OD} \times 1$,

➔ $\hat{Q}_k$: $k.n_{OD} \times 1$,

$H_{(k-1)}$: $(k-1).n_{OD} \times (k-1).n_{OD}$,

U_k : $n_{OD} \times n_{OD}$,

➔ $\hat{H}_k$: $k.n_{OD} \times k.n_{OD}$.

- Etape d'ajustement utilise les observations sur les arcs $\partial\Delta V(T_k)$ pendant intervalle k pour ajuster les résultats de l'étape auxiliaire $\hat{Q}_k$ et $\hat{H}_k$. Dans cette étape, les tailles des matrices sont les suivantes :

R_k : $(A \times A)$,

➔ H_k : $k.n_{OD} \times k.n_{OD}$,

➔ Q_k : $k.n_{OD} \times 1$.

D.5.1.4 Algorithme de calcul de la matrice d'affectation

Pour chaque intervalle observé k, on a p'_k matrices de taille $n_{OD} \times A$, c'est-à-dire, chaque élément de la matrice « affectation » donne la proportion du volume OD (i) pendant intervalle p chargé à l'arc a pendant intervalle k.

Pour chaque OD (i), la proportion de partages entre les chemins $r \in R_i$ pendant intervalle p est $P^p_{r \in R_i}$.

L'algorithme de calcul de la matrice d'affectation est la suivant :

1. Début. Poser $k \leftarrow k_0$

2. Calcul par intervalles de distribution $p \in [(k - p'_k)\ k]$. Poser $p = (k - p'_k)$

3. Calcul par OD. Poser $i \leftarrow 1$; $a^p_k \leftarrow \varnothing$

4. Calcul par arcs. Poser $l \leftarrow 1$

5. Calcul des éléments de la matrice d'affectation :

$a^p_{k,i,l} = \sum_{r \in R_i \cap l} P^p_r \times \pi^p_{k,r,l}$ où $\pi^p_{k,r,l} \in [0\ 1]$

Ajouter l'élément $a^p_{k,i,l}$ à la matrice a^p_k

6. Terminaison pour les arcs :

Si $l = A$, poser $i \leftarrow i + 1$, retourner à l'étape 4.

Sinon, poser $l \leftarrow l + 1$, retourner à l'étape 5.

7. Terminaison pour les OD :

Si $i = n_{OD}$, poser $p \leftarrow p + 1$, retourner à l'étape 3.

Sinon, poser $i \leftarrow i + 1$, retourner à l'étape 4.

8. Terminaison pour les intervalles *p* :

Si $p = p'_k$, poser $k \leftarrow k + 1$, retourner à l'étape 2.

Sinon, poser $p \leftarrow p+1$, retourner à l'étape 3.

9. Terminaison pour les intervalles *k* :

Si $k = K$, aller à la Terminaison.

Sinon, poser $k \leftarrow k+1$, retourner à l'étape 2

10. Terminaison.

Afin de calculer la proportion du volume entrant d'un chemin *r* pendant *p* sur l'arc *l* pendant *k* noté $\pi_{k,r,l}^{p}$, nous avons fait une propagation du flux des chemins aux arcs en sachant le temps de traversée des chemins du réseau.

D.5.2 Programmation d'un estimateur

Le travail de programmation est réalisé d'après les étapes suivantes :

1. La topologie du réseau est représentée comme variable Network utilisant la fonction Make_Net

2. Le temps de traversée des arcs est représenté comme variable Prof utilisant la fonction Make_Prof

3. Détermination des chemins par OD avec leurs attributs. Pour cela, la fonction de formation des services du prototype de modèle LADTA a été modifiée, afin de trouver tous chemins possibles (ne pas tester les chemins efficaces/non-efficaces), en ajoutant des paramètres nécessaires au chaque itinéraire pour le calcul de la matrice d'affectation (les sous-chemins d'un itinéraire).

4. Détermination de la matrice « affectation » par une nouvelle fonction programmée d'après l'algorithme présenté dans la section D.5.1.4.

5. Estimation de la matrice OD actuelle par une fonction d'estimation par intervalle de temps utilisant l'algorithme « filtre de Kalman » présenté dans la section D.5.1.2.

6. Représentation des résultats par une procédure qui permet de sauvegarder et de représenter les hypothèses, les résultats sous format d'EXCEL.

8.5.3 Test d'exemple numérique

Afin de valider le fonctionnement de l'estimateur, un exemple d'estimation simple a été traité. Le réseau d'exemple est présenté dans la section D.2.3 (cf. figure D.2). Les hypothèses du problème d'estimation d'une matrice OD dynamique pour ce réseau sont les suivantes :

- l'estimation pour 5 intervalles de temps, dont la durée est de 2 heures (2 intervalles d'amorçage).

- les profils de temps de traversée des arcs avec la congestion.

- les proportions de partage du volume entre les chemins par OD pour chaque période selon le temps de traversée.

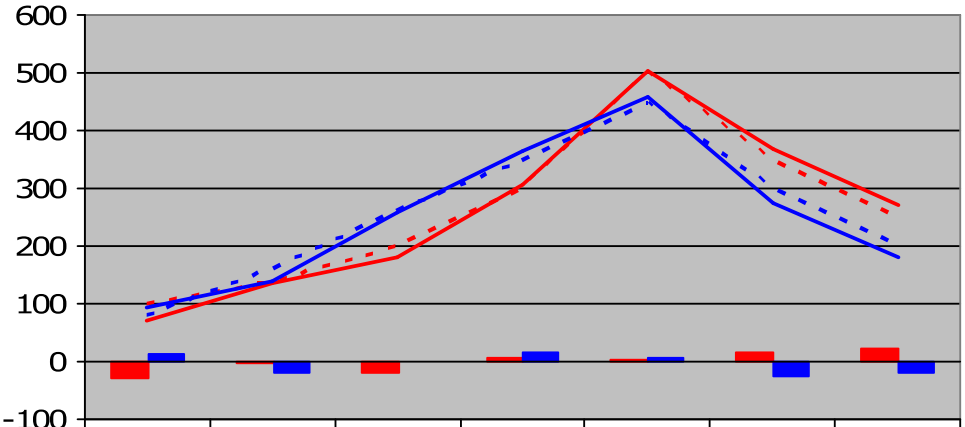
- la matrice OD historique.

- les matrices variace-covariance posées à l'avance.

- les comptages sur 3 arcs du réseau pour 5 périodes d'estimation.

- les matrices d'amorçage posées à l'avance pour 2 intervalles.

Le résultat d'estimation est présenté dans la figure D.3 :



	0_2	2_4	4_6	6_8	8_10	12_14	14_16
Ecart OD1	-30	-4	-18	7	4	17	22
Ecart OD2	13	-20	-1	15	8	-26	-18
OD1 His	100	140	200	300	500	350	250
OD1 Act	70	136	182	307	504	367	272
OD2 His	80	160	260	350	450	300	200
OD2 Act	93	140	259	365	458	273	182

Figure D.3 Matrice OD dynamique estimée par l'estimateur KF

Bibliographie

[CIM93] Cascetta, E., Inaudi, D. and Marquis, G. (1993) Dynamic estimators of origin-destination matrices using traffic counts. *Transportation Science*, **27**, 363-373.

[ZM02] Zhang X. and D. Mustard (2002) Methods for the estimation of time-dependent origin-destination matrices using traffic flow data on road links. *Paper produced by TRL Limited, as part of a contract placed by the Highways Agency*. Napier University.

[CW94] Chang G. and Wu J. (1994) Recursive Estimation of Time-varying Origin-Destination Flows from Traffic Counts in Freeway Corridors. *Transportation Research B*, **28/2**, 141-160.

[Bell91] Bell M. G.H. (1991) The real time estimation of origin-destination flows in the presence of platoon dispersion. *Transportation Research B*, **25, 2/3**, 115-125.

[ND87] Nihan, L. N. and Davis, G. (1987) Recursive estimation of Origin-Destination matrices from input/output counts. *Transportation Research Part B*, **21/2**, 149-163.

[VZ96] Van der Zijpp N. (1996) Dynamic origin-destination matrix estimation on motorway networks. Ph.D. Thesis, Delft University of Technology.

[KK87] Cremer, M. and Keller, H. (1987) A new class of dynamic methods for identification of Origin-Destination flows. *Transportation Research Part B*, **21/2**, 117-132.

[WDL04] Wu Y., Davis G. and Levinson D. (2004*) Improving the Estimation of Travel Demand for Traffic Simulation : Part II*. Final Report, Department of Civil Engineering, Intelligent Transportation Systems Institute, University of Minnesota.

[ABA93] Ashok, K and Ben-Akiva, M. (1993). Dynamic origin-destination matrix estimation and prediction for real-time traffic management systems, in C. Daganzo (ed.), *Transportation and Traffic Theory*, Elsevier Publishing Company Inc. Proceedings of the 12th ISTTT.

[PS82] Paige, C. C. and Saunders, M. A. (1982). LSQR : an algorithm for sparse linear equations and sparse least squares, ACM Transactions on Mathematical Software **8** : 43-71.

[BC02] Bierlaire, M. and Crittin, F. (2002). An efficient algorithm for real-time estimation and prediction of dynamic OD table, Proceedings of the 2nd Swiss *Transportation Research Conference,* Ascona, Switzerland.

[OS84] Okutani, I. and Stephanades, Y. (1984). Dynamic prediction of traffic volume through Kalman filtering theory*, Transportation Research Part B*, **18/1**, 1-11.

[YMB01] Yang, H., Q. Meng, and M.G.H. Bell. (2001). Simultaneous Estimation of the Origin-Destination Matrices and Travel Cost Coefficient for Congested Networks in a Stochastic User Equilibrium. *Transportation Science*, **35**, 107–123.

[Bert95] Bertsekas, D. P. (1995). Nonlinear Programming, Athena Scientific, Belmont.

[GV96] Golub, G. H. and Van Loan, C. F. (1996). Matrix Computations, third ed., Johns Hopkins, Baltimore and London.

E.1 Introduction 322

E.2 Le réseau et le trafic sur le territoire étudié 323

E.2.1 Réseau étudié 323
E.2.2 Les trafics : situation en 2000 et évolution à l'horizon 2020 323

E.3 Constitution des données pour une affectation dynamique ... 326

E.3.1 Acquisition des données 326
E.3.2 Génération des données 327

E.4 Premier résultat 328

Bibliographique 331

E.1 Introduction

Face à la forte évolution du trafic sur le réseau routier de la Vallée du Rhône (VdR), les bureaux d'études, notamment les CETES[1], ont réalisé plusieurs études de trafic sur la Vallée du Rhône et l'arc languedocien afin de prévoir le trafic futur. Les résultats de ces études permettent d'évaluer les scénarios d'aménagement et d'envisager de nouveaux projets de planification et d'exploitation sur le réseau étudié afin de satisfaire la demande croissante. Dans le chapitre 2, nous avons rapporté que les bureaux d'études de trafic en France n'utilisent que les modèles statiques (ex. TRANSCAD) des modèles microscopiques aux petites échelles (ex. DYNAMIT, AIMSUN). L'utilisation de modèles statiques ne répond pas toujours aux questions posées précédemment de manière satisfaisante. Il est donc nécessaire de procéder à des études de trafic de manière dynamique, c'est-à-dire : prenant en compte l'évolution du trafic au cours de la journée. Pour une étude à grande échelle comme le réseau VdR, nous avons besoin d'un outil capable de traiter des applications de grande taille en temps de calcul raisonnable. Un tel modèle dynamique et macroscopique doit être simple et réaliste. Afin de répondre à ces exigences, un modèle théorique, nommé LADTA, a été investigué en 2003 par Fabien Leurent, dont une version restreinte a été présentée dans le chapitre 3. Puis, un prototype du modèle a été fabriqué afin de tester la convergence de l'algorithme d'équilibrage ; sa programmation a été présentée dans le chapitre 5, et la question de la convergence a été discutée dans la partie 3 de la thèse. A partir de 2005, grâce à ces travaux, notamment les expériences numériques par prototype et les études de la question de convergence, Vincent Aguiléra a rempli une partie du travail de codage d'un simulateur performant du modèle LADTA en C^{++}. Cet outil de simulation a été testé par une première application rapide sur le réseau VdR pour sa performance et sa convergence. Les données nécessaires à cette application ont été récupérées à partir des études du CETE Méditerranée sur la Vallée du Rhône et l'arc languedocien [CETE04] pour une prévision à l'horizon 2020. Le réseau et les trafics de la Vallée du Rhône sont présentés dans la section E.2. La constitution des données pour une affectation dynamique est exposée dans la section E.3. Le premier résultat est résumé dans la section E.4.

[1] CETE : les Centres d'Etudes Techniques de l'Equipement

E.2 Le réseau et le trafic sur le territoire étudié

E.2.1 Réseau étudié

Le réseau routier de la Vallée du Rhône et de l'arc languedocien comprend les autoroutes A7 et A9, ainsi que des axes alternatifs, tels que l'A75 ou la liaison Grenoble-Sisteron. Dans le secteur restreint de l'étude, entre Lyon et Narbonne, le réseau est plus détaillé, y compris les principales RN (RN7, RN86, RN9, RN113, RN75 et 85). Ce réseau dessert les grands flux entre des régions en France (Nord, Centre, Région parisienne, Nord Est), l'étranger (Suisse, Angleterre, Belgique, Allemagne), et des régions du sud de la France, L'Espagne et la moitié de l'Italie au Sud-Ouest. Le réseau étudié est illustré dans la figure E.1.

E.2.2 Les trafics : situation en 2000 et évolution à l'horizon 2020

Le Cété Méditerranée a utilisé les comptages JMA en 2000 (sources SETRA : cf. la carte de comptages E.2) pour prévoir une évolution du trafic dans la Vallée du Rhône à l'horizon 2020. L'évolution tendancielle des trafics à l'horizon 2020 est illustrée dans la figure E.3.

Concernant les grands flux, les trafics actuels ont été analysés afin de caractériser la répartition des flux entre les grands axes du réseau (A75, A7+RN7, RN75+RN85 sur 2 sens). Les résultats sont présentés dans les figures E.4.

Figure E.1 Réseau étudié avec principaux projets (Source Cété Méditerrané)

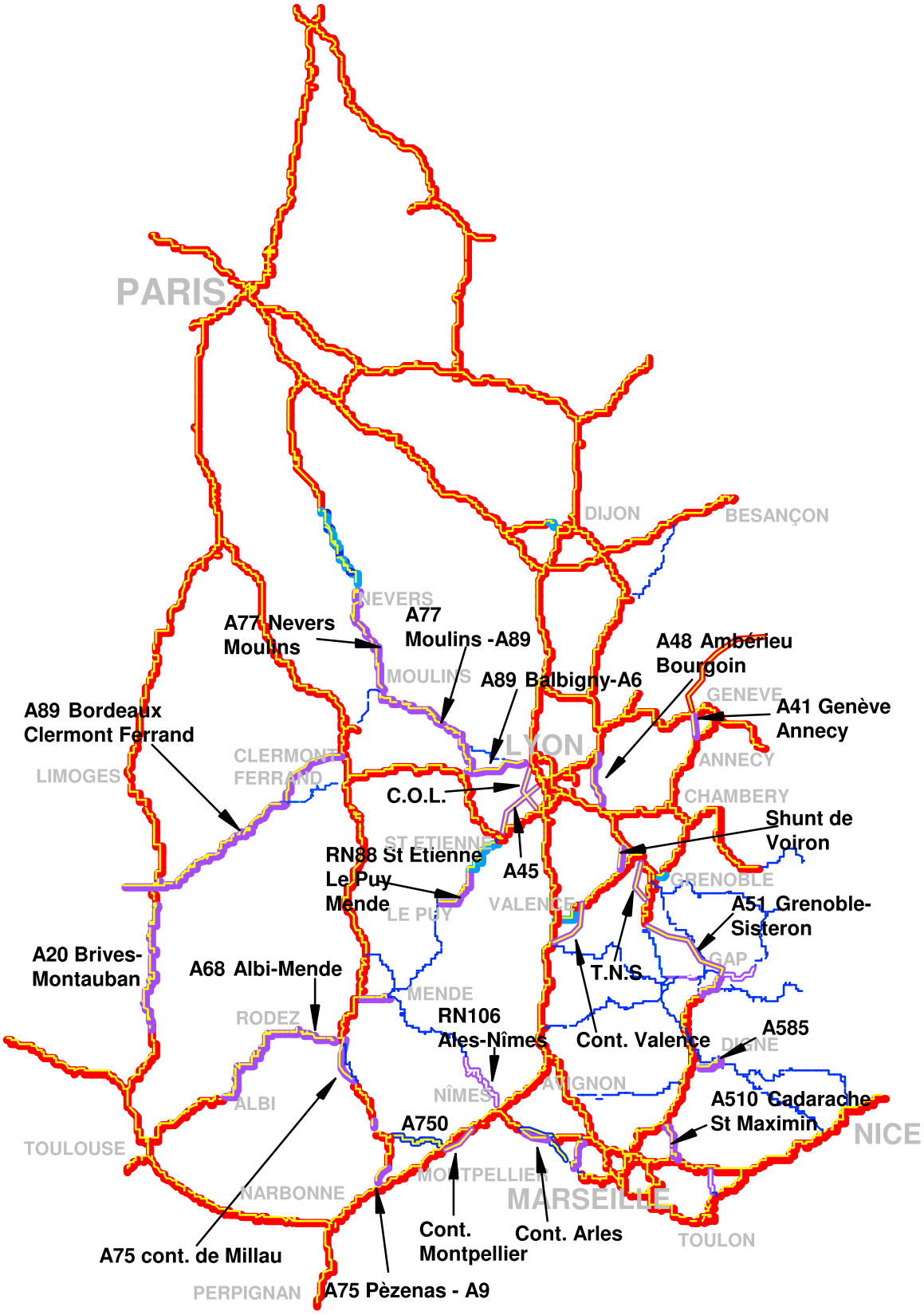


Figure E.2 Carte de comptages 2000

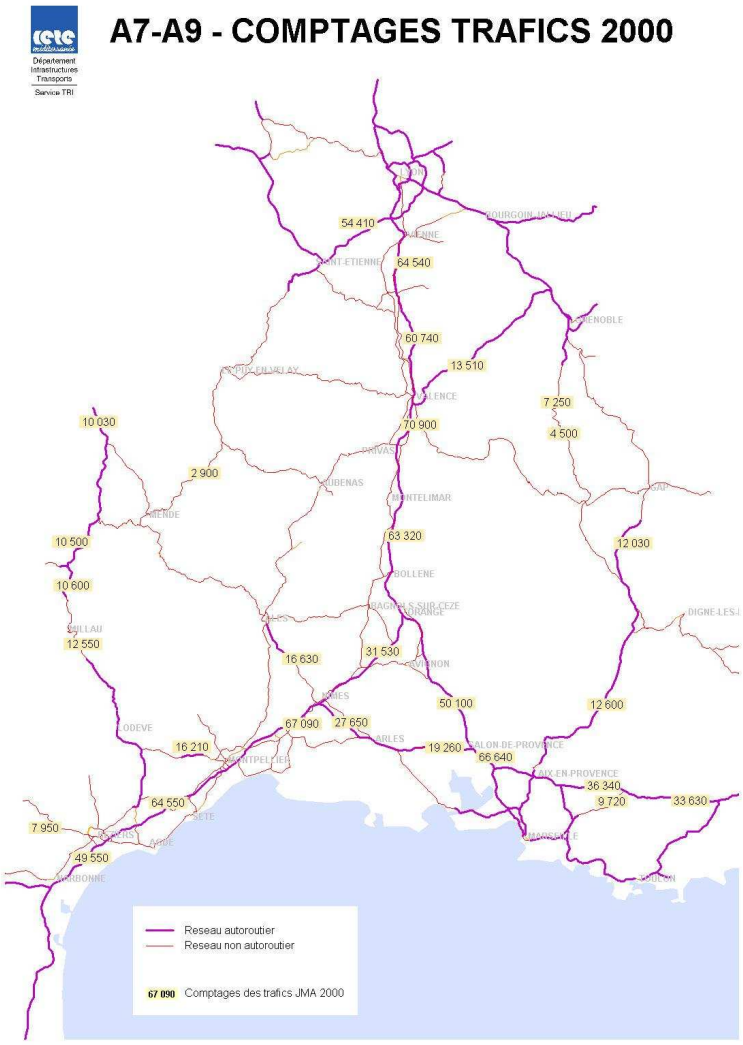


Figure E.3 Evolution tendancielle des trafics à l'horizon 2020

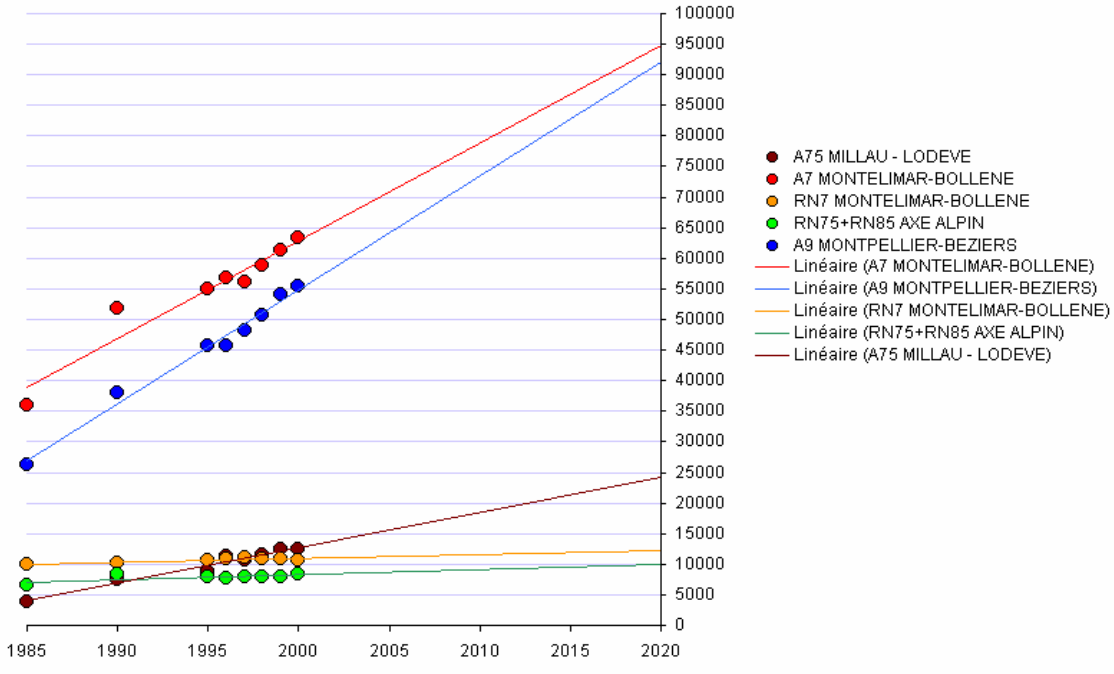
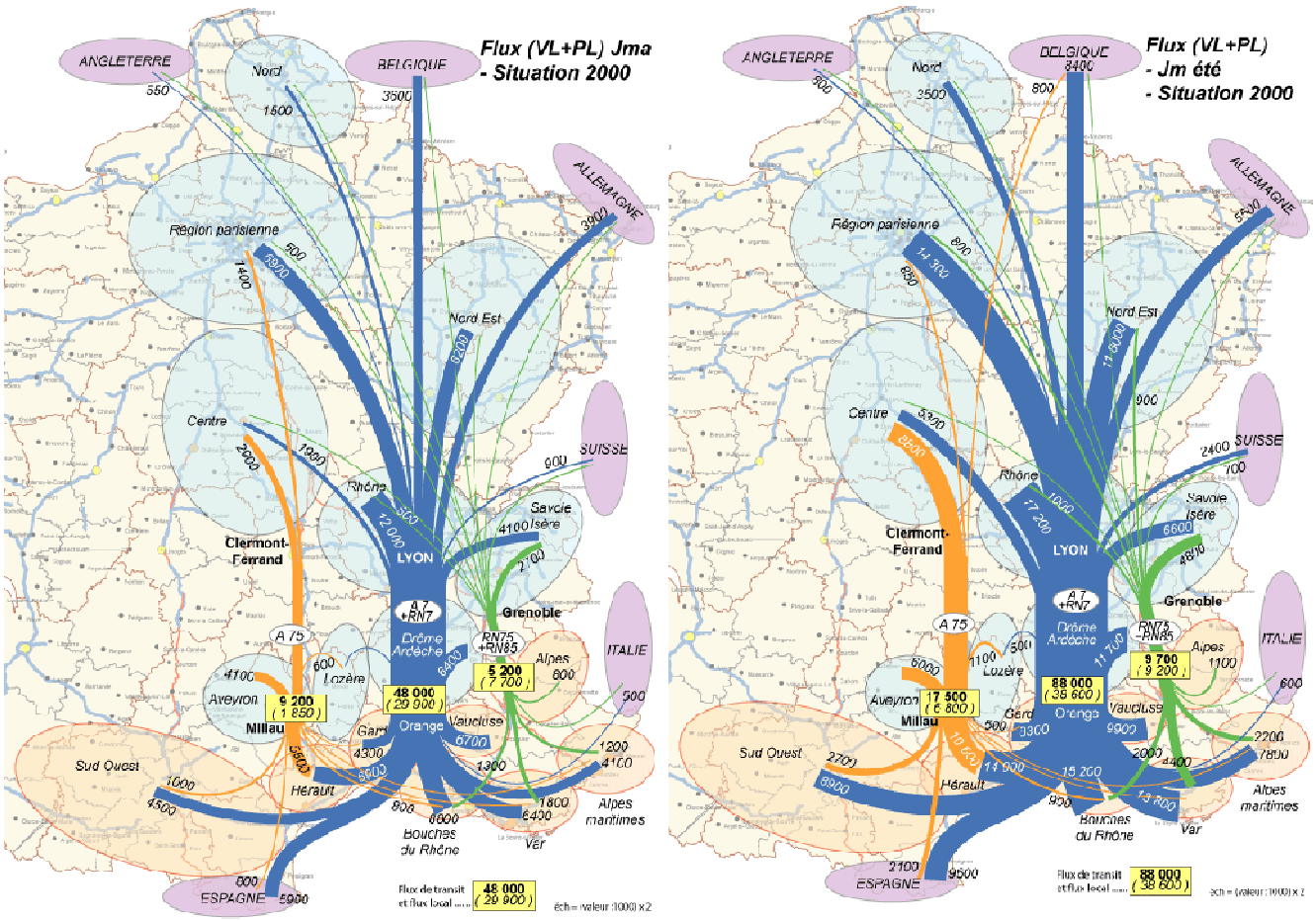


Figure E.4 Flux Nord-Sud (VL+PL) en TMJA et en TMJE situation 2000

(Source Cété Méditerranée)



E.3 Constitution des données pour une affectation dynamique

Afin de lancer une affectation du trafic sur le réseau VdR à l'aide du code en C++, nous avons utilisé les données fournies par le SETRA et le CETE Méditerranée. Ces données nous permettent de recoder la topologie du réseau, de passer d'une matrice statique è une matrice temporalisée de manière fictive, et de former les caractères physiques et économiques des services du réseau. Cette affectation est réalisée pour un horizon 2020.

E.3.1 Acquisition des données

Deux acteurs du RST (SETRA et CETE Méditerranée) ont nous fourni les données suivantes :

1. la topologie du réseau sous format de MAPINFO. Le réseau comprend de 1464 arcs (732 arcs indirects), de 595 nœuds avec 139 centroïdes et 1137 relations O-D.

2. pour chaque arc, sa capacité q_a et sa longueur l_a, avec sa vitesse libre en moyenne mesurée.

3. pour la capacité des arcs, il y a cinq niveaux différents :

- 1800 véh/h : 471 arcs
- 2600 véh/h : 26 arcs
- 3600 véh/h : 145 arcs (correspondent à la section autoroutière de 2 voies par sens)
- 5400 véh/h : 81 arcs (correspondent à la section autoroutière de 3 voies par sens)
- La capacité infinie est appliquée aux connecteurs (vers Belgique, Allemande, Angleterre, Espagne, Italie, etc.)

4. une matrice O-D statique représentant la moyenne journalière de la demande pour 1137 relations O-D en 2000, à la fois les PL et les VL, y compris la demande des connecteurs.

E.3.2 Génération des données

Pour notre expérimentation, les données ci-dessus sont utilisées afin de générer les inputs du simulateur après les étapes suivantes :

Etape 1 : génération des arcs directs. Pour chaque arc indirect a qui joint l'ensemble des nœuds $\{m,n\}$, deux arcs directs $a_1=(m,n)$ et $a_2=(n,m)$ ont été créés.

Etape 2 : génération des données des arcs. Pour chaque arc indirect, le profil du temps de parcours libre et de la capacité des arcs direct associés a_1 et a_2 sont établis comme :

$$t_{a_1,0}(h)=t_{a_2,0}(h)=\frac{t_a}{v_a}$$

$$K_{a_1}(h)=K_{a_2}(h)=q_a\cdot h\,.$$

Etape 3 : génération de la demande. Pour chaque couple O-D i, le profil du volume cumulé $X_i^+(h)$ a été établi. Le volume moyen de la demande journalière Q_i a été distribué sur la tranche 0-24 selon l'équation suivante :

$$\frac{dX_i^+}{dt}(h)=w_{k(h)}Q_i\text{, avec } k(h)=k \text{ pour } k-1\le h<k \text{ et } \sum_{k=1}^{24}w_k=1{,}8$$

Les valeurs des coefficients w_k sont présentées dans la figure E.5, dont la somme de 1,8 correspond à un samedi en été à l'horizon 2020. Le choix de la valeur de cette somme est fondé sur l'évolution de trafic prévue par l'étude de CETE Méditerranée présentée dans la figure E.3. Cette méthode de distribution du volume de la demande permet de constituer une matrice O-D temporalisée afin de simuler le trafic d'une journée avec deux périodes de l'heure de point du martin et du soir. Il semble que cette méthode ne permet pas d'établir l'évolution réelle de la demande, mais cela ne fait pas le conflit avec notre principal objectif.

E.4 Premier résultat

Le problème de fragmentation des profils temporels discuté dans la section 5.6 devient très important dans le cas où nous traitons une affectation du trafic sur un réseau de grande taille. Cette remarque a deux mauvaises conséquences en pratique :

- le temps de réponse par itération augmente avec le nombre d'itérations.

- l'occupation de la mémoire augmente en fonction du nombre d'itérations.

Afin de mieux illustrer la performance du simulateur du modèle LADTA, nous avons enregistré le temps de réponse et l'occupation de la mémoire au cours des itérations pendant une affectation du trafic sur le réseau VdR. Cette affectation est opérée par un ordinateur portable équipé d'un processeur Intel Pentium de 1.6 GHz, 512MB de mémoire, sous système d'exploitation Windows XP. Le code du modèle, programmé en langage C++, est compilé avec g++, avec toutes les optimisations standard ouvertes. Les résultats, illustrés en figures E.6 et E.7 permettent de conclure que le code est performant. Une situation de trafic à l'heure de point du martin, qui résulte d'une simulation par Code en C++, est présentée dans la figure E.8.

Figure E.5 valeurs des coefficients w_k utilisées pour génération du profil de la demande profiles.

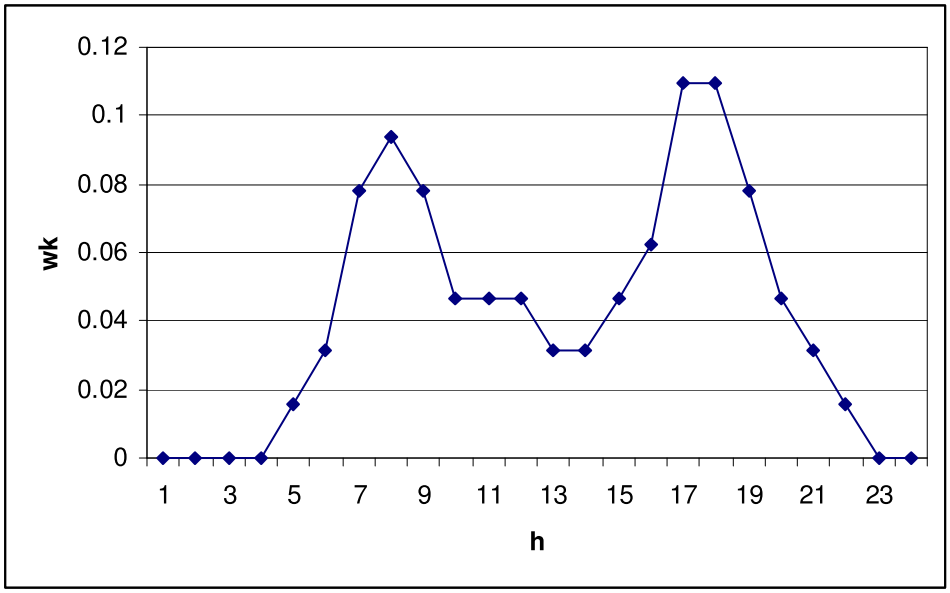


Figure E.6 évolution du temps de réponse au cours des itérations

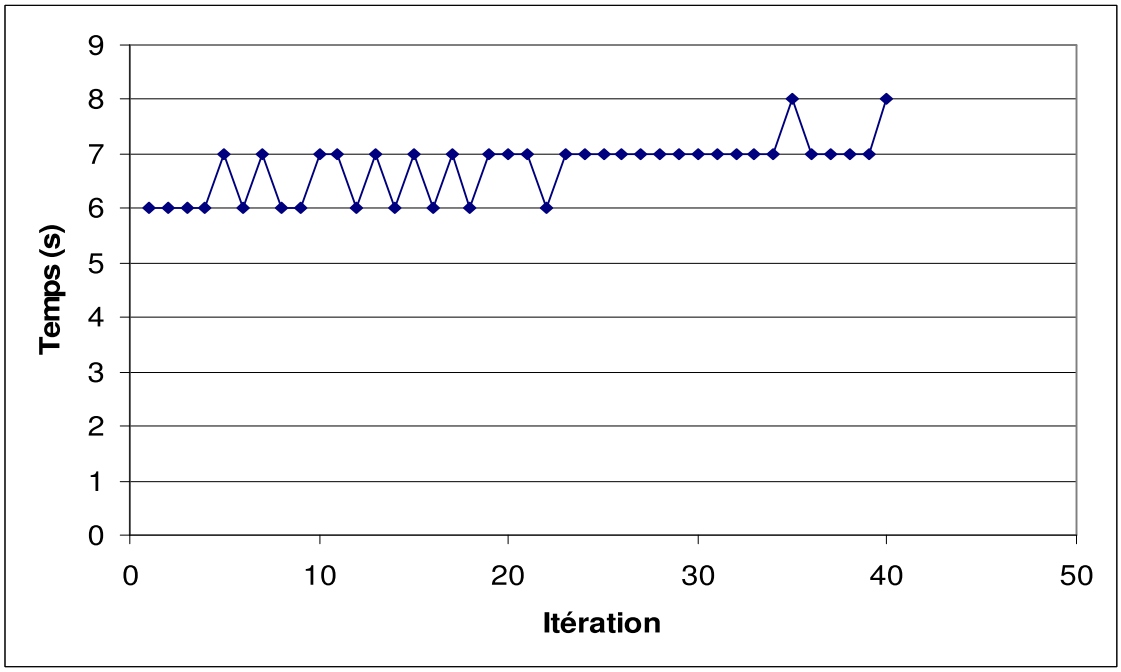


Figure E.7 évolution due l'occupation de la mémoire au cours des itérations

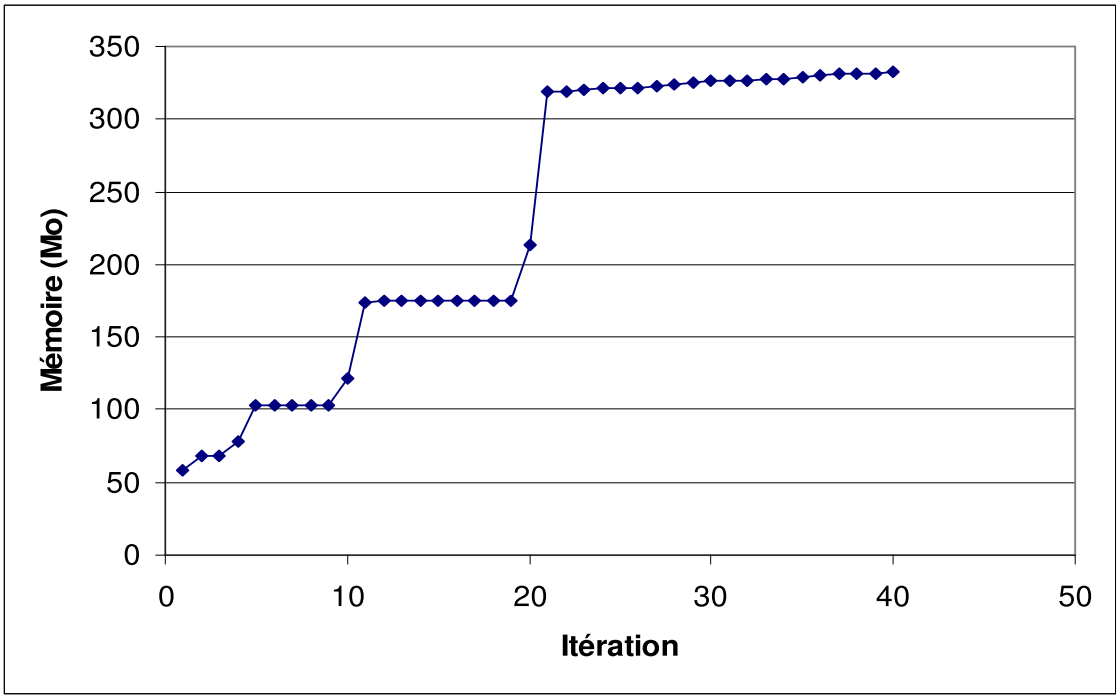
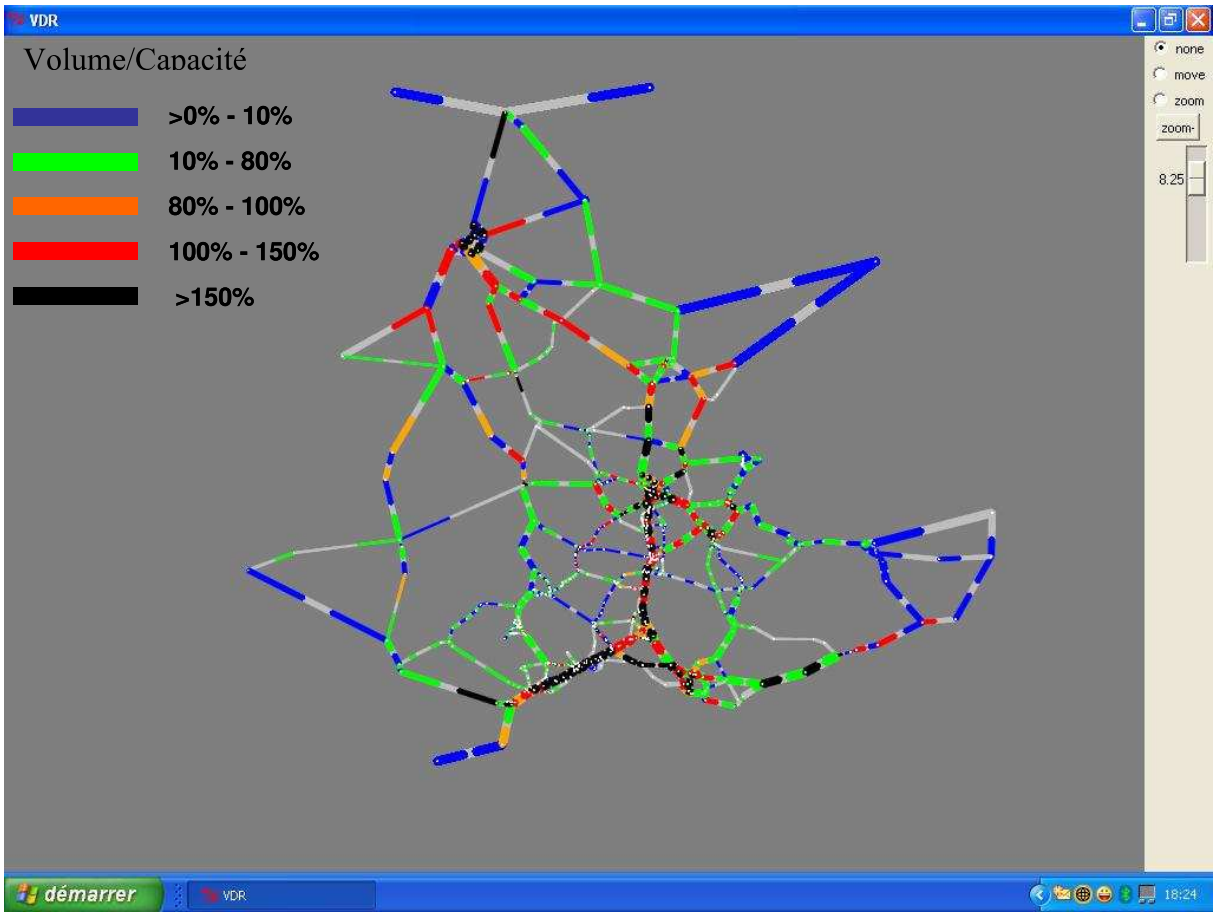


Figure E.8 une situation de trafic à la Vallée du Rhône simulée par Code en C++



Bibliographique

[CETE04] CETE Méditerranée (2004). Les études de trafic sur la vallée du Rhône et l'arc languedocien : prévision à l'horizon 2020. Document de référence, maître d'ouvrage : SETRA.