

Estimation du temps de parcours d'un réseau urbain par fusion de données de boucles magnétiques et de véhicules traceurs : Une approche stochastique avec mise en oeuvre d'un filtre de Kalman sans parfum

Ré-Mi Hage

▶ **To cite this version:**

Ré-Mi Hage. Estimation du temps de parcours d'un réseau urbain par fusion de données de boucles magnétiques et de véhicules traceurs : Une approche stochastique avec mise en oeuvre d'un filtre de Kalman sans parfum. Traitement du signal et de l'image [eess.SP]. Université de Nantes, 2012. Français. NNT : . tel-00919589

HAL Id: tel-00919589

https://theses.hal.science/tel-00919589

Submitted on 17 Dec 2013

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.

UNIVERSITE DE NANTES

ÉCOLE DOCTORALE

« SCIENCES ET TECHNOLOGIES DE L'INFORMATION ET DE MATHEMATIQUES »

Année : 2012

Thèse de Doctorat de l'Université de Nantes

Spécialité : AUTOMATIQUE
Numéro de la thèse : ED 503-170

Présentée et soutenue publiquement par

Ré-Mi HAGE

le 25 Septembre 2012

à IFSTTAR-Nantes

Estimation du temps de parcours d'un réseau urbain par fusion de données de boucles magnétiques et de véhicules traceurs

Une approche stochastique avec mise en œuvre d'un filtre de Kalman sans parfum

Jury

Président	: Nour-Eddin EL-FAOUZI	Directeur de recherche, IFSTTAR
Rapporteurs	: Jaume BARCELÓ	Professeur, Université Polytechnique de Catalogne
	Thierry DENOEUX	Professeur, Université de Technologie de Compiègne
Examinateurs	: Nour-Eddin EL-FAOUZI	Directeur de recherche, IFSTTAR
	Eric LE CARPENTIER	Maitre de conférences, Ecole Centrale de Nantes
	Yide WANG	Professeur, Université de Nantes
	François PEYRET	Directeur de recherche, IFSTTAR
	Dominique MEIZEL	Professeur, XLIM, Université de Limoges
	David BETAILLE	Chargé de recherche, IFSTTAR

Directeur de Thèse : François PEYRET
Laboratoire : IFSTTAR
Co-Directeur : Dominique MEIZEL
Laboratoire : XLIM
Encadrant : David BETAILLE
Laboratoire : IFSTTAR

Résumé

La notion de temps de parcours est une information simple à intégrer par les usagers des transports et a le potentiel, via des systèmes d'assistance, de réduire la congestion aussi bien de manière temporelle que spatiale. Cette thèse porte sur la problématique de l'estimation du temps de parcours en ville, où la circulation se caractérise par un débit variable et un flux significatif de/vers des voies transversales non équipées de capteurs de trafic, flux qui agit comme perturbation de l'équation de conservation des véhicules. La bibliographie présente des approches déterministes et stochastiques, utilisant, selon les situations expérimentales, soit des boucles inductives, soit des véhicules traceurs, soit les deux. Nous proposons une méthode stochastique du type filtrage de Kalman sans parfum. Son modèle dynamique est basé sur la méthode analytique classique des stocks qui considère le décalage temporel entre les cumuls de véhicules entrants et sortants dans chaque tronçon du réseau. La formulation de ce modèle n'est pas explicite, ce qui justifie l'utilisation d'un filtre sans parfum. Les observations utilisées sont à la fois les boucles magnétiques à la sortie des tronçons et les véhicules traceurs après *map-matching* sur le réseau. L'algorithme est applicable en temps réel, sans connaissance *a priori* des affectations aux carrefours, et sous conditions éventuelles de perturbation. En sortie de filtre, on dispose des variances d'estimation, ce qui mesure la confiance dans le temps de parcours estimé et permet de rejeter des mesures aberrantes. La validation a été montrée sur la base de simulations sur un réseau simple à trois intersections, avec des flux variables et des perturbations, et pour quelques pourcents de véhicules traceurs.

Mots clés : trafic, temps de parcours, boucle magnétique, véhicule traceur, map-matching, filtrage de Kalman sans parfum, réseau urbain.

Abstract

Travel time information is easy to be perceived by users and has the potential to reduce congestion on both temporal and spatial scales thanks to traveller assistance systems. This thesis addresses the problem of estimating the travel time in city, where traffic is characterized by variable flow and also significant perturbation from/to mid-link source/sink that are not equipped with traffic sensors. This flow disturbs the equation of conservation of vehicles. The bibliography contains deterministic and stochastic approaches, using, in specific experimental situations, either inductive loops or probe vehicles, or both. We propose a stochastic approach based on unscented Kalman filtering. Its dynamic model is based on the classical analytical procedure that considers the time difference between the total of incoming and outgoing vehicles for each segment of the network. The formulation of this model is not explicit, which justifies the use of an unscented filter. The observations are both loop counters at the exit sections and probe vehicles map-matched to the network. The algorithm is applicable in real time, without prior information on the percentage of turning movements at intersections, and under any condition of perturbation. The variance estimation output by the filter measures the confidence in the estimated travel time and allows the rejection of outliers. The validation was shown based on simulations on a simple network with three intersections, variable entering flows, variable percentage of perturbation, and a variable percentage of probe vehicles.

Key words: traffic, travel time, loop detector, probe vehicles, map-matching, unscented Kalman filter, urban network.

À mes parents et mes sœurs

À mon fiancé Saad

Remerciements

Je remercie vivement Monsieur François Peyret, Chef du groupe géolocalisation du département mesure, auscultation et calcul scientifique (MACS) dans le laboratoire IFSTTAR et directeur de ma thèse, pour m'avoir permis de travailler au sein du laboratoire et à qui je voue tant d'estime et de reconnaissance pour sa grande expérience.

Je remercie Monsieur Dominique Meizel, responsable de la spécialité Mécatronique à XLIM et mon co-directeur, de m'avoir guidée et accordé sa confiance dans mon travail.

Au terme de ce travail, je tiens à exprimer ma gratitude et mon respect envers Monsieur David Bétaille, chargé de recherche du groupe géolocalisation à IFSTTAR et encadrant de ma thèse, pour sa disponibilité et le temps qu'il a su me consacrer, pour tous ses conseils et toutes ses idées sans lesquels ce travail de thèse n'aurait pu aboutir, pour son suivi au quotidien de l'avancée de mon travail et surtout pour ses encouragements.

Je n'oublie pas non plus toutes les personnes qui ont contribué, d'une façon directe ou indirecte, continue ou ponctuelle, à l'achèvement de ce travail. Un grand merci à Nour-Eddin El-Faouzi de l'IFSTTAR/LICIT et Eric Le Carpentier de l'ECN/IRCCyN pour leurs précieux conseils.

J'adresse également mes remerciements à Monsieur Jaume Barceló et Monsieur Thierry Denœux qui m'ont fait l'honneur d'être les rapporteurs de ce travail.

J'associe à ces remerciements un regard particulier aux collègues du groupe MACS à IFSTTAR : Sylvie Chambon, Jean-Marc Molliard, Philippe Nicolle, Christian Gourraud, Valérie Renaudin, Ahmed Selloum, Miguel Ortiz et William Guillemaud pour leur amitié et pour m'avoir assuré les conditions les plus agréables de travail.

Un grand merci à mes parents qui m'ont donné la vie, m'ont inspiré et m'ont servi de modèles. J'espère que vous êtes fiers de moi. Je remercie également mes sœurs Ilige, Hiba et Rita-Héléna, ma nièce Alexandra et mon neveu Christopher pour leur amour et leur soutien.

Finalement, je réserve une place singulière à mon fiancé Saad pour sa confiance, son soutien et son amour.

Table des matières

RÉSUMÉ - 1 -

CHAPITRE 1. INTRODUCTION - 5 -

1.1 Contexte et problématique - 7 -

1.2 Objectif du travail de thèse - 9 -

1.3 Plan du mémoire - 10 -

CHAPITRE 2. RAPPEL SUR LE RÉSEAU ET LE TRAFIC ROUTIER - 13 -

2.1 Description d'un réseau routier et spécificité d'un réseau routier urbain - 15 -

2.2 Variable du trafic - 16 -
2.2.1 Variables microscopiques - 16 -
2.2.2.1 Vitesse individuelle - 17 -
2.2.2 Variables macroscopiques - 17 -
2.2.2.1 Débit du trafic - 17 -
2.2.2.2 Densité du trafic - 18 -
2.2.2.3 Taux d'occupation - 18 -
2.2.2.4 Relation entre les variables : vitesse du flot - 19 -
Cas isovélocique - 19 -
Cas général - 19 -

2.3 Théorie du trafic - 20 -
2.3.1 Équation de conservation - 21 -
2.3.2 Diagramme fondamental - 21 -
2.3.3 Méthode analytique classique des stocks - 23 -
2.3.3.1 Principe de fonctionnement - 23 -
2.3.3.2 Limitation dans les réseaux urbains - 25 -
2.3.3.3 Application - 26 -
Estimation du temps de parcours sous la loi de conservation des véhicules - 26 -
Estimation du temps de parcours d'un tronçon urbain - 28 -
Effet des valeurs aberrantes - 29 -

2.4 Principaux capteurs de mesure de trafic - 30 -
2.4.1 Capteurs fixes : boucles magnétiques - 30 -
2.4.1.1 Principe de fonctionnement - 31 -
Boucles magnétiques simples - 31 -
Boucles magnétiques doubles - 32 -
2.4.1.2 Caractéristiques des erreurs - 33 -
2.4.2 Capteurs mobiles : véhicules traceurs - 33 -
2.4.2.1 Principe de fonctionnement - 34 -
2.4.2.2 Caractérisation des erreurs - 34 -
2.4.3 Synthèse et conclusion - 37 -

CHAPITRE 3. LE TEMPS DE PARCOURS - 39 -

3.1 Définition - 41 -

3.2 Principales méthodes d'estimation - 41 -
3.2.1 À partir des boucles magnétiques - 42 -
3.2.1.1 Principes d'utilisation - 42 -
3.2.1.2 Approches déterministes - 42 -
Méthode de Nam et Drew - 43 -
Description et procédure - 43 -
Analyse et critique - 45 -
Méthode d'Oh et al - 46 -
Description et procédure - 46 -
Analyse et critique - 47 -
Méthode de Bonvalet et Robin-Prévallée (BRP) - 47 -
Description et procédure - 47 -
Analyse et critique - 48 -
3.2.1.3 Approches stochastiques - 50 -
Méthode de Dailey - 50 -
Description et procédure - 50 -
Analyse et critique - 51 -
Méthode de Ye et al. - 52 -
Description et procédure - 52 -
Analyse et critique - 52 -
Méthode d'Hinsbergen et al. - 53 -
Description et procédure - 53 -
Analyse et critique - 54 -
3.2.2 À partir des véhicules traceurs - 55 -
3.2.2.1 Principe d'utilisation - 55 -
3.2.2.2 Méthode d'Herring et al. - 55 -
Description et procédure - 55 -
Analyse et critique - 56 -
3.2.2.3 Problème du nombre relatif de véhicules traceurs - 57 -
3.2.3 Par fusion des données des boucles magnétiques et des véhicules traceurs - 59 -
3.2.3.1 Approches déterministes - 60 -
Méthode d'Ashish Bhaskar : CUPRITE - 60 -
Description et procédure - 60 -
i Estimation des courbes cumulées - 60 -
ii Ajustement des courbes cumulées par les véhicules traceurs - 61 -
Analyse et critique - 62 -
Application de CUPRITE - 63 -
i. Estimation du temps de parcours par CUPRITE et CUPRITE temps réel - 64 -
ii. Sensibilité de CUPRITE à l'aléa des véhicules traceurs - 66 -
iii. Effet des valeurs aberrantes - 67 -
3.2.3.2 Méthodes stochastiques - 69 -
Méthode d'El Faouzi et al. - 69 -
Description et procédure - 69 -
Analyse et critique - 70 -
Méthode de Choi et Chung - 70 -
Description et procédure - 70 -
Analyse et critique - 71 -
3.2.4 Synthèse et conclusion - 72 -

CHAPITRE 4. LE FILTRAGE DE KALMAN ET SES EXTENSIONS - 77 -

4.1 Introduction - 79 -

4.2 Le filtre de Kalman - 79 -

4.3 Le filtre de Kalman étendu - 81 -

4.4 Le filtre de Kalman sans parfum - 82 -
4.4.1 La transformation sans parfum - 83 -

4.4.2 Le filtrage sans parfum - 84 -

4.5 Conclusion - 85 -

CHAPITRE 5. MÉTHODE PROPOSÉE - 87 -

5.1 Introduction - 89 -

5.2 Choix des capteurs - 90 -

5.3 Type de réseau - 92 -

5.4 Principe de notre approche - 92 -

5.5 Description de la modélisation proposée - 93 -
5.5.1 Introduction - 93 -
5.5.2 Vecteur d'état - 95 -
5.5.3 Modèle d'évolution - 96 -
5.5.4 Erreur d'évolution - 98 -
5.5.5 Modèle d'observation - 99 -
5.5.6 Erreur d'observation - 99 -
5.5.6.1 Principales caractéristiques des données de position - 100 -
5.5.6.2 Principales caractéristiques de la carte numérique - 101 -
5.5.6.3 Map-matching - 103 -
5.5.6.4 Estimation des erreurs - 103 -
5.5.7 L'algorithme - 104 -

5.6 Choix de la méthode d'estimation - 105 -

5.7 Test de performance - 106 -
5.7.1 Erreur - 107 -
5.7.2 Erreur absolue - 107 -
5.7.3 Erreur absolue relative - 107 -
5.7.4 Pourcentage d'erreur absolue moyenne - 107 -
5.7.5 Précision - 107 -

CHAPITRE 6. APPLICATION DE LA MÉTHODE EN SIMULATION SUR UN TRONÇON SANS ET AVEC PERTURBATIONS - 109 -

6.1 Introduction - 111 -

6.2 Description du simulateur utilisé - 111 -

6.3 Description des scénarios choisis - 112 -

6.4 Méthodes appliquées - 114 -

6.5 Mise au point de notre algorithme - 115 -

6.6 Résultats et comparaisons - 117 -
6.6.1 Cas d'un tronçon sans perturbations - 117 -
6.6.1.1 Comparaison de l'UKF à la méthode analytique classique des stocks - 118 -
6.6.1.2 Comparaison de l'UKF à CUPRITE - 119 -
Sensibilité de l'UKF et de CUPRITE à l'aléa des véhicules traceurs - 119 -
Comparaison des résultats - 120 -
6.6.1.3 Estimation de la perturbation - 121 -
6.6.1.4 Erreur d'estimation - 122 -
6.6.1.5 Effet des valeurs aberrantes - 123 -

6.6.2 Cas d'un tronçon avec perturbations - 125 -
6.6.2.1 Comparaison des résultats - 125 -
6.6.2.2 Estimation de la perturbation - 129 -

6.7 Analyse et synthèse - 130 -

CHAPITRE 7. MÉTHODE ÉTENDUE À UN RÉSEAU - 131 -

7.1 Introduction - 133 -

7.2 Description du modèle générique - 134 -
7.2.1 Vecteur d'état - 134 -
7.2.2 Modèle d'évolution - 135 -
7.2.3 Modèle d'observation - 135 -
7.2.4 Erreur d'observation - 136 -
7.2.5 L'algorithme - 136 -

7.3 Description du modèle dans le cas d'un divergent - 138 -

7.4 Description du modèle dans le cas d'un convergent - 141 -

CHAPITRE 8. APPLICATION DE LA MÉTHODE ÉTENDUE À UN DIVERGENT, À UN CONVERGENT ET À UN RÉSEAU - 145 -

8.1 Introduction - 147 -

8.2 Application pour le cas d'un divergent - 147 -
8.2.1 Description des scénarios choisis - 147 -
8.2.2 Le comportement de CUPRITE - 149 -

8.3 Application pour le cas d'un convergent - 155 -
8.3.1 Description des scénarios choisis - 155 -
8.3.2 Le comportement de CUPRITE - 155 -
8.3.3 Estimation par l'UKF du nombre cumulé en amont - 156 -
8.3.4 Estimation du temps de parcours - 156 -

8.4 Application dans un réseau - 159 -
8.4.1 Description des scénarios choisis - 159 -
8.4.2 Estimation du temps de parcours - 160 -
8.4.3 Estimation de la perturbation - 162 -

8.5 Synthèse et analyse - 163 -

CHAPITRE 9. CONCLUSION ET PERSPECTIVES - 165 -

9.1 Conclusion et contribution - 167 -

9.2 Perspectives - 169 -

ANNEXE : MAP-MATCHING - 173 -

RÉFÉRENCES - 177 -

PUBLICATIONS - 185 -

Liste des figures

Figure 2. 1 : Représentation d'un réseau routier - 15 -
Figure 2. 2 : Diagramme fondamental de Greenshield - 22 -
Figure 2. 3 : Diagramme fondamental - 22 -
Figure 2. 4 : Diagramme fondamental triangulaire - 23 -
Figure 2. 5 : Nombre cumulé des véhicules - 24 -
Figure 2. 6 : Estimation du temps de parcours par la méthode analytique classique des stocks - 25 -
Figure 2. 7 : Tronçon étudié - 26 -
Figure 2. 8 : Variation du débit - 27 -
Figure 2. 9 : Temps de parcours de référence - 27 -
Figure 2. 10 : Erreur commise par la méthode analytique classique de stock - 28 -
Figure 2. 11 : Tronçon étudié avec source/fuite à mi-lien - 28 -
Figure 2. 12 : Erreur commise par la méthode analytique classique de stock avec (a) 1% (b) 5% de fuite/source. - 29 -
Figure 2. 13 : Temps de parcours estimé par la méthode analytique classique (MACS) des stocks avec 2 valeurs aberrantes - 30 -
Figure 2. 14 : Boucle magnétique intégrée dans le sol - 31 -
Figure 2. 15 : Schéma d'une boucle magnétique simple - 31 -
Figure 2. 16 : Véhicule de sonde au cas de tourne-à-gauche - 36 -

Figure 3. 1 : Figure représentant le temps de parcours total entre t_{n-1} et t_n en condition normale de circulation. - 44 -
Figure 3. 2 : Figure représentant le temps de parcours total entre t_{n-1} et t_n en condition de congestion - 45 -
Figure 3. 3 : Temps de parcours estimé par BRP - 49 -
Figure 3. 4 : Correction de la courbe cumulée en amont - 61 -
Figure 3. 5 : Courbe cumulée d'amont ajustée par CUPRITE - 64 -
Figure 3. 6 : Courbe cumulée en amont ajustée par CUPRITE temps réel - 65 -
Figure 3. 7 : Courbes cumulées d'amont ajustées par CUPRITE et CUPRITE temps réel - 65 -
Figure 3. 8 : Courbe cumulée an amont ajustée avec perturbation - 66 -
Figure 3. 9 : Sensibilité de CUPRITE - 67 -
Figure 3. 10 : Effet d'une valeur aberrante de véhicule traceur sur le temps de parcours - 67 -
Figure 3. 11 : Effet d'une valeur aberrante de boucle magnétique sur le temps de parcours - 68 -

Figure 5. 1 : Propriétés des sources de données - 91 -
Figure 5. 2 : Comparaison entre réseau inter-urbain et réseau urbain - 92 -
Figure 5. 3 : Tronçon avec un pourcentage connu d'affectation - 94 -
Figure 5. 4 : Tronçon à étudier - 94 -
Figure 5. 5 : Tronçon avec une source/fuite à mi-lien - 95 -
Figure 5. 6 : Calcul du temps de parcours - 97 -
Figure 5. 7 : Véhicule VERT et sa centrale inertielle - 100 -
Figure 5. 8 : Extrait de pleine densité (aussi appelé : map-caché) : 1736 segment, 630 liens - 102 -
Figure 5. 9 : Extrait limité à la trajectoire pratiquée lors des tests : 305 segments, 131 liens - 102 -
Figure 5. 10 : Erreur de temps de parcours calculé par GPS et temps de parcours référence - 103 -
Figure 5. 11 : Algorithme proposé - 104 -
Figure 5. 12 : Algorithme de l'UKF - 106 -

Figure 6. 1 : Estimation du temps de parcours par la méthode analytique classique des stocks - 114 -
Figure 6. 2 (a) : TP UKF vs TP' UKF avec 1% de véhicule traceur - 116 -
Figure 6. 3 : Comparaison entre les temps de parcours estimés par la méthode analytique classique des stocks et par UKF - 118 -
Figure 6. 4 : (a) L'erreur de la méthode analytique classique des stocks en minute. (b) L'erreur de l'UKF en minute - 119 -
Figure 6. 5 : Sensibilité du temps de parcours estimé avec l'aléa des véhicules traceurs - 120 -
Figure 6. 6 : Erreur commise par CUPRITE en minute - 120 -
Figure 6. 7 : Erreur commise par CUPRITE temps réel en minute - 121 -
Figure 6. 8 : Erreur commise par UKF en minute - 121 -
Figure 6. 9 : Estimation de la perturbation - 122 -
Figure 6. 10 : Enveloppes d'erreurs - 123 -

Figure 6. 11 : Effet d'une valeur aberrante de boucle magnétique sur le temps de parcours ..... - 124 -
Figure 6. 12 : Effet d'une valeur aberrante de véhicule traceur sur le temps de parcours ..... - 124 -
Figure 6. 13 : Estimation du temps de parcours par MACS (a) fuite, (b) source ..... - 125 -
Figure 6. 14 : Estimation du temps de parcours par MACS, UKF, CUPRITE et CUPRITE temps réel avec 1% de véhicules traceurs (a) fuite, (b) source ..... - 126 -
Figure 6. 15 : Estimation du temps de parcours par MACS, UKF, CUPRITE et CUPRITE temps réel avec 5% de véhicules traceurs (a) fuite, (b) source ..... - 126 -
Figure 6. 16 : Estimation du temps de parcours par MACS, UKF, CUPRITE et CUPRITE temps réel avec 10% de véhicules traceurs (a) fuite, (b) source ..... - 126 -
Figure 6. 17 : Erreur de la méthode analytique classique des stocks (a) 1% fuite/source, (b) 5% fuite/source. - 127 -
Figure 6. 18 : Erreur CUPRITE (a) 1% fuite/source, (b) 5% fuite/source ..... - 127 -
Figure 6. 19 : Erreur CUPRITE temps réel (a) 1% fuite/source, (b) 5% fuite/source ..... - 128 -
Figure 6. 20 : Erreur UKF (a) 1% fuite/source, (b) 5% fuite/source ..... - 128 -
Figure 6. 21 : Estimation de la perturbation avec (a) 1% source (b) 5% source (c) 1% fuite (d) 5% fuite ..... - 129 -

Figure 7. 1 : Réseau urbain ..... - 133 -
Figure 7. 2 : Algorithme général du modèle proposé ..... - 137 -
Figure 7. 3 : Tronçons divergents ..... - 138 -
Figure 7. 4 : Tronçons convergents ..... - 141 -

Figure 8. 1: Correction des nombres cumulés de véhicules en amont par CUPRITE ..... - 150 -
Figure 8. 2 : Estimation du nombre cumulé de véhicules à l'amont du deuxième tronçon ..... - 150 -
Figure 8. 3 : Erreur d'estimation du nombre cumulé de véhicules à l'amont du deuxième tronçon ..... - 151 -
Figure 8. 4 : Estimation du temps de parcours avec (a) 0%, (b) 1%, (c) 5%, (d) 10% de véhicules traceurs ..... - 152 -
Figure 8. 5 : Effet de la prédiction ..... - 152 -
Figure 8. 6 : Erreur d'estimation du temps de parcours pour le premier tronçon ..... - 153 -
Figure 8. 7 : Erreur d'estimation du temps de parcours pour le deuxième tronçon ..... - 153 -
Figure 8. 8 : Erreur d'estimation du temps de parcours pour le troisième tronçon ..... - 154 -
Figure 8. 9 : Erreur d'estimation du nombre cumulé de véhicules à l'amont du troisième tronçon ..... - 156 -
Figure 8. 10 : Estimation du temps de parcours avec (a) 0%, (b) 1%, (c) 5%, (d) 10% de véhicules traceurs - 157 -
Figure 8. 11 : Erreur d'estimation du temps de parcours pour le premier tronçon ..... - 158 -
Figure 8. 12 : Erreur d'estimation du temps de parcours pour le deuxième tronçon ..... - 158 -
Figure 8. 13 : Erreur d'estimation du temps de parcours pour le troisième tronçon ..... - 159 -
Figure 8. 14 : Erreur d'estimation du temps de parcours pour (a) TR(1), (b) TR(2), (c) TR(3), (d) TR(4), (e) TR(5), (f) TR(6) ..... - 161 -
Figure 8. 15 : Estimation de la perturbation pour (a) 3% source TR(3), (b) 3% source TR(4), (c) 1% fuite TR(2), (d) 1% fuite TR(5). ..... - 163 -

Liste des tableaux

Tableau 3. 1 : Résumé des méthodes d'estimation du temps de parcours .....- 74 -

Tableau 6. 1 : Précision moyenne (100-MAPE%) en %. .....- 129 -

Tableau 8. 1 : Précision moyenne (100-MAPE%) en % pour un divergent. .....- 154 -
Tableau 8. 2 : Précision moyenne (100-MAPE%) en % pour un convergent.....- 159 -
Tableau 8. 3 : Précision moyenne (100-MAPE%) en % pour un réseau. .....- 162 -

Résumé

La notion de temps, sa perception, son utilisation et sa gestion ont pris une importance essentielle dans l'organisation et le fonctionnement de nos sociétés. Avec l'augmentation du trafic, les déplacements sont de plus en plus affectés par des retards. Des systèmes intelligents d'information pour le transport « *Advanced Traveler Information Systems* » ont tenté de réduire l'incertitude sur l'estimation du temps de parcours. En effet, l'efficacité de l'exploitation des réseaux de transport dépend des performances des stratégies de gestion du trafic et du développement en temps réel des systèmes de guidage des véhicules, afin que les automobilistes puissent adapter leur stratégie de conduite en fonction du temps de parcours. En conséquence, une bonne estimation du temps de parcours s'avère essentielle pour l'efficacité et la capacité du réseau routier. Une mauvaise estimation du temps de parcours entraîne des coûts indirects pour les conducteurs en termes de temps perdu, d'inconfort et d'insatisfaction, et des coûts directs en termes de consommation de carburant. La notion de temps de parcours est une information simple à intégrer pour les utilisateurs et a le potentiel de réduire la congestion aussi bien de manière temporelle que spatiale. En réduisant la congestion, les émissions des véhicules et la consommation d'énergie seront également réduites et, ainsi, l'effet des transports sur le réchauffement climatique diminuera.

Différentes techniques sont utilisées pour estimer les temps de parcours routiers. Ces techniques dépendent du type de système utilisé pour recueillir des données relatives au trafic. Ces systèmes sont soit des capteurs fixes, soit des capteurs mobiles. Les capteurs fixes, tels que les détecteurs à boucles magnétiques, sont les plus anciens et les plus largement utilisés comme sources de données de trafic. Ils fournissent des informations temporelles sur l'état du trafic relatives à l'emplacement des détecteurs. Sous certaines hypothèses, des chercheurs ont proposé des méthodes d'estimation du temps de parcours qui peuvent être classées en deux catégories distinctes : les approches déterministes et les approches stochastiques.

Les capteurs mobiles, comme les véhicules sondes appelés aussi véhicules traceurs, sont des véhicules embarquant des équipements de localisation de véhicule (par exemple « *Global Positioning System* », (GPS)). Ils fournissent des données sur la trajectoire telles que les coordonnées de position ou la vitesse. Ces véhicules particuliers représentent un échantillon aléatoire de la population des véhicules dans le réseau. Par conséquent, la précision de l'estimation du temps de parcours par les véhicules traceurs est liée à leur nombre. Des

chercheurs se sont concentrés sur la détermination du nombre minimal de véhicules nécessaires pour une estimation significative du temps de parcours.

Les propriétés de ces deux sources de données sont complémentaires. Par conséquent, elles peuvent être exploitées en développant une solution de fusion des données multi-capteurs pour le problème de l'estimation du temps de parcours. La grande majorité des recherches sur l'estimation du temps de parcours portent sur les autoroutes. Or, le contexte urbain est différent du contexte inter-urbain en raison des fluctuations et des perturbations du débit dues, entre autres, aux feux de signalisation et au débit non négligeable provenant des rues transversales, des parkings, les propriétés privées, etc. qui agissent comme fuite ou source, ou les deux à la fois. De plus dans de tels réseaux les pourcentages des mouvements tournants aux carrefours sont généralement inconnus.

Dans ce contexte, l'algorithme proposé dans le cadre de cette thèse est basé sur la méthode analytique classique des stocks pour l'estimation du temps de parcours : elle estime les temps de parcours à partir des courbes cumulatives des nombres de véhicules provenant des capteurs en amont et en aval du lien étudié. Cependant, cette méthode est sujette à des erreurs principalement causées par le débit fluctuant et les erreurs de mesure des capteurs. De plus, la majorité des réseaux urbains sont équipés de capteurs à la sortie du tronçon et non pas à l'entrée. Aux carrefours, souvent les affections des mouvements tournants sont inconnues. Il est donc difficile d'appliquer cette méthode dans de telles conditions. Dans ce travail, une méthode stochastique du type filtrage de Kalman sans parfum est développée pour estimer le temps de parcours et les problèmes évoqués sont pris en considération. Son évolution est basée sur la méthode analytique classique des stocks. Les observations sont issues de données multi-sources : les boucles magnétiques à la sortie des tronçons et les véhicules traceurs.

L'originalité de cet algorithme est son application en temps réel grâce à l'utilisation des données *a priori* à leur instant exact d'occurrence. De plus, grâce au formalisme de ce modèle, on dispose des variances d'estimation, ce qui permet d'indiquer une mesure de la confiance dans le temps de parcours estimé par le filtre. Grâce à ce filtre, le nombre cumulé de véhicules en amont des liens des mouvements tournants est estimé sans connaître les affectations ainsi que celui des sources/fuites à mi-lien. Et finalement notre algorithme contient un test de χ^2 qui permet de tester directement les données des véhicules traceurs et les données des boucles magnétiques et de rejeter celles qui sont aberrantes.

Ce filtre a été testé avec des données simulées. Les pourcentages de véhicules traceurs sont variables (1% à 10%). Deux stratégies ont été établies connaissant ou non le pourcentage

d’affectation des mouvements tournants aux intersections. Chacune des stratégies est appliquée avec et sans sources et fuites. Les tests de validation de ce filtre ont montré que cette méthode peut être appliquée pour estimer le temps de parcours en zone urbaine en temps réel et avec une bonne fiabilité, quel que soit le pourcentage de véhicules traceurs (0%, 1%, 5%).

Chapitre 1. Introduction

SOMMAIRE

1.1	**Contexte et problématique**	**- 7 -**
1.2	**Objectif du travail de thèse**	**- 9 -**
1.3	**Plan du mémoire**	**- 10 -**

1.1 Contexte et problématique

La notion de temps, sa perception, son utilisation et sa gestion ont pris une importance essentielle dans nos sociétés. Dans le domaine des transports de personnes aussi bien que de marchandises, les gains de temps constituent une motivation majeure. C'est pour cela que le problème d'estimation du temps de parcours a été un problème étudié par de nombreux chercheurs.

Le temps de parcours est considéré comme l'une des mesures du degré de performance les plus importantes dans les études de transport. La congestion du trafic est une situation inévitable dans presque toutes les grandes villes. Cette congestion croissante du réseau routier a un impact essentiel sur l'environnement. Les résultats de la congestion sont les circulations de type accordéons (« *stop and go* »), ce qui non seulement augmente la consommation d'énergie mais ainsi provoque de la pollution de l'air. Ainsi, en réduisant l'encombrement, on assure la fluidité du trafic et on réduit la consommation d'énergie.

Le travail de cette thèse porte sur l'estimation du temps de parcours en milieu urbain. Essentiellement, l'estimation du temps de parcours peut être divisée en deux grandes catégories, à savoir, les méthodes indirectes et les méthodes directes suivant les types des sources de données utilisées.

Dans les méthodes indirectes, le temps de parcours est estimé ou calculé à partir des paramètres intermédiaires, comme par exemple la vitesse, grâce aux capteurs fixes tels que les détecteurs à boucles magnétiques (simples et doubles) qui sont les capteurs les plus utilisés. La vitesse d'un flux de trafic est égale au rapport du débit sur la concentration. Cette formule fournit une méthode de mesure possible de cette vitesse. Les boucles magnétiques simples ne fournissent ni la concentration ni la vitesse mais, en revanche, elles fournissent le débit et le taux d'occupation qui dans certaines conditions et sous certaines hypothèses peut être utilisé comme un estimateur de la concentration. Alors que les boucles magnétiques doubles fournissent directement la vitesse. Un des avantages des boucles est que les réseaux routiers européens sont équipés de ce type de détecteur. En outre, elles fournissent une base de données sur une longue période de temps. Toutefois, l'une des principales critiques sur les données des boucles est la forte probabilité d'erreur due, par exemple, à la défaillance de l'équipement. Ainsi, il est nécessaire de vérifier l'exactitude des données recueillies par les

boucles avant de les utiliser. De plus, ces mesures ne représentent l'état du trafic qu'à l'endroit où le capteur est placé et non pas sur l'ensemble d'un itinéraire ni même sur un tronçon de cet itinéraire.

Dans le cas des méthodes directes, le temps de parcours est collecté directement sur le terrain à partir des capteurs mobiles tels que les véhicules sondes ou les véhicules traceurs (*probe vehicles*). Historiquement, la méthode manuelle a été utilisée pour le recueil des données de temps de parcours. Dans ce cas, le conducteur (ou un passager) enregistre l'information des temps de parcours entre des points de contrôle prédéfinis. Actuellement, Le GPS, *Global Positioning System*, est l'outil technologique le plus utilisé pour collecter les données de temps de parcours et il permet l'automatisation de la méthode manuelle précitée. Ces données expriment l'état du trafic sur l'ensemble des tronçons d'un itinéraire. L'un des désavantages principaux des véhicules traceurs est que seuls quelques véhicules jouent ce rôle dans toute la population. Il faut donc estimer le nombre minimum nécessaire de ce type de véhicules pour une estimation acceptable du temps de parcours.

Ces deux sources de données ont des propriétés de complémentarité et de redondance qui peuvent être mises à profit en élaborant des méthodes de fusion de données pour estimer le temps de parcours urbain. Un atout majeur de la fusion de données est la possibilité d'exploiter au mieux les avantages de chacune des sources d'information, en palliant leurs limitations individuelles respectives (données manquantes, aberrantes...) pour fournir une meilleure image (plus globale et complète) de l'état du trafic.

Les recherches sur l'estimation du temps de parcours sont généralement adaptées et applicables sur des réseaux inter-urbains. Or, le contexte urbain est différent du contexte inter-urbain. La circulation urbaine est souvent interrompue par les feux de signalisation, panneaux d'arrêt, passages pour piétons, arrêts de bus, etc. Ceci entraîne, d'une part, une variation significative du temps de parcours entre deux véhicules consécutifs et, d'autre part, une fluctuation de la vitesse d'un véhicule (en fonction du temps). De plus, l'hypothèse de conservation des véhicules, hypothèse vérifiée dans un réseau inter-urbain, ne l'est plus dans un réseau urbain à cause des perturbations dans les flux. C'est pour cela que nous souhaitons adapter les stratégies d'estimation afin qu'elles répondent le mieux possible aux exigences urbaines.

Dans ce contexte, l'une des méthodes les plus récentes pour l'estimation du temps de parcours pour un lien urbain est CUPRITE [Bhaskar, 2009b] qui est basée sur la méthode analytique classique des stocks. CUPRITE corrige de façon déterministe les courbes cumulées

en amont en utilisant les temps de parcours des véhicules traceurs. La correction est établie à chaque occurrence des véhicules traceurs, non seulement à partir de cette occurrence, mais aussi antérieurement, en remontant dans le temps jusqu'à l'occurrence précédente. Dans un contexte d'estimation temps réel, une correction avant l'occurrence du véhicule traceur n'est pas possible, ou plus exactement pas utile puisqu'elle affecte les temps de parcours passés alors que ce sont les temps de parcours présents, voire futurs, qui nous intéressent. En outre, cette méthode est sensible au bruit sur les temps de parcours des véhicules traceurs. En effet, lorsque l'écart entre le temps de parcours de référence et le temps de parcours des véhicules sonde est significatif, la correction des courbes cumulées en amont sera biaisée par cet écart. CUPRITE ne peut pas être appliquée dans les réseaux avec un pourcentage inconnu de mouvements tournants sans un nombre minimal de véhicules traceurs pour corriger correctement les courbes cumulées en amont.

La majorité des recherches portent sur les routes inter-urbaines et ne peuvent pas être appliquées en ville, où l'estimation du temps de parcours est plus complexe, principalement dues aux flux significatifs qui proviennent des sources et fuites à mi-lien et qui ne sont pas prises en compte. En outre, la plupart de ces méthodes ne peuvent pas être appliquées pour une estimation en temps réel du temps de parcours, ni dans des réseaux avec un pourcentage inconnu des mouvements tournants. Cette thèse présente donc un algorithme qui répond à ces faiblesses en développant une approche fondée sur filtrage de Kalman sans parfum.

1.2 Objectif du travail de thèse

L'objectif principal de cette recherche est de développer une méthodologie d'estimation du temps parcours moyen urbain prenant en considération les problèmes déjà mentionnés en utilisant des données multi-sources. Les observations utilisées sont issues des deux types de capteurs mobiles et fixes. Les capteurs mobiles fournissent des temps de parcours individuels réalisés. Dans l'algorithme proposé, ces données sont utilisées à leur instant exact d'occurrence. L'une des données fournies par les capteurs fixes est le nombre de véhicules. Ces données sont agrégées chaque minute (voire 6 minutes).

Pour estimer le temps de parcours d'un lien, la méthode analytique classique des stocks utilise les courbes cumulatives du nombre de véhicules qui passent sur les boucles situées en amont et en aval du lien étudié. Cette méthode est mise en défaut dans les réseaux urbains principalement à cause des flux importants vers et depuis les fuites et sources à mi-lien. En

outre, la plupart des réseaux urbains ne sont équipés de détecteurs qu'à la fin de chaque lien. C'est pourquoi, sans aucune information sur le pourcentage des mouvements tournants aux carrefours, la méthode classique d'analyse n'est pas applicable. L'algorithme proposé dans cette recherche pour estimer le temps de parcours des liens urbains est basé sur un filtre de Kalman sans parfum (UKF). Cet algorithme intègre suivant une approche stochastique les données de comptage de véhicules à partir de détecteurs à boucles magnétiques à la fin de chaque lien et les temps de parcours réalisés par les véhicules traceurs. Dans la méthodologie proposée, l'évolution de ce filtre est basée sur la méthode analytique classique des stocks.

La performance de l'algorithme proposé est évaluée par des simulations en adoptant différents scénarios de trafic (avec/sans perturbations, avec des boucles uniquement à la sortie, avec un pourcentage connu/inconnu d'affectation dans les mouvements tournants), des pourcentages variables de perturbations et des pourcentages variables de véhicules traceurs dans le flot total (0%, 1%, 5% et 10%). Finalement, sa précision est comparée à celle de la méthode des stocks et à celle des méthodes basées sur cette dernière comme CUPRITE.

1.3 Plan du mémoire

Ce travail de thèse vise donc à élaborer une méthodologie d'estimation d'un temps de parcours dans un milieu urbain. Pour élaborer cet algorithme, une première étape nécessite de définir les éléments constituants un réseau routier et son trafic.

Ainsi, le chapitre 2 est un rappel sur le trafic routier avec les définitions des variables utilisées et les relations entre ces variables. De plus, dans ce chapitre, la théorie du trafic, le diagramme fondamental et la méthode analytique classique des stocks sont expliqués. Le comportement de cette dernière dans un réseau urbain contenant des sources/fuites à mi-lien est étudié en détails en utilisant des données simulées par le logiciel AIMSUM. Finalement, ce chapitre introduit le principe de fonctionnement des capteurs (respectivement fixes et mobiles) ainsi que la caractérisation de leurs erreurs.

Le chapitre 3 est consacré au développement des méthodes d'estimation du temps de parcours avec leurs avantages et leurs limitations :

- estimation par les boucles magnétiques qui généralement s'appuie sur le principe « first in first out » tout en négligeant les erreurs des boucles ;

- estimation par véhicules traceurs (instrumenté d'un récepteur GPS en utilisant une méthode de « *map-matching* ») qui impose un nombre minimal de ces derniers ;

- estimation par la fusion des données issues des véhicules traceurs et des boucles magnétiques. Dans cette partie, nous mettrons en évidence l'instabilité de CUPRITE par rapport au temps de parcours des véhicules traceurs, son comportement dans une estimation en temps réel et dans un réseau urbain sans aucune information sur le pourcentage des mouvements tournants dans un carrefour.

Le chapitre 4 introduit le filtrage de Kalman avec ses extensions : le filtrage de Kalman étendu et le filtrage de Kalman sans parfum.

Le chapitre 5 propose une formulation d'état (formulation markovienne) du problème étudié avec :

- un modèle d'évolution qui est basé sur la méthode analytique classique des stocks ;

- un modèle d'observation dont les composants sont le temps de parcours des véhicules traceurs et le nombre de véhicules comptés par les boucles magnétiques ;

- une méthode d'estimation de type filtrage de Kalman sans parfum.

Le chapitre 6 introduit l'application de l'algorithme proposé dans le chapitre précédent dans le cas où le pourcentage des mouvements tournants est connu. Alors, l'estimation du temps de parcours d'un réseau peut être découplée lien par lien. Le temps de parcours est estimé dans un premier temps pour un tronçon non perturbé (cas d'un tronçon inter-urbain) et dans un deuxième temps pour un tronçon perturbé (cas d'un tronçon urbain) dont le pourcentage de perturbation est variable. Dans ce chapitre, la précision de l'algorithme est comparée à celle de la méthode analytique classique des stocks et à celle de CUPRITE.

Le chapitre 7 présente une extension de l'algorithme proposé dans le cas où le pourcentage des mouvements tournants est inconnu. Alors, on présente l'estimation du temps de parcours dans le cas où un tronçon diverge à gauche et à droite. Puis, l'estimation du temps de parcours dans le cas où deux tronçons convergent en un tronçon. Les tronçons dans les deux cas ne sont pas découplés et sont considérés sans et avec perturbations.

Finalement le chapitre 8 introduit l'application de l'algorithme sur un divergent non perturbé, un convergent non perturbé et un réseau perturbé.

Les données utilisées dans les applications numériques des chapitres 6 et 8 sont produites par le logiciel de simulation microscopique de trafic AIMSUN (voir §.6.2). Dans chacun de ces chapitres, le contexte de simulation et l'analyse des données sont décrits en détails.

Chapitre 2. Rappel sur le réseau et le trafic routier

SOMMAIRE

2.1 Description d'un réseau routier et la spécificité d'un réseau routier urbain - 15 -

2.2 Variable du trafic - 16 -
2.2.1 Variables microscopiques - 16 -
2.2.2 Variables macroscopiques - 17 -

2.3 Théorie du trafic - 20 -
2.3.1 Équation de conservation - 21 -
2.3.2 Diagramme fondamental - 21 -
2.3.3 Méthode analytique classique des stocks - 23 -

2.4 Principaux capteurs de mesure de trafic - 30 -
2.4.1 Capteurs fixes : boucles magnétiques - 30 -
2.4.2 Capteurs mobiles : véhicules traceurs - 33 -
2.4.3 Synthèse et conclusion - 37 -

2.1 Description d'un réseau routier et spécificité d'un réseau routier urbain

Un réseau est un ensemble d'itinéraires caractérisé par une topologie donnée, des propriétés relatives à l'écoulement telles que la demande origine-destination ainsi que les contraintes relatives à la capacité, au choix d'itinéraires, etc.

Le choix d'un modèle pour étudier un tel système est lié à sa décomposition en un ensemble d'éléments et à la description de chacun de ces éléments et leurs interactions. D'un point de vue graphique, un réseau routier est considéré comme un graphe composé d'un ensemble de liens (ou tronçons) et d'un ensemble de nœuds.

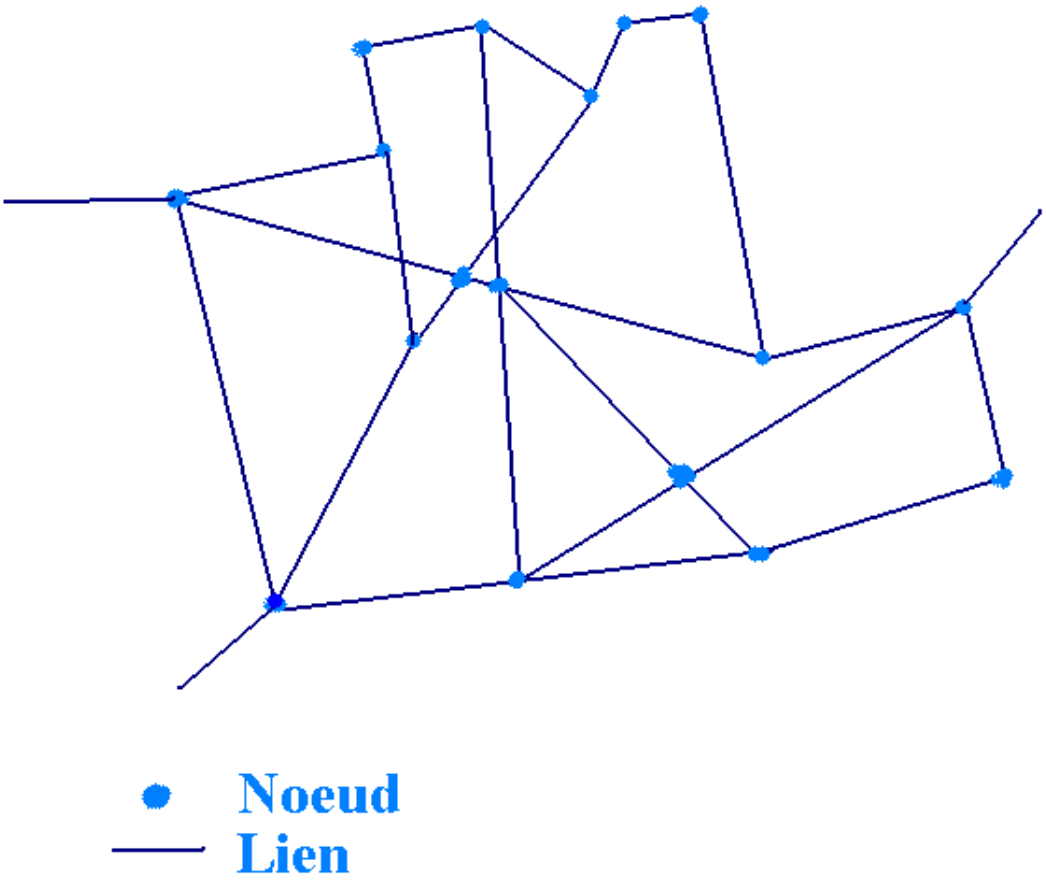


Figure 2. 1 : Représentation d'un réseau routier

Les liens désignent des routes et rues (urbaines ou inter-urbaines, avec une ou plusieurs voies). Les nœuds matérialisent les intersections et les rampes d'entrée / sortie. Les intersections se situent le plus souvent et en plus grand nombre à l'intérieur des zones urbaines et assurent le lien entre les différentes routes. Les rampes d'entrée / sortie assurent la transition entre les grands axes routiers.

Dans un réseau urbain l'écoulement est sujet à des interruptions surtout aux intersections dues à des contrôles de réseau tels que les feux de signalisations, les panneaux stop, etc. Les véhicules qui circulent « en accordéon » enregistrent des retards. Ceci conduit à une variation significative dans le temps de parcours entre deux véhicules consécutifs en fonction du temps auquel ils arrivent à une intersection. Par exemple, si le véhicule de tête arrive alors que le signal est au vert et le véhicule suivant alors qu'il est au rouge, alors le véhicule suivant doit s'arrêter à l'intersection ce qui entraîne un temps de parcours plus élevé. Les retards enregistrés au niveau des intersections sont alors considérés comme une partie importante du temps de parcours en milieu urbain.

En plus des retards aux intersections, les véhicules sont également sujets à des retards dans le lien, dus à un certain nombre de raisons telles que, les piétons, les véhicules qui entrent des rues transversales, etc.

2.2 Variable du trafic

L'élaboration de la théorie du trafic a permis de mieux comprendre et prévoir les interactions inter-véhiculaires ainsi que le mouvement global du flux au niveau des réseaux routiers. Chaque niveau (véhicule ou global) est caractérisé par des variables et des paramètres spécifiques. Pour cela, on doit prendre en compte un certain nombre de caractéristiques : lorsque celles-ci concernent le véhicule individuel ou le comportement de chaque conducteur, on parle de variables « microscopiques » alors que si elles concernent la description du comportement global du flux de véhicules, cela fait appel à des variables dites « macroscopiques ». Le lecteur pourra se référer à [Buisson, 1996] pour plus de détails sur la classification des modèles.

2.2.1 Variables microscopiques

Pour la description du trafic au niveau granulaire, l'accent est mis sur le véhicule individuel et son interaction avec son environnement. Les variables sont donc exprimées pour

chaque véhicule et sont étroitement liées à sa dynamique. La vitesse individuelle d'un véhicule est la principale variable microscopique.

2.2.2.1 Vitesse individuelle

Au niveau microscopique, la vitesse moyenne v_{im} d'un véhicule *i* est définie sur une durée *T* par :

$$v_{im} = \frac{1}{T} \int_{t_0}^{t_0+T} v_i(t)dt \tag{2.1}$$

où v_i exprime la vitesse instantanée du véhicule *i*.

La mesure de la vitesse instantanée d'un véhicule individuel *i* nécessite des observations sur un intervalle d'espace $[x_1, x_2]$ et un intervalle de temps $[t_1, t_2]$. Cette vitesse est définie par :

$$v_i = \frac{dx}{dt} = \lim_{(t_2 - t_1) \to 0} \frac{x_2 - x_1}{t_2 - t_1} \tag{2.2}$$

Généralement, l'enregistrement dans le temps de la vitesse instantanée d'un véhicule le long d'un parcours sert à calculer les vitesses moyennes au cours de différents cycles de fonctionnement. L'intérêt de ces vitesses moyennes est de fournir des indications sur la qualité du trafic. Elles permettent, par exemple, de révéler la proportion de temps d'arrêt d'un véhicule sur une route donnée.

2.2.2 Variables macroscopiques

Au niveau macroscopique, les variables mises en jeu permettent de décrire l'état du flux du trafic comme le débit, la densité, le taux d'occupation et la vitesse moyenne du flux.

2.2.2.1 Débit du trafic

Le débit du trafic, appelé aussi intensité du trafic [Hall, 1997], représente la quantité moyenne du flux $\overline{q}$ mesurée en nombre de véhicules N qui passent au point x de la section de route considérée, entre deux instants de temps t_1 et t_2. Il permet de décrire la distribution des véhicules dans le temps.

$$\bar{q}\left(x,t_1,t_2\right)=\frac{N}{T};T=t_1-t_2 \tag{2.3}$$

En pratique, cette variable peut être calculée par un simple comptage des véhicules sur la section de la route étudiée.

2.2.2.2 Densité du trafic

La densité moyenne (appelée aussi concentration) ne peut être mesurée que sur une certaine longueur. Ainsi, la densité moyenne $\bar{k}\left(t,x_1,x_2\right)$ est définie par le nombre de véhicules N qui se trouvent à l'instant *t* dans la section délimitée par les points x_1 et x_2. Elle décrit la distribution des véhicules dans l'espace [El Faouzi et Lessort, 2001].

$$\bar{k}\left(t,x_1,x_2\right)=\frac{N\left(t,x_1,x_2\right)}{x_2-x_1} \tag{2.4}$$

D'une manière générale, la densité du trafic s'exprime en nombre de véhicules par kilomètre.

2.2.2.3 Taux d'occupation

La concentration n'est pas mesurée directement. Pour l'estimer, on utilise souvent une grandeur plus facilement mesurable, qui est le taux d'occupation, définie par la fraction du temps durant laquelle le capteur de présence (par exemple une boucle magnétique) détecte un véhicule [El Faouzi et Lessort, 2001].

Soit une boucle magnétique de longueur l_0, sur laquelle on observe les passages de N véhicules pendant un temps Δt. Chaque véhicule *i* possède une longueur l_i et une vitesse v_i. La boucle est ainsi excitée par le passage du véhicule *i* pendant un temps t_i :

$$t_i=\frac{l_i+l_0}{v_i} \tag{2.5}$$

On définit le taux d'occupation *o* pendant la période Δt par :

$$o=\frac{1}{\Delta t}\sum_{i=1}^{N}t_i=\frac{1}{\Delta t}\sum_{i=1}^{N}\frac{l_i+l_0}{v_i} \tag{2.6}$$

En supposant que la longueur des véhicules est une constante L (cette hypothèse est à peu près valide dans un environnement urbain où les poids lourds sont généralement peu nombreux), le taux d'occupation est égal à :

$$o = \frac{L + l_0}{\Delta t} \sum_{i=1}^{N} \frac{1}{v_i} = (L + l_o) q \frac{1}{q \Delta t} \sum_{i=1}^{N} \frac{1}{v_i} = (L + l_0) \frac{q}{V} = (L + l_0) k \tag{2.7}$$

avec V la vitesse du flot (voir équation 2.12).

2.2.2.4 Relation entre les variables : vitesse du flot

Il existe deux définitions possibles (spatiale et temporelle), de la vitesse moyenne à un point x à l'instant t [El Faouzi et Lessort, 2001].

Cas isovélocique

Dans le cas où l'écoulement est uniforme (tous les véhicules ont une vitesse identique v) la vitesse moyenne V du flot est de façon triviale égale à v. Supposons qu'une portion de chaussée Δx contienne à tout instant un nombre de véhicules $N = k \times \Delta x$ où k est la concentration. La vitesse du flot (égale à celle de tous les véhicules) étant égale à V, il faut un temps $\Delta t = \Delta x / V$ pour que ces N véhicules quittent la zone considérée. Le débit sortant de cette zone est donc égal à :

$$q = \frac{N}{\Delta t} = k \Delta x . \frac{V}{\Delta x} = kV \tag{2.8}$$

Cas général

On définit la vitesse de flot en un point x et à un temps t par le rapport entre le débit instantané et la densité spatiale instantanée [Wardrop, 1952] et elle s'exprime de la façon suivante :

$$V(x;t) = \frac{q(x;t)}{k(x;t)} \tag{2.9}$$

Généralement, ***la vitesse moyenne temporelle (V_T) n'est pas égale à la vitesse du flot*** [El Faouzi et Lessort, 2001]. La relation qui existe entre les deux formes de vitesse moyenne est connue sous le nom de formule de [Wardrop, 1952] et s'exprime :

$$\overline{V}_T = V + \frac{\sigma_E^2}{V} \tag{2.10}$$

où σ_E^2 désigne la variance spatiale de la vitesse.

Ceci implique que la vitesse moyenne temporelle est toujours supérieure à la vitesse du flot (dans le cas isovélocique, où σ_E^2 =0, les deux moyennes sont confondues).

La vitesse du flot est égale à la moyenne harmonique des vitesses mesurées en un point. Si *N* véhicules passent au point *x* pendant Δt [El Faouzi et Lessort, 2001] :

$$\frac{1}{V} = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{v_i} \tag{2.11}$$

Cette propriété est importante puisqu'elle donne un moyen d'estimer la vitesse du flot à partir d'un instrument de mesure ponctuel (radar ou boucle).

2.3 Théorie du trafic

En théorie du trafic [Buisson, 1996], deux types principaux de modèles ont été développés : les modèles macroscopiques qui décrivent l'évolution du trafic comme un écoulement dans un milieu continu en établissant une relation entre le débit, la densité et la vitesse du flux [Lighthill et Whitham, 1955], [Richards, 1956], [Jiang, 2002], [Payne 1979], [Rascle, 2002], [Zhang, 2002] et les modèles microscopiques qui se fondent sur une description du comportement individuel des véhicules et l'utilisation de modèles représentant les interactions inter-véhiculaires et les interactions entre le véhicule et son environnement (modèles de poursuite) [Ben Akiva, 1994], [Yang et al., 1999]. Dans le cadre de cette thèse, nous nous intéressons à l'aspect macroscopique de trafic (à l'estimation moyenne du temps de parcours et pas à l'estimation individuelle du temps de parcours). Les premiers travaux en la matière ont été ceux de [Lighthill et Whitham, 1955] et [Richards, 1956]. Ils ont développé la théorie macroscopique de l'écoulement hydrodynamique du trafic à partir des trois équations qui relient les trois variables macroscopiques du trafic, le débit, la densité et la vitesse du flux :

$$\begin{cases} q(x,t) = k(x,t) \times V(x,t) \\ \partial_x q(x,t) + \partial_t k(x,t) = 0 \\ V_{eq}(x,t) = V\left(k(x,t)\right) \end{cases} \tag{2.12}$$

La première équation est l'équation de la vitesse du flot, la deuxième équation est l'équation de conservation et la troisième équation est l'équation d'équilibre tirée du

« diagramme fondamental » présenté ci-dessous. Nous présentons ci-dessous les deux dernières équations (l'équation du flot est déjà expliquée dans §.2.2.2.4). La résolution de ces équations peut être réalisée à partir d'une approche analytique ou d'une approche numérique. Pour une description de la résolution des ces équations, le lecteur pourra se reporter aux références [Godlewski et Raviart, 1990] et [Leclercq, 2002].

2.3.1 Équation de conservation

Le principe de conservation des véhicules intervient entre deux points x_A et x_B. L'évolution du nombre de véhicules est égale à la différence entre le flux entrant en x_A et le flux sortant en x_B. Si x_A et x_B sont des carrefours, on somme les flux entrant d'une part et sortant d'autre part, avant d'en faire la différence.

2.3.2 Diagramme fondamental

L'équation de la vitesse à l'état d'équilibre proposée par [Greenshield, 1935] à partir de données expérimentales est la suivante :

$$V_{eq}(x,t) = V_f\left(1 - \frac{k(x,t)}{k_{\max}}\right) \tag{2.13}$$

où

- V_{eq} : est la vitesse à l'état d'équilibre ;
- V_f: est la vitesse libre du flux ;
- k_{max} : la densité maximale.

Il existe trois relations liant deux des trois variables débit, densité et vitesse : soit entre le débit et la densité, soit entre la vitesse et la densité (équation 2.13), soit entre la vitesse et le débit, sachant que la première reste toujours égale au produit des deux autres. Ces trois relations portent le nom de « diagramme fondamental ». La relation entre le débit et la densité est la plus employée dans le domaine du trafic. Le premier diagramme fondamental est issu de la relation de Greenshield et conduit à une forme parabolique de la courbe représentant le débit en fonction de la densité (voir figure 2.2). A titre d'exemple, le diagramme fondamental

sous ses 3 formes pour une vitesse libre de circulation de 36 km/h et une densité maximale de 140 véh/km est illustré figure 2.2.

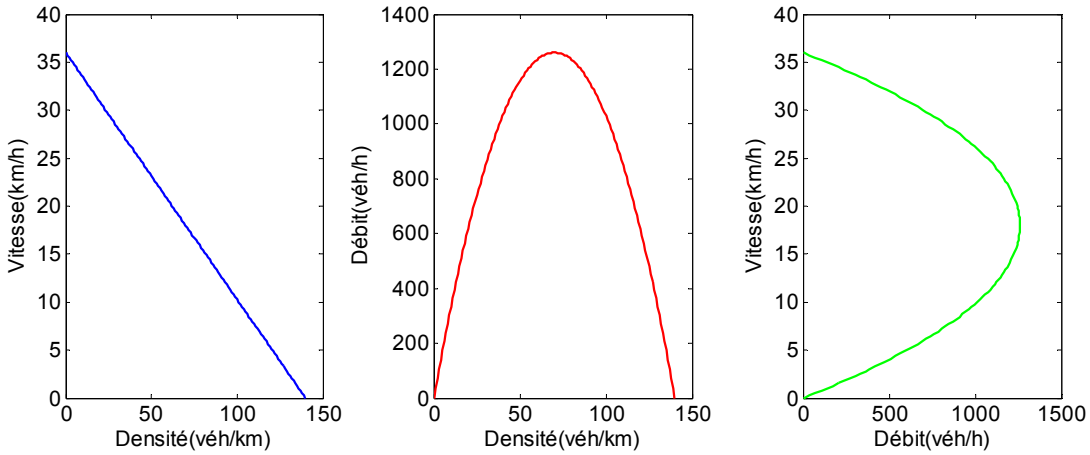


Figure 2. 2 : Diagramme fondamental de Greenshield

D'autres auteurs ([Buisson, 2010]) ont proposé d'autres formes pour ce diagramme en modulant celui de [Greenshield, 1935] et en utilisant des mesures de trafic (voir figure 2.3).

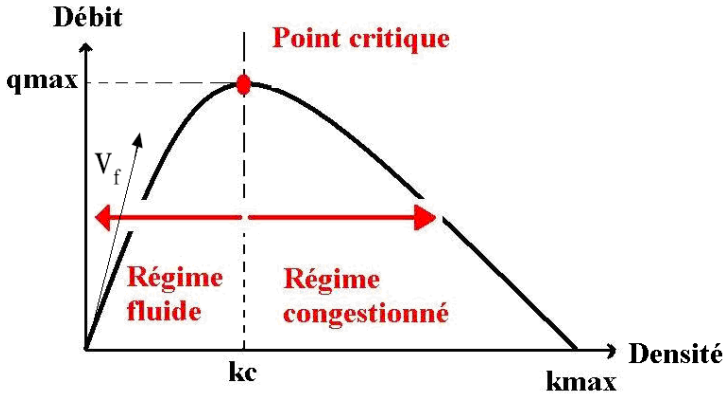


Figure 2. 3 : Diagramme fondamental

Quand la concentration est faible, les véhicules roulent à leur vitesse maximale qui est égale à la vitesse libre du flux, notée V_f, et on remarque que le débit passant en un point a une croissance presque linéaire. Plus la concentration des véhicules dans une zone de longueur constante située en amont d'un point donné augmente, plus le débit augmente. Cette augmentation se poursuivra jusqu'à atteindre le débit maximal désignant la capacité de la section étudiée. Ce point est appelé point critique (k_c). Si la densité continue de croître, on atteint la saturation de la section de route étudiée. Cette saturation se traduit par une

diminution du débit et, par conséquent, de la vitesse. À ce moment là, il ne s'agit plus d'un régime de trafic fluide. Au-delà de ce point critique, quand la concentration augmente, le débit passant en un point diminue. La densité maximale est obtenue dans le cas où tous les véhicules sont immobiles, ce qui désigne le nombre maximal de véhicules en arrêt complet qu'une route peut contenir. Souvent le diagramme fondamental est approximé par un triangle, comme illustré par la figure 2.4.

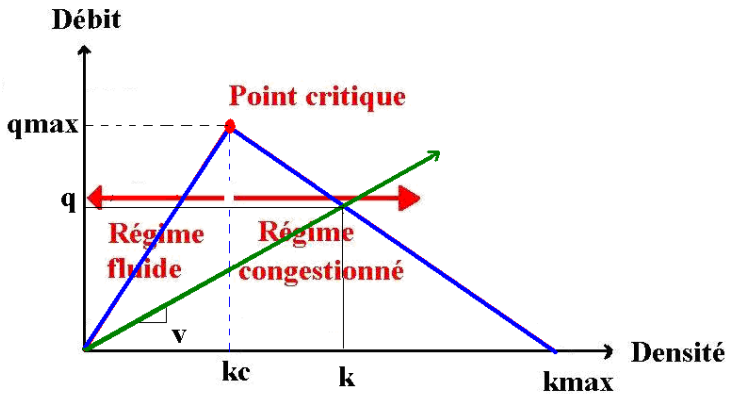


Figure 2. 4 : Diagramme fondamental triangulaire

2.3.3 Méthode analytique classique des stocks

2.3.3.1 Principe de fonctionnement

Une autre approche d'analyse qui permet d'étudier les phénomènes de trafic est la méthode analytique classique de stocks. Cette méthode permet de calculer le temps de parcours et de le prédire [Bhaskar, 2009b], [Buisson et Lessort, 2010].

La courbe cumulative représente le graphique du nombre cumulé de véhicules qui passent devant un observateur (ou un détecteur) en un lieu donné au cours du temps *t*, à partir d'un nombre de véhicules initial arbitraire.

La courbe obtenue est discrète parce qu'en pratique les détecteurs sont lus généralement toutes les 6 minutes (on n'a pas les instants de passage de tous les véhicules). Dans la figure 2.5, l'évolution du nombre de véhicules est représentée par une courbe continue alors qu'en réalité cette évolution est constante par morceaux (forme d'escalier) où chaque « marche » est une lecture du compteur (par exemple au pas de 6 minutes). Généralement, il y a plusieurs passages/unités par intervalle de 6 minutes.

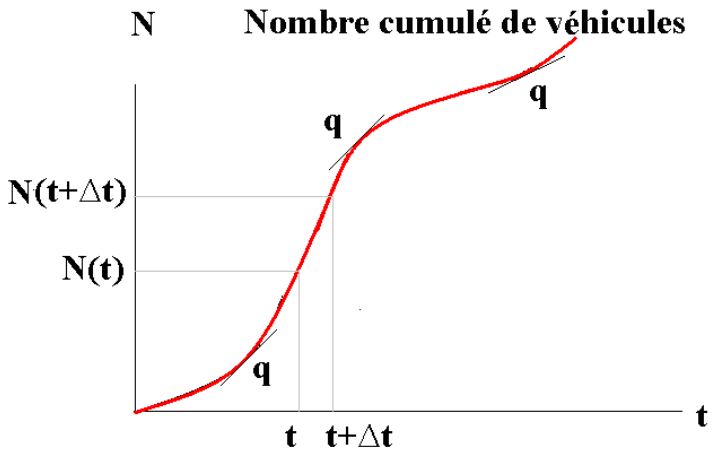


Figure 2. 5 **: Nombre cumulé des véhicules**

La fonction cumulative est supposée différentiable par rapport au temps. La pente de la courbe à l'instant *t* est le débit de la circulation instantanée à l'instant *t*. La valeur de la fonction cumulative au temps *t* est *N(t)*. Le débit moyen entre *t* et *t*+Δ*t* est égal à :

$$\left(N\left(t+\Delta t\right)-N\left(t\right)\right)/\Delta t \qquad (2.14)$$

La figure 2.6 montre deux courbes cumulatives $N_u(t)$ et $N_d(t)$ obtenues respectivement en amont (*Upstream*) et en aval (*Downstream*) d'un tronçon donné, en supposant que :

- la discipline *First-In-First-Out* (FIFO) est respectée pour tous les véhicules traversant de l'amont vers l'aval, c'est-à-dire qu'il n'y a pas de dépassement de véhicules ;
- les véhicules sont conservés, c'est-à-dire qu'il n'y a pas de perte ou de gain de véhicules dans le tronçon.

La distance verticale (le long de l'axe Y) entre les deux courbes $N_u(t)$ et $N_d(t)$ à l'instant *t* définit l'écart de comptage instantané (*n*) entre les deux endroits. La distance horizontale (le long de l'axe X) pour le nombre de comptage *i* définit le temps de parcours (TT_i) pour le $i^{ème}$ véhicule (en supposant que les passages des véhicules sont tous comptés individuellement).

Le temps de parcours total entre l'instant t_1 et l'instant t_3 est la somme du temps de parcours des *N* véhicules sortant du tronçon entre l'instant t_2 et l'instant t_4 c'est-à-dire l'aire (*A*) de la surface entre les deux courbes cumulées. Le temps de parcours moyen par véhicule est le rapport : *A* /*N*.

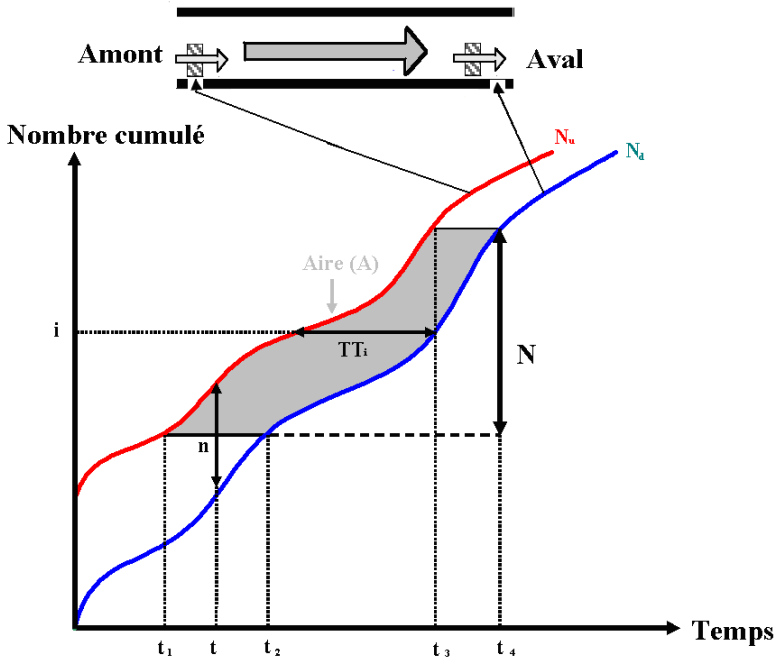


Figure 2. 6 : Estimation du temps de parcours par la méthode analytique classique des stocks

La méthode des stocks est simple et efficace car elle permet d'estimer le temps de parcours à partir de données standard de comptage en quelques points. Cependant, il est nécessaire que les pertes ou les gains de véhicules sur l'itinéraire concerné soient négligeables, que l'hypothèse de non dépassement soit respectée et que la mesure des nombres de véhicules soit fiable.

2.3.3.2 Limitation dans les réseaux urbains

Pour l'application de la méthode classique mentionnée ci-dessus, les courbes cumulées doivent être évaluées avec précision : il ne doit pas y avoir de divergence entre les courbes. La situation est idéale lorsque les détecteurs sont fiables et parfaits (c'est-à-dire qu'ils fournissent des informations précises véhicule par véhicule) et les véhicules sont conservés entre les détecteurs situés en amont et en aval. Toutefois, ces conditions sont difficiles à obtenir dans la pratique, en particulier dans les réseaux urbains pour les raisons suivantes :

- l'hypothèse de conservation des véhicules entre l'amont et l'aval n'est pas respectée du fait des sources et fuites telles que : les parkings, les liens transversaux, les zones résidentielles et commerciales, etc. ;

- il peut y avoir des dysfonctionnements des boucles magnétiques qui peuvent sur-compter ou sous-compter. De plus, comme les données sont cumulées, les erreurs se cumulent également ce qui peut entraîner des écarts importants entre les courbes cumulées ;

- la majorité des réseaux urbains sont équipés de boucles seulement à la sortie des liens, ce qui cause un problème dans l'estimation des courbes cumulées pour les mouvements tournants : un réseau urbain peut présenter des combinaisons complexes de flux vers et à partir d'un lien. C'est pour cela que les affectations des mouvements tournants doivent être connues *a priori*.

2.3.3.3 Application

L'objectif ici est de mettre en évidence les limitations de la méthode des stocks. Le premier paragraphe introduit le fonctionnement de cette méthode sous la loi de conservation de véhicules. Le deuxième paragraphe met en évidence que cette méthode dysfonctionne pour estimer le temps de parcours d'un tronçon urbain. Finalement le troisième paragraphe montre l'effet des valeurs aberrantes sur l'estimation du temps de parcours par cette procédure.

Estimation du temps de parcours sous la loi de conservation des véhicules

Le temps de parcours est estimé par la méthode de stocks pour 10 simulations réalisés avec le logiciel AIMSUM (voir §.6.2). Les simulations sont réalisées pour un tronçon sans source ni fuite à mi-lien, c'est-à-dire que les véhicules sont conservés. De plus, ce tronçon, illustré dans la figure 2.7, est équipé d'une boucle à son entrée et d'une boucle à sa sortie. Ces boucles sont supposées fournir des données exactes.

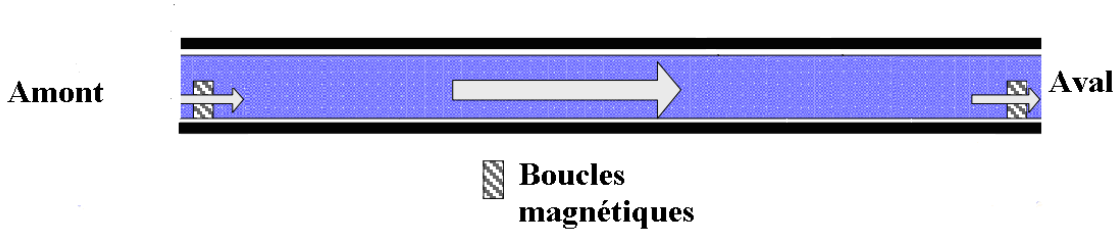


Figure 2. 7 : Tronçon étudié

La variation du débit en amont injecté en tant que commande dans le logiciel de simulation est illustrée par la figure 2.8.

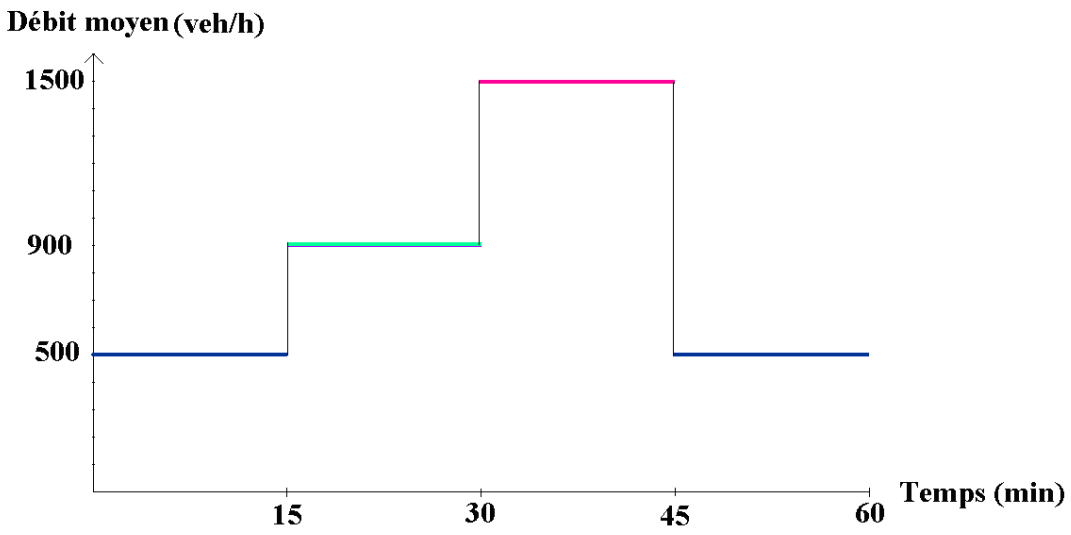


Figure 2. 8 : Variation du débit

Le temps de parcours moyen de référence à l'instant *t* est la moyenne des temps des parcours de tous les véhicules qui sont sortis du tronçon entre l'instant *t-1* et *t*. Le calcul du temps de parcours de référence est réalisé à chaque minute. Le temps de référence pour l'une des simulations est illustré dans la figure 2.9.

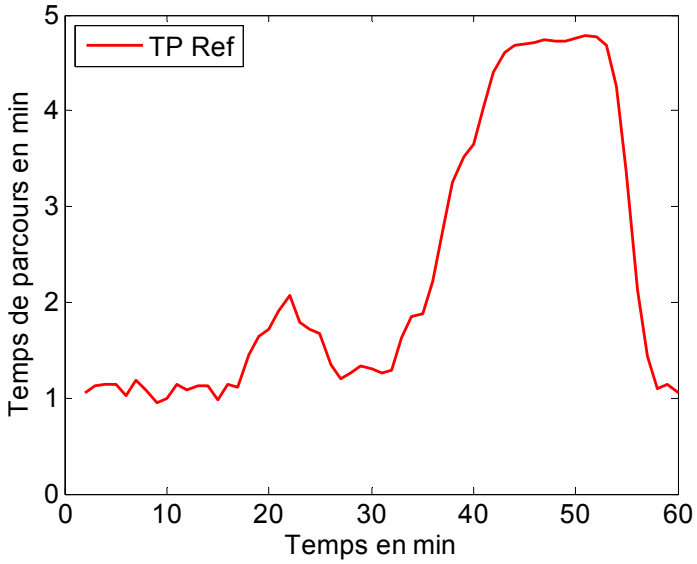


Figure 2. 9 : Temps de parcours de référence

L'erreur commise par la méthode analytique classique des stocks exprimée en minute est la différence entre le temps de parcours de référence et le temps de parcours estimé par cette procédure. La figure 2.10 illustre cette différence en minute.

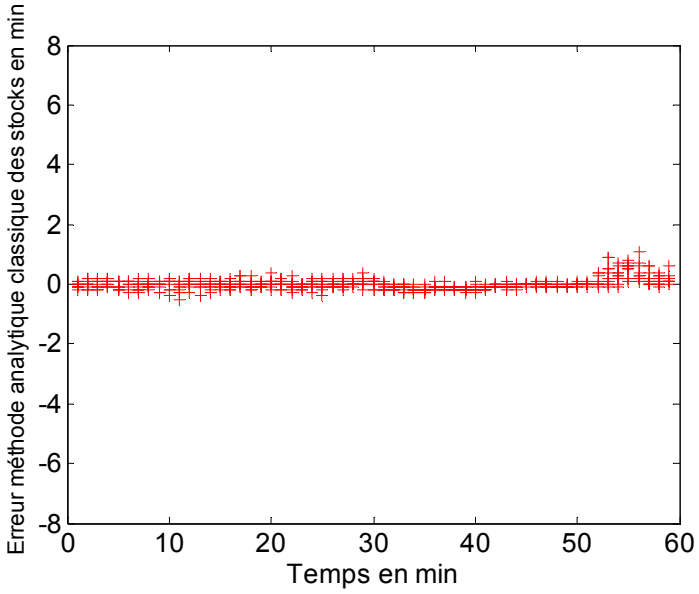


Figure 2. 10 : Erreur commise par la méthode analytique classique de stock

L'erreur maximale commise est d'une minute et la majorité des erreurs appartiennent à l'intervalle [-0,5 min ; + 0,9 min].

Estimation du temps de parcours d'un tronçon urbain

Dans un tronçon urbain, l'hypothèse de conservation des véhicules entre l'amont et l'aval n'est pas vérifiée. La perturbation est représentée simplement par une source/fuite à mi-lien (voir figure 2.11).

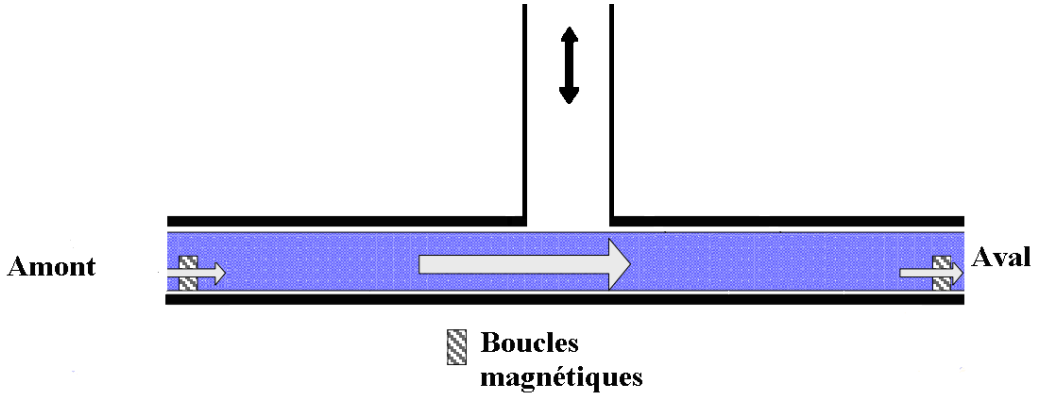


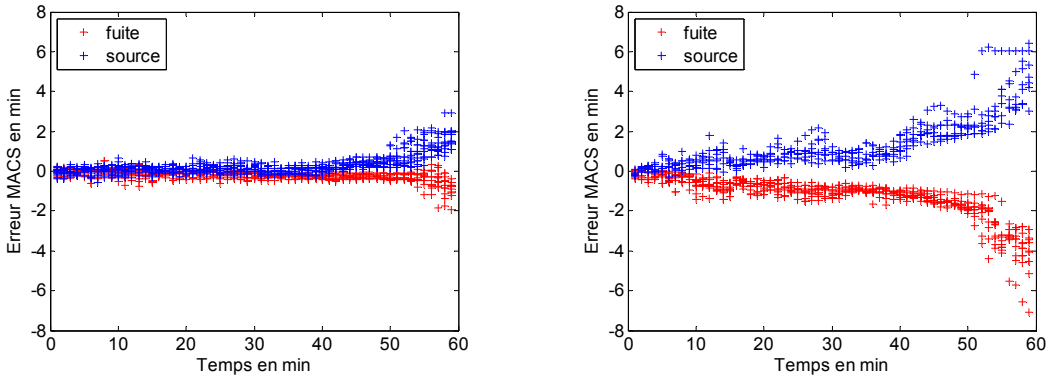
Figure 2. 11 : Tronçon étudié avec source/fuite à mi-lien

Quatre scénarios sont testés :

- 1% des véhicules entrent dans le tronçon à mi-lien ;

- 1% des véhicules sortent du tronçon à mi-lien ;
- 5% des véhicules entrent dans le tronçon à mi-lien ;
- 5% des véhicules sortent du tronçon à mi-lien.

Les figures 2.12 (a) et (b) illustrent respectivement l'erreur commise par cette méthode (MACS) pour 1% et 5% de source/fuite.



(a) (b)

Figure 2. 12 : Erreur commise par la méthode analytique classique de stock avec (a) 1% (b) 5% de fuite/source

L'erreur maximale atteinte avec 1% de perturbation est de 2 minutes et avec 5% de 8 minutes. Nous remarquons l'effet cumulé de la perturbation dans les deux cas 1% et 5% de source/fuite à mi-lien, en particulier, sur les 10 dernières minutes. Ceci met en évidence le dysfonctionnement de cette méthode dans le cas d'un tronçon urbain présentant des fuites ou des sources.

Effet des valeurs aberrantes

Dans le cas d'un tronçon satisfaisant la loi de conservation de véhicules, les valeurs cumulées des nombres de véhicules en aval pour les instants 29 et 30 sont volontairement bloquées à la même valeur. Le graphe 2.13 donne le temps de parcours de référence (TP Ref) ainsi que le temps de parcours calculé par la méthode analytique classique des stocks (TP MACS).

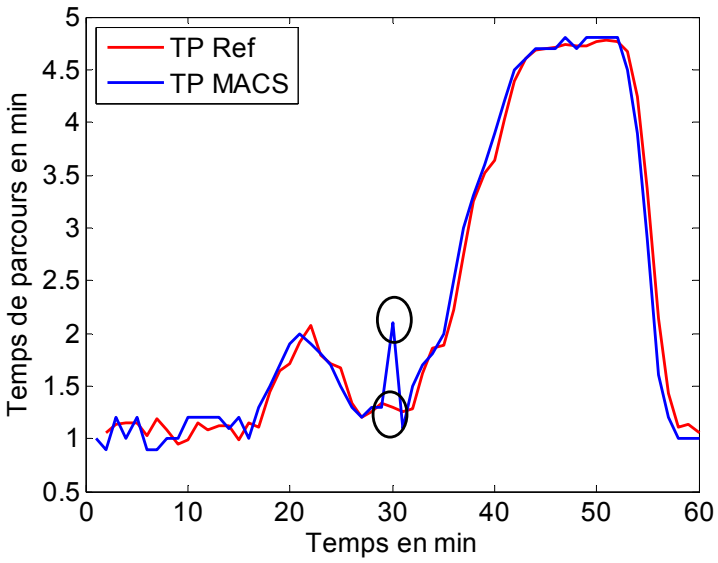


Figure 2. 13 : Temps de parcours estimé par la méthode analytique classique (MACS) des stocks avec 2 valeurs aberrantes

L'erreur d'estimation commise par MACS à l'instant 30 est significative. Ceci met en évidence l'effet des valeurs aberrantes fournies par les boucles magnétiques sur l'estimation du temps de parcours par la méthode analytique classique des stocks.

2.4 Principaux capteurs de mesure de trafic

Dans le domaine du trafic, plusieurs instruments ont été mis en œuvre pour la mesure des variables. Nous pouvons les classer en capteurs fixes et capteurs mobiles [Keen et al., 1990], [Kell et al., 1990], [Johnson, 1999]. Ce paragraphe présente ces différents types de capteurs, met en évidence leur complémentarité, ce qui nous amènera finalement à proposer leur fusion.

2.4.1 Capteurs fixes : boucles magnétiques

Plusieurs capteurs sont considérés comme fixes parmi lesquels les boucles magnétiques, les capteurs infrarouges, les capteurs à ultrasons, les capteurs vidéo, etc. Dans le cadre de cette thèse seules les boucles magnétiques sont présentées car elles sont les plus anciennes, les plus largement utilisées comme sources de données de trafic sur les réseaux routiers, notamment européens.

Les boucles magnétiques ont été introduites dans le domaine du trafic routier depuis les années 60 et sont, à ce jour, les capteurs les plus utilisés principalement dans la commande du

trafic en temps réel [Middleton et al., 2002], [Papageorgiou et al., 1990], [Ellouni, 1996]. Elles rentrent dans la famille des détecteurs de véhicules.

2.4.1.1 Principe de fonctionnement

Les boucles inductives sont implémentées dans la chaussée. Quand un véhicule passe au-dessus de la boucle, le champ électromagnétique subit une variation et en conséquence un courant électrique est généré dont on mesure la tension induite. Il existe deux types de boucles magnétiques : les simples et les doubles.



Figure 2. 14 : Boucle magnétique intégrée dans le sol

Boucles magnétiques simples

Ce type de capteur est constitué d'une boucle inductive traversée par un courant alternatif de haute fréquence. Il est le plus souvent situé juste sous le revêtement de surface d'une route.

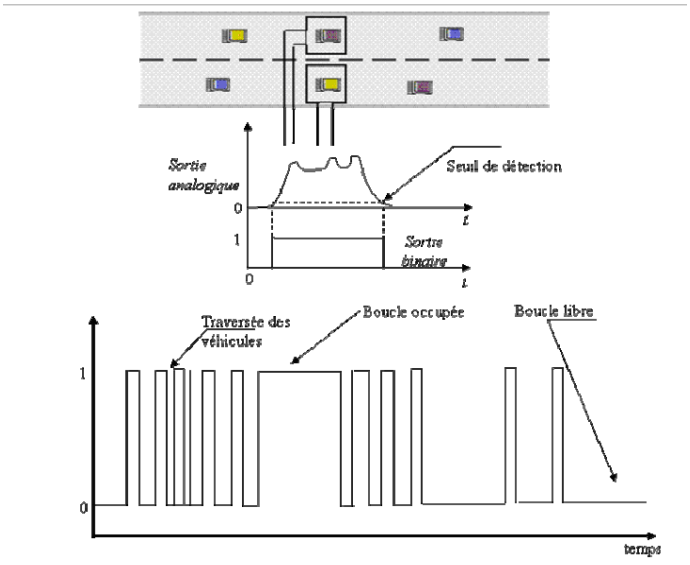


Figure 2. 15 : Schéma d'une boucle magnétique simple

La boucle est reliée à un circuit électronique produisant une information binaire au passage d'un véhicule. Lorsqu'un véhicule est proche de la boucle magnétique, il provoque une diminution de l'inductance qui passe en dessous d'un seuil. Le capteur génère un signal binaire qui passe alors d'un niveau 0 à 1 pour signaler le passage du véhicule. De même lors de l'éloignement du véhicule, l'inductance augmente et le capteur repasse à un niveau 0, ce qui correspond à une absence de véhicule. Avec ce type de capteur, le débit et le taux d'occupation sont mesurés sur la portion de route où les boucles magnétiques se trouvent.

Boucles magnétiques doubles

Dans un système à boucles magnétiques doubles, deux boucles magnétiques simples consécutives, appelées « boucle M » et « boucle S », sont installées à quelques mètres d'intervalle. Avec une telle conception, une minuterie se déclenche au passage d'un véhicule sur la première boucle et se termine avec le passage sur la deuxième. Ainsi, en plus du débit des véhicules et du taux d'occupation, les boucles magnétiques doubles donnent la vitesse des véhicules comme suit [Woods et al., 1994] :

$$V = \frac{D}{t_{on}^{1} - t_{on}^{2}} \tag{2.15}$$

où :

- V est la vitesse du véhicule ;
- D est la distance entre les deux boucles ;
- t_{on}^{1} et t_{on}^{2} sont respectivement les instants où la première puis la deuxième boucle détectent le véhicule.

Une autre formule est utilisée pour le calcul de la vitesse [Wilshire et al., 1985] :

$$V = \frac{1}{2}\left(\frac{D}{t_{on}^{1} - t_{on}^{2}} + \frac{D}{t_{off}^{1} - t_{off}^{2}} \right) \tag{2.16}$$

où :

- t_{off}^{1} et t_{off}^{2} sont respectivement les instants où la première puis la deuxième boucle s'éteignent.

2.4.1.2 Caractéristiques des erreurs

Les boucles magnétiques sont généralement utilisées pour fournir des données pour la surveillance des routes, le contrôle de la circulation, etc. Avec les boucles magnétiques simples, on peut mesurer le débit et le taux d'occupation. En plus de ces mesures, les boucles magnétiques doubles mesurent la vitesse. Ces mesures sont agrégées par intervalle d'une minute (voire 6 minutes). Les boucles magnétiques fournissent des informations sur la portion du tronçon où elles se trouvent, c'est-à-dire des informations ponctuelles qui ne correspondent pas forcement aux conditions de trafic sur le reste du tronçon étudié. Un des avantages des boucles est leur forte dissémination sur les réseaux routiers européens. En outre, elles fournissent une base de données sur une longue période de temps.

Les taux de panne ou de défaillance des boucles magnétiques sont relativement importants. On estime que, sur le parc total de boucles, environ 10 à 20% sont hors d'usage ou en panne. Par exemple, si le réglage de fréquence des détecteurs n'est pas fait correctement, on va avoir des « accrochages », c'est-à-dire des générations spontanées de véhicules : un exemple de ce type de problème peut être brusquement vers 2h ou 3h du matin, alors que le trafic est faible normalement, le débit en 6 minutes est de 1000 véhicules. Il peut y avoir également un détecteur qui « fait des rebonds », il double tous les véhicules qui passent. Ainsi, il y a de nombreuses possibilités d'erreur, au niveau des capteurs mais aussi au niveau de la station de comptage. Si les capteurs sont bons, le taux d'erreur sur le débit est généralement de l'ordre de 3 à 5% [Borsu, 2004].

2.4.2 Capteurs mobiles : véhicules traceurs

Les capteurs de trafic mobiles, appelés les véhicules traceurs, sont les véhicules équipés d'un système de géolocalisation tel que le système de positionnement global GPS ou le téléphone mobile. En communiquant régulièrement leur position, leur vitesse, etc. ils sont capables de fournir la trajectoire spatio-temporelle du véhicule équipé. Les véhicules traceurs (VT) aussi appelés véhicules sondes (VS) peuvent mesurer directement la vitesse et le temps de parcours entre deux endroits et donc sont considérés comme l'une des sources de données les plus prometteuses pour l'estimation du temps de parcours.

2.4.2.1 Principe de fonctionnement

Un véhicule traceur transmet (généralement par GSM) ses coordonnées géographiques (longitudes et latitudes), sa vitesse et éventuellement le cap suivi par le véhicule à différents intervalles de temps (défini par la fréquence de transmission des données). On appelle *map-matching*[1] le processus qui permet de passer d'une position globale (en coordonnées géographiques WGS 84 ou en coordonnées « Lambert », qui est la projection légale nationale) à une position sur la carte. Par ailleurs, les techniques de *map-matching* permettent de compenser certaines imprécisions du positionnement global, mais cela est fortement dépendant de l'exactitude de la carte numérique, de la précision et de la fréquence des données des véhicules traceurs.

Généralement, il existe deux manières d'estimer le temps de parcours à partir des données des véhicules traceurs :

1. le temps de parcours est obtenu directement : soit un segment de route de longueur L entre l'amont A et l'aval B, on suppose qu'on peut repérer le passage d'un véhicule i en A et en B. Soit t_{iA} l'instant auquel le véhicule i est en A et soit t_{iB} l'instant auquel le véhicule i est en B, alors le temps de parcours réalisé par ce véhicule à l'instant t_{iA} est t_{iB}-t_{iA}. C'est un temps de parcours réalisé. L'estimation du temps de parcours moyen du tronçon pour tous les véhicules (la population) traversant ce tronçon pendant un certain intervalle est obtenu en sommant le temps de parcours individuel de chaque véhicule traceur et en divisant la somme par le nombre total ;
2. si l'instant de passage du véhicule en amont et en aval n'est pas détecté, le temps de parcours peut être obtenu à partir de la vitesse du véhicule attribué au tronçon. La vitesse moyenne du flux de chaque tronçon peut être calculée en utilisant la vitesse moyenne harmonique[2] des véhicules traceurs. Le temps de parcours du tronçon est alors le rapport entre sa longueur et la vitesse moyenne du flux.

2.4.2.2 Caractérisation des erreurs

Contrairement aux capteurs fixes, les véhicules traceurs décrivent l'état du trafic sur l'ensemble du réseau circulé et représentent une couverture spatiale illimitée. Les mesures des véhicules traceurs sont non exhaustives car elles n'intègrent qu'une partie des véhicules qui

[1] Voir l'annexe : map-matching

ont parcouru le tronçon. En fait, ces véhicules représentent un échantillon aléatoire de la population des véhicules dans le réseau. Par conséquent, la précision de l'estimation du temps de parcours par les véhicules traceurs est liée au nombre de ces derniers.

Les données des véhicules traceurs aujourd'hui sont de plus en plus nombreuses, et tendent à fournir suffisamment d'informations pour déduire directement l'état macroscopique du trafic à une grande échelle spatio-temporelle.

Dans le cas du GPS, la disponibilité des signaux satellitaires est un facteur crucial pour obtenir en temps réel des données provenant des véhicules traceurs. En fait, le système GPS nécessite au moins quatre satellites pour estimer la position du véhicule. En raison de l'indisponibilité de signaux (problèmes de propagation), il y a des cas où les positions ne peuvent pas être calculées par les récepteurs GPS. Ces problèmes de propagation peuvent se produire lorsque le véhicule traceur est dans le voisinage de structures élevées telles que les bâtiments, ou dans les tunnels. Les erreurs du GPS, et consécutivement les erreurs de *map-matching*, seront dépendantes de l'emplacement du tronçon : dans un centre ville (dense), l'ordre de grandeur de ces erreurs est de quelques dizaines de mètres, tandis que dans une zone ouverte, il est seulement de quelques mètres. Dans une toute première approximation, nous pouvons classer les liens de cette façon : centre ville d'une part, et espace ouvert d'autre part, où les erreurs des observations du temps de parcours sont au maximum de 10 et de quelques secondes respectivement. Nous avons pu confirmer cette approximation pour le centre ville par une étude où nous avons déterminé l'erreur commise sur le temps de parcours due à l'imprécision du GPS et du *map-matching*. Le lecteur pourra se reporter pour cette étude à §.5.5.4.

Les données obtenues à partir du téléphone mobile proviennent de tous les utilisateurs, qu'ils soient dans les véhicules et à l'extérieur. Les différencier est compliqué. Par exemple, si il y a une ligne de métro parallèle à la route, les données du téléphone mobile d'un passager dans le métro ou même à pied sur le trottoir peuvent être interprétées comme celles d'un utilisateur de la route, conduisant à l'estimation erronée du temps de parcours.

Les véhicules traceurs fournissent uniquement des informations pour les tronçons et les itinéraires parcourus par ces véhicules ainsi que dans l'intervalle de temps dans lequel ils sont disponibles. Par conséquent, pour une estimation spatio-temporelle du temps de parcours, les véhicules traceurs doivent être complétés par d'autres sources.

[2] La vitesse du flot est égale à la moyenne harmonique des vitesses mesurées (voir § 2.5.2.4)

Un autre point important à signaler est que le temps de parcours estimé par les véhicules traceurs n'est disponible que lorsque ces véhicules ont effectivement traversé le lien étudié : c'est donc un temps de parcours expérimenté ou réalisé. Ce n'est pas le cas pour le temps de parcours prédit. À un temps *t*, le temps de parcours mesuré sur un lien par un véhicule traceur est *TP*. Cela signifie qu'au temps *t* - *TP*, le temps de parcours du lien à prévoir est *TP*. Selon la longueur du lien et la variabilité des conditions de circulation (trafic en processus d'accumulation ou de dissipation), le temps réel de parcours au temps *t* pour un véhicule entrant pourra être significativement différent de *TP*, c'est-à-dire de ce qu'il était un certain temps auparavant.

Les mesures des véhicules traceurs ne sont pas toujours fiables, notamment lorsqu'ils sont *map-matchés* à la voie près. Par exemple, dans le cas d'une intersection comme dans la figure ci-dessous, le véhicule traceur est *map-matché* sur la voie de gauche utilisée pour tourner à gauche. En conséquence de cette erreur du GPS, on pourrait en déduire un temps beaucoup plus long que le « vrai » temps de parcours moyen nécessaire pour parcourir ce tronçon sur la voie qui continue tout droit.

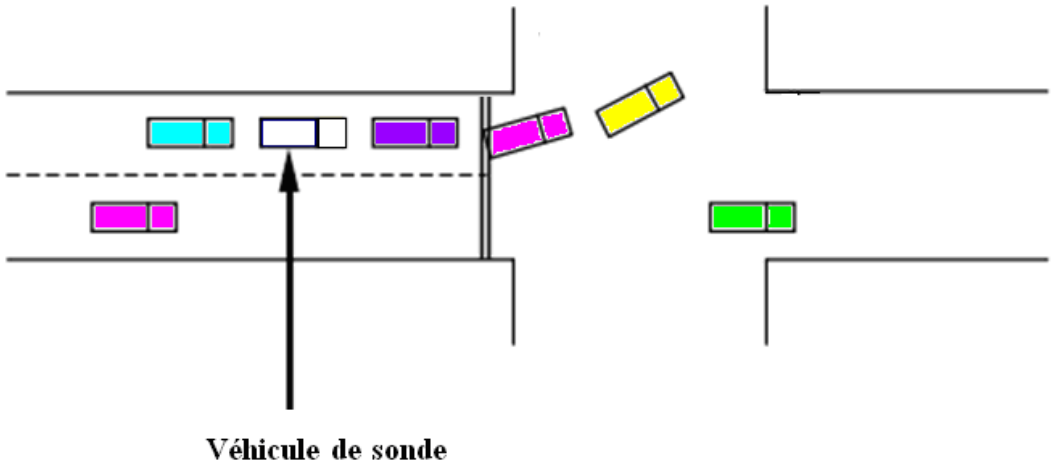


Figure 2. 16 : Véhicule de sonde au cas de tourne-à-gauche

2.4.3 Synthèse et conclusion

Les sources de données mobiles et fixes sont deux sources possédant chacune leurs propres caractéristiques et leurs propres erreurs. Chacune de ces sources a besoin d'être complétée et elles présentent l'avantage d'avoir des caractéristiques complémentaires. Afin d'exploiter au mieux ces deux sources et de compenser leurs erreurs, la solution de les fusionner semble naturelle.

Des chercheurs ont proposé des méthodes d'estimation des temps de parcours par les boucles magnétiques, d'autres par les véhicules traceurs, d'autres en fusionnant ces deux types de mesures. Le chapitre 3 détaille ces recherches en les analysant et les critiquant. Dans ce travail de thèse, nous proposons une méthode de fusion qui combine ces deux types de mesures, et sa mise en œuvre sur un réseau urbain.

Chapitre 3. Le temps de parcours

SOMMAIRE

3.1 Définition - 41 -

3.2 Principales méthodes d'estimation - 41 -
3.2.1 À partir des boucles magnétiques - 42 -
3.2.1.1 Principes d'utilisation - 42 -
3.2.1.2 Approches déterministes - 42 -
Méthode de Nam et Drew - 43 -
Description et procédure - 43 -
Analyse et critique - 45 -
Méthode d'Oh et al - 46 -
Description et procédure - 46 -
Analyse et critique - 47 -
Méthode de Bonvalet et Robin-Prévallée (BRP) - 47 -
Description et procédure - 47 -
Analyse et critique - 48 -
3.2.1.3 Approches stochastiques - 50 -
Méthode de Dailey - 50 -
Description et procédure - 50 -
Analyse et critique - 51 -
Méthode de Ye et al. - 52 -
Description et procédure - 52 -
Analyse et critique - 52 -
Méthode d'Hinsbergen et al. - 53 -
Description et procédure - 53 -
Analyse et critique - 54 -
3.2.2 À partir des véhicules traceurs - 55 -
3.2.2.1 Principe d'utilisation - 55 -
3.2.2.2 Méthode d'Herring et al. - 55 -
Description et procédure - 55 -
Analyse et critique - 56 -
3.2.2.3 Problème du nombre relatif de véhicules traceurs - 57 -
3.2.3 Par fusion des données des boucles magnétiques et des véhicules traceurs - 59 -
3.2.3.1 Approches déterministes - 60 -
Méthode d'Ashish Bhaskar : CUPRITE - 60 -
Description et procédure - 60 -
i Estimation des courbes cumulées - 60 -
ii Ajustement des courbes cumulées par les véhicules traceurs - 61 -
Analyse et critique - 62 -
Application de CUPRITE - 63 -
i. Estimation du temps de parcours par CUPRITE et CUPRITE temps réel - 64 -
ii. Sensibilité de CUPRITE à l'aléa des véhicules traceurs - 66 -
iii. Effet des valeurs aberrantes - 67 -
3.2.3.2 Méthodes stochastiques - 69 -
Méthode d'El Faouzi et al. - 69 -
Description et procédure - 69 -
Analyse et critique - 70 -

Méthode de Choi et Chung - 70 -
Description et procédure - 70 -
Analyse et critique - 71 -
3.2.4 Synthèse et conclusion - 72 -

3.1 Définition

Le temps de parcours d'un usager, appelé aussi temps de parcours individuel, sur une route ou un tronçon de route, est le temps écoulé depuis l'entrée jusqu'à la sortie. Le temps de parcours moyen sur un itinéraire pendant un intervalle de temps correspond à la moyenne arithmétique des temps de parcours individuels des véhicules sortis de cet itinéraire pendant cette période. Il existe plusieurs types de temps de parcours [El Faouzi et Leclercq, 2001]. Chaque type fait appel à ses propres caractéristiques et implique des modes de mesure ou d'estimation spécifiques. Dans le cadre de cette recherche, deux types de temps de parcours sont définis et le lecteur pourra se reporter à [El Faouzi et Lessort, 2001] :

- temps de parcours prédit : c'est le temps de parcours que va réaliser un véhicule qui vient d'entrer en amont pour sortir à l'aval d'un itinéraire ;
- temps de parcours réalisé ou expérimenté : c'est le temps de parcours réalisé par un véhicule qui sort de l'itinéraire à l'instant courant. Les véhicules traceurs fournissent ce type de temps de parcours.

3.2 Principales méthodes d'estimation

L'estimation du temps de parcours est un domaine de recherche actif depuis les années 1950. Certains chercheurs ont estimé le temps de parcours à partir des capteurs fixes, d'autres à partir des temps de parcours expérimentés des véhicules traceurs, et d'autres par la fusion des deux sources de données. C'est la nature des données disponibles qui oriente le choix de la méthode pour l'estimation du temps de parcours.

Dans ce chapitre, les méthodes d'estimation du temps de parcours sont détaillées ainsi que les méthodes de fusion de données. Ce chapitre est donc divisé en trois parties :

- la première partie développe la méthode indirecte c'est-à-dire l'estimation du temps de parcours à partir des boucles magnétiques et les limitations d'une telle estimation ;
- la deuxième partie explique la méthode directe, en d'autres termes l'estimation du temps de parcours à partir des véhicules traceurs. De plus, elle expose les limitations de cette méthode ;

- finalement, la troisième partie introduit le concept de fusion de données multi-sources, en expliquant les bénéfices de cette méthode.

3.2.1 À partir des boucles magnétiques

3.2.1.1 Principes d'utilisation

Les boucles magnétiques ont été introduites dans le domaine du trafic routier depuis les années 60 et sont à ce jour, les capteurs les plus utilisés principalement dans l'analyse du trafic en temps réel [Papageorgiou et al., 1990], [Ellouni, 1996] et [Middleton et al., 2002]. Elles entrent dans la famille des détecteurs de véhicules. Les détecteurs de véhicules peuvent être utilisés pour fournir des données pour la surveillance des routes, le contrôle de la circulation, etc. Des détecteurs de véhicules divers ont été développés et utilisés pour la collecte des données de trafic. Les boucles magnétiques sont le type d'équipement le plus répandu pour estimer le temps de parcours. Les données fournies par les boucles magnétiques décrivent l'état du trafic à l'endroit où les capteurs se trouvent et la couverture spatiale dépend de la densité de leur déploiement. Ces données sont quasi exhaustives parce qu'elles prennent en compte tous les véhicules qui ont parcouru le tronçon.

Nous distinguons deux types d'approches :

- approches déterministes ;
- approches stochastiques.

3.2.1.2 Approches déterministes

Une modélisation déterministe établit des relations entre variables qui sont strictement fonctionnelles. Une relation définie d'une manière déterministe peut être, par exemple, décrite comme une fonction du temps. En effet, les approches déterministes sont explicitées par des lois mathématiques bien connues. On peut donc prévoir exactement l'évolution des variables dans le temps. Parmi les approches déterministes en trafic, nous mentionnerons la méthode de Nam et Drew et la méthode d'Oh et al. qui se basent sur la méthode analytique classique des stocks, et la méthode de Bonvalet et Robin-Prévallée (BRP) utilisée comme un indicateur pour le temps de parcours inter-urbain.

Méthode de Nam et Drew

Description et procédure

[Nam et Drew, 1999] ont reformulé la théorie du trafic pour estimer le temps de parcours pour un lien inter-urbain. Le lien est défini entre deux emplacements de détecteurs et sans source ni fuite à mi-lien. Dans une première étape, Nam et Drew ont reformulé l'équation de conservation des véhicules en se basant sur la méthode analytique classique des stocks (les courbes cumulées à partir des détecteur à l'amont et à l'aval du lien) et en faisant l'hypothèse de conservation de véhicules (entre les emplacements des deux détecteurs). La variation du débit (q) entre l'amont (*upstream* indicé u) et l'aval (*downstream* indicé d) par rapport à la distance est égale à la variation de la densité (k) moyenne dans le temps :

$$\frac{q(x_u,t)-q(x_d,t)}{\Delta x}=\frac{k(t+\Delta t)-k(t)}{\Delta t} \tag{3.1}$$

Dans une seconde étape, ils ont reformulé la relation entre la vitesse, le débit, et la densité pour calculer le temps de parcours. Si les véhicules sont conservés, alors au temps t, la différence entre les courbes cumulées définit le nombre de véhicules qui traversent le lien. La densité est le ratio entre le nombre de véhicules et la longueur du lien. La durée totale du parcours de l'amont vers l'aval est la zone comprise entre les courbes cumulées en amont et en aval. Cette zone, qui est une fonction du débit et de la densité, est considérée comme trapézoïdale.

Soient $Q(x_u,t_n)$ et $Q(x_d,t_n)$ les nombres cumulés de véhicules en amont et en aval entre t_0 et t_n. Ils sont représentés par les équations suivantes :

$$Q(x_u,t_n)=\Delta t\sum_{i=1}^{n}q(x_u,t_i) \tag{3.2}$$

$$Q(x_d,t_n)=\Delta t\sum_{i=1}^{n}q(x_d,t_i) \tag{3.3}$$

avec :

- x_u et x_d représentent respectivement l'emplacement du détecteur en amont et en aval ;
- $q(x,t)$ est le flux mesuré à l'emplacement x et à l'instant t ;
- Δt est l'échantillonnage temporel.

Cas 1 : condition normale de circulation

Sous cette condition les courbes cumulées sont représentés comme suit (graphe issu de [Nam et Drew, 1999]) :

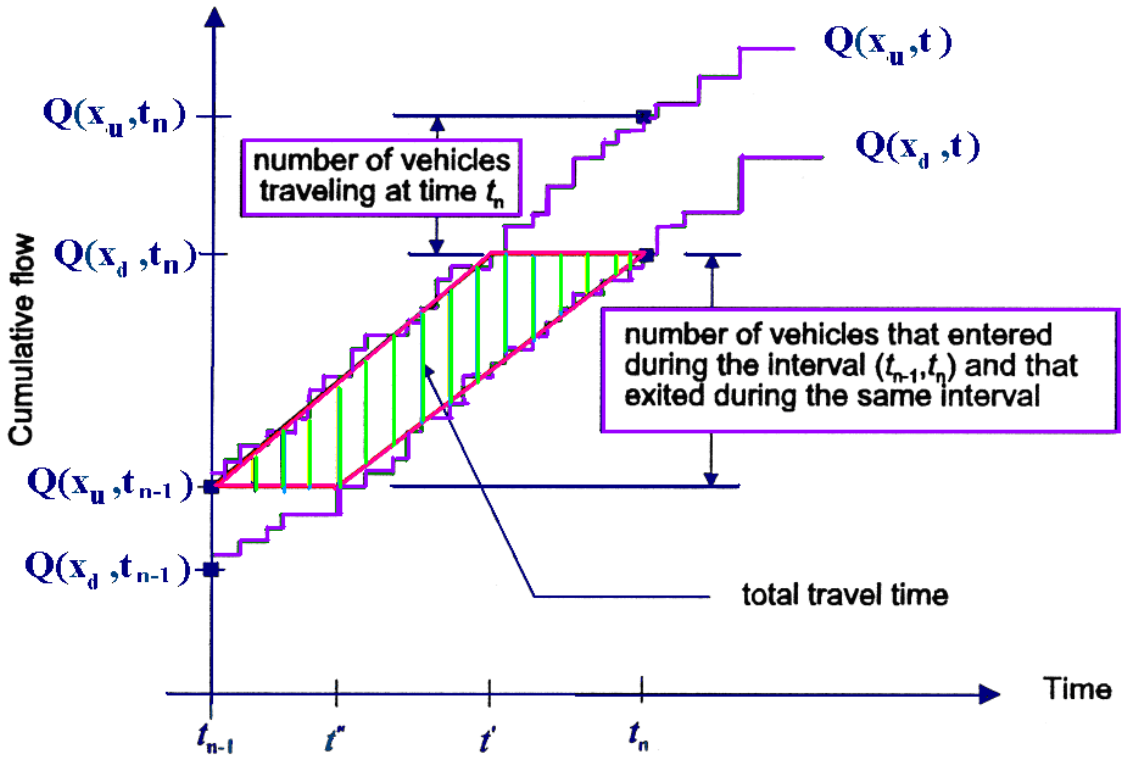


Figure 3. 1 : Figure représentant le temps de parcours total entre t_{n-1} et t_n en condition normale de circulation

et l'équation du temps de parcours est :

$$TP(t_n) = \frac{\Delta x}{2} \frac{q(x_u,t_n)k(t_{n-1}) + q(x_d,t_n)k(t_n)}{q(x_u,t_n)q(x_d,t_n)} \leq TP_{ff} \qquad (3.4)$$

- Δx est la distance entre les deux compteurs d'amont et d'aval ;
- TP_{ff} est le temps de parcours en « ***free flow*** » ou en **vitesse libre**.

Cas 2 : condition de congestion

Sous cette condition les courbes cumulées sont représentés comme suit (graphe issu de [Nam et Drew, 1999] :

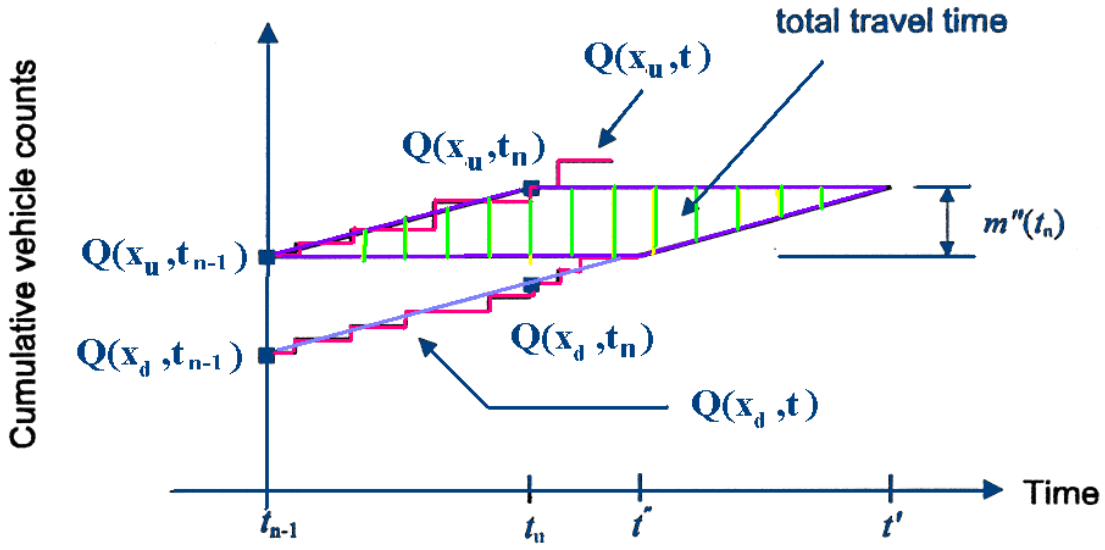


Figure 3. 2 : Figure représentant le temps de parcours total entre t_{n-1} et t_n en condition de congestion

et l'équation du temps de parcours est :

$$TP(t_n) = \frac{k(t_{n-1})\Delta x + \left[q(x_u, t_n) - q(x_d, t_n)\right]\Delta t}{2q(x_d, t_n)} \tag{3.5}$$

Analyse et critique

Nam et Drew ont validé leur méthode sur des données réelles (Queen Elizabeth Way, à Toronto, Canada), où la performance de la méthode est évaluée en comparant la vitesse estimée avec la vitesse de parcours réelle. Leur précision était proche de 90%. La façon dont ils ont estimé la vitesse de parcours réelle n'est valable que lorsqu'il n'y a pas de congestion, et, par ailleurs, leur évaluation des performances de la méthode ne fait pas de distinction entre l'état du trafic congestionné et non congestionné.

La base des équations (3.4 et 3.5) est la différence entre les compteurs cumulés en amont et en aval des liens qui sont sensibles à des erreurs. Même en absence de perturbations, il y a des risques de déviation et d'erreur de détection et de comptage. Nam et Drew ont constaté que les nombres en amont étaient de 3% plus élevés que les nombres en aval. Pour tenir compte de cette différence, ils ont appliqué un facteur d'ajustement pour chaque demi-heure et les mesures de débit en aval ont été multipliées par ce facteur d'ajustement. L'explication de la façon dont les facteurs d'ajustement sont déterminés n'a pas été documentée.

La méthode est limitée à un tronçon inter-urbain (absence de perturbations) en vertu de la discipline *first in first out* (FIFO) et ne peut pas être appliquée sur les réseaux urbains suite aux raisons suivantes :

- le principe de conservation de véhicule n'est généralement pas valide sur le réseau urbain, cela est dû entre autres, aux liens transversaux, parkings, voies privées, etc. provoquant des sources et des fuites de véhicules ;

- les équations 3.4 et 3.5 n'y sont pas applicables, premièrement parce que le flux est interrompu. Deuxièmement, ces équations sont dérivées tout en considérant une forme trapézoïdale entre les deux courbes. Or, dans un réseau urbain, cette forme dépend fortement de l'emplacement du détecteur et des circulations en accordéon, d'où le fait que la forme trapézoïdale n'est pas toujours valide.

Méthode d'Oh et al

Description et procédure

[Oh et al., 2003] ont défini le temps de parcours (équation 3.6) pour un lien inter-urbain comme une fonction de la densité (équation 3.7) et du flux en amont et en aval du lien. La densité et le débit sont estimés d'après les données provenant des détecteurs à l'entrée et à la sortie du lien. En effet, Oh et al. ont considéré que la loi de conservation des véhicules n'est pas respectée et qu'ils peuvent mesurer le débit entrant des sources à mi-lien et le débit sortant des fuites à mi-lien. En plus ils ont supposé qu'un lien est formé de *l* voies.

$$tt_s = \frac{\Delta x}{v} = \frac{\Delta x k}{q} = \frac{\Delta x\left[k\left(t_{n+1}\right)+k\left(t_n\right)\right]}{\left[q\left(x_u,t_n\right)+q\left(x_d,t_n\right)\right]} \tag{3.6}$$

$$k\left(t_{n+1}\right) = k\left(t_n\right) + \frac{\Delta t}{l\Delta x}\left[\left(lq_u\left(t\right)+l_{source}q_{source}\left(t\right)\right)-\alpha\left(lq_d\left(t\right)+l_{fuite}q_{fuite}\left(t\right)\right)\right] \tag{3.7}$$

où :

- Δx est la longueur du lien inter-urbain ;

- tt_s est le temps de parcours ;

- q_u, q_d sont respectivement le flux à l'amont et à l'aval du lien ;

- l est le nombre de voies sur le lien inter-urbain ;
- q_{source}, q_{fuite}, l_{source}, l_{fuite}, sont respectivement le débit et le nombre de voies pour les liens sources et fuites ;
- α est le paramètre de calibrage qui prend en compte les erreurs systématiques dans le comptage du détecteur.

Tout ce qui entre dans un lien doit sortir de ce lien alors les flux totaux des véhicules entrant sont égaux aux flux totaux sortant du lien. Ainsi, α est obtenu en estimant le rapport entre les flux totaux entrant et les flux totaux sortant :

$$\alpha = \frac{\sum_t \left(l q_u(t) + l_{source} q_{source}(t) \right)}{\sum_t \left(l q_q(t) + l_{fuite} q_{fuite}(t) \right)} \tag{3.8}$$

Analyse et critique

La méthode est applicable uniquement pour des sections entre deux détecteurs et elle nécessite des détecteurs sur les sources et fuites à mi-lien. La correction de l'erreur de comptage de détecteur α (équation 3.8) suppose que l'erreur est systématique pour une période importante. Ceci n'est pas applicable par exemple lorsque les véhicules sont étroitement espacés, car les détecteurs ont tendance à sous-compter. Par conséquent, l'estimation de la densité de la section pendant des intervalles d'estimation petits (chaque 5 min) est erronée et ceci entraîne des erreurs dans l'estimation du temps de parcours.

Méthode de Bonvalet et Robin-Prévallée (BRP)

Description et procédure

L'objectif de la recherche de Bonvalet et Robin-Prévallée (BRP) [Bonvalet et Robin-Prévallée, 1987] et [Buisson et Lessort, 2010] est d'étudier la relation entre le taux d'occupation et le temps de parcours en supposant qu'à chaque point de mesure disponible sur le réseau considéré est associée une zone représentative (un tronçon ou une portion du tronçon). Ces zones sont disjointes et leur union recouvre la totalité du réseau.

Sur chaque zone, et d'après leur étude, la relation entre le taux d'occupation mesuré par le capteur de la zone i et le temps de parcours moyen associé à cette zone i est linéaire :

$$TP_i = a_i \times o_i + b_i \tag{3.9}$$

où :

- TP_i : temps de parcours moyen de la zone *i* ;
- o_i : taux d'occupation mesuré par le capteur de la zone *i* ;
- a_i, b_i : coefficients à déterminer.

Pour calculer a_i et b_i, des conditions aux limites sont utilisées :

Condition fluide ou « free flow »

Sous cette condition, les véhicules circulent avec la vitesse de « *free flow* », alors le temps de parcours tend vers une valeur TP_{ff} tandis que le taux d'occupation tend vers zéro. Ainsi :

$$b_i = TP_{ff} \tag{3.10}$$

Condition de congestion complète

Sous cette condition, le débit tend vers 0 et le taux d'occupation tend vers 1. La zone est considérée comme saturée. Ainsi, le temps de parcours égale le temps d'écoulement du nombre maximum de véhicules que peut contenir le tronçon, N_i^{max}.

Ainsi :

$$TP_i = \frac{N_i^{\max}}{q_i} \times \left(o_i\right) + TP_{ff} \times \left(1 - o_i\right) \tag{3.11}$$

D'après la formule 3.11, nous remarquons que le temps de parcours est un barycentre entre la condition de fluidité et la condition de congestion, pondéré par le taux d'occupation.

Cette méthode a été testée pour les autoroutes urbaines parisiennes (périphérique) depuis les années 1987. Plusieurs gestionnaires de trafic tels que celui de Nantes-Métropole utilisent cette méthode.

Analyse et critique

Les mesures des boucles dans cette estimation sont considérées comme représentatives de liens entiers. C'est la raison pour laquelle, cette estimation extrapole spatialement les données ponctuelles des boucles. Ce type de méthode essaie d'estimer un temps de parcours en

effectuant une pondération entre un temps de parcours dans le cas de circulation libre et un temps de parcours dans le cas de circulation congestionnée. La pondération est estimée grâce au taux d'occupation. Cette méthode ne permet donc pas de décrire l'état réel du trafic. Elle est généralement applicable aux réseaux inter-urbains où la variation de l'état du trafic n'est pas significative contrairement aux réseaux urbains. Afin de bien comprendre le comportement de cette méthode, nous l'avons testée dans nos simulations sur un tronçon sans perturbations dont la variation du débit comme illustrée figure 2.7.

L'estimation du temps de parcours par la méthode BRP sur la simulation décrite §.2.3.3.3 est réalisé à partir des données issues d'un détecteur situé au milieu du tronçon. La figure 3.4 donne le résultat pour une réplication.

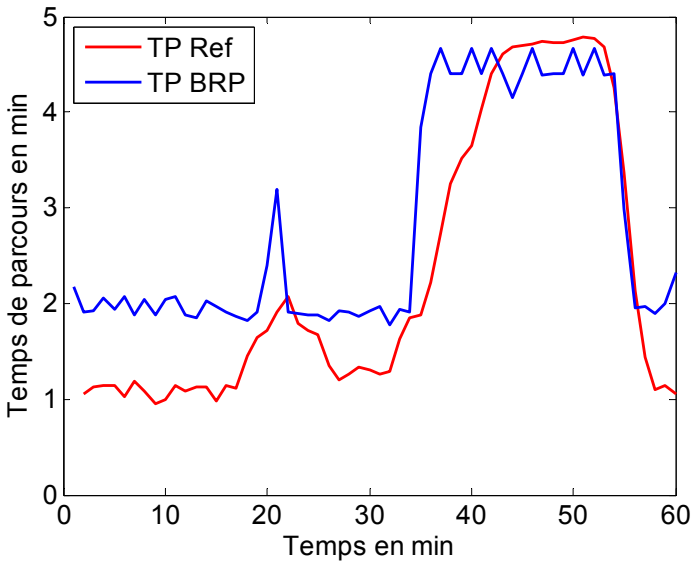


Figure 3. 3 : Temps de parcours estimé par BRP

D'après le graphe de la figure 3.3, pendant les 15 premières minutes, le trafic était en régime fluide. Même dans ce régime l'erreur commise par l'estimation BRP atteint 100% : la pondération proposée est discutable. Par ailleurs les variations du temps de parcours sont généralement déphasées par rapport aux variations vraies, ceci selon l'emplacement du détecteur utilisé par BRP par rapport aux détecteurs amont et aval.

3.2.1.3 Approches stochastiques

Les fluctuations aléatoires de certaines grandeurs sont souvent le reflet d'un processus stochastique. Dans ce contexte, les erreurs de modèle, ou toute autre quantité influant sur son évolution (paramètres de contrôle, perturbations de mesures) peuvent être représentées par des variables aléatoires. Leur nature probabiliste est souvent dictée par une connaissance *a priori* du système. Enfin, dans le cadre des algorithmes stochastiques, les composantes aléatoires du système permettent d'augmenter les capacités d'exploration et d'adaptation du modèle.

Dans la littérature, il existe plusieurs approches qui utilisent des données issues de capteurs fixes. Certains chercheurs [Rice et Van Zwet, 2004] ont observé une régression linéaire entre le futur temps de parcours et le temps de parcours à l'instant courant. La pente de cette relation linéaire change en fonction de l'heure de la journée. Les méthodes de [Hamed et al., 1995] et [Davis et al., 1990] sont fondées sur l'analyse des séries chronologiques telles que *auto-regressif integrated moving average* (ARIMA) pour la prédiction des paramètres de trafic comme le nombre de véhicule, la vitesse et l'occupation.

Ici, nous présentons deux méthodes d'estimation du temps de parcours par filtrage de Kalman développées par Dailey , Ye et al. et Hinsbergen et al.

Méthode de Dailey

Description et procédure

Dans le problème d'estimation de la vitesse posé par [Dailey, 1999], la fonction *f*, qui désigne la relation entre l'état actuel et l'état précédent, est linéaire, alors que *h*, la relation entre l'état et les observations, est non linéaire (équation 3.14). Dailey a considéré que l'état est la vitesse moyenne (v_k), et les observations sont le taux d'occupation (o_k) et le nombre de véhicules (N_k) issus des boucles magnétiques. Les équations suivantes définissent l'état et le modèle de mesure :

$$v_k = \begin{bmatrix} \overline{v}_{k-1} \\ \overline{v}_{k-2} \end{bmatrix} \tag{3.12}$$

$$y_k = \begin{bmatrix} o_k / N_k \\ o_{k-1} / N_{k-1} \end{bmatrix} \tag{3.13}$$

$$h(v_k) = \frac{L_k}{T}\begin{bmatrix} \dfrac{\sigma_v^2 + \overline{v}_{k-1}^2}{\overline{v}_{k-1}^3} \\ \dfrac{\sigma_v^2 + \overline{v}_{k-2}^2}{\overline{v}_{k-2}^3} \end{bmatrix} \tag{3.14}$$

La fonction *h* représente une relation non linéaire entre l'état et les mesures qui prend en considération la variance de la vitesse et est issue de l'équation développée par [Dailey, 1999] :

$$\frac{N_i}{o_i} = \left(\frac{\overline{v}T}{\overline{l}}\right)\left[\frac{\overline{v}^2}{\sigma_v^2 + \overline{v}^2}\right] \begin{cases} N_i \ \textit{nombre des vehicules} \\ o_i \ \textit{occupation} \\ T \ \textit{durée} \\ \overline{v} \ \textit{vitesse moyenne} \\ \overline{l} \ \textit{longueur moyenne des véhicules} \end{cases} \tag{3.15}$$

La fonction *f* qui régit l'évolution de l'état est une fonction d'auto-régression d'ordre 2 dont les paramètres sont estimés par un apprentissage préalable.

Le calcul de la vitesse est obtenu par filtrage de Kalman étendu (EKF).

Analyse et critique

Les mesures ont été prises à partir du système de gestion du trafic « *Washington State Department of Transportation (WSDOT) - central traffic management system (TMS)* » à Seattle. Les sites choisis pour les essais sont équipés de boucles magnétiques doubles qui agissent comme des détecteurs de vitesse d'une part et comme détecteurs de débit et de taux d'occupation d'autre part (et ce comme si elles étaient des boucles magnétiques simples). Les valeurs des débits et des taux d'occupation sont agrégées sur un intervalle de 20s. Dans cet algorithme, il est nécessaire de définir *a priori* la longueur moyenne $\overline{l}$ des véhicules ainsi qu'une estimation de la variabilité de la vitesse σ_v^2. Pour obtenir la longueur moyenne, Dailey utilise les données de *TMS* sur une durée totale de six jours. La variabilité de la vitesse est estimée à partir des vitesses des boucles magnétiques doubles considérées comme référence. La vitesse estimée est comparée avec la référence. Cette estimation a une variance plus grande que les données mesurées, mais suit généralement le caractère de la vitesse mesurée.

La longueur moyenne doit être estimée tronçon par tronçon et par intervalles de temps dans une étude préalable. Cependant, il y a plusieurs questions dans la mise en œuvre de l'EKF pour l'estimation de la vitesse. Comme l'a souligné Julier et al. [Julier et Uhlmann,

1997], la linéarisation dans l'EKF va produire des filtres très instables si certaines hypothèses ne sont pas remplies, et la dérivation des matrices jacobiennes conduit souvent à des difficultés de mise en œuvre importantes.

Méthode de Ye et al.

Description et procédure

Au lieu d'estimer la vitesse par le filtrage de Kalman étendu, Ye et al [Ye et al., 2006] ont proposé le filtrage de Kalman sans parfum (UKF). Nous avons déjà mentionné que l'équation d'observation est non-linéaire et contient la longueur moyenne des véhicules et la variance de la vitesse. Cette longueur moyenne varie souvent dans le temps et dans le réseau (par exemple, à cause des véhicules lourds). Cette variable est généralement difficile à estimer. Mais dans leur étude, ils ont supposé que cette longueur est constante. Ils ont démontré que le rapport entre de la variance de la vitesse et la vitesse est négligeable si les conditions de trafic sont normales. Mais dans les conditions de congestion, la vitesse diminue, et ce rapport devient significatif. Cette variance est déterminée par des expérimentations. Enfin, les paramètres de la fonction d'évolution f sont supposés constants et égaux.

Analyse et critique

Les mesures ont été prises à partir des données « *Texas Transportation Institute (TTI)* ». L'algorithme est appliqué en trois locations en inter-urbain. En plus des boucles magnétiques simples et doubles, les auteurs ont utilisé les données issues de ADR-6000, également connu comme Idris ou système de boucle intelligente. Idris est une technologie de détection et de classification automatique des véhicules. Les données issues de ADR-6000 sont par exemple la vitesse, la longueur et le temps de présence pour chaque véhicule. La méthode proposée par Ye et al. résout les problèmes de la non-linéarité. Les résultats de l'UKF sont meilleurs que ceux de l'EKF. Comme dans le cas de l'EKF, la longueur moyenne des véhicules doit être estimée. Cet algorithme requiert des détecteurs Idris qui ne sont pas souvent présents dans les réseaux.

Méthode d'Hinsbergen et al.

Description et procédure

[Hinsbergen et al., 2010] proposent une nouvelle implémentation, le filtre de Kalman étendu localisé (EKF-L) qui est beaucoup plus rapide sur les grands réseaux. Elle simplifie l'opération inverse inévitable lors du calcul du gain de Kalman.

En résumé pour décrire leur filtre, les auteurs ont discrétisé les liens par des cellules. Dans chaque cellule *i* d'un lien *j*, l'état du trafic à l'instant *t* est particulièrement bien décrit par la densité, qui est une fonction de l'état antérieur et du flux entrant et sortant de la cellule :

$$k_{ijt} = k_{ijt-1} + \frac{\Delta t}{\Delta l_j}\left(q^{in}_{ijt-1} - q^{out}_{ijt-1}\right) \tag{3.16}$$

avec Δt le pas de temps et Δl_j la longueur de la cellule du lien *j*.

Lorsque la densité dans chaque cellule est connue, d'autres variables comme la vitesse ou le débit peuvent être obtenues en utilisant le diagramme fondamental (voir figure 2.3).

Dans le filtre, le vecteur d'état est le vecteur contenant toutes les densités de toutes les cellules à l'instant *t*. L'étape de prédiction est définie par :

$$k_t^- = g\left(k_{t-1}\right) \tag{3.17}$$

$$P_t^- = A_t P_{t-1} A_t^T + Q \tag{3.18}$$

avec :

- $g\left(k_{t-1}\right)$ est l'équation d'évolution donnée dans la formule (3.16) pour chaque cellule ;
- Q est la matrice de covariance d'erreur d'évolution du modèle qui est prise constante dans le temps ;
- P_t^- est l'estimateur *a priori* de la matrice de variance-covariance de l'erreur du vecteur d'état.

La linéarisation du modèle introduit la dérivée du modèle d'état :

$$A_t = \frac{\partial g\left(k_{t-1}\right)}{\partial k_{t-1}} \tag{3.19}$$

Il est important de noter que la matrice de covariance d'erreur contient généralement de nombreuses valeurs qui sont proches de zéro. D'un point de vue physique, ceci est causé par le fait que deux endroits sur le réseau routier sont éloignés l'un de l'autre.

Dans cette étude, les mesures sont les vitesses et le diagramme fondamental est utilisé comme fonction d'observation. Aussi, les mesures sont linéarisées autour de l'état courant en dérivant la fonction d'observation de l'état.

Les covariances sont influencées par la linéarisation du diagramme fondamental et le gain de Kalman. En raison de la non-linéarité du système et la stochasticité du modèle et des données, il est très difficile de prouver analytiquement les conditions dans lesquelles les covariances diminuent avec la distance séparant deux endroits du réseau. Cependant, grâce à l'expérimentation de Hinsbergen et al., il a été en effet confirmé que la plupart des valeurs dans la matrice de covariance sont généralement très proches de zéro.

Dans cette étude, il est donc proposé d'utiliser une mesure d'un détecteur pour corriger l'état des cellules dans le voisinage de ce détecteur. Le schéma résultant est nommé filtre de Kalman étendu localisé pour indiquer le caractère local des corrections.

Dans ce filtre, de nombreux EKF locaux sont appelés successivement pour chaque cellule qui contient des mesures, au lieu de construire un grand filtre pour l'ensemble du réseau. Pour terminer, dans cette étude, les covariances des bruits sont supposés constantes.

Analyse et critique

L'EKF-L a été testé sur un petit réseau de 6 liens divisés en 59 cellules et a été comparé avec un EKF. La précision de l'EKF est légèrement meilleure que celle de l'EKF-L pour les premières 20 cellules puis, pour les suivantes, les résultats sont similaires. L'EKF-L en temps de calcul est plus rapide que l'EKF.

Ce filtre est applicable sur un réseau inter-urbain et non pas sur un réseau urbain puisqu'il ne considère pas la non-conservation des véhicules. En plus, il nécessite l'équipement de boucles magnétiques sur chaque cellule.

3.2.2 À partir des véhicules traceurs

3.2.2.1 Principe d'utilisation

Les systèmes de positionnement global par satellites (GNSS, *Global Navigation Satellite System*, dont bien sûr le GPS), mais aussi les téléphones mobiles (qui se localisent par rapport aux relais du réseau) sont capables de fournir la trajectoire (espace-temps) des mobiles équipés. Les véhicules avec de tels équipements sont appelés véhicules sondes ou traceurs. Les données obtenues de tels véhicules sont les coordonnées de position (longitudes et latitudes) à différents intervalles de temps (définis par la fréquence de transmission des données entre les véhicules et un central) et (le cas échéant) leur vitesse. Ainsi, il est possible d'accéder directement au temps de parcours d'un point à un autre le long du chemin parcouru par chaque véhicule.

3.2.2.2 Méthode d'Herring et al.

L'une des recherches les plus récentes sur l'estimation du temps de parcours par les véhicules traceurs est celle de l'équipe de recherche de l'université de Berkeley.

Description et procédure

[Herring et al., 2010] propose une modélisation probabiliste comme cadre de travail, en utilisant les véhicules traceurs pour la modélisation de l'évolution des états du trafic comme un couplage de modèles de Markov cachés (*Coupled Hidden Markov Model,* CHMM). En particulier, ils ont proposé un algorithme avec des mesures reçues à des endroits et à des moments aléatoires. L'auteur définit la distribution du temps de parcours de chaque observation (temps entre les mesures consécutives) en fonction :

- des distributions des temps de parcours des liens traversés ;
- de la répartition spatiale de la position des véhicules sur chaque lien traversé.

Les hypothèses utilisées pour formuler leur algorithme sont les suivantes :

1. le digramme fondamental triangulaire (voir figure 2.4) ;
2. la stationnarité du trafic : au cours de chaque intervalle d'estimation, les paramètres des cycles des feux (rouge et vert) ne varient pas. La densité d'arrivée est constante. La

dynamique du trafic évolue régulièrement avec la période du cycle des feux. En particulier, il n'y a aucune augmentation ou diminution systématique de la longueur de la file d'attente ;

3. le principe *(FIFO)* : le dépassement sur le réseau routier est négligé ;

4. les états de congestion sont discrets : pour chaque jour et chaque intervalle de temps, les conditions de circulation sur le lien sont représentées par une valeur discrète qui indique le niveau de congestion ;

5. l'indépendance conditionnelle du temps de parcours du lien : la distribution du temps de parcours du lien conditionnée par son état est indépendante de toutes les variables du trafic ;

6. l'indépendance conditionnelle des états de transitions : chaque lien est en corrélation avec une certaine série de liens voisins (le voisinage pris en compte est petit), mais elle est indépendante du reste du réseau.

Les conditions de circulation varient dans l'espace et le temps. Compte tenu des hypothèses 4, 5 et 6, les conditions spatio-temporelles de trafic sont modélisées en utilisant une chaîne de Markov cachée. L'état de chaque lien à n'importe quel instant dans le temps n'est pas observé donc Herring et al. le considèrent comme caché. La distribution des temps de parcours sur chaque lien est conditionnée par son état caché à partir duquel ils disposent d'observations de véhicules traceurs à travers le réseau. En effet, ils estiment les probabilités d'état initial pour chaque lien ainsi que les fonctions de distribution discrètes de probabilité de transition des liens (pour chaque lien et chaque intervalle de temps t, la probabilité pour qu'il soit dans l'état s à l'instant $t+1$ tout en connaissant l'état de ses voisins à l'instant t est donnée par la fonction de distribution discrète de probabilité de transition). Ces estimations leur permettent d'estimer la distribution du temps de parcours de ce lien à partir de son état.

Analyse et critique

Cette méthode a été évaluée en utilisant des données à partir d'une flotte de taxis (500 véhicules) à San Francisco, CA. Ces taxis envoient les données GPS à leur serveur toutes les minutes. L'échantillon ne fournit pas d'informations détaillées sur l'endroit où les véhicules ont rencontré du retard ou la raison de ce retard (soit un retard dû aux feux de signalisation soit dû à la congestion, etc.). Leur méthode fournit une augmentation de la précision de l'estimation de 35% par rapport à d'autres approches.

Il est clair que toutes ces hypothèses ne peuvent pas être tout le temps satisfaites dans un réseau réel sous trafic réel mais il est de la conviction des auteurs que c'est une bonne approximation. Dans leurs futures études, ils examineront comment cette approximation affecte les résultats et si ces hypothèses peuvent être assouplies, tout en ayant un modèle possible à résoudre.

Ils ont supposé aussi que les fonctions de distribution du temps de parcours d'un lien, tout en connaissant son état, sont des gaussiennes indépendantes. En général, le choix d'une distribution gaussienne restreint la flexibilité du modèle (il ne peut saisir qu'une modalité de trafic unique), mais d'un autre côté offre une vraie simplicité de résolution. D'autres distributions pourraient être considérées (gamma ou log-normale), moins simples mais toutefois d'utilisation possible.

3.2.2.3 Problème du nombre relatif de véhicules traceurs

Les véhicules traceurs représentent un échantillon aléatoire de la population des véhicules dans le réseau. Par conséquent, la précision de l'estimation du temps de parcours avec des véhicules traceurs est liée à leur nombre. [El Faouzi et Lessort, 2001] ont cherché à déterminer le nombre minimal de véhicules traceurs nécessaire afin que le temps de parcours estimé de ces données respect une précision garantie. Ils ont quantifié l'intervalle de confiance (au niveau de confiance $(1-\alpha)$) de la moyenne des temps de parcours des véhicules traceurs $\hat{TP}$ avec le temps de parcours moyen obtenu à partir de l'ensemble des véhicules composant le flot du trafic TP :

$$P\left(\left|TP-\hat{TP}\right| \le \varepsilon_a(\alpha)\right) = 1-\alpha \tag{3.20}$$

Soit $v(\Delta x,t)$ la vitesse du flot à l'instant t sur un tronçon Δx et $v_1, v_2, ..., v_p$ les vitesses individuelles des véhicules traceurs, alors, $v(\Delta x,t)$ est approximée par la moyenne arithmétique des vitesses des véhicules traceurs. Chaque mesure v_i est entachée d'une incertitude δ_i. Les auteurs ont supposé que les $(\delta_i)_{i=1,...,P}$ sont de moyenne nulle et de variance $\sigma^2 = \mathrm{var}(\delta_i)$ ce qui implique que les $(\delta_i / p)_{i=1,...,P}$ sont des variables aléatoires indépendantes de moyenne nulle et de variance :

$$\mathrm{var}(\delta_i / p)_{i=1,...,P} = \sigma^2 / p^2 \tag{3.21}$$

La somme de ces p erreurs est l'erreur aléatoire totale de v_E. La variance de la distribution d'échantillonnage est $\text{var}(v_E) = \sigma_{v_E}^2 = \sigma^2 / p$.

En supposant σ^2 connue, la taille de l'échantillon est :

$$p = \left(\frac{\sigma}{\sigma_{v_E}} \right)^2 \tag{3.22}$$

L'intervalle de confiance de la moyenne montre que le seuil $\varepsilon_a(\alpha)$ est égal à $u_{\alpha/2}\sigma_{v_E}$ où $u_{\alpha/2}$ désigne la marge d'erreur issue dans la table de Gauss-Laplace en fonction de la probabilité attribuée au paramètre $(u_{0,05/2} = 1,96,\ u_{0,1/2} = 1,645)$.

Alors p devient :

$$p = \left(\frac{\sigma}{\sigma_{v_E}} \right)^2 = \left(\frac{u_{\alpha/2}\sigma}{u_{\alpha/2}\sigma_{v_E}} \right)^2 = \left(\frac{u_{\alpha/2}\sigma}{\varepsilon(\alpha)} \right)^2 \tag{3.23}$$

Généralement σ (l'écart type de la population) est inconnu, alors il est estimé à partir de l'échantillon étudié. D'après l'équation (3.23) la taille d'échantillon pour un critère de fiabilité donné dépend directement de l'écart type de la population. Plus cet écart type augmente, plus la taille de l'échantillon augmente. L'écart type des liens pour les zones urbaines devrait être plus grand que sur autoroutes en raison des circulations en accordéon « *stop-and-go* ». Les valeurs observées pour cet écart type pour des intervalles de temps de 5 min sur un lien urbain à Lucerne Suisse varient de 26% à 55% [Bhaskar, 2009b]. [Turner et Holderner, 1995] ont remarqué que cet écart type pour une autoroute à Houston pendant le même intervalle de 5 min varie de 5% à 15%. Ils estiment que pour 5 min d'intervalle de temps avec un niveau de confiance de 95%, la taille des échantillons des véhicules traceurs varie d'un véhicule toutes les 5 min dans des conditions de circulation fluide à 6 véhicules pour les conditions de circulation congestionnée.

3.2.3 Par fusion des données des boucles magnétiques et des véhicules traceurs

Les données fournies par les boucles magnétiques et les véhicules traceurs ont des propriétés complémentaires parmi lesquelles :

- l'état du trafic décrit par les boucles magnétiques est représenté à l'endroit où les capteurs se trouvent tandis que les véhicules traceurs le décrivent sur tout leur itinéraire ;

- les mesures de boucles magnétiques sont quasi exhaustives parce qu'elles prennent en compte tous les véhicules qui ont parcouru le tronçon alors que les mesures des véhicules traceurs sont non exhaustives car elles n'intègrent qu'une partie des véhicules qui ont parcouru le tronçon ;

- finalement, les données des boucles présentent une couverture spatiale « statique » qui dépend de la densité et de l'emplacement des capteurs alors que celles des véhicules traceurs représentent une couverture spatiale « dynamique ». Les véhicules traceurs sont potentiellement présents partout, alors que les boucles n'existent pas sur tout le réseau.

Ces deux sources de données peuvent montrer des incohérences et ont besoin d'être complétées l'une par l'autre. Quand on prend en compte la qualité des données fournies par chaque source, nous pouvons détecter et éliminer les incohérences dans le but d'augmenter la précision, la fiabilité et la robustesse de l'estimation et de la prédiction [El Faouzi, 2000b] [El Faouzi, 2004]. La fusion permet de fournir une meilleure image (globale et complète) de l'état du trafic et offre des solutions aux imperfections des mesures.

On peut classer aussi les méthodes en :

- approche déterministe ;

- approche stochastique.

3.2.3.1 Approches déterministes

Méthode d'Ashish Bhaskar : CUPRITE

Description et procédure

L'estimation du temps de parcours par CUPRITE est divisée en deux étapes :

- estimation des courbes cumulées du nombre de véhicules par les boucles magnétiques [Bhaskar et al., 2008], [Bhaskar et al., 2009b] ;
- correction des courbes cumulées par les véhicules traceurs [Bhaskar, 2009b], [Bhaskar et al., 2011].

i Estimation des courbes cumulées

Dans un carrefour à feux et au cours de la phase de signal rouge, les véhicules s'arrêtent et la file d'attente augmente. Au cours de la phase de signal vert, les véhicules de la file d'attente sont dissipés au débit de saturation et le reste suit selon un certain motif d'arrivée. C'est pour cela que dans CUPRITE, la courbe cumulée à l'emplacement du détecteur est estimée en fonction de la disponibilité des données du détecteur, des données du feu (signal vert / signal rouge) et du débit de saturation, de la manière suivante :

CUPRITE estime la courbe cumulée des nombres des véhicules en aval « *downstream cumulative plot* » ($N_d(t)$) à partir de la courbe cumulée des nombres des véhicules en amont « *upstream cumulative plot* » ($N_u(t)$). La demande est déduite par translation horizontale de $N_u(t)$ du temps de parcours fluide (*free-flow)* (t_{ff}) du tronçon, i.e. : $N_u(t\text{-}t_{ff})$. Le débit est défini comme étant de zéro pour la phase de rouge. Pour la phase de vert, si la demande est plus grande que le nombre cumulé en aval alors le débit est saturé sinon le modèle d'écoulement est le même que celui de la demande. Les hypothèses sont : la discipline FIFO et la conservation du nombre de véhicules.

$$\begin{aligned} &Durant\ la\ phase\ rouge \Rightarrow q = 0 \\ &Durant\ la\ phase\ verte \Rightarrow \\ &\begin{cases} si\ CPdemande(t) > N_d(t) & q = saturation \\ \sin on & q = \left(\frac{d(CPdemande(t))}{dt} \right) et\ CPdemande(t) = N_u(t - tff) \end{cases} \end{aligned} \tag{3.24}$$

ii Ajustement des courbes cumulées par les véhicules traceurs

Dans CUPRITE, les véhicules sondes sont définis comme un horodateur aux intersections où se situent les boucles (et où les courbes cumulées sont générées). De plus, l'auteur a supposé qu'il n'y a pas d'erreur sur l'instant où le véhicule sonde se trouve en amont (t_u) et en aval (t_d) (précision infinie).

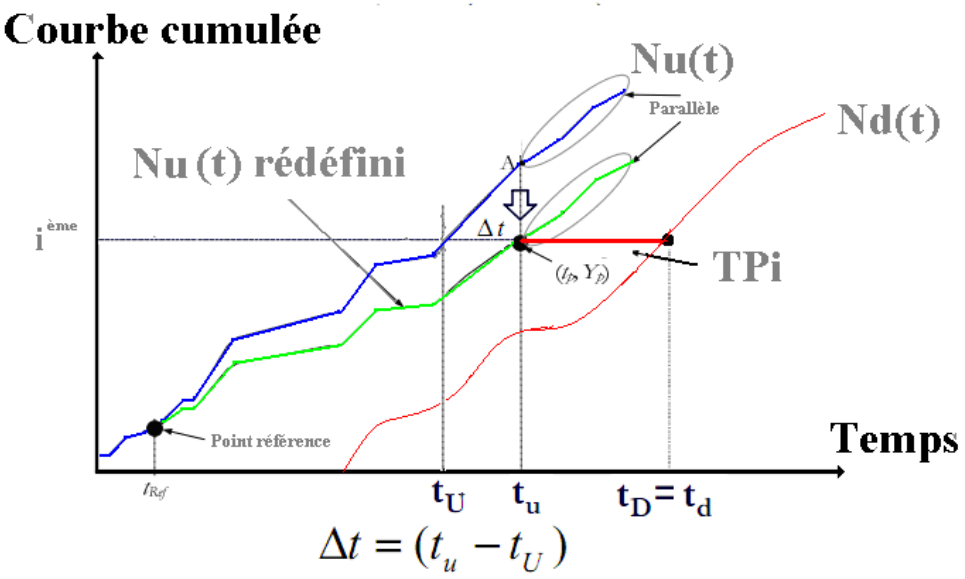


Figure 3. 4 : Correction de la courbe cumulée en amont

CUPRITE suggère une correction déterministe de la courbe cumulée en amont, $N_u(t)$, qui corresponde à la durée du parcours instantané mesuré (TP_i), sous les mêmes hypothèses déjà mentionnées où la courbe cumulée d'aval reste fixe. En effet, les données relatives aux véhicules sondes sont utilisées dans CUPRITE pour définir les points où $N_u(t)$ devrait passer en supposant connaître le temps où un véhicule traceur est à l'amont et à l'aval. Cette information est fixée à la courbe cumulée en aval, c'est-à-dire $t_D = t_d$. Alors le véhicule sonde est le i^{ème} véhicule des courbes cumulées, et il devrait être passé en amont à t_u=t_d-TP_i, ce qui rend $\Delta t = t_u - t_U$ corrigeable. Ceci est représenté dans la figure 3.4. En d'autres termes et en vertu de la discipline du trafic FIFO, la distance horizontale entre les courbes donne le temps de parcours pour le i^{ème} véhicule et les moments où il est en amont, (t_U), et en aval, (t_D). Sous les hypothèses mentionnées, Δt doit être nulle mais ce n'est fréquemment pas le cas par exemple à cause d'entrées ou sorties perturbant la conservation des véhicules ou d'erreurs de détection et comptage par les boucles. Ainsi, à partir de :

a) un point de référence (t_{ref}, $N_u(t_{ref})$) digne de confiance sur la courbe cumulée (par exemple le point de correction précédent par véhicule traceur ou le point initial) ;

b) un point (t_p, Y_p) où $N_u(t)$ devrait passer ;

CUPRITE redéfinit $N_u(t)$ et en appliquant les corrections affines suivantes :

- pas de correction avant t_{Ref};
- entre t_{Ref} et t_p, $N_u(t)$ est corrigée verticalement par le coefficient suivant :

$$\left(\frac{y_p - N_u\left(t_{ref}\right)}{N_u\left(t_p\right) - N_u\left(t_{ref}\right)} - 1\right) \times \left(N_u\left(t\right) - N_u\left(t_{ref}\right)\right) \text{ si } N_u\left(t_p\right) \neq N_u\left(t_{ref}\right) \tag{3.25}$$

- après t_p, la courbe est décalée verticalement de sorte qu'elle soit parallèle à $Nu(t)$. Le décalage égale :

$$\left(y_p - N_u\left(t_p\right)\right) \tag{3.26}$$

Analyse et critique

La validation est obtenue par des données réelles sur un site en ville de Lucerne, en Suisse. Le site d'étude est un réseau urbain typique disposant des caractéristiques suivantes :

- un trafic mixte (avec des bus) ;
- des arrêts de bus ;
- des fuites et sources à mi-lien ;
- de nombreux piétons qui passent fréquemment.

Les détecteurs à boucles sur le site ne sont pas parfaits. Le test est effectué pour les conditions de circulation saturées ou non. La précision est de plus de 95% et c'est l'intégration des données de véhicules sondes qui a amélioré considérablement la précision.

Dans CUPRITE les données des véhicules traceurs sont :

- utilisées *a posteriori* puisqu'il corrige les courbes cumulées avant le passage des données de véhicules traceurs ;
- ne sont pas utilisées à leur instant exact de passage mais arrondi à la minute supérieure ;
- supposées exactes.

Cette méthode est sensible à l'aléa des véhicules traceurs, qui peuvent ne pas être représentatifs du temps de parcours de référence, sans pour autant apparaitre comme aberrants. Lorsque la différence entre le temps de parcours des véhicules traceurs et le temps de parcours de référence est significative, CUPRITE corrige $N_u(t)$ par des coefficients erronés. Ceci est d'autant plus vrai si l'on considère un véhicule traceur aberrant.

De plus, des comptages par boucles peuvent aussi être aberrants et nous verrons comment cela impacte CUPRITE.

Enfin, dans un réseau urbain, les liens sont généralement équipés de boucles à la sortie du lien mais pas à l'entrée et les affectations des mouvements tournants aux intersections sont inconnues. Nous renvoyons au chapitre 8 (voir §.8.3.2) l'analyse de CUPRITE dans les conditions d'un réseau.

Dans ce qui suit, CUPRITE est testé par des simulations pour analyser son comportement.

Application de CUPRITE

Nous soulignons ici :

- l'adaptation de CUPRITE à une application en temps réel ;
- la sensibilité de CUPRITE à l'aléa des véhicules traceurs ;
- le comportement de CUPRITE dans le cas de données aberrantes.

i. Estimation du temps de parcours par CUPRITE et CUPRITE temps réel

Les données des simulations déjà décrites §.2.3.3.3 (un tronçon équipé d'une boucle à son entrée et à sa sortie avec un flux entrant variable dans le temps) sont réutilisées. Pour bien comprendre comment CUPRITE redéfinit la courbe $N_u(t)$, un véhicule traceur au premier instant est considéré comme étant le véhicule traceur de référence et un autre véhicule traceur est aléatoirement sélectionné parmi l'ensemble des véhicules. Ce véhicule choisi est sorti du tronçon à l'instant *37,2* min. Pour souligner la différence entre CUPRITE et son application en temps réel, le tronçon est considéré sans perturbations.

La figure 3.5 illustre le processus de correction de la courbe d'entrée par la méthode de CUPRITE sans perturbations. La courbe redéfinie par CUPRITE est notée $N_u'(t)$.

A l'instant 37,2 min, le véhicule traceur sort du tronçon et a eu besoin de 3,1 min pour le traverser. Cette information est utilisée à l'instant *38* car le calcul du $N_d(37,2)$ est l'interpolation linéaire de $N_d(38)$ et $N_d(37)$, or l'information de $N_d(38)$ n'est pas disponible à l'instant *37,2* mais seulement à l'instant 38.

$N_d(38)=479$ et $N_u(38-3,1)=484,4$ donc après l'instant *38* la nouvelle courbe est la translation horizontale de l'ancienne courbe d'une valeur de *484,4 - 479=5,4*. Avant l'instant *38* et jusqu'à l'instant *0* où le temps de parcours initial était supposé connu la courbe est translatée en sifflet (verticalement par des valeurs calculées au fur et à mesure).

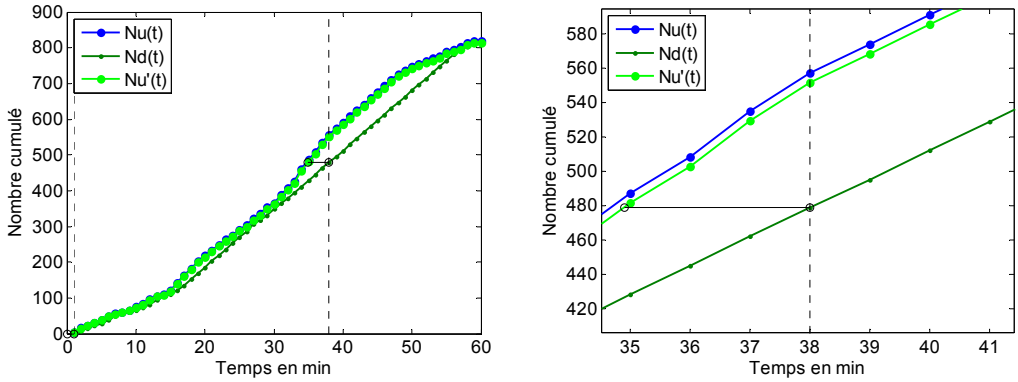


Figure 3. 5 : Courbe cumulée d'amont ajustée par CUPRITE

CUPRITE corrige correctement à l'occurrence du VT et après, mais aussi avant l'occurrence du VT. Pour une application en temps réel, une telle correction n'aurait pas été utile puisqu'elle affecte les temps de parcours passés alors que ce sont les temps de parcours présents, voire futurs, qui nous intéressent. En plus toute correction portant sur $N_u(t)$ avant

l'occurrence du véhicule utilise des informations *a posteriori*. Pour un contexte d'estimation en temps réel, la courbe en amont doit être corrigée seulement à et après l'instant de sortie du véhicule traceur. La figure 3.6 illustre la correction de la courbe en amont par CUPRITE temps réel, notée $N_u''(t)$. A partir de l'instant d'occurrence du véhicule traceur, $N_u''(t)$ dévie de $N_u(t)$.

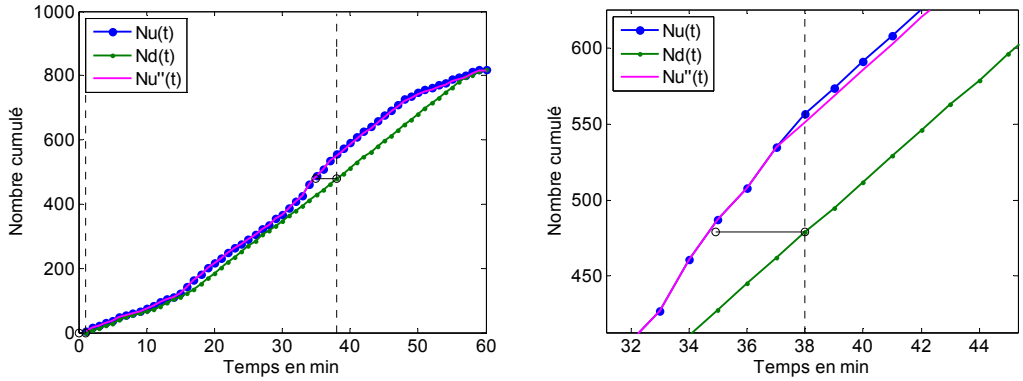


Figure 3. 6 : Courbe cumulée en amont ajustée par CUPRITE temps réel

La seule différence entre CUPRITE et CUPRITE temps réel est qu'avant l'occurrence du véhicule traceur la première corrige $N_u(t)$ alors que la deuxième ne la corrige pas.

La figure 3.7 ci-dessous comprend les trois courbes en amont $N_u(t)$ non corrigée, la courbe $N_u'(t)$ corrigée par CUPRITE et $N_u''(t)$ corrigée par CUPRITE temps réel en indiquant l'instant de passage du véhicule traceur.

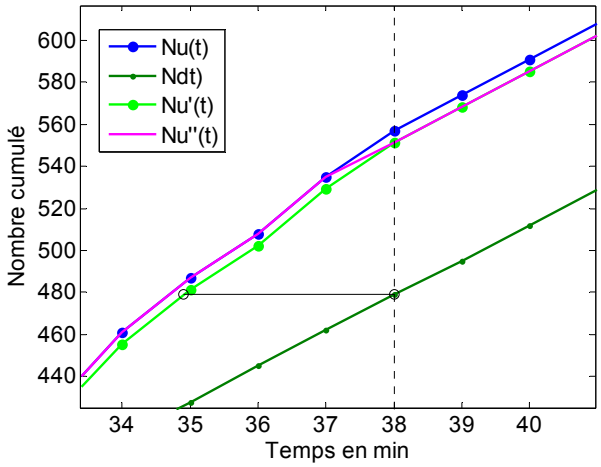


Figure 3. 7 : Courbes cumulées d'amont ajustées par CUPRITE et CUPRITE temps réel

La figure 3.8 illustre la correction de $N_u(t)$ par CUPRITE et CUPRITE temps réel quand la perturbation est introduite (1% source).

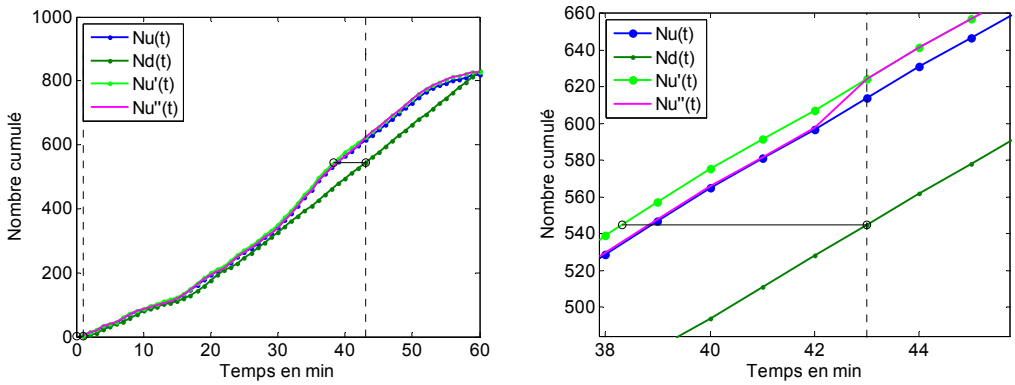


Figure 3. 8 : Courbe cumulée an amont ajustée avec perturbation

ii. Sensibilité de CUPRITE à l'aléa des véhicules traceurs

La méthode analytique classique de stock donne une estimation non biaisée du temps de parcours lorsque les véhicules dans le lien étudié sont conservés et si les données des boucles sont exactes. Dans un tel cas, la correction CUPRITE ainsi que CUPRITE temps réel devrait être zéro. Mais dans le cas où l'écart entre le temps de parcours de référence (noté TP Ref) et le temps de parcours du véhicule traceur est significatif, cette correction va biaiser l'estimation du temps de parcours puisque la correction de $N_u(t)$ est en relation directe avec ce temps de parcours. Dans CUPRITE temps réel, ce biais aura lieu seulement après l'instant de véhicule traceur puisque la correction n'aura lieu qu'après cet instant. La figure 3.9 illustre la sensibilité de CUPRITE à l'aléa d'un véhicule traceur parmi tous les véhicules. La différence entre le temps de parcours de ce véhicule et le temps de parcours de référence est de *0,4* min (*3,3-2,9*). Cet écart a biaisé la correction de $N_u(t)$ qui, à son tour, a biaisé l'estimation du temps de parcours. De même, comme attendu, TP CUPRITE temps réel (CUPRITE TR) colle à TP CUPRITE après l'instant du véhicule traceur et avec TP estimé par la méthode analytique classique des stocks (MACS) avant cet instant. Nous remarquons sur le graphe l'écart entre TP CUPRITE et TP Ref.

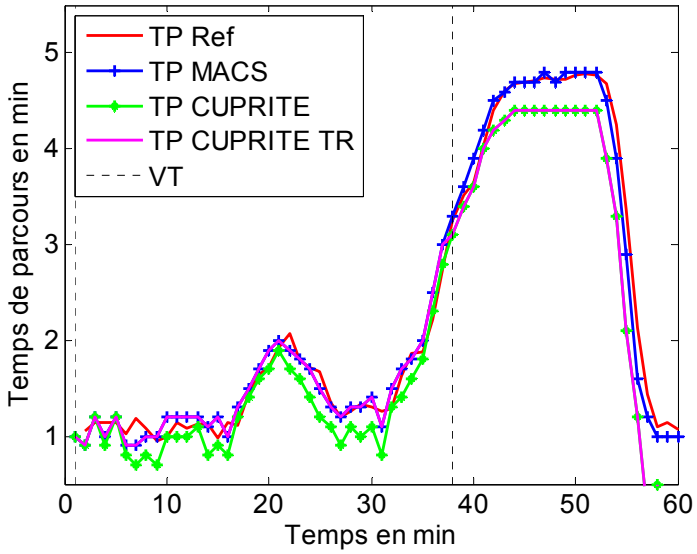


Figure 3. 9 : Sensibilité de CUPRITE

iii. Effet des valeurs aberrantes

La figure 3.10 illustre l'effet d'une valeur aberrante de véhicule sonde (2,8 min au lieu de 1,1 min à l'instant 29) sur l'estimation du temps de parcours.

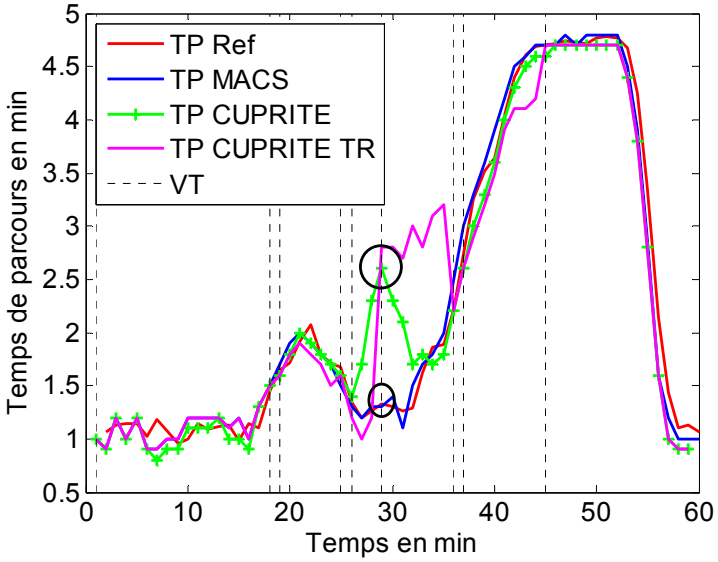


Figure 3. 10 : Effet d'une valeur aberrante de véhicule traceur sur le temps de parcours

Le TP CUPRITE à l'instant 29 est celui du véhicule traceur. CUPRITE corrige la courbe cumulée d'amont $N_u(t)$ en fonction du TP du véhicule traceur avant son occurrence et

jusqu'au véhicule traceur précédent (instant *26*), alors entre les instants *26* et *29* l'écart entre TP CUPRITE et TP Ref augmente progressivement. En revanche dans CUPRITE TR l'écart est brutal juste à l'instant *29* du VT. De même, on remarque qu'entre les instants *29* et *36* l'écart entre TP Ref et CUPRITE TR est supérieur à celui entre TP Ref et CUPRITE puisque ce dernier a corrigé $N_u(t)$ entre ces deux instants avec le temps de parcours du véhicule traceur suivant à l'instant *36*, ce qui n'est pas le cas de CUPRITE TR.

La figure.3.11 illustre l'effet d'une valeur aberrante du détecteur, où la valeur comptée en aval a été fixée pendant 2 minutes ($N_d(29)=N_d(30)$). Ici, CUPRITE et CUPRITE temps réel ont le même comportement puisqu'ils fixent la courbe cumulée en aval, $N_d(t)$, et corrigent celle en amont, or les valeurs aberrantes sont dans celle en aval.

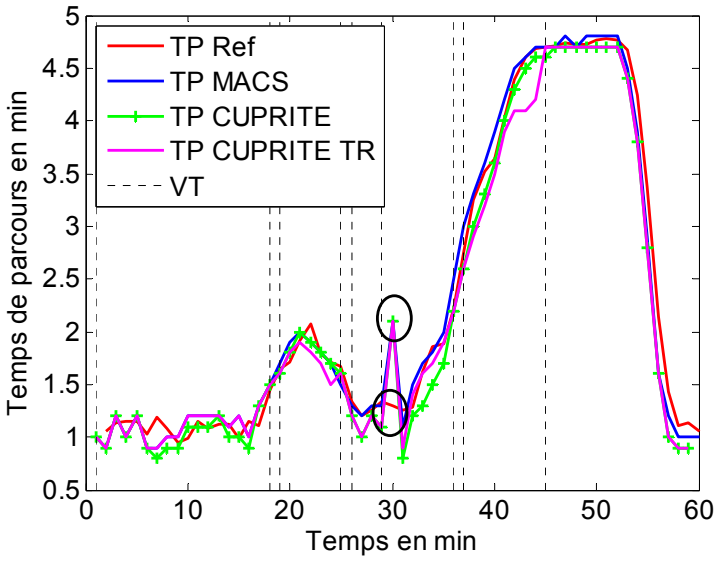


Figure 3. 11 : Effet d'une valeur aberrante de boucle magnétique sur le temps de parcours

3.2.3.2 Méthodes stochastiques

Méthode d'El Faouzi et al.

Description et procédure

[El Faouzi, 2000a], [Lefevre et El Faouzi, 2006] et [El Faouzi, 2006] ont abordé dans leur recherche le problème de l'estimation du temps de parcours sur un axe urbain par des méthodes de classification basées sur la théorie de l'évidence (théorie des fonctions de croyance).

L'objectif de leur travail est de proposer un cadre méthodologique et des solutions au problème d'estimation du temps de parcours en présence de données issues de sources hétérogènes : soit des boucles magnétiques, qui permettent d'obtenir une estimation du temps de parcours moyen ; soit des véhicules traceurs, qui fournissent les temps de parcours réalisés. Ces auteurs ont abordé le problème de l'estimation du temps de parcours comme un problème typique de fusion de classifieurs en utilisant la théorie des croyances pour sa capacité de modélisation de la connaissance.

La théorie de croyance est basée sur les travaux de [Dempster, 1968], et puis a été formalisée par [Shafer, 1976]. Cette théorie constitue une généralisation de l'approche bayésienne. L'idée de base de cette théorie est qu'à partir de l'ensemble des classes appelées « cadre de discernement », on définit sur l'ensemble des parties « d'une fonction de masse ». Une fois ce dernier défini pour chacune des sources en présence, la décision finale est élaborée comme suit :

- d'une part en combinant les masses par une règle propre à cette théorie, appelée règle de combinaison orthogonale de Dempster ;
- d'autre part en adoptant une des multiples règles de décision qu'offre la théorie des croyances.

Dans le cadre de cette théorie, ils ont présenté deux approches afin de modéliser les informations des boucles magnétiques et des véhicules traceurs en fonctions de croyance.

- la première technique repose sur le concept de fusion de classifieurs. Dans ce cas, chaque source d'information est considérée comme un classifieur. Pour chacun de ces

classifieurs, ils ont défini une matrice de confusion qui reflète le pouvoir de discrimination des sources. Cette matrice leur permet par la suite de construire les fonctions de croyance. Les erreurs de chaque classifieur sont usuellement consignées dans la matrice de confusion ;

- la seconde approche emploie une technique classique de classification basée sur un calcul de distance pour construire les fonctions de croyance.

Analyse et critique

Cette méthode est testée par des données issues de la campagne de mesures réalisée à la ZELT (Zone Expérimentale, Laboratoire de Trafic). Elle a montré une certaine efficacité en comparaison à des méthodes d'estimation qui se basent soit sur les véhicules traceurs seuls soit les boucles magnétiques seules. Toutefois, les résultats ne sont pas généralisés et sont spécifiques au site. Dans cette méthode, les auteurs ont pris en compte les incertitudes liées aux différentes informations traitées et ils ont démontré les avantages de la fusion. La théorie des fonctions de croyance propose plusieurs critères de décision et la construction de ces fonctions dépend de l'application envisagée. Pour plus d'informations, le lecteur pourra se reporter à [Denœux, 2000].

Méthode de Choi et Chung

Description et procédure

[Choi et Chung, 2002] se sont intéressés au développement d'un algorithme de fusion de données basé sur les trois techniques : « *voting technique* », « *fuzzy regression* » et « *Bayesian pooling* » pour l'estimation des temps de parcours d'un lien encombré dans un réseau routier urbain.

Voting technique est une méthode de calcul de poids pour les diverses sources d'information [Olkin, 1992].

En analyse par régression traditionnelle, les résidus sont supposés être dus à des erreurs aléatoires. Cependant, ces résidus sont parfois dus à l'imprécision de la structure du modèle ou des observations imprécises. L'incertitude dans ce type de modèle de régression devient floue (*fuzzy*), et non plus aléatoire [Zadeh, 1965]. Des recherches ont tenté d'examiner l'analyse par régression floue et elles peuvent être divisées en deux catégories. La première

est basée sur une approche linéaire [Tanaka et al., 1982], [Tanaka, 1987]. La deuxième est basée sur une approche par moindres carrés [D'Urso et Gastaldi, 2000].

La méthode de *Bayesian pooling* est une méthode qui permet de fusionner et d'agréger plusieurs paramètres après qu'ils sont transformés dans un même espace [Givens et Roback, 1999].

L'algorithme [Choi et Chung, 2002] suit :

- après la collection des données des boucles, la vitesse moyenne est calculée par la formule de Dailey ;
- un temps parcours moyen par intervalle de 5 min est calculé par *voting technique* ;
- un temps de parcours moyen par intervalle de 5 min à partir des véhicules traceurs est calculé simultanément par *fuzzy regression* en considérant l'historique du temps de parcours du lien ;
- finalement les deux temps de parcours sont fusionnés par *bayesian pooling* et le temps de parcours du lien est estimé.

Analyse et critique

La méthodologie est testée en utilisant des données réelles collectées sur quatre mois consécutifs pour des liens urbains, à Suwon. Ils ont utilisé à la fois les données des détecteurs des liens transversaux et aux intersections. Le pourcentage des moyennes absolues des erreurs (*Mean Average Pourcentage Error* MAPE) pour les quatre liens est compris dans la fourchette de 15% à 26,5%. L'algorithme est testé pour les conditions de circulation non saturées et doit être testé pour les conditions de circulation saturées. Pour cette méthode, il faut toujours tester la performance de cette méthode pour d'autres conditions de trafic. En plus, l'équation de Dailey, utilisée dans cet algorithme, suppose une longueur moyenne constante des véhicules. Cette hypothèse n'est pas toujours vérifiée en fonction du temps et en fonction du site.

3.2.4 Synthèse et conclusion

Plusieurs chercheurs ont proposé des méthodes pour estimer le temps de parcours. Le tableau 3.1 résume les méthodes présentées dans ce chapitre. La majorité de la littérature sur l'estimation du temps de parcours est spécifique aux autoroutes et ne peut pas être applicable aux réseaux urbains en raison de la différence de trafic sur les deux types de réseaux. Les complexités liées au réseau urbain comprennent :

- un trafic interrompu (feux de circulation, piétons, arrêts de bus) qui conduit à des discontinuités importantes ;
- un flux significatif provenant des sources et fuites à mi-lien ;
- une différence significative entre le temps de parcours pour les différents mouvements tournants dans un carrefour ;
- la majorité des réseaux urbains sont équipés de détecteurs à la sortie des tronçons ;
- le pourcentage des mouvements tournants est souvent inconnu.

Dans le contexte urbain, la majorité des méthodes d'estimation du temps de parcours basées sur les boucles magnétiques ne considère pas spécialement le flux au niveau des fuites ou des sources à mi-lien. En outre, la performance des méthodes en ce qui concerne l'erreur de comptage de détection n'est pas évaluée. Les méthodes basées sur les données de véhicules traceurs supposent qu'il existe un nombre suffisant de véhicules traceurs. En fait, et surtout dans les pays européens, le nombre requis de véhicules traceurs n'est pas toujours disponible.

Des chercheurs ont également appliqué la technique de fusion des données des détecteurs et des données des véhicules traceurs. Cette technique d'intégration des données multi-sources a le potentiel d'améliorer la précision et la fiabilité des estimations. CUPRITE est une méthode récente qui résout les questions citées avant mais les données des véhicules traceurs y sont supposées d'erreur nulle. De plus et pour estimer le temps de parcours des mouvements tournants, CUPRITE implique qu'il faut connaître *a priori* les pourcentages d'affectations, c'est-à-dire le pourcentage de véhicules qui vont à gauche, à droite, et tout droit. Dans le cas où ces pourcentages sont inconnus, CUPRITE requiert un nombre minimal de véhicules pour estimer correctement le temps de parcours (voir chapitre 8).

Dans le cadre de cette thèse, nous allons tenter d'estimer le temps de parcours dans un milieu urbain en prenant en considération les problèmes soulevés. L'algorithme est basé sur le filtre de Kalman sans parfum. Les observations de ce filtre sont le nombre de véhicules issu des boucles magnétiques et les temps de parcours des véhicules traceurs. En plus, cet algorithme a la capacité d'être appliqué à la fois dans les réseaux urbains et inter-urbains. De plus, il est applicable en temps réel. Avant de détailler cet algorithme, il est essentiel de faire un rappel sur le filtrage de Kalman et ses extensions (voir chapitre 4).

Tableau 3. 1 : Résumé des méthodes d'estimation du temps de parcours

Méthode	Type de réseau	Source des données	Caractéristique / modèle	Résumé
[Nam et Drew, 1999]	Inter-urbain	Boucle	Méthode déterministe basé sur la méthode analytique classique des stocks.	Méthode sous discipline FIFO et en supposant que la loi de conservation de véhicules est respectée.
[Oh et al., 2003]	Inter-urbain	Boucle	Méthode déterministe basé sur la méthode analytique classique des stocks.	Nécessite des boucles sur les sources et fuites à mi-lien.
[Bonvalet et Robin-Prévallée, 1987]	Inter-urbain	Boucle	Méthode déterministe qui n'intègre pas le modèle d'écoulement.	Barycentre entre un état fluide et un état congestionné ne décrivant pas l'état du trafic réel.
[Dailey, 1999]	Inter-urbain	Boucle	Méthode stochastique (filtre de Kalman étendu) et un modèle d'écoulement basé sur et la théorie du trafic.	Reformulation de l'équation de la vitesse utilisée comme équation d'observation qui a besoin d'être calibrée (longueur moyenne des véhicules supposée constante).
[Ye et al., 2006]	Inter-urbain	Boucle	Méthode stochastique (Filtre de Kalman sans parfum) basé le modèle de Dailey.	Suppose que la loi de conservation de véhicule est respectée.
[Hinsbergen et al., 2010]	Urbain	Boucle	Méthode stochastique (Filtre de Kalman Localisé) basé sur le diagramme fondamental.	Le modèle d'observation est le diagramme fondamental. A besoin des observations sur chaque cellule du lien.
[Herring et al, 2010]	Urbain	Véhicules traceurs	Méthode stochastique (chaîne de Markov cachée).	Estime les conditions du trafic des liens urbain supposées cachées. Les conditions des liens sont corrélées et évoluent avec le temps.

[El Faouzi et al., 2000, 2006]	Urbain	Fusion	Méthode stochastique (théorie de croyance).	Complexité de la détermination des poids.
[Choi et Chung, 2002]	Urbain	Fusion	Méthode stochastique (pas de modèle d'écoulement).	Algorithme basé sur voting technique, fuzzy regression et bayesian pooling méthod. Spécifique au site et requiert une grande base de donnée et suppose une longueur moyenne des véhicule constante.
[Bhaskar, 2009]	Urbain	Fusion	Méthode déterministe (pas de considération des erreurs des véhicules traceurs).	Méthode qui corrige la courbe des nombres de véhicules cumulés en connaissant le TP des véhicules traceurs *a posteriori*. Requiert avoir un nombre minimal de véhicules traceurs dans le cas d'un réseau. N'est pas applicable en temps réel. Sensible à l'aléa des véhicules traceurs et à des données aberrantes.

Chapitre 4. Le filtrage de Kalman et ses extensions

SOMMAIRE

4.1 Introduction - 79 -

4.2 Le filtre de Kalman - 79 -

4.3 Le filtre de Kalman étendu - 81 -

4.4 Le filtre de Kalman sans parfum - 82 -
4.4.1 La transformation sans parfum - 83 -
4.4.2 Le filtrage sans parfum - 84 -

4.5 Conclusion - 85 -

4.1 Introduction

Dans le cadre de cette thèse, le temps de parcours et le nombre de véhicules cumulés en aval sont des composantes de l'état d'un modèle markovien. L'état va être estimé en ligne par une technique de filtrage bayésien approché qui fournira en outre une matrice de covariance de l'erreur d'estimation. L'équation d'évolution de l'état est inspirée par la méthode analytique classique des stocks. L'équation obtenue est non-linéaire. On rappellera plus loin les équations du filtrage bayésien dans le cas linéaire, c'est-à-dire le filtrage de Kalman (KF). Dans le cas non-linéaire, il est courant d'approcher le filtrage bayésien par le filtrage de Kalman étendu (EKF), obtenu par linéarisation des équations non-linéaires autour de l'estimation courante de l'état. Il est alors nécessaire que les équations soient différentiables, et il faut expliciter les matrices jacobiennes. Cependant, ces matrices peuvent être parfois difficiles à obtenir. Afin de pallier ces inconvénients, d'autres estimateurs bayésiens approchés tels que le filtre de Kalman sans parfum (Unscented Kalman Filter ou UKF) ont été développés. Le filtre de Kalman sans parfum utilise une technique d'échantillonnage déterministe connu sous le nom de « transformation sans parfum » qui consiste à choisir un ensemble minimal de points d'échantillonnage (appelés sigma-points) autour de l'estimation courante. Ces sigma-points sont ensuite propagés à travers les fonctions non-linéaires. La nouvelle estimation et sa covariance d'erreur sont alors calculées par moyennage sur les images des sigma-points. De plus, cette technique élimine l'obligation de calculer explicitement les matrices jacobiennes, qui pour des fonctions complexes peut s'avérer être une tâche difficile.

4.2 Le filtre de Kalman

Le filtre de Kalman est souvent utilisé dans divers domaines comme celui de l'automatique et du traitement du signal. Par exemple, il est utilisé pour l'estimation de la position d'un objet à partir d'une série d'observations relative à sa position, incluant éventuellement des erreurs de mesures.

Le filtre de Kalman utilise un contrôle en *feedback* : il estime l'état du système puis améliore son résultat avec les informations bruitées fournies par les capteurs.

Le filtre de Kalman (KF) permet d'estimer en ligne l'état d'un système markovien décrit par deux équations, une équation de transition d'état et une équation de mesure linéaires [Maybeck, 1979] et [Bozic, 1994]. Ces deux équations sont respectivement :

$$x_{t+1} = Fx_t + Bu_t + v_t \tag{4.1}$$

$$y_t = Hx_t + w_t \tag{4.2}$$

avec :

- x_t : valeur estimée de l'état à l'instant *t* ;
- y_t : mesure à l'instant *t*;
- u_t : entrée ou commande (éventuellement) ;
- v_t : bruit gaussien indépendant de moyenne nulle et de matrice de covariance *Q* qui modélise l'erreur d'évolution ;
- w_t : bruit gaussien indépendant de moyenne nulle et de matrice de covariance *R* qui modélise l'erreur sur les observations ;
- F : est la matrice de transition connue permettant le passage de l'état x_t à l'état x_{t+1} ;
- H : est la matrice d'observation connue liant l'état à l'observation.

Dans ce qui suit, $\hat{x}_t^-$ désigne le vecteur d'état estimé à l'instant *t* connaissant les observations jusqu'à l'instant *t-1* (*a priori,* prédiction), et $\hat{x}_t$ désigne le vecteur d'état estimé *a posteriori* connaissant les observations jusqu'à l'instant *t* (*a posteriori,* correction). $\hat{p}_t^-$ et $\hat{p}_t$ désignent la covariance de l'erreur respectivement *a priori* et *a posteriori*. L'algorithme du filtre de Kalman est le suivant :

- initialisation :

  Initialisation de l'état du système et de sa matrice de covariance :

$$x_0 = X_0 \text{ et } p_0 = P_0 \tag{4.3}$$

- Prédiction :

$$\hat{x}_t^- = F\hat{x}_{t-1} + Bu_{t-1} \tag{4.4}$$

$$\hat{p}_t^- = F\hat{p}_{t-1}F^T + Q \tag{4.5}$$

- Correction :

Calcul du gain de Kalman :

$$K_t = P_t^- H^T \left(HP_t^- H^T + R \right)^{-1} \tag{4.6}$$

Le gain de Kalman est choisi de manière à minimiser la covariance de l'erreur *a posteriori*.

Réactualisation de l'estimation de l'état :

$$\hat{x}_t = \hat{x}_t^- + K_t \left(y_t - H\hat{x}_t^- \right) \tag{4.7}$$

Mise à jour de la matrice de covariance de l'état :

$$P_t = \left(1 - K_t H\right) P_t^- \tag{4.8}$$

4.3 Le filtre de Kalman étendu

Comme mentionné ci-dessus, le filtre de Kalman peut être appliqué uniquement aux systèmes linéaires pour estimer l'état d'un processus. Toutefois, dans de nombreux cas, le système d'état ou de mesure ou les deux à la fois ne sont pas linéaires. Ainsi, le développement du filtre de Kalman étendu (EKF) [Julier et al., 1995], [Brown et Hwang, 1997], [Julier et Uhlmann, 1997], [Simon, 2006] a pour objectif de rendre le filtre de Kalman applicable aux systèmes non linéaires par linéarisation de la moyenne actuelle et de la covariance. À la place des équations linéaires (4.1) et (4.2), le processus et la mesure sont désormais régis par des équations non linéaires :

$$x_t = f_t \left(x_{t-1} \right) + v_{t-1} \tag{4.9}$$

$$y_t = h_t \left(x_t \right) + n_t \tag{4.10}$$

où f et h sont des fonctions non linéaires. Elles peuvent être appliquées aux équations de prédiction de l'état et de l'observation mais en revanche ne peuvent pas être appliquées pour calculer les matrices de covariance. On définit alors les matrices jacobiennes f_t et h_t :

$$F_t = \left.\frac{\partial f_t}{\partial X}\right|_{x_{t-1/t-1}} \tag{4.11}$$

$$H_t = \left.\frac{\partial h_t}{\partial X}\right|_{x_{t/t}} \tag{4.12}$$

Les équations de

Prédiction :

$$\hat{x}_t^- = f_t\left(\hat{x}_{t-1}\right) \tag{4.13}$$

$$\hat{p}_t^- = F\hat{p}_{t-1}F^T + Q_t \tag{4.14}$$

Correction :

- Calcul du gain de Kalman :

$$K_t = P_t^- H_t^T \left(H_t P_t^- H_t^T + R_t\right)^{-1} \tag{4.15}$$

- Mise à jour de la matrice de covariance de l'état :

$$P_t = \left(1 - K_t H_t\right) P_t^- \tag{4.16}$$

- Réactualisation de l'estimation de l'état :

$$\hat{x}_t = \hat{x}_t^- + K_t\left(y_t - h\hat{x}_t^-\right) \tag{4.17}$$

4.4 Le filtre de Kalman sans parfum

Au lieu de faire une linéarisation des fonctions non linéaires f_t et h_t, le filtrage de Kalman sans parfum [Julier et al., 2000], [Wan et Vander Merwe, 2000], [Wan et Vander Merwe, 2001] utilise la technique de la transformation sans parfum [Julier, 2002], [Julier et Uhlmann, 2002].

Etant donné une variable aléatoire de moyenne et covariance connues et une fonction non-linéaire, on applique une technique d'échantillonnage déterministe autour de la valeur moyenne pour obtenir des échantillons appelés sigma-points, qui sont ensuite transformés par

la fonction non linéaire. Moyenne et covariance de la fonction non linéaire de la variable aléatoire sont ensuite approchées par moyennage sur les sigma-points.

4.4.1 La transformation sans parfum

Ici, nous détaillons une des transformées sans parfum [Julier, 2002], [Julier et Uhlmann, 2002]. Considérons x un variable aléatoire de dimension L, ayant la moyenne $\bar{x}$ et la covariance P, et supposons qu'elle se propage dans une fonction arbitraire non-linéaire f. La transformée sans parfum crée un vecteur de $2L+1$ sigma-points pondérés par w. Etant donné le vecteur d'état à l'instant t, nous calculons une collection de sigma-points, stockés dans les colonnes de la matrice x_t des sigma-points de dimension $L \times (2L+1)$ où L est la dimension du vecteur d'état. Les colonnes sont calculées de la manière suivante :

$$(\chi_t)_0 = \bar{x} \tag{4.18}$$

$$(\chi)_i = \bar{x} + \left(\sqrt{(L+\lambda)P}\right)_i \tag{4.19}$$

$$(\chi)_{i+L} = \bar{x} - \left(\sqrt{(L+\lambda)P}\right)_i \tag{4.20}$$

et pondérés par :

$$w_0 = \frac{\lambda}{(L+\lambda)} \tag{4.21}$$

$$w_i = \frac{1}{2(L+\lambda)} \tag{4.22}$$

$$w_{i+L} = \frac{1}{2(L+\lambda)} \tag{4.23}$$

où $\left(\sqrt{(L+\lambda)P}\right)_i$ est la i^{ème} colonne de la racine carrée de la matrice et elle est obtenue à partir de la matrice triangulaire inférieure par factorisation de Cholesky.

Les vecteurs des sigma-points sont propagés à travers la fonction non linéaire pour produire un ensemble de sigma-points transformés,

$$y_i = f\left((x_t)_i\right), i = 0, ..., 2L \tag{4.24}$$

La moyenne et la covariance sont approchées par une moyenne et une covariance pondérée des sigma-points transformés comme suit :

$$\bar{y} = \sum_{i=0}^{2L} w_i \, y_i \tag{4.25}$$

$$P_y = \sum_{i=0}^{2L} w_i \left(y_i - \bar{y}\right)\left(y_i - \bar{y}\right)^T \tag{4.26}$$

4.4.2 Le filtrage sans parfum

Après la génération des sigma-points $\left(\xi_{t-1,i}, w_{t-1,i}\right)$ autour de la moyenne $\hat{x}_{t-1}$ et de la covariance p_{t-1}, les équations de prédictions suivantes sont utilisées :

$$\hat{x}_t^- = \sum_{i=0}^{2L} w_{t-1,i} f_t\left(\xi_{t-1,i}\right) \tag{4.27}$$

$$p_t^- = \sum_{i=0}^{2L} w_{t-1,i}\left[f_t\left(\xi_{t-1,i}\right) - \hat{x}_t^-\right].\left[f_t\left(\xi_{t-1,i}\right) - x_t^-\right]^T + Q \tag{4.28}$$

On redéfinit des sigma-points $\left(\xi_{t,i}^-, w_{t,i}^-\right)$ autour de la nouvelle moyenne $\hat{x}_t^-$ de la covariance p_t^-. L'étape suivante est d'initier chacun des points de la prévision par un modèle d'observation :

$$\hat{y}_t^- = \sum_{i=0}^{2L} w_{t,i}^- h_t\left(\xi_{t,i}^-\right) \tag{4.29}$$

Les équations de corrections sont :

$$p_{yy} = \sum_{i=0}^{2L} w_{t,i}\left[h_t\left(\xi_{t,i}^-\right) - \hat{y}_t^-\right].\left[h_t\left(\xi_{t,i}^-\right) - \hat{y}_t^-\right]^T + R \tag{4.30}$$

$$p_{xy} = \sum_{i=0}^{2L} w_{t,i}\left[\xi_{t,i}^- - \hat{x}_t^-\right].\left[h_t\left(\xi_{t,i}^-\right) - \hat{y}_t^-\right]^T \tag{4.31}$$

$$K_t = P_{xy} P_{yy}^{-1} \tag{4.32}$$

$$\hat{x}_t = \hat{x}_t^- + K_t\left(y_t - \hat{y}_t^-\right) \tag{4.33}$$

$$p_{xx} = P_{xx}^- - K_t P_{yy} K_t^T \tag{4.34}$$

4.5 Conclusion

Dans ce chapitre le filtrage de Kalman, le filtrage de Kalman étendu et le filtrage de Kalman sans parfum sont définis avec leurs conditions d'application. Dans le chapitre suivant, nous détaillerons l'estimation du temps de parcours par un modèle markovien.

Chapitre 5. Méthode proposée

SOMMAIRE

5.1 Introduction - 89 -

5.2 Choix des capteurs - 90 -

5.3 Type de réseau - 92 -

5.4 Principe de notre approche - 92 -

5.5 Description de la modélisation proposée - 93 -
5.5.1 Introduction - 93 -
5.5.2 Vecteur d'état - 95 -
5.5.3 Modèle d'évolution - 96 -
5.5.4 Erreur d'évolution - 98 -
5.5.5 Modèle d'observation - 99 -
5.5.6 Erreur d'observation - 99 -
5.5.6.1 Principales caractéristiques des données de position - 100 -
5.5.6.2 Principales caractéristiques de la carte numérique - 101 -
5.5.6.3 Map-matching - 103 -
5.5.6.4 Estimation des erreurs - 103 -
5.5.7 L'algorithme - 104 -

5.6 Choix de la méthode d'estimation - 105 -

5.7 Test de performance - 106 -
5.7.1 Erreur - 107 -
5.7.2 Erreur absolue - 107 -
5.7.3 Erreur absolue relative - 107 -
5.7.4 Pourcentage d'erreur absolue moyenne - 107 -
5.7.5 Précision - 107 -

5.1 Introduction

Le temps de parcours est une information indispensable dans la gestion en temps réel du trafic urbain, en particulier en période de congestion et d'accidents, et il constitue une mesure complémentaire des niveaux de saturation sur les réseaux. Cette information commune à tous les modes de déplacement nous permet de comparer des solutions alternatives pour un déplacement sur la base d'un même critère (entre par exemple un point de départ et un point d'arrivée donnés) et de mettre en place des stratégies adaptées : itinéraire, changement de mode de transport…

La majorité des recherches sur l'estimation du temps de parcours sont limitées aux autoroutes et ne peuvent pas être appliquées aux réseaux urbains où l'estimation du temps de parcours est plus difficile, principalement à cause des proportions significatives de flux en provenance de sources et fuites à mi-lien. Ces recherches, qui dépendent de l'instrument de recueil de données, sont souvent applicables sous certaines hypothèses et spécifiques au site de l'étude. En outre, la plupart de ces méthodes ne peuvent pas être appliquées pour une estimation en temps réel, et ne sont pas adaptées aux réseaux avec un pourcentage inconnu de mouvements tournants.

Ce travail de thèse vise à estimer en temps réel le temps de parcours spécialement en milieu urbain. Notre objectif est de fournir une méthode simple à mettre en œuvre, peu coûteuse (qui requiert un minimum de données), applicable dans n'importe quel type de réseau (urbain ou inter-urbain), et finalement qui n'a pas besoin de calibrer ses paramètres. Ce travail pallie les lacunes déjà mentionnées en utilisant une approche fondée sur un filtrage stochastique, le filtrage de Kalman sans parfum (UKF).

Dans un premier temps, nous analyserons les caractéristiques des données que nous avons choisi d'utiliser ainsi que la différence entre les différents types de réseau. Puis nous décrirons la méthode proposée et expliquerons son origine : la méthode analytique classique des stocks adaptée.

L'application de l'algorithme proposé est faite étapes par étapes pour démontrer que cet algorithme peut être appliqué quelles que soient les conditions du réseau et de sa circulation. Les applications sont les suivantes (chapitre 6) :

- application sur un tronçon non perturbé avec un pourcentage connu de mouvements tournants ;

- application sur un tronçon perturbé avec un pourcentage connu de mouvements tournants.

Cette méthode est ensuite étendue pour un réseau (chapitre 7). Les applications de cette méthode étendue sont les suivantes (chapitre 8) :

- application sur un divergent avec un pourcentage inconnu de mouvements tournants et dans un convergent, non perturbés ;

- application sur un réseau typique urbain, perturbé et avec un pourcentage inconnu de mouvements tournants.

5.2 Choix des capteurs

Historiquement, les recherches sur l'estimation du temps de parcours routier ont été la plupart du temps limitées aux autoroutes et se sont appuyées sur les données des dispositifs présents dans l'infrastructure, qui comprennent souvent des détecteurs à boucles inductives, des radars, et des caméras vidéo. Les boucles magnétiques sont les plus anciennes et les plus largement utilisées comme sources de données de trafic. Elles fournissent des informations temporelles sur l'état du trafic, mais seulement à l'endroit où elles sont implémentées. Les chercheurs ont proposé des méthodes qui peuvent être caractérisées selon leur approche : déterministe ou stochastique.

Récemment, les capteurs mobiles sont apparus comme une nouvelle source de données de trafic. Ces capteurs, tels que les « véhicules traceurs » (ou « véhicule sonde »), sont des véhicules équipés de systèmes de localisation (par exemple GPS (*Global Positioning System*) ou réseau téléphonique cellulaire). Ils fournissent des données de trajectoire (les coordonnées de position datées, les vitesses, etc.) et, en conséquence permettent de calculer un temps de parcours réalisé. Les véhicules sondes représentent un échantillon aléatoire de la population des véhicules présents dans le réseau. Par conséquent, la précision de l'estimation du temps de parcours avec des véhicules sondes est liée au nombre de ces derniers. Dans ce contexte, plusieurs recherches se sont attaquées au problème de détermination du nombre minimal de

véhicules traceurs nécessaires pour une estimation du temps de parcours statistiquement significatif.

Les propriétés de ces deux sources de données sont complémentaires et sont résumées dans la figure 5.1.

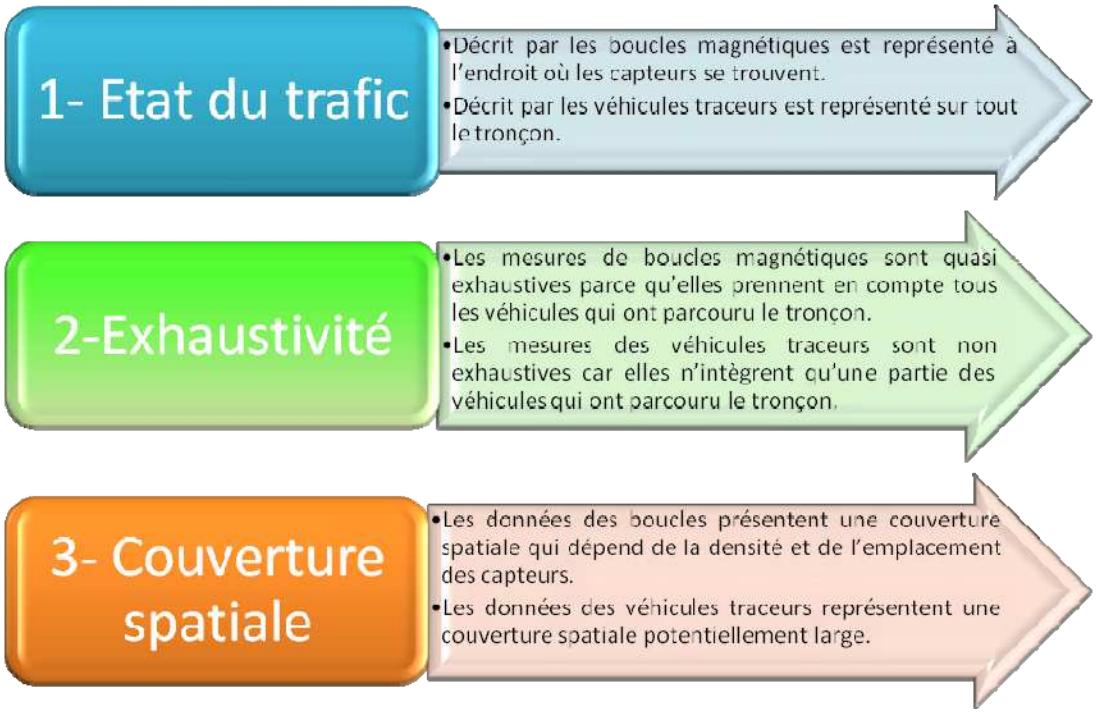


Figure 5. 1 : Propriétés des sources de données

Par conséquent, ces deux sources peuvent être exploitées en développant une solution de fusion de données multi-capteurs pour le problème de l'estimation du temps de parcours, notamment dans les zones urbaines.

Les deux sources de données utilisées dans le cadre de cette thèse sont :

- les boucles magnétiques ;
- les véhicules traceurs.

Les données issues des boucles magnétiques sont le nombre de véhicules et le taux d'occupation (la vitesse si les boucles sont doubles). Nous n'utiliserons que le nombre de véhicules. En ce qui concerne les véhicules traceurs, ceux-ci nous fournissent un temps de parcours individuel et réalisé.

5.3 Type de réseau

Pour estimer le temps de parcours il est important de connaître les caractéristiques du réseau où le temps de parcours va être estimé, d'où l'importance de signaler la différence entre réseaux urbain et inter-urbain. Les différences les plus importants sont résumées dans la figure 5.2.

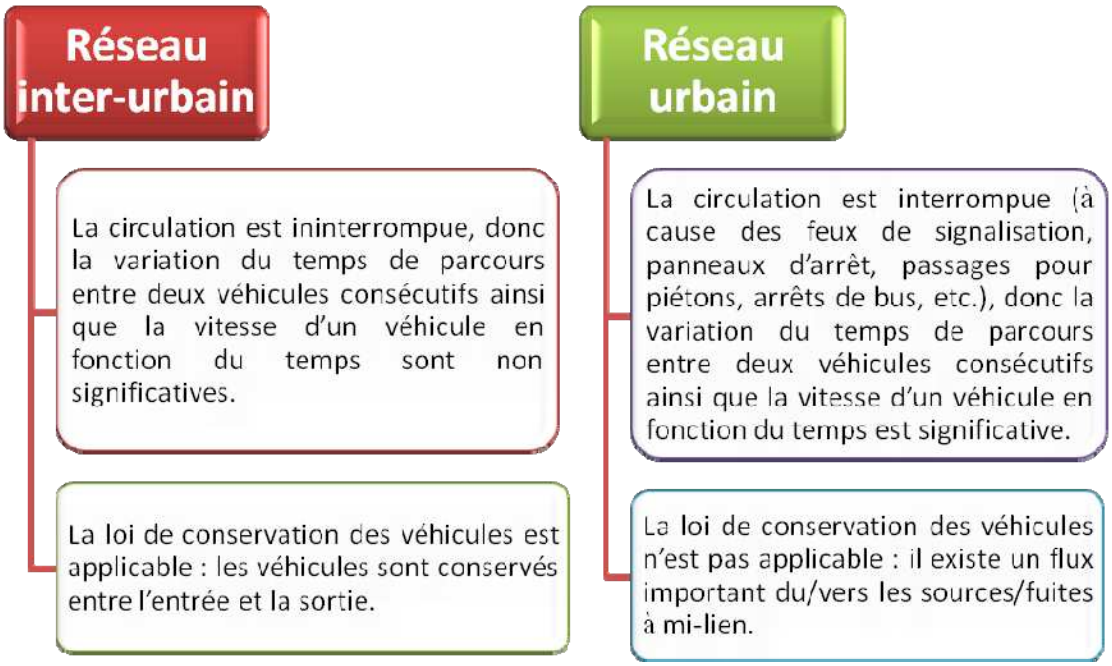


Figure 5. 2 : Comparaison entre réseau inter-urbain et réseau urbain

5.4 Principe de notre approche

Dans le cadre de ce travail de thèse nous avons choisi d'adapter la méthode analytique classique des stocks parce que c'est une méthode simple et rapide en termes de temps de calcul. En effet cette méthode utilise le nombre de véhicules en amont et en aval d'un lien pour estimer le temps de parcours, en respectant la loi de conservation des véhicules. Mais, cette méthode n'est pas applicable dans un réseau urbain qui n'en respecte généralement pas les hypothèses. Dans le cadre de cette thèse, cette méthode qui est déterministe a été adaptée sous la forme d'une méthode stochastique de type filtrage de Kalman. Dans ce contexte, nous devons modéliser l'évolution du nombre de véhicules (et considérer que cette évolution est généralement bruitée) ainsi que l'évolution du temps de parcours des véhicules traceurs (eux

aussi bruités spécialement à cause du processus de *map-matching*). L'adaptation proposée peut être appliquée à un réseau routier.

Le filtre de Kalman est une méthode d'estimation des états d'un système évoluant dans le temps à partir de mesures bruitées. Pour appliquer un filtre de Kalman, il faut avant tout modéliser le système. Si cette modélisation est non linéaire, des extensions du filtre de Kalman existent pour la prise en compte de ces non-linéarités (filtre de Kalman étendu et filtre de Kalman sans parfum) (voir chapitre 4).

L'un des avantages de ce filtre est sa capacité de prédiction des composantes de l'état. De plus, il considère les propriétés statistiques des erreurs, non seulement des capteurs, mais aussi du modèle d'évolution lui-même. Finalement, un des points fort du filtre de Kalman est aussi sa capacité à déterminer la matrice de covariance de l'erreur qui renseigne sur la précision de l'estimation.

5.5 Description de la modélisation proposée

5.5.1 Introduction

Par définition, le temps de parcours d'un tronçon par un véhicule est le temps nécessaire pour parcourir de l'amont à l'aval de ce tronçon. Cette recherche se concentre sur l'estimation du temps de parcours moyen pour tous les véhicules qui sortent en aval pendant un intervalle de temps donné, par exemple par minute.

Dans ce chapitre, nous supposons que chaque tronçon est équipé de boucles magnétiques à sa sortie et que le pourcentage des mouvements tournants aux intersections est connu.

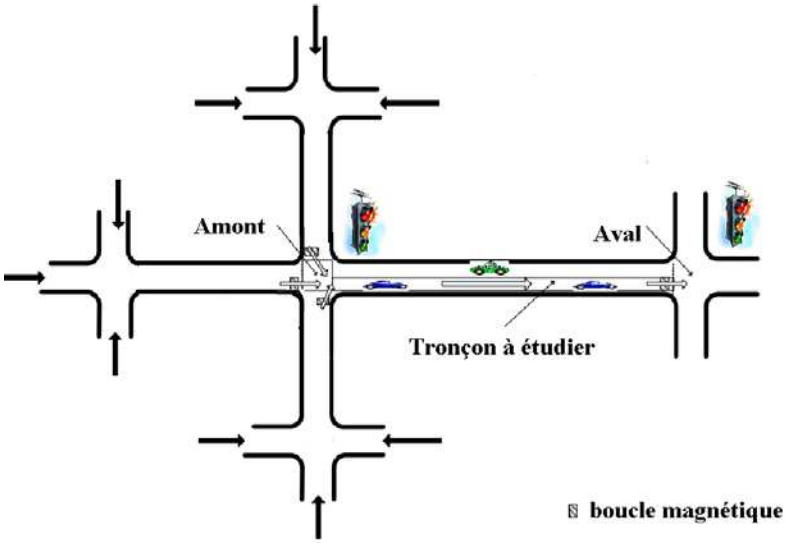


Figure 5. 3 : Tronçon avec un pourcentage connu d'affectation

Comme le pourcentage de mouvements tournants est connu, le nombre de véhicules à l'entrée du tronçon à étudier est égal à la somme du produit des nombres de véhicules en aval des tronçons adjoints par leur proportion respective d'affectation. On connaît donc les flux aux différentes entrées, comme si des boucles y étaient implémentées. En conséquence, ce tronçon peut être représenté plus simplement par la figure 5.4.

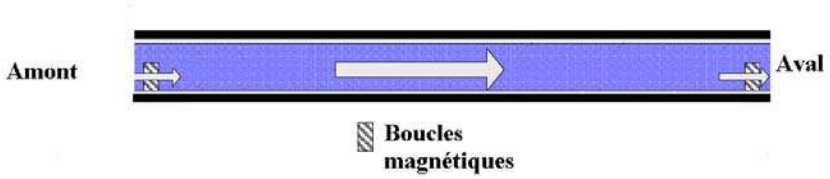


Figure 5. 4 : Tronçon à étudier

Cependant, un tronçon urbain présente souvent des voies transversales à mi-lien non équipées de boucles, telles que des rues secondaires, des parkings, des propriétés privées, etc. agissant comme fuite ou source, ou les deux à la fois. Dans notre modélisation, ces voies sont représentées par une source/fuite à mi-lien comme illustré figure 5.5.

Dans le cadre de ce travail, nous voulons estimer le nombre de véhicules en amont et en aval du lien et le temps de parcours. En plus, nous voulons, s'ils existent, estimer le nombre de véhicules qui entrent/sortent des sources/fuites à mi-lien. Le vecteur d'état, l'évolution, l'observation et l'algorithme du filtrage proposé sont détaillés ci-dessous.

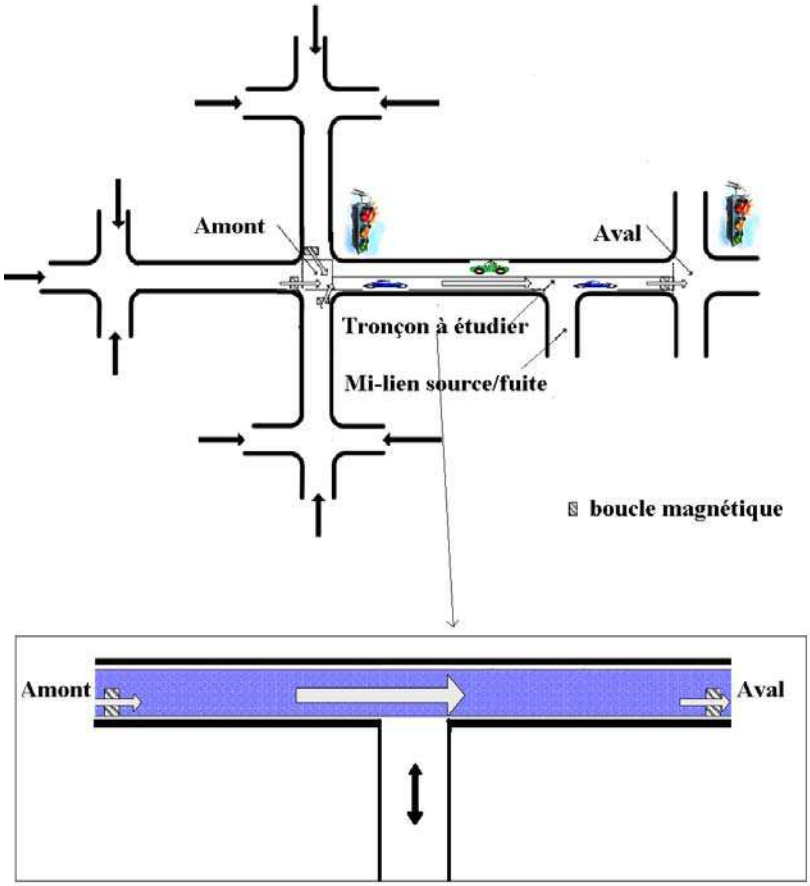


Figure 5. 5 : Tronçon avec une source/fuite à mi-lien

5.5.2 Vecteur d'état

Nous définissons T_S comme étant le temps ou la période d'échantillonnage ou le pas d'estimation.

Le vecteur d'état pour le lien k est défini comme suit :

- TP : le temps de parcours ;
- N_d : le nombre cumulé de véhicules en aval ;
- q_u : le flux (débit) en amont ;
- N_u : le nombre cumulé de véhicules en amont ;

- p : le nombre cumulé de véhicules qui entrent/sortent de/vers les sources/fuites à mi-lien, sachant qu'on compte positivement les véhicules entrant et négativement les véhicules sortant ;

- l'historique du nombre cumulé de véhicules en amont, c'est-à-dire les nombres cumulés N_u aux instants précédents. Cet historique compile un nombre fixe h de ces derniers, ce nombre étant un paramètre *a priori* du modèle, spécifique pour chaque lien.

Par conséquent, le vecteur d'état s'écrit comme suit :

$$x(t) = \begin{bmatrix} TP(t) \\ N_d(t) \\ N_u(t) \\ q_u(t) \\ p(t) \\ N_u(t - T_S) \\ .... \\ N_u(t - h \times T_S) \end{bmatrix} \tag{5.1}$$

5.5.3 Modèle d'évolution

Nous supposons que l'état à l'instant t se déduit de l'état à l'instant t-T_S (généralement avec un temps d'échantillonnage T_S qui correspond aux agrégations des boucles magnétiques). L'évolution de chaque élément du vecteur d'état suit :

- d'après la méthode analytique classique des stocks et à un instant t, le nombre cumulé de véhicules en amont du lien doit être égal au nombre cumulé de véhicules à son aval après un temps de parcours moyen TP. Alors le temps de parcours à l'instant t est égal à la différence entre t et l'instant t_1 en lequel le nombre de véhicules en amont ($N_u(t_1)$) est égal au nombre de véhicules en aval à l'instant t ($N_d(t)$), comme illustré figure 5.6 ;

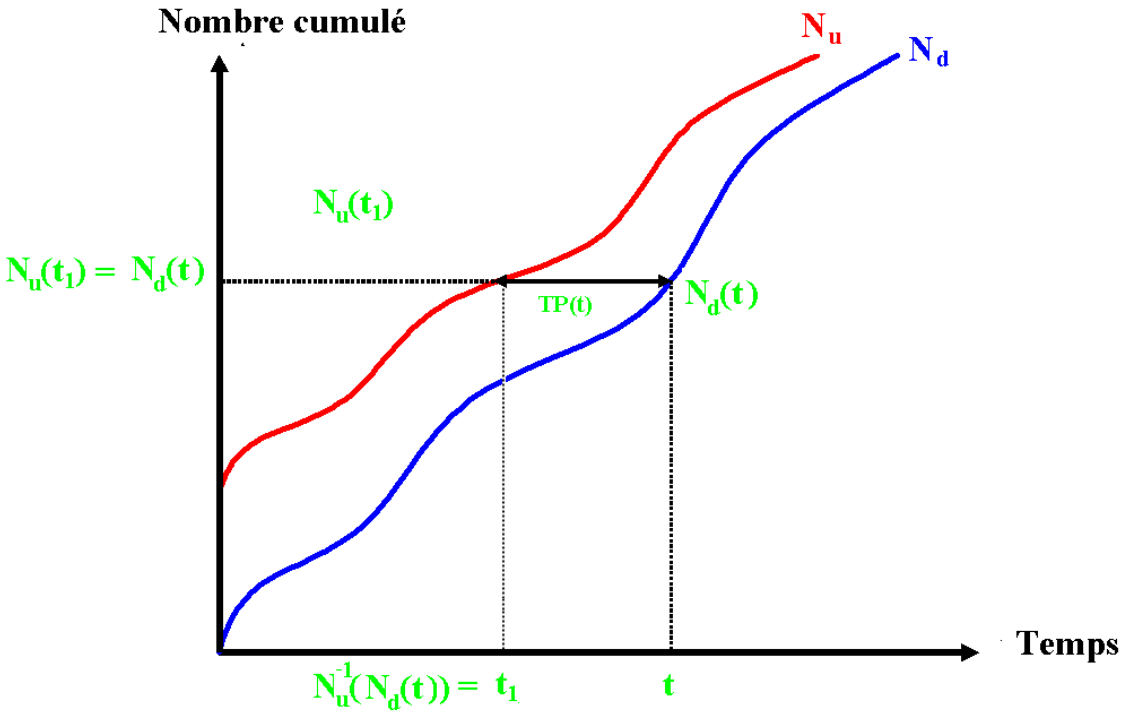


Figure 5. 6 : Calcul du temps de parcours

- le nombre cumulé à l'aval à l'instant *t* est égal au nombre cumulé à l'amont à l'instant *t-TP* (à l'instant $t-T_S$), modulé par le nombre de véhicules issus de sources/fuites à mi-lien. Le nombre cumulé à l'amont à l'instant *t-TP* est tiré de l'historique des nombres cumulés à l'amont ; donc ce qu'on observe à la sortie est le nombre de véhicules à l'entrée *TP* avant, plus la perturbation *(+p*, *-p*, ou *0)* ;

- le nombre cumulé à l'amont à l'instant *t* est le nombre cumulé à l'amont à l'instant $t-T_S$, incrémenté par le flux à l'amont multiplié par T_S ;

- le débit à l'amont à l'instant *t* est supposé constant ;

- le nombre cumulé de véhicules issus des sources/fuites à mi-lien à l'instant *t* est supposé constant. Ce nombre est positif (respectivement négatif) si le nombre de véhicules entrant est supérieur (respectivement inférieur) au nombre de véhicules sortant. S'il n'y a ni fuite ni source (autrement dit : conservation totale des véhicules entre l'entrée et la sortie), ce nombre doit être nul.

Par conséquent, le modèle d'évolution s'écrit comme suit :

$$
\begin{aligned}
&TP(t) = \mathrm{t} - N_u^{-1}\left(N_d(\mathrm{t})\right) \\
&N_d(t) = N_u\left(t - TP\left(t - T_S\right)\right) + p(t) \\
&N_u(t) = N_u(t - Ts) + q_u(t) \times Ts \\
&q_u(t) = q_u(t - Ts) \\
&p(t) = p(t - Ts)
\end{aligned}
\tag{5.2}
$$

5.5.4 Erreur d'évolution

Par la suite, T_S est égal à 1 minute. La matrice de covariance des erreurs d'évolution est de taille ($h+5 \times h+5$). Cette matrice est supposée nulle sauf pour les éléments suivants :

- Pour *TP*, la variance de l'erreur est fixée à 30 s au carré (variation du temps de parcours entre deux minutes consécutives) ;
- Pour q_u, la variance de l'erreur est fixée à 100 véh/h au carré conformément à la variabilité du débit simulé autour de sa valeur de consigne (ceci correspond à la variation du débit entrant dans la simulation et est supposé proportionnel à T_S) ;
- Pour N_u, l'erreur est déduite de l'erreur de q_u par intégration et sa variance est égale au tiers de celle de q_u (i.e. $\propto T_S^3/3$, la variance de l'erreur pour q_u étant $\propto T_S$, avec T_S=1) ;
- La covariance entre q_u et N_u est égale à la moitié de la variance de q_u (i.e. $\propto T_S^2/2$) ;
- La variance de N_d est égale à la somme des variances de p et N_u ;
- Les covariances entre N_u et N_d et entre N_d et p ont été réglées par essais-erreurs ;
- La variance de l'erreur de p est fixée à un demi-véhicule au carré.

$$
Q = \overbrace{\begin{pmatrix}
(1/2)^2 & 0 & 0 & 0 & 0 & 0 & .... & 0 \\
0 & (100/60)^2/3 + 0.5^2 & 0.1^2 & 0 & 0.1^2 & 0 & ... & 0 \\
0 & 0.1^2 & (100/60)^2/3 & (100/60)^2/2 & 0 & 0 & ... & 0 \\
0 & 0 & (100/60)^2/2 & (100/60)^2 & 0 & 0 & ... & 0 \\
0 & 0.1^2 & 0 & 0 & 0.5^2 & 0 & ... & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & ... & 0 \\
... & ... & ... & ... & ... & ... & ... & ... \\
0 & 0 & 0 & 0 & 0 & 0 & ... & 0
\end{pmatrix}}^{h+5}
\tag{5.3}
$$

5.5.5 Modèle d'observation

Les observations sont :

- les temps de parcours des véhicules sondes ;
- les nombres de véhicules comptés par les boucles magnétiques.

Par conséquent, au temps *t* une observation ou une mesure y_t de l'état réel x_t s'écrit de la façon suivante selon les cas :

Cas 1 : le temps de parcours des véhicules sondes qui est l'observation du temps de parcours (*TP*) à l'instant *t* :

$$y_t = H_1 x_t + w_{1t} \tag{5.4}$$

où : H_1 est le modèle d'observation. Dans ce cas, ce modèle est $\begin{bmatrix}1 & 0 & 0 & 0 & \dots & 0\end{bmatrix}$. w_{1t} est le bruit d'observation supposé être un bruit blanc gaussien de moyenne nulle et de covariance R_{1t}.

Cas 2 : une paire de compteurs à boucles, c'est-à-dire une observation des nombres de véhicules cumulés en aval N_d et en amont N_u :

$$y_t = H_2 x_t + w_{2t} \tag{5.5}$$

où : H_2 est le modèle d'observation. Dans ce cas, ce modèle est $\begin{bmatrix}0 & 1 & 0 & 0 & \dots & 0\end{bmatrix}$ ou $\begin{bmatrix}0 & 0 & 1 & 0 & \dots & 0\end{bmatrix}$ si les observations sont respectivement les nombres cumulés de l'aval ou les nombres cumulés de l'amont. w_{2t} est le bruit d'observation supposé être un bruit blanc gaussien de moyenne nulle et de covariance R_{2t}.

5.5.6 Erreur d'observation

Les covariances sont évidemment différentes si l'on considère les compteurs des boucles magnétiques ou le temps de parcours des véhicules traceurs. L'écart type des compteurs sera fixé à 1 véhicule (voir §.2.4.1 : ceci correspond à un taux d'erreur de 5% pour un débit de 20 véhicules par minute, ce qui est représentatif de nos conditions de circulation simulées au chapitre 6). Pour les véhicules sondes, la covariance doit caractériser les erreurs éventuelles

du processus de *map-matching* des positions GPS. L'écart type du temps de parcours des véhicules traceurs est fixé à 3 secondes selon les résultats de l'étude que nous allons présenter dans la suite de ce chapitre, à partir de données réelles et du calcul de l'erreur entre le temps de parcours de référence et le temps de parcours calculé par GPS et *map-matching*. Avant de présenter les résultats, nous expliquons les principales caractéristiques des données de position, les principales caractéristiques de la carte numérique utilisée pour cette étude, le processus de *map-matching* et finalement la répartition de la différence entre le temps de parcours de référence et le temps de parcours estimé par GPS.

5.5.6.1 Principales caractéristiques des données de position

Les principales caractéristiques des données de position sont :

- les solutions GPS standard, fournies à 4 Hz par un récepteur haute-sensibilité Ublox LEA 6T connecté à une antenne patch, sont exactement synchrones aux quarts de seconde UTC (*Coordinated Universal Time*), et présentent des erreurs jusqu'à plusieurs dizaines de mètres (typiques en environnement urbain du fait des problèmes de trajets multiples et de masquage) ;

- la référence, à 10 Hz, est cadencée selon l'horloge propre du système inertiel IXSEA du véhicule expérimental VERT (aussi équipé d'un GPS cinématique et d'un odomètre), voir figure 5.7, et elle n'est donc pas exactement synchrone aux dixièmes de seconde UTC ; sa précision est de l'ordre de grandeur du décimètre.





Figure 5. 7 : Véhicule VERT et sa centrale inertielle

5.5.6.2 Principales caractéristiques de la carte numérique

Les principales caractéristiques de la carte numérique sont :

- elle est extraite (pour Nantes qui nous intéresse) de la Base de données topographique nationale de l'IGN, Institut National de l'Information Géographique et Forestière, et est de précision métrique (figure 5.8). Les rues et les routes sont représentées par des polylignes, c'est-à-dire des séries de segments, sans distinction des voies. Toutefois, quand les chaussées sont physiquement séparées, plusieurs polylignes sont utilisées. Par exemple, le long du busway à Nantes, une polyligne décrit les voies centrales réservées aux transports en commun et deux polylignes décrivent de part et d'autre les voies ouvertes aux autres usagers. On parle de lien, ou d'arc, quand on désigne la série des segments situés entre deux nœuds consécutifs ;

- l'extrait n'est pas aussi dense que la base standard, car nous avons limité l'extraction aux seuls liens proches de la trajectoire suivie, afin de faciliter le *map-matching* et en limiter *a priori* les erreurs (figure 5.9). Ainsi, les erreurs obtenues sont donc plus petites que celles potentiellement obtenues quand tous les liens de la zone urbaine circulée sont considérés ;

- dans chaque lien, les éventuels segments de moins de 50 cm de longueur sont concaténés avec le segment suivant. Ceux-ci sont rares (4 pour 1740 dans l'extrait dense, 3 pour 307 dans l'extrait proche de la trajectoire). On note ainsi qu'aucun lien fait moins de 50 cm de longueur. Les segments ayant donc une longueur d'au moins 50 cm, le passage d'un segment à un autre qui ne lui soit pas connexe doit rarement se produire, en particulier pour la trajectoire de référence à 10 Hz.

Nous gérons deux tables de connexité : une pour les segments (chaque segment connait les segments voisins) et une pour les liens (chaque lien connait les liens voisins).

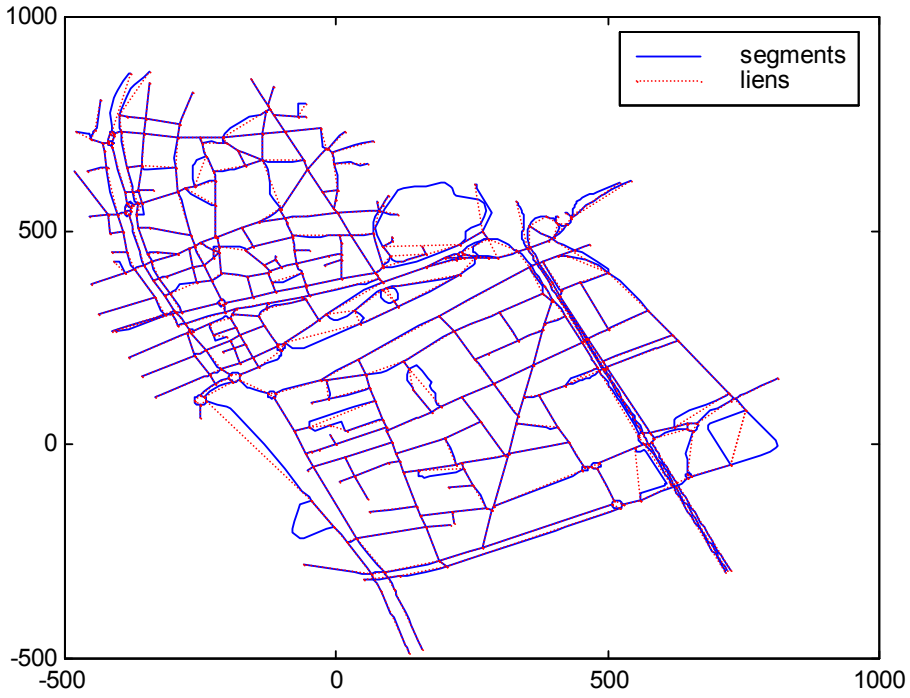


Figure 5. 8 : Extrait de pleine densité (aussi appelé : *map-caché*) : 1736 segment, 630 liens

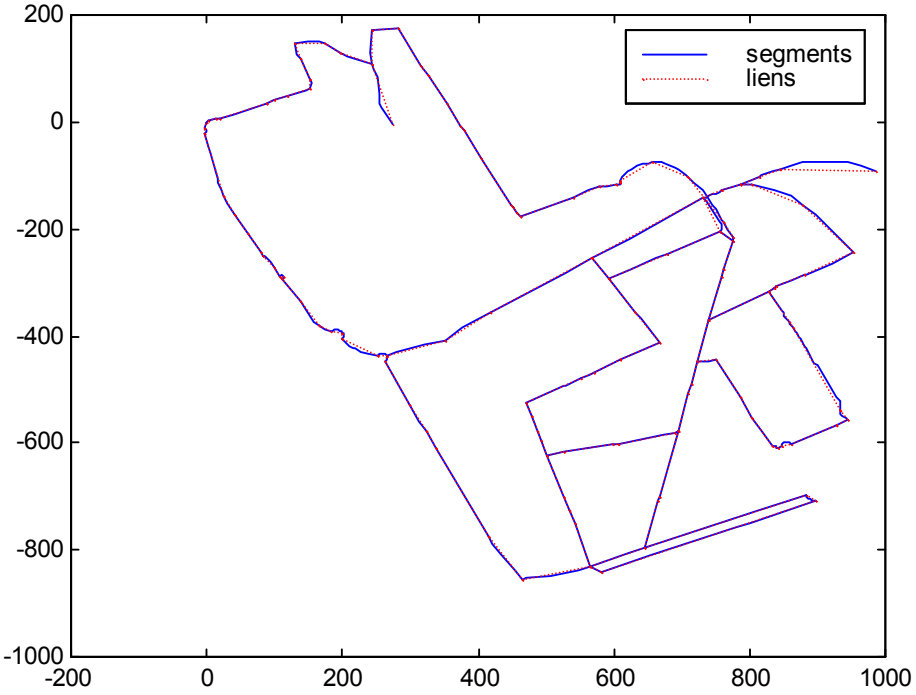


Figure 5. 9 : Extrait limité à la trajectoire pratiquée lors des tests : 305 segments, 131 liens

5.5.6.3 Map-matching

Le processus de *map-matching* recherche d'abord le plus proche segment dans toute la carte (dans tout l'extrait de la carte en fait), à partir de la première solution GPS. Puis cette recherche se limite au segment courant, à ses connexes, aux connexes de ses connexes, etc. jusqu'à un niveau d'exploration des connexes tel que la profondeur d'exploration atteigne 50 mètres[3].

Les transitions d'un lien à un autre sont identifiées, chaque segment ayant un attribut identifiant son lien d'appartenance. Ainsi, deux séries de liens parcourus sont calculées, d'une part pour les solutions GPS standard et d'autre part pour la trajectoire de référence. Les liens en trop dans la série GPS standard (par rapport à la référence) sont supprimés. Les liens manquants sont ajoutés avec un temps de parcours nul. La comparaison des temps de parcours est ensuite possible et elle est montrée dans la figure 5.10 pour tous liens parcourus (244 au total).

5.5.6.4 Estimation des erreurs

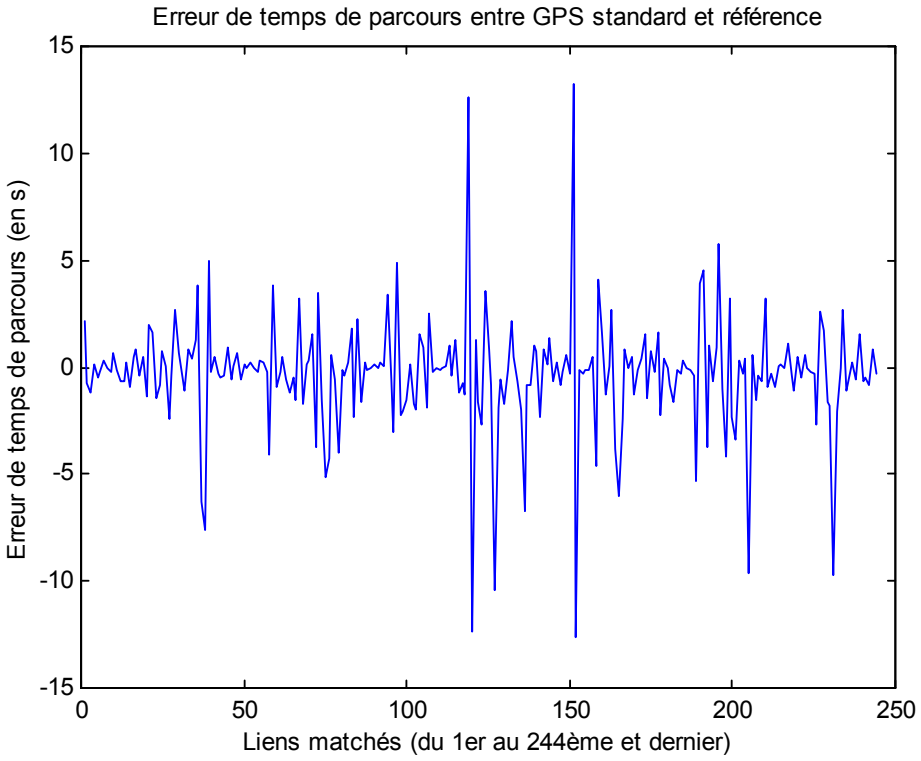


Figure 5. 10 : Erreur de temps de parcours calculé par GPS et temps de parcours référence

[3] Ce calcul est à la fois géométrique (point à segment) et topologique, voir annexe : map-matching.

En moyenne, l'erreur moyenne est proche de zéro (-0.3 s), pour un écart-type de l'erreur de 3 secondes environ (2.7 s). Les erreurs atteignent 13 secondes crête-crête.

5.5.7 L'algorithme

Les données provenant de détecteurs sont agrégés chaque minute, alors que les données des véhicules traceurs sont disponibles à des instants aléatoires entre deux minutes consécutives *t* et *t+1*. Afin d'utiliser chaque information à son instant exact d'occurrence, une étape intermédiaire est faite entre *t* et *t+1*. La figure 5.11 résume l'algorithme proposé et la suite nous permet d'expliquer l'introduction de cette étape intermédiaire.

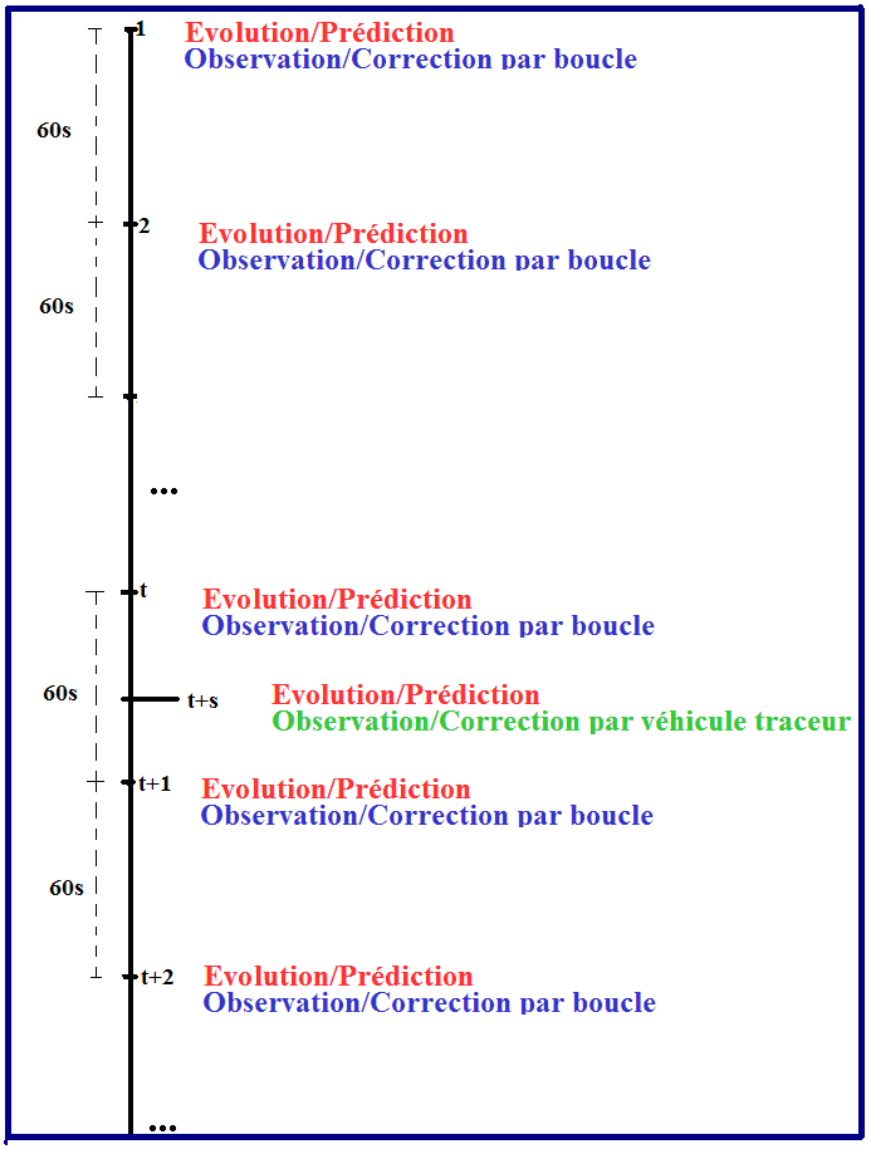


Figure 5. 11 : Algorithme proposé

A chaque pas de calcul, le pas étant égal à une minute, chaque élément du vecteur d'état est prédit par le modèle d'évolution. Comme les observations des boucles magnétiques sont agrégées à la minute, alors à chaque minute l'étape de prédiction est corrigée par les observations de ces boucles. Supposons que l'information des véhicules traceurs soit accessible à l'instant *t* minutes et s secondes. Alors, pour l'estimation en temps réel, nous proposons de réaliser une étape intermédiaire à *t+s* (entre *t* et *t+1* minute) :

- à l'instant *t*, nous appliquons l'étape de prédiction et l'étape de correction par boucle ;
- à l'instant *t+s*, nous appliquons l'étape de prédiction et l'étape de correction par véhicules traceurs ;
- à l'instant *t+1*, nous appliquons l'étape de prédiction et l'étape de correction par boucle.

Et ainsi de suite cette étape intermédiaire est appliquée à chaque fois que nous avons une information qui parvient des véhicules traceurs. Parfois entre deux minutes consécutives, il existe deux ou plus occurrences de véhicules traceurs. Du coup l'étape intermédiaire est applicable à chaque occurrence c'est-à-dire à l'instant *t+s* où le premier véhicule traceur est sorti du tronçon et à l'instant *t+s+s'* où le deuxième véhicule traceur est sorti, etc.

5.6 Choix de la méthode d'estimation

Le filtrage de Kalman est applicable quand l'évolution et l'observation sont linéaires. Dans le cas où au moins l'une des deux fonctions n'est pas linéaire, le filtrage de Kalman étendu est une solution alternative. Mais quand la matrice jacobienne est difficile à exprimer, le filtrage de Kalman sans parfum est la solution la plus adaptée.

Dans l'algorithme proposé, le vecteur d'état dépend de lui-même. En effet, les deux points clés du modèle d'évolution (expliqué §.5.5.2) sont :

- le calcul du temps de parcours requiert de chercher dans l'historique du nombre cumulé de l'entrée l'instant pour lequel le nombre cumulé de sortie à l'instant *t* est égal au nombre cumulé d'entrée ;

- le nombre cumulé de sortie à l'instant *t* est égale au nombre cumulé en aval *TP* avant. Donc, ce nombre cumulé est tiré de l'historique du nombre cumulé en amont en fonction de *TP*. En conséquence le vecteur d'état dépend de lui-même.

Ces deux points clés rendent impossible d'expliciter la matrice jacobienne. Le filtre de Kalman sans parfum (UKF) apporte une solution adéquate. L'algorithme de l'UKF est illustré dans le graphe 5.12.

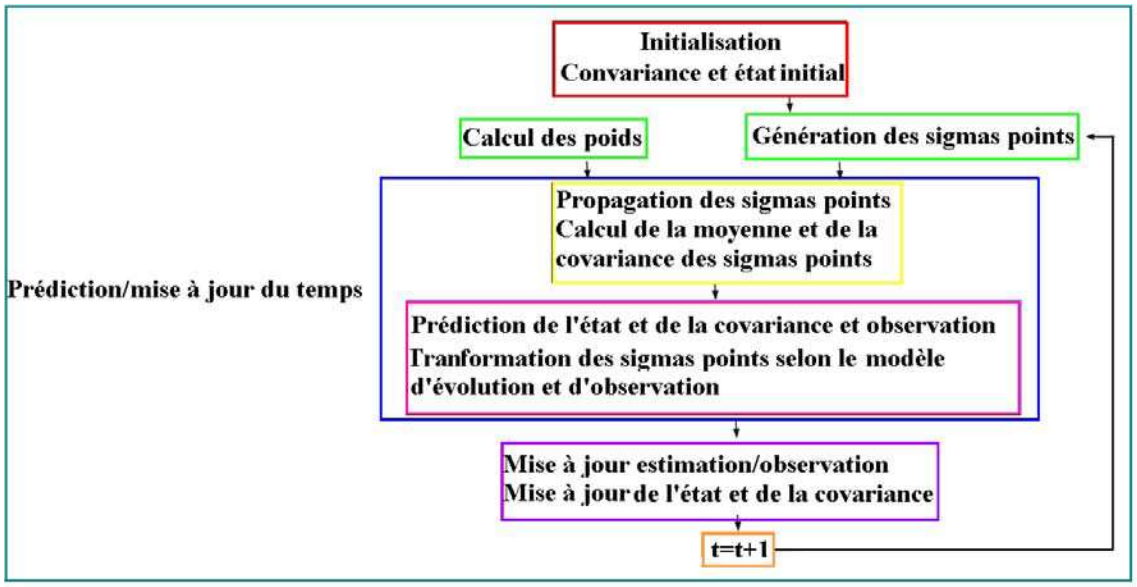


Figure 5. 12 : Algorithme de l'UKF

5.7 Test de performance

Les indicateurs statistiques suivants seront utilisés pour tester la performance de la modélisation et du filtrage proposés ainsi que pour la comparer à celle de la méthode analytique classique des stocks et à celle de CUPRITE. Nous notons :

- *valeur référence*$_i$ le temps de parcours moyen de référence pour l'i^{ème} intervalle d'estimation ;
- *valeur estimée*$_i$ le temps de parcours moyen estimé pour l'i^{ème} intervalle d'estimation ;
- *n* le nombre total d'estimations. Dans nos expérimentations, nous avons choisi 60 (une heure de simulation).

5.7.1 Erreur

C'est la différence entre la valeur estimée et la valeur de référence :

$$erreur_i = valeur\ estimée_i - valeur\ référence_i \tag{5.6}$$

5.7.2 Erreur absolue

C'est la valeur absolue de la différence entre la valeur estimée et la valeur de référence :

$$erreur\ absolue_i = \left| valeur\ estimée_i - valeur\ référence_i \right| \tag{5.7}$$

5.7.3 Erreur absolue relative

C'est la valeur absolue du rapport entre l'erreur et la valeur de référence :

$$erreur\ absolue\ relative_i = \left| \frac{valeur\ estimée_i - valeur\ référence_i}{valeur\ référence_i} \right| \tag{5.8}$$

5.7.4 Pourcentage d'erreur absolue moyenne

Le pourcentage d'erreur absolue moyenne « *mean absolute percentage error MAPE* » est la moyenne des erreurs absolues relatives sur les *n* estimations :

$$MAPE = \frac{\sum_{i=1}^{n} erreur\ absolue\ relative_i}{n} \tag{5.9}$$

5.7.5 Précision

La précision mesure à quel la valeur estimée est proche de la valeur de référence :

$$précision = 1 - MAPE * 100 \tag{5.10}$$

Chapitre 6. Application de la méthode en simulation sur un tronçon sans et avec perturbations

SOMMAIRE

6.1 Introduction - 111 -

6.2 Description du simulateur utilisé - 111 -

6.3 Description des scénarios choisis - 112 -

6.4 Méthodes appliquées - 114 -

6.5 Mise au point de notre algorithme - 115 -

6.6 Résultats et comparaisons - 117 -
6.6.1 Cas d'un tronçon sans perturbations - 117 -
6.6.1.1 Comparaison de l'UKF à la méthode analytique classique des stocks - 118 -
6.6.1.2 Comparaison de l'UKF à CUPRITE - 119 -
Sensibilité de l'UKF et de CUPRITE à l'aléa des véhicules traceurs - 119 -
Comparaison des résultats - 120 -
6.6.1.3 Estimation de la perturbation - 121 -
6.6.1.4 Erreur d'estimation - 122 -
6.6.1.5 Effet des valeurs aberrantes - 123 -
6.6.2 Cas d'un tronçon avec perturbations - 125 -
6.6.2.1 Comparaison des résultats - 125 -
6.6.2.2 Estimation de la perturbation - 129 -

6.7 Analyse et synthèse - 130 -

6.1 Introduction

L'objet de ce chapitre est de mettre en œuvre la méthodologie proposée avec des données simulées. Notre but est de démontrer la capacité de cette méthodologie à estimer le temps de parcours en temps réel avec comptage à l'entrée et à la sortie, ou pourcentage des mouvements tournants aux carrefours connu.

Dans le cadre de ce travail de thèse, les données utilisées sont les résultats des simulations issues du logiciel AIMSUN. Nous intégrerons dans le simulateur des données d'entrée pour avoir une représentation proche de la réalité. La simulation permet de connaître le nombre de véhicules entrés et sortis tronçon par tronçon ainsi que les instants d'entrée et de sortie de chaque véhicule de sorte de construire un temps de parcours de référence. Un nombre aléatoire de véhicules est considéré comme traceurs. A partir de là, l'algorithme de la méthode proposée est implémenté sous MATLAB.

Dans ce chapitre, nous présentons le simulateur AIMSUN et la description des scénarios ; puis les méthodes appliquées (méthode analytique classique des stocks, CUPRITE et notre filtrage UKF). En synthèse et conclusion de ce chapitre, nous comparerons les résultats de l'algorithme proposé à ceux de la méthode analytique des stocks et de CUPRITE.

6.2 Description du simulateur utilisé

Les logiciels de simulation de trafic ont la capacité à émuler la variabilité temporelle des phénomènes de trafic. Ils permettent de saisir la complexité des systèmes de trafic. Il existe une grande variété de logiciels de simulation [Barceló, 2010]. Dans le cadre de cette thèse, l'outil de simulation choisi pour récupérer des données de boucles magnétiques et de véhicules traceurs sur un réseau est AIMSUN. AIMSUN est l'abréviation de « *Advanced Interactive Microscopic Simulator for Urban and Non-Urban Networks* ». AIMSUN est un simulateur de trafic microscopique et la structure de base est la suivante : les véhicules pénètrent sur le réseau aux points d'entrée du réseau et leurs mouvements à travers le réseau sont déterminés par des modèles comportementaux tels que, véhicule isolé, poursuite de véhicule, peloton, changement de voie, etc. AIMSUN applique un modèle stochastique où les véhicules circulent à travers le réseau sur la base de tableaux Origine/Destination et de probabilités de tourner aux carrefours (modèle d'affectation du trafic).

L'entrée du logiciel AIMSUN comprend un scénario et un jeu de paramètres qui définissent l'expérience. Le scénario est composé de quatre types de données : description du réseau, un plan de contrôle du trafic, demande de trafic et type de véhicules en circulation (dans notre cas il n'y a que des voitures). Les paramètres de simulation sont : les valeurs fixes qui décrivent l'expérience comme la durée de simulation, l'intervalle d'échantillonnage des données (dans notre cas elles sont agrégées chaque minute), etc. et des paramètres variables utilisés pour calibrer les modèles tels que les temps de réaction, la zone de changement de voie, etc.

AIMSUN peut fournir en continu une représentation graphique de la circulation du réseau, les données de sortie et des statistiques (débit, vitesse, temps de trajet, retards, arrêts, etc.) et les données recueillies par les détecteurs simulés (densité, occupation, vitesse).

Les résultats d'un modèle de simulation ne sont fiables que lorsque ses paramètres sont correctement calibrés conformément à la représentation du monde réel et aux différents états du trafic. Il est important que les sorties du modèle de simulation soient validées par des données de terrain, pour en confirmer la calibration. Un modèle de simulation correctement calibré a le potentiel pour fournir des données pour des scénarios de trafic différents et peut donc être utilisé en analyse de trafic et à des fins de recherche. L'utilisation de logiciels de simulation pour la recherche est en augmentation.

6.3 Description des scénarios choisis

La simulation est établie sur un tronçon de 600 mètres de longueur. Ce tronçon n'est formé que d'une seule voie, et est équipé de deux détecteurs : le premier est à l'amont du lien (entrée) et le second est à l'aval du lien (sortie). Les données d'entrées sont les suivantes :

- la vitesse libre de circulation est de 36 km/h ;
- la densité maximale est de 140 véh/km ;
- le débit est limité à 1008 véh/h. En effet la densité critique est observée à 1/5 de la densité maximale d'où la valeur du débit (voir figure 2.3). Ce calcul est basé sur le diagramme fondamental applicable dans le centre de Nantes-Métropole pour la gestion du trafic.

Pendant une heure de simulation le flux qui pénètre dans le lien est comme suit : pour le premier quart d'heure le débit moyen est de 500 véh/h, pour le deuxième quart d'heure le débit moyen est de 900 véh/h, pour le troisième quart d'heure le débit moyen va augmenter pour atteindre 1400 véh/h, ce qui est supérieur au débit maximal (1008 véh/h), et pour le dernier quart d'heure le débit moyen va diminuer à 500 véh/h (voir figure 2.8).

Pour chaque véhicule, l'instant de son entrée et l'instant de sa sortie sont enregistrés. Le *TP* instantané réalisé est la différence entre ces instants. Ce *TP* est associé à l'instant de sortie. Nous considérons que le temps de parcours moyen de référence à l'instant *t* est la moyenne des temps de parcours de tous les véhicules qui sont sortis du tronçon entre l'instant *t-1* et *t*. Le calcul du temps de parcours de référence est réalisé à chaque minute.

Les véhicules traceurs sont un échantillon aléatoire de la population, et quatre scénarios ont été préparés :

1. aucun de ces véhicules n'a été considéré comme véhicule sonde ;
2. 1% des véhicules étaient considérés comme des véhicules sondes ;
3. 5% des véhicules étaient considérés comme des véhicules sondes ;
4. 10% des véhicules étaient considérés comme des véhicules sondes.

Les données mises à disposition pour l'estimation du temps de parcours sont :

- le nombre de véhicules à l'amont et à l'aval pour chaque intervalle de temps ;
- le temps de parcours réalisé par les véhicules traceurs correspondant à la sortie du véhicule.

Afin de calculer l'erreur d'estimation, on dispose du temps de parcours moyen de référence (TP réf) à chaque minute.

De plus, deux cas sont considérés :

- les scénarios décrits observent la loi de conservation des véhicules, c'est-à-dire sans fuite ni source de véhicules et tous les véhicules qui sont entrés dans le tronçon sont sortis après leur temps de parcours *TP*. 10 réplications de la simulation sont réalisées ;
- les scénarios décrits présentent des fuites et des sources de véhicules qui perturbent la loi de conservation des véhicules. 10 réplications sont réalisées pour chacune des

conditions suivantes : 1% de fuite, 1% de source, 5% de fuite et finalement 5% de source.

6.4 Méthodes appliquées

Dans le contexte de ce chapitre les trois méthodes suivantes sont appliquées :

- la méthode analytique classique des stocks (MACS) ;
- CUPRITE ;
- l'algorithme proposé (UKF).

Pour les scénarios introduits précédemment, le tronçon est équipé d'une boucle en amont et en aval. Dans le cas des simulations où le tronçon est sans perturbations (sans source ni fuite à mi-lien), la méthode analytique classique des stocks est applicable comme estimateur du temps de parcours. Les nombres de véhicules en amont et en aval sont agrégés chaque minute. Donc à chaque minute t il suffit de rechercher t' pour lequel $N_d(t)$ est égal à $N_u(t')$. Le temps de parcours à l'instant t est égal à la différence entre les instants t et t' (voir figure 6.1).

Dans le cas où le tronçon est perturbé, des véhicules sont comptés en entrée mais pas en sortie s'il y a une fuite et vice versa s'il y a une source. Alors la méthode analytique classique des stocks estime le temps de parcours avec un biais.

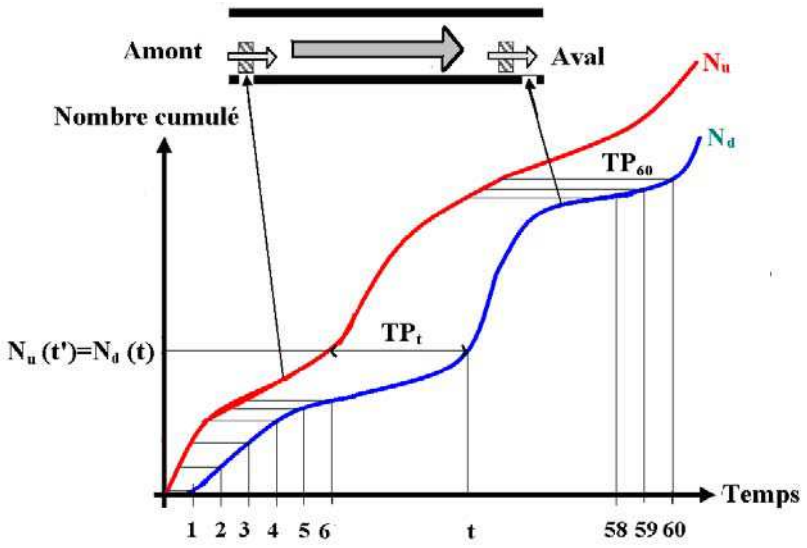


Figure 6. 1 : Estimation du temps de parcours par la méthode analytique classique des stocks

CUPRITE a traité ce problème de biais en corrigeant le nombre cumulé de véhicules en amont de chaque lien du réseau. Ces nombres (ou courbes cumulatives) sont corrigés de façon déterministe par les données provenant des véhicules traceurs. CUPRITE corrige les courbes cumulées en amont à l'instant arrondi à la minute supérieure de l'occurrence du véhicule traceur, ainsi qu'avant et après cet instant. Ceci n'est pas compatible avec la contrainte du temps réel : aussi avons-nous calculé une variante temps réel de CUPRITE (voir § 3.2.3.1).

Les résultats de CUPRITE sont fondamentaux et il est important que l'algorithme que nous proposons lui soit comparé. Ainsi, les résultats de l'estimation du temps de parcours de l'algorithme UKF seront comparés à ceux de CUPRITE et CUPRITE temps réel avec 1%, 5% et 10% de véhicule traceurs.

6.5 Mise au point de notre algorithme

Il est important d'utiliser les temps de parcours des véhicules traceurs à leur instant exact d'occurrence. En effet, deux cas sont ici comparés :

- Cas 1 : dans l'algorithme proposé, le temps de parcours des véhicules traceurs est utilisé comme observation à leurs instants exacts d'occurrence. Donc des pas de calcul intermédiaires sont établis au cours de l'estimation. Le temps de parcours estimé par cet algorithme est noté « TP UKF » ;

- Cas 2 : dans ce cas, le temps de parcours des véhicules est utilisé à la minute entière supérieure, comme dans CUPRITE. Dans ce cas là, nous n'établirons pas d'étape intermédiaire. Par exemple, si un véhicule traceur est sorti à l'instant 37.2 minute, son temps de parcours est utilisé comme observation à la minute 38. Le temps de parcours estimé par cet algorithme est noté « TP' UKF ».

La figure 6.2 (a), (b) et (c) représente TP Ref, TP UKF et TP' UKF avec respectivement 1%, 5% et 10% de véhicules traceurs.

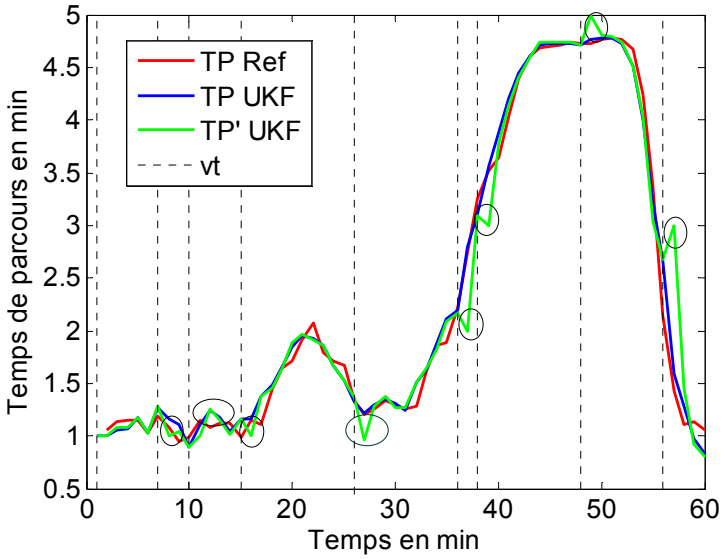


Figure 6. 2 (a) : TP UKF vs TP' UKF avec 1% de véhicule traceur

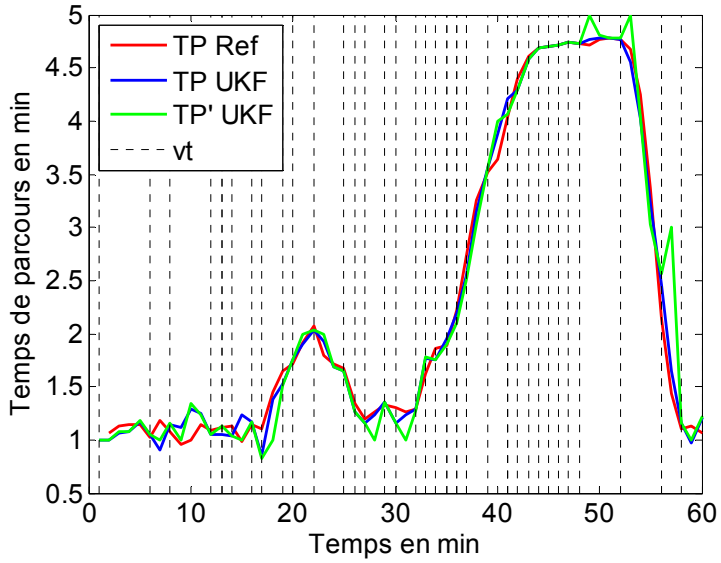


Figure 6.2 (b) : TP UKF vs TP' UKF avec 5% de véhicule traceur.

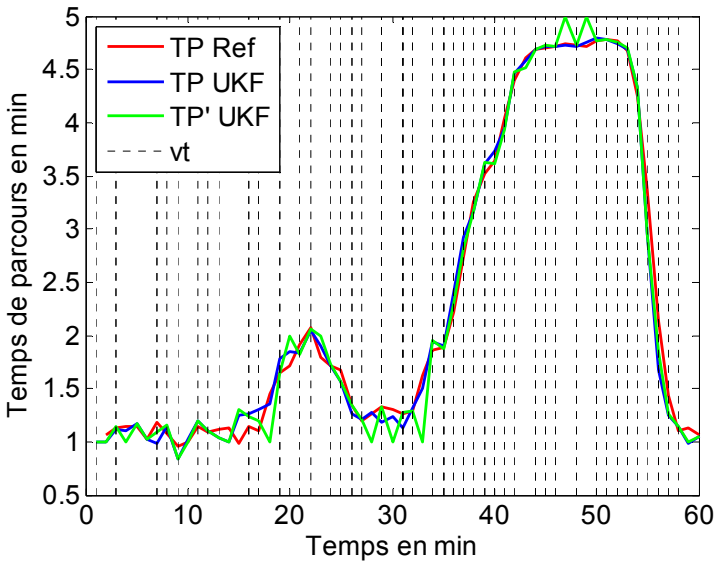


Figure 6.2 (c) : TP UKF vs TP' UKF avec 10% de véhicule traceur.

Dans la courbe de TP' UKF, le temps de parcours corrigé à la minute *t* par le véhicule traceur, entraîne un temps de parcours bruité à la minute *t+1* (encerclé dans la figure 6.5 (a)) ce qui n'est pas le cas dans TP UKF. En effet dans TP' UKF les observations des véhicules traceurs corrigent le temps de parcours à un instant incorrect ce qui entraîne un bruit pour l'instant d'après. Ceci met en évidence l'importance d'utiliser les observations à leurs instants exacts.

6.6 Résultats et comparaisons

6.6.1 Cas d'un tronçon sans perturbations

10 réplications de la simulation sur le tronçon sans perturbations ont été établies. Les résultats de l'algorithme proposé sont comparés à ceux de la méthode analytique classique des stocks et à ceux de CUPRITE. De plus, nous montrerons que la perturbation (ici nulle) est bien estimée. Finalement, nous prouverons la capacité de cet algorithme à rejeter des valeurs aberrantes.

6.6.1.1 Comparaison de l'UKF à la méthode analytique classique des stocks

La figure 6.3 est un graphe superposant le temps de parcours de référence (TP Ref), le temps de parcours estimé par la méthode analytique classique des stocks (TP MACS) et finalement le temps de parcours estimé par l'UKF (TP UKF), pour une des réplications. Les trois courbes sont proches. Nous mettons ici en évidence que, sous la loi de conservation des véhicules et sans erreurs de comptage des boucles, MACS peut estimer correctement le temps de parcours, tout comme l'algorithme proposé, et ce sans véhicules traceurs.

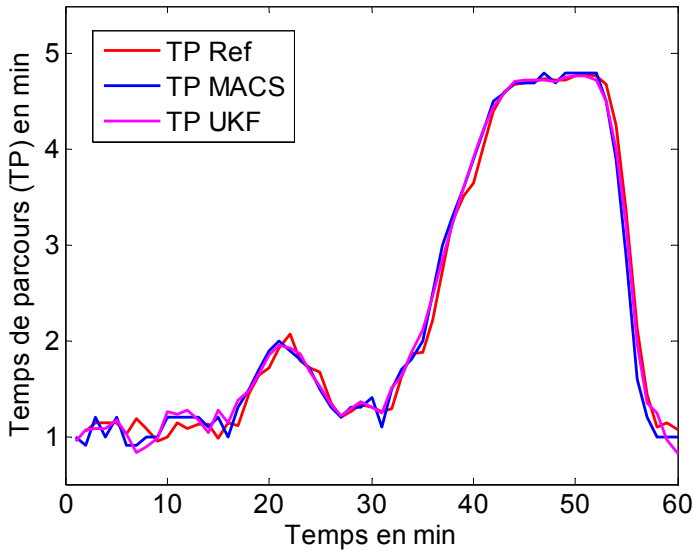
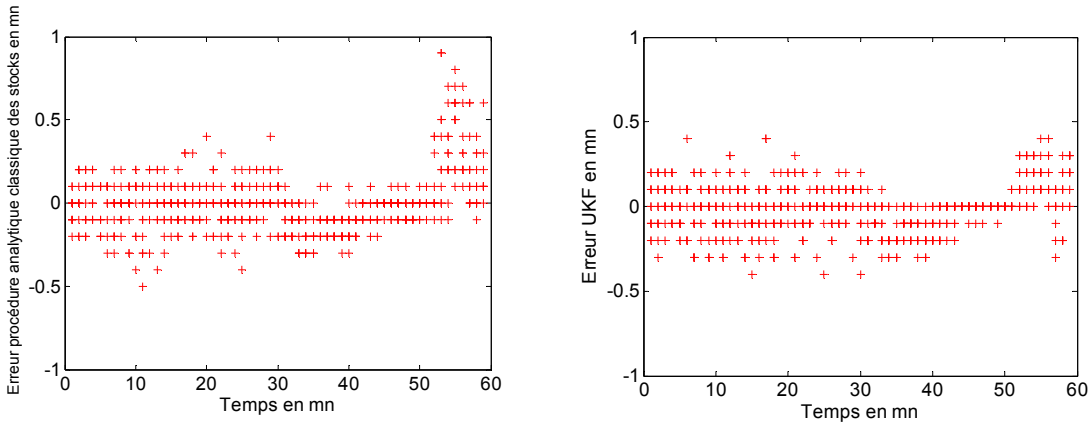


Figure 6. 3 : Comparaison entre les temps de parcours estimés par la méthode analytique classique des stocks et par UKF

La figure 6.4 (a) représente la différence entre le temps de parcours de référence et le temps de parcours estimé par MACS et (b) celle de la différence entre le temps de parcours référence et le temps de parcours estimé par UKF. L'intervalle d'erreur de l'UKF est [-0,4 – 0,4] et de MACS est [-0,5 – 0,9].



(a) (b)

Figure 6. 4 : (a) L'erreur de la méthode analytique classique des stocks en minute. (b) L'erreur de l'UKF en minute

La précision moyenne de l'UKF est 93% avec un minimum de 92% et un maximum de 95% et celle de MACS est aussi de 93% avec un minimum de 91% et un maximum de 94%, sur les 10 réplications.

6.6.1.2 Comparaison de l'UKF à CUPRITE

Sensibilité de l'UKF et de CUPRITE à l'aléa des véhicules traceurs

Nous avons démontré (voir §.3.2.3.1) que si la différence entre le temps de parcours du véhicule traceur (tiré aléatoirement) et le temps de parcours de référence est significative, alors le temps de parcours estimé par CUPRITE (et CUPRITE temps réel) est biaisé. Cette différence a un effet très atténué avec l'algorithme proposé. La figure 6.5 superpose le temps de parcours de référence, le temps de parcours CUPRITE, le temps de parcours CUPRITE temps réel, et le temps de parcours UKF. Il n'y a qu'un véhicule traceur. La différence entre le temps de parcours de ce véhicule et le temps de parcours de référence est de 0,4 min. Avec l'UKF, cet écart a principalement un effet à l'instant de correction par ce véhicule sonde. Après cette correction, l'évolution et la correction par les détecteurs permettra d'atténuer l'effet précédent, alors que CUPRITE (temps réel on non) reste biaisé tant qu'aucun nouveau véhicule sonde ne passe. Nous remarquons sur le graphe l'écart entre TP CUPRITE et TP Ref.

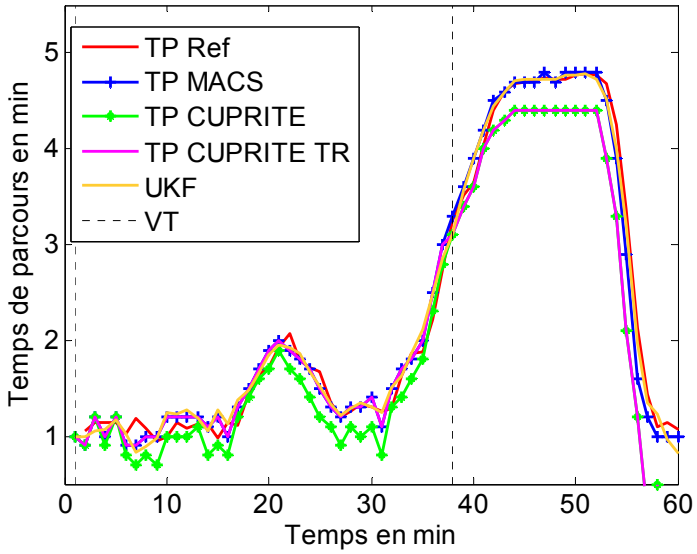


Figure 6. 5 : Sensibilité du temps de parcours estimé avec l'aléa des véhicules traceurs

Comparaison des résultats

Le temps de parcours estimé par l'algorithme proposé est comparé à celui estimé par CUPRITE et CUPRITE temps réel avec 1%, 5% et 10% de véhicules traceurs. Les figures 6.6, 6.7 et 6.8 illustrent la différence entre le temps de parcours de référence et respectivement le temps de parcours CUPRITE, le temps de parcours CUPRITE temps réel et le temps de parcours UKF pour les 10 réplications établies.

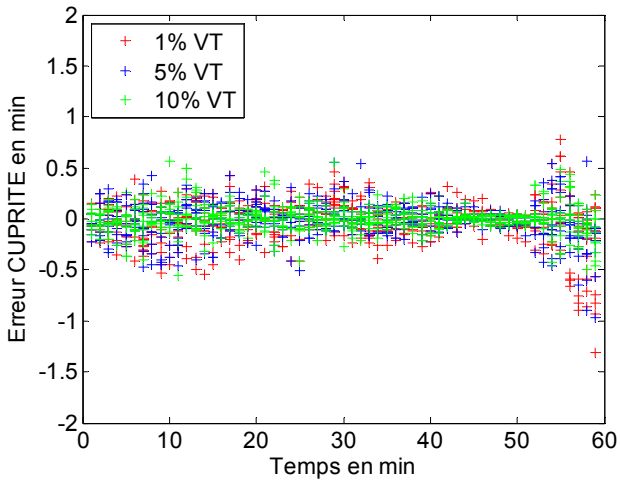


Figure 6. 6 : Erreur commise par CUPRITE en minute

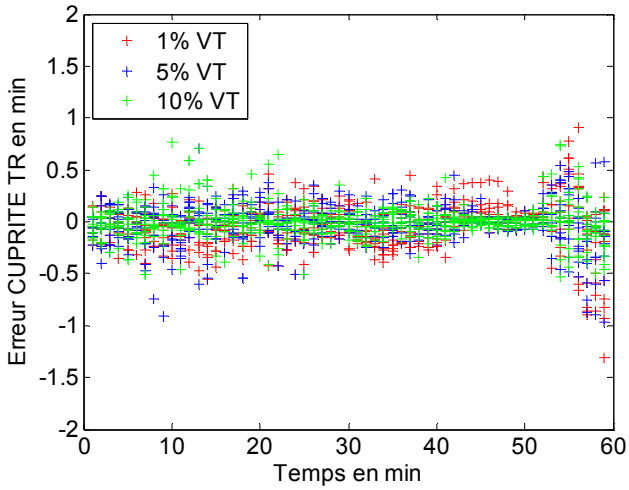


Figure 6. 7 : Erreur commise par CUPRITE temps réel en minute

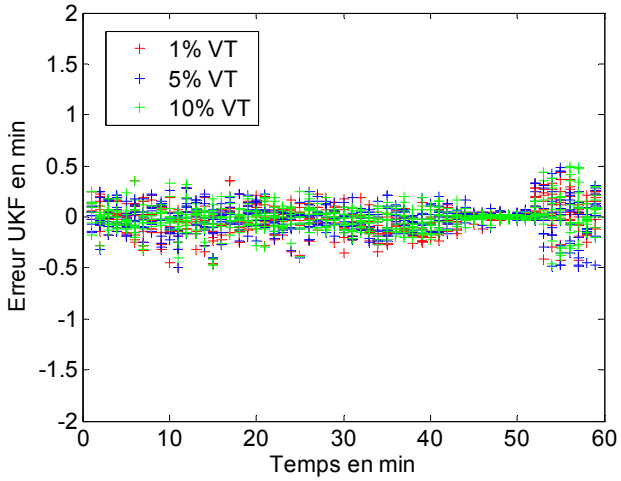


Figure 6. 8 : Erreur commise par UKF en minute

D'après les graphes ci-dessus, nous remarquons que le temps de parcours UKF est le moins bruité. L'écart type de l'UKF et de CUPRITE est respectivement 0,1 et 0,2 avec une moyenne nulle.

6.6.1.3 Estimation de la perturbation

L'une des composantes du vecteur d'état est la perturbation, que l'algorithme proposé estime. La figure 6.4 illustre l'estimation de la perturbation pour une des réplications avec 0%, 1%, 5% et 10% de véhicules traceurs. Ici, cette perturbation n'est qu'un bruit dont la moyenne est proche de zéro et l'écart type proche de 0,1. On retrouve donc ici que la loi de

conservation des véhicules ne subit pas de perturbation, ce que confirme le fonctionnement de l'algorithme proposé.

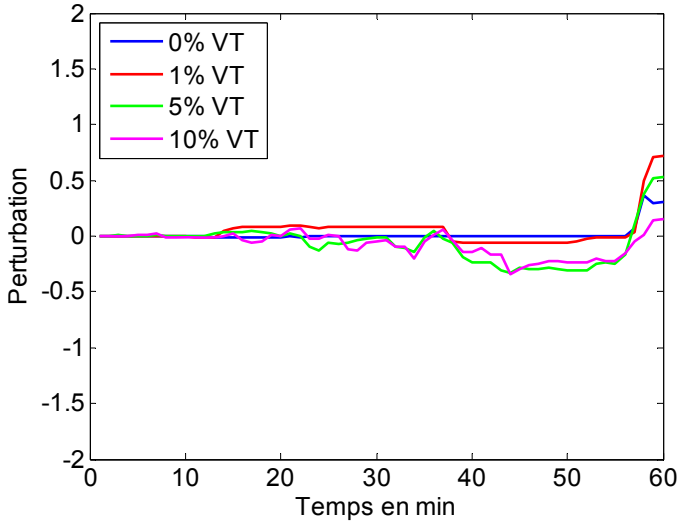


Figure 6. 9 : Estimation de la perturbation

6.6.1.4 Erreur d'estimation

Un des avantages du filtre de Kalman est la possibilité d'estimer l'écart type de chacune des composantes du vecteur d'état. Dans la figure 6.10 les graphes rouges représentent l'enveloppe à deux écarts types estimés par l'UKF et les graphes bleus représentent l'erreur, c'est-à-dire la différence entre la valeur estimée par l'UKF et la valeur de référence des composantes du vecteur d'état (TP, N_d et N_u). Les courbes rouges enveloppent bien la courbe bleue pour les trois composantes.

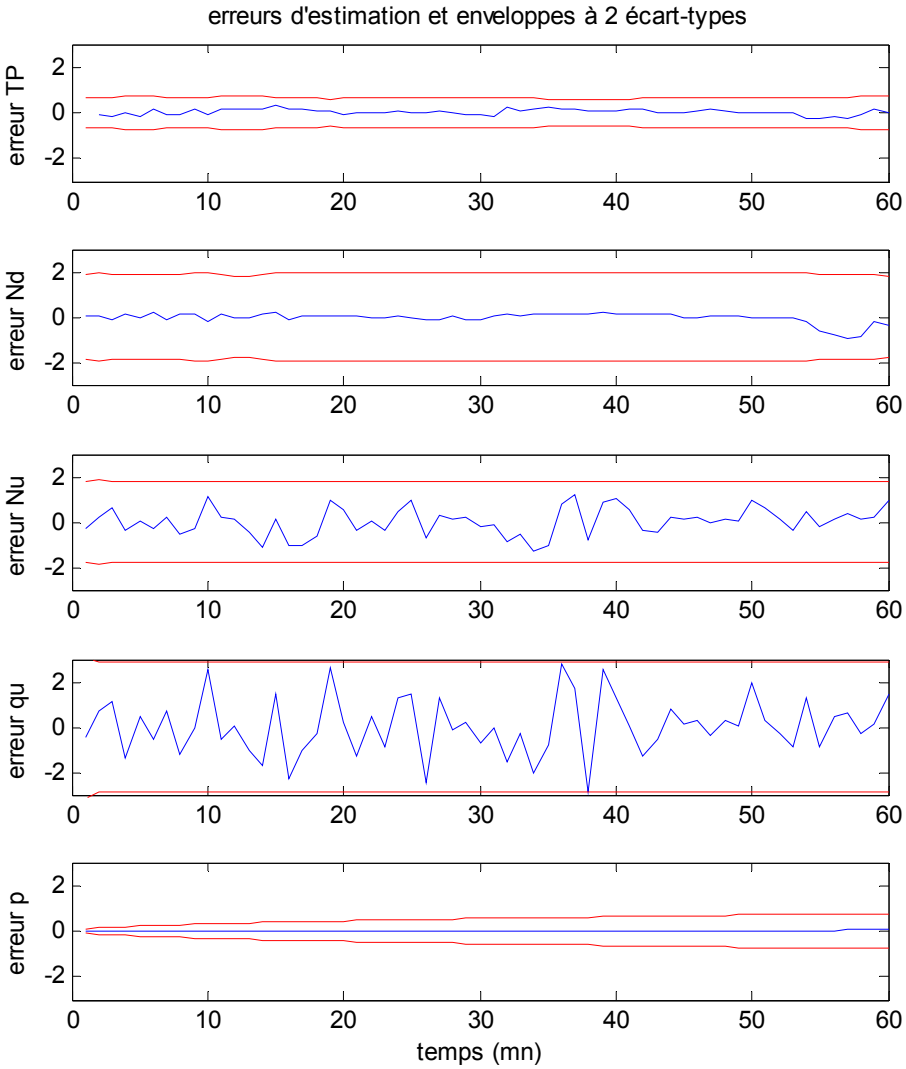


Figure 6. 10 : Enveloppes d'erreurs

6.6.1.5 Effet des valeurs aberrantes

La figure 6.11 illustre l'effet d'une valeur aberrante de véhicule sonde (2,8 min au lieu de 1,1 min) sur l'estimation du temps de parcours, tandis que la figure 6.12 illustre l'effet d'une valeur aberrante du détecteur, où la valeur comptée en aval a été fixée pendant 2 minutes consécutives. Au moyen d'un test du χ^2 de l'innovation carrée normalisée (aussi appelée la distance de Mahalanobis), l'UKF a rejeté la valeur aberrante, alors que MACS et CUPRITE (ainsi que CUPRITE temps réel) ne pouvaient pas (voir §.2.3.3.3, paragraphe : effet des

valeurs aberrantes et §.3.2.3.1, paragraphe : application de CUPRITE) conduisant à une estimation du temps de parcours aberrante.

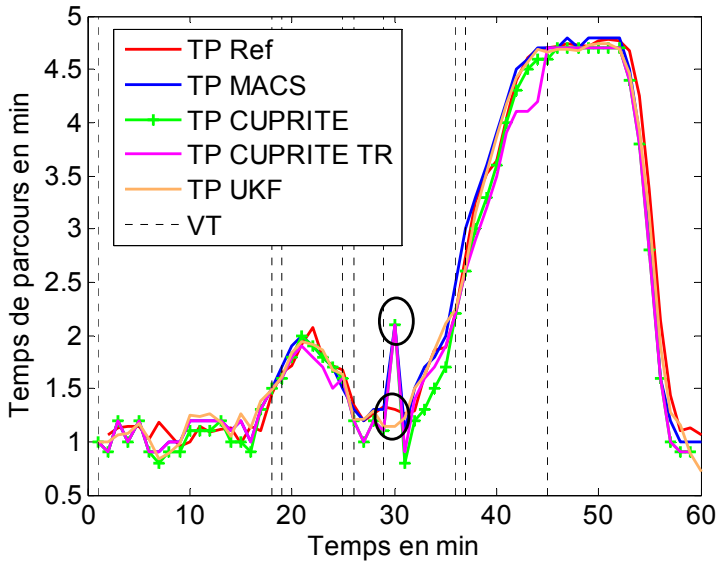


Figure 6. 11 : Effet d'une valeur aberrante de boucle magnétique sur le temps de parcours

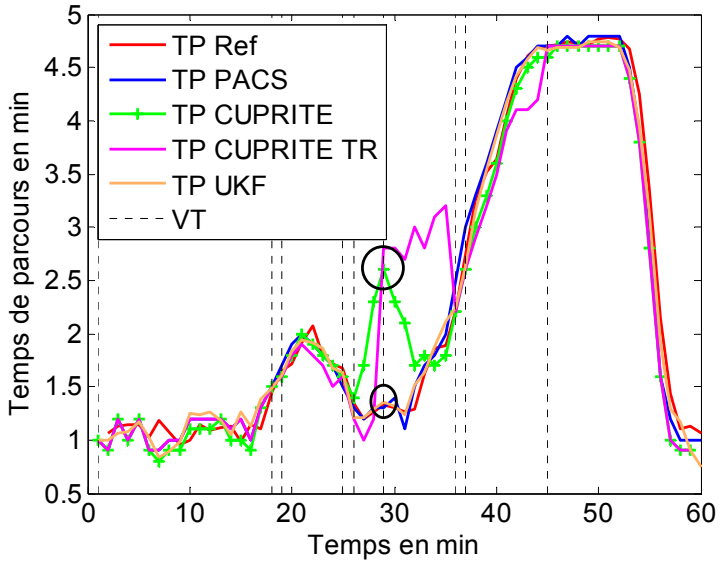


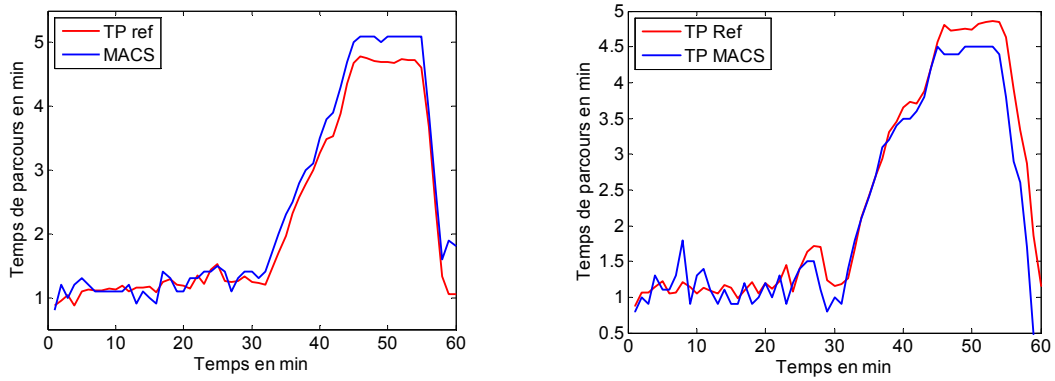
Figure 6. 12 : Effet d'une valeur aberrante de véhicule traceur sur le temps de parcours

6.6.2 Cas d'un tronçon avec perturbations

10 réplications de la simulation sur le tronçon avec 1%, 5% de fuite et 1%, 5% de source ont été réalisées (soit 40 au total). Dans ces conditions, nous comparerons aussi l'UKF à la méthode analytique classique des stocks et à CUPRITE. L'importance ici est de montrer la capacité de l'algorithme proposé à estimer les nombres cumulés de véhicules qui entrent/sortent à travers les sources/fuites à mi-lien sans les observer.

6.6.2.1 Comparaison des résultats

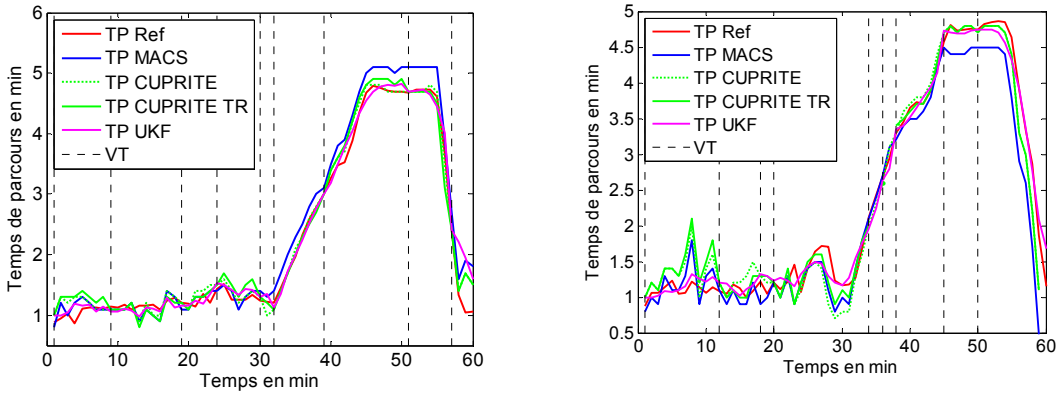
La figure 6.13 (a) et (b) illustrent le temps de parcours de référence et le temps de parcours estimé par la méthode analytique classique des stocks dans respectivement le cas d'une fuite et d'une source. Dans le cas d'une fuite, il existe des véhicules qui sont entrés en amont du tronçon mais ne sont pas sorti à l'aval. Alors, il est évident que la méthode analytique classique des stocks surestime le temps de parcours. Inversement dans le cas d'une source, il existe des véhicules comptés en sortie mais pas en entrée, et MACS sous-estime.



(a) (b)

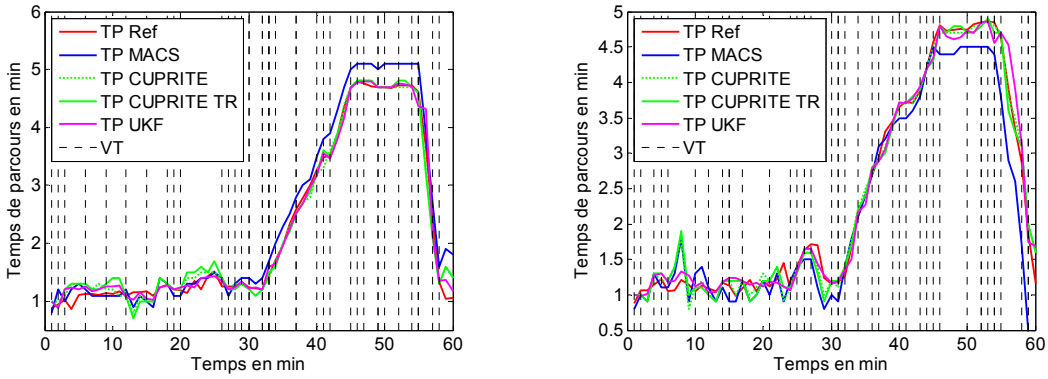
Figure 6. 13 : Estimation du temps de parcours par MACS (a) fuite, (b) source

Les figures 6.14, 6.15 et 6.16 représentent les temps de parcours estimés par la méthode analytique classique des stocks, CUPRITE, CUPRITE temps réel et l'algorithme proposé superposés au temps de parcours de référence pour respectivement 1%, 5% et 10% de véhicules traceurs pour les deux cas fuites et sources. Même avec 1% de véhicules traceurs, l'algorithme proposé, CUPRITE et CUPRITE temps réel corrigent cet écart et estiment correctement le temps de parcours.



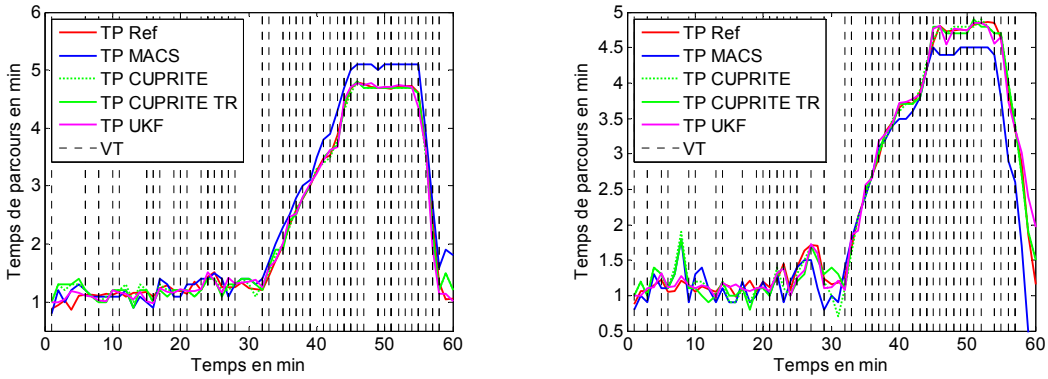
(a) (b)

Figure 6. 14 : Estimation du temps de parcours par MACS, UKF, CUPRITE et CUPRITE temps réel avec 1% de véhicules traceurs (a) fuite, (b) source



(a) (b)

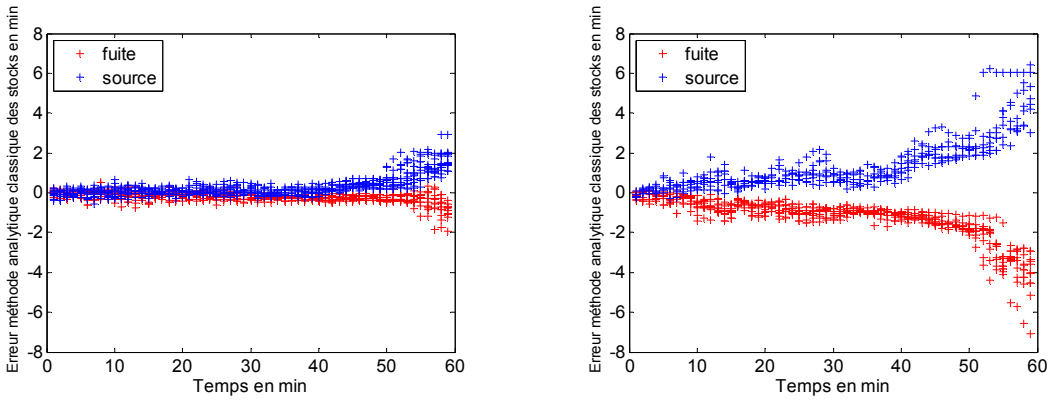
Figure 6. 15 : Estimation du temps de parcours par MACS, UKF, CUPRITE et CUPRITE temps réel avec 5% de véhicules traceurs (a) fuite, (b) source



(a) (b)

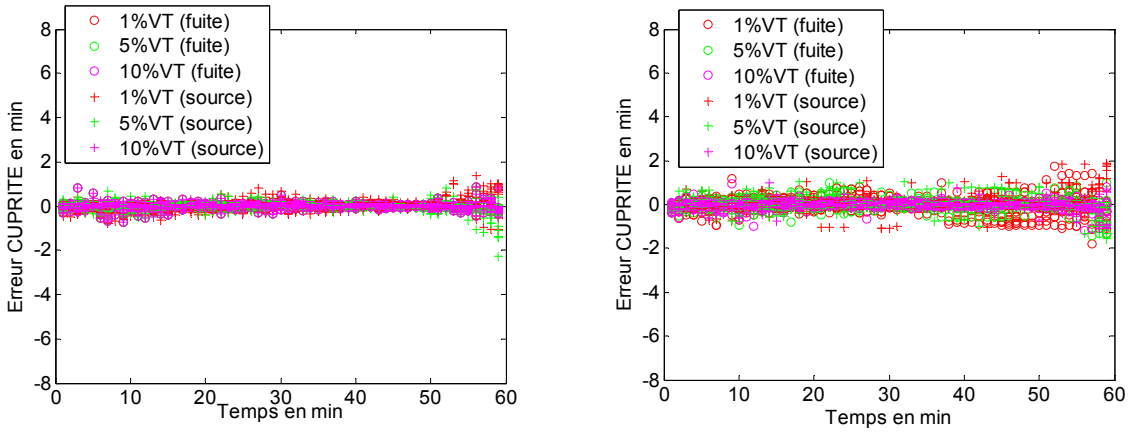
Figure 6. 16 : Estimation du temps de parcours par MACS, UKF, CUPRITE et CUPRITE temps réel avec 10% de véhicules traceurs (a) fuite, (b) source

Les figures 6.17 (a) et (b) représentent respectivement l'erreur commise par la méthode analytique classique des stocks en minute pour 1% et 5% de source/fuite à mi-lien. Nous remarquons l'effet de la perturbation qui s'amplifie par cumul en particulier sur les 10 dernières minutes. Les figures 6.18 (a) et (b) représentent respectivement l'erreur de CUPRITE en minute pour 1% et 5% de source/fuite à mi-lien et les figures 6.19 (a) et (b) représentent respectivement l'erreur de CUPRITE temps réel en minute pour 1% et 5% de source/fuite à mi-lien. Elles montrent que pour certaines réplications, l'erreur atteint ± 2 min. Ici, nous observons à nouveau la sensibilité de CUPRITE. Les figures 6.20 (a) et (b) représentent respectivement l'erreur de l'UKF en minute pour 1% et 5% de source/fuite à mi-lien. L'écart-type du temps de parcours estimé par UKF et par CUPRITE diminue avec l'augmentation du pourcentage de véhicules sondes. Pour conclure, l'estimation UKF est moins bruitée que CUPRITE.



(a) (b)

Figure 6. 17 : Erreur de la méthode analytique classique des stocks (a) 1% fuite/source, (b) 5% fuite/source



(a) (b)

Figure 6. 18 : Erreur CUPRITE (a) 1% fuite/source, (b) 5% fuite/source

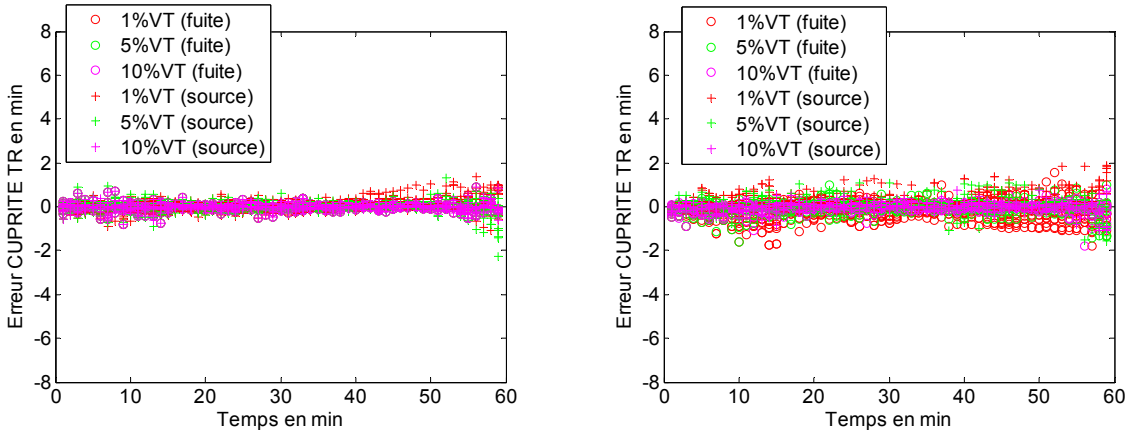


Figure 6. 19 : Erreur CUPRITE temps réel (a) 1% fuite/source, (b) 5% fuite/source

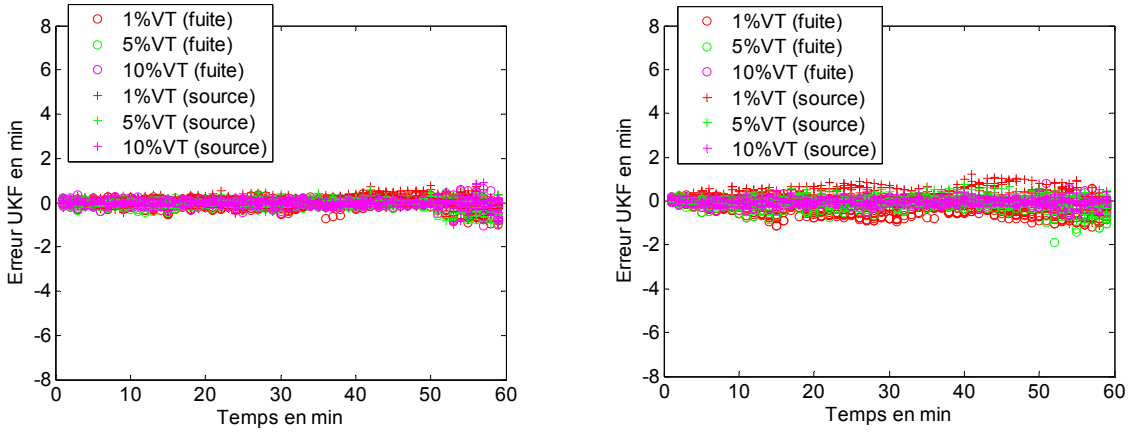


Figure 6. 20 : Erreur UKF (a) 1% fuite/source, (b) 5% fuite/source

Le tableau 6.1 résume la précision moyenne en pourcentage avec 1% et 5% de source/fuite à mi-lien, pour 10 réplications, pour la méthode analytique classique des stocks, CUPRITE, CUPRITE temps réel et UKF avec 1%, 5% et 10% de véhicules sondes. Il est évident que UKF et CUPRITE corrigent le biais de la méthode analytique classique des stocks. Comme prévu, leurs précisions augmentent avec l'augmentation du pourcentage de véhicules sondes.

Tableau 6. 1 : Précision moyenne (100-MAPE%) en %.

		Précision moyenne			
		Fuite à mi-lien		Source à mi-lien	
	% des véhicules traceurs	1%	5%	1%	5%
MACS		85	45	86	48
UKF	1	90	80	91	84
	5	93	90	93	91
	10	94	93	94	94
CUPRITE temps reel	1	87	77	89	85
	5	90	88	92	90
	10	92	90	92	92
CUPRITE	1	88	79	90	85
	5	91	89	92	90
	10	93	91	93	93

6.6.2.2 Estimation de la perturbation

Les figures 6.21 (a), (b), (c) et (d) illustrent l'estimation de la perturbation avec 1%, 5% et 10 % de véhicules traceurs pour respectivement 1% source, 5% source, 1% fuite et 5% fuite.

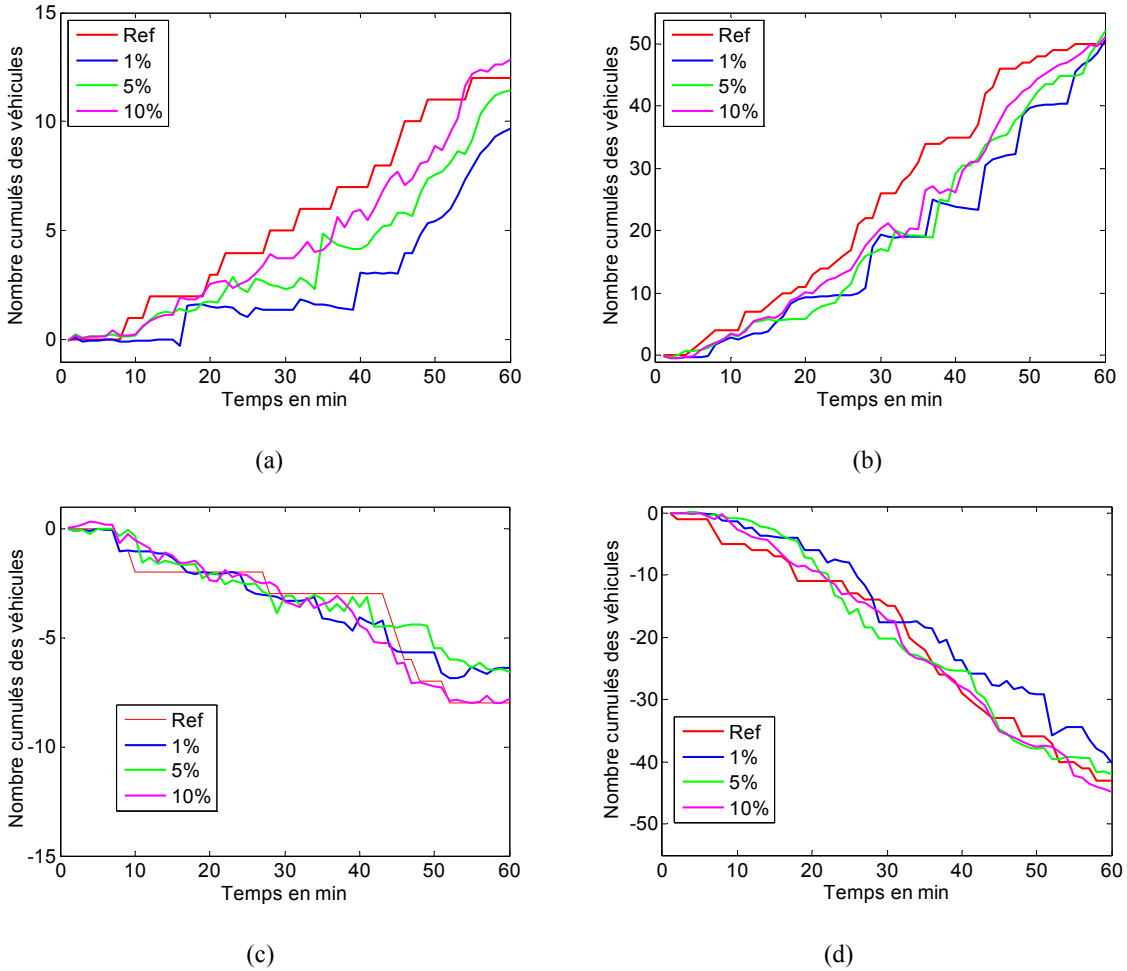


Figure 6. 21 : Estimation de la perturbation avec (a) 1% source (b) 5% source (c) 1% fuite (d) 5% fuite.

6.7 Analyse et synthèse

L'algorithme développé dans ce chapitre donne des résultats encourageants pour l'estimation du temps de parcours d'un lien, y compris avec des fuites et sources à mi-liens avec un pourcentage de mouvements tournants connu (ce qui est équivalent à une boucle magnétique à l'entrée du tronçon). Le modèle d'évolution de cet algorithme est basé sur la méthode analytique classique des stocks. Les observations sont les comptages des véhicules à partir de détecteurs à boucles magnétiques situés au début et à la fin de chaque lien et le temps de parcours des véhicules traceurs après qu'ils ont été associés aux liens appropriés par *map-matching*. En plus du temps de parcours, ce filtre estime le nombre de véhicules en amont, en aval et les sources/fuites à mi-liens.

De nombreuses simulations ont été effectuées avec un débit variable, un pourcentage variable de véhicules qui sont choisis au hasard et considéré comme sondes et un pourcentage variable de source/fuite à mi-lien.

Dans le cas d'un tronçon non perturbé, l'algorithme proposé a des résultats semblables à la méthode analytique classique des stocks et à CUPRITE et CUPRITE temps réel. Dans le cas d'un tronçon perturbé, la méthode analytique classique des stocks n'est plus applicable sans biais. Dans ce cas, l'UKF et CUPRITE corrigent ce biais et leurs résultats sont semblables.

La principale contribution de l'approche stochastique proposée est qu'elle surmonte la sensibilité de l'approche déterministe de CUPRITE par rapport aux véhicules sondes. En fait, CUPRITE suppose que les données utilisées sont exactes alors que le filtre UKF offre la possibilité d'introduire un modèle d'erreur pour le temps de parcours obtenu par les véhicules traceurs *map-matchés* ainsi que le comptage par les détecteurs à boucles magnétiques. Ainsi, l'estimation est lissée. En plus, les tests statistiques rendus possibles par le formalisme de Kalman et un réglage préalable des variances dans le filtre, permettent la détection et l'exclusion de valeurs aberrantes, à cause de données GPS mal *map-matchées* ou à cause de la défaillance de boucles par exemple. La faisabilité de la détection et l'exclusion de valeurs aberrantes a été démontrée. Finalement, l'UKF proposé peut être appliqué dans un contexte temps réel.

L'application de cette méthode à un réseau urbain est développée dans les chapitres 7 et 8 : il s'agit d'une extension de la méthode proposée dans le cas où les informations sur la proportion des mouvements tournants aux intersections est inconnue.

Chapitre 7. Méthode étendue à un réseau

SOMMAIRE

7.1 Introduction - 133 -

7.2 Description du modèle générique - 134 -
7.2.1 Vecteur d'état - 134 -
7.2.2 Modèle d'évolution - 135 -
7.2.3 Modèle d'observation - 135 -
7.2.4 Erreur d'observation - 136 -
7.2.5 L'algorithme - 136 -

7.3 Description du modèle dans le cas d'un divergent - 138 -

7.4 Description du modèle dans le cas d'un convergent - 141 -

7.1 Introduction

Dans ce chapitre, nous supposons que les pourcentages des mouvements tournants aux intersections d'un réseau sont inconnus et que nous cherchons à estimer le temps de parcours de chaque lien du réseau illustré dans la figure 7.1. Ce réseau est formé de 6 tronçons et chaque tronçon est équipé d'une boucle magnétique à son aval. Les deux entrées de ce réseau sont (E1) et (E2), et sa sortie est (S). En plus, il est équipé d'une boucle magnétique à chacune de ses entrées. Deux types d'intersection existent dans ce réseau, une divergente (le tronçon TR(1) alimente les deux tronçons TR(3) et TR(4) et deux convergentes (le tronçon TR(5) est alimenté par les deux tronçons TR(2) et TR(4), le tronçon TR(6) est alimenté par les deux tronçons TR(3) et TR(5)). De plus, des sources/fuites à mi-lien existent sur les tronçons TR(2), TR(3), TR(4) et TR(5). Ce petit réseau est un extrait d'un réseau typique urbain. Dans ce chapitre nous expliquons la modélisation générale. Pour en simplifier la présentation, elle sera détaillée pour le cas d'un divergent d'une part et pour le cas d'un convergent d'autre part. Dans le chapitre 8, l'application de chacun des cas (divergent, convergent, et réseau) sera faite en simulation.

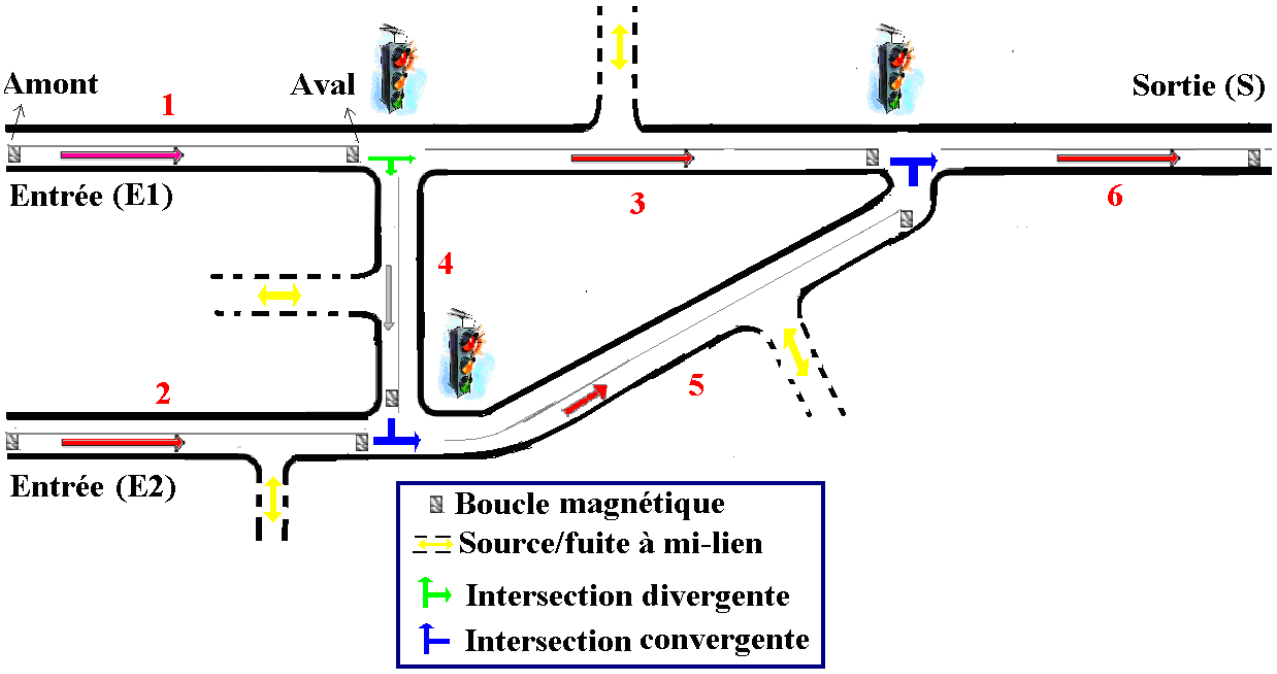


Figure 7. 1 : Réseau urbain

7.2 Description du modèle générique

Dans le cadre de ce travail, nous voulons estimer le temps de parcours et le nombre de véhicules comptés en amont et en aval de chaque lien k du réseau. En plus, nous voulons, s'il existe, estimer le nombre de véhicules qui entrent/sortent à mi-lien. Le vecteur d'état, l'évolution, l'observation et l'algorithme de filtrage proposés sont détaillés ci-dessous.

7.2.1 Vecteur d'état

Le vecteur d'état pour chaque lien k est composé comme suit :

- TP_k : le temps de parcours du lien k ;
- N_{dk} : le nombre cumulé de véhicules en aval du lien k ;
- q_{uk} : le flux en amont du lien k ;
- N_{uk} : le nombre cumulé de véhicules en amont du lien k ;
- P_k : le nombre cumulé de véhicules qui entrent/sortent du lien k à mi-lien ;
- l'historique des nombres de véhicules cumulés en amont. Cet historique compile un nombre fixe h_k de ces derniers, ce nombre étant un paramètre *a priori* du modèle spécifique pour chaque lien du réseau.

Par conséquent, le vecteur d'état (pour K liens) s'écrit comme suit :

$$x_k(t) = \begin{bmatrix} TP_k(t) \\ N_{dk}(t) \\ N_{uk}(t) \\ q_{uk}(t) \\ p_k(t) \\ N_{uk}(t - T_S) \\ .... \\ N_{uk}(t - h_k \times T_S) \end{bmatrix} k = 1, ..., K \qquad (5.11)$$

7.2.2 Modèle d'évolution

L'évolution de chaque élément du vecteur d'état est comme suit (voir §.5.5.3) :

$$\begin{aligned} TP_k(t) &= \mathrm{t} - N_{uk}^{-1}\left(N_{dk}(\mathrm{t})\right) \\ N_{dk}(t) &= N_{uk}\left(t - TP_k(t - Ts)\right) + p_k(t) \\ N_{uk}(t) &= N_{uk}(t - Ts) + q_{uk}(t) \times Ts \\ q_{uk}(t) &= q_{uk}(t - Ts) \\ p_k(t) &= p_k(t - Ts) \end{aligned} \tag{5.12}$$

7.2.3 Modèle d'observation

Les observations prises en compte ici sont à la fois le temps de parcours des véhicules traceurs et le nombre compté de véhicules par les boucles magnétiques. D'où les vecteurs d'observations suivants :

- cas 1 : un temps de parcours d'un véhicule traceur traversant le lien k (k=1-K) est une observation de TP_k avec un bruit d'observation w_{1t} supposé être un bruit blanc gaussien dont la moyenne est égale à zéro et la covariance est R_{1t} ;

- cas 2 : la lecture du compteur associé à un détecteur à l'aval d'un lien k est une observation de N_{dk} avec un bruit d'observation w_{2t} supposé être un bruit blanc gaussien de moyenne nulle de matrice de covariance R_{2t} ;

- cas 3 : la lecture du compteur associé à un détecteur à l'amont d'un lien l, lorsqu'il existe (ici l=1 et 2), est une observation de N_{ul} avec un bruit d'observation w_{2t} supposé être un bruit blanc gaussien de moyenne nulle et de matrice de covariance R_{2t} (généralement les réseaux sont équipés de boucles à leurs entrées).

En plus des observations ci-dessus, la loi de conservation des véhicules aux intersections du réseau introduit une contrainte qui peut être considérée comme une observation. En effet, cette loi stipule que (équation 7.3), à une intersection et à chaque instant *t*, le nombre total des véhicules des liens sortants (avec des indices $k = i$) est égal au nombre total des véhicules des liens entrants (avec des indices $k = j$). Si l'on considère que les pourcentages de mouvements tournants aux intersections sont connus *a priori*, alors le N_u du lien entrant est défini par la proportion correspondante du N_d total des liens sortants. Dans ce cas, le problème dans le réseau peut être considéré comme un problème découplé entre ses différents liens, et chaque

lien peut être traité par son propre algorithme. Par conséquent, il n'est pas nécessaire de prendre en considération l'équation de contrainte.

$$\sum_i N_{di} - \sum_j N_{uj} = 0 \tag{5.13}$$

7.2.4 Erreur d'observation

L'erreur des observations des boucles et véhicules traceurs sont expliquées dans le chapitre 5. Comme la contrainte aux intersections est considérée comme une observation, la covariance de l'erreur associée est égale à $\varepsilon << 1$, en raison de la grande confiance assumée dans la loi de conservation des véhicules

7.2.5 L'algorithme

La figure 7.2 résume l'algorithme proposé.

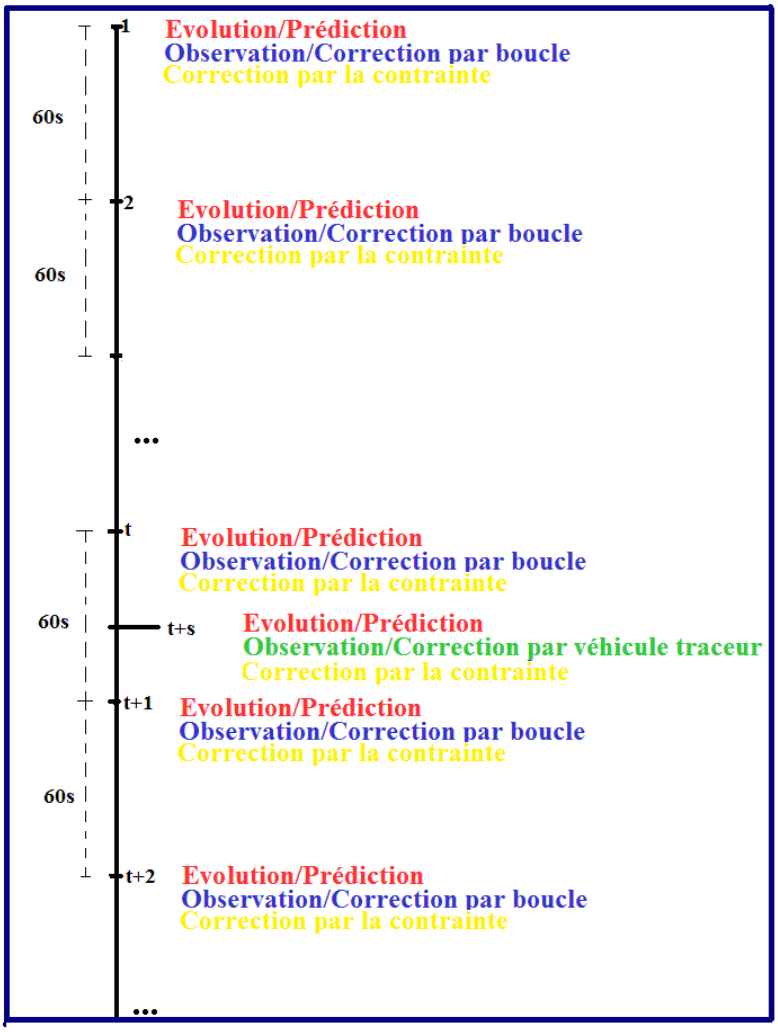


Figure 7. 2 : Algorithme général du modèle proposé

A chaque pas de calcul, égal à une minute, chaque élément du vecteur d'état est prédit par le modèle d'évolution. Comme les observations des boucles magnétiques sont agrégées à la minute, alors à chaque minute l'étape de prédiction est corrigée par les observations de ces boucles. Dans le cas où les pourcentages des mouvements tournants sont inconnus, alors la contrainte (équation 7.3) est applicable systématiquement. Enfin, supposons que l'information des véhicules traceurs soit accessible à l'instant t minutes et s secondes. Pour l'estimation en temps réel, nous proposons de faire une étape intermédiaire à $t+s$ (entre t et $t+1$ minute). Donc :

- à l'instant t, nous appliquons l'étape de prédiction et l'étape de correction par boucle et par contrainte ;

- à l'instant *t+s*, nous appliquons l'étape de prédiction et l'étape de correction par véhicules traceurs et par contrainte ;

- à l'instant *t+1*, nous appliquons l'étape de prédiction et l'étape de correction par boucle et par contrainte.

7.3 Description du modèle dans le cas d'un divergent

Dans le cas d'un tronçon TR(1) qui diverge en deux tronçons TR(2) et TR(3), comme illustré dans la figure 7.3, le nombre des véhicules qui sortent du TR(1) est égale à la somme des véhicules qui entrent dans TR(2) et TR(3). Ce divergent est équipé d'une boucle à son entrée et d'une boucle à la sortie de chaque tronçon.

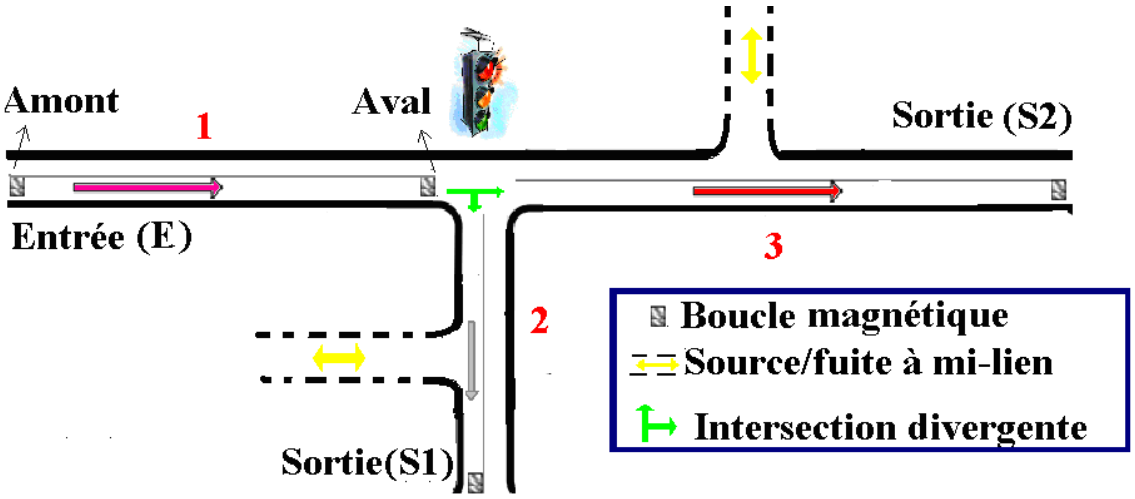


Figure 7. 3 : Tronçons divergents

Le temps de parcours des tronçons 1, 2 et 3 ainsi que les nombres des véhicules en amont, et en aval sur les 3 tronçons, et les source/fuites à mi-lien sont à estimer. Le vecteur d'état est le suivant :

$$x(t) = \begin{pmatrix} TP_1(t) \\ N_{d1}(t) \\ N_{u1}(t) \\ q_{u1}(t) \\ p_1(t) \\ N_{u1}(t-T_S) \\ \dots \\ N_{u1}(t-h_1 \times T_S) \\ TP_2(t) \\ N_{d2}(t) \\ N_{u2}(t) \\ q_{u2}(t) \\ p_2(t) \\ N_{u2}(t-T_S) \\ \dots \\ N_{u2}(t-h_2 \times T_S) \\ TP_3(t) \\ N_{d3}(t) \\ N_{u3}(t) \\ q_{u3}(t) \\ p_3(t) \\ N_{u3}(t-T_S) \\ \dots \\ N_{u3}(t-h_3 \times T_S) \end{pmatrix}$$

L'équation générale d'observation est :

$$y_t = Hx_t + v_t$$

Les observations sont :

Le temps de parcours des véhicules traceurs. Pour un véhicule donné, y_t est alors une observation de TP_1, TP_2 ou TP_3 selon que ce véhicule traverse TR(1), TR(2) ou TR(3)

- Dans le cas où TP_1 provient d'un véhicule sonde qui traverse le premier tronçon, H est :

$$\left[\underbrace{1 \quad 0 \quad 0 \quad 0 \quad \dots \quad 0}_{h_1+5} \underbrace{0 \quad 0 \quad 0 \quad 0 \quad \dots \quad 0}_{h_2+5} \underbrace{0 \quad 0 \quad 0 \quad 0 \quad \dots \quad 0}_{h_3+5} \right]$$

- Dans le cas où TP_2 provient d'un véhicule sonde qui traverse le deuxième tronçon, H est :

$$\left[\underbrace{0\ 0\ 0\ 0\ \ldots\ 0}_{h_1+5}\underbrace{1\ 0\ 0\ 0\ \ldots\ 0}_{h_2+5}\underbrace{0\ 0\ 0\ 0\ \ldots\ 0}_{h_3+5}\right]$$

- Dans le cas où TP_3 provient d'un véhicule sonde qui traverse le troisième tronçon, H est :

$$\left[\underbrace{0\ 0\ 0\ 0\ \ldots\ 0}_{h_1+5}\underbrace{0\ 0\ 0\ 0\ \ldots\ 0}_{h_2+5}\underbrace{1\ 0\ 0\ 0\ \ldots\ 0}_{h_3+5}\right]$$

Les nombres cumulés en aval de chaque tronçon sont agrégés au même instant. Donc à chaque instant le N_d de chaque tronçon est observé. y_t sera l'observation des N_{d1}, N_{d2} et N_{d3} et H sera :

$$\left[\underbrace{0\ 1\ 0\ 0\ \ldots\ 0}_{h_1+5}\underbrace{0\ 1\ 0\ 0\ \ldots\ 0}_{h_2+5}\underbrace{0\ 1\ 0\ 0\ \ldots\ 0}_{h_3+5}\right]$$

Les nombres cumulés en amont de TR(1). y_t sera l'observation de N_{u1} et H sera :

$$\left[\underbrace{0\ 0\ 1\ 0\ \ldots\ 0}_{h_1+5}\underbrace{0\ 0\ 0\ 0\ \ldots\ 0}_{h_2+5}\underbrace{0\ 0\ 0\ 0\ \ldots\ 0}_{h_3+5}\right]$$

La contrainte : le nombre de véhicules qui sortent du premier tronçon moins la somme des nombres de véhicules qui entrent dans les deuxième et troisième tronçons est égal à zéro. Donc H sera :

$$\left[\underbrace{0\ 1\ 0\ 0\ \ldots\ 0}_{h_1+5}\underbrace{0\ 0\ -1\ 0\ \ldots\ 0}_{h_2+5}\underbrace{0\ 0\ -1\ 0\ \ldots\ 0}_{h_3+5}\right]$$

7.4 Description du modèle dans le cas d'un convergent

Dans le cas où les deux tronçons TR(1) et TR(2) qui convergent au tronçon TR(3), comme illustré dans la figure 7.4, la somme des véhicules qui sortent de TR(1) et TR(2) est égale au nombre de véhicules qui entrent dans TR(3). Ce convergent est équipé d'une boucle dans chacune des ses entrées (E1 et E2) et d'une boucle à la sortie de chaque tronçon.

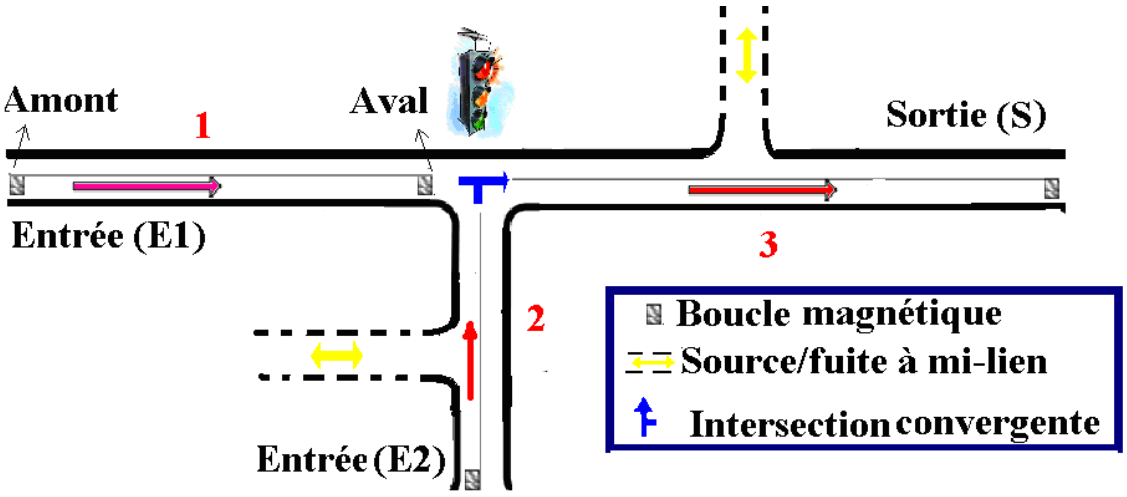


Figure 7. 4 : Tronçons convergents

Le temps de parcours des tronçons 1, 2 et 3, ainsi que les nombres des véhicules en amont et en aval sur les trois tronçons, et les source/fuites à mi-lien sont à estimer. Le vecteur d'état est le suivant :

$$x(t) = \begin{pmatrix} TP_1(t) \\ N_{d1}(t) \\ N_{u1}(t) \\ q_{u1}(t) \\ p_1(t) \\ N_{u1}(t-T_S) \\ \dots \\ N_{u1}(t-h_1 \times T_S) \\ TP_2(t) \\ N_{d2}(t) \\ N_{u2}(t) \\ q_{u2}(t) \\ p_2(t) \\ N_{u2}(t-T_S) \\ \dots \\ N_{u2}(t-h_2 \times T_S) \\ TP_3(t) \\ N_{d3}(t) \\ N_{u3}(t) \\ q_{u3}(t) \\ p_3(t) \\ N_{u3}(t-T_S) \\ \dots \\ N_{u3}(t-h_3 \times T_S) \end{pmatrix}$$

Les observations sont :

Le temps de parcours des véhicules traceurs. Pour un véhicule donné, y_t est alors une observation de TP_1, TP_2 ou TP_3 selon que ce véhicule traverse TR(1), TR(2) ou TR(3)

- Dans le cas où l'observation de TP_1 provient d'un véhicule sonde qui traverse le premier tronçon, H est :

$$\left[\underbrace{1 \quad 0 \quad 0 \quad 0 \quad \dots \quad 0}_{h_1+5} \underbrace{0 \quad 0 \quad 0 \quad 0 \quad \dots \quad 0}_{h_2+5} \underbrace{0 \quad 0 \quad 0 \quad 0 \quad \dots \quad 0}_{h_3+5} \right]$$

- Dans le cas où l'observation de TP_2 provient d'un véhicule sonde qui traverse le deuxième tronçon, H est:

$$\left[\underbrace{0\ 0\ 0\ 0\ \ldots\ 0}_{h_1+5}\ \underbrace{1\ 0\ 0\ 0\ \ldots\ 0}_{h_2+5}\ \underbrace{0\ 0\ 0\ 0\ \ldots\ 0}_{h_3+5}\right]$$

- Dans le cas où l'observation de TP_3 provient d'un véhicule sonde qui traverse le troisième tronçon, H est:

$$\left[\underbrace{0\ 0\ 0\ 0\ \ldots\ 0}_{h_1+5}\ \underbrace{0\ 0\ 0\ 0\ \ldots\ 0}_{h_2+5}\ \underbrace{1\ 0\ 0\ 0\ \ldots\ 0}_{h_3+5}\right]$$

Les nombres cumulés en aval de chaque tronçon sont agrégés au même instant. Donc à chaque instant le N_d de chaque tronçon est observé. y_t sera l'observation des N_{d1}, N_{d2} et N_{d3} et H sera :

$$\left[\underbrace{0\ 1\ 0\ 0\ \ldots\ 0}_{h_1+5}\ \underbrace{0\ 1\ 0\ 0\ \ldots\ 0}_{h_2+5}\ \underbrace{0\ 1\ 0\ 0\ \ldots\ 0}_{h_3+5}\right]$$

Les nombres cumulés en amont de TR(1) et TR(2) sont agrégés au même instant. y_t sera l'observation des N_{u1} et N_{u2} et H sera :

$$\left[\underbrace{0\ 0\ 1\ 0\ \ldots\ 0}_{h_1+5}\ \underbrace{0\ 0\ 1\ 0\ \ldots\ 0}_{h_2+5}\ \underbrace{0\ 0\ 0\ 0\ \ldots\ 0}_{h_3+5}\right]$$

La contrainte : la somme des nombres de véhicules qui sortent du premier tronçon et du deuxième tronçon moins le nombre de véhicules qui entrent dans le troisième tronçon est égale à zéro. Donc H sera :

$$\left[\underbrace{0\ 1\ 0\ 0\ \ldots\ 0}_{h_1+5}\ \underbrace{0\ 1\ 0\ 0\ \ldots\ 0}_{h_2+5}\ \underbrace{0\ 0\ -1\ 0\ \ldots\ 0}_{h_3+5}\right]$$

Chapitre 8. Application de la méthode étendue à un divergent, à un convergent et à un réseau

SOMMAIRE

8.1 Introduction - 147 -

8.2 Application pour le cas d'un divergent - 147 -
8.2.1 Description des scénarios choisis - 147 -
8.2.2 Le comportement de CUPRITE - 149 -

8.3 Application pour le cas d'un convergent - 155 -
8.3.1 Description des scénarios choisis - 155 -
8.3.2 Le comportement de CUPRITE - 155 -
8.3.3 Estimation par l'UKF du nombre cumulé en amont - 156 -
8.3.4 Estimation du temps de parcours - 156 -

8.4 Application dans un réseau - 159 -
8.4.1 Description des scénarios choisis - 159 -
8.4.2 Estimation du temps de parcours - 160 -
8.4.3 Estimation de la perturbation - 162 -

8.5 Synthèse et analyse - 163 -

8.1 Introduction

Ce chapitre examine le comportement de la nouvelle approche adaptée au réseau. Trois parties composent ce chapitre : la première concerne le cas où les tronçons divergent, la deuxième le cas où les tronçons convergent, et la troisième partie le cas du petit réseau complet.

Dans ce chapitre, comme dans le chapitre 7, le pourcentage des mouvements tournants est inconnu. On verra dans ce chapitre comment l'application de la méthode analytique classique des stocks conduit à un biais dans le calcul du temps de parcours. Dans ce chapitre également, nous examinerons le comportement de CUPRITE, pour finalement montrer l'apport de notre approche.

8.2 Application pour le cas d'un divergent

8.2.1 Description des scénarios choisis

La simulation est établie sur un divergent comme illustré figure 7.3, dont chaque tronçon est de 600 mètres de longueur. Chaque tronçon n'est formé que d'une seule voie, et est équipé d'un détecteur à l'aval (sortie). L'entrée du divergent est aussi équipée à l'amont. Les paramètres des tronçons sont les suivants :

- la vitesse libre de circulation est de 36 km/h ;
- la densité maximale est de 140 véh/km ;
- le débit est limité à 1008 véh/h.

Pendant une heure de simulation le flux qui pénètre TR(1) est comme suit : pour les vingt premières minutes le débit moyen est de 600 véh/h, pour les vingt minutes suivantes le débit moyen est de 1200 véh/h, pour les vingt dernières minutes le débit moyen diminue pour revenir à 600 véh/h.

Les pourcentages d'affectation aussi varient en fonction du temps et le flux se répartit comme suit : 60% des véhicules iront vers TR(2) et 40 % vers TR(3) pour les 20 premières minutes. 50% des véhicules iront vers TR(2) et TR(3) pour les 20 minutes suivantes et finalement 40% des véhicules iront vers TR(2) et 60 % vers TR(3) pour les 20 dernières minutes.

Pour chaque véhicule, l'instant de son entrée et l'instant de sa sortie sont enregistrés. Le *TP* instantané réalisé est la différence entre l'instant de sortie du véhicule et l'instant de son entrée. Ce *TP* est associé à l'instant de sortie. Nous considérons que le temps de parcours moyen de référence à l'instant *t* est la moyenne des temps de parcours de tous les véhicules qui sont sortis du tronçon entre l'instant *t-1* et *t*. Le calcul du temps de parcours de référence est réalisé à chaque minute.

Les véhicules traceurs sont un échantillon aléatoire de la population, et quatre scénarios ont été préparés :

1. aucun de ces véhicules n'a été considéré comme véhicule sonde ;
2. 1% des véhicules ont été considérés comme des véhicules sondes ;
3. 5% des véhicules ont été considérés comme des véhicules sondes ;
4. 10% des véhicules ont été considérés comme des véhicules sondes.

Les données mises à disposition pour l'estimation du temps de parcours sont :

- le nombre de véhicules à l'aval pour chaque intervalle de temps pour chaque tronçon ;
- le nombre de véhicules à l'amont pour chaque intervalle de temps pour le premier tronçon (tronçon d'entrée) ;
- les temps de parcours réalisés par les véhicules traceurs aux temps correspondants à la sortie du véhicule et pour chaque tronçon.

Afin de calculer l'erreur d'estimation, on dispose du temps de parcours moyen de référence (TP réf) à chaque minute.

Les scénarios décrits observent la loi de conservation des véhicules, c'est-à-dire sans fuite ni source de véhicules et tous les véhicules qui sont entrés dans le tronçon TR(1), sont sortis après leur temps de parcours *TP* soit du tronçon TR(2) soit du tronçon TR(3).

Six réplications de cette simulation ont été réalisées.

8.2.2 Le comportement de CUPRITE

Le problème est qu'on ne sait pas répartir le flux de TR(1) vers TR(2) et TR(3) : à défaut de cette information, on peut diviser le flux en deux, ou même supposer (et c'est ce que propose CUPRITE) que la totalité du flux alimente à la fois TR(2) et TR(3). Ceci mène (sauf par chance…) à un biais d'estimation du temps de parcours par la méthode analytique classique des stocks.

CUPRITE est sensé corriger ladite méthode. Dans le cas où CUPRITE estime le temps de parcours, par exemple de TR(2), il suppose que le nombre cumulé de véhicules à l'amont de TR(2) (N_{u2}) est égal au nombre cumulé de véhicules en aval de TR(1) (N_{d1}) multiplié par le pourcentage de véhicules qui entrent dans TR(2) (voir chapitre 6), puis il applique la méthode de correction de N_{u2} par les véhicules traceurs sortant de TR(2). Mais dans le cas ou le pourcentage des mouvements tournants est inconnu, CUPRITE suppose que $N_{d1} = N_{u2}$. La figure 8.1 montre l'estimation de N_{u2}. Dans ce cas, 10% des véhicules sont supposés sondes. Leur répartition n'est pas homogène, et quand ils sont assez fréquents, la correction CUPRITE est satisfaisante (c'est ce qu'on voit sur la deuxième demi-heure de simulation). Cependant, même avec ce pourcentage, il arrive qu'entre deux véhicules la correction de N_{u2} par CUPRITE se situe sous N_{d2}, ce qui conduit à une impossibilité de calcul du temps de parcours. Avec CUPRITE temps réel, nous remarquons que N_{u2} est corrigé à chaque véhicule traceur. Mais comme il n'y a pas d'anticipation de cette correction avant son occurrence, il y a sur-comptage à l'entrée (N_{u2} corrigé est trop grand), conduisant à un biais dans le calcul du temps de parcours. Celui-ci croît systématiquement à partir des véhicules sondes. A partir de 20% de véhicule traceurs CUPRITE a néanmoins corrigé correctement N_{u2}. Le même comportement est observé pour N_{u3}.

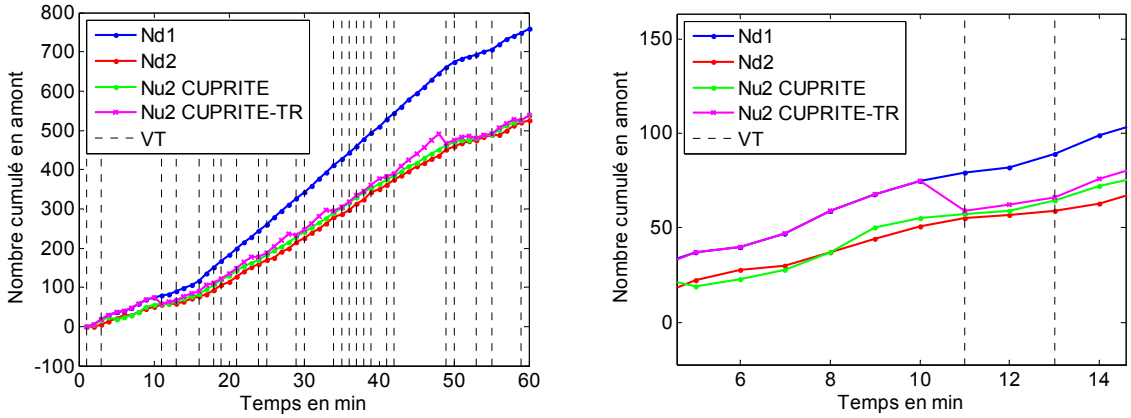


Figure 8. 1: Correction des nombres cumulés de véhicules en amont par CUPRITE

Les figures 8.2 et 8.3 représentent respectivement l'estimation du nombre cumulé de véhicules en amont pour l'une des réplications et la différence entre le nombre cumulé en amont de référence, et le nombre cumulé en amont estimé par UKF pour le tronçon TR(2). Dans l'algorithme proposé, ces nombres cumulés ne sont pas observés, mais l'algorithme rend possible leur estimation avec une erreur maximale de ±8 véhicules, pour un cumul initialement nul et atteignant quelques centaines de véhicules (500 véhicules). Les mêmes résultats sont obtenus pour le tronçon TR(3).

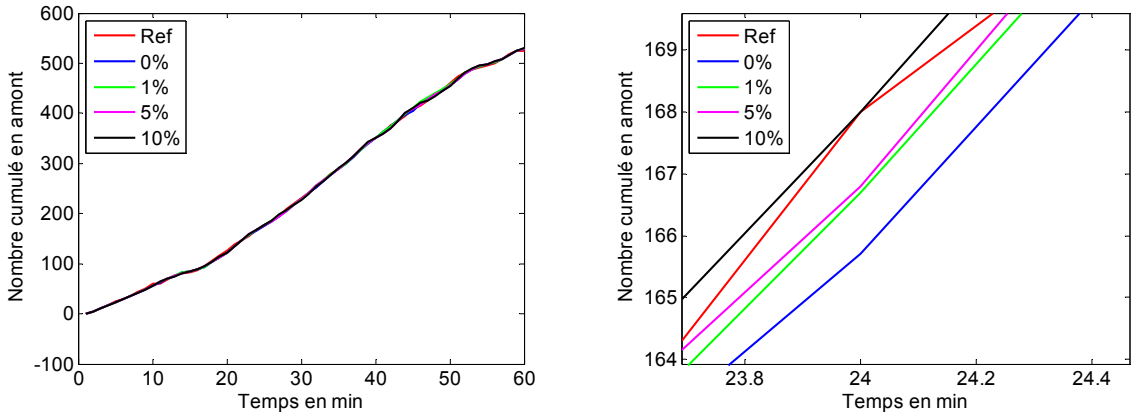


Figure 8. 2 : Estimation du nombre cumulé de véhicules à l'amont du deuxième tronçon

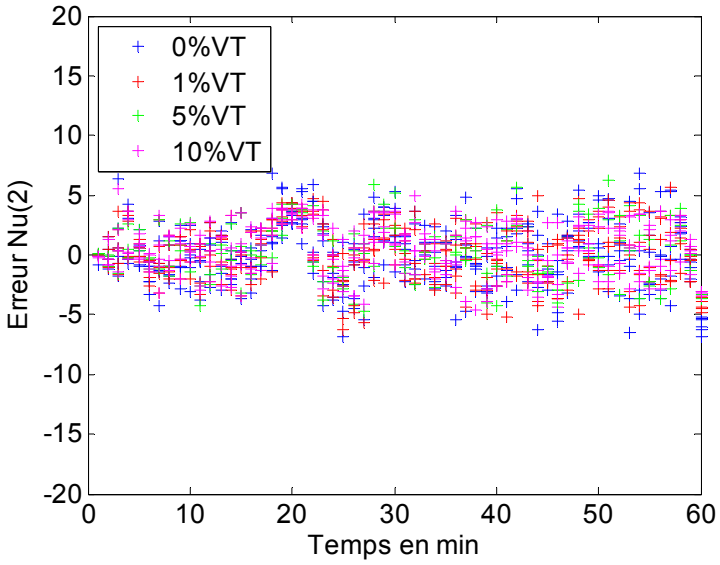
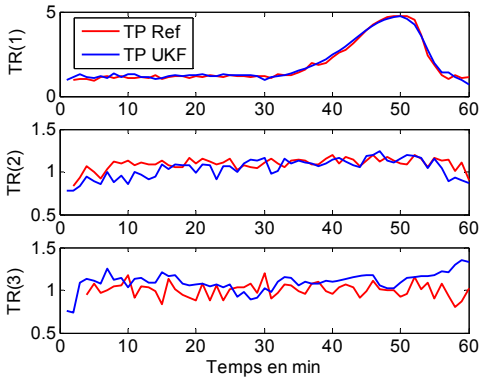
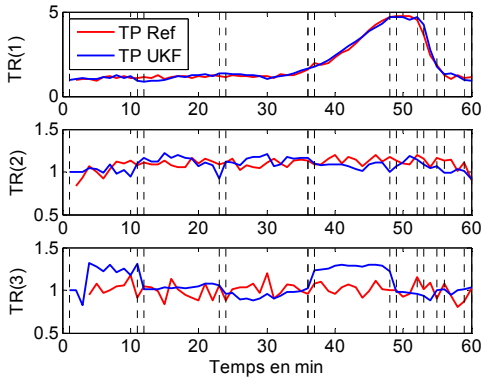


Figure 8. 3 : Erreur d'estimation du nombre cumulé de véhicules à l'amont du deuxième tronçon

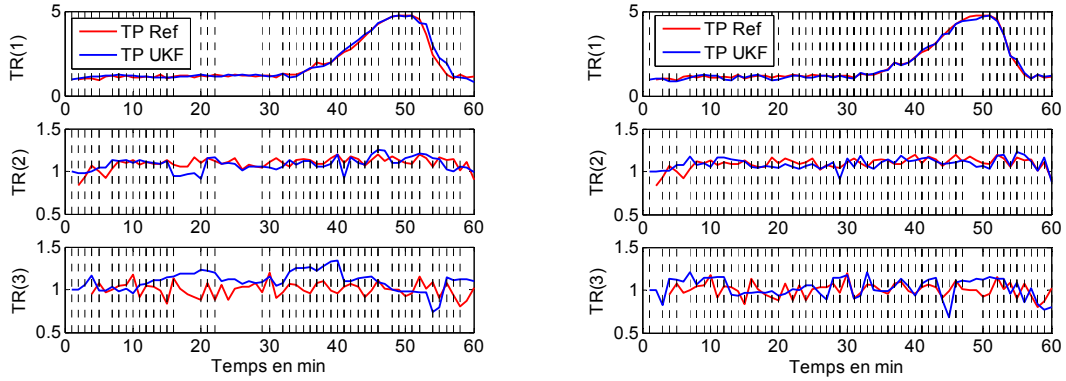
Les figures 8.4 (a), (b), (c) et (d) sont les graphes du temps de parcours de référence et du temps de parcours estimé par UKF pour les trois tronçons respectivement avec 0%, 1%, 5% et 10% de véhicules traceurs.



(a)



(b)



(c) (d)

Figure 8. 4 : Estimation du temps de parcours avec (a) 0%, (b) 1%, (c) 5%, (d) 10% de véhicules traceurs

On note qu'au cours de la simulation et selon comment les véhicules choisissent TR(2) ou TR(3), il est possible qu'aucun véhicule simulé n'entre pendant une minute, voire deux, dans l'un ou l'autre des liens (notamment quand le débit d'entrée est faible). Ainsi, le calcul de TP Ref est discontinu (il peut manquer des valeurs dans la courbe rouge) et les compteurs restent inchangés pendant ces minutes là. Ceci étant, grâce au formalisme de l'algorithme proposé, le temps de parcours est systématiquement prédit par le modèle d'évolution et corrigé au moins par les observations des boucles. La figure 8.5 illustre le temps de parcours estimé par UKF et le temps de parcours de référence, particulièrement discontinu pour cette réplication.

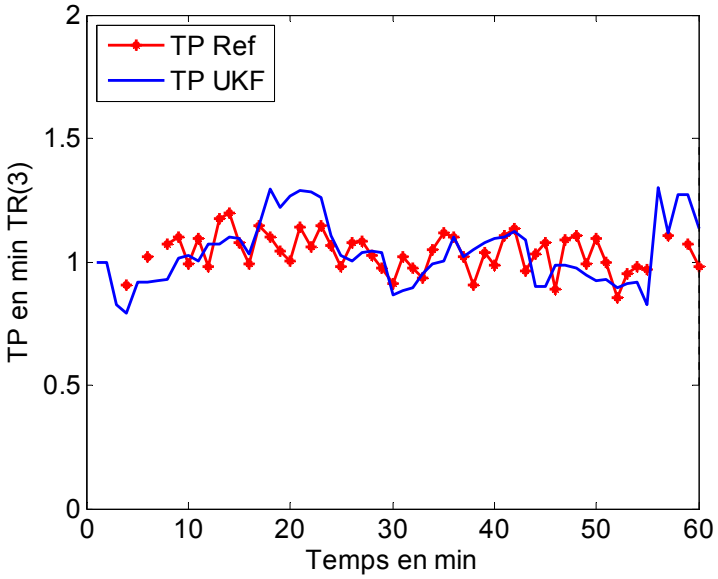


Figure 8. 5 : Effet de la prédiction

Nous montrons ci-dessous les graphes de la différence entre le temps de parcours de référence (TP Ref) et le temps de parcours estimé par UKF (TP UKF) en minute pour chacun des tronçons TR(1), TR(2) et TR(3). L'erreur pour les trois tronçons diminue avec l'augmentation du pourcentage de véhicules traceurs. Même sans aucune information sur le pourcentage des mouvements tournants, l'UKF est capable d'estimer le temps de parcours avec une erreur maximale de 1 minute.

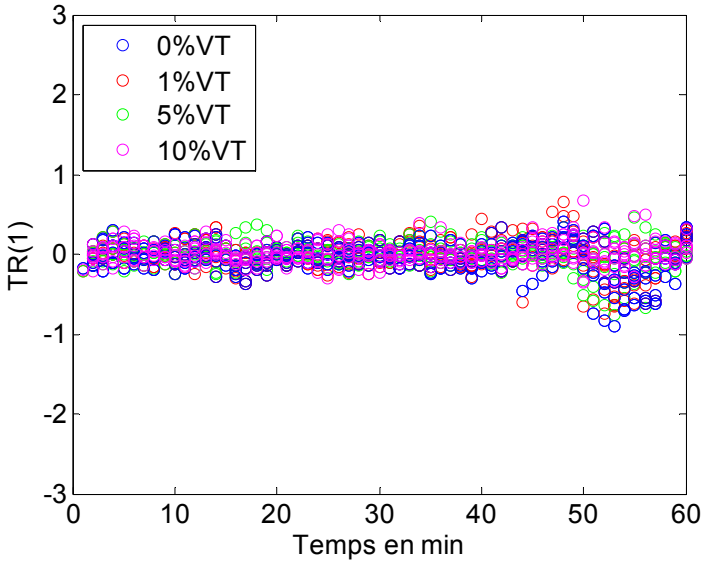


Figure 8. 6 : Erreur d'estimation du temps de parcours pour le premier tronçon

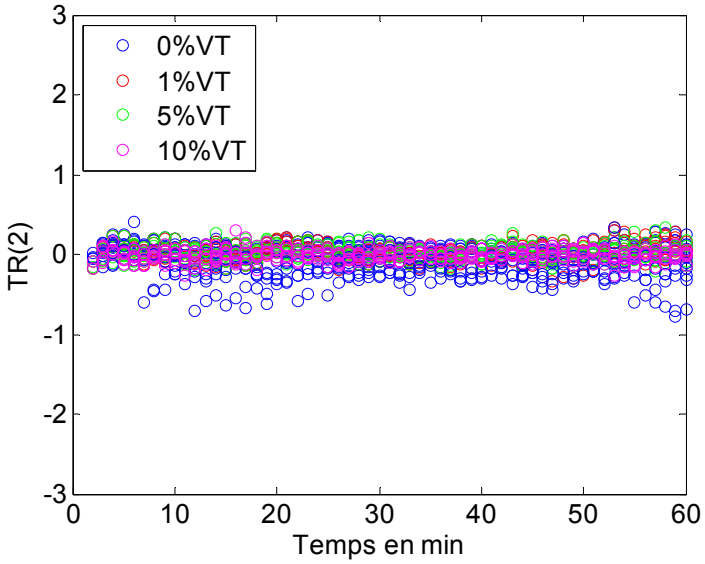


Figure 8. 7 : Erreur d'estimation du temps de parcours pour le deuxième tronçon

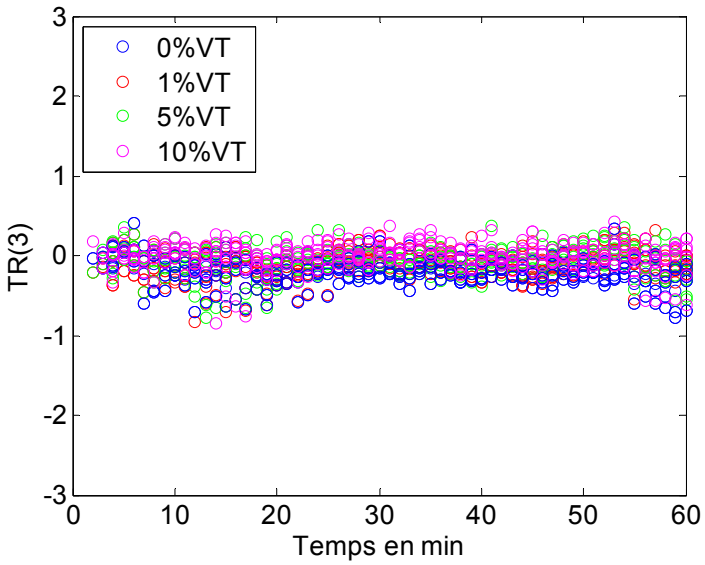


Figure 8. 8 : Erreur d'estimation du temps de parcours pour le troisième tronçon

Le tableau 8.1 résume la précision moyenne en pourcentage de l'UKF pour chaque lien, pour 6 réplications, avec respectivement 0%, 1%, 5% et 10% des véhicules sondes. La précision augmente avec le pourcentage de véhicules comme prévu.

Tableau 8. 1 : Précision moyenne (100-MAPE%) en % pour un divergent.

	Tronçon		
% de véhicules traceurs	**1**	**2**	**3**
0	**90**	**89**	**85**
1	**92**	**90**	**87**
5	**93**	**91**	**89**
10	**94**	**92**	**90**

8.3 Application pour le cas d'un convergent

8.3.1 Description des scénarios choisis

La simulation est maintenant établie sur un convergent comme illustré figure 7.4. La seule différence par rapport au divergent est qu'il existe deux entrées, toutes deux équipées d'un détecteur.

Pendant une heure de simulation le flux qui pénètre dans chacun des liens TR(1) et TR(2) est comme suit : pour les vingt premières minutes le débit moyen est de 600 véh/h, pour les vingt deuxièmes minutes le débit moyen est de 1200 véh/h, pour les vingt dernières minutes le débit moyen diminue pour atteindre 600 véh/h. Ces deux flux sont les mêmes et s'ajoutent au convergent TR(3).

Les données mis en disposition pour l'estimation du temps de parcours sont :

- le nombre de véhicules à l'aval pour chaque intervalle de temps pour chaque tronçon ;
- le nombre de véhicules à l'amont pour chaque intervalle de temps pour le premier et le deuxième tronçon ;
- le temps de parcours réalisé par les véhicules traceurs correspondant à la sortie du véhicule et pour chaque tronçon.

Les scénarios décrits sont soumis à la loi de conservation des véhicules, c'est-à-dire sans fuite ni source de véhicules et tous les véhicules qui sont entrés dans les tronçons TR(1) et TR(2) sont sortis après leur temps de parcours *TP* du tronçon TR(3). 6 réplications de cette simulation ont été réalisées.

8.3.2 Le comportement de CUPRITE

Dans le cas d'un convergent le nombre cumulé en amont de TR(3) est la somme des nombres cumulés mesurés en sortie de TR(1) et TR(2). Donc la méthode analytique classique des stocks ainsi que CUPRITE fonctionnent comme un tronçon équipé d'une boucle en entrée (voir chapitre 6).

8.3.3 Estimation par l'UKF du nombre cumulé en amont

On montre ci-dessous le graphe de la différence entre le nombre cumulé en amont de référence, et le nombre cumulé en amont estimé par UKF pour TR(3). Dans l'algorithme proposé ce nombre cumulé est observable avec l'équation de la contrainte et est estimé avec une erreur maximale de 5 véhicules.

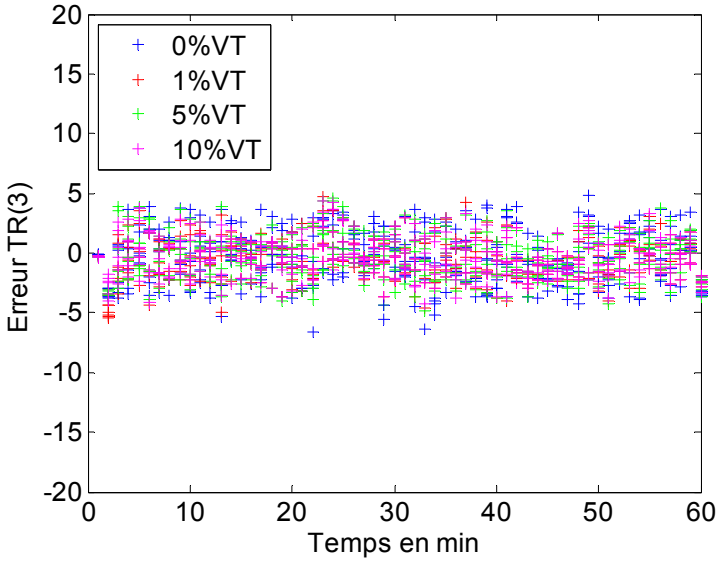


Figure 8. 9 : Erreur d'estimation du nombre cumulé de véhicules à l'amont du troisième tronçon

8.3.4 Estimation du temps de parcours

Les figures 8.10 représentent le temps de parcours de référence (TP Ref) et le temps de parcours estimé par UKF (TP UKF) pour chacun des tronçons TR(1), TR(2) et TR(3) avec respectivement 0%, 1%, 5% et 10% de véhicules traceurs.

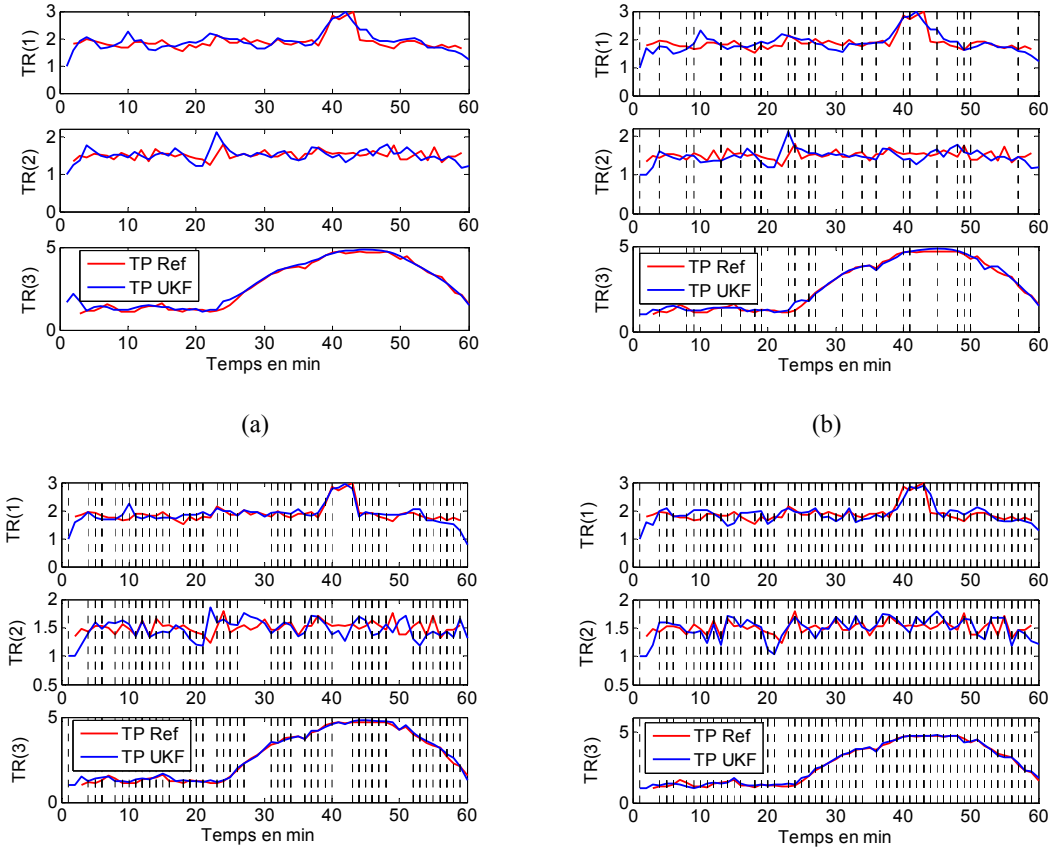


Figure 8. 10 : Estimation du temps de parcours avec (a) 0%, (b) 1%, (c) 5%, (d) 10% de véhicules traceurs

Nous montrons ci-dessous les graphes de la différence entre le temps de parcours de référence (TP Ref) et le temps de parcours estimé par UKF (TP UKF) en minute pour chacun des tronçons TR(1), TR(2) et TR(3). L'erreur pour les trois tronçons diminue avec l'augmentation du pourcentage de véhicules traceurs avec une erreur maximale de 0,75 minute.

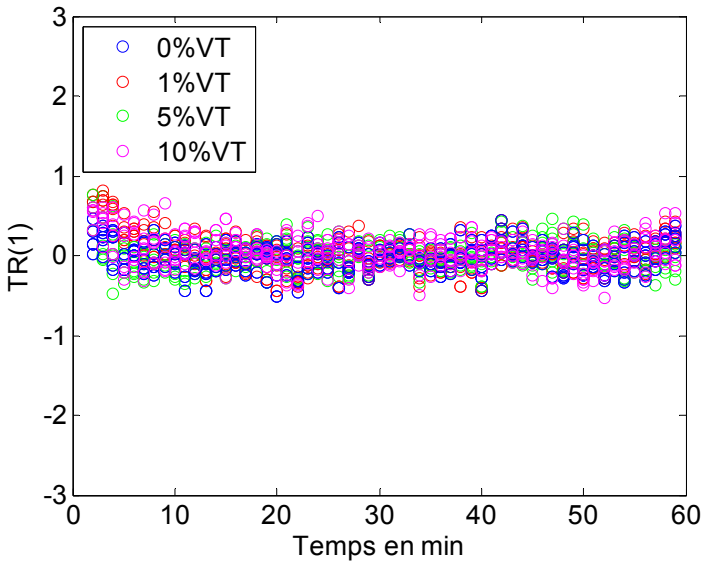


Figure 8. 11 : Erreur d'estimation du temps de parcours pour le premier tronçon

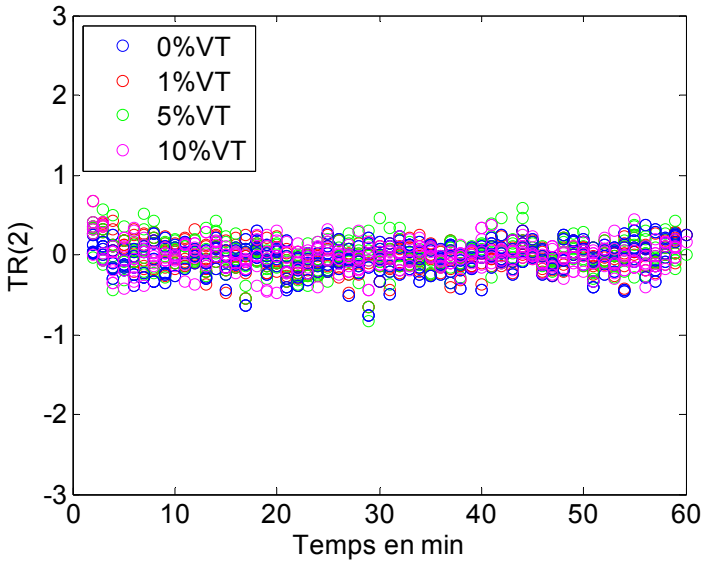


Figure 8. 12 : Erreur d'estimation du temps de parcours pour le deuxième tronçon

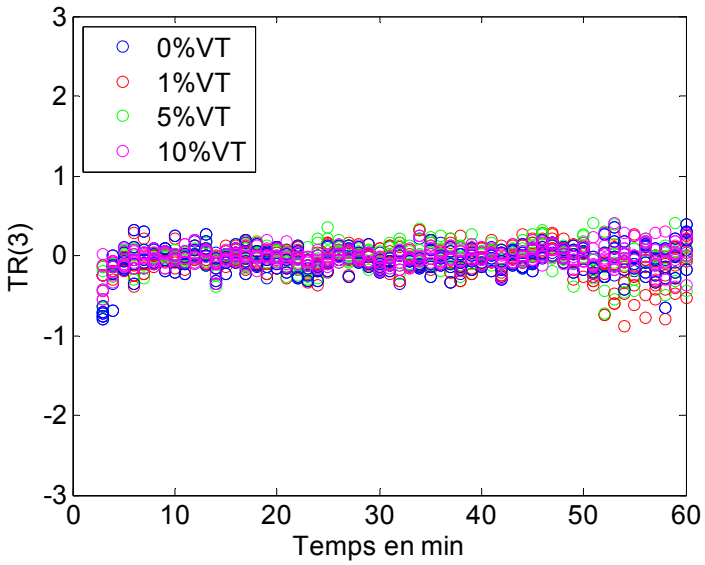


Figure 8. 13 : Erreur d'estimation du temps de parcours pour le troisième tronçon

Le tableau 8.2 résume la précision moyenne en pourcentage de l'UKF pour chaque lien, pour 6 réplications, avec respectivement 0%, 1%, 5% et 10% des véhicules sondes. La précision augmente avec le pourcentage de véhicules comme prévu.

Tableau 8. 2 : Précision moyenne (100-MAPE%) en % pour un convergent.

	Tronçon		
% de véhicules traceurs	**1**	**2**	**3**
0	**90**	**90**	**91**
1	**91**	**90**	**92**
5	**92**	**91**	**93**
10	**92**	**91**	**93**

8.4 Application dans un réseau

8.4.1 Description des scénarios choisis

La simulation est établie sur un réseau comme illustré figure 7.1, dont chaque tronçon est de 600 mètres de longueur. Chaque tronçon n'est formé que d'une seule voie, et est équipé

d'un détecteur à l'aval (sortie). Les deux entrées du réseau sont aussi équipées à l'amont. Les paramètres des tronçons sont les mêmes que dans le cas du divergent (voir §.8.2.1).

Pendant une heure de simulation le flux qui pénètre dans chacun des liens TR(1) et TR(2) est comme suit : pour les vingt premières minutes le débit moyen est de 600 véh/h, pour les vingt deuxièmes minutes le débit moyen est de 1200 véh/h, pour les vingt dernières minutes le débit moyen diminue pour atteindre 600 véh/h. La répartition du flux sortant de TR(1) se fait dans TR(3) et TR(4) comme dans le cas du divergent (voir §.8.2.1). Les deux autres intersections sont des convergents.

Les données mis en disposition pour l'estimation du temps de parcours sont :

- le nombre de véhicules à l'aval pour chaque intervalle de temps pour chaque tronçon ;
- le nombre de véhicules à l'amont pour chaque intervalle de temps pour le premier et le deuxième tronçon ;
- le temps de parcours réalisé par les véhicules traceurs correspondant à la sortie du véhicule et pour chaque tronçon.

Les tronçons 1 et 6 sont soumis à la loi de conservation des véhicules, c'est-à-dire sans fuite ni source de véhicules et les autres tronçons sont perturbés avec une proportion qui varie de 1% à 3% (1% de fuite pour TR(2) et TR(5), et 3% de source pour TR(3) et TR(4)). 10 réplications de cette simulation ont été réalisées.

8.4.2 Estimation du temps de parcours

Le tableau 8.1 résume la précision moyenne en pourcentage de l'UKF pour chaque lien avec 1% à 3% de source/fuite à mi-lien, pour les 10 réplications, avec respectivement 0%, 1%, 5% et 10% de véhicules sondes. La performance est conforme à tous les liens, qu'ils soient perturbés ou non, et la précision augmente avec le pourcentage de véhicules sondes. La figure 8.15 montre la différence entre le temps de parcours de référence (TP ref) et le temps de parcours estimé par UKF en minute pour chacun des tronçons du réseau. Même sans informations sur la proportion des mouvements tournants aux intersections, l'UKF est capable d'estimer correctement le *TP* des tronçons avec une erreur maximale de 1 minute.

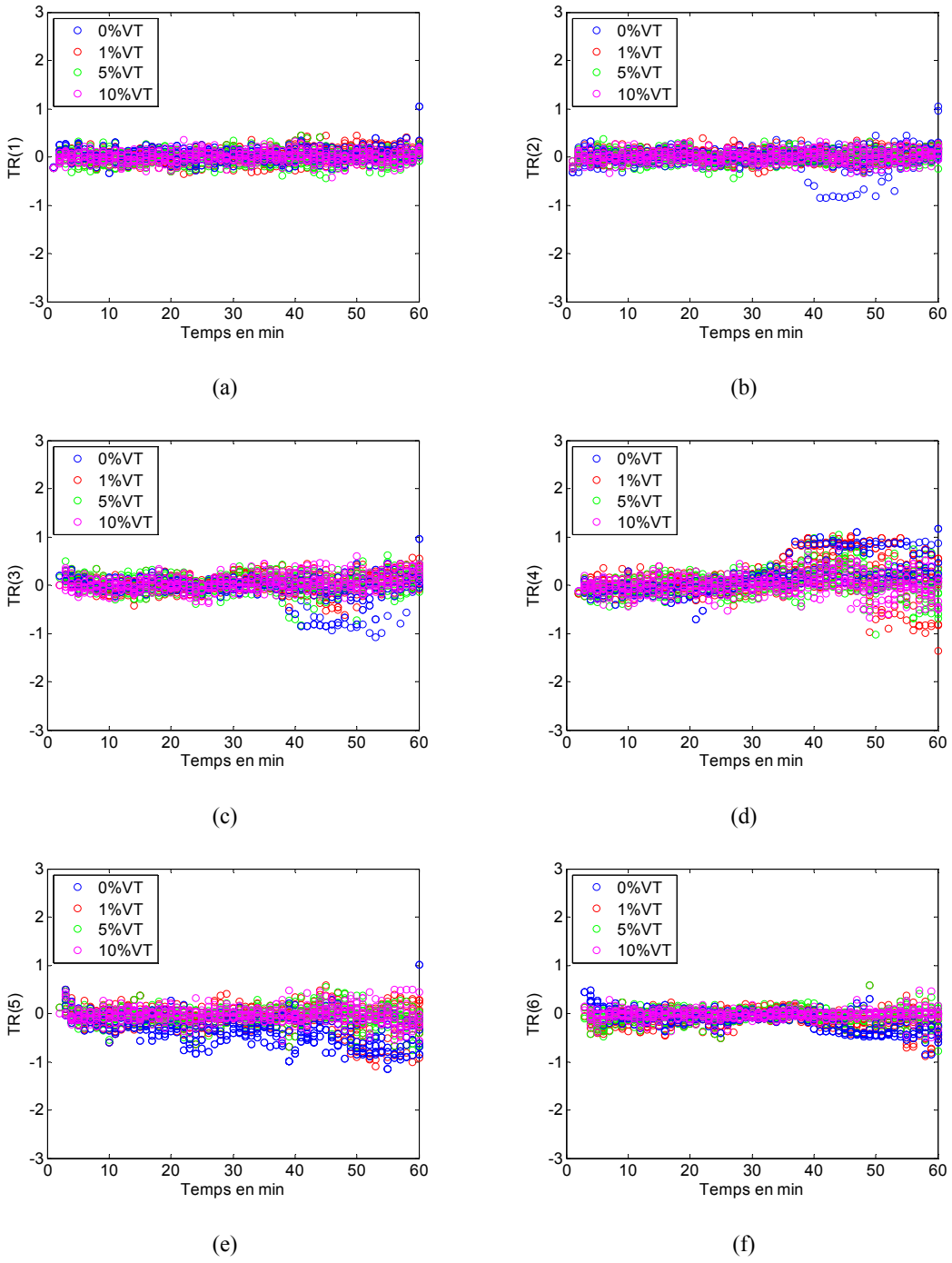


Figure 8. 14 : Erreur d'estimation du temps de parcours pour (a) TR(1), (b) TR(2), (c) TR(3), (d) TR(4), (e) TR(5), (f) TR(6)

Tableau 8. 3 : Précision moyenne (100-MAPE%) en % pour un réseau.

	Tronçon					
% de véhicules traceurs	**1**	**2**	**3**	**4**	**5**	**6**
0	**89**	**89**	**86**	**86**	**90**	**91**
1	**90**	**90**	**87**	**87**	**91**	**94**
5	**91**	**91**	**90**	**90**	**92**	**96**
10	**92**	**92**	**91**	**91**	**94**	**97**

8.4.3 Estimation de la perturbation

Les figures 8.15 (a), (b), (c) et (d) illustrent l'estimation de la perturbation avec 0%, 1%, 5% et 10% de véhicules traceurs pour respectivement 3% de source pour TR(3) et TR(4), 1% de fuite pour TR(2) et TR(5). Ici, nous remarquons aussi l'apport des véhicules traceurs : à chaque fois que le pourcentage des véhicules traceurs augmente, la courbe de la perturbation estimée s'approche de la courbe de référence.

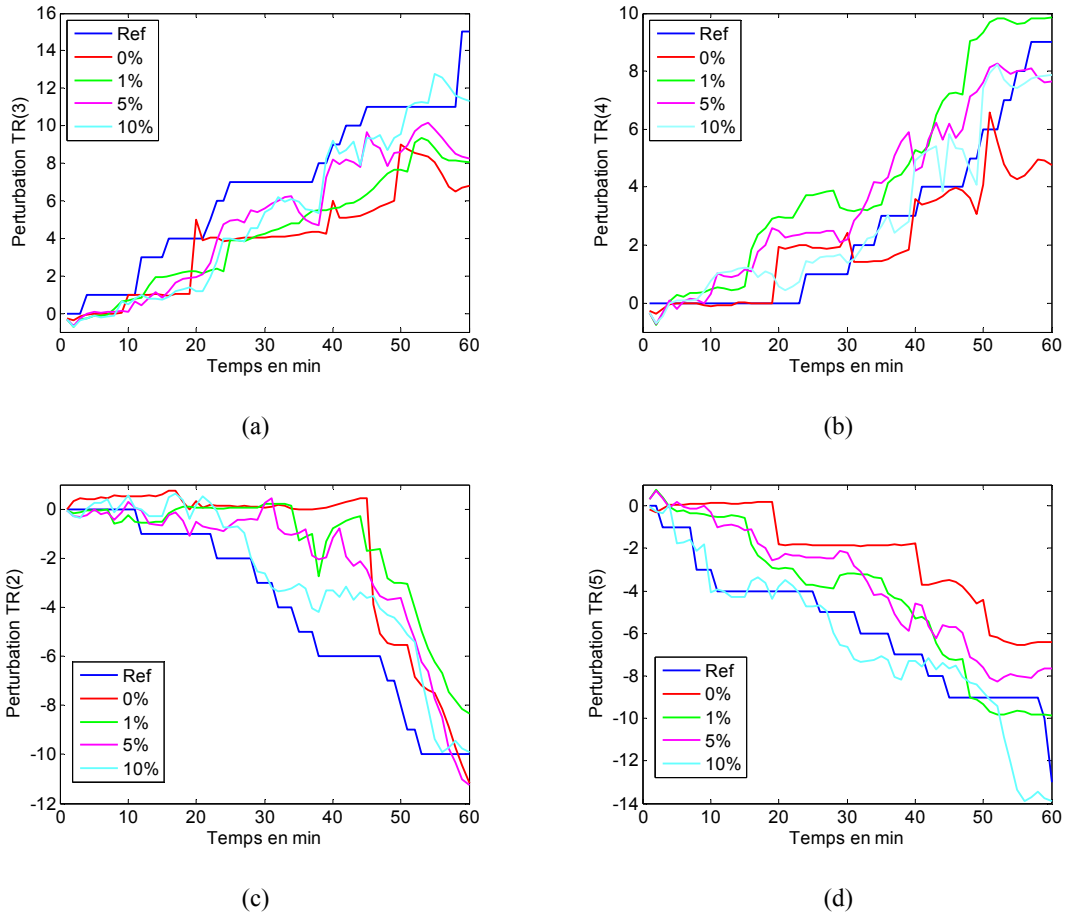


Figure 8. 15 : Estimation de la perturbation pour (a) 3% source TR(3), (b) 3% source TR(4), (c) 1% fuite TR(2), (d) 1% fuite TR(5).

8.5 Synthèse et analyse

Le filtre UKF développé ici donne des résultats encourageants pour l'estimation de temps de parcours dans un réseau urbain avec des fuites et des sources à mi-lien et un pourcentage inconnu des mouvements tournants. La méthode analytique classique des stocks biaise l'estimation du temps de parcours dans de telles conditions. CUPRITE a besoin d'un nombre minimal de véhicules traceurs pour corriger correctement le nombre de véhicules en amont afin d'estimer le temps de parcours.

Le modèle d'évolution de notre algorithme est basé sur la méthode analytique classique des stocks. Les observations sont les comptages des véhicules en provenance des détecteurs à boucles, qui sont situés à la fin de chaque lien, et les temps de parcours des véhicules sondes

après qu'ils aient été *map-matchés* sur le lien approprié. De plus, la contrainte de la loi de conservation des véhicules est traitée comme une observation de variance quasi nulle.

Dans ce chapitre, de nombreuses simulations ont été effectuées avec un débit variable, un pourcentage variable de véhicules qui ont été choisis au hasard et considérés comme des véhicules sondes, un pourcentage variable de mouvements tournants et finalement un pourcentage variable de perturbations à mi-lien. La principale contribution dans ce chapitre est que l'approche stochastique UKF, même avec un petit nombre de véhicules sondes (1%), permet d'estimer le nombre de véhicules à l'amont de chaque lien et par conséquent le temps de parcours du réseau. En outre, l'UKF contient un modèle de prédiction qui, pour quelques minutes voire davantage tant que les flux restent constants, prédit le temps de parcours.

Chapitre 9. Conclusion et perspectives

SOMMAIRE

9.1 Conclusion et contribution - 167 -

9.2 Perspectives - 167 -

9.1 Conclusion et contribution

L'algorithme d'estimation du temps de parcours sur les réseaux urbains développé dans le cadre de cette thèse a été conçu pour apporter des réponses aux cas suivants, qui sont souvent mal traités par les méthodes classiques, généralement conçues pour les réseaux inter-urbains :

- proportion des flux significative de/vers des sources/fuites à mi-lien ;
- pourcentage inconnu des affectations aux intersections ;
- estimation en temps réel ;
- présence d'erreurs de comptage des boucles magnétiques ;
- présence d'erreur sur les temps de parcours des véhicules traceurs.

Une méthode simple et efficace sous certaines conditions pour estimer les temps de parcours à partir des informations de comptage de boucles magnétiques est la méthode analytique classique des stocks. Mais cette méthode est structurellement vulnérable à une déviation des courbes cumulatives dues aux fuites et sources à mi-lien et aux erreurs de comptage des détecteurs.

Ces problèmes peuvent être traités en fusionnant des données de type temps de parcours de véhicules traceurs avec les données de comptage des boucles magnétiques.

Une amélioration de la méthode des stocks exploitant des véhicules traceurs a déjà été proposée par [Bhaskar, 2009b] et porte le nom de CUPRITE. CUPRITE estime les temps de parcours en deux étapes. D'abord, les nombres de véhicules des détecteurs sont intégrés pour obtenir les courbes cumulatives à des emplacements différents sur le réseau, et ensuite les données des véhicules traceurs sont utilisées pour améliorer la précision de l'estimation du temps de parcours. Mais CUPRITE est sensible aux biais des données des véhicules traceurs. De plus, CUPRITE corrige la courbe cumulée en amont à l'instant d'occurrence du véhicule sonde, ainsi qu'avant et après cet instant. Toute correction avant l'occurrence des véhicules traceurs ne correspond pas à la contrainte d'estimation en temps réel. Egalement, CUPRITE exige un nombre minimum de véhicules traceurs pour être efficace, en particulier lorsque les pourcentages de mouvements tournants ne sont pas connus. Finalement, CUPRITE étant une

méthode déterministe, il est impossible de tenir compte des estimations statistiques des erreurs de mesure pour optimiser l'estimation finale, comme c'est le cas avec les méthodes de fusion bayésiennes.

Nous avons proposé dans cette thèse une approche stochastique afin d'apporter des améliorations aux points faibles mentionnés ci-dessus. Nous avons choisi une méthode d'estimation/fusion de données de type bayésienne, en l'occurrence un filtre de Kalman sans parfum (UKF), du fait du caractère implicite de l'expression du vecteur d'état qui rend impossible le calcul de la jacobienne nécessaire au filtre de Kalman étendu.

Dans ce filtre, les données de comptage des boucles et les données de temps de parcours des véhicules traceurs sont utilisées comme observations. Seules ces dernières permettent d'observer le temps de parcours moyen pour le tronçon de route considéré et ainsi de compenser les erreurs apportées par des défauts de comptage des boucles ou des sources/fuites à mi-lien. Les éléments du vecteur d'état du filtre sont : le temps de parcours, les nombres cumulés en amont, en aval, et des sources/fuites à mi-lien, le débit en amont, et l'historique des nombres cumulés en amont. Son modèle d'évolution est basée sur la méthode analytique classique des stocks.

Au cours de chaque étape de son développement, la performance de l'algorithme proposé a été soigneusement testée en utilisant des données simulées à l'aide du logiciel AIMSUN.

De nombreuses simulations ont ainsi été effectuées avec un débit variable en entrée du tronçon ou du réseau, un pourcentage variable de véhicules qui ont été choisis au hasard et considérés comme des véhicules traceurs, un pourcentage variable de mouvements tournants, et finalement un pourcentage variable de perturbations (quand elles sont simulées). La détermination de l'erreur d'observation à partir de données GPS *map-matchées* a été aussi développée.

Les tests de validation ont été effectués avec un pourcentage connu et inconnu des mouvements tournants aux intersections. Les résultats de notre filtre ont été systématiquement comparés à ceux de CUPRITE. Ces tests ont démontré que notre algorithme pouvait être appliqué avec succès pour une estimation précise et robuste du temps de parcours sur les réseaux urbains et que, par conséquent, les buts et les objectifs de cette recherche ont été atteints.

Les principales contributions de cette recherche peuvent être résumées comme suit :

- une nouvelle approche stochastique de type filtre de Kalman sans parfum (UKF) pour l'estimation de temps de parcours sur les réseaux urbains, exploitant les avantages des deux sources de données que sont les boucles magnétiques et les véhicules traceurs, et traitant leurs faiblesses respectives en les fusionnant de façon optimale. En tant qu'approche stochastique, notre UKF offre la possibilité d'introduire un modèle d'erreur pour le temps de parcours des véhicules traceurs, ainsi que pour les nombres de véhicules des boucles magnétiques, et d'en tenir compte pour la fusion ;

- le principal avantage de cette nouvelle méthode par rapport aux méthodes de fusion déterministes est qu'elle permet d'estimer, même avec un petit nombre de véhicules traceurs (1%), le nombre de véhicules à l'amont de chaque lien du réseau sans connaître les pourcentages d'affectation aux intersections et par conséquent sans observation directe de ces nombres ;

- de même, cette méthode permet aussi d'estimer le nombre de véhicules des sources et fuites à mi-lien sans aucune information sur ces nombres ;

- L'algorithme a prouvé sa robustesse vis-à-vis des fuites et des sources à mi-lien, des erreurs des détecteurs et des erreurs des véhicules traceurs ;

- l'estimation peut encore être rendue plus robuste par des tests statistiques, rendus possibles par le formalisme UKF. En effet, ce formalisme permet la détection et l'exclusion des valeurs aberrantes, comme des positions GPS qui seraient mal *map-matchées* ou des pannes des boucles ;

- l'UKF peut être appliquée en temps réel sur des réseaux urbains comme inter-urbains ;

- l'UKF est capable, sans observations, de prédire des temps de parcours, sous réserve que les débits restent constants.

9.2 Perspectives

La poursuite de ce travail consisterait à appliquer cet algorithme sur des données réelles en provenance d'un réseau urbain, comme par exemple celui de « Nantes Métropole ». Cette

application sur des données réelles nécessiterait bien entendu de développer des modèles réalistes des erreurs réellement commises par les 2 types de capteurs (boucles et véhicules traceurs) sur le réseau urbain considéré.

Une autre piste de poursuite consisterait à améliorer les performances en prédiction de notre méthode. Quand on veut utiliser un Filtre de Kalman en prédiction pure, seul le modèle d'évolution est alors utilisé étant donné que l'on ne connait pas les mesures futures. Dans la version actuelle de notre filtre, le modèle d'évolution est basé sur la méthode des stocks et considère que le débit a l'entrée des tronçons reste constant. Dans un contexte urbain très instable, cette prédiction peut s'avérer très fausse quand l'horizon est trop lointain.

Une façon d'améliorer cette prédiction, sous l'hypothèse bien entendu d'absence d' « accidents » de circulation imprévisibles, pourrait être d'exploiter l'historique de la semaine précédente par exemple pour améliorer l'équation d'évolution du débit en tenant compte des variations quotidiennes observées et en les modélisant avec des outils de type analyse paramétrique des séries chronologiques.

Annexe

Annexe : Map-matching

Il existe plusieurs méthodes de *map-matching*, et nous allons aborder l'approche géométrique et l'approche topologique [Quddus et Ochieng, 2007].

Approche géométrique

Cette approche consiste à prendre une représentation très simple du problème mettant en correspondance les données de position calculées par le récepteur GPS, (représentées par des points) avec les tronçons de route situés aux alentours de ces points (représentés par des segments) et susceptibles de correspondre à la route sur laquelle le véhicule est en train de circuler.

a. La méthode de la distance point à point

Les cartes numériques sont essentiellement constituées d'ensembles de points reliés entre eux. Cette méthode consiste à comparer le point correspondant aux coordonnées GPS fournies avec les différents points de la carte, et permet ainsi de sélectionner le candidat le plus proche du point GPS comme position supposée du véhicule sur la carte.

Cette approche qui est très simple a cependant de nombreux inconvénients :

- la résolution du positionnement se limite à l'espacement des différents nœuds de la carte, ce qui est en général très grossier ;
- l'apport de faux appariements dans des situations « *a priori* » peu problématiques. Nous illustrons cette proposition avec l'exemple figure 1 :

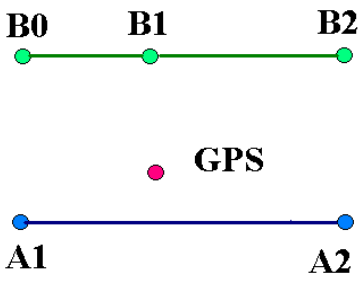


Figure 1 : Méthode de la distance point à point

En utilisant cette approche, le point GPS sera associé au point le plus proche qui est B1, alors que le véhicule se trouve plus vraisemblablement sur le segment [A0;A1]. Les résultats du positionnement dépendront donc beaucoup de la façon dont la cartographie est réalisée.

b. La méthode de la distance point à segment

Cette méthode consiste à mesurer la distance du point GPS aux différents segments ; on considère que le véhicule circule sur le segment le plus proche.

Cette méthode corrige effectivement le problème observé dans l'exemple précédent, puisque cette fois-ci le point GPS sera attribué au segment [A0;A1] et non plus au segment [B0;B2].

Cependant cette méthode présente des défauts. On peut par exemple observer de fréquentes instabilités dans l'attribution à un segment donné, comme illustré figure 2 :

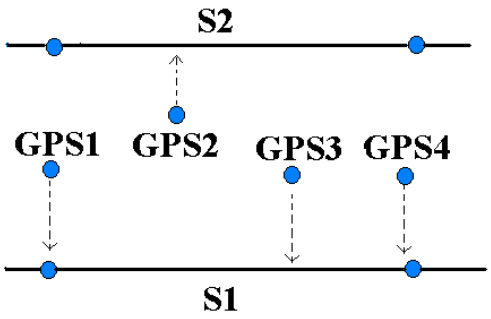


Figure 2 : Méthode de la distance point à segment.

La plupart des points GPS se voient affectés au segment S1, mais le point GPS2 qui se trouve plus près du segment S2 sera affecté à tort, puisqu'il est matériellement impossible au véhicule de passer temporairement d'un segment à l'autre.

De façon générale cette méthode ne tient pas assez compte de la cohérence spatiale et temporelle des données : en traitant les points GPS un à un sans tenir compte des données précédentes, il faut s'attendre à voir apparaître des incohérences dans le positionnement. La figure ci dessous (figure 3) en est une autre illustration :

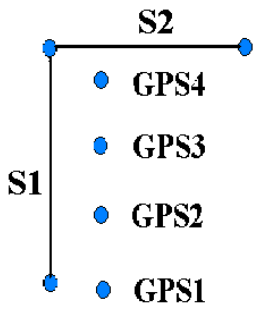


Figure 3 : méthode de la distance point à segment.

En observant les points GPS 1 à 4, on devine facilement que le véhicule est en train de circuler le long du segment S1. Or le point GPS4 pouvant se trouver plus proche de S2 que de S1, il y a un risque pour qu'il soit affecté à ce segment S2 au mépris de toute cohérence.

c. La méthode de la distance courbe à courbe

Dans cette méthode on garde en mémoire plusieurs points GPS reçus précédemment, et on les considère comme un ensemble à mettre en correspondance avec la carte. Alors dans un premier temps, on définit une mesure de distance entre arcs, sachant que l'espacement entre les nœuds des segments cartographiques peut être très différent de celui entre les points GPS. La distance peut alors être définie par la somme des distances entre chaque point et sa projection sur l'arc correspondant.

La prise en compte de la courbe ne règle pas les problèmes tellement mieux que précédemment, pour cela on passe aux approches topologiques.

Approche topologique

L'approche topologique consiste à tenir compte de la connectivité des différents arcs pour éventuellement résoudre certaines ambiguïtés, en supprimant les hypothèses qui ne seraient pas matériellement réalisables compte tenu de la topologie de la cartographie.

Cette méthode nécessite cependant un "point d'appui" sous la forme d'une position de référence que l'on considère comme fiable, et à partir de laquelle se développent les raisonnements. Disons que l'initialisation d'une approche topologique est critique.

Considérons, par exemple, le cas de la figure 4 : les méthodes géométriques auront tendance à attribuer la plupart des points GPS au segment S3. Mais si l'on considère que le

véhicule se trouve initialement et avec certitude au point P0, il est alors matériellement au segment S1.

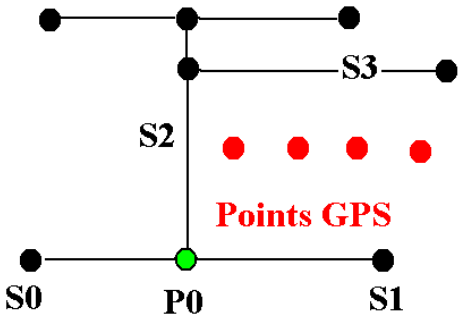


Figure 4 : méthode topologique

Références

[Barceló, 2010] Barceló J., « Fundamentals of traffic simulation », Springer, 2010.

[Barceló, 2010] Barceló J., « A Kalman filter approach for dynamic OD estimation in corridors based on Bluetooth and Wifi data collection », 12th WCTR, July 11-15, Lisbon, Portugal, 2010.

[Ben Akiva, 1994] Ben Akiva M.E., « A dynamic traffic model system for ARMS/ATIS operations », IVHS journal, vol. 2. pp. 1-19 MITTNS, 1994.

[Bosru, 2004] Bosru M., « les apports des nouvelles technologies pour le recueil de données traffic », travail de fin d'étude, ENTPE, CERTU, 2004.

[Bhaskar et al., 2008] Bhaskar A., Chung E., Dumont A.G. & de Mouzon O., « Analytical modelling and sensitivity analysis for travel time estimation on signalized urban networks », Transportation Research Board, 2008.

[Bhaskar et al., 2009a] Bhaskar A., Chung E. & Dumont A.G., « Estimation of travel time on urban networks with midlinks sources and sinks », Transportation Research Board, 2121, pp. 41-54, 2009.

[Bhaskar, 2009b] Bhaskar A., « A methodology (CUPRITE) for urban network travel time estimation by integrating multisource data », travail de thèse, École polytechnique fédérale de Lausanne, Suisse, 2009.

[Bhaskar et al., 2011] Bhaskar A., Chung E. & Dumont A.G., « Arterial travel time estimation: revisiting the classical procedure », in Australasian Transport Research Forum Proceedings, Adelaide, Australia, 2011.

[Bonvalet et Robin-Prévallée, 1987] Bonvalet F. & Robin-Prévallée Y., « Mise au point d'un indicateur permanent des conditions de circulation en Île-de-France », Transport Environnement Circulation, pp. 84-85, 1987.

[Bozic, 1994] Bozic S.M., « Digital and Kalman filtering: an introduction to discrete-time filtering and optimum linear », Halsted Press, New York, 1994.

[Brown et Hwang, 1997] Brown R.G. & Hwang P.Y.C., « Introduction to random signals and applied Kalman filtering », Wiley, New York, 1997.

[Buisson, 1996] Buisson C., « Analyse bibliographique et première classification des algorithmes de simulation de l'écoulement du trafic routier », Recherche Transports Sécurité, vol. 53, pp. 35-48, 1996.

[Buisson et Lessort, 2010] Buisson C. & Lessort J.B., « Comprendre le trafic routier, méthodes et calculs », Edition Certu, 2010.

[Choi et Chung, 2002] Choi K. & Chung A., « Data fusion algorithm for estimatingl travel time », in Journal of Intelligent Transportation Systems, 7, pp. 235–260, 2002.

[D'Urso et Gastaldi, 2000] D'Urso P. & Gastaldi T., « A least-squares approach to fuzzy linear regression analysis », Computational Statistics and Data Analysis, 34, pp. 427-440, 2000.

[Dailey, 1999] Dailey D.J., « A statistical algorithm for estimating speed from single loop volume and occupancy measurements », in Transportation Research Part B: Methodological, 33, pp. 313-322, 1999.

[Davis et al., 1990] Davis G.A., Nihan N.L., Hamed M.M. & Jacobson L.N., « Adaptive forecasting of freeway traffic congestion », Transportation Research Record, 1287, pp. 29-33, 1990.

[Dempster, 1968] Dempster A., « A generalization of bayesian inference », Journal of Royal Statistical Society, Serie B 30, pp. 205-247, 1968.

[Denœux, 1997] Denœux T. « Analysis of evidence-theoretic decision rules for pattern classification », Pattern Recognition, 30(7), pp. 1095-1107, 1997.

[Denœux, 2000] Denœux T. « A neural network classifier based on Dempster-Shafer theory », IEEE Transactions on Systems, Man and Cybernetics, Part A, 30(2), 2000.

[El Faouzi, 2000a] Faouzi N.-E., « Fusion de données pour l'estimation des temps de parcours via la théorie de l'évidence », Recherche Transport Sécurité (68), pp. 15-30, 2000.

[El Faouzi, 2000b] El Faouzi N.-E., « Fusion de données : Concepts et méthodes», Rapport de Synthèse, n° 2002, INRETS-LICIT, 2000.

[El Faouzi et Lessort, 2001] El Faouzi N.-E. & Lessort J.B., « Fusion de données pour l'estimation du temps de parcours, réflexions théoriques et méthodologies », Rapport LICIT, n° 0101, 2001.

[El Faouzi, 2004] El Faouzi N.-E., « Data fusion in road traffic engineering : an overview », in Proceedings of the SPIE - International Society for Optical Engineering, Orlando, Florida, 2004.

[El Faouzi et Lefèvre, 2006] El Faouzi N.-E. & Lefèvre E. « Bayesian and evidential approaches for traffic data fusion : methodological issues and case study », presented at the 85th Transportation Research Board Meeting, 2006.

[Ellouni, 1996] Ellouni N., « Modélisation et commande du trafic sur un corridor : application de la méthode LP », Thèse de l'Université Paris XI Orsay, 1996.

[Givens et Roback, 1999] Givens G.H. & Roback P.J., « Logarithmic pooling of priors linked by a deterministic simulation model », Journal of Computational and Graphical Statistics, pp. 452–478, 1999.

[Godlewski et Raviart, 1990] Godlewski E. & Raviart P.A., « Hyperbolic systems of conservation laws », Paris, France : Ellipses, 1990.

[Greenshield, 1935] Greenshield B.D., « A study of traffic capacity », Proceedings of the Highway Research Board, vol. 14, pp. 448-477, 1935.

[Hall, 1997] Hall F.L., « Traffic stream characteristics », in Traffic flow theory: A state of the art report - revised monograph on traffic flow theory, N. H. Gartner, C. Messer and A. K. Rathi, Eds., Oak Ridge, Tennessee, Oak Ridge National Laboratory, 1997.

[Hamed et al., 1995] Hamed M.M., Al-Masaeid H.R. & Bani Said Z.M., « Short-term prediction of traffic volume in urban arterials », Journal of Transportation Engineering, pp. 249-254, 1995.

[Herring et al., 2010] Herring R., Hofleitner A, Abbeel P. & Bayen A. « Estimating arterial traffic conditions using sparse probe data », IEEE, pp.929-936, 2010.

[Hinsbergen et al., 2010] Hinsbergen C.P., Schreiter T., Zuurbier F.S.,V Lint J.W.C., & Van Zuylen H.J., « Fast traffic state estimation with the localized extended Kalman filter », in Proceedings of the IEEE, ITSC, pp. 917-922, Madeira, Portugal, 2010.

[Jiang et al., 2002] Jiang R., Wu Q.S. & Zhu Z.J. « A new continuum model for traffic flow and numerical testes », Transportation Research Part B, n° 36,. pp. 405-419, 2002.

[Johnson, 1999] Johnson B., « Keeping the world flowing », Traffic Technology International, pp.38-40, Juin-Juillet, 1999.

[Julier et al., 1995] Julier S.J., Uhlmann J.K. & Durrant-whyte H.F., « A new approach for filtering non linear system », in Proceeding of the American Control Conference, pp. 1628-1632, 1995.

[Julier et Uhlmann, 1997] Julier S.J. & Uhlmann J.K., « A new extension of the Kalman filter to non linear systems », in Proceedings of the 11th International Symposium on Aerospace/Defence Sensing, Simulation and Controls, pp. 54-65, 1997.

[Julier et al., 2000] Julier S.J., Uhlmann J.K. & Durrant-Whyte H.F., « A new method for the non linear transformation of means and covariances in filters and estimators », IEEE, Trans Automat Control, pp. 477-482, 2000.

[Julier, 2002] Julier S.J., « The scaled unscented transformation", In Proceedings of the American Control Conference Anchorage, pp. 4555-4559, 2002.

[Julier et Uhlmann, 2002] Julier S.J. & Uhlmann J.K., « Reduced sigma point filters for the propagation of means and covariances through non linear transformations », in Proceeding of the American Control Conference, pp. 887-892, 2002.

[Keen et al., 1990] Keen K.G., Schofield & G.C. Hay, « Ramp metering access control on M6 motorways », Proceeding of the 2nd IEE Intern., International Conference on Road Traffic Control, London, 1986.

[Kell et al., 1990] Kell J.H., Fullerton I.J. & Mills M.K., « Traffic detector handbook », 2nd Edition, Federal Highway Administration, U.S. Department of Transportation, Washington D.C, 1990.

[Leclercq, 2002] Leclercq L., « Modélisation dynamique du trafic et applications à l'estimation du bruit routier », Thèse en Génie Civil, Villeurbanne, INSA de Lyon, 2002.

[Lefèvre et El Faouzi, 2006] Lefèvre E. & El Faouzi N.-E., « Estimation et fusion des temps de parcours routiers par la théorie de l'évidence », 2006.

[Lighthill et Whitham, 1955] Lighthill M.J. & Whitham G.B., « A theory of traffic flow on long crowded roads », in Proceedings of the Royal Society A, vol. 229 , pp. 317-345, 1955.

[Maybeck, 1979] Maybeck P.S., « Stochastic models, estimation, and control », Academic Press, Inc., New York, 1979.

[Middleton et al., 2002] Middleton D., Gopalakrishna D. & Raman M., « Advances in traffic data collection and management », Technical Report BAT-02-006, Texas Transportation Institute, Cambridge systematics, Inc., Washngton, DC, December, 2002.

[Nam et Drew, 1999] Nam D.H. & Drew D.R., « Automatic measurement of traffic variables for intelligent transportation systems applications », in Transportation Research Part B: Methodological, n° 33, pp. 437-457, 1999.

[Oh et al., 2003] Oh J.S., Jayakrishan R. & Rechker W., « Section travel time estimation from point detection data », in proceedings of the 82th Annual Meeting of Transportation Research Board, Washington, D.C., U.S.A, 2003.

[Olkin, 1992] Olkin I., « Meta-analysis: Methods for combining independent studies », Statistical Science, n° 7(2), pp. 226-236, 1992.

[Papageorgiou et al., 1990] Papageorgiou M., Blosseville J.-M. & Haj-Salem H., « Dynamic modeling, assignment, and route guidance in traffic networks », Transportation Research part B, vol. 24B, n° 6, pp. 471-495, 1990.

[Payne, 1979] Payne H.J., « A critical Review of a Macroscopic freeway model », Res. Direc. in Comput. Cont. of Urb. Traf. Sys. Am. Soc. Cov. Eng., New York, pp. 251-265, 1979.

[Quddus et Ochieng, 2007] Quddus M.A. & Ochieng W.Y., « Current map matching algorithms for transport applications: State-of-the-art and future research directions », in Transportation Research Part C: Emerging Technologies, n° 15, pp. 312-328, 2007.

[Rascle, 2002] Rascle M., « An improved macroscopic model of traffic flow: derivation and links with Lighthill-Whitham Model », Mathematical and Computer Modelling, n° 35. pp. 581-590, 2002.

[Rice et Van Zwet, 2004] Rice J. & Van Zwet E. « A simple and effective method for predicting travel times on freeways », IEEE, Transactions on Intelligent Transportation Systems, n° 5, pp. 200-207, 2004.

[Richards, 1956] Richards P.I., « Shockwaves on the highway », Operations Research B, n° 22, pp. 81-101, 1956.

[Shafer,1976] Shafer G., « A Mathematical Theory of Evidence », Princeton, New Jersey, PrincetonUniversity Press, 1976.

[Simon, 2006] Simon D., « Optimal state estimation, Kalman, H∞, and nonlinear approaches », Wiley, New York, 2006.

[Srinivasan et Jovanis, 2004] Srinivasan K.K. & Jovanis P.P., « Determination of number of probe vehicles required for reliable travel time measurement in urban network », in Transportation Research Record, pp. 15-22, Washington, D.C, 2004.

[Tanaka et al., 1982] Tanaka H., Uejima S. & Asai K., « Fuzzy linear regression model », IEEE, Trans. Systems Man Cybernet n° 12, pp. 903-907, 1982.

[Tanaka, 1987] Tanaka H., « Fuzzy data analysis by possibilistic linerar models », Fuzzy Sets and Systems n° 24, pp. 363-375, 1987.

[Turner et Holderner, 1995] Turner S. M. & Holderner D. J., « Probe vehicle sample sizes for real-time information », the Houston experience, in Dailey Daniel, J. & Haselkorn Mark, P. (Eds.), Vehicle Navigation and Information Systems Conference (VNIS), Seattle, WA, USA, IEEE, 1995.

[Wan et Vander Merwe, 2000] Wan E.A. & Vander Merwe R., « The unscented Kalman filter for non linear estimation », in Proceedings of Adaptive Systems for Signal Processing, Communication and Control (AS-SPCC) Symposium, Alberta, pp. 153-156, 2000.

[Wan et Vander Merwe, 2001] Wan E.A. & Vander Merwe R., « The unscented Kalman filter », In Haykin S (ed) Kalman Filtering and Neural Networks, Wiley, NewYork, 2001.

[Woods et al., 1994] Woods D.L., Cronin B.P. & Hamm R.A., « Speed measurement with inductance loop speed traps. », Texas Transportation Institute, Research Report FHWA/TX-95/1392-8, Texas A&M, College Station, 1994.

[Yang et al., 1999] Yang Q., Koutsopoulos H.N. & Ben Akiva M.E., « A simulating laboratory for evaluating dynamic traffic management systems », ASCE Journal of Transportation Engeering, July/August, 1999.

[Ye et al., 2006] Ye Z., Zhang Y. & Middleton D., « An unscented Kalman filter method for speed estimation using single loop detector data », Transportation Research Record, Transportation Research Board, Washington, DC, pp. 117-25, 2006.

[Zhang, 2002] Zhang H.M., « A non-equilibrium traffic model devoid of gas-like behavior », Transportation Research B, n° 36, pp. 275-290, 2002.

Publications

[Hage et al., 2011] Hage R.M., Bétaille D., Le Carpentier E., Peyret F. & Meizel D. « Link travel time estimation in urban areas by detectors and probe vehicles fusion», 2nd International Conference on Models and Technologies for Intelligent Transportation Systems, Louvain, Belgique, Juin, 2011.

[Hage et al., 2012] Hage R.M., Bétaille D., Peyret F. & Meizel D. Unscented Kalman filter for urban network travel time estimation », 15th meeting of the EURO Working Group on Transportation, Paris, September 2012.

[Hage et al., 2012] Hage R.M., Bétaille D., Peyret F., Meizel D. & Smal J.C., « Unscented Kalman filter for urban link travel time estimation with mid-link sinks and sources », Intelligent Transportation Systems Conference - ITSC, IEEE, Anchorage, USA, September 2012.

Estimation du temps de parcours d'un réseau urbain par fusion de données de boucles magnétiques et de véhicules traceurs – Une approche stochastique avec mise en œuvre d'un filtre de Kalman sans parfum

La notion de temps de parcours est une information simple à intégrer par les usagers des transports et a le potentiel, via des systèmes d'assistance, de réduire la congestion aussi bien de manière temporelle que spatiale. Cette thèse porte sur la problématique de l'estimation du temps de parcours en ville, où la circulation se caractérise par un débit variable et un flux significatif de/vers des voies transversales non équipées de capteurs de trafic, flux qui agit comme perturbation de l'équation de conservation des véhicules. La bibliographie présente des approches déterministes et stochastiques, utilisant, selon les situations expérimentales, soit des boucles inductives, soit des véhicules traceurs, soit les deux. Nous proposons une méthode stochastique du type filtrage de Kalman sans parfum. Son modèle dynamique est basé sur la méthode analytique classique des stocks qui considère le décalage temporel entre les cumuls de véhicules entrants et sortants dans chaque tronçon du réseau. La formulation de ce modèle n'est pas explicite, ce qui justifie l'utilisation d'un filtre sans parfum. Les observations utilisées sont à la fois les boucles magnétiques à la sortie des tronçons et les véhicules traceurs après map-matching sur le réseau. L'algorithme est applicable en temps réel, sans connaissance a priori des affectations aux carrefours, et sous conditions éventuelles de perturbation. En sortie de filtre, on dispose des variances d'estimation, ce qui mesure la confiance dans le temps de parcours estimé et permet de rejeter des mesures aberrantes. La validation a été montrée sur la base de simulations sur un réseau simple à trois intersections, avec des flux variables et des perturbations, et pour quelques pourcents de véhicules traceurs.

Mots clés : trafic, temps de parcours, boucle magnétique, véhicule traceur, map-matching, filtre de Kalman sans parfum, réseau urbain.

Urban network travel time estimation by loop detectors and probe vehicles data fusion – A stochastic approach by unscented Kalman filtering

Travel time information is easy to be perceived by users and has the potential to reduce congestion on both temporal and spatial scales thanks to traveller assistance systems. This thesis addresses the problem of estimating the travel time in city, where traffic is characterized by variable flow and also significant perturbation from/to mid-link source/sink that are not equipped with traffic sensors. This flow disturbs the equation of conservation of vehicles. The bibliography contains deterministic and stochastic approaches, using, in specific experimental situations, either inductive loops or probe vehicles, or both. We propose a stochastic approach based on unscented Kalman filtering. Its dynamic model is based on the classical analytical procedure that considers the time difference between the total of incoming and outgoing vehicles for each segment of the network. The formulation of this model is not explicit, which justifies the use of an unscented filter. The observations are both loop counters at the exit sections and probe vehicles map-matched to the network. The algorithm is applicable in real time, without prior information on the percentage of turning movements at intersections, and under any condition of perturbation. The variance estimation output by the filter measures the confidence in the estimated travel time and allows the rejection of outliers. The validation was shown based on simulations on a simple network with three intersections, variable entering flows, variable percentage of perturbation, and a variable percentage of probe vehicles

Key words: traffic, travel time, loop detector, probe vehicle, map-matching, unscented Kalman filter, urban network.