



**HAL**  
open science

## Traitements des données de trafic : besoins, état de l'art, exemple de mise en oeuvre

Laurent Bréheret, Frédéric Schettini, Eric Bernauer, Magali Barbier

### ► To cite this version:

Laurent Bréheret, Frédéric Schettini, Eric Bernauer, Magali Barbier. Traitements des données de trafic : besoins, état de l'art, exemple de mise en oeuvre. [Rapport de recherche] Centre d'études sur les réseaux, les transports, l'urbanisme et les constructions publiques (CERTU). 2000, 100 p., figures, tableaux, graphiques, 33 références bibliographiques. hal-02163332

**HAL Id: hal-02163332**

**<https://hal-lara.archives-ouvertes.fr/hal-02163332>**

Submitted on 24 Jun 2019

**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.

PREDIT

Programme national de recherche et  
d'innovation dans les transports terrestres

# Traitements des données de Trafic

*Besoins*

*État de l'art*

*Exemple de mise en œuvre*



## NOTICE ANALYTIQUE

<b>Organismes commanditaires :</b> DRAST : Direction de la Recherche et des Affaires Scientifiques et Techniques Tour Pascal B – 92055 LA DEFENSE CEDEX Lettre de commande N° 98 MT 125			
<b>Titre :</b> <b>Traitements des données de trafic</b>			
<b>Sous-titre :</b> Besoins, État de l'art, Exemple de mise en œuvre		<b>Date d'achèvement :</b> <b>Janvier 2000</b>	<b>Langue :</b> <b>Français</b>
<b>Organisme auteur</b> <b>SODIT</b> 2, avenue Édouard Belin 31400 Toulouse		<b>Rédacteurs</b> Laurent BRÉHERET, Frédéric SCHETTINI, et Éric BERNAUER, SODIT  Magali BARBIER, ONERA - CERT	<b>Relecture assurance qualité faite par :</b> Sylvie CHAMBON, CERTU
<b>Remarques préliminaires :</b> Cette étude a été réalisée dans le cadre du Programme national de Recherche et d'Innovation dans les Transports Terrestres - PREDIT / 1996 - 2000 pour le compte du groupe thématique 4.2 Route Intelligente, sur le thème de l'amélioration des traitements de données de trafic.			
<b>Résumé :</b> Cette étude s'est intéressée particulièrement à la mise en œuvre d'un outil de traitement de données de trafic pour les besoins du système global de gestion des déplacements de l'agglomération toulousaine. Cependant, la problématique du traitement de données et les solutions proposées sont largement transposables à d'autres contextes locaux. Ce rapport présente les différents acteurs gestionnaires de déplacements et leurs relations. Leurs besoins sont analysés, qui couvrent à la fois les services aux usagers et les besoins propres aux exploitants. Une revue des systèmes de traitements de données utilisés aujourd'hui permet de présenter différentes techniques de recueil et de fusion de données. Les spécifications techniques d'un système de traitement de données sont ensuite largement détaillées avec les méthodes utilisées dans les trois modules clés que sont : la qualification des données, le filtrage et l'estimation par fusion de données, la prédiction des états du réseau. La mise en œuvre du système proposé est ensuite présentée, avec le contenu de la base de données, les fichiers utilisés et l'implantation matérielle. Le document est complété par des références bibliographiques et des approfondissements techniques sur les méthodes de traitement de données trafic proposées. Ce rapport s'adresse particulièrement aux gestionnaires de réseaux de circulation, qui reconnaîtront dans cette étude la problématique du traitement des données à laquelle ils sont quotidiennement confrontés. Ce document constitue pour eux des pistes de réflexion et un point de vue actuel sur les systèmes de traitement de données de trafic.			
<b>Mots clés :</b> données de trafic, exploitation de la route, algorithme, fusion de données, filtrage, qualification, recherche et développement		<b>Diffusion :</b> Libre	
<b>Nombre de pages :</b> 92 pages.	<b>Prix :</b> 100 FF – 15,24 €	<b>Confidentialité :</b> NON	<b>Bibliographie :</b> OUI



# Sommaire

<b>1</b>	<b>OBJECTIFS DE L'ÉTUDE</b>	<b>1</b>
<b>1.1</b>	<b>TITRE ET CONTEXTE DE L'ÉTUDE</b>	<b>1</b>
<b>1.2</b>	<b>ORGANISATION DE L'ÉTUDE</b>	<b>1</b>
1.2.1	ANALYSE DES BESOINS	1
1.2.2	SPÉCIFICATIONS DES MÉTHODES ET OUTILS	1
<b>1.3</b>	<b>STRUCTURE DU RAPPORT</b>	<b>2</b>
<b>2</b>	<b>ORGANISATION DES TRANSPORTS DANS L'AGGLOMÉRATION DE TOULOUSE</b>	<b>3</b>
<b>3</b>	<b>ANALYSE DES BESOINS</b>	<b>5</b>
<b>3.1</b>	<b>MÉTHODE : ENTRETIENS ET QUESTIONNAIRE</b>	<b>5</b>
3.1.1	CONTENU DU QUESTIONNAIRE	5
<b>3.2</b>	<b>RÉSULTATS DE L'ANALYSE DES ENTRETIENS</b>	<b>6</b>
3.2.1	LES FONCTIONS D'EXPLOITATION	6
3.2.2	LES USAGERS DES RÉSEAUX	7
3.2.3	LES DIFFÉRENTES DONNÉES MESURÉES	8
3.2.4	LES TRAITEMENTS RÉALISÉS SUR LES DIFFÉRENTES DONNÉES	10
3.2.5	MESSAGES ET DIFFUSION D'INFORMATION	11
3.2.6	LES BESOINS EXPRIMÉS PAR LES OPÉRATEURS	11
<b>3.3</b>	<b>COMPLÉMENT SUR LES BESOINS EN TRAITEMENT DE DONNÉES</b>	<b>12</b>
<b>3.4</b>	<b>SYNTHÈSE SUR LES BESOINS DES EXPLOITANTS</b>	<b>15</b>
<b>4</b>	<b>ÉTAT DE L'ART SUR LES SYSTÈMES DE FUSION DE DONNÉES</b>	<b>17</b>
<b>4.1</b>	<b>DÉFINITION ET PROBLÉMATIQUE</b>	<b>17</b>
<b>4.2</b>	<b>LE RECUEIL DE DONNÉES</b>	<b>18</b>
<b>4.3</b>	<b>LES MÉTHODES EXISTANTES</b>	<b>19</b>
<b>4.4</b>	<b>LES SYSTÈMES EXISTANTS</b>	<b>21</b>
<b>5</b>	<b>SPÉCIFICATIONS D'UN SYSTÈME DE TRAITEMENT DE DONNÉES</b>	<b>23</b>
<b>5.1</b>	<b>CAHIER DES CHARGES</b>	<b>23</b>
5.1.1	CONCEPT GÉNÉRAL	24
5.1.2	LA BASE DE DONNÉES TRAFIC	25
5.1.3	LES FONCTIONS PRINCIPALES	27
5.1.4	LE SYSTÈME DE FUSION DE DONNÉES ÉTUDIÉ	29
5.1.5	LES NOTATIONS UTILISÉES	30
<b>5.2</b>	<b>LES FONCTIONS</b>	<b>30</b>
5.2.1	LA QUALIFICATION DES DONNÉES	31
5.2.2	LE FILTRAGE ET L'ESTIMATION PAR FUSION DE DONNÉES	36
5.2.3	LA PRÉDICTION DES ÉTATS DU RÉSEAU	39
<b>6</b>	<b>MISE EN ŒUVRE</b>	<b>42</b>

<b>6.1</b>	<b>DESCRIPTION ET CONSTITUTION DES BASES DE DONNÉE</b>	<b>42</b>
6.1.1	LES DONNÉES D'ENTRÉE	42
6.1.2	LES DONNÉES DE SORTIE	43
<b>6.2</b>	<b>CALIBRAGE</b>	<b>45</b>
6.2.1	CALIBRAGE DES VARIABLES CLÉS	46
6.2.2	CONSTITUTION DE LA BASE DE DONNÉES HISTORIQUES	47
<b>6.3</b>	<b>VALIDATION DES MODÈLES</b>	<b>47</b>
<b>6.4</b>	<b>IMPLANTATION INFORMATIQUE</b>	<b>48</b>
6.4.1	LA CONFIGURATION LOGICIELLE	48
6.4.2	LA CONFIGURATION MATÉRIELLE	49
6.4.3	ÉLÉMENTS DE PERFORMANCES	49
<b>7</b>	<b>CONCLUSIONS</b>	<b>51</b>

---

## **8 ANNEXES** **53**

---

<b>8.1</b>	<b>QUESTIONNAIRE</b>	<b>53</b>
<b>8.2</b>	<b>ÉLÉMENTS DE BASE DE LA THÉORIE DU TRAFIC</b>	<b>58</b>
<b>8.3</b>	<b>COMPLÉMENTS SUR LA MÉTHODE DE QUALIFICATION DES DONNÉES DE COMPTAGES</b>	<b>60</b>
8.3.1	TESTS SUR LES DONNÉES ABERRANTES	60
8.3.2	ALGORITHME DE DÉTECTION ET DE CORRECTION DES MESURES	61
8.3.3	RÉSULTATS D'EXPÉRIMENTATION	62
<b>8.4</b>	<b>COMPLÉMENTS SUR LA MÉTHODE DE QUALIFICATION DES DONNÉES DE VÉHICULES TRACEURS</b>	<b>65</b>
8.4.1	EXEMPLE D'UN CHAÎNON À DEUX VOIES	65
8.4.2	RÉSULTATS DES TESTS	67
<b>8.5</b>	<b>COMPLÉMENTS SUR LA MÉTHODE DE FILTRAGE ET D'ESTIMATION</b>	<b>68</b>
8.5.1	MODÈLES UTILISÉS EN ZONE URBAINE	68
8.5.2	MODÈLES DE TRAFIC ADAPTÉS AUX TRAJETS	70
8.5.3	MÉTHODES DE RÉOLUTION	72
<b>8.6</b>	<b>COMPLÉMENTS SUR LA MÉTHODE DE PRÉDICTION</b>	<b>77</b>
8.6.1	MÉTHODE BASÉE SUR LA COMBINAISON KALMAN / DONNÉES HISTORIQUES	77
8.6.2	MÉTHODE NON CONVENTIONNELLE BASÉE SUR LA THÉORIE DES RÉSEAUX DE NEURONES	78
<b>8.7</b>	<b>DESCRIPTION DU CONTENU DE LA BASE DE DONNÉES</b>	<b>82</b>
8.7.1	LA BASE DE DONNÉES GÉOGRAPHIQUES	82
8.7.2	LA BASE DE DONNÉES HISTORIQUES	84
8.7.3	LA BASE DE DONNÉES DYNAMIQUES	85
8.7.4	LA BASE DE DONNÉES DE SORTIE	85

## **9 RÉFÉRENCES** **88**

---

## **10 SIGLES ET ACRONYMES** **90**

---

## **LISTE DES FIGURES** **91**

---

## **LISTE DES TABLEAUX** **92**

---





## **Remerciements**

Les auteurs remercient la ville de Toulouse et la Direction Départementale de l'Équipement de la Haute-Garonne pour leur contribution l'exécution de cette étude.

# 1 Objectifs de l'étude

## 1.1 Titre et contexte de l'étude

**Mise en œuvre d'un outil de traitement de données trafic en vue de la mise en place du système global d'information sur les déplacements dans l'agglomération toulousaine.**

Les autorités de transport de l'agglomération toulousaine collaborent par voie de convention, dans le souhait de partager des objectifs stratégiques bénéfiques aux usagers (convention <sup>1</sup>SGGD). Dans le cadre des actions du SGGD, ces organismes ont initié plusieurs réflexions, concernant l'architecture des systèmes de gestion d'une part, et d'autre part, la réalisation d'un système global d'information sur les déplacements. Les études menées concluent a priori sur la mise en place d'un système centralisé, recueillant les données des différents opérateurs, avant de les redistribuer vers des organes d'exploitation ou de diffusion, afin de constituer ainsi les prémisses d'une fonction «agence de presse» pour les informations sur les déplacements.

L'objectif de cette étude est de définir, sur le plan technique, les fonctions de fusion de données et de traitement d'informations nécessaires aux fonctions d'exploitation et permettant une continuité dans la fonction d'information.

## 1.2 Organisation de l'étude

Cette étude est décomposée en deux phases : une phase d'analyse des besoins et une phase de spécifications.

### 1.2.1 Analyse des besoins

Deux partenaires de ce projet sont directement et quotidiennement impliqués par le problème du traitement des données relatives au trafic : la Ville de Toulouse et la Direction Départementale de l'Équipement de la Haute-Garonne. Ces deux organismes sont partenaires du SGGD et sont impliqués dans divers projets communs, tels que le Campus Trafic, la mise au point d'un système global d'information et la gestion des crises. Bien que très similaires, leurs besoins diffèrent dans les approches ou leurs modes de résolution. L'objet de cette première phase est de déterminer précisément les besoins des deux exploitants en terme de traitement de données, soit dans un but de gestion du trafic, soit dans un but d'information.

### 1.2.2 Spécifications des méthodes et outils

L'objet de cette phase est de fournir les spécifications pour les méthodes et les outils à développer. Elle s'appuie sur l'analyse des besoins effectuée lors de la phase précédente.

Ces spécifications détaillent en particulier les fonctions de fusion de données, celles de filtrage des données et de qualification des mesures, les modèles de trafic applicables, la détection des pannes et les modes de corrections, l'élaboration d'indicateurs de trafic et la prédiction de ces indicateurs à court terme.

---

<sup>1</sup> SGGD : Système de Gestion Globale des Déplacements

### **1.3 Structure du rapport**

Le rapport est organisé de la façon suivante :

- Le chapitre 2 présente l'organisation des transports dans l'agglomération toulousaine,
- Le chapitre 3 consiste en une analyse des besoins des deux partenaires du projet qui sont directement et quotidiennement concernés par le traitement des données de trafic : la ville de Toulouse et la Direction Départementale de l'Équipement de la Haute-Garonne.
- Le chapitre 4 présente un état de l'art des systèmes de recueil et de fusion des données,
- Le chapitre 5 décrit les spécifications techniques pour le développement et la mise en œuvre d'un système de traitement de données capable de proposer des fonctions d'exploitation avancées aux gestionnaires de trafic et permettant de diffuser des informations fiables aux usagers de la route,

Le chapitre 6 conclut ce rapport.

Les chapitres 7 et 8 contiennent respectivement des références bibliographiques et des compléments techniques sur les méthodes proposées pour le système de traitement de données.

## 2 Organisation des transports dans l'agglomération de Toulouse

Dans le cadre de ce projet, il nous semble important de présenter les différents acteurs impliqués dans les transports ainsi que leurs relations.

Il n'y a pas d'autorité de transport supérieure couvrant toute l'agglomération de Toulouse pour tous les modes. Toutefois, le cas de Toulouse est exemplaire de par l'existence d'une structure de coopération commune : le Système de Gestion Globale des Déplacements (**SGGD**), qui regroupe en son sein la plupart des acteurs locaux du transport. Ceux-ci peuvent être classés selon leur rôle dans l'organisation des transports :

- **Les gestionnaires de voiries** : l'état (représentant en particulier ERATO) ; le Département de la Haute-Garonne ; la Ville de Toulouse ; la Direction Départementale de l'Equipe (DDE 31) et les Autoroutes du Sud de la France (ASF) ;
- **Les autorités de police** : la Police Nationale et la Gendarmerie Nationale
- **L'autorité organisatrice des transports collectifs** : le SMTC
- **Les organismes de service ou de recherche** : le CETE du sud-ouest, l'INRETS et l'ONERA-CERT partenaires dans l'organisation de la ZELT
- **Les autorités financières** : la région Midi-Pyrénées ; le district de Toulouse
- **L'opérateur de transport public** : la SEMVAT

Le projet SGGD a été lancé en 1993. La plupart des aspects organisationnels, tels que les liens à établir entre les différents acteurs ont été identifiés. D'un point de vue légal, un accord de coopération a été signé par tous les partenaires en août 1996, fixant les objectifs globaux communs. L'information sur les déplacements pour les usagers est un des domaines d'intérêt pour les membres du SGGD, et fait partie des objectifs fixés dans cette convention.

Les relations entre les différents acteurs locaux sont gérées dans le cadre de la convention SGGD. D'un point de vue organisationnel, les partenaires du SGGD se sont accordés pour coordonner leurs activités, en particulier en regroupant les individus dans un site commun : le Campus Trafic de Montblanc. Ce Campus Trafic abrite :

- les bureaux du service de circulation de la Ville de Toulouse avec le PC CAPITOUL,
- les bureaux du CIGT (Centre d'Ingénierie et de Gestion du Trafic) de la DDE avec son PC,
- le PC ZELT et ses bureaux techniques,
- le PC de gestion des taxis toulousains
- les locaux pour les forces de l'ordre

Le PC SEMVAT de régulation des bus et le PC ASF sont situés à proximité.

D'un point de vue technique, l'architecture système du SGGD est en cours de définition dans le cadre de groupes de travaux. A priori, les différents postes et ordinateurs seront reliés à un ordinateur central, qui fera fonction de serveur de données commun.

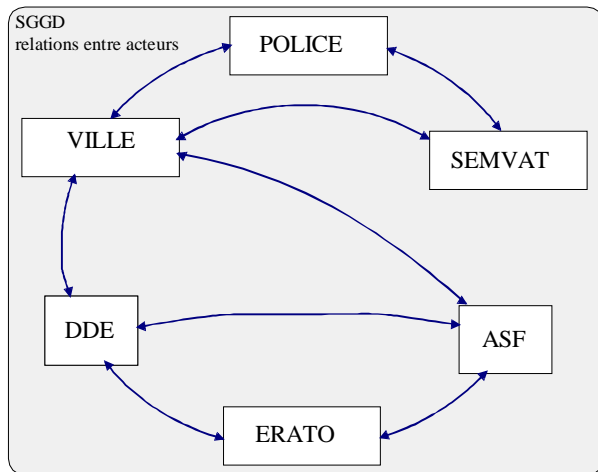


Figure 1 : Relations entre les acteurs du transport dans l'agglomération toulousaine

Le schéma ci-dessous présente la couverture géographique des gestionnaires de voiries.

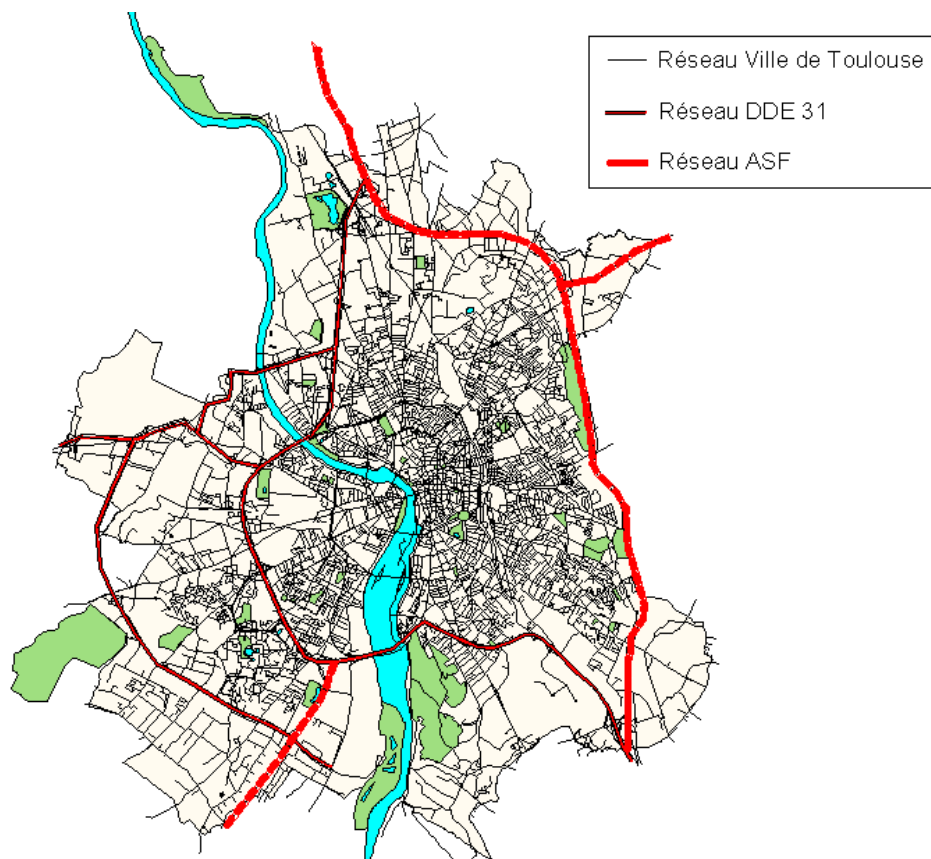


Figure 2 : Couverture géographique des gestionnaires de voirie

## 3 Analyse des besoins

Cette partie décrit l'analyse des besoins des deux partenaires de ce projet qui sont directement et quotidiennement concernés par le problème du traitement des données du trafic : la Ville de Toulouse et la Direction Départementale de l'Équipement de la Haute-Garonne.

Ce chapitre présente tout d'abord la méthode utilisée pour identifier les besoins des exploitants. Ensuite, un deuxième paragraphe précise ces besoins à partir de l'analyse des réponses des différents entretiens effectués. Cette analyse est également complétée par les résultats de travaux et d'enquêtes effectués dans le cadre de récents projets. Enfin, un dernier paragraphe propose une synthèse de cette analyse permettant de dégager les principaux éléments pour l'élaboration du cahier des charges au chapitre 5.1.

### 3.1 Méthode : Entretiens et Questionnaire

Une série d'entretiens ont été réalisés avec un groupe d'exploitants du trafic de l'agglomération toulousaine, au cours desquels ont été discuté en groupe la méthodologie de l'étude, les questions à soulever et un certain nombre de besoins a priori. Ces entretiens avaient pour but d'identifier au mieux les besoins des exploitants en matière de traitement de données afin de dégager un ensemble de fonctions et d'opérations pour une meilleure exploitation et connaissance du trafic.

La méthode utilisée a consisté à :

1. élaborer un questionnaire pour identifier le fonctionnement global des organismes concernés et ainsi que leurs besoins.
2. soumettre ce questionnaire à la Ville de Toulouse et la DDE 31
3. analyser les réponses
4. obtenir des compléments d'information par des entretiens complémentaires

#### 3.1.1 Contenu du questionnaire

Un questionnaire unique a donc été élaboré et diffusé aux exploitants de l'agglomération. Ce questionnaire est fourni en annexe. Il a permis à la fois de préciser le fonctionnement global de chaque organisme interrogé et d'identifier les opérations d'exploitation du trafic. Pour cela, différents thèmes ont été abordés afin :

- d'expliquer les fonctions exercées en matière d'exploitation comme par exemple les fonctions de supervision du trafic, celles de gestion des événements et de diffusion d'information vers les usagers ainsi que les fonctions destinées au partage des données entre les exploitants.
- d'indiquer le type d'utilisateur qui fréquente les réseaux urbains, périurbains ainsi que de préciser ses principales attentes en terme de gestion, de sécurité et d'information.
- de décrire les données statiques (ou données d'archive) disponibles avec les méthodes de recueil associées.
- d'identifier les données dynamiques recueillies par les systèmes automatiques de mesure (stations des comptages, caméras, etc.).

- de préciser les sources de données supplémentaires mais à ce jour non utilisées soit pour l'exploitation soit pour l'information des usagers. Il s'agit là d'identifier les données susceptibles d'être incorporées dans un système de traitement de données.
- d'expliquer les différents traitements effectués sur les données dynamiques et statiques.
- de décrire les types d'informations accessibles aux différents médias et diffusées aux usagers de la route.
- de préciser les besoins qui ont été directement exprimés par les opérateurs.

## 3.2 Résultats de l'analyse des entretiens

Ce paragraphe propose une analyse des entretiens effectués auprès de deux exploitants de l'agglomération toulousaine. Nous proposons une analyse commune des résultats obtenus afin de dégager les synergies en terme de traitements de données de trafic pour l'exploitation de la route et d'information des usagers. En effet, la Ville de Toulouse et la DDE 31 ont, dans le cadre du SGGD, des actions communes ou complémentaires et qui se sont reflétées au travers des réponses formulées. Toutefois, chaque exploitant a des besoins en terme d'exploitation de la route qui lui sont propres. Pour cela, nous nous attacherons à bien les préciser.

### 3.2.1 Les fonctions d'exploitation

Les principales fonctions d'exploitation sont avant tout la gestion à partir du PC de la circulation et du fonctionnement du matériel.

En matière de gestion du trafic, les exploitants s'intéressent plus particulièrement à la détermination des congestions et ont une mission d'urgence en cas d'événements. C'est la fonction principale du CIGT (de la DDE) qui est l'exploitant des voies rapides urbaines. Les états de circulation sont essentiellement observés à partir de caméras vidéos ou indiqués par les patrouilleurs, par les automobilistes à partir du réseau de Postes d'Appel d'Urgence (PAU) ou bien par la Police Nationale. La Ville de Toulouse, responsable de la circulation dans la partie intra-muros, exerce en plus des fonctions de régulation liées à l'existence des carrefours à feux. Ces fonctions sont la mise en place de stratégies de régulation (plan de feux fixes par plages horaires), l'anticipation de l'événement par le système expert CLAIRE et la mise en place de stratégies de priorité au bus sur certains axes du réseau.

Les fonctions de supervision du matériel sont également primordiales pour assurer la gestion du trafic et des événements et donc une continuité de services. Il s'agit dans ce cas de vérifier l'état des carrefours à feux (lampes de signalisation, contrôleurs, etc.), du système de recueil de données (les stations de comptages automatiques), des systèmes de communication et de traitement (CAPITOU par exemple) et les systèmes d'information (Panneaux à Messages Variables). Cette supervision du matériel est effectuée par l'opérateur quasiment en temps réel à partir de logiciels de supervision qui retournent des alarmes matérielles ou des informations inscrites dans les journaux d'exploitation.

Les fonctions d'information et d'échange de données entre les gestionnaires sont des fonctions nouvelles et apparaissent donc moins prioritaires que celles proprement dédiées à la gestion du trafic. Toutefois, le positionnement du SGGD pour assurer les fonctions de recueil des données et d'agence de presse indique que ces fonctions seront importantes dans la diffusion d'information aux usagers.

En ce qui concerne les fonctions d'information aux usagers, il s'agit d'éditer manuellement (jusqu'à l'acquisition d'un système semi-automatique) des messages à

partir des états de la circulation observés (niveaux de circulation, accidents, etc.) ou prévus (travaux, fermeture de voiries, etc.). Actuellement, les informations sont proposées à l'utilisateur par différents médias :

- par voies de presse. Certaines radios locales reçoivent automatiquement par fax des bulletins sur les états du trafic, qu'ils diffusent plusieurs fois par jour lors des flashes d'information.
- par l'intermédiaire de Panneaux à Messages Variables (PMV) situés sur les voies rapides urbaines ou à une dizaine de carrefours importants dans le centre ville. Les premiers PMV sont pilotés par la DDE 31 ou les ASF, tandis que les seconds sont commandés par le PC de la Ville de Toulouse.
- par le serveur audiotel Allo Trafic (08 36 67 31 01) qui donne des renseignements sur les états de la circulation par trajet et/ou par quartier (Toulouse Nord, Sud, Est et Ouest) et par rocades. Des conseils pratiques pour mieux circuler, des informations sur les parkings ainsi que des services sur abonnement sont également accessibles sur ce média.

Le partage des informations entre les différents exploitants du réseau et les autorités (Ville de Toulouse, DDE 31, ASF, Centre Régional d'Information Routière) concerne principalement les informations sur les événements et sert à la gestion globale des déplacements dans l'agglomération de Toulouse. Cet échange s'effectue soit par téléphone, soit par fax, soit par l'intermédiaire de liaisons numériques spécialisées. Certaines passerelles informatiques ont été mises en place pour des besoins spécifiques (échange des données de comptage entre la Ville de Toulouse et la DDE 31 ; système ERATO qui permet la coordination des réseaux entre la DDE 31 et ASF). Toutefois, les moyens informatiques actuels en matière de communication ne permettent pas à un exploitant d'utiliser des données en temps réel en provenance d'un autre centre d'exploitation pour la gestion du trafic. On peut cependant noter que dans le cadre du SGGD, les différents partenaires réfléchissent à ce problème et étudient la mise en œuvre d'un système de partage et d'échange de données.

### 3.2.2 Les usagers des réseaux

Afin de bien appréhender les fonctions de traitements de données de trafic qui sont utiles aux exploitants, il nous a semblé important de caractériser les usagers des réseaux de l'agglomération. En effet, suivant la nature du trafic les besoins peuvent être différents.

On observe 1 million de déplacements par jour sur l'agglomération de Toulouse. Les modes de déplacement les plus utilisés sont la voiture particulière (63% des trajets), la marche à pied (22%), les transports collectifs urbains (8%). Le reste des déplacements ne représente que 7% (deux-roues motorisés ou non, taxis, etc.). L'usage de la voiture est donc prépondérant, évolution qui est confirmée par une augmentation de la circulation (déplacements des voitures) de 30% depuis 1990.

On peut signaler que 96 % des déplacements sont réalisés avec un seul mode. La durée du déplacement en 1996 est de 16 minutes en moyenne pour l'automobiliste. L'utilisateur des transports collectifs passe le double de temps en déplacement ! Les déplacements domicile travail représentaient 20% des déplacements en 1996. Concernant les flux de déplacements, on a assisté depuis une décennie à une véritable explosion des déplacements dans les liaisons internes à la périphérie (+58% entre 1990 et 1996).

Si l'on regarde plus précisément l'usage des voies rapides par des automobilistes réguliers, on s'aperçoit que 49% l'utilisent pour le motif domicile travail. Un trajet dure en moyenne 16 minutes et mesure 18 km.



Lorsqu'il est demandé aux usagers de citer les informations qu'ils souhaitaient obtenir, ils indiquent par ordre d'importance des informations sur les bouchons et les ralentissements, les accidents, l'état du trafic ainsi que les itinéraires possibles.

Ces données sont issues de l'enquête sur les déplacements des ménages en 1996 et de l'enquête ERATO réalisée en 1997 auprès de la clientèle régulière du périphérique (1000 interviews téléphoniques et 300 questionnaires).

En conclusion, les besoins des usagers qu'ils soient directement exprimés par les automobilistes lors d'enquêtes ou explicités par l'intermédiaire des exploitants, sont liés :

- **à l'objet du déplacement** : la mobilité qui peut s'exprimer par le fait que l'utilisateur veut bénéficier des installations de transports lui permettant d'atteindre n'importe quel endroit à n'importe quel moment.
- **aux conditions du déplacement** : la commodité qui comprend la notion de confort et de sécurité. Comme les récentes enquêtes l'ont montré les usagers sont très sensibles à la fluidité du trafic surtout sur les axes périphériques. En centre ville, ces besoins sont complétés par d'autres besoins tels que le stationnement par exemple.

### 3.2.3 Les différentes données mesurées

On distingue les données statiques et les données dynamiques. Les premières sont des données archivées qui servent essentiellement à caractériser les déplacements lors d'analyses de la circulation. Les secondes sont des données utilisées au fil de l'eau pour la surveillance du trafic.

#### 3.2.3.1 Les données statiques

Les données statiques sont issues :

- des capteurs posés temporairement sur certains axes. Les capteurs fournissent des données sur les débits de véhicules, la vitesse et la composition du trafic.
- des mesures de la congestion à partir de véhicules traceurs équipés du système MiTemps permettant de mesurer les vitesses et les temps de parcours sur des trajets prédéfinis.
- de l'observatoire du centre ville qui effectue des enquêtes sur les déplacements du centre ville sur une base annuelle.
- d'enquêtes régulières auprès des usagers (enquête ménage tous les 5 ans) ou plus sporadiques suivant les besoins des exploitants (étude ERATO en 1997).

D'une manière générale, ces informations restent très difficilement exploitables dans le sens où elles existent le plus souvent sous le format papier. De plus, les mesures effectuées sont bien souvent ponctuelles et donc forcément limitées dans le temps et l'espace pour un usage systématique dans l'exploitation des réseaux.

#### 3.2.3.2 Les données dynamiques

On distingue différents types d'informations qui peuvent être utilisées pour caractériser les états de trafic en temps réel, il s'agit :

- **des systèmes automatiques de recueil** (comme les stations de comptages) qui transmettent des données relatives à la circulation par des liaisons numériques spécialisées ou non. Ces données arrivent au fil de l'eau au niveau du PC d'exploitation et sont utilisées généralement pour la supervision et la régulation du trafic.

- **des systèmes de vidéo surveillance,**
- **des informations sur les événements** transmises par les usagers via les bornes d'appels téléphoniques du PAU et par des patrouilleurs. Les mains courantes informatisées (MCI) ou mains courantes électroniques (MCE) sont utilisées pour saisir de manière informatique les événements constatés sur la voirie ainsi que les anomalies de fonctionnement des systèmes. Le contenu des mains courantes peut être transmis vers les opérateurs voisins.
- **des informations sur les autres modes de transports,** comme la SEMVAT,
- **des informations émises par des véhicules traceurs** et équipés d'un système automatique de communication entre le véhicule et le central. Certains systèmes de guidage embarqués dans le véhicule utilisant des technologies à base de balises infrarouges ou de GSM offrent la possibilité d'obtenir le temps de parcours et le nombre d'arrêts des véhicules équipés. Les flottes de taxis et les usagers équipés d'un téléphone mobile peuvent également transmettre des informations.
- **des données qualitatives** comme par exemple les manifestations, les occupations des parkings, les prévisions météorologiques,

Au niveau de l'agglomération de Toulouse, il existe :

- **un système de recueil automatique** pour le comptage des véhicules sur les voies rapides urbaines et un autre pour la voirie du centre ville.

Sur le réseau des voies rapides urbaines qui est exploitée en partie par la DDE 31, les mesures proviennent toutes les 6 minutes de 60 stations SIREDO. Une station est composée d'une ou plusieurs boucles électromagnétiques. Les valeurs mesurées correspondent à chaque boucle et sont agrégées pour fournir des mesures pour la station. Toutes ces mesures sont stockées au niveau du PC DDE au CIGT dans une base de données.

Pour le réseau des voies urbaines exploitées par la Ville de Toulouse, le système de recueil actuel comprend un ensemble de 400 Postes de Mesures (PM). Les capacités du nouveau PC CAPITOUL permettent le déploiement de 1000 PM. Des valeurs de débits de véhicules et de taux d'occupation sont mesurées par cet ensemble de PM toutes les 3 minutes et sont archivées dans une base de données relationnelles (ORACLE™) au niveau de CAPITOUL.

Toutes ces données sont utilisées à la fois en temps réel pour le suivi du trafic mais aussi en temps différé pour l'étude de phénomènes de trafic ou pour le calcul de statistiques.

- **Un système de baie vidéo** aussi bien au PC DDE qu'au PC CAPITOUL permet de suivre l'évolution du trafic sur les principaux axes du réseau. Actuellement, il existe 40 caméras de surveillance des axes périphériques. Le système de vidéo surveillance Pégase permet de télécommander les 30 caméras installées en milieu urbain et est utilisée par les exploitants du PC CAPITOUL. Cette supervision est assurée par les différents opérateurs de trafic afin d'intervenir en cas d'événements majeurs.
- **Des événements** relatifs aux accidents ou aux congestions sont transmis par les patrouilleurs certifiés, par la Police Municipale ou par les usagers eux-mêmes à partir des postes d'appel du RAU situées sur les voies rapides par exemple. Ces événements sont reçus au niveau du pupitre centralisateur pour la DDE 31 et

généralement «consommés» très rapidement lors de missions d’urgence. Ils sont enregistrés dans des journaux d’événements.

- **Le système de jalonnement des parkings** du centre ville permet d’afficher sur des panneaux dédiés le taux de remplissage. Ces informations diffusées sont centralisées au niveau du PC CAPITOU.
- **Les informations** (position du Bus, les demandes de priorité) **élaborées par le système SITERE** de la SEMVAT, dont la fonction est l’information clientèle, sont également échangées avec le PC CAPITOU.

### 3.2.3.3 Les données disponibles mais non utilisées

Nous listerons comme données celles recueillies par les différents systèmes, qui sont susceptibles d’être intégrées dans un organe de commande, de régulation ou de surveillance aussi bien pour l’exploitation du trafic que pour l’information des usagers mais qui ne sont toutefois pas utilisées. Suite aux entretiens, il ressort que les données utilisables sont :

- les événements observés ou programmés par les partenaires du SGGD qui intéressent la Ville de Toulouse pour adapter les stratégies de régulation en fonction des événements survenus ou ceux programmés.
- la localisation des véhicules d’intervention pour sélectionner la ou les unités proches de l’incident sont des données potentielles. Actuellement, les véhicules patrouilleurs de la DDE 31 sont localisés par simple appel radio.
- Les données fournies (comme par exemple les temps de parcours) par des véhicules traceurs comme les flottes de taxis, le système de suivi des bus et les systèmes embarqués dans le véhicule pour le guidage des automobilistes.

### 3.2.4 Les traitements réalisés sur les différentes données

Le tableau ci-dessous résume la nature des traitements effectués à ce jour sur les données statiques et dynamiques ainsi que la méthode utilisée.

Exploitant	Données Statiques	Données Dynamiques
DDE 31	Traitements statistiques des données de comptage archivées des stations SIREDO	Traitements en ligne des données de comptages des stations SIREDO
	Vérification de la cohérence des mesures et agrégation des données par tranches de temps et par stations	Vérification de la cohérence des mesures et du fonctionnement des équipements
	Utilisation de logiciels spécifiques	Utilisation de logiciels spécifiques
Ville de Toulouse	Traitements statistiques des données de comptages des postes de mesure et des enquêtes diverses (Système MiTemps)	Exploitation du trafic
	Vérification du fonctionnement du système de recueil Analyse de phénomènes de trafic	Suivi des alarmes matérielles Suivi et actions pour remédier à la saturation.

		Suivi du remplissage des parkings
	Logiciels de trafic utilisés : <ul style="list-style-type: none"> <li>développé par la société CGA avec le PC actuel</li> <li>Business Object et logiciel MAITRE Analyse</li> </ul>	Logiciels utilisés : <ul style="list-style-type: none"> <li>systèmes informatiques de régulation et de surveillance CAPTITOU II</li> <li>CLAIRE (système expert)</li> <li>Application de suivi de l'occupation des parkings</li> </ul>

### 3.2.5 Messages et diffusion d'information

Dans cette partie, nous nous intéressons aux différents types d'informations qui sont proposées aux usagers ou bien aux différents opérateurs. Le tableau suivant précise les informations diffusées à ce jour, les médias concernés, la méthode de diffusion employée et autant que possible la fréquence de diffusion de ces informations.

Messages diffusés	Médias utilisés	Méthode de diffusion	Fréquence
Informations relatives aux événements : <ul style="list-style-type: none"> <li>chantiers,</li> <li>accidents,</li> <li>manifestations,</li> <li>etc.</li> </ul>	Radios Télévision	Les messages sont : <ul style="list-style-type: none"> <li>édités par les pupitreurs du PC DDE et du PC CAPITOU</li> <li>transmis automatiquement du PC aux médias par fax ou par téléphone (uniquement sur appel d'un tiers)</li> </ul>	Point sur le trafic aux heures de pointes
	PMV du réseau périphérique (3-4 lignes de texte, pictogrammes)	Les messages sont : <ul style="list-style-type: none"> <li>édités par les pupitreurs du PC DDE et du PC CAPITOU.</li> <li>affichés automatiquement après validation de l'opérateur</li> </ul>	Sur événements
	PMV du centre ville (3 lignes de 12 caractères)	Messages préétablis et enregistrés.	Non précisé
Informations sur les niveaux de congestion	Serveur audiotel AlloTrafic 08 36 67 31 01		
	PMV du centre ville (3 lignes de 12 caractères)	Messages automatiquement édités par le système CLAIRE à partir des seuils de saturation détectés	Après détection de la saturation
	Serveur AlloTrafic	+ messages sur les états de trafic du périphérique DDE 31	Non précisé

### 3.2.6 Les besoins exprimés par les opérateurs

Cette partie s'attache à décrire les besoins qu'ont pu exprimer les gestionnaires durant les différents entretiens. Les besoins sont relatifs à la supervision, la gestion des événements, la gestion du trafic, l'information des usagers et la fonction de partage de l'information entre les différents membres du SGGD.

Tout d'abord, chaque exploitant a exprimé le besoin de connaître les stratégies mises en place par les partenaires lors de situations de crise (événements) ou en fonctionnement dégradé (dysfonctionnement d'un système) afin de déployer les actions ad hoc (régulation, information des usagers, etc.). De la même façon, la détection et la validation des événements sont des préoccupations car ceux-ci ont des conséquences directes et indirectes sur les réseaux concernés mais aussi sur les autres réseaux. Par exemple, la fermeture d'un échangeur d'accès du périphérique peut engendrer des perturbations aussi bien sur le réseau urbain et que sur le réseau périphérique.

En ce qui concerne la gestion du trafic, les gestionnaires du PC trafic urbain sont relativement demandeurs d'informations complémentaires à leur système de recueil de données. Par exemple, l'utilisation de données trafic (comptages des véhicules et saturation des voies) des autres centres d'exploitation leur permettrait de mieux anticiper les phénomènes de circulation.

Pour l'information des usagers, la vision du SGGD consiste à promouvoir une "agence de presse" centralisant à la fois des informations concernant la circulation, les transports en commun ainsi que des informations pratiques (tarification des parkings, conditions météorologiques). Les exploitants se placeraient tout d'abord en fournisseurs de données (élaborées ou brutes). Le partage des informations s'intégrerait dans ce schéma dans le sens où chaque exploitant pourrait être un consommateur de données. De plus, le rapprochement des PC de la ville de Toulouse et de la DDE au nouveau Campus du Montblanc doit permettre le déploiement d'un système de centralisation et d'échange de l'information.

### 3.3 Complément sur les besoins en traitement de données

Les partenaires du SGGD mais plus particulièrement la Ville de Toulouse, ont été ou sont impliqués dans différentes études européennes. Certains des projets ont permis le développement ainsi que l'évaluation de nouveaux systèmes pour la régulation et la supervision de la circulation de régulation. Parmi les projets les plus récents, on peut citer les projets ANTARES, CLEOPATRA et QUARTET +.

#### **QUARTET +**

Le projet QUARTET+ a consisté à mettre en place et à évaluer une stratégie de priorité aux bus basée sur le système PRODYN BUS. Une enquête auprès d'un petit nombre d'usagers de l'agglomération toulousaine a été effectuée. Les résultats obtenus sont difficilement exploitables pour cette étude, d'une part du fait de la petite taille de l'échantillon (14 personnes). A titre indicatif et à partir des remarques que les usagers ont exprimées, on peut citer deux objectifs attendus en matière d'informations sur les déplacements : une information en temps réel sur les conditions de circulation et ce, y compris pour les usagers du bus et une information sur les moyens de réduire le temps du déplacement par des recommandations d'itinéraires adaptés (ce qui ne concerne que les usagers des véhicules particuliers).

#### **ANTARES**

Dans le cadre d'une meilleure gestion et régulation des flux urbains, le projet européen ANTARES a eu pour objet de mettre en place une nouvelle approche des transports en rapport avec l'économie d'énergie. Les objectifs sous-jacents étaient :

- la réduction de la consommation,
- de stimuler la pénétration des technologies européennes,
- de réduire les rejets de polluants,

- de promouvoir la cohésion économique et sociale de la communauté européenne.

Les villes de Barcelone, Bologne, Dublin, Leipzig et de Toulouse ont été impliquées dans ce projet. Le travail mené sur le site de Toulouse a consisté en la mise en œuvre de l'outil informatique CLAIRE de supervision de la régulation du trafic routier intégré au PC Capitoul. Ce système offre à l'exploitation un ensemble de capacités fonctionnelles permettant de remédier à la congestion surtout en milieu urbain telles que :

- la perception et l'interprétation de la saturation (diagnostic),
- proposer et appliquer des remèdes contre la saturation (aide à la décision et automatisme),
- reconnaître les saturations et s'en souvenir (mémoire à long terme),
- élaborer lui-même ses propres remèdes.

Les principaux résultats du projet ANTARES sont l'intégration de l'activité de l'homme en gestion du trafic, l'élaboration de modèles symboliques pour le diagnostic et le contrôle de la congestion, l'élaboration d'une base de connaissance pour la congestion du trafic (observatoire de la congestion) et enfin l'intégration du système expert CLAIRE aux organes de régulation et d'information des usagers (PMV et serveur AlloTrafic) du PC CAPITOUL.

### **CLEOPATRA**

Le projet CLEOPATRA de 1995 à 1999 est un projet de recherche et de développement technologiques dans le domaine ITS (Intelligent Transport Systems). Ce projet s'est focalisé sur le guidage dynamique et l'information des usagers en milieu urbain et périurbain en utilisant des systèmes d'information individuels (embarqués dans le véhicule) et collectifs (au bord des routes). Ce projet a consisté à développer et à évaluer en laboratoire et surtout sur site des algorithmes, des modèles et des stratégies pour alimenter les systèmes d'information et de guidage. Les algorithmes évalués sont des méthodes d'estimation et de prédiction dans le futur proche des états des réseaux routiers en temps réel.

Les sites impliqués dans ce projet sont Berlin, Londres, Lyon, Stockholm, Toulouse et Turin. Six applications/algorithmes différents ont été évalués : estimation des flux de véhicules et des matrices OD ; prédiction des temps de parcours ; prédiction des effets des incidents ; stratégies de guidage collectif et d'information basées sur l'affichage de messages sur différents types de PMV ; stratégies d'information à partir d'unités embarquées dans le véhicule (In-Vehicle Information System -IVIS) et intégration des systèmes UTC (Urban Traffic Control), PMV et IVIS.

Dans le cadre de ce projet, une méthode de prédiction jusqu'à 15 minutes des temps de parcours (Journey Time Prediction JTP) sur une petite partie du réseau de l'agglomération (boulevards circulaires du Pont Saint-Michel, Grand Rond, Place Jeanne d'Arc, Place Héraclès, Place Saint-Cyprien ainsi que la croix rue du Languedoc, rue Alsace Lorraine et rue de Metz). Une stratégie de guidage des usagers à partir des 11 PMV télécommandés par la Ville de Toulouse a également été testée sur ce réseau. Les applications JTP et stratégie de guidage par PMV ont fait l'objet d'une implémentation au PC CAPITOUL et ont été connectées aux données du trafic (comptage des débits et taux d'occupation) en temps réel.

Les enquêtes réalisées sur le site de Toulouse ont permis à la fois de tester les performances des algorithmes et de mesurer le degré de satisfaction des usagers (exploitants et automobilistes). Nous nous intéresserons seulement aux résultats relatifs à la satisfaction des usagers de la route et des exploitants :

- les exploitants du PC CAPITOUL ont pu évaluer les performances de la méthode de prédiction des temps de parcours en fonctionnement en ligne et surtout appliquée en milieu urbain. En effet, les tests effectués ont porté sur la comparaison entre des temps de parcours prédits sur des chaînons du réseau et des temps de parcours historiques ainsi qu'avec des temps de parcours mesurés par des véhicules traceurs. Ces tests ont montré que les temps de parcours estimés et prédits étaient assez fiables pour une utilisation possible pour l'exploitation, la détection des niveaux de circulation et donc de la congestion pour des chaînons et des trajets.

Les besoins pour les exploitants en matière de prédiction, c'est-à-dire qu'à partir d'un ensemble de mesures passées et courantes on essaye de prévoir l'évolution du trafic dans les instants futurs, sont :

- ◇ la prédiction à court terme (de 3 minutes à ½ heure) pour donner des consignes de guidage en fonction de l'état de la circulation actuel mais aussi à venir,
- ◇ la prédiction à un jour pour la mise en place de stratégies en fonction des événements récurrents, programmés, etc.

Les exploitants ont émis des besoins complémentaires comme par exemple l'intégration d'autres données (événements, données météorologiques, données VRU) pour affiner à la fois l'estimation et la prédiction.

En matière de guidage des usagers de la route par PMV, les exploitants sont tout à fait satisfaits de l'expérimentation car d'une part la stratégie a montré l'utilité des PMV urbains (à l'heure actuelle sous-exploités) et d'autre part, la stratégie pouvait dévier une partie du flux de véhicules et ainsi être un outil pour remédier à la congestion.

Les entretiens effectués ont indiqué que les exploitants du PC CAPITOUL semblent être intéressés par l'utilisation des PMV pour le guidage vers l'extérieur de la ville et par l'affichage d'informations sur les événements.

- Une enquête auprès des conducteurs de l'agglomération toulousaine a également été menée. Elle concerne uniquement l'évaluation d'une stratégie d'affichage d'informations de guidage sur le Panneau à Messages Variables installé sur l'avenue de Muret. L'échantillon est plus conséquent que pour les enquêtes qui viennent d'être citées puisque 500 questionnaires ont été distribués avec un taux de réponse de 32.6% (163 réponses). Des résultats partiels qui concernent l'appréciation de l'information communiquée aux conducteurs indiquent que :

Un tiers des automobilistes estiment le message insuffisant pour indiquer le conseil de guidage, car ils veulent connaître : les raisons du guidage (57%), l'itinéraire (17%), le temps ou la durée du trajet (17%)

Les principaux arguments faisant changer de direction les conducteurs sont : un accident visible (18%), une information sur un accident (26%), un embouteillage visible (31%), une information d'embouteillage (39%) et un temps de parcours plus faible (28%).

Cette enquête indique donc que l'utilisateur perçoit une information de guidage comme une proposition et qu'il souhaite disposer d'informations complémentaires pour décider de suivre ou non le conseil. Notons également que l'information sur un accident ou sur un embouteillage est plus suivie que l'événement lui-même, ce qui témoigne de l'importance que les usagers accordent à l'information.

### 3.4 Synthèse sur les besoins des exploitants

Les principaux besoins explicités par les exploitants de la DDE 31 et de la Ville de Toulouse en matière de traitement dynamique des données sont :

- **d'assurer une continuité de traitements et de services**

Pour effectuer leurs fonctions temps réel (régulation des flux de circulation, surveillance du trafic et diffusion d'informations routières), les exploitants ont recours aux mesures dynamiques du trafic par boucles électromagnétiques. Malheureusement, les bases de données trafic générées par ces systèmes sont souvent incomplètes, du fait des pannes fréquentes des éléments de détection, et ne permettent pas d'assurer une continuité dans les traitements.

L'un des besoins des exploitants est d'utiliser des nouvelles méthodes qui remédient aux pannes des postes de mesure pour se substituer aux méthodes empiriques des exploitants. Il s'agit d'utiliser une bibliothèque d'algorithmes de qualification et de reconstitution de données pour assurer un fonctionnement optimal ou quasi optimal des organes de gestion du trafic. L'idée sous-jacente est de minimiser la durée de fonctionnement en mode dégradé de ces organes.

- **de quantifier en temps réel les états de la circulation sur les différents types de réseau**

Les indicateurs les plus cités sont le temps de parcours et le niveau de circulation qui permettent de quantifier les conditions de circulation.

L'estimation des **temps de parcours en temps réel** constitue un enjeu important en matière d'information routière notamment en période de congestion et d'accidents. Cet indicateur est **commun à tous les modes de déplacement** et permet de comparer des solutions alternatives pour un déplacement sur la base d'un même critère.

De façon analogue, il est souvent important de connaître précisément les niveaux de saturation sur les réseaux urbains et périurbains. L'estimation des niveaux de saturation, également appelé niveaux de services, permet aux exploitants de mettre en place les stratégies adaptées.

D'autres types d'informations sont aussi attendus par les exploitants des réseaux. Il s'agit des données sur les débits de véhicules (sur tous les chaînons du réseau), les longueurs des files d'attente et les vitesses.

- **de connaître l'évolution des conditions de circulation dans le futur proche**

La connaissance du futur proche en terme de trafic permet aux exploitants des réseaux d'améliorer et d'adapter le service offert aux usagers. Fournir à l'automobiliste la meilleure route tenant compte des conditions à venir semble être une composante vitale pour les systèmes de guidage dynamique. Les relations entre les conditions présentes et futures du trafic sont complexes et impliquent à la fois des dimensions spatiales et temporelles. L'écoulement du trafic routier est souvent caractérisé de «non linéaire», les causes en étant souvent les congestions dues à des accidents, des travaux...

De plus, les fonctions de traitement effectuées à ce jour sont appréhendées uniquement à partir de mesures issues de boucles magnétiques. Il semble important de pouvoir intégrer



et de fusionner d'autres sources de mesure dans le processus de traitement des données. Sur l'agglomération toulousaine, cette fusion peut s'exprimer en utilisant :

à la fois les données de comptage des différents réseaux,

mais aussi l'observation des temps de parcours de véhicules traceurs (flottes de taxis, véhicules équipés de terminaux ou de téléphone GSM, etc.).

## 4 Etat de l'art sur les systèmes de fusion de données

### 4.1 Définition et problématique

Les systèmes modernes d'information sur le trafic comprennent trois fonctions principales. La première est celle du recueil des données provenant des différentes sources de mesures des conditions du trafic (postes de mesure, caméras vidéo, appels téléphoniques, etc.). La seconde est celle de la fusion de données qui élabore automatiquement des informations relatives à la situation de trafic présente. La dernière est celle de la diffusion des informations qui permet aux automobilistes d'en disposer en temps voulu. Par ailleurs, ces systèmes peuvent aussi fournir des informations de nature pratique (localisation des stations de services, des hôtels, des sites touristiques, etc.).

D'autre part, les informations traitées par ces systèmes peuvent être également utilisées par les opérateurs des réseaux urbains pour assurer la surveillance globale du trafic sur la ville. Actuellement la surveillance du trafic (utilisation de caméras vidéo, bornes d'appel d'urgence, etc.) se fait de façon non systématique, en certains points du réseau, et de manière qualitative.

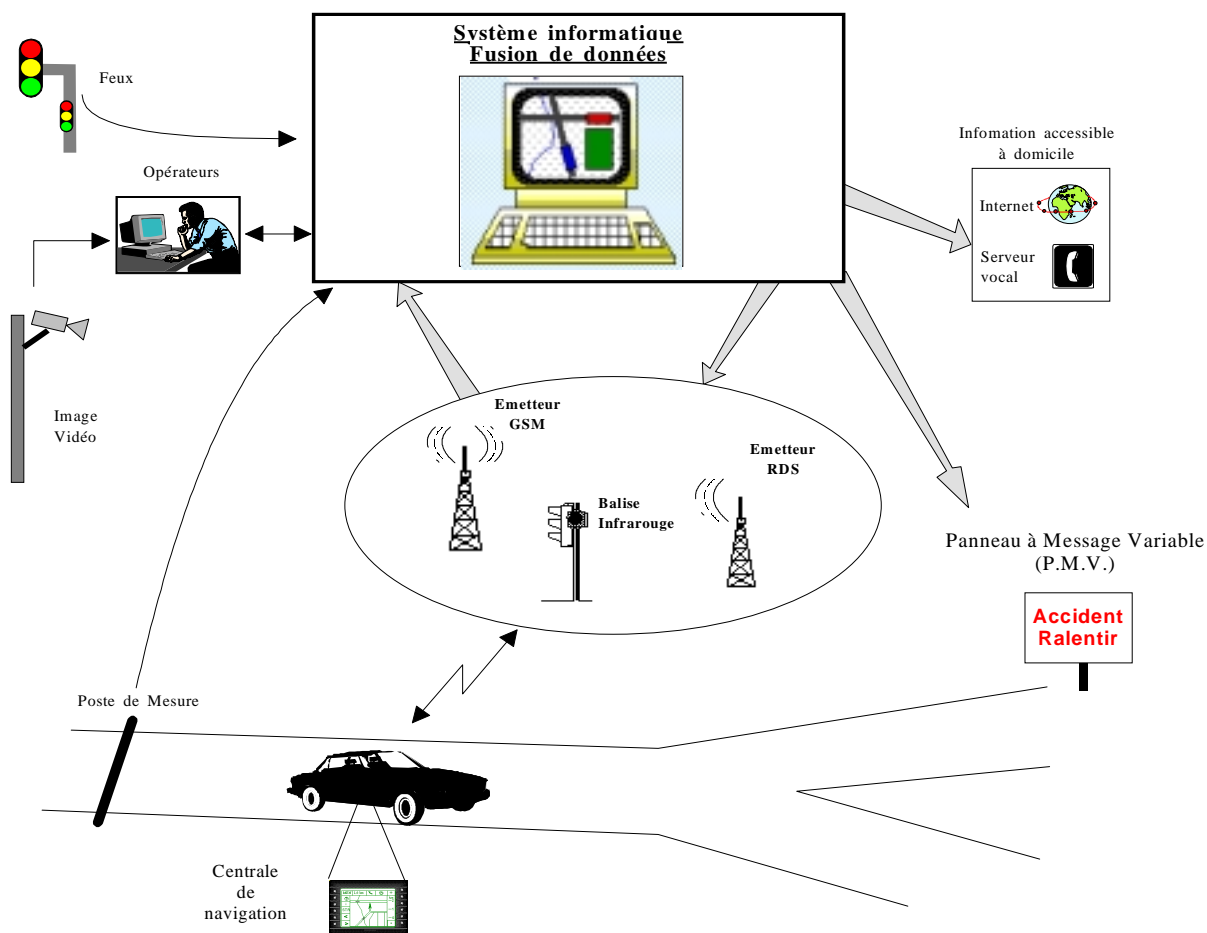


Figure 3: Support physique pour l'information routière

Il est donc essentiel que la génération, en temps réel, des informations relatives au trafic soit continuellement disponible et fiable. Ce besoin a par ailleurs été fortement exprimé par les exploitants des réseaux de l'agglomération de Toulouse. C'est l'objet d'un système de fusion que d'utiliser les différentes sources de mesures du trafic disponible sur une agglomération pour élaborer un ensemble d'indicateurs pour répondre aux exigences des usagers.

## 4.2 Le recueil de données

L'objectif de ce paragraphe est de rappeler les différents systèmes qui permettent de recueillir les données trafic utiles aux exploitants pour la gestion des réseaux. Pour une information complète sur le sujet, le lecteur peut se référer à une étude récente réalisée pour le PREDIT (DATA Plus, 1999) qui porte sur l'évolution des technologies et des pratiques en matière de recueil de données.

Il convient tout d'abord de rappeler que le recueil de données constitue une chaîne d'informations comportant les maillons *capteur*, *détecteur*, *transmission* et *traitement*. Dans cette section nous nous intéressons seulement aux capteurs et aux détecteurs.

Les grandeurs habituellement mesurées dans le domaine du trafic sont variées et destinées à être utilisées pour la connaissance du trafic et sa régulation, l'information des conducteurs, l'exploitation des péages ou la répression des infractions. Les mesures les plus classiques concernent :

- Les variables fondamentales du trafic comme le débit, la vitesse, la densité, le taux d'occupation, les mouvements tournants dans les carrefours, la présence et la longueur de files d'attente,
- Les événements affectant le trafic comme les bouchons et les accidents,
- Les caractéristiques des véhicules comme le poids, les dimensions, l'identifiant,
- Les conditions météorologiques,
- Les infractions comme l'excès de vitesse, le franchissement de feu rouge,
- Les grandeurs relatives à l'environnement comme les nuisances sonores et la pollution.

La mesure de ces variables peut être réalisée au moyen de diverses technologies :

- Capteurs à boucle,
- Magnétomètres,
- Capteurs acoustiques et ultrasons,
- Capteurs optiques,
- Capteurs micro-ondes et radars,
- Capteurs de pression,
- Capteurs vidéo,
- Lecture de plaques minéralogiques,
- Fusion de données,
- Véhicules "traceurs".

Le tableau suivant présente la correspondance entre les différentes variables qui peuvent être mesurées et les types de capteurs susceptibles d'être employés pour ces mesures. La notation suivante a été adoptée :

a : le capteur est bien adapté à la mesure,

b : le capteur peut éventuellement servir à la mesure, mais ses performances ne sont pas optimales.

Tableau 1 : correspondance entre les différentes variables et les types de capteurs (source DATAPLUS)

	boucles	tube pneum.	magnétom.	accou. passif	US pulsé	IR passif	IR actif	radar Doppl.	radar	video	lecture plaq.	piézo	bascule dyn.	fibres opt.
Q	a	a	a		a		a	b	a	a		a		a
V	a	a	a		a		a	a	a	a		a		a
T.O	a		a		a		a		b	b	a			
intervalle	a	b	a		a		a			a				
densité K									a	a				
prés./pass	a	b	a		a		a	a	a	a				
Lqueues									a	a				
niveau traf.	a				a		a		a	a				
tps trajet	a								a	a	a			
O/D											a			
DAI	b		b		b		b	b	a	a	b			
DAB	b		b		b		b	b	a	a				
mvts tourn.	b				b					a				
chgt voie										a				
poids												a	a	a
gabarit					a		a							
classif.	b				b		a			a		a	a	a
identif.											a			
nb. passag														
dét. 2R								a	a	a				
dét. piéton								a	a	a				
dét. objet									a	a				
infrac. feu	a							a	a	a				
infrac. stop	b									b				
infrac.sens	b							a	a	a				
infrac.ligne	b									b				

### 4.3 Les méthodes existantes

Des travaux ont abouti à des systèmes temps réel de surveillance ou de traitement de la saturation. Des avancées importantes sont faites aux Etats-Unis, notamment avec le projet ADVANCE (Advanced Driver and Vehicle Advisory Navigation Concept). En Europe, on trouve le projet LLAMD (Londres, Lyon, Amsterdam, Munich, Dublin) où les algorithmes développés dans ce cadre sont essentiellement destinés à l'affectation des

véhicules sur des réseaux urbains et au guidage des usagers. Le projet CLEOPATRA (City Laboratories Enabling Organisation Of Particularly Advanced Telematics Research and Assesment) a pour objectif d'évaluer sur site des méthodes d'estimation, d'affectation et de prédiction du trafic. Le projet DACCORD étudie un outil d'estimation et de prédiction à court terme du trafic dans la région francilienne.

Toutefois, les systèmes et les méthodes actuels se heurtent à de nombreuses difficultés comme par exemple :

- le peu d'informations dynamiques disponibles (nombre de postes de mesures faible par rapport à taille du réseau routier),
- le fonctionnement des systèmes en mode dégradé (panne du système de recueil de données),
- l'utilisation optimale de données hétérogènes (mesures du trafic, conditions météorologiques, événements, etc.),
- la contrainte forte imposée par le temps réel.

Face à ces difficultés, certains compromis ont été adoptés pour améliorer artificiellement les résultats des estimations, par exemple en agrégeant dans le temps (utilisation de tranche de calcul plus longue) ou en simplifiant le réseau (agrégation ou réduction). De plus, la mise en place des systèmes de surveillance s'appuyant sur les nouvelles techniques de l'informatique, a permis le développement d'un certain nombre de fonctionnalités de base permettant de pallier des problèmes rencontrés sur site. C'est le cas de la fonction de fusion de données qui s'articule autour des applications :

- **de filtrage des données** qui reconstitue les données de capteurs manquants ou de capteurs dégradés afin d'assurer un fonctionnement optimal ou quasi optimal des organes de commandes.

L'article coréen (Choi D.B., 1998) présente ainsi divers algorithmes utilisés pour reconstituer des données dans des parties d'un réseau non couvertes par des détecteurs : « evidential reasoning », logique floue, hiérarchie Bayésienne ; il présente un peu plus en détail une démarche fondée sur la première de ces méthodes, sans fournir toutefois de résultats. Un article de l'INRETS (Hadj Salem H., 1997) aborde la question des détecteurs dégradés, et examine des cas où 1,3 ou 9 stations seraient hors service sur une autoroute, et de 0 à 6 stations seraient hors service en urbain. Il montre qu'en utilisant un filtrage de Kalman sur autoroute, et un filtrage statistique en urbain il est possible de reconstituer avec une précision satisfaisante, de l'ordre de 5 à 10% les données manquantes.

Des algorithmes résultants d'un travail de recherche (Schettini F., 1998), basés sur des méthodes statistiques des nuées dynamiques et de recherche de modèles macroscopiques liant les valeurs de débits de véhicules, de vitesse, de taux d'occupation, permettent de corriger un nombre important de bruit de mesures des capteurs installés en milieu urbain. Des études ont été menées sur deux réseaux urbains pour des données échantillonnées à 3 et 6 minutes et fournissent des résultats probants.

- **d'estimation des différents états du trafic** (temps de parcours, niveau de circulation).

Les méthodes de détermination des temps de trajet par fusion de données «hétérogènes» à la fois de capteurs au sol et de véhicules «traceurs» ou d'autres sources (lecteurs de plaques minéralogiques, bases de données historiques, observations qualitatives sur les niveaux de trafic, informations sur les travaux et obstructions) ont été explorées dans le cadre de nombreuses recherches. Les

articles (El Faouzi N.E., 1997), (Kühne R.D., 1997), (Nuttall I., 1998), (El Faouzi N.E., 1995), (Choi K., 1997), (Abe A., 1997), (Schettini F., 1998) fournissent des méthodes pour l'estimation des «temps de trajet». Ces documents décrivent des méthodes et modèles pouvant être appliqués pour déterminer des valeurs de temps de parcours.

- **de la connaissance à court terme de l'évolution du trafic** qui permet aux exploitants des réseaux d'améliorer et d'adapter le service offert aux usagers. Fournir à l'automobiliste la meilleure route tenant compte des conditions à venir semble être une composante vitale pour les systèmes de guidage dynamique. Les relations entre les conditions présentes et futures du trafic sont complexes et impliquent à la fois des dimensions spatiales et temporelles. L'écoulement du trafic routier est «non linéaire», les causes en étant souvent les congestions dues à des accidents, des travaux...

Les expériences passées en prévision sont essentiellement basées sur la prédiction du temps de parcours ou du débit de véhicules dans le futur proche à  $t+1, \dots, t+Hp$  ( $H_p$  étant l'horizon de prédiction) à partir des données courantes et historiques. Cependant, le temps de parcours n'est pas directement observable par les postes de mesure (PM), il est alors estimé à partir de modèles déterministes et la prédiction de celui-ci reste très délicate. Bien que les modèles, statistiques ou heuristiques, développés à ce jour pour prédire le débit à  $t+1, \dots, t+Hp$  fournissent des résultats corrects lorsque le trafic est fluide, ceux-ci se dégradent vite lorsque la circulation atteint la saturation. De plus, la relation fondamentale du trafic indique que pour un même débit, la circulation sur un chaînon peut être dans deux états (fluide ou saturée), donc la prédiction du débit seul ne détermine pas parfaitement la situation du trafic à venir.

Les méthodes potentielles pouvant être utilisées comme outil de prévision des séries chronologiques tels que le débit de véhicules et le taux d'occupation sont décrites brièvement ici. Un classement des différentes méthodes donne lieu à deux grandes catégories d'approches prévisionnelles, les approches conventionnelles et celles non conventionnelles. Dans la catégorie des approches prévisionnelles conventionnelles, on distingue les méthodes d'analyse des séries temporelles (Box & Jenkins, 1970), les techniques de filtrage (LLAMD, 1994, Gaillet, 1992 et CLEOPATRA, 1999), et les méthodes statistiques de Holt-Winter (Statistique, 1995). En ce qui concerne les approches prévisionnelles non conventionnelles, on trouve des méthodes basées sur les prédicteurs non paramétriques (El Faouzi, 1996, Rosenblatt, 1956, Bosq et Lecoutre, 1992), un prédicteur combinant des données historiques et mesurées (Gabard, 1986) et plus récemment des techniques neuronales (Dochy et Al, 1992, Dougherty et al. 1994, Sastri, 1994, Schettini et Bréheret, 1996 et 1997).

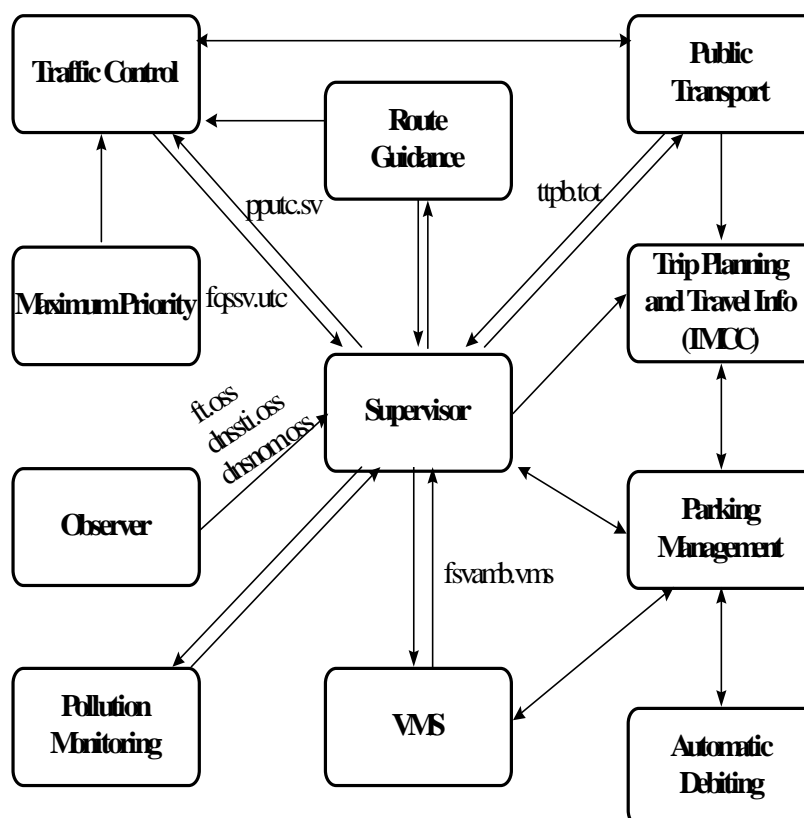
#### 4.4 Les systèmes existants

L'objectif de ce paragraphe est de citer quelques exemples de systèmes automatiques de traitement de données opérationnels.

Les systèmes d'exploitation à ce jour sont supervisés par des opérateurs. En effet, un opérateur peut corriger les réglages des différents algorithmes en fonction d'informations événementielles (accidents, manifestations, priorités, etc.). Pour que l'opérateur soit pleinement efficace, il faut en particulier qu'il soit informé au mieux par un système de traitement des données et de surveillance qui a aussi vocation d'être utilisé par les systèmes d'information des usagers de la route.

Les centres d'exploitation du trafic en milieu urbain ou périurbain sont équipés, au moins pour les grandes agglomérations, de systèmes automatiques de surveillance en temps réel de la circulation. Ces systèmes peuvent être plus ou moins sophistiqués (le système SIRIUS en Ile-de-France). De récents projets ont donné le jour à des systèmes opérationnels de régulation et de surveillance. A titre d'exemple, on peut citer :

- le système CLAIRE (T. Dallery, 1996), développé par l'Institut National de Recherche sur les Transports et leur Sécurité (INRETS) et industrialisé par CGA, qui est un système expert pouvant se greffer sur certains systèmes de régulation. Il détecte la saturation, en détermine les causes, les mémorise et propose aux opérateurs des actions contre la saturation (favoriser ou retenir un mouvement donné). A ce jour, CLAIRE est installé à Paris, Rennes, Toulouse et à Londres où il constitue un niveau supérieur du système de régulation SCOOT.
- le système CONCERT qui est le fruit de recherches et de développements de SIEMENS, propose des outils permettant de réguler, de détecter ainsi que traiter la saturation. Il intègre également des fonctions plus évoluées comme l'estimation des flux de circulation entre les origines et destinations, la recherche des meilleurs itinéraires, de la prédiction des états du trafic à long terme (d'un jour à l'autre, etc.). Toutes ces fonctions peuvent servir aussi bien à la surveillance du trafic qu'à la diffusion d'information.
- le système 5T (Integrated Road transport Environment) développé à Turin en Italie, dans le cadre de projets européens (projet Quartet Plus par exemple) s'appuie sur l'intégration physique et fonctionnelle de 10 systèmes de commande et de surveillance du transports et du trafic. Ces systèmes sont représentés sur la figure ci-dessous :



## 5 Spécifications d'un système de traitement de données

Ce chapitre décrit les spécifications techniques pour le développement et la mise en œuvre d'un système de traitement de données capable de proposer des fonctions d'exploitation avancées aux gestionnaires de trafic et permettant de diffuser des informations fiables aux usagers de la route.

Le cahier des charges du système de traitement de données appliqué au cas de l'agglomération toulousaine est tout d'abord présenté. La partie suivante précise les fonctions à intégrer ainsi que les modèles théoriques à utiliser. La dernière partie spécifie les éléments techniques pour la mise en œuvre informatique mais aussi les étapes nécessaires pour le déploiement in situ. Ces spécifications sont adaptées au PC Campus de la ville de Toulouse mais elles peuvent facilement s'étendre à tout autre centre d'exploitation de réseaux urbains et périurbains.

### 5.1 Cahier des charges

L'analyse des besoins des exploitants Ville de Toulouse et DDE 31, en matière de traitement de données a fait ressortir l'utilité d'étudier et de développer de nouveaux systèmes de traitement de données. En effet, les sources de mesure du trafic (stations de comptage installées dans la chaussée, données historiques, données des véhicules traceurs, etc.) sont variées et fournissent des informations à la fois quantitatives et qualitatives. De plus, le développement des moyens de communication donne une nouvelle dimension à l'usage des mesures de trafic en temps réel dans la chaîne de traitement de l'information.

L'un des besoins exprimés par les gestionnaires de trafic est de disposer d'outils informatiques de supervision du trafic afin :

- de connaître automatiquement et à chaque instant les niveaux de circulation sur les différents types de réseaux (urbain et périurbain),
- d'avoir des indicateurs permettant de connaître l'évolution du trafic à court terme.

Ces deux points permettraient à l'exploitant de mettre de place des stratégies adaptées à la situation de circulation observée en tenant compte des conditions à venir.

Le système de traitement de données, défini dans ce rapport, concerne la mise en place d'un système informatique répondant aux deux points précédents. Pour cela, cette étude vise à spécifier des méthodes numériques et un système informatique pour l'utilisation des données fournies par des systèmes usuels de mesures du trafic (stations de comptage composées de boucles électromagnétiques) et par des informations du type «temps de parcours» fournies par des véhicules traceurs sur certaines sections du réseau



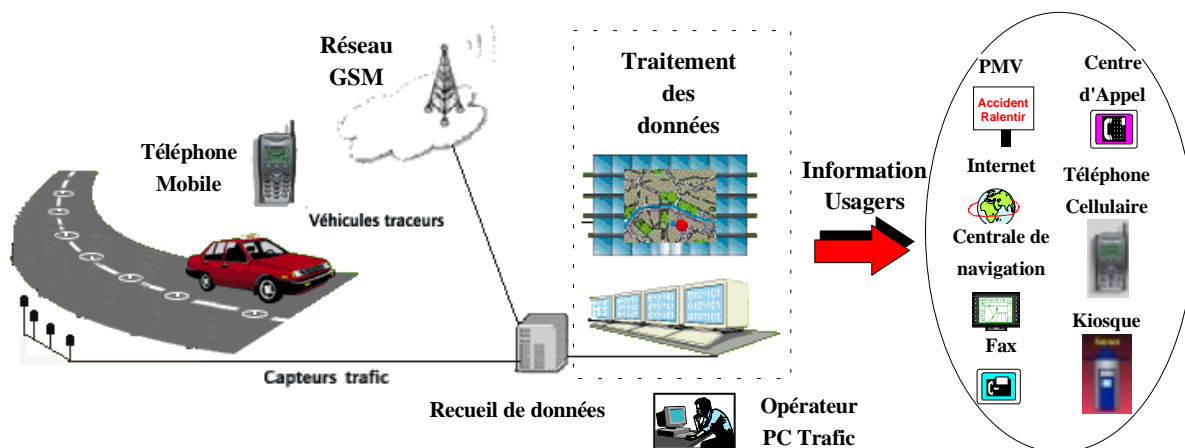


Figure 4 : Synoptique général du système de traitement des données

### 5.1.1 Concept général

La revue des systèmes de surveillance montre que les nouvelles techniques de l'informatique ont permis le développement d'un certain nombre de fonctionnalités. Un système de traitement de données s'articule autour de trois fonctions principales :

- la qualification des données observées,
- le filtrage et l'estimation des états du trafic par fusion de données,
- la prédiction dans le futur proche.

La figure suivante schématise le système de traitement de données étudié. On distingue les modes de fonctionnements hors-ligne et en ligne. Le mode hors-ligne consiste en une configuration du système, en particulier pour certains modèles utilisés en ligne, et nécessite une identification ou un calibrage préalable. Le mode en ligne correspond aux fonctions étudiées dans ce rapport.

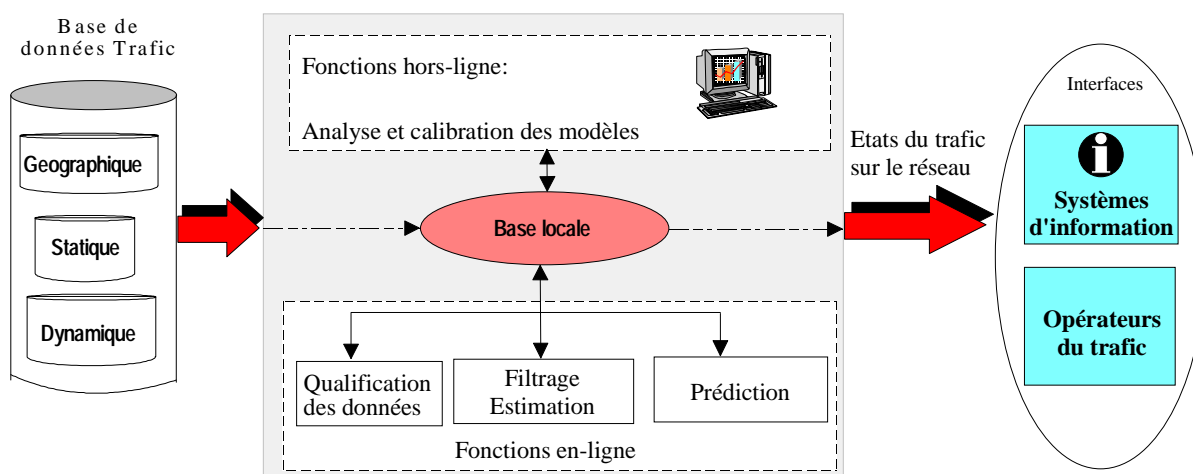


Figure 5 : Schéma bloc fonctionnel du système de traitement de données

### 5.1.2 La base de données trafic

Le système de recueil du PC trafic a pour objectif de centraliser les données dynamiques dans différentes bases de données trafic ; en pratique, celles-ci sont le plus souvent relationnelles. La première fonction du système de traitement de données, est d'extraire les informations les plus importantes pour alimenter l'ensemble des modules et permettre l'échange de données entre eux. Toutes ces informations ainsi que les résultats temporaires des traitements sont sauvegardés dans une base propre au système. Les fonctions d'extraction des données et de gestion de la base de données locale seront abordées dans le chapitre Mise en œuvre.

Les données nécessaires au fonctionnement des modèles de qualification des données, de filtrage/estimation et de prédiction sont de plusieurs natures : géographique, dynamique et statique.

#### *Les données géographiques*

La représentation du réseau routier dépend fortement du niveau d'agrégation de la cartographie. Pour la surveillance du trafic, une modélisation macroscopique du trafic est suffisante. C'est-à-dire que l'objet élémentaire est le chaînon qui décrit un segment de route. Il peut comporter plusieurs voies de circulation. Chaque chaînon est délimité par deux nœuds qui correspondent aux intersections amont et aval. Celles-ci peuvent présenter des feux. Les principales informations sont la position géographique des nœuds, les caractéristiques de chaque chaînon (longueur, nœuds amont et aval, type de voie, etc.), la position des feux tricolores et la position des points de mesures du trafic.

La base de données géographiques est décomposée en trois couches différentes et suivant une approche orientée objet. Chaque couche comprend des entités différentes qui servent à décrire le réseau routier. Un ensemble de relations permet par ailleurs de faire le lien entre chaque couche. Le contenu et le rôle de chaque couche sont les suivants :

- La **couche I** sert principalement à décrire les caractéristiques techniques de chaque poste de mesure (PM) et des carrefours à feux. On y trouve des informations telles que le numéro, le nombre de boucles, et le type des variables mesurées pour chaque PM ainsi que la position des carrefours à feux et la liste des phases associées à chacun.

Cette couche sert au recueil des données provenant des différents systèmes de mesure.

- La **couche II** modélise le réseau routier sous la forme d'un graphe orienté par l'intermédiaire de chaînons et de nœuds. Les données contenues dans cette couche sont la géométrie de la voirie (longueur, sens de circulation et nombre de voies des chaînons élémentaires) et la relation existante avec la couche I (PM, carrefour à feux et phases associés aux chaînons). La couche II spécifie aussi les variables macroscopiques clés qui décrivent les flots de circulation (capacité maximale, vitesse libre et vitesse à la saturation).

Cette couche propose des traitements élémentaires (affectation des mesures aux chaînons et sélection des meilleurs historiques pour une situation donnée) mais surtout des traitements plus évolués (qualification des données, estimation du temps de parcours (TP) élémentaire et prédiction à court terme).

- La **couche III** propose une représentation plus macroscopique du réseau routier que la couche II. Cette couche modélise sous la forme d'un graphe orienté des pôles qui représentent des choix d'itinéraires ou des origines/destination sur le réseau et des trajets qui modélisent des itinéraires entre ces points de choix. Un trajet comprend un ou plusieurs chaînons éléments de la couche II et de la même manière un pôle symbolise un conglomérat de nœuds.

Cette couche assure des fonctions d'agrégation de certains indicateurs du trafic. Ces niveaux d'agrégation peuvent être spatiaux pour estimer les flux de circulation et permettent de prendre en compte des phénomènes tels que les ondes vertes sur les axes importants.

Le tableau et la figure ci-dessous résument les éléments caractéristiques de la base de données géographiques.

Eléments caractéristiques du réseau

Couche	Elément
I	poste de mesure
	feux de circulation
II	nœud
	chaînon
III	pôle
	trajet

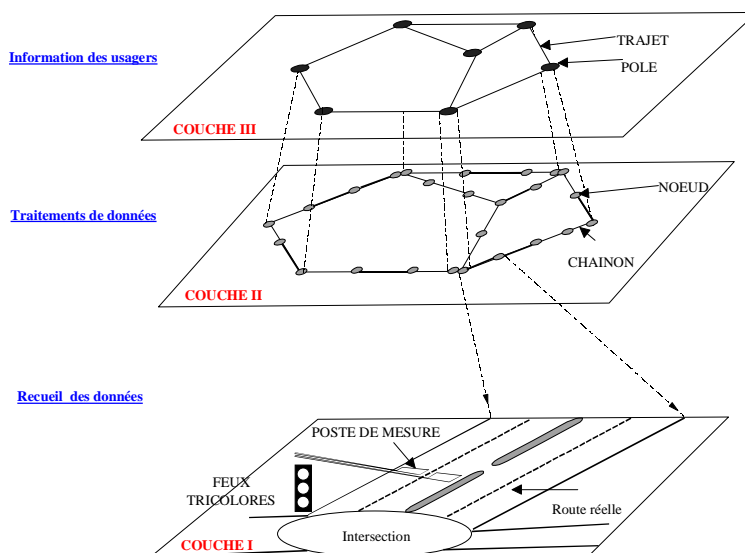


Figure 6 : Relation entre les différents éléments de la base de données géographiques

### Les données dynamiques de comptage

Ce sont essentiellement celles qui sont fournies par les postes de mesure. Ces données sont primordiales car elles permettent de connaître, en certains points du réseau, le débit de véhicules, le taux d'occupation et parfois la vitesse. La période de rafraîchissement de ces données varie suivant le site d'exploitation. Elle est généralement comprise entre 1 minute et 15 minutes.

Des informations complémentaires peuvent être utilisées pour calculer dynamiquement l'état du trafic. C'est le cas des durées de cycles des feux de circulation qui fluctuent durant la journée (application de différents plans de feux pour les heures de pointes, ou optimisation des plans en ligne). Comme nous nous plaçons dans le cas d'un trafic décrit de façon macroscopique, les informations sur les temps de cycle, lissées sur plusieurs cycles, permettent d'estimer le temps perdu à la fin d'un chaînon.

### Les données dynamiques provenant des véhicules traceurs

Plusieurs fournisseurs de données provenant de véhicules traceurs peuvent être envisagés sur l'agglomération toulousaine : les taxis, les bus et les téléphones mobiles.

Les données provenant des taxis permettent d'obtenir une bonne couverture de l'ensemble de la voirie mais seulement 200 à 300 véhicules sont équipés. Le système de localisation par zone est actuellement utilisé pour la gestion des appels et l'affectation des taxis aux appels ainsi que pour la sécurité des chauffeurs. Il n'y a pas d'enregistrement des temps de parcours et l'obtention de cette variable par calcul présente des incertitudes. Les taxis peuvent en effet utiliser des voies réservées, leur trajet est inconnu et ils n'ont pas un comportement standard par rapport aux véhicules particuliers.

Le système de localisation utilisée pour les bus permet le calcul des temps de parcours et l'informations des usagers. Le temps de parcours calculé inclut les arrêts aux arrêts de bus (montées et descentes). D'autre part, le trajet des bus est fixé et peut emprunter des voies réservées. La couverture géographique du réseau est incomplète et de limite aux lignes de bus.

La téléphonie mobile offre également des possibilités de recueil de données. Le nombre de véhicules équipés est important et croissant (30% de la population française est équipée de téléphone portable). Cette technologie évolue vite et pour l'instant la localisation n'est pas instantanée et n'est pas accessible par un système externe aux opérateurs de télécommunications. L'ouverture de ce type de système est néanmoins possible par un accord avec l'opérateur de télécommunications.

Dans un objectif à moyen et long terme, nous avons fait abstraction de la source de données temps de parcours et nous nous sommes basés sur l'hypothèse que ces données provenaient du GSM (voir paragraphe 5.2.1.2).

### *Les données statiques*

Dans cette catégorie, différents types de données peuvent être utilisés pour substituer l'information dynamique en cas de perte de données. Elles permettent aussi d'interpoler l'état du trafic sur l'ensemble du réseau là où les données dynamiques ne sont pas disponibles. Il existe :

- Des données historiques qui représentent des jours standards de circulation, c'est-à-dire qu'en situation normale, le trafic peut être reproductible d'un jour à l'autre. Ces données sont considérées comme statiques car elles n'évoluent que très lentement. Elles servent essentiellement à pallier le manque d'informations (pannes, erreurs de mesure, etc.) pour assurer une continuité dans la surveillance du trafic.
- Une ou plusieurs matrices qui permettent d'identifier la demande de mobilité des usagers sur un réseau routier. Une matrice Origine Destination (OD) est réalisée généralement à partir d'enquêtes terrains pour connaître les déplacements journaliers des usagers. L'information contenue dans cette matrice est le nombre de véhicules circulant d'une origine vers une destination pour un jour et une heure déterminés. D'autres informations peuvent enrichir la modélisation du trafic. Les plus fréquentes sont les pourcentages directionnels, qui indiquent le taux de mouvements tournants des véhicules sur les carrefours importants, ainsi que la mesure des temps de parcours sur certains itinéraires particuliers.

### **5.1.3 Les fonctions principales**

Ce paragraphe détaille uniquement le rôle des fonctions en ligne (qualification des données, filtrage et estimation par fusion de données et prédiction des états du trafic) qui doivent être intégrées dans le système de traitement des données. La figure ci-dessous précise l'interaction des trois fonctions en ligne dans le temps. A un instant courant  $t$ , la fonction de qualification des données permettra de détecter et de corriger les erreurs de mesure du système de recueil de données à partir des mesures du proche passé sur un horizon  $H_f$ . La fonction de filtrage et d'estimation utilise les mesures fournies par les différentes sources pour déterminer les états de trafic à  $t$  en se basant sur les mesures corrigées de l'instant courant et des instants précédents. La fonction de prédiction permet de déterminer l'évolution sur un horizon de prédiction  $H_p$  des états de trafic à partir d'un instant donné  $t$ .

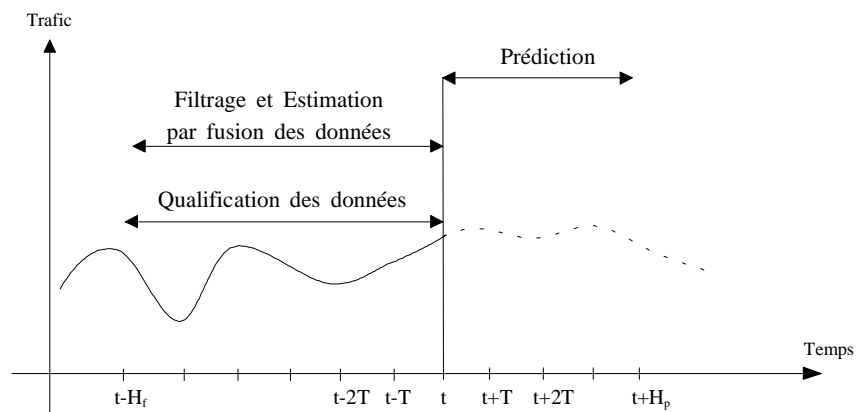


Figure 7 : Interaction des fonctions dans le temps

### **La fonction de qualification des données dynamiques**

La maîtrise de la disponibilité et de la qualité des données dynamiques dans la base de données locale est un problème important dans les systèmes modernes de traitement de données du trafic. La qualification des données dynamiques précède les autres fonctions du système pour s'assurer que des données erronées ne sont pas utilisées dans les fonctions d'estimation et de prédiction des niveaux de circulation. En pratique, les problèmes proviennent soit d'un mauvais fonctionnement des postes de mesure, soit d'une erreur de transmission ou d'une situation de trafic incongrue (stationnement illicite d'un véhicule sur une boucle, travaux, voie fermée, etc.).

A ce jour, peu de méthodes se sont intéressées à la mesure de la qualité des données de comptage du trafic. Les quelques techniques décrites dans la littérature, utilisent des algorithmes basés sur la vérification de la cohérence des mesures et proposent une méthode de correction en cas d'anomalie.

### **La fonction de filtrage et d'estimation des états du trafic**

Deux types d'informations sont attendus à la fois par les exploitants des réseaux et les usagers de la route. Il s'agit des états permanents de la circulation (temps de parcours et niveau de service) mais aussi des données sur les débits de véhicules, les longueurs de files d'attente, le niveau de saturation et la vitesse.

Cette étude s'intéresse plus particulièrement au développement d'une méthode de filtrage et d'estimation d'états du trafic, pour calculer à la fois, les temps de parcours sur les axes routiers et les niveaux de trafic, à partir de la mesure des variables macroscopiques (flux de véhicule, taux d'occupation et vitesse) et des données temps de parcours émises par des véhicules traceurs.

Les algorithmes de filtrage et d'estimation des états du trafic cherchent à déterminer des indicateurs du trafic non observables. Il s'agit, avant tout, du temps de parcours entre points importants du réseau. En effet, le temps de parcours est un indicateur assez révélateur de l'état de la circulation et constitue un enjeu important en matière d'information routière notamment en période de congestion et d'incident. Il permet notamment de détecter les niveaux de trafic en comparant les temps de parcours estimés à des situations de référence. Bien qu'il existe de nombreux modèles analytiques d'estimation du temps de parcours, ces modèles sont souvent peu satisfaisants soit parce qu'ils représentent mal la réalité, soit parce qu'ils sont mal utilisés.

### **La fonction de prédiction**

Cette fonction sert à prédire l'évolution des principaux indicateurs du trafic dans le futur proche. Elle a pour objet de calculer les valeurs des variables du trafic (débit de véhicules,

taux d'occupation, files d'attente, vitesse, temps de parcours) dans le futur proche à  $t+T$ ,  $t+2T$ , ...,  $t+h$  où  $t$  est l'instant courant,  $T$  la période d'échantillonnage et  $h$  est l'horizon de prédiction qui doit être un multiple de  $T$ .

La plupart des méthodes relatives à la prédiction du trafic sont basées sur l'analyse des séries temporelles que forment les observations du trafic. En s'appuyant sur les études passées et grâce aux nouvelles techniques numériques, un modèle adapté aux phénomènes du trafic routier urbain et périurbain, donnera l'évolution du trafic pour les instants futurs à partir des observations jusqu'à  $t$ .

#### 5.1.4 Le système de fusion de données étudié

Pour conclure cette section, le tableau ci-dessous résume l'ensemble des données qui sont requises en entrée de chaque module ainsi que les données de sortie qu'ils produisent. Ce tableau montre également la méthode de calcul envisagée pour la réalisation de chaque fonction.

Récapitulatif des fonctions en ligne du système de traitement de données

Fonctions	Données d'entrée	Données de sortie	Méthodes étudiées
<i>Qualification des données</i>	Données dynamiques : <ul style="list-style-type: none"> <li>débit, taux d'occupation et vitesse pour l'ensemble des PM de <math>t-H_f</math> à <math>t</math></li> <li>temps de parcours des véhicules traceurs</li> </ul> Données géographiques <ul style="list-style-type: none"> <li>couches I et II</li> </ul> Données statiques <ul style="list-style-type: none"> <li>profils historiques par PM</li> </ul>	Données dynamiques corrigées à $t$ : <ul style="list-style-type: none"> <li>Débit, taux d'occupation et vitesse pour l'ensemble des PM</li> <li>temps de parcours des véhicules traceurs</li> </ul>	Procédure heuristique qui reconstitue les données manquantes ou détecte et corrige les données erronées
<i>Filtrage et Estimation par fusion de données</i>	Données dynamiques corrigées : <ul style="list-style-type: none"> <li>débit, taux d'occupation et vitesse pour l'ensemble des PM de <math>t-H_f</math> à <math>t</math></li> <li>Temps de parcours des véhicules traceurs</li> </ul> Données géographiques <ul style="list-style-type: none"> <li>couches I, II et III</li> </ul>	Données estimées à $t$ : <ul style="list-style-type: none"> <li>longueur des files d'attente, débit de véhicule, temps de parcours pour les chaînons équipés de PM</li> <li>Temps de parcours pour l'ensemble des chaînons</li> <li>Temps de parcours et niveau de trafic estimés pour l'ensemble des trajets du réseau</li> </ul>	Méthode d'optimisation basée sur des modèles analytiques
<i>Prédiction</i>	Données dynamiques : <ul style="list-style-type: none"> <li>débit et taux d'occupation, vitesse filtrés/estimés pour l'ensemble des PM de <math>t-H_f</math> à <math>t</math></li> <li>temps de parcours filtrés/estimés ou corrigés pour tous les chaînons du réseau de <math>t-H_f</math> à <math>t</math></li> </ul> Données géographiques <ul style="list-style-type: none"> <li>couches II et III</li> </ul> Données statiques <ul style="list-style-type: none"> <li>profils historiques par PM</li> <li>profils historiques des temps de parcours des chaînons</li> </ul>	Données prédites de $t$ à $H_p$ l'horizon de prédiction : <ul style="list-style-type: none"> <li>longueur des files d'attente, débit de véhicule, temps de parcours pour les chaînons équipés de PM</li> <li>Temps de parcours pour l'ensemble des chaînons</li> <li>Temps de parcours et niveau de circulation pour l'ensemble des trajets du réseau</li> </ul>	Analyse et prédiction de séries temporelles à partir de : <ul style="list-style-type: none"> <li>méthode basée sur le filtre de Kalman</li> <li>méthode basée sur l'intelligence artificielle (optimisation par réseau de neurone)</li> </ul>

### 5.1.5 Les notations utilisées

Les notations utilisées pour les équations mathématiques sont les suivantes :

- chaînon = portion de route entre deux intersections, les véhicules roulant dans le même sens (couche II)

$Q$  = débit, en véh/s

$TO$  = taux d'occupation, en %

$V$  = vitesse, en m/s

$TP$  = temps de parcours, en s

$X$  = nombre de véhicules en queue, en véh

$L$  = longueur, en m

$f$  = nombre de voies

$cap$  = capacité, en véh/s

$s$  = débit de saturation, en véh/s

$c$  = durée de cycle feux, en s

$r$  = durée de rouge feux, en s

$TPl$  = temps de parcours à vitesse libre

- $d$  = longueur moyenne en mètre d'un véhicule (en zone urbaine)
- $l$  = longueur en mètre d'une boucle magnétique (sens chaînon)
- $T$  = durée de la période d'échantillonnage, en s (égale à 3 minutes soit 180 s)
- $Hf$  = horizon de filtrage, en nombre de périodes, paramètre de la méthode de filtrage
- $Hp$  = horizon de prédiction, en nombre de périodes, paramètre de la méthode de prédiction

#### Remarques :

- Pour la mise en œuvre in situ, les unités (véh/h, km/h, km, %) plus usuelles pour les exploitants seront utilisées.
- L'indice  $c$  d'une variable indique qu'il s'agit d'une variable corrigée par la fonction de qualification des données. Tandis que l'indice  $p$  précise qu'il s'agit d'une variable prédite.

## 5.2 Les fonctions

Cette partie décrit pour chacune des fonctions identifiées précédemment (qualification des données, filtrage et estimation, prédiction) les données d'entrée nécessaires, les modèles de trafic utilisés avec leur méthode de résolution associée et enfin les données de sortie générées par la fonction.

## 5.2.1 La qualification des données

### 5.2.1.1 Les données de comptage

Cette fonction vérifie, en temps réel, la cohérence des observations faites par les PM et remédie aux pannes de recueil des données du trafic. Le taux d'erreur de mesure est non négligeable dans les sites d'exploitation. En effet, il est fréquent de trouver des taux d'erreur de 5% ou plus dans les systèmes de recueil de données actuels. Ces erreurs proviennent essentiellement de pannes de transmission entre le PM et le système de recueil ou directement de la mise en alarme du PM lui-même. D'autres erreurs s'expriment par le fait que les PM peuvent fournir des mesures aberrantes qui ne permettent pas un fonctionnement optimal des autres modules qui composent le système de traitement des données. Il est donc très important de détecter et de corriger les observations du trafic et de remplacer les valeurs anormales le cas échéant.

Le principe de la qualification des données de comptage revient à décomposer le processus en deux étapes distinctes :

- La première pour détecter les données erronées fournissant un journal d'alarme des postes de mesure (à corriger, en panne, valides),
- La seconde pour reconstituer les données manquantes ou aberrantes.

Ce processus est activé à chaque nouvelle période d'échantillonnage  $t$  et pour l'ensemble des PM. La Figure 8 schématise ce processus de qualification des données où  $Y$  représente l'ensemble des mesures d'un PM.

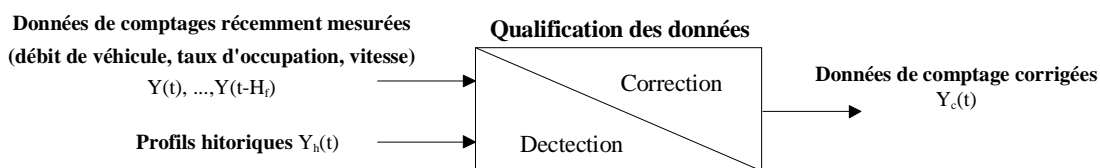


Figure 8 : Principe de la qualification des données

Les données dynamiques d'entrée sont les suivantes :

- les valeurs mesurées toutes les 3 minutes de débits de véhicules  $Q$  et de taux d'occupation  $TO$  pour chaque poste de mesure du réseau de la Ville de Toulouse
- les valeurs mesurées toutes les trois minutes de débits de véhicules  $Q$ , de taux d'occupation  $TO$  et de vitesse  $V$  pour chaque poste de mesure du réseau de la DDE 31

Les données géographiques utilisées sont celles de la couche I (liste des PM du réseau) et de la couche II (liaison entre les PM et les chaînons).

Les données statiques sont les suivantes :

- les valeurs historiques de débits de véhicules  $Q$  et de taux d'occupation  $TO$  par chaque poste de mesure du réseau urbain (Ville de Toulouse)
- les valeurs historiques de débits de véhicules  $Q$ , de taux d'occupation  $TO$  et de vitesse  $V$  par chaque poste de mesure du réseau de voies rapides (DDE 31)

Le principe de base de tout redressement de données est que l'information contenue dans la base de données est fortement redondante. En effet, il existe des régularités et des périodicités facilement mesurables d'un jour à l'autre ou d'une semaine à l'autre. Il existe



également des corrélations spatiales entre différents points de réseau. C'est cette redondance de l'information qui va permettre de tester la vraisemblance d'une donnée par rapport à son environnement et de remplacer toute donnée anormale par une valeur vraisemblable.

### **La détection des erreurs de mesures**

Les problèmes principaux des systèmes de recueil des données sont soit l'absence de certaines données, soit la présence de données aberrantes. Tout d'abord, il est nécessaire de définir précisément ces deux notions propres au domaine du trafic :

- les **données manquantes** correspondent essentiellement à des pannes du système de détection ou à des pertes de données lors de transmission.

La procédure de détection des données manquantes revient à détecter des valeurs particulières indiquant que le PM est en alarme ou à vérifier que les variables macroscopiques sont en dehors d'un intervalle admissible physiquement. A contrario, la mesure admissible doit être à l'intérieur d'un intervalle. Par exemple :  $0 \leq Q \leq Q_{max}$ ,  $0 \leq TO \leq TO_{max}$  et  $0 \leq V \leq V_{max}$ .

- Les **données aberrantes diverses** (rebond de détecteur, problèmes de télécommunication, etc.). Leur détection est un problème très délicat, car il faut faire la différence entre une mesure qui ne correspond à aucun phénomène physique, et la mesure d'un phénomène physique exceptionnel (accident par exemple).

La question de la détection des données aberrantes est plus difficile à résoudre. La solution envisagée ici, réside dans la redondance de l'information disponible : il est raisonnable de penser qu'un phénomène de trafic ne peut se produire isolément dans le temps et l'espace. On considérera comme aberration de mesure toute variation significative se produisant :

- pendant une seule période de mesure,
- en un seul point.

Une vérification temporelle (Test 1) des données est effectuée pour s'assurer que les nouvelles mesures sont cohérentes par rapport aux dernières mesures enregistrées. Cette vérification détermine si une variation a lieu pendant un instant donné. Une seconde vérification (test 2), dite spatiale, détermine si la variation est simplement locale au PM où si celle-ci est observable sur les PM voisins de celui considéré. Enfin, un troisième test (test 3) permet de déterminer si la donnée respecte les lois du trafic. Ces dernières s'expriment par la relation fondamentale du trafic.

La méthode basée sur des techniques statistiques, revient à calculer trois estimateurs d'états pour chaque PM,  $Y_{c1}(k)$ ,  $Y_{c2}(k)$  et  $Y_{c3}(k)$  ainsi que leur intervalle de confiance. Cette méthode vérifie ensuite si la mesure  $Y(k)$  est comprise ou non dans chacun des intervalles de confiance des estimateurs. L'estimateur  $Y_{c1}(k)$  fournit une estimation des variables  $Q$ ,  $TO$  et  $V$  à l'instant  $k$ , en utilisant la corrélation temporelle des termes issus du proche passé et ceux provenant d'un historique. Tandis que  $Y_{c2}(k)$  donne une estimation de chaque variable macroscopique à partir des observations des PM voisins. Enfin,  $Y_{c3}(k)$  estime un point situé sur la courbe théorique identifiée lors de la phase de calibrage. Le détail de ces estimateurs est fourni dans l'annexe 8.3.

### La correction des mesures erronées

Les corrections peuvent être de trois natures :

- Pour une donnée manquante, on la remplace dans un premier temps par son estimateur temporel  $Y_{c1}(k)$ .
- Pour une donnée ne satisfaisant pas à la fois les tests 1 et 2, on la remplace par son estimateur spatial  $Y_{c2}(k)$ .
- Pour une donnée ne satisfaisant pas le test 3, on la remplace par sa projection sur la courbe  $(y, to)$ .

On doit noter qu'à partir du moment où il n'y a qu'un seul PM qui ne satisfait pas les tests 1 et 2, l'estimateur spatial relatif à un PM est fonction des corrections de données effectuées au niveau des autres PM. On cherche à ajuster cette correction définie implicitement au travers d'un processus de calcul itératif.

### Données de sortie

Pour chaque PM, les données générées par la fonction de qualification sont les données corrigées ainsi que les états du PM. Le tableau ci-dessous indique les différentes données générées à chaque nouvelle période de mesure

Tableau 2: Données de sortie générées par la fonction de qualification des données pour chaque PM

Etat PM	Données PM	Commentaire
Normal	Q, TO, V	Cas des données cohérentes
Corrigées	Qc, TOc, Vc	Cas où l'une des données a été corrigée
Alarme	-1,-1,-1	Cas où les données erronées n'ont pas pu être corrigées (PM en panne depuis une période trop longue)

#### 5.2.1.2 Les données des véhicules traceurs

Nous rappelons que dans cette étude les véhicules traceurs considérés sont ceux pour lesquels l'utilisateur est équipé de terminaux GSM.

La fonction de qualification des données revient à utiliser les données des positions des téléphones mobiles GSM (localisation par les balises GSM), afin de déduire des informations sur le trafic (temps de parcours). La figure ci-dessous propose un schéma bloc des traitements effectués.

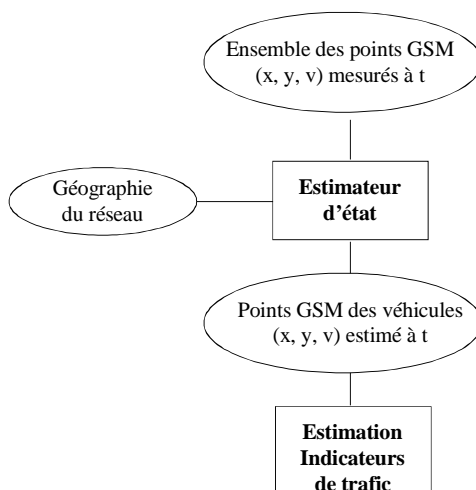
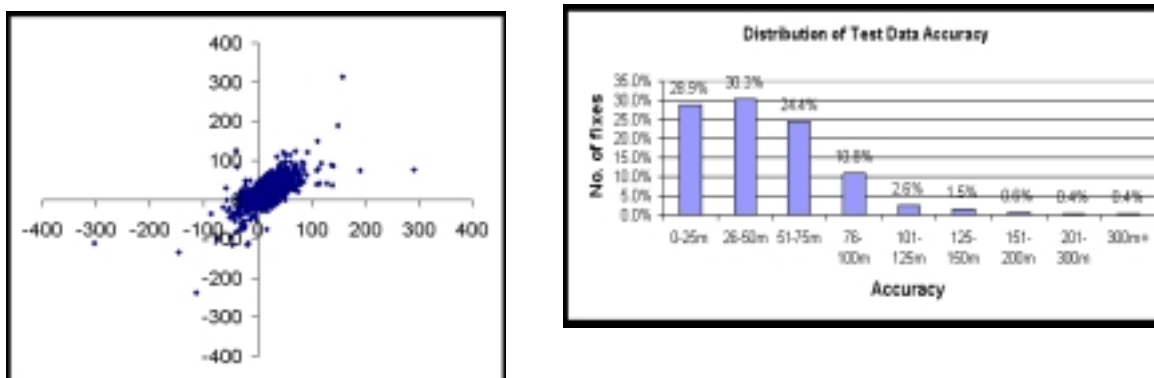


Figure 9: Schéma bloc fonctionnel de la qualification des données véhicules traceurs

### Données d'entrée

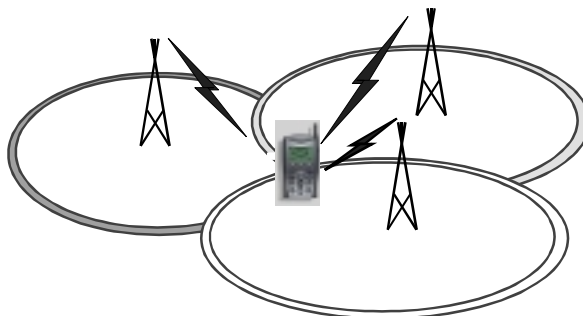
La localisation des téléphones mobiles est un domaine en plein essor et de nombreux systèmes ont déjà été mis au point surtout aux Etats-Unis. Les performances, en terme de précision de localisation, sont très variables (de 50 m à 125 m). La mise en service de ces systèmes est prévue entre la fin de l'année 1999 et l'année 2001.

Par exemple, la précision du système Cursor est représentée par les statistiques suivantes.



Par l'intermédiaire des opérateurs GSM, on peut obtenir<sup>2</sup> la position du mobile localisée par intersection entre les trois cercles dont les rayons correspondent aux distances approchées.

<sup>2</sup> A noter que l'on ne considère dans cette étude que la faisabilité technique. Les aspects institutionnels pour traiter des problèmes de confidentialité et de protection des libertés, ne sont pas explicités, bien que les auteurs soient tout à fait conscient qu'il s'agit là d'un préalable à toute utilisation de telle données.



### *Méthode et modèle utilisés*

A chaque période d'échantillonnage, un ensemble de  $n$  séries de données est centralisé par l'opérateur GSM. Les temps de parcours des véhicules traceurs par chaînon traversé sont estimés à partir de cet ensemble de points GSM provenant de mobiles (piéton, véhicule, etc.). La procédure de qualification des temps de parcours revient :

- à filtrer tous les points GSM aberrants ne correspondant pas à des points de véhicules,
- à estimer les temps de parcours à partir des points GSM filtrés.

La méthode mise en place dans cette étude consiste à prendre l'hypothèse que l'état du trafic (très encombré ou non...) est connu et que cet état reste stable durant la période d'échantillonnage  $T$ .

### *Filtrage préliminaire*

A la fin d'une période d'échantillonnage, on classe les points GSM pour les mobiles qui ont fourni au moins deux points suivant les  $m$  vitesses et positions mesurées successives.

- Les vitesses supérieures aux bornes maximales sont éliminées, car leur probabilité est pratiquement nulle sur des axes urbains ou périurbains ; ces vitesses élevées peuvent être introduites par des déplacements plus rapides (train, métro par exemple).
- On supprime les mobiles dont la vitesse est nulle durant toute la période, c'est à dire dont les  $n$  vitesses sont nulles, en considérant que ce sont des voitures arrêtées ou des téléphones se trouvant dans les habitations alentour ; ceci à condition que l'on ne soit pas dans un cas de vrai bouchon.
- On attribue une enveloppe à chaque route (ici 300 mètres). Si le mobile est sorti au moins une fois de l'enveloppe durant la période, on l'élimine également. D'où l'importance des performances du dispositif de localisation. En effet, dans le cas d'une route très proche du chaînon étudié, et d'une zone de localisation limitée, il n'est pas possible de repérer le mobile très au-delà du chaînon pour déterminer s'il a quitté l'enveloppe ou non.

On stocke pour chaque chaînon les vitesses de tous les mobiles qui l'ont éventuellement traversé. Puis on classe ces vitesses par tranches de vitesse (5 Km/h par exemple), et on obtient donc une distribution des vitesses sur ce chaînon. Un exemple pour un chaînon à deux voies est donné en annexe 8.4.

### *Calcul du temps de parcours*

On calcule le temps de parcours à partir de la moyenne harmonique, et non pas arithmétique, des vitesses. On obtiendra un temps de parcours global pour les deux voies dans le cas d'une seule distribution et un par voie dans le cas de deux distributions. Pour le calcul des indicateurs réels, on obtient un temps de parcours par voie systématiquement, que l'on ramène à un temps de parcours global si nécessaire pour la comparaison. Des résultats de tests figurent en annexe 8.4.

### **Données de sortie**

A chaque période d'échantillonnage, les données générées par la fonction de qualification des données des véhicules traceurs sont les temps de parcours des chaînons traversés avec leur écart type. On obtient donc à l'instant  $k$ , le temps de parcours corrigé  $TP_{c,i}(k)$  où  $i$  est un chaînon du réseau sur lequel au moins un véhicule a fourni deux informations de position entre  $k-1$  et  $k$ .

## **5.2.2 Le filtrage et l'estimation par fusion de données**

L'objectif de cette fonction est de reconstituer les états du réseau sur le proche passé en fonction des données reçues dynamiquement à chaque période d'échantillonnage et d'estimer l'état actuel.

La méthode de fusion de données doit permettre l'estimation :

- des temps de parcours des chaînons élémentaires (couche II),
- des temps de parcours pour les trajets (couche III),
- des niveaux de circulation pour les trajets (couche III).

La méthode de fusion de données proposée permet de prendre en compte différents types de données. Des modèles de trafic permettent de relier les différentes variables du trafic entre elles.

### *5.2.2.1 Données d'entrée*

La fonction de qualification des mesures a déjà corrigé les données aberrantes et reconstitué les données manquantes sur la période précédente, et donc sur l'horizon de filtrage. De même, les données temps de parcours ont été filtrées par l'opérateur GSM. On sait par expérience que la qualité de ces données n'est pas parfaite. La valeur d'un type de donnée est connue avec une précision exprimée par un écart type. Ces écarts types sont connus étant donné le moyen de mesure :

- pour les débits, les taux d'occupation et les vitesses, valeurs fournies par la Ville de Toulouse et par la DDE31 en fonction du type de poste de mesure sur chaque chaînon
- pour les temps de parcours GSM, valeurs fournies en temps réel par l'opérateur GSM

Au niveau d'un chaînon, plusieurs scénarios (Tableau 3) sont identifiés en fonction des mesures disponibles sur l'horizon de filtrage.

Tableau 3 : Scénarios pour le filtrage et l'estimation au niveau chaînon

Mesures disponibles sur l'horizon de filtrage	urbain	périurbain
poste de mesure (+ feux)	$Q, TO, (+ r, c)$	$Q, TO, V$
poste de mesure + GSM	$Q, TO, TP$	$Q, TO, V, TP$

GSM seulement	$TP$	$TP$
aucune mesure	$\emptyset$	$\emptyset$

Les données de la période passée ne sont pas suffisantes à elles seules pour qualifier les différents cas. En effet, la méthode de filtrage travaille sur plusieurs périodes continues avec une évolution de la file d'attente dans le temps. Si une mesure manque sur l'horizon, la méthode peut travailler avec les autres données et les autres périodes.

Lorsqu'un chaînon urbain possède des feux sur son intersection aval, le modèle de temps de parcours utilise les données des feux, c'est-à-dire les durées de rouge et de cycle. Celles-ci peuvent être centralisées dynamiquement et utilisées par la méthode de fusion ou une valeur moyenne peut être calibrée pour les calculs.

Pour chaque chaînon, la vitesse libre ou le temps de parcours à vitesse libre doit avoir été évalué (la longueur du chaînon étant connue).

### 5.2.2.2 Modèles de trafic utilisés

Plusieurs modèles de trafic routier sont utilisés pour relier entre elles les variables de trafic intervenant dans la méthode de fusion de données pour le filtrage. Le comportement du trafic étant différent en zone urbaine et périurbaine, les modèles sont détaillés pour les trois cas suivants :

- les modèles de trafic pour les chaînons en zone urbaine
- les modèles de trafic pour les chaînons en zone périurbaine
- les modèles de trafic adaptés aux trajets

#### **Modèles de trafic pour les chaînons situés en zone urbaine**

La Figure 10 exprime les relations entre les variables  $Q$ ,  $X$ ,  $TO$  et  $TP$  dans le cas urbain. A partir du débit sur un chaînon et de la valeur de la file d'attente en début de période, ces relations permettent de calculer le temps de parcours sur le chaînon, le taux d'occupation et la valeur de la file d'attente en fin de période. Les méthodes appliquées ainsi que leur évaluation figurent en annexe 8.5.

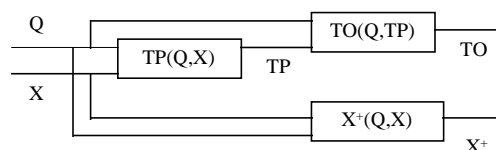


Figure 10 : Relations entre variables pour un chaînon urbain

#### **Modèles de trafic pour les chaînons situés en zone périurbaine**

La Figure 11 exprime les relations entre les variables  $Q$ ,  $TO$ ,  $TP$  et  $V$  dans le cas périurbain. A partir du débit et du taux d'occupation sur un chaînon, ces relations permettent de calculer le temps de parcours sur le chaînon et la vitesse moyenne. Le modèle est ici statique, il n'y a pas d'équations d'état, le temps de parcours à une période ne dépend pas des périodes précédentes.

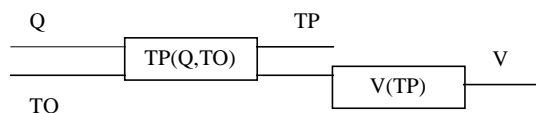


Figure 11 : Relations entre variables pour un chaînon périurbain

Le modèle BRP (eq. 1) reconstitue la vitesse moyenne à partir des données  $Q$  et  $TO$  mesurées. Il fournit de bons résultats sur voies rapides.

$$eq. 1 \quad TP(Q,TO) = TPI ( 1 - TO ) + ( L . f / dp ) . ( TO / Q )$$

Le modèle de vitesse utilisé pour les chaîons périurbains est directement donné par l'eq. 2.

$$eq. 2 \quad V(TP) = L / TP$$

### Modèles de trafic adaptés aux trajets

Un des besoins explicité par les exploitants des réseaux est d'obtenir des indicateurs sur les états de la circulation au niveau des trajets (couche II du réseau géographique) qui composent le réseau. On distingue deux indicateurs :

- le temps de parcours,
- le niveau de circulation ou niveau de service qui est utilisé par l'exploitant pour qualifier des conditions de circulation.

Ces indicateurs sont déterminés à partir d'une ou plusieurs variables de trafic en prenant comme référence une situation nominale. Les méthodes de détermination sont données en annexe (8.5.2).

#### 5.2.2.3 Méthode de résolution proposée

La méthode de filtrage proposée est décrite en annexe 8.5

#### 5.2.2.4 Données de sortie

En sortie de la fonction de fusion des données pour le filtrage, on obtient pour les chaîons sur lesquels on dispose de mesures corrigées, pour chaque période de l'horizon de filtrage :

- **cas 1 (chaîons urbains avec postes de mesure)**, le débit, le taux d'occupation et le temps de parcours,
- **cas 2 (chaîons périurbains avec postes de mesure)**, le débit, le taux d'occupation, la vitesse et le temps de parcours,
- **cas 3 (chaîons avec données GSM)** seulement, le temps de parcours,
- **cas 4 (chaîons sans données)**, le débit estimé et temps de parcours.
- **Pour les cas 1, 2 et 3**, le nombre de véhicules en queue en fin d'horizon de filtrage est également disponible sur les chaîons du réseau.

### 5.2.3 La prédiction des états du réseau

L'approche décrite dans cette section, porte sur la prédiction des séries temporelles que sont les variables macroscopiques fournis par les postes de mesures, les temps de parcours par chaînons et par trajets. Le modèle prévisionnel (Figure 12) devra être capable d'intégrer des données variées telles que les mesures des PM ( $Q, TO, V$ ) et/ou des TP sur le chaînon considéré ainsi que celles des chaînons adjacents, les données historiques afin de fournir un modèle fiable et robuste. Aussi, la prédiction des variables macroscopiques s'effectuera sur plusieurs pas d'échantillonnage avec un horizon  $H_p$  borné et multiple du pas d'échantillonnage. La fonction de prédiction doit également s'appliquer au cas des chaînons urbains et à celui des chaînons en périurbain. Enfin, le module prévisionnel est intégré dans un système complet de traitement des données du trafic et est activé après la fonction de filtrage et d'estimation par fusion de données.

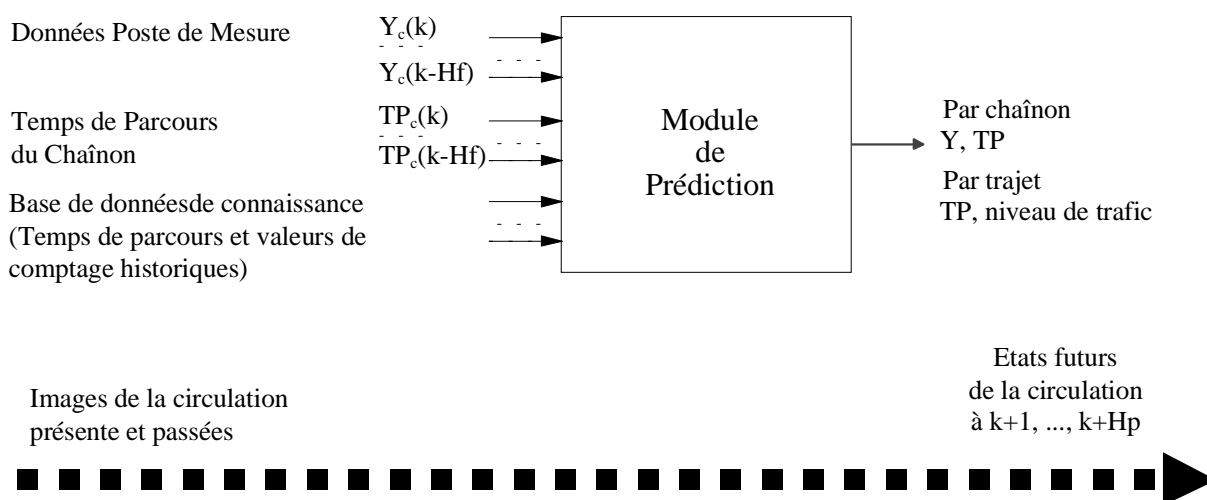


Figure 12 : Schéma de principe de la fonction de prédiction

#### 5.2.3.1 Données d'entrée

Les données d'entrée de la fonction de prédiction des temps de parcours à court terme sont les sorties de la fonction de fusion de données pour le filtrage, variables calculées sur l'horizon de filtrage  $H_f$ . Par prédiction à court terme, on entend prédiction sur au plus l'heure qui suit l'application de la fonction. Un événement inconnu survenant sur cet horizon de prédiction ne peut évidemment pas être prévu. Par contre, l'effet d'un événement survenu sur l'horizon de filtrage doit apparaître sur l'horizon de prédiction.

Pour prédire le temps de parcours, deux solutions sont possibles :

- prédiction directe à partir des temps de parcours filtrés,
- prédiction d'abord des variables de débits et de queues, puis utilisation du modèle trafic pour calculer les temps de parcours.

Le choix entre ces deux solutions dépend de la précision connue sur les types de données :

- pour un chaînon urbain, on dispose de données postes de mesure seulement. Le temps de parcours sur l'horizon de filtrage a été calculé à partir des débits et taux d'occupation filtrés. Il est plus judicieux d'effectuer la prédiction sur les débits et sur les queues que sur les temps de parcours.



- pour un chaînon quelconque, on dispose de données postes de mesure et GSM. Le choix entre la prédiction sur les données postes de mesure et sur les données GSM dépend de la précision existant sur les deux types de données.
- pour un chaînon quelconque, on dispose de données temps de parcours GSM. La prédiction va alors directement travailler sur les temps de parcours filtrés.

Le Tableau 4 résume les variables sur lesquelles sont réalisées le filtre de Kalman en fonction de la disponibilité des différents types de mesure.

Tableau 4 - Variables utilisés dans les modules de prédiction

Mesures disponibles sur l'horizon de filtrage	Urbain	Périurbain
Poste de mesure	Prédiction des $Q$ et des $X$ puis application du modèle $TP(Q,X)$	Prédiction des $Q$ et des $TO$ puis application du modèle $TP(Q,TO)$
Poste de mesure + Temps de parcours GSM	- soit prédiction des $Q$ et des $X$ puis application du modèle $TP(Q,X)$ - soit prédiction des $TP$	- soit prédiction des $Q$ et des $TO$ puis application du modèle $TP(Q,TO)$ - soit prédiction des $TP$
Temps de parcours GSM	Prédiction des $TP$	Prédiction des $TP$

### Les données historiques

L'importance des données historiques dans le domaine du trafic est bien connue. En effet, la prédiction des états de trafic le lundi à l'instant 8h pour les instants futurs 8h5, 8h10, ..., 8h30 dépendra de ce qui s'est passé entre 7h30 et 8h et du trafic habituel des lundis entre 8h à 8h30 (hypothèse d'horizon de filtrage et de prédiction de 30 minutes).

Pour tenir compte des phénomènes réguliers du trafic, la méthode de prédiction envisagée combine des données filtrées issues du proche passé et des données historiques. Si on dispose d'une base de données suffisamment conséquente, la base de données historiques pour chaque chaînon doit être distribuée par jour de la semaine et par période d'échantillonnage. Le poids entre les données historiques et les données filtrées est calibré hors ligne sur un jeu de données. Ce poids pourrait être ajusté dynamiquement dans la méthode en fonction de l'écart entre les données réellement mesurées et les données historiques.

Que l'on dispose ou non de données historiques lors de la mise en œuvre de la méthode de prédiction, la base de données doit être construite et enrichie au fur et à mesure de l'arrivée de nouvelles mesures. Le poids des données historiques dans la méthode doit alors être adapté en fonction de la taille de la base de données. Ce sera notamment le cas pour les données temps de parcours fournies par l'opérateur GSM.

### 5.2.3.2 Méthode de résolution

Deux approches de prévision de séries temporelles sont proposées pour l'implémentation, il s'agit :

- d'une méthode conventionnelle combinant un filtre de Kalman et des données historiques,
- d'une méthode non conventionnelle basée sur la théorie des réseaux des neurones.

Ces deux méthodes sont exposées en annexe (8.6)

### 5.2.3.3 Données de sortie

En sortie de la fonction de prédiction, on obtient à chaque période  $k$  et pour les instants futurs  $k+1$  à  $k+Hp$  les résultats suivants :

- **pour les chaînons** :  $Q_p$ ,  $TO_p$ ,  $V_p$ ,  $X_p$  et  $TP_p$
- **pour les trajets** :  $TP_p$  et les niveaux de services prédits

## 6 Mise en œuvre

Cette partie a pour but de spécifier les étapes nécessaires pour la mise en œuvre informatique du système de traitement de données ainsi que son déploiement in situ. Pour cela, les aspects suivants sont abordés :

- la description du contenu de la base de données,
- les prérequis en matière de calibrage des modèles de trafic et méthodes numériques,
- les méthodes de validation des données de sortie générées par les fonctions de qualification des données, de filtrage et d'estimation ainsi que la méthode de prédiction,
- la configuration matérielle nécessaire pour l'implémentation numérique,

### 6.1 Description et constitution des bases de donnée

Dans cette partie, nous précisons les données utiles pour la modélisation du réseau géographique, celles pour la constitution de la base de données historiques ainsi que les données de sortie générées par le système de traitement des données.

#### 6.1.1 Les données d'entrée

Comme il a été vu précédemment, on distingue trois bases de données en entrée du système : la base géographique, la base historique et la base dynamique. On peut également classer les données en fonction de leur période de rafraîchissement ; les données sont considérées comme statiques lorsque la période de rafraîchissement est de plusieurs jours, semaines ou mois et elles sont dites dynamiques si la période de rafraîchissement est de quelques minutes.

##### 6.1.1.1 La base de données géographiques

La modélisation des différentes couches est une tâche importante et souvent fastidieuse. Dans cette partie, on décrit brièvement la façon dont les données d'entrée peuvent être obtenues.

- La **couche I**, description du système de mesure, est constituée à partir des informations fournies par les gestionnaires des réseaux routiers. En d'autres termes, la liste et la position des postes de mesures sont connues des exploitants et devront être fournies sous forme numérique.

Les tables suivantes décrivent les éléments devant être saisis pour décrire chaque poste de mesure et carrefour à feux du réseau.

- La **couche II** décrit les chaînons, les nœuds et les interdictions de tourner.

Autant que possible, nous nous appuyons sur des bases de données numériques existantes (par ex. Michelin, TéléAtlas) ou des bases de données existantes au niveau des exploitants des réseaux de l'agglomération de Toulouse pour obtenir les informations sur les longueurs, les sens de circulation, la largeur des voies et les noms des chaînons. Les nœuds sont également extraits de ces bases de données.

Cette couche est enrichie par des informations propres à l'exploitation du trafic (vitesse, capacité des chaînons, etc.).

Les caractéristiques d'une voirie comprennent le nom de la rue, le sens de circulation et l'importance de la rue (Axe principal, secondaire, etc.). Ces caractéristiques sont communes à plusieurs chaînons. Pour éviter de dupliquer l'information, la liste des voiries est enregistrée dans un fichier différent de celui décrivant l'ensemble des chaînons. Le champ *voirie* défini pour chaque chaînon permet de faire le lien avec la table des voiries.

Une table des interdictions de tourner au nœud est facultative et peut-être décrite pour préciser les mouvements autorisés dans les carrefours.

- La **couche III** qui contient les trajets et les pôles, est produite à partir de modélisations existantes ou est complètement construite.

#### 6.1.1.2 La base de données historiques

Ces données représentent des jours standards de semaines types de la circulation. En effet, le trafic présente un caractère journalier car il est souvent reproductible d'un jour à l'autre. Les données historiques servent à pallier le manque d'informations (pannes, erreurs de mesures) pour assurer une continuité dans les traitements de la chaîne de traitement des données (qualification des données et prédiction des états du trafic).

De façon analogue, une base de données historique est constituée pour un ensemble  $T$  de trajet.

#### 6.1.1.3 La base de données dynamiques

On distingue différentes données dynamiques :

- les données de comptage pour un ensemble  $M$  de PM. La période d'échantillonnage des mesures sur la ville de Toulouse (urbain et périurbain) est égale à 3 minutes.
- les données de temps de parcours de véhicules traceurs pour un ensemble  $C$  de chaînons.

Ces données sont enregistrées dans deux fichiers distincts. Chaque fichier est réécrit à chaque nouvelle période.

### 6.1.2 Les données de sortie

Les résultats du système de traitement des données sont générés à chaque tranche de temps :

- dans des fichiers circulaires pour l'échange des données entre les différents modules ou pour accéder à des résultats d'un traitement particulier. Ils contiennent les résultats des traitements effectués lors d'une seule tranche de temps.
- dans des fichiers de sauvegarde qui enregistrent les principaux résultats. Ces fichiers contiennent les résultats de chaque tranche de temps (classification chronologique) ou des résultats fournis par le système de traitement des données. Ils sont utilisés pour une analyse a posteriori des conditions de circulation.

Les données de sortie sont enregistrées dans un ensemble de fichiers distincts. Ces données sont relatives aux données dynamiques corrigées pour l'ensemble des PM, aux estimations (temps de parcours, longueurs des files d'attente, etc.) et aux prédictions.

### 6.1.2.1 Les fichiers circulaires

Les fichiers de données pour l'instant  $tx$  sont :

- $pmc_{tx}.txt$  pour les mesures de comptage corrigées par le module de qualification des données (voir Tableau 28),
- $chainon_{tx}.txt$  qui fournit les résultats du module de filtrage et d'estimation relatifs aux chaînons de la couche II (voir Tableau 29),
- $trajet_{tx}.txt$  qui fournit les résultats du module de filtrage et d'estimation relatifs aux trajets de la couche III (voir Tableau 30),
- $chainon_{tx}.tkp$  et  $trajet_{tx}.tkp$  pour les résultats du module de prédiction.  $Kp$  correspond à l'instant de prédiction. Par conséquent, à chaque tranche de temps un ensemble de  $Kp$  fichiers est généré à la fois pour les chaînons et les trajets.  $Kp$  est égal à l'horizon de prédiction  $H_p$  modulo la période d'échantillonnage  $T$ . Par exemple, si  $H_p = 36$  minutes et  $T = 6$  minutes, le module de prédiction génère à l'instant  $tx$  6 fichiers de résultats pour les chaînons plus 6 autres pour les trajets. Le format des fichiers de résultats est identique à ceux du Tableau 29 et du Tableau 30.

### 6.1.2.2 Les fichiers de sauvegardes

Comme précisé précédemment, les principaux résultats des traitements sont archivés dans des fichiers par jour. Chaque fichier possède la même structure qui est définie dans le Tableau 31 et le Tableau 32 (annexe 8.7.4). Dans ces tableaux, la période d'échantillonnage est de 15 minutes. Les fichiers contiennent une ligne d'en-tête et des lignes de données. L'intérêt du premier tableau est d'enregistrer dans le fichier les résultats de chaque traitement ligne après ligne. Le second tableau permet l'enregistrement des données pour des listes importantes d'éléments. Le premier type de sauvegarde est plus simple et moins «gourmand» en termes de données à conserver en mémoire vive mais pose le problème de lecture des fichiers par des tableaux usuels (MS Excel par exemple) lorsque la liste des objets est supérieure à 256 objets.

Le tableau suivant spécifie les noms des fichiers de sauvegarde relatifs aux données mesurées, corrigées et à celles prédites. Il est à noter que les résultats sont enregistrés pour chaque instant de mesure ou de traitement. Dans le cas où le nombre d'éléments (PM, Trajet) est important, il est peut être intéressant, pour limiter la taille des fichiers, de sauvegarder les résultats agrégés par tranche de temps (toutes les 10, 15, ... , 30 minutes).

Tableau 5 : Structure des noms de fichier de sauvegarde

Objet	Données	Mesuré à k	Corrigé à k	Prédit à k+1 pas	...	Prédit à k+H <sub>p</sub> pas
Poste de Mesure	Débit de véhicules	q_jjmmaaaa.txt	qc_jjmmaaaa.txt	q_jjmmaaaa.k1		q_jjmmaaaa.khp
	Taux d'occupation	t_jjmmaaaa.txt	tc_jjmmaaaa.txt	t_jjmmaaaa.k1		t_jjmmaaaa.khp
	Vitesse	v_jjmmaaaa.txt	vc_jjmmaaaa.txt	v_jjmmaaaa.k1		v_jjmmaaaa.khp
	Etat	etatpmjjmmaaaa.txt				
Trajet	Temps de parcours		tpjjmmaaaa.txt	tpjjmmaaaa.k1		tpjjmmaaaa.khp
	Niveau de services		etatrajetjjmmaa.txt	etatrajetjjmmaa.k1		etatrajetjjmmaa.khp

## 6.2 Calibrage

Cette partie précise les étapes à effectuer pour ajuster les paramètres des différents modèles et méthodes intégrés dans le système de traitement des données. En effet, les modèles et méthodes de résolution proposés nécessitent un calibrage sur site de certains paramètres.

Fonction	Éléments à paramétrer	Méthode
<b>Qualification des mesures (données de comptage)</b>	<p>Nombre d'instants passés (<math>H_f</math>) à utiliser pour corriger les données</p> <p>Méthode de correction :</p> <ul style="list-style-type: none"> <li>• seuil de convergence (<math>\mathcal{E}_{min}</math>)</li> <li>• nombre maximal d'itération</li> <li>• intervalle de confiance des valeurs mesurées en % pour les tests 1, 2 et 3</li> </ul> <p>Pour chaque PM :</p> <ul style="list-style-type: none"> <li>• <math>Q_{max}</math>, <math>V_{max}</math>, <math>T_{omax}</math></li> <li>• modèle théorique <math>V=f(TO)</math> et ses paramètres</li> <li>• valeurs historiques des PM</li> </ul>	<p>Certaines informations (<math>Q_{max}</math>, <math>V_{max}</math>, <math>T_{omax}</math>) peuvent être fournies par les exploitants des réseaux.</p> <p>Cependant, la procédure de calibrage revient à effectuer une analyse statistique en termes de trafic et des données de comptage archivées.</p> <p>En pratique, la détermination du modèle théorique d'un PM ne peut être obtenue que par une procédure semi-automatique.</p>
<b>Qualification des mesures (temps de parcours des véhicules traceurs)</b>	<p>Vitesses maximales des véhicules en fonctions des types de voirie</p> <p>Erreur de localisation du téléphone en mètres</p>	<p>Utilisation des vitesses libres ou des informations fournies par les exploitants des réseaux.</p> <p>Caractéristiques des systèmes de localisation des téléphones mobiles</p>
<b>Filtrage et Estimation</b>	<p>L'horizon de filtrage <math>H_f</math></p> <p>Les coefficients <math>\alpha_i</math> de la méthode de filtrage</p>	<p><math>H_f</math> peut être calibré hors ligne en faisant fonctionner la méthode de filtrage sur des données réelles. L'horizon de prédiction est donné par l'exploitant.</p> <p>Les coefficients <math>\alpha_i</math> sont inversement égaux aux carrés de la précision donnée pour chaque type de mesure. Par défaut, on peut utiliser les valeurs suivantes pour la précision : (<math>\alpha_1</math> : débit urbain, <math>\alpha_2</math> : débit périurbain, <math>\alpha_3</math> : taux d'occupation urbain, <math>\alpha_4</math> : taux d'occupation périurbain, <math>\alpha_5</math> : vitesse périurbaine)</p>
<b>Prédiction</b>	<p>Horizon de filtrage <math>H_f</math></p> <p>Horizon de prédiction <math>H_p</math></p> <p>Liste des variables à prédire</p> <p>Méthode de prédiction utilisée</p>	<p>Analyse statistique des séries temporelles observées (comptage et mesure des temps de parcours en fonction du temps). Cette analyse permet de déterminer le nombre de</p>

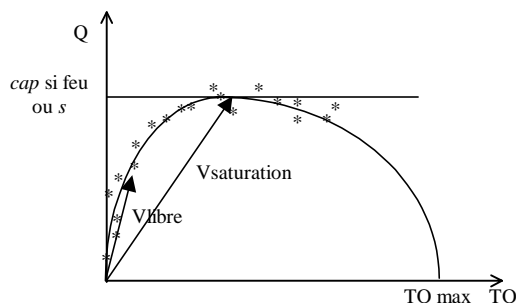
	<ul style="list-style-type: none"> <li>• filtre de Kalman</li> <li>• réseau de neurones</li> </ul> <p>Paramètres des méthodes :</p> <ul style="list-style-type: none"> <li>• paramètres initiaux</li> <li>• seuil d'erreur de prédiction</li> </ul> <p>Valeurs historiques :</p> <ul style="list-style-type: none"> <li>• des comptages des PM et des temps de parcours</li> <li>• poids des valeurs historiques (% fonction du pas de prédiction)</li> </ul>	<p>mesures du proche passé à utiliser</p> <p>Utilisation d'au moins 15 jours d'observations pour calibrer les modèles de prédiction.</p> <p>Le coefficient <math>\beta</math> de la méthode de prédiction pondère les données historiques. Il doit être calibré hors ligne sur des données réelles pour chaque période de prédiction possible. Une valeur initiale de 30% de données historiques (<math>\beta = 70%</math>) peut être considérée dans l'algorithme.</p>
--	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

### 6.2.1 Calibrage des variables clés

Un ensemble de paramètres caractérisant la circulation sur les différents chaînons doit être ajusté en fonction de la configuration du terrain. Il s'agit de :

- **la capacité ou débit à la saturation**

La capacité en présence de feux est calibrée par chaînon en fonction du comportement moyen observé sur les données mesurées de débit et de taux d'occupation. Il en est de même pour le débit de saturation dans le cas d'un chaînon sans feu.



- **la vitesse libre et le temps de parcours**

La vitesse libre peut être déterminée à partir des types de voiries (voie rapide, boulevard, avenue, rue etc.). Toutefois, la vitesse libre peut être ajustée pour les chaînons à partir de mesures empiriques de la vitesse par des stations de comptage de véhicules.

Le temps de parcours à vitesse libre pour un chaînon est calculé à partir de la longueur du chaînon et de la vitesse maximale autorisée (en urbain,  $TPI = L / 50km/h$  et en périurbain rocade  $TPI = L / 120km/h$ , sauf indication contraire sur la voirie).

- **la vitesse à la saturation**

La vitesse à la saturation indique la vitesse limite à partir de laquelle le trafic pour un chaînon particulier bascule d'un état fluide à un état saturé.

- **la densité maximale**

La densité maximale (km/véh.) peut être déterminée à partir de la longueur moyenne d'un véhicule, de l'ordre de 4,5 m en urbain et 5,5 m en périurbain et sur certaines voies rapides urbaines (circulation de poids lourds). Ces informations peuvent être obtenues à partir de certains capteurs qui indiquent la proportion de véhicules légers et de poids lourds.

Le caractère journalier du trafic a été introduit dans le processus de traitement en ajoutant une composante historique pour pallier le manque des données et pour la prédiction des états de la circulation. Une base de données historiques doit donc être construite.

### 6.2.2 Constitution de la base de données historiques

Cette base de données est construite à partir d'un algorithme de classification qui génère un ensemble d'historiques  $Y_{i,h}$  où  $i$  est un PM,  $h \in [1, H]$  avec  $H$  le nombre maximal d'historiques.  $Y$  est en fait une matrice à deux dimensions dont la première dimension donne tous les éléments de la série temporelle représentant le profil journalier de 0h jusqu'à 24h. La seconde dimension donne la liste des variables classées. Les variables sont le débit de véhicules  $Q$ , le taux d'occupation  $TO$  et parfois la vitesse  $V$  si le PM permet de mesurer cette dernière.

En pratique, cette base de données contient un profil pour chaque jour d'une semaine type ( $H=7$ ) et pour chaque PM. Elle est obtenue avec, au minimum, 2 ou 3 mois d'observations. En règle générale, plus le nombre de jours d'observations est grand et plus le résultat de la classification sera indépendant de tous phénomènes ponctuels et mouvements saisonniers.

De la même manière, une base de données des temps de parcours peut être constituée au fur et à mesure de la mise en exploitation du système à partir des informations fournies par les véhicules traceurs.

## 6.3 Validation des modèles

La Figure 13 présente la méthode permettant de valider d'une part la fonction de filtrage et d'estimation et d'autre part la fonction de prédiction. Deux jeux de données sont nécessaires pour garantir l'indépendance des résultats.

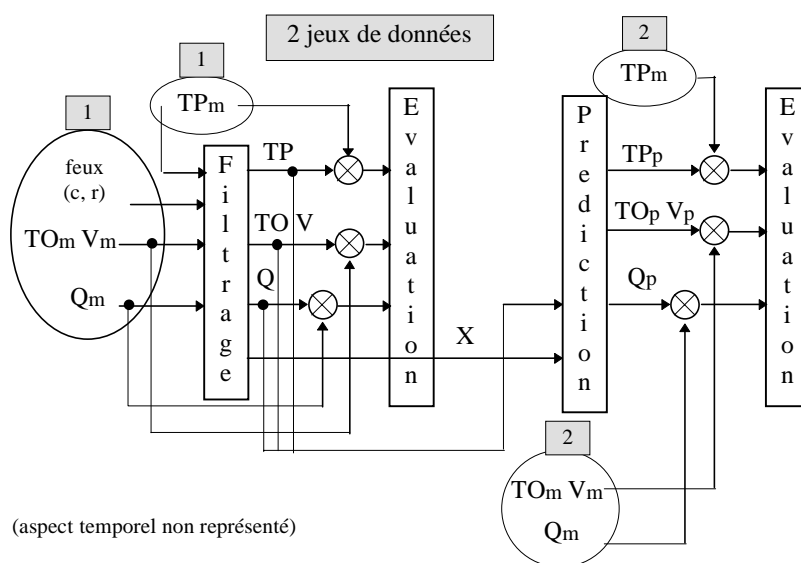


Figure 13 : Méthode de validation de la fonction de filtrage, d'estimation et de prédiction



On peut ainsi calculer pour chaque type de variables  $V$  dont on possède un nombre suffisant de mesures  $n$  l'erreur moyenne absolue relative (ou *MAPE* en anglais pour Mean Absolute Percentage Error).

$$MAPE(nV) = \frac{1}{n} \sum \frac{|V_m - V_p|}{V_m}$$

On peut ainsi noter qu'une validation des temps de parcours prédits nécessite un recueil de données suffisant sur le terrain, soit par l'intermédiaire de véhicules flottants, soit par une enquête minéralogique. A partir de  $n=100$ , les résultats statistiques obtenus sont valides. Pour  $n=50$ , on a une bonne idée du comportement de la fonction.

## 6.4 Implantation informatique

Cette partie spécifie les moyens informatiques nécessaires pour l'implémentation du système de traitement de données. Les informations indiquées ici doivent servir à la configuration le matériel pour une installation sur site.

### 6.4.1 La configuration logicielle

Le système de traitement des données doit fonctionner dans un environnement Windows™ NT 4.0 Workstation ou Server. Ce système intègre les composants suivants :

- une application (exécutable) qui assure le rôle de superviseur pour synchroniser tous les modules. Ce superviseur intègre un serveur de communication pour assurer le dialogue avec les différents modules.

Le superviseur comprend également une interface Homme/Machine pour le contrôle et le suivi du fonctionnement du système. Des fonctionnalités avancées doivent être intégrées comme par exemple :

- gestion des processus (initialisation et arrêt des modules),
- gestion de la communication,
- fonctions pour la manipulation (sélection, zoom avant et arrière, etc.) des différentes couches et éléments du réseau géographique. Pour cela, il est préconisé d'utiliser des composants de Système d'Information Géographique.
- outils de visualisation des résultats (évolution temporelle et tableaux de résultats des variables),
- impression
- des modules de traitement (qualification des données, filtrage/estimation et prédiction) qui sont développés sous la forme de processus indépendants (exécutable ou librairie). Ce qui signifie que les modules doivent pouvoir être répartis sur différentes machines si la consommation de la mémoire ou du processeur est critique. Pour cela, deux modes de communication sont retenus :
  - un mode de communication local dans le cas où le superviseur et une application sont sur le même ordinateur. Le principe est d'intégrer dans une seule application (exécutable) le Superviseur et l'application et de lancer chaque module comme une fonction.
  - un mode de communication distant si le superviseur et une application sont sur des ordinateurs différents. Dans ce cas, le mode de communication retenu est l'utilisation des socket sur réseau TCP/IP.

- un système de gestion des données (lecture et écriture) qui est utilisé pour manipuler les différentes bases de données.

La gestion des données peut s'effectuer à partir de fichiers externes (format ASCII/TEXTE). Cette solution permet de consulter les résultats à partir de traitements de texte et tableurs standards. Une autre possibilité est d'utiliser un SGBD (Système de Gestion de Base de Données) standard (MS ACCESS, ORACLE, etc.). Dans ce cas, les modules doivent intégrer des requêtes d'accès au SGBD en SQL. L'intérêt de cette solution est de faciliter la manipulation des données, de simplifier l'archivage et de rendre les manipulations plus rapides.

#### 6.4.2 La configuration matérielle

La période d'échantillonnage des mesures sur la ville de Toulouse (urbain et périurbain) est égale à 3 minutes. Pour être le plus dynamique possible, le système de traitement des données doit donc réaliser la qualification des mesures, le filtrage, l'estimation et la prédiction dans la tranche de temps imparti.

Pour cela, les caractéristiques minimales pour faire fonctionner le système de traitement de données sont de disposer d'un ordinateur PC pentium III 500 Mhz avec 256 Moctets de mémoire vive, 5 Goctets de disque dur SCSI, d'une carte réseau TCP/IP - Fast Ethernet, d'un lecteur de CD ROM et d'un écran 17 pouce ou plus.

Toutefois, il peut être envisager de répartir les modules sur plusieurs ordinateurs ou d'utiliser des ordinateurs à plusieurs processeurs.

#### 6.4.3 Eléments de performances

Cette partie donne des éléments de performance des différents modules intégrés dans le système de traitement de données. Le tableau ci-dessous donne des résultats d'évaluation du module de qualification des données et de prédiction par réseau de neurones.

<i>Critère</i>	<i>Module</i>	<i>1250 chaînes et 125 PM</i>	<i>160 chaînes et 75 PM</i>
<i>Temps de calcul</i>	Qualification des données	23.5	68.4
	Autres processus	19.1	31.6
Rapport du temps d'utilisation du processeur par un module sur le temps d'utilisation total du système (%)		57.4	Dont 20.4% sont considérés pour la prédiction et 37 pour de nouveaux calibrages
<i>Mémoire requise (Moctets)</i>	Qualification des données	12.45	3.68
	Filtrage	6.3	2.87
	Autres processus	4.1	1.49
	Prédiction réseaux de neurones	10.6	

La méthode de filtrage mono-variable avec réception de deux types de données de mesure et minimisation quadratique sur ces données, puis prédiction des temps de parcours sur un horizon de deux périodes a fonctionné sur une station de travail SUN Spark 5 et sur un

PC Pentium 200. Pour 15 chaînes, le temps de calcul et d'affichage sur PC pour une période d'échantillonnage est inférieur à 10 secondes. Sur une station SUN, le module de filtrage sur 8 périodes passées avec une prédiction sur 2 périodes donne des temps inférieurs à la seconde. Ces résultats montrent que le temps de calcul n'est pas une contrainte dans les méthodes proposées ici.

## 7 Conclusions

*Cette étude a été réalisée dans le cadre du Programme national de Recherche et d'Innovation dans les Transports Terrestres - PREDIT / 1996 - 2000 - dans le thème Route Intelligente avec pour sujet l'amélioration des traitements de données de trafic*

*Elle s'est intéressée particulièrement à la mise en œuvre d'un outil de traitement de données de trafic en vue de la mise en place du système global d'information sur les déplacements dans l'agglomération toulousaine.*

*La phase d'analyse des besoins, qui s'est appuyée sur des études récentes et des entretiens avec les exploitants, s'est attachée à recenser les fonctions d'exploitation actuellement opérationnelles, les données dont disposent les exploitants et celles qu'ils pourraient exploiter, les traitements qui sont effectués sur ces données et enfin l'information diffusée à ce jour aux usagers et les moyens utilisés.*

*Cette analyse des besoins fait apparaître la nécessité pour les exploitants de disposer d'outils pour le traitement des données. Les systèmes informatiques et de communications permettent actuellement aux exploitants de recueillir sur le terrain une quantité importante de données et de les centraliser. Mais les systèmes qui utilisent ces données sont essentiellement dédiés à la sécurité et à l'exploitation de trafic et pas à l'information des usagers.*

*Les fonctions nouvelles attendues par les exploitants concernent tout d'abord l'assurance de la continuité de service d'exploitation et d'informations et la possibilité de pallier les pannes des systèmes de mesures. Il s'agit également de disposer d'outils pour quantifier les états de circulation en temps réel et pour prévoir l'évolution de trafic à court terme dans l'objectif d'informer les usagers sur la situation à venir.*

*La phase de spécifications du système de traitement de données a consisté à proposer les fonctions, les méthodes et l'architecture qui répondent aux besoins exprimés. Le système proposé permet d'une part d'utiliser les données de comptages provenant des systèmes de mesures installés et d'autre part d'utiliser les données issues de véhicules traceurs. Il convient de noter que ces données concernent à la fois les réseaux urbain et périurbain.*

*Les fonctions principales de traitement des données sont au nombre de trois :*

- *qualification et filtrage des mesures,*
- *estimation d'indicateurs,*
- *prédiction des niveaux de circulation.*

*Chaque fonction a donné lieu à l'étude des données d'entrées, des modèles de trafic et des méthodes de résolution utilisés et des données de sortie.*

*Il est important de noter que les fonctions et les méthodes proposées ont été étudiées en vue d'une implémentation réelle. Ce souci d'aboutir à une solution réaliste a conduit pour les fonctions de filtrage et de prédiction à envisager des cas concrets comme les tronçons avec mesures de débit, de taux d'occupation et/ou de vitesse, ceux possédant en plus des mesures issues de véhicules traceurs et ceux sans données mesurées.*

*Une mise en œuvre du système a été proposée dans le but de fournir aux exploitants des éléments techniques concernant le déploiement sur site. Le contenu des bases de données, le calibrage des modèles, les moyens de calibrer les méthodes, la configuration matérielle et l'interface homme machine du système ont été exposés. Ce rapport fournit*

*aux exploitants un point de vue théorique sur les méthodes de résolution et les modèles de trafic et un point de vue pratique sur les éléments pour rendre le système de traitement de données opérationnel.*

*Ce travail a ainsi permis d'étudier un système de traitement de données de trafic dans le but de fournir des indicateurs de temps de parcours et de niveaux de circulation en temps réel adaptés au cas toulousain. Cette étude ouvre alors la possibilité d'une réalisation pratique sur Toulouse. Toutefois, la démarche proposée (depuis l'analyse des besoins, jusqu'à la mise en œuvre) est générique et peut s'appliquer à d'autres agglomérations françaises.*

*Les systèmes de recueil par postes de mesures actuels ne suffisent pas seuls à élaborer de façon satisfaisante ces indicateurs. C'est pourquoi le travail s'est orienté vers d'autres sources de mesures comme les données issues des véhicules traceurs, notamment par les moyens de téléphonie mobile, qui permettent une extension de la couverture géographique (extension spatiale des mesures). L'un des intérêts du système proposé réside alors dans la possibilité de fusion des données qu'il offre.*

*Enfin, cette étude ouvre la possibilité d'explorer des nouvelles voies dans l'utilisation des données autres que celles provenant des postes de mesure. Les systèmes de communication modernes comme la téléphonie mobile et les systèmes de gestion des flottes de véhicules ne manqueront pas de proposer de nouveaux moyens d'obtenir des informations sur le trafic.*

## 8 Annexes

### 8.1 Questionnaire

#### Questionnaire

Les réponses qui seront fournies dans ce questionnaire devront refléter le fonctionnement global de votre organisme. Ce questionnaire est plus un guide de réflexion, et vous pouvez y adjoindre tous les documents nécessaires à la bonne compréhension des réponses : n'hésitez pas le compléter dans des feuilles à part.

Les questions qui suivent sont volontairement ouvertes pour élargir le champ des réflexions. Une fois ce questionnaire complété, nous reviendrons vers vous avec des questions directes plus précises.

#### Identification de l'organisme

Organisme : .....  
Nom, Prénom : .....  
Téléphone : .....  
Fax : .....  
Email : .....

#### Fonctions d'exploitation exercées

##### Supervision : acquisition et mise en forme

Description de la fonction : .....  
.....

##### Gestion des événements

Description de la fonction : .....  
.....

##### Gestion de trafic

Description de la fonction : .....  
.....

##### Information vers les usagers

Description de la fonction : .....  
.....

##### Partage d'information avec les autres exploitants

Description de la fonction : .....  
.....

## Usagers

### Composition du trafic, nombre d'usagers, etc.

.....  
.....

### Attentes identifiées a priori de ces usagers

En terme de gestion : .....

.....

En terme de sécurité : .....

.....

En terme d'information : .....

.....

## Données statiques disponibles

### Description du réseau de surface

.....  
.....

### Enquêtes de déplacements réalisés

Type : .....

Fréquence : .....

Données recueillies : .....

.....

.....

## Données dynamiques recueillies

### Moyens de mesures automatiques

Nombre : .....

Position : .....

Format de données : .....

Fréquence : .....

.....

.....

### Caméras, vidéo, etc.

Nombre : .....

Position : .....

Format : .....

Utilisation : .....

.....

**Contacts usagers : bornes d'appels et autre média**

Type : .....

Fréquence d'appel : .....

Traitement : .....

.....

**Echanges avec les autres opérateurs ou partenaires**

Type : .....

Nombre : .....

Position : .....

Format : .....

Utilisation ou traitement : .....

.....

**Données non utilisées**

**Données recueillies mais non utilisé**

Fournisseur potentiel : (ex. taxis) .....

Description des données : .....

Format de données : .....

Raison de la non utilisation : .....

.....

**Autres données (non disponibles)**

Type de donnée : (ex. localisation de véhicules d'urgences, etc.) .....

Utilisation souhaitée : .....

Mode de recueil de ces données : .....

Fréquence : .....

.....

**Traitements des données réalisés**

**Données statiques**

Outils de recueil, de stockage : .....

Utilisation : .....

Méthodes de traitement : .....

.....

**Données dynamiques**

Outil de recueil, de stockage : .....

Utilisation : .....

Méthodes de traitement : .....



.....  
**Indicateurs utilisés**

Indicateurs élaborés : .....

Méthodes de calculs : .....

Utilisation : .....

Formats : .....

**Messages et diffusion d'information**

(diffusion vers les usagers ou les opérateurs privés)

**Données brutes diffusées / diffusables**

Type : .....

Destinataires : .....

Format : .....

Fréquence : .....

.....  
**Messages élaborés diffusés**

Type : .....

Destinataires : .....

Format : .....

Fréquence : .....

.....  
**Accessibilité de l'information**

Pour les opérateurs de diffusion : .....

Pour les usagers VP : .....

Pour les Toulousains : .....

**Besoins identifiés par les opérateurs**

**Supervision : acquisition et mise en forme**

.....  
.....  
.....  
.....

**Gestion des événements**

.....  
.....  
.....

.....  
**Gestion de trafic**

.....  
.....  
.....  
.....

**Information vers les usagers**

.....  
.....  
.....  
.....

**Partage d'information avec les autres exploitants**

.....  
.....  
.....

**Commentaires**

## 8.2 Eléments de base de la théorie du trafic

L'écoulement de la circulation que ce soit en milieu urbain ou mixte (périphérique et urbain) est envisagé ici sous l'angle macroscopique. C'est-à-dire que les variables de trafic décrivent globalement des flots de véhicules sur une section de route. Les variables du trafic sont : le débit de véhicules  $Q$  représentant l'écoulement des véhicules dans le temps, la densité ou concentration  $C$  représentant la répartition de véhicules dans l'espace et la vitesse moyenne des véhicules  $V$ . Ces trois variables sont régies par une relation fondamentale telle qu'à un point d'abscisse  $x$  et pour un instant  $k$ , on a :

$$Q(x, k) = V(x, k) \cdot C(x, k) \tag{8.1}$$

On fait l'hypothèse que  $V$  et  $C$  sont liés par une fonction monotone et décroissante  $V=f(C)$ . Cette hypothèse est fondamentale. Un diagramme fondamental découle de la relation intervenant dans l'hypothèse fondamentale, il est représenté par les plans  $(V, Q)$ ,  $(V, C)$  et  $(Q, C)$ . De ce diagramme, quatre valeurs clés caractérisent l'écoulement du trafic sur un chaînon, ce sont la vitesse libre  $V_f$ , la vitesse à la capacité  $V_{sat}$ , le débit à la capacité  $Q_{sat}$  et la densité maximale  $C_{max}$ . La densité à la capacité  $C_{sat}$  est obtenue à partir de  $V_{sat}$  et  $Q_{sat}$  d'après la relation (8.1). La Figure 14 ci-dessous représente l'allure du diagramme fondamental pour une fonction  $V(C)$  non linéaire.

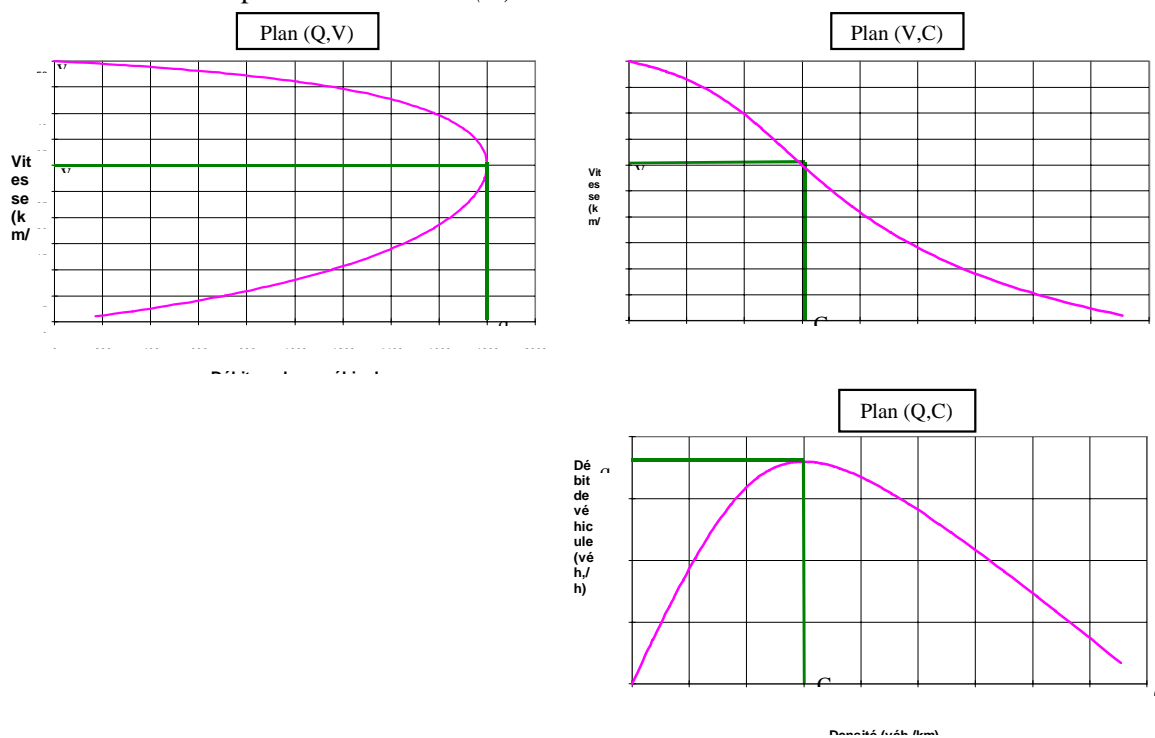


Figure 14 : Représentation du diagramme fondamental

Si l'on considère que ce qui se passe dans le temps en un point est représentatif de ce qui se passe autour ce point alors on peut lier la concentration  $C$  avec le taux d'occupation  $TO$  par la relation suivante ( $d$  et  $l$  étant respectivement la longueur d'une boucle et la longueur moyenne d'un véhicule) :

$$TO = (d + l) \cdot C \tag{8.2}$$

L'hypothèse d'une concentration  $C$  proportionnelle au taux d'occupation  $TO$  permet de s'intéresser à cette dernière variable et, lorsque c'est nécessaire, déduire la concentration

à partir de la formule 8.2. Les paramètres clés  $TO_{sat}$  et  $TO_{max}$  représentent respectivement le taux d'occupation à saturation et le taux d'occupation maximal.

De plus, en considérant les relations (8.1) et (8.2), on a  $Q/TO$  proportionnel à  $V$ . On s'intéresse à une variable enrichie  $V'$  définie par  $V'=Q/TO$ . La relation liant les vitesses  $V$  et  $V'$  est donnée ci-dessous.

$$Q = TO \cdot (d + l) \Rightarrow V' = Q/TO = V/(d + l)$$

## 8.3 Compléments sur la méthode de qualification des données de comptages

### 8.3.1 Tests sur les données aberrantes

Cette partie détaille les tests mises en œuvre pour détecter si une mesure de trafic est déclarée aberrante ou non.

#### Test 1 : Vérification temporelle

Ce test consiste à estimer les variables macroscopiques  $Y_{c1}(k, m)$  d'un PM  $m$  à partir des dernières mesures  $Y(k-1, m), Y(k-2, m), \dots, Y(k-n, m)$  où  $n$  est le nombre de mesures issues du proche passé et d'une valeur historique provenant d'un profil standard  $Y_h(k, m)$  de ce même PM. L'estimateur est donné par la relation ci-dessous où  $Y_{passé}(k, m)$  est une estimation de  $Y(k, m)$  à partir d'un processus autorégressif et  $\alpha$  un coefficient de lissage :

$$Y_{c1}(k, m) = Y_{passé}(k, m) + \alpha \cdot Y_h(k, m)$$

$$\text{avec } \begin{cases} Y_{passé}(k, m) = \sum_{i=1}^{Hh} a_i \cdot Y(k-i, m) + \varepsilon(h) \\ a_i = \text{coefficient de pondération} \\ \varepsilon(h) = \text{bruit blanc tel que } E[\varepsilon] = 0 \end{cases}$$

A cet estimateur, on associe un intervalle de confiance permettant d'assurer que la valeur réelle  $Y(k, m)$  se situe dans un encadrement à plus de 95% de la valeur estimée  $Y_{c1}(k, m)$ . Le test de validité dans l'espace temporel dans l'hypothèse d'une distribution normale, où  $SE_1$  est l'écart type de l'estimateur  $Y_{c1}(k, m)$  et  $t_{.025}$  la valeur du test du  $t$  à 2.5/100, est :

$$Y_{c1} - t_{.025} SE_1 \leq Y(k, m) \leq Y_{c1} + t_{.025} SE_1 \quad (\text{Test 1})$$

#### Test 2 : Vérification spatiale

Lorsqu'une variation temporelle est détectée, c'est-à-dire que le test 1 n'est pas vérifié, cette étape a pour but de faire la différence entre une erreur de mesure et un phénomène de trafic récurrent. Pour cela, on vérifie que les PM voisins présentent la même variation de trafic. La méthode consiste à calculer l'estimation de la donnée vraisemblable  $Y_{c2}(k, m)$  à partir du tableau  $Y_j(k, m)$  où  $k$  est l'instant courant et  $m$  le point de mesure, puis à calculer son résidu standard  $e_j(k, m)$  par la formule

$$Y_{c2}(k, m) = \frac{Y_j(k, \cdot) \cdot Y_j(\cdot, m)}{Y_j(\cdot, \cdot)}$$

$$e_j(k, m) = \frac{Y_j(k, m) - Y_{c2}(k, m)}{\sqrt{Y_{c2}(k, m)}}$$

où chaque valeur est considérée comme une variable aléatoire normalement distribuée de moyenne nulle dont la variance est estimée par

$$v_j(k, m) = \left(1 - \frac{Y_j(k, \cdot)}{Y_j(\cdot, \cdot)}\right) \cdot \left(1 - \frac{Y_j(\cdot, m)}{Y_j(\cdot, \cdot)}\right)$$

On peut ainsi déterminer si le résidu  $e(k,m)$  est dans l'intervalle de confiance :

$$-\lambda\sqrt{v_j(k,m)} \leq e_j(k,m) \leq +\lambda\sqrt{v_j(k,m)} \quad (\text{Test 2})$$

avec  $\lambda$  le seuil de confiance désiré (fixé empiriquement).

### Test 3 : Vérification autour de la courbe théorique

Les formes fonctionnelles qui décrivent le diagramme fondamental du trafic (voir annexe 8.2), a permettent de qualifier la précision des PM à partir des nuages de points  $Y(Q,TO,V)$  observés par rapport à une courbe théorique. Certaines mesures peuvent être à la fois admissibles dans l'espace et le temps mais physiquement inadmissibles lorsqu'elles sont représentées dans le plan  $(Q, TO)$   $(V, TO)$  et  $(Q, V)$ . La méthode consiste à mettre en œuvre un estimateur à partir d'un modèle théorique liant les variables  $Q$ ,  $TO$  et  $V$

$$Y_{c3}(k,m) = [Q_{c3}(k,m), TO_{c3}(k,m), V_{c3}(k,m)]^T$$

$$\text{avec } \begin{cases} Q_c(k,m) = TO_c(k,m) \cdot V_c(k,m) \\ V_c(k,m) = f(TO_c(k,m)) \\ \text{tel que } \min_{TO_c, V_c} [(TO - TO_c)^2 + (V - V_c)^2] \end{cases}$$

De la même façon que pour la vérification temporelle et spatiale, une bande de confiance est construite à 95% ( $SE_3$  écart type de l'estimateur et  $t_{.025}$  la valeur du test du t à 2.5/100) autour du point estimé pour s'assurer que la mesure réelle est acceptable dans les plans  $(V,Q)$ ,  $(V,TO)$  et  $(Q, TO)$ .

$$Y_{c3} - t_{.025}SE_3 \leq Y(h,k) \leq Y_{c3} + t_{.025}SE_3 \quad (\text{Test 3})$$

### 8.3.2 Algorithme de détection et de correction des mesures

Cette section présente l'algorithme de détection et de correction d'une mesure du trafic faite à un instant donné  $k$  pour un ensemble  $M$  de PM.

1. Début du processus itératif de filtrage (itération  $i=0$ )
2. Pour tous les PM,  $m \in [1,M]$
3. Test de donnée manquante  
Si la donnée est manquante alors  $Y^i(k,m) = Y_{c1}(k,m)$
4. Test de variation temporelle (test 1)  
Si le test est faux, saut à l'étape 5  
Sinon, on continue à l'étape 6
5. Test de variation spatiale (test 2)  
Si le test est faux alors  $Y^i(k,m) = Y_{c2}^i(k,m)$
6. Test de variation autour d'une courbe théorique (test 3)  
Si le test est faux alors  $Y^i(k,m) = Y_{c3}(k,m)$
7. Test si  $m < M$   
Si le test est vrai, sélection du PM suivant, retour en 2  
Sinon, on continue en 8.
8. Si le test de convergence ou si le nombre d'itérations max. sont satisfaits

⇒ Arrêt du processus.  
Sinon, itération suivante  $i=i+1$  et retour en 2.

Pour un ensemble de données manquantes ou aberrantes  $M$ , le test de convergence est le suivant :

$$\varepsilon = \sum_M \left( \frac{Y^i(k,m) - Y^{i-1}(k,m)}{Y^i(k,m)} \right)^2 < \varepsilon_{\min}$$

### 8.3.3 Résultats d'expérimentation

Il est apparu important de vérifier le comportement général de la méthode sur plusieurs sites, où les caractéristiques sont différentes (géométrie du terrain, comportement des automobilistes, système de mesure du trafic, etc.). Les résultats présentés ici sont relatifs à deux sites ; le réseau urbain de la ville de Lyon (réseau I) et le réseau urbain de la ville de Toulouse (réseau II). Les expériences menées sur ces deux réseaux, se sont inscrites dans le cadre du projet européen CLEOPATRA (City Laboratories Enabling Organisation Of Particularly Advanced Telematics Research and Assesment) pour le premier réseau (Bréheret et Schettini, 1997), et dans le cadre du projet TITUS (Toulouse Information Traffic Usagers) pour le second (Bréheret et Schettini, 1998).

Le tableau ci-dessous indique au lecteur le nombre de PM utilisés pour l'évaluation, la période de mesure des données de comptage ainsi que celle de calcul.

Tableau 6: Description du système de recueil de données des réseaux expérimentaux

Réseau	Poste de mesure	Périodes	
		Recueil	Calcul
I	150	6 minutes	6 minutes
II	75	3 minutes	3 minutes

L'évaluation a permis de mesurer à la fois :

- **des indicateurs de performance globaux**

Les résultats obtenus s'intéressent à un critère global intégré dans le système de fusion de données (Tableau 7). Le critère de qualification consiste à calculer le pourcentage d'erreurs de mesures détectées par rapport à celles réellement observées (moyenne pour l'ensemble des PM). Un critère de 100 % signifie que toutes les erreurs ont été détectées. Les pourcentages mesurés sont très bons, de l'ordre de 99 % pour les deux types de recueil. Les résultats semblent cohérents car la grande majorité des problèmes est due au manque de données ce qui est donc très simple à détecter.

Tableau 7 : Evaluation de la qualité des résultats

Critères	Réseau I	Réseau II
Pourcentage des erreurs détectées	98,61	99,1

- **des indicateurs de performance individuels**

Le Tableau 8 détaille pour un échantillon arbitraire de PM (10 pour le réseau I et 5 pour le réseau II), des indications telles que le nombre de fausses erreurs détectées et celui

d'erreurs non détectées. Sur ces échantillons, 96,9 % et 98 % de données erronées ont été détectées par l'algorithme de qualification pour chacun des réseaux tests. Toutefois, les données détectées contiennent des mesures cohérentes mais pourtant traitées comme des anomalies. Une analyse détaillée de ces anomalies souligne que ce phénomène apparaît lorsque le PM est isolé, c'est à dire lorsqu'il ne dispose pas de PM dans un voisinage très proche. Si une variation de trafic survient sur un PM isolé, cette variation ne produit, dans le temps et dans l'espace, aucun effet sur les autres. Dans ce cas, si les seuils de détection d'erreurs ne sont pas adaptés au PM isolé, la mesure est considérée comme non vraisemblable et donc corrigée.

Le nombre d'erreurs non détectées n'est pas négligeable : 91 pour le réseau I et 102 pour le réseau II. En terme de pourcentage, cela représente 3,7% et 3,4% respectivement pour les deux réseaux, par rapport au nombre d'erreurs non détectées. Là aussi, une analyse plus fine a montré que la majorité de ces mesures erronées ne sont pas admissibles dans le plan  $(q, to)$  sans que le test 3 de la procédure de qualification n'ait détecté cette anomalie. Cela signifie que le modèle de courbe théorique n'est pas adapté au nuage de points expérimentaux et par conséquent le test, construit autour de ce modèle, n'est pas valide pour toutes les mesures. La superposition de la courbe estimée sur les nuages de points expérimentaux montre que pour certains PM, le test 3 ne peut s'appliquer. En effet, ces PM particuliers présentent de telles variations de débit de véhicules et de taux d'occupation, qu'il est impossible d'extraire un modèle théorique adéquat. C'est notamment le cas de certains PM à plusieurs boucles de mesure dont l'une est fréquemment saturée (cas de stationnements illicites, etc.).

Tableau 8 : Evaluation du module de qualification des données

Réseau	Nb de PM	Nb d'erreurs observées	Nb d'erreurs détectées	Nb de fausses erreurs détectées	Nb d'erreurs non détectées
I	10	2398	2325	18	91
II	5	2990	2930	42	102

### Exemple de données corrigées

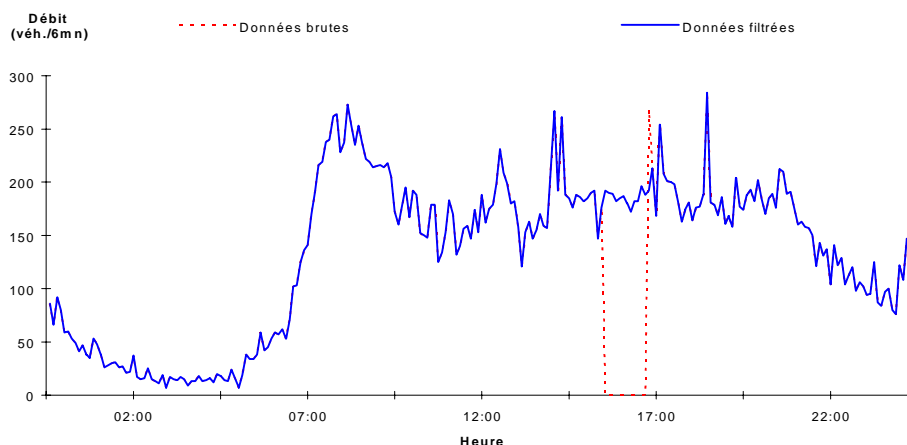


Figure 15 : Détection et correction des données manquantes



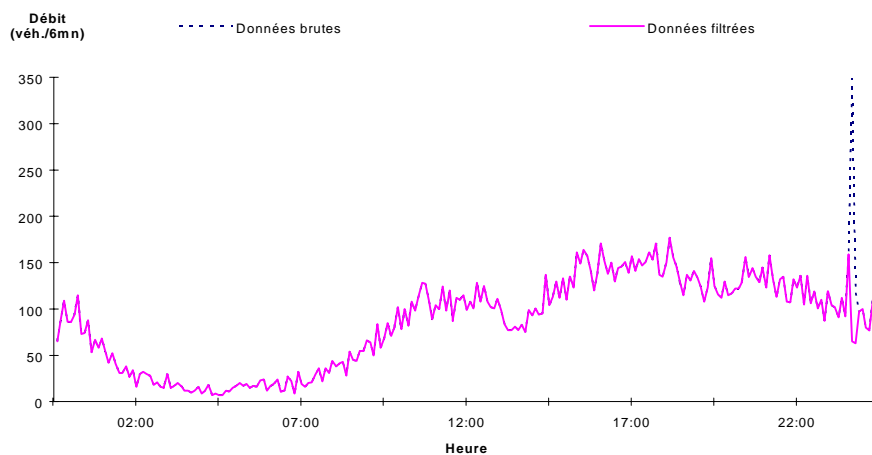


Figure 16 : Détection et correction d'une valeur aberrante (tests 1 et 2)

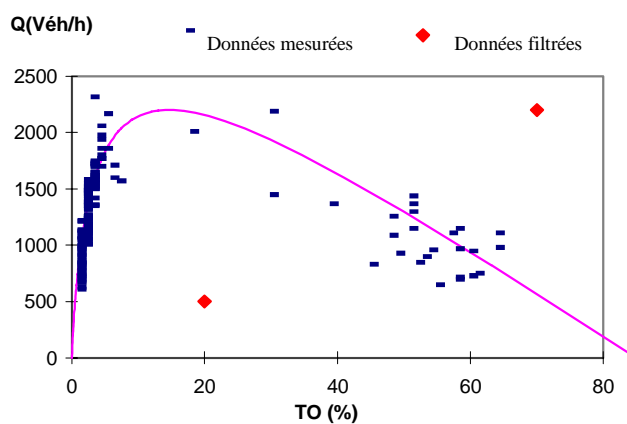


Figure 17: Détection et correction d'une valeur aberrante (Test 3)

## 8.4 Compléments sur la méthode de qualification des données de véhicules traceurs

### 8.4.1 Exemple d'un chaînon à deux voies

Différents cas de figure peuvent se présenter :

- Cas d'une seule distribution (Figure 19 cas 1), lente ou rapide, qui traduit que l'état est sensiblement le même sur les deux voies. On prend en compte toutes les vitesses pour le calcul du temps de parcours.
- Cas de deux distributions distinctes (Figure 19 cas 2), correspondant à une distribution par voie. On peut calculer un temps de parcours par voie et utiliser un temps de parcours moyen sur le chaînon.
- Cas de deux distributions dont les moyennes sont proches (Figure 19 cas 3), et qui se chevauchent sur certains intervalles. Ce cas n'est pas gênant car les moyennes de vitesse sur chaque voie étant proches, on prend en compte les deux voies pour le calcul du temps de parcours ; on a donc pas besoin de dissocier les deux distributions.
- Cas de deux distributions correspondant aux deux voies du périphérique (Figure 19 cas 4), et d'une troisième distribution introduite par l'existence d'une route annexe. Le graphique représente un cas de figure relativement rare car en général la taille en effectif de la distribution de vitesse de la route annexe est très inférieure à celle des distributions du périphérique. Donc finalement, quelques mobiles joueront le rôle de parasites, mais pas suffisamment pour fausser de manière significative l'estimation du temps de parcours.

Une fois la répartition effectuée, il s'agit de repérer les maximums locaux d'effectif, puis de se fixer un intervalle autour de ce maximum, intervalle qui est fonction de la moyenne considérée, et d'éliminer les vitesses n'appartenant à aucun intervalle. Ce repérage des distributions a été simplifié par l'hypothèse adoptée et précisant que l'état du trafic est connu. La figure ci-dessous est une illustration de ce repérage.

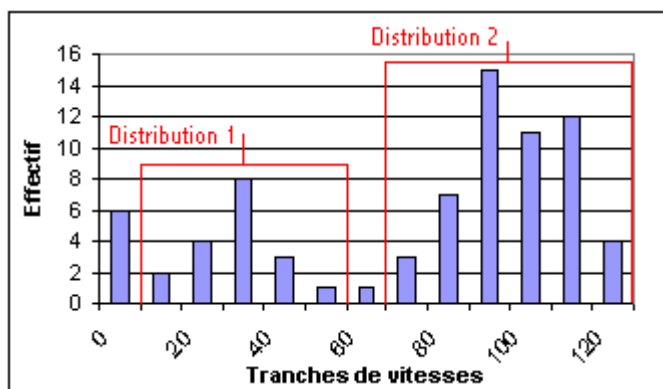


Figure 18 : exemple de distributions multiples

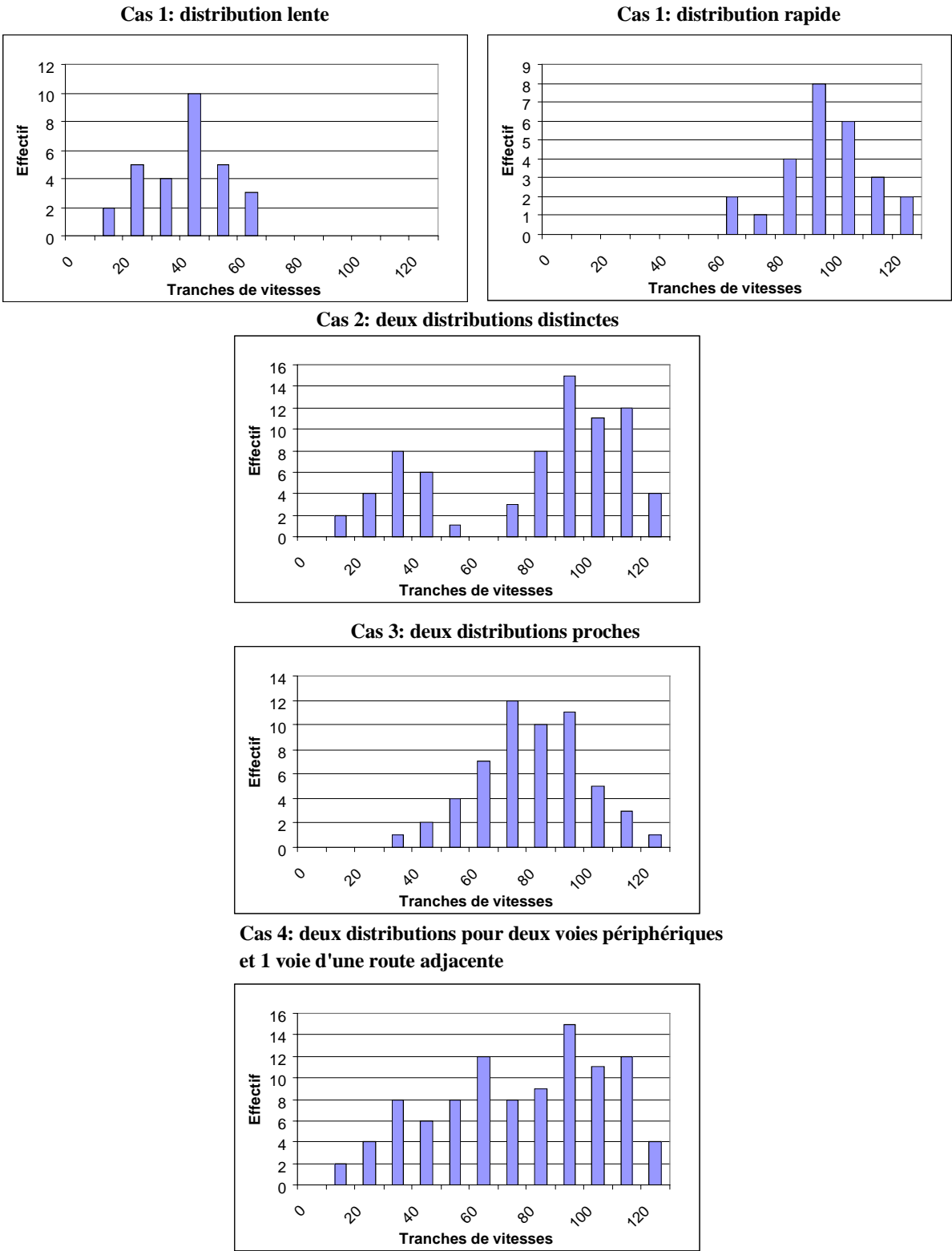


Figure 19: Exemples de distributions de vitesses

### 8.4.2 Résultats des tests

Différents tests ont été effectués pour évaluer les performances de l'algorithme de qualification des temps de parcours des véhicules traceurs. Ces tests ont porté sur différentes configurations de la topologie du réseau :

- Si l'on considère un axe périphérique seul, on constate que les taux d'incertitude obtenus restent faibles et inférieurs à 5%. On remarque que plus la distribution de vitesse a une moyenne faible, et plus l'incertitude est petite. Ceci peut être dû au fait d'une part qu'à faible vitesse, les véhicules restent plus longtemps dans la zone de localisation et sont donc repérés plus souvent, et d'autre part au fait que, dans le test effectué, l'étalement de la distribution diminue avec sa moyenne.

Tableau 9 : Moyennes des écarts des temps de parcours estimés et simulés pour un écoulement fluide suivant le taux d'équipement de GSM (cas d'un axe périphérique seul)

Taux GSM	5%	10%	20%	30%	40%	50%
Moyenne des écarts	2,9%	2,2%	1,7%	1,5%	1,8%	1,2%
Ecart type des écarts	2,1%	1,4%	1,1%	1,2%	1,2%	0,7%

Tableau 10: Moyennes des écarts des temps de parcours estimés et simulés pour un écoulement lent suivant le taux d'équipement de GSM (cas d'un axe périphérique seul)

Taux GSM	5%	10%	20%	30%	40%	50%
Moyenne des écarts	3,8%	3,8%	1,8%	1,2%	1,2%	1,3%
Ecart type des écarts	6,4%	3,4%	1,2%	0,9%	1,0%	1,3%

- L'addition de perturbations (axe périphérique et route adjacente) entraîne des variations sensibles dans les résultats sur l'incertitude ; on obtient en effet des taux un peu plus élevés, dépassant quelquefois le seuil des 5 %, mais ne franchissant jamais la barre des 10%.

Si la vitesse moyenne de la route adjacente est très proche de celle du périphérique, elle influe peu sur le temps de parcours, car les quelques points pris en compte appartenant à cette route vont se fondre dans la population repérée sur le chaînon.

Si cette vitesse moyenne est très inférieure ou très supérieure à celle du périphérique, un système bien élaboré repérera une distribution pauvre en effectif, et l'éliminera.

En revanche, si cette vitesse moyenne se trouve en bout d'intervalle d'étalement de la distribution de vitesse du périphérique, elle aura pour effet d'augmenter cet étalement, et donc d'introduire une erreur plus significative que dans les précédents cas.

On constate également que la taille de la zone de localisation joue un grand rôle dans le filtrage lorsqu'il y a une route adjacente, car il est évident que plus on peut repérer loin, plus on est apte à observer l'éloignement d'un mobile par rapport à l'axe étudié.

Tableau 11 : Moyennes des écarts des temps de parcours estimés et simulés suivant le taux d'équipement de GSM (cas d'un axe périphérique et d'une route adjacente)

Taux GSM	5%	10%	20%	30%	40%	50%
Moyenne des écarts	7,10%	7,20%	7,10%	7,10%	6,20%	6,70%
Ecart type des écarts	4,00%	4,40%	3,50%	3,80%	3,90%	4,30%

## 8.5 Compléments sur la méthode de filtrage et d'estimation

### 8.5.1 Modèles utilisés en zone urbaine

#### 8.5.1.1 Evolution des queues

En zone urbaine, le trafic dans un chaînon est modélisé suivant un modèle à queue verticale : chaque véhicule traverse le chaînon à vitesse libre et s'empile à la ligne de feux dans une queue verticale. Ce modèle permet de calculer la queue d'un chaînon en fonction de la queue à l'instant précédent, du nombre de véhicules entrés dans le chaînon pendant la période précédente et du nombre de véhicules qui en sont sortis. Si  $k$  et  $k+1$  représentent deux périodes successives, l'eq. 3 traduit l'état de la queue.

$$\text{eq. 3} \quad X[k+1] = \max ( 0 ; X[k] + (Q[k] - \text{cap}[k]). T )$$

Lorsqu'il y a des feux en fin de chaînon, la capacité s'exprime en fonction du débit de saturation et des durées de cycle et de rouge (eq. 4). Sans feu, la capacité est égale à la saturation.

$$\text{eq. 4} \quad \text{cap} = s.(c - r)/c$$

#### 8.5.1.2 Temps de parcours

Un modèle de temps de parcours cohérent avec le modèle à queue verticale est défini pour les chaînons urbains. Le modèle  $TP(Q,X)$  utilisé en urbain s'adapte à l'existence de feux en fin de chaînon. Il permet ainsi de prendre en compte les problèmes de saturation pouvant survenir.

Le temps de parcours d'un véhicule sur un chaînon (eq. 5) est la somme du temps de parcours à vitesse libre nécessaire pour traverser le chaînon et du temps d'attente (ou de retard) occasionné par les files d'attente. Ce temps d'attente est une valeur moyenne sur la période considérée.

$$\text{eq. 5} \quad TP (Q,X) = TPI + \frac{1}{T} \int_0^T TA(t) dt$$

où  $TA(t)$  est le temps d'attente pour un véhicule arrivant à l'instant  $t$  dans la queue. Le débit est supposé constant sur la période.

Trois types de trafic pour un chaînon ont été identifiés (Figure 20) : le cas fluide, le cas dessaturant, et le cas saturé. Ces trois cas sont déterminés par une intégration récursive de la queue.

- pour le cas fluide, la queue en début de période est nulle et le débit d'entrée est inférieur à la capacité,
- pour le cas dessaturant, la queue en début de période est non nulle et la capacité est supérieure au débit d'entrée,
- pour le cas saturé, la queue en fin de période est non nulle.

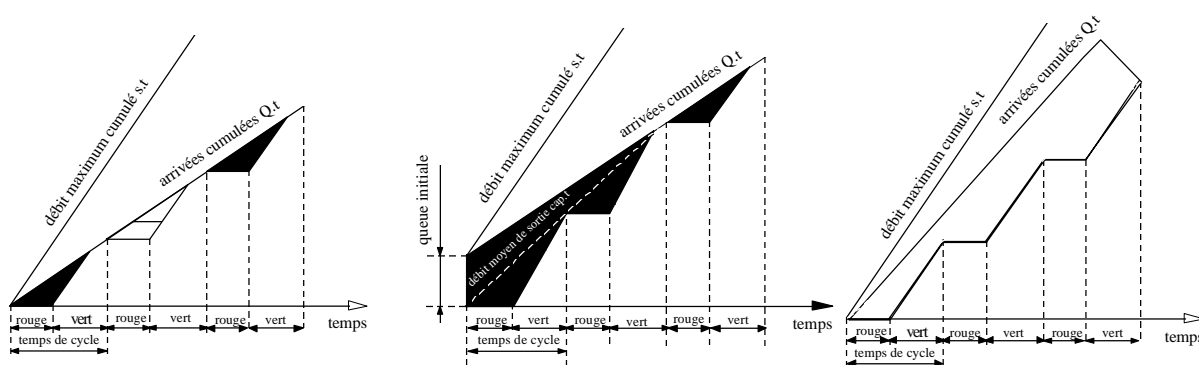


Figure 20: Evolution de la file d'attente pour les 3 cas fluide, dessaturant et saturé

Le temps d'attente  $TA(t)$  peut alors être calculé analytiquement pour ces trois cas possibles puis intégré pour obtenir le temps de parcours. Deux exemples de courbes  $TP(Q,X)$  sont donnés sur la Figure 21, la courbe de gauche pour un chaînon avec feux, celle de droite pour un chaînon sans feu. La courbe la plus basse correspond à  $X=0$ , puis  $X=3$ ,  $X=5...$

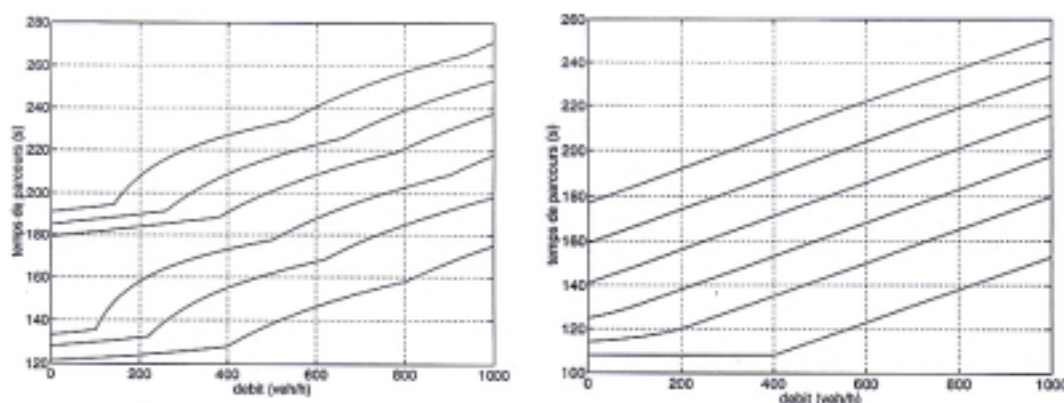


Figure 21 :  $TP(Q,X)$  pour deux chaînons, l'un avec feux ( $cap = 400 \text{ véh/h}$ ), l'autre sans feux

Le temps de parcours calculé par le modèle  $TP(Q,X)$  a été comparé avec celui donné par simulation avec le logiciel SITRA-B+ (simulateur microscopique de trafic routier) pour un chaînon avec feux et un débit variable dans le temps par période. La comparaison a également été faite avec les temps de parcours analytiques calculés avec les modèles classiques de Wardrop et de Webster. Par rapport aux données simulées, l'erreur moyenne pour  $TP(Q,X)$  est de 3%. Elle est de 26% avec le modèle de Wardrop (bon en fluide mais pas en saturé) et de 17% avec le modèle de Webster (mais ce dernier modèle donne une valeur infinie à la saturation). Le modèle  $TP(Q,X)$  se comporte donc bien, notamment lorsque le chaînon est saturé.

### 8.5.1.3 Taux d'occupation

Le modèle du taux d'occupation utilisé pour les chaînons urbains ne prend pas en compte les contraintes liées aux feux. L'eq. 6 donne le taux d'occupation en fonction du débit et du temps de parcours d'un chaînon.

$$eq. 6 \quad TO(Q,TP) = Q \cdot (d+ 1) \cdot TP / L$$

## 8.5.2 Modèles de trafic adaptés aux trajets

### 8.5.2.1 Temps de parcours

Pour avoir le temps de parcours estimé sur un trajet, il faut sommer les temps de parcours des différents chaînons composant le trajet. Les temps de parcours des chaînons de la couche II du réseau géographique sont déterminés par la méthode de résolution décrite dans le paragraphe 5.2.2.3 en intégrant les modèles définis ci-dessus.

La formulation suivante permet de calculer le temps de parcours,  $TP_{trajet}$ , sur un trajet  $j$  composé d'un ensemble de chaînon  $E_j$ .

$$eq. 7 \quad TP_{trajet,j} = \sum_i TP_{chaînon,i}$$

### 8.5.2.2 Niveaux de circulation

Une définition possible de 6 niveaux de circulation est :

- Le niveau fluide 1 correspond à un écoulement libre avec des débits faibles et des vitesses élevées.
- Le niveau 2 correspond à un écoulement stable et des vitesses légèrement diminuées ne gênant pas trop la conduite des automobilistes.
- Le niveau 3 correspond toujours à un écoulement stable mais les conducteurs ne sont plus libres dans leur manœuvre.
- Le niveau 4 se rapproche de l'écoulement instable, avec des vitesses baissant sérieusement et des fluctuations de débits.
- Dans le niveau 5, les débits sont de l'ordre de la capacité de la route, avec des arrêts temporaires des véhicules.
- Dans le niveau 6, les vitesses sont très faibles et le débit est inférieur à la capacité.

La fonction de calcul des niveaux de service a pour entrée certaines variables de trafic associées aux chaînons et pour sortie une qualification du service, soit sur un chaînon, soit sur un trajet. Nous proposons d'utiliser la classification de 4 niveaux de service qui semble la plus connue des usagers. Si le choix final des exploitants est différent, une correction des seuils devra être réalisée.

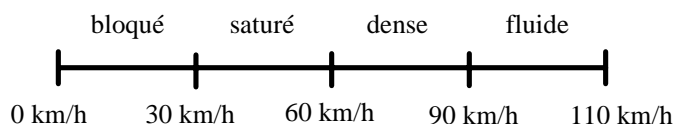
Les correspondances entre variables de trafic (temps de parcours, vitesses, débits, taux d'occupation) et niveaux de circulation dépendent du type de chaînon (urbain ou périurbain). On ne circule pas en effet de la même façon sur voie rapide et dans les rues du centre urbain.

Pour un chaînon quelconque, des valeurs par défaut pour les seuils sont proposées. Elles sont modifiables par l'exploitant. Des enquêtes sur le terrain pourraient également permettre de quantifier le niveau de service en fonction de la perception de l'utilisateur. L'indice de satisfaction d'un usager est lié à sa connaissance de la circulation sur la portion de route considéré.

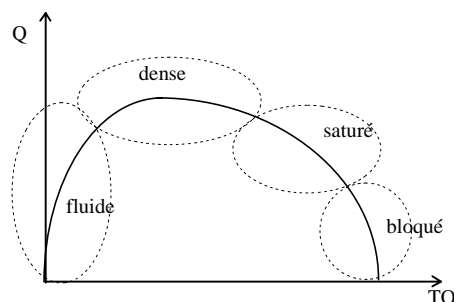
Le niveau fluide correspond à un temps de parcours à vitesse libre ou à une vitesse maximale autorisée.

Calcul des seuils par défaut pour un chaînon périurbain :

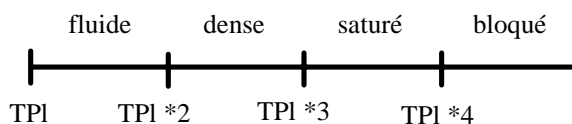
- en utilisant les vitesses et une répartition quasi uniforme entre 0 km/h et 110 km/h



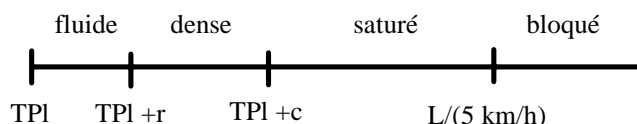
- en utilisant les données historiques du comportement moyen connu  $Q(TO)$



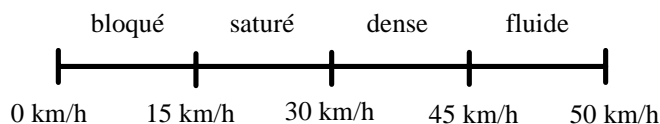
- en utilisant les temps de parcours



Pour un chaînon urbain avec feux, le niveau de service est lié au passage au niveau du feu (un rouge maximum si fluide, au moins un cycle si dense) et au gain par rapport à la vitesse d'un piéton (5 km/h).



Pour un chaînon urbain sans feu, on peut utiliser une des trois méthodes développées en périurbain (et modification des seuils par défaut de la première méthode).



Calcul du niveau de service pour un trajet, comportant plusieurs chaînons donnant des informations de niveaux de service différents : il faut agréger les niveaux des chaînons. On peut être plutôt optimiste (meilleur niveau affiché), plutôt pessimiste (plus mauvais niveau affiché) ou moyen. Dans ce dernier cas, on peut affecter un chiffre à chaque niveau de 1 à 4, sommer les chiffres des différents chaînons composant le trajet, diviser par le nombre de trajets et prendre la valeur la plus proche (à 0,5 près). Sur un exemple :



5 chaînons, fluide, fluide, dense, dense, bloqué, 1+1+2+2+4=10 et 10/5=2 donc affichage dense.

### 8.5.3 Méthodes de résolution

L'objectif de la méthode de filtrage est d'obtenir des débits et des taux d'occupation le plus proche possible de leur mesure pour tous les chaînons et sur tout l'horizon de filtrage. Les relations entre variables sont décrites par les modèles de trafic. Si les relations sont simples le problème d'optimisation peut être résolu par des méthodes classiques comme le gradient. Des relations non linéaires rendent plus complexes la résolution.

La méthode proposée ne permet pas de filtrer directement les variables de trafic pour les chaînons ne possédant aucune mesure sur l'horizon. Une étude spatiale peut cependant la compléter.

Le critère à minimiser est une combinaison quadratique sur l'horizon de filtrage et sur les chaînons des écarts entre les valeurs mesurées puis corrigées et les variables associées. On note avec l'exposant  $i$  le numéro du chaînon et avec l'indice  $c$  les variables corrigées après leur qualification. Par exemple,  $Q_c^i[k]$  est le débit corrigé sur le chaînon  $i$  à l'instant  $k$ .

Le critère général est donné par l'eq.8.

$$\begin{aligned}
 \text{eq. 8 } \sum_{k \in \text{Hf}} [ & \alpha_1 \cdot \{ \sum_{\text{chaînons urbains avec postes de mesure } i} (Q_c^i[k] - Q^i[k])^2 \} \\
 & + \alpha_2 \cdot \{ \sum_{\text{chaînons périurbains avec postes de mesure } i} (Q_c^i[k] - Q^i[k])^2 \} \\
 & + \alpha_3 \cdot \{ \sum_{\text{chaînons urbains avec postes de mesure } i} (TO_c^i[k] - TO^i[k])^2 \} \\
 & + \alpha_4 \cdot \{ \sum_{\text{chaînons périurbains avec postes de mesure } i} (TO_c^i[k] - TO^i[k])^2 \} \\
 & + \alpha_5 \cdot \{ \sum_{\text{chaînons périurbains avec postes de mesure } i} (V_c^i[k] - V^i[k])^2 \} \\
 & + \{ \sum_{\text{chaînons avec données GSM } i} \alpha_6 \cdot (TP_c^i[k] - TP^i[k])^2 \} ]
 \end{aligned}$$

Les valeurs recherchées sont les débits et les taux d'occupation sur les chaînons urbains et périurbains équipés de postes de mesure, les vitesses sur les chaînons périurbains et les temps de parcours pour tout chaînon possédant au moins une mesure sur l'intervalle de filtrage.

Les variables du problème d'optimisation global sont donc les débits  $Q$  pour les chaînons urbains et les débits  $Q$  et les taux d'occupation  $TO$  pour les chaînons périurbains. Les relations reliant les différentes variables en fonction du type de chaînon sont les modèles décrits dans le paragraphe précédent et donnés dans le Tableau 12.

Tableau 12: Modèles utilisés en urbain et périurbain en présence de données postes de mesure

En urbain	En périurbain
$TP(Q,X)$	$TP(Q,TO)$
$TO(Q,TP)$	$V(TP)$
$X+(Q,X)$	

Les coefficients  $\alpha_j$  permettent de pondérer les différents types de données en fonction de la précision connue (voir le paragraphe sur les données d'entrée). Le coefficient  $\alpha_6$  dépend du chaînon (c'est-à-dire du nombre de mesures GSM disponibles sur la période considérée).

Si on suppose que les données sont connues pour chaque chaînon indépendamment des autres chaînons, la minimisation du critère global peut être résolue pour chaque chaînon. Le Tableau 13 présente les variables du problème d'optimisation ainsi que le critère associé à chaque scénario en fonction des données d'entrée.

Tableau 13 : Variables et critères à minimiser pour chaque scénario de disponibilité de mesures

Mesures disponibles sur l'horizon de filtrage	En urbain	En périurbain
Poste de mesure	<p>Recherche de <math>Q</math> sur <math>H_f</math></p> <p>Minimisant le critère</p> $\sum_{k \in H_f} [ \alpha_1 \cdot (Q_c[k] - Q[k])^2 + \alpha_3 \cdot (TO_c[k] - TO[k])^2 ]$	<p>Recherche de <math>Q, TO</math> sur <math>H_f</math></p> <p>Minimisant le critère</p> $\sum_{k \in H_f} [ \alpha_2 \cdot (Q_c[k] - Q[k])^2 + \alpha_4 \cdot (T_c[k] - TO[k])^2 + \alpha_5 \cdot (V_c[k] - V[k])^2 ]$
Poste de mesure + GSM	<p>Recherche de <math>Q</math> sur <math>H_f</math></p> <p>Minimisant le critère</p> $\sum_{k \in H_f} [ \alpha_1 \cdot (Q_c[k] - Q[k])^2 + \alpha_3 \cdot (TO_c[k] - TO[k])^2 + \alpha_6 \cdot (TP_c[k] - TP[k])^2 ]$	<p>Recherche <math>Q, TO</math> sur <math>H_f</math></p> <p>Minimisant le critère</p> $\sum_{k \in H_f} [ \alpha_2 \cdot (Q_c[k] - Q[k])^2 + \alpha_4 \cdot (TO_c[k] - TO[k])^2 + \alpha_5 \cdot (V_c[k] - V[k])^2 + \alpha_6 \cdot (TP_c[k] - TP[k])^2 ]$

Quand, pour un chaînon quelconque, on dispose seulement de données GSM, on peut utiliser directement les données en ayant conscience de la précision affichée par l'opérateur GSM. Il n'y a pas de filtrage effectif, mais seulement un lissage sur les différentes mesures en fonction de la précision associée à chaque mesure.

### 8.5.3.1 Résolution locale dans le cas urbain

Deux périodes successives sont uniquement reliés par l'eq. 3 qui donne la queue à une période en fonction de la queue à la période précédente. Il est donc possible de résoudre la minimisation pour un chaînon urbain sur l'horizon de filtrage en minimisant successivement sur chaque période de l'horizon le critère correspondant à la période. Ces problèmes d'optimisation mono-dimensionnel peuvent être facilement résolus par une méthode basée sur le nombre d'or.

Parmi les méthodes analytiques itératives de résolution de problème d'optimisation non linéaire mono-variable, la méthode du nombre d'or est légèrement moins performante que la méthode de Fibonacci, mais elle est plus facile à mettre en œuvre et à implémenter sur calculateur. Avec cette méthode, on réduit itérativement l'intervalle de recherche en calculant le critère en deux points dans l'intervalle courant, ces deux points étant situés à la distance  $(-1 + \sqrt{5})/2$  des points extrémaux courants, et en choisissant le nouvel intervalle en fonction de la valeur optimale du critère sur les deux nouveaux points. La précision de la méthode dépend directement du nombre d'itérations.

#### Résultats issus de projets de recherche

Cette méthode de filtrage par minimisation quadratique a été utilisée dans le cadre du projet européen PROMETHEUS. Une partie de la ZELT (Zone Expérimentale et Laboratoire de Trafic de Toulouse) a été modélisée avec le logiciel de simulation

microscopique SITRA-B+. Des mesures de type postes de mesure étaient disponibles sur certains chaînons, des temps de parcours étaient transmis soit par des véhicules équipés d'un système de transmission par balises infrarouge (cas 1), soit par des véhicules équipés d'un système de communication radio-cellulaire (cas 2). Le modèle TP(Q,X) a été utilisé dans les deux cas. Le problème bi-variables du cas 1 a été résolu globalement d'une manière itérative : résolution par rapport à la première variable, puis par rapport à la seconde variable, ceci tant que le critère diminuait. Le problème mono-variable du cas 2 a été résolu avec la méthode du nombre d'or.

Des taux de pénétration de 1 à 20% de véhicules équipés ont été simulés sur l'ensemble du réseau. Sur les 52 chaînons du réseau, les erreurs moyennes sur les débits et les temps de parcours sont cohérents avec les précisions demandées. Pour la communication par balise (cas 1), l'erreur est inférieure à 9 véh/h pour les débits et à 2 s pour les temps de parcours. Pour la communication radio-cellulaire (cas 2), l'erreur est inférieure à 8 véh/h pour les débits et à 2 s pour les temps de parcours. Si la durée de calcul pour le radio-cellulaire était très rapide (quelques secondes), le critère global pour les balises a mis trop longtemps à converger (2 heures en moyenne), n'autorisant pas une implémentation sur site tel quel.

La méthode de filtrage a été validée sur 15 chaînons urbains équipés de postes de mesure dans le cadre du projet européen ATT/CLEOPATRA (données débits et taux d'occupation collectées sur le site de la Ville de Toulouse), certains possédant des feux en sortie. L'erreur absolue moyenne par comparaison entre les débits filtrés et les débits mesurés est de 3% seulement. Cette méthode a été appliquée sans correction préalable par la fonction de qualification des données, mais les valeurs incohérentes ont été partiellement corrigées par la méthode de par les modèles utilisés. Les courbes fondamentales Q(TO) sont ainsi parfaitement cohérentes, pour les chaînons avec et sans feux (Figure 22).

Aucune mesure dynamique de temps de parcours n'était disponible lors de cette étude, et le temps de parcours a été directement calculé à partir du modèle. Sa comparaison avec quelques mesures provenant de véhicules flottants n'est pas suffisante.

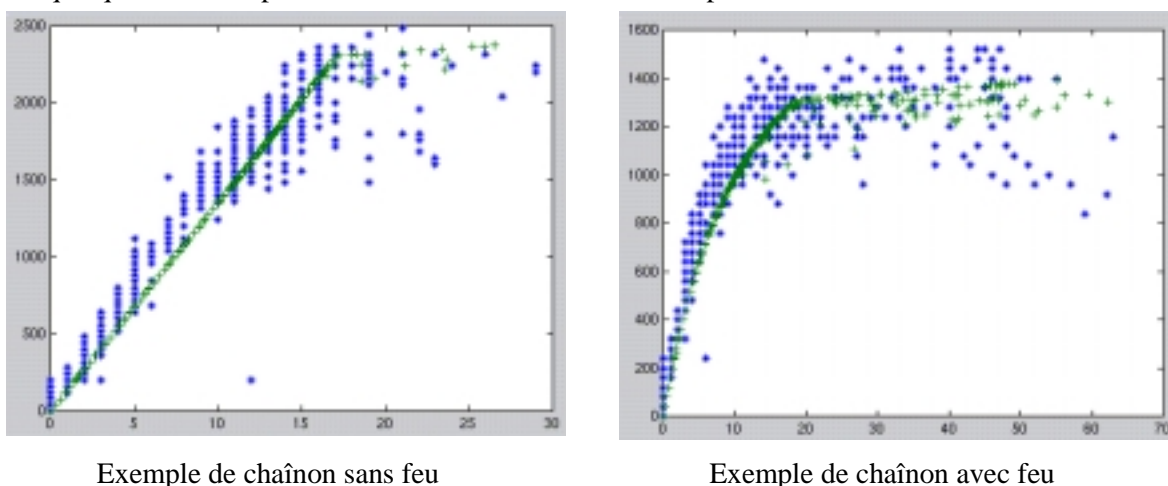


Figure 22: Courbes Q(TO): comparaison entre données mesurées (x) et variables filtrées (+)

La méthode de filtrage a été intégrée dans une stratégie de guidage par panneaux à message variable également dans le projet CLEOPATRA sur la Ville de Toulouse. Les résultats sont bons sur les 73 chaînons renseignés, avec une sous-estimation faible des taux d'occupation (erreur absolue moyenne de 3%) et une surestimation faible des débits (erreur de 2%).

### 8.5.3.2 Résolution locale dans le cas périurbain

Les relations entre variables ne concernent qu'une période et, sous l'hypothèse d'indépendance spatiale, on peut résoudre le problème de filtrage pour la dernière période passée.

C'est un problème à deux variables ( $Q$  et  $TO$ ) avec des modèles de trafic relativement simple. On peut donc utiliser une méthode d'ordre 1 comme la méthode du gradient basée sur le développement du critère au premier ordre au sens de Taylor. La dérivée analytique du critère par rapport aux variables est directe. On peut ainsi se diriger à chaque itération dans la direction opposée à la plus grande augmentation du critère.

#### Reconstitution de l'état sur des chaînons sans données

Pour obtenir les débits de chaînons dont aucune donnée n'est recueillie, on peut utiliser les débits filtrés sur les chaînons avec postes de mesure. La contrainte de conservation des débits à une intersection (la somme des véhicules sortant des chaînons amont est égale à la somme des véhicules sortant des chaînons aval) permet de reconstituer l'état de chaînons n'ayant ni postes de mesure, ni information de temps de parcours.

Le problème peut être posé pour chaque nœud du réseau, avec comme variables connues les débits des chaînons avec variables filtrées et comme inconnues les débits des chaînons sans données. Si le problème à une intersection n'a pas d'inconnue, l'équation de conservation n'est pas forcément vraie mais le filtrage pour tous les chaînons adjacents a été résolu. Si le problème à une intersection a une inconnue, le débit peut être calculé directement. S'il y a deux ou plus inconnues, une méthode globale est nécessaire : identification de toutes les intersections avec une inconnue, résolution, reidentification, etc. Le réseau peut ne pas être couvert totalement si le nombre de mesures n'est pas suffisant (la méthode ne permet pas de reconstruire l'état d'un chaînon sans mesure et isolé des points de mesure).

Des méthodes de reconstitution globale des débits sur les chemins peuvent aussi être développées, avec le calcul d'une matrice de débits entre toutes les entrées et les sorties du réseau et l'affectation de ces débits sur les différents chemins possibles, ceci en collant aux mesures avec un seuil de tolérance.

#### Résultats issus de projets de recherche

Une reconstitution de l'état sur des chaînons sans données basée sur la conservation des débits aux intersections a été implémentée dans la stratégie de guidage par panneaux à message variable dans le cadre du projet CLEOPATRA (utilisation des  $Q$  et  $TO$  de la Ville de Toulouse). 65% des chaînons du réseau modélisé sont équipés de postes de mesure, 35% des données ont donc été reconstituées. Une démonstration sur le site et l'analyse de l'impact de la stratégie par l'étude des postes de mesure en aval des points de décision a montré l'intérêt de la stratégie. Mais une analyse de l'impact de la méthode de reconstitution aurait nécessité une enquête plus précise au niveau des chaînons non équipés.

### 8.5.3.3 Résolution globale de l'optimisation

Si l'hypothèse d'indépendance spatiale n'est pas prise, le problème de l'éq. 8 ne peut pas être décomposé par chaînon. Il faut alors rajouter des équations de relations entre les variables de débits des différents chaînons (conservation aux intersections). Le problème d'optimisation est complexe à résoudre.

Par exemple, pour 100 chaînons dans un réseau avec 30 intersections, avec 3 variables de trafic à identifier sur un horizon passé de 15 minutes (5 périodes de 3 minutes), le nombre

de variables est égal à 1500, le nombre d'équations à 500 (TP(Q,X)) + 500 (TO(Q,TP) ou V(Q,TP)) + 500 (X(k)) + 150 (conservation aux intersections).

Résultats issus de projets de recherche

Le premier résultat de résolution locale dans le cas urbain met en avant la durée trop longue de la convergence d'un gradient sur deux variables d'optimisation, un horizon de filtrage non réduit à une période et l'utilisation de modèles de trafic complexes.

## 8.6 Compléments sur la méthode de prédiction

### 8.6.1 Méthode basée sur la combinaison Kalman / données historiques

L'objectif de la méthode de prédiction est d'obtenir des variables de trafic sur l'horizon de prédiction demandé. Ces variables sont principalement des débits et des temps de parcours sur les chaînons du réseau. La méthode fournit également les queues en début de chaque période future.

L'équation de mesure (eq. 9) dans l'estimateur discret de Kalman exprime la nouvelle variable mesurée puis corrigée en fonction des variables passées. Soit  $Y$  la variable générale pouvant être un débit, un taux d'occupation, une vitesse ou un temps de parcours.  $C[k]$  est la matrice des valeurs passées de  $Y$ .  $v[k]$  est le bruit de mesure qui dépend du type de données.

$$\text{eq. 9} \quad Y_c[k] = a[k] \cdot C[k] + v[k]$$

Le processus de Kalman génère les estimées optimales de  $a[k]$  et de  $b[k]$ . La meilleure prédiction à  $k$  de  $a[k+1]$  est la variable estimée  $b[k]$ . L'équation 10 permet de calculer la variable prédite  $Y_{pp}[k+1]$ .

$$\text{eq. 10} \quad Y_{pp}[k+1] = b[k] \cdot C[k+1]$$

Il suffit ensuite de combiner cette valeur avec la valeur historique  $Y_h$  suivant l'eq. 11.

$$\text{eq. 11} \quad Y_p[k+1] = \alpha \cdot Y_{pp}[k+1] + (1 - \alpha) \cdot Y_h$$

avec  $0 < \alpha < 1$ .

Ce processus appliqué d'une période à l'autre peut l'être sur tout l'horizon de prédiction. Les mesures  $Y_p$  n'étant pas disponibles, on utilise la valeur précédemment prédite dans la matrice  $C$ . Le poids historique dépend de la position de la période sur l'horizon, il augmente avec l'incertitude sur la prédiction.

#### Résultats issus de projets de recherche

Cette méthode de prédiction a été validée sur 10 chaînons urbains équipés de postes de mesure dans le cadre du projet européen ATT/CLEOPATRA (données débits et taux d'occupation fournies par la Ville de Toulouse). Le filtrage utilisait 8 périodes passées et la prédiction a été réalisée sur les débits pour deux périodes (+3 et +6 mn). 70% des périodes disposent effectivement de mesures. La base de données ( $Q$ ,  $TO$ ) a été enrichie au fur et à mesure de l'avancée du temps. Le poids respectif des données historiques est de 30% et 40%. L'erreur absolue moyenne sur les débits pour les deux périodes est de 22%. Si on utilise seulement les données historiques pour la prédiction, on trouve une erreur de 29%. Le gain de la méthode est donc de 7% sur la prédiction des débits (entre 5 et 12% en fonction du chaînon). L'erreur absolue moyenne sur les temps de parcours est de 6% sur la première période et de 9% sur la seconde période. Ces valeurs plus faibles s'expliquent par l'allure de la courbe TP(Q,X) et donne plus de robustesse à la prédiction des temps de parcours.

La méthode de prédiction a été également validée sur des données de temps de parcours disponibles sur des chaînons de la Ville de Berlin (projet CLEOPATRA). Les temps de parcours ont été recueillis par l'intermédiaire de véhicules équipés d'un système de communication par balises. La période d'échantillonnage est de 5 mn, les données sur 217 jours ont permis de constituer une base de données historiques de 150 jours, et la prédiction a été réalisée directement sur les temps de parcours (seules données disponibles) sur les 67 autres jours, ceci pour 6 chaînons. Seulement 17% à 22% de périodes possédant des données, un algorithme d'estimation a permis de reconstituer les

données manquantes (utilisation des données prédites). La prédiction a été réalisée sur deux périodes, +5 et +10 mn. Le poids des valeurs historiques est de 30% pour ces deux périodes. L'erreur absolue moyenne sur les temps de parcours est comprise entre 14% et 28% (moyenne de 20%). Par comparaison avec une prédiction utilisant les données historiques seules, le gain est de l'ordre de 5%.

### 8.6.2 Méthode non conventionnelle basée sur la théorie des réseaux de neurones

Pour tous les PM, les mesures les plus récentes ainsi que des tendances historiques sont combinées dans un modèle non linéaire utilisant la théorie des réseaux de neurones (Box et Bourret, 1991). De la même manière que pour les modules décrits précédemment, deux modes de fonctionnement (hors-ligne et en ligne) sont nécessaires pour garantir une prédiction précise des variables macroscopiques du trafic.

- **Le mode de fonctionnement hors-ligne** a pour but de déterminer les caractéristiques statistiques (moyenne et variance des mesures par exemple) des séries temporelles relatives aux variables  $Y$  de chaque PM et de sélectionner les PM utiles lors du fonctionnement en ligne.

La procédure consiste à analyser les séries temporelles et à classer tous les PM en trois grandes catégories : trafic régulier et prévisible, trafic aléatoire mais prévisible enfin trafic chaotique et difficilement prévisible.

Ensuite, le nombre d'entrées issues du proche passé  $k-1$ ,  $k-2$ , ... et nécessaires pour la prédiction, est déterminé pour chaque PM, à partir d'une estimation des degrés d'auto-corrélation empiriques à la fois pour  $Q$ ,  $TO$ ,  $V$  et  $TP$ .

Lorsque le nombre de termes utiles et la classe du PM sont connus, le choix du modèle neuronal (nombre de couches, réseaux directs ou récurrents, etc.) dépendra des ressources systèmes allouées et de la précision souhaitée.

Une fois tous les points précédents spécifiés, un jeu de paramètres initiaux est affecté à chaque type de prédicteur. Il convient toutefois de les ajuster à la configuration du terrain en utilisant des jeux de données réelles dans une phase de calibrage

- **Le mode de fonctionnement en ligne** permet de prédire  $Y_p(k+1), \dots, Y_p(k+Hp)$  à chaque période d'échantillonnage  $k$ .  $Hp$  est l'horizon de prédiction.  $Y$  est un vecteur caractérisant les variables macroscopiques du trafic ( $Q$ ,  $TO$  et  $V$ ) pour les chaînons équipés d'un PM ainsi que le temps de parcours des chaînons disposant d'une mesure de temps de parcours fournie par les véhicules traceurs.

La procédure de prédiction du trafic à  $k+1$  pour un PM consiste à utiliser les données filtrées de  $k$ ,  $k-1$ ,  $k-2$ , ..., à sélectionner ensuite les valeurs historiques à  $k+1$ ,  $k$ ,  $k-1$ ,  $k-2$ , ..., les mieux corrélées avec les dernières mesures, et enfin à effectuer une propagation du vecteur d'entrée à travers toutes les couches du réseau de neurones. Pour prédire le trafic à  $k+2$  à partir de  $k$ , une procédure hiérarchique décale d'un pas toutes les données d'entrée puis réintroduit les valeurs prédites pour l'instant  $k+1$  dans le modèle. La procédure se poursuit jusqu'à ce que les  $Hp$  pas de prédiction soient atteints. La Figure 23 représente la prédiction hiérarchique.

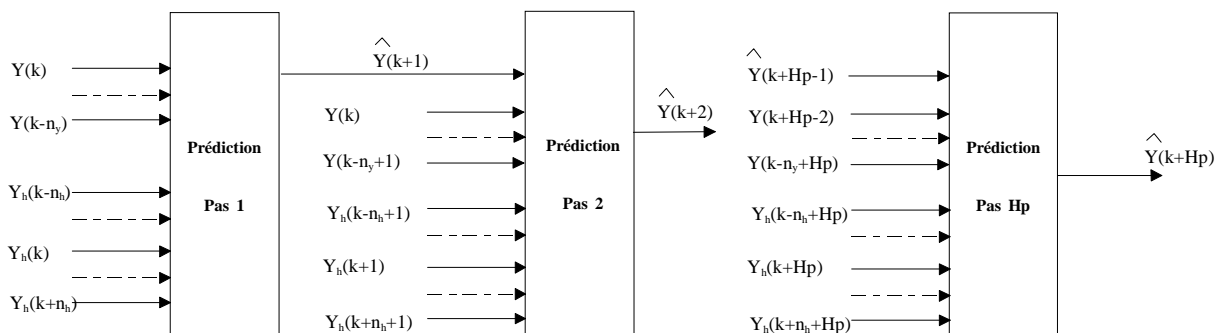


Figure 23 : Concept de prédiction sur plusieurs pas d'échantillonnage

Enfin, cette méthode de prédiction intègre une procédure d'ajustement automatique du modèle prévisionnel lorsqu'une dérive de la prédiction est constatée dynamiquement. En effet, les paramètres initiaux, qui sont calibrés sur des jeux de données historiques, ne correspondent pas forcément aux fluctuations journalières (situation de trafic non apprise par le modèle) ou aux évolutions sur une constante de temps beaucoup plus longue (modification du nombre de boucles, déplacement sur la voirie du PM).

L'objectif de cette procédure est de réduire l'erreur de prédiction dans les situations de trafic fortement perturbées tout en minimisant le nombre d'ajustements du modèle. Pour cela, la procédure calcule à chaque tranche de temps un taux d'erreur de prédiction (différence entre les valeurs prédites et mesurées sur les dernières tranches de temps) et vérifie si ce taux est acceptable. Dans le cas contraire, le modèle utilisé par défaut est automatiquement ajusté à partir des dernières mesures. Lors du retour à une situation stable, le modèle par défaut est réutilisé. La Figure 24 ci-dessous schématise la règle d'adaptation des modèles.

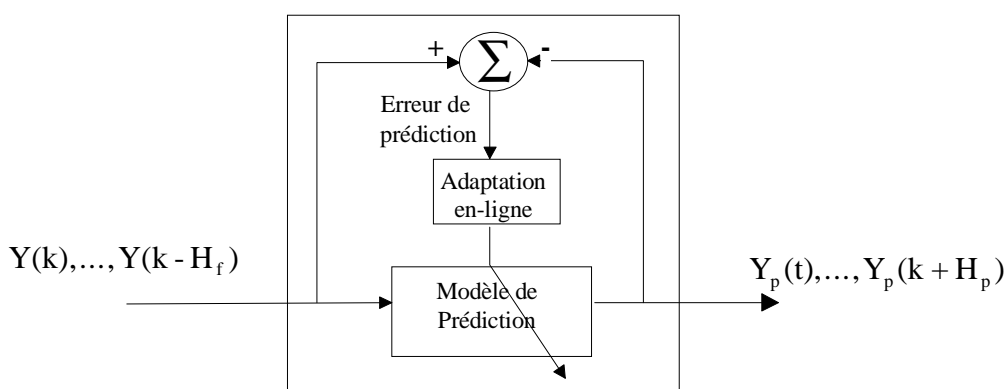


Figure 24 : Principe d'ajustement du modèle prévisionnel

### Résultats d'expérimentations

Cette méthode de prédiction basée sur l'utilisation de la théorie des réseaux de neurones permet d'intégrer des variables exogènes (mesures de comptage, temps de parcours et historiques) dans un modèle adapté aux conditions du trafic routier urbain et périurbain. Les études menées ces dernières années (F. Schettini et Laurent Bréheret, 1996 et 1997) ont permis de dégager les modèles les plus adaptés aux conditions de circulation ainsi que de définir les domaines de validité des modèles. Ces études ont été effectuées sur un réseau expérimental à partir de données réelles.



- Un critère global relatif à la prédiction a été calculé sur un jeu de test de 15 jours avec 240 valeurs journalières. Celui-ci fournit l'erreur moyenne relative à la prédiction et à l'observation du  $Q$ , et du  $TO$  à  $k+12\text{ mn}$ ,  $k+24\text{ mn}$  et  $k+36\text{ mn}$  pour un ensemble des PM. De la même manière, le taux moyen d'erreurs est proposé pour le temps de parcours des trajets ( $TP_{\text{trajet}}$ ). Il est déterminé par la différence entre un temps de parcours prédit à  $k_1=k_0+12\text{ mn}$  et estimé à  $k_1$ . Les erreurs de prédiction à  $k+12\text{ mn}$  sont de 9,7 %, 12,1 % et 13,3 % respectivement pour  $Q$ ,  $TO$  et  $TP_{\text{trajet}}$ . Ces résultats sont très satisfaisants compte tenu du fait que 10 PM sur 150 ne disposent pas de base de données historiques. Il est à noter que les performances des prédicteurs se dégradent avec l'horizon de prédiction.

Tableau 14: Evaluation du module de prédiction - Erreur moyenne relative (%) des indicateurs globaux

	Erreur moyenne de prédiction des Q (%)	Erreur moyenne de prédiction des TO (%)	Erreur moyenne de prédiction des TP (%)
<b>k+12 minutes</b>	9,7	12,1	13,3
<b>k+24 minutes</b>	13,4	16,1	17,1
<b>k+36 minutes</b>	18,4	23,8	25,1

- Des indicateurs individuels de performance ont été également calculés. Ces indicateurs sont les erreurs moyennes relatives pour chaque variable ( $Q$ ,  $TO$  et  $TP$ ) et pour chaque pas de prédiction déterminées en fonction des heures considérées ; HPM (Heures de pointe du matin), HPS (Heures de pointe du soir), HC (Heures creuses). L'erreur moyenne relative est calculée pour un ensemble de 10 PM et d'un trajet par comparaison des données prédites pour une tranche de temps future  $k_1 = k_0 + Hp\text{ pas}$  et des valeurs réellement observées ou estimées à la même tranche de temps  $k_1$ . Par exemple, un débit prédit à  $k + 36\text{ minutes}$  où  $k = 8\text{ h }54$  est comparé à la valeur observée à  $k = 9\text{ h }30$ .

Le Tableau 15 détaille les résultats des erreurs relatives obtenues, en fonction des pas de prédiction, ainsi qu'en fonction des différentes périodes de trafic (HPM, HPS et HC).

En première analyse, la prédiction est meilleure durant les périodes les plus chargées (HPM et HPS). Le critère d'erreur concerne l'amplitude relative de l'observation. Il est donc normal que pour une erreur constante, celui-ci décroisse à mesure que les variables augmentent. On observe un taux d'erreurs à  $k+12\text{ mn}$  de l'ordre de 10% pour  $Q$  et  $TO$ . Ce taux est un peu plus important (11,3%) pour la prédiction du  $TP_{\text{trajet}}$ . Pour prédire le temps de parcours sur un trajet, on procède d'abord à la prédiction des états des PM, pour calculer ensuite l'évolution des  $TP_1$  (temps de parcours par chaînon). Enfin, le  $TP_{\text{trajet}}$  prédit résulte d'une agrégation des prédictions relatives à tous les  $TP_1$  associés au trajet. En d'autres termes, l'erreur de prédiction des états des PM se propage et s'amplifie dans la chaîne de traitements. Une solution consisterait à introduire un terme de correction du  $TP_{\text{trajet}}$  prédit pour chaque trajet ; ce terme serait optimisé avec les estimations du  $TP_{\text{trajet}}$  pour la même heure.

Cependant, il convient de tempérer ces remarques car on constate que cette différence reste constante avec l'horizon de prédiction.

Tableau 15 : Evaluation du module de prédiction - Erreur moyenne relative (%) des indicateurs individuels

<b>Variables</b>	<b><i>t + 12mn</i></b>			<b><i>t + 24mn</i></b>			<b><i>t + 36mn</i></b>		
	<b><i>HPM</i></b>	<b><i>HPS</i></b>	<b><i>HC</i></b>	<b><i>HPM</i></b>	<b><i>HPS</i></b>	<b><i>HC</i></b>	<b><i>HPM</i></b>	<b><i>HPS</i></b>	<b><i>HC</i></b>
<b><i>q</i></b>	8.2	8.9	10.3	10.1	11.2	14.9	15.6	16.1	20.3
<b><i>to</i></b>	10.4	11.3	14.2	13.7	14.5	19	18.4	19.7	25.3
<b><i>TP<sub>trajet</sub></i></b>	11.3	11.6	14.8	15.1	16.3	18.4	20.8	22.1	27.3

## 8.7 Description du contenu de la base de données

### 8.7.1 La base de données géographiques

#### Couche I

Tableau 16 : Caractéristiques des champs du fichier des postes de mesure

<i>Champ</i>	<i>Description</i>
1	Identifiant du poste de mesure
2	Nombre de boucles
3	Utilisé pour la micro-régulation
4	Position par rapport au carrefour (m)
5	Libellé

Tableau 17 : Caractéristiques des champs du fichier des carrefours à feux

<i>Champ</i>	<i>Description</i>
1	Identifiant du carrefour
2	Nombre de phases associées
3+	Nom de chaque phase
8	Libellé

#### Couche II

Tableau 18 : Caractéristiques des champs du fichier des chaînons

<i>Champ</i>	<i>Description</i>
1	Identifiant du chaînon
2	Numéro terrain (utilisé seulement pour garder une correspondance avec la base de données initiale)
3	Identifiant du nœud amont
4	Identifiant du nœud aval
5	Point de mesure associé (0= pas de poste de mesure, n°= identifiant du poste de mesure)
6	Longueur (m)
7	Nombre de voies
8	Capacité maximale (véh./h/voies)
9	Vitesse libre (km/h)
10	Vitesse à la saturation (km/h)
11	N° du carrefour à feux associé (0 ou indice table des carrefours à feux)

12	N° de la phase
13	Identifiant du Trajet associé (0 ou indice table des trajets)
14	Voirie (0 ou indice table voirie)
15	Informé (0=non utilisé ou 1=utilisé pour le traitement)
16	Libellé

Tableau 19 : Caractéristiques des champs du fichier des nœuds

<b>Champ</b>	<b>Description</b>
1	Identifiant du nœud
2	Coordonnées en X
3	Coordonnées en Y
4	Identifiant du pôle
5	Libellé

Tableau 20 : Caractéristiques des champs du fichier des Voiries

<b>Champ</b>	<b>Description</b>
1	Identifiant de la Voie
2	Sens (Double, sens unique, Alterné, Interdit)
3	Importance (de 1 à 4)
4	Libellé

Tableau 21 : Caractéristiques des champs du fichier des Interdictions

<b>Champ</b>	<b>Description</b>
1	Identifiant du chaînon provenance
2	Identifiant du chaînon destination
3	Identifiant du nœud
4	Code interdiction (de tourner, de sortie ou d'entrer)

### Couche III

Tableau 22 : Caractéristiques des champs du fichier des trajets

<b>Champ</b>	<b>Description</b>
1	Identifiant du trajet
2	Longueur (m)
3	Nombre de chaînons connectés
4	Libellé

Tableau 23 : Caractéristiques des champs du fichier des pôles

<b>Champ</b>	<b>Description</b>
1	Identifiant du pôle
2	Nombre de nœuds connectés
3	Libellé

### 8.7.2 La base de données historiques

Tableau 24 : Caractéristiques d'un profil historique d'une variable X (débit, taux d'occupation ou vitesse)

<b>Ligne</b>	<b>Champ</b>	<b>Description</b>
1	1	Libellé de la colonne «Heure»
	2	Identifiant du poste de mesure n°1
	...	...
	1+M	Identifiant du poste de mesure n°M
2	1	Heure (format hh :mm)
	2	X <sub>1</sub> (t) en veh/h, % ou km/h
	...	....
	1+M	X <sub>M</sub> (t) en veh/h, % ou km/h

Tableau 25 : Caractéristiques d'un profil historique d'une variable TP par trajet

<b>Ligne</b>	<b>Champ</b>	<b>Description</b>
1	1	Libellé de la colonne «Heure»
	2	Identifiant du trajet n°1
	...	...
	1+T	Identifiant du trajet n°T
2	1	Heure (format hh :mm)
	2	TP <sub>1</sub> en secondes
	...	....
	1+T	TP <sub>C</sub> en secondes

### 8.7.3 La base de données dynamiques

Tableau 26 : Caractéristiques des données de comptage (débit, taux et vitesse)

<i>Ligne</i>	<i>Champ</i>	<i>Description</i>
1	1	Horodate (jj/mm/aa hh:mm:ss)
2	1	Identifiant du poste de mesure n°1
	2	Q en véh/h
	3	TO en %
	4	V en km/h
1+M	1	Identifiant du poste de mesure n°M
	2	Q en véh/h
	3	TO en %
	4	V en km/h

Tableau 27 : Caractéristiques des temps de parcours mesurés par chaînon

<i>Ligne</i>	<i>Champ</i>	<i>Description</i>
1	1	Horodate (jj/mm/aa hh:mm:ss)
2	1	Identifiant du chaînon n°1
	2	Temps de parcours (en secondes) ou vitesse (en km/h)
1+C	1	Identifiant du chaînon n°C
	2	Temps de parcours (en secondes)

### 8.7.4 La base de données de sortie

Tableau 28: Caractéristiques du fichier circulaire des données de comptage corrigées

<i>Ligne</i>	<i>Champ</i>	<i>Description</i>
1	1	Horodate (jj/mm/aa hh:mm:ss)
2	1	Identifiant du PM n°1
	2	Débit de véhicules (en véh/h)
	3	Taux d'occupation (%)
	4	Vitesse (en km/h)
	5	Etat du PM (Normal, Corrigé, Alarme)
1+M	1	Identifiant du PM n°M
	2	Débit de véhicules (en véh/h)
	3	Taux d'occupation (%)
	4	Vitesse (en km/h)
	5	Etat du PM (Normal, Corrigé, Alarme)

Tableau 29 : Caractéristiques du fichier circulaire des estimations relatives aux X chaînons

<b>Ligne</b>	<b>Champ</b>	<b>Description</b>
1	1	Horodate (jj/mm/aa hh:mm:ss)
2	1	Identifiant du chaînon n°1
	2	Temps de parcours (en secondes)
	3	Débit de véhicule (véh/h)
	4	Taux d'Occupation (%)
	5	Vitesse (en km/h)
	6	Longueur de la file d'attente (km)
1+X	1	Identifiant du chaînon n°X
	..	....
	6	Longueur de la file d'attente (km)

Tableau 30 : Caractéristiques du fichier circulaire des estimations relatives aux X trajets

<b>Ligne</b>	<b>Champ</b>	<b>Description</b>
1	1	Horodate (jj/mm/aa hh:mm:ss)
2	1	Identifiant du trajet n°1
	2	Temps de parcours (en secondes)
	3	Vitesse (en km/h)
	4	Niveau de trafic
1+X	1	Identifiant du trajet n°X
	2	Temps de parcours (en secondes)
	3	Vitesse (en km/h)
	4	Niveau de trafic

Tableau 31 : Structure des fichiers de sauvegarde (tableau en ligne)

<b>Ligne</b>	<b>Champ 1</b>	<b>Champ 2</b>	<b>Champ 3</b>	<b>...</b>	<b>Champ 1+N</b>	<b>Commentaire</b>
1	Heure	Indice objet n°1	Indice objet n°2	...	Indice objet n°N	Ligne d'en-tête
2	00:00	Valeur	Valeur	...	Valeur	Ligne de données
3	00:15	Valeur	Valeur	...	Valeur	Ligne de données
...	...	...	...	...	...	...
1+96	23:45	Valeur	Valeur	...	Valeur	...

Tableau 32 : Structure des fichiers de sauvegarde (tableau en colonne)

<b>Ligne</b>	<b>Champ 1</b>	<b>Champ 2</b>	<b>Champ 3</b>	<b>...</b>	<b>Champ 1+96</b>	<b>Commentaire</b>
1	Objet	00:00	00:15	...	23:55	Ligne d'en-tête
2	Indice objet n°1	Valeur	Valeur	...	Valeur	Ligne de données pour l'objet 1
3	Indice objet n°2	Valeur	Valeur	...	Valeur	Ligne de données
...	...	...	...	...	...	...
1+N	Indice objet n°N	Valeur	Valeur	...	Valeur	...



## 9 Références

- Abe A. (1997) The correction of the forecasting travel time by using AVI data- IVHS 97
- Barbier M., J.L. Farges (1997). Flow and Travel Time Fusion in Bi-directional Route Guidance Systems. *Proceedings of the 8th IFAC Symposium on Transportation Systems*, Vol 1, pp. 421-426.
- Beccaria G., P. Bonsall, N. Hounsell, B. Stoneman, M. Barbier, M. Bell and J. Janko (1995). MARGOT book, The Contribution of Basic Research to the Assessment of Route Guidance Systems. *V2033 DRIVE Project*.
- Bonvallet, F., Y. Robin-Prévallée. Comprehensive network traffic indicators based on centralized data from an operating system. *English Issue No 6 of Recherche Transports Sécurité*.
- Bosq D. et Lecoutre J.P.(1992). Analyse et prévision des séries chronologiques, Méthodes paramétriques et non paramétriques, *Masson (Ed.)*.
- Bourret P, Reggia J. et Samuelides M. (1991) Réseaux neuronaux. *Teknea (Ed.)*.
- Box G. et Jenkins G. (1970) Times Series Analysis, Forecasting and Control. *Holden Day (Ed.)*.
- Breheret L. et Schettini F. (1997) Cleopatra : Application of Traffic Monitorig in Lyon. *Proceeding of 4<sup>th</sup> World Congress on Intelligent Traffic Systems*.
- Breheret L. et Schettini F. (1998) TITUS : Expérimentation d'un système d'information trafic dans l'agglomération de Toulouse. *Recueil de communication du séminaire des Télématiques des déplacements urbains*, Toulouse.
- Choi K. (1997) Data fusion for generating the link travel time with insufficient data sources- IVHS 97
- Choi D.B. (1998) On multisensor data fusion using attribute association for intelligent traffic congestion information inference - IVHS 98
- DATA Plus - Etude prospective dans le domaine du recueil des données de trafic – Rapport d'étude u PREDIT 1996-2000 - Collections du Certu - Octobre 1999.
- Dochy T., Danech-Pajouh M. et Lechevallier Y. (1992) Prévision à court terme du trafic routier par réseau de neurones. *Recherche Transport Sécurité n°42*, pp 35-45.
- Di Taranto C., P. Firmin, J. Janko, L. Breheret, F. Schettini, M. Barbier, M. Bell, S. Grosso, N. Hounsell, K. Chatterjee, A. Lofqvist, A. Peterson (1999). Recommendations for future implementations, CLEOPATRA book. *ATT CLEOPATRA TR1012 project*.
- Dougherty M, Kirby H. et Boyle R. (1994) Using neural networks to recognise, predict and model trafic. *Artif. Intell. Applic. to Traffic Engng.*, Bielli M., Ambrosino G. et Boero M. (Ed.), pp 233-250.
- El Faouzi N.E. (1995) Travel time estimation on urban networks from traffic data and on-board trip characteristics- IVHS 95
- El Faouzi N.E. (1996) Nonparametric Traffic Flow Prediction Using Kernel Estimator. *Proceeding of the 13<sup>th</sup> International Symposium on Transportation and Traffic Theory*, pp 41-54.

- EL Faouzi N.E. (1997) Heterogeneous data sources fusion for impedance indicators- IFAC 97- p. 1375- 1380
- Gabard J.F.(1986) Calculation of Fixed Plans. *IFAC Control in Transportation Systems*, Vienna, pp 233-237.
- Gaillet J-F. (1992) Base de données et Traitement d'information dans les PC de circulation routière. *Rapport de DEA Automatique- CERT-DERA*.
- Hadj Salem H. (1997) Real data screening : problem statement, methods and field results, IFAC 97, pp 1369-1374
- Janko J. (1993) Comparison of journey time prediction methods. *Advanced telematics in road transport, Proceedings of the TECHNICAL DAYS*, Vol. 2, project reports, part, pp 172-175.
- Kühne R.D. (1997) Data fusion for DRG systems- IFAC 97- p. 1386-1390
- LLAMD (1994) Improved Journey Time Prediction. *DRIVE II Project V2033*, Deliverable 6014.
- Nutall I. (1998) Someone to watch over me- An eye on surveillance applications- TTI - Déc. 98- Janv. 99
- Rosenblatt (1956) Remark on some non parametric estimates of a density function. *Ann. Math. Statistic n°27*, pp 832-837.
- Sastri T.(1994) Multi-Functional Neural Networks for System Identification. *Neural Networks in design and manufacturing*, World Scientific, pp 185-245.
- Schettini F. et Breheret L. (1996), Recurrent Neural Network as Models for Traffic Forecasting in urban Network. *Proceedings of the 4<sup>th</sup> Euro Working Group on Transportation*, Newcastle.
- Schettini F. Et Breheret L. (1997), Application of Artificial Neural Network in Urban Network. *Proceedings of the 5<sup>th</sup> European Congress on Intelligent Techniques and Soft Computing, EUFIT*, Aachen.
- Schettini F. (1998), Fusion de données pour la surveillance du trafic et l'information des usagers Doctorat en Informatique Industrielle de l'Ecole Nationale Supérieure de l'Aéronautique et de l'Espace de Toulouse.
- Wardrop J.G. (1952) Some theoretical aspects of road traffic research. *Proc. Inst. Civ. Eng.*, Part. II 1, 2 pp. 325-362.
- Webster F.V. (1958) Traffic Signal Settings. Road Research Tech. Paper n°39. *Great Britain Road Research Laboratory*, London.
- Weymann J., J.L. Farges and J.J. Henry (1994). Dynamic Route Guidance with Queue- and Flow-Dependent Travel Time. *Transportation Research C*, Vol. 2 N° 3, pp. 165-183.

## **10 Sigles et acronymes**

ASF : Autoroutes du Sud de la France

CERT-ONERA : Centre d'Étude et de Recherche de Toulouse

CIGT : Centre d'Ingénierie et de Gestion de trafic

CLAIRE : Système de gestion et de supervision de la congestion

DDE : Direction Départementale de l'Équipement

ERATO : Système de gestion du réseau des voies rapides urbaines de Toulouse

GSM : Global System for Mobile communications

INRETS : Institut National de Recherche sur les Transports et leur Sécurité

ITS : Intelligent Transport Systems

JTP : Journey Time Prediction

ONERA : Office National d'Études et de Recherches Aérospatiales

PAU : Poste d'Appel d'Urgence

PM : Poste de Mesure

PMV : Panneau à Message Variable

RAU : Réseau d'Appel d'Urgence

SEMVAT : Société anonyme d'Économie Mixte des transports publics de Voyageurs de l'Agglomération Toulousaine

SGGD : Système de Gestion Globale des Déplacements

SIREDO : Système Informatisé de REcueil de Données

SMTC : Syndicat Mixte des Transports en Commun de l'agglomération toulousaine

ZELT : Zone Expérimentale et Laboratoire de Trafic

## Liste des figures

Figure 1 : Relations entre les acteurs du transport dans l'agglomération toulousaine.....	4
Figure 2 : Couverture géographique des gestionnaires de voirie .....	4
Figure 3: Support physique pour l'information routière .....	17
Figure 4 : Synoptique général du système de traitement des données .....	24
Figure 5 : Schéma bloc fonctionnel du système de traitement de données .....	24
Figure 6 : Relation entre les différents éléments de la base de données géographiques .....	26
Figure 7 : Interaction des fonctions dans le temps .....	28
Figure 8 : Principe de la qualification des données .....	31
Figure 9: Schéma bloc fonctionnel de la qualification des données véhicules traceurs .....	34
Figure 10 : Relations entre variables pour un chaînon urbain.....	37
Figure 11 : Relations entre variables pour un chaînon périurbain .....	38
Figure 12 : Schéma de principe de la fonction de prédiction .....	39
Figure 13 : Méthode de validation de la fonction de filtrage, d'estimation et de prédiction.....	47
Figure 14 : Représentation du diagramme fondamental.....	58
Figure 15 : Détection et correction des données manquantes .....	63
Figure 16 : Détection et correction d'une valeur aberrante (tests 1 et 2).....	64
Figure 17: Détection et correction d'une valeur aberrante (Test 3).....	64
Figure 18 : exemple de distributions multiples .....	65
Figure 19: Exemples de distributions de vitesses.....	66
Figure 20: Evolution de la file d'attente pour les 3 cas fluide, dessaturant et saturé.....	69
Figure 21 : $TP(Q,X)$ pour deux chaînons, l'un avec feux ( $cap = 400$ véh/h), l'autre sans feux .....	69
Figure 22: Courbes $Q(TO)$ : comparaison entre données mesurées (x) et variables filtrées (+).....	74
Figure 23 : Concept de prédiction sur plusieurs pas d'échantillonnage.....	79
Figure 24 : Principe d'ajustement du modèle prévisionnel.....	79

## Liste des tableaux

Tableau 1 : correspondance entre les différentes variables et les types de capteurs (source DATAPLUS) ..... 19

Tableau 2: Données de sortie générées par la fonction de qualification des données pour chaque PM ... 33

Tableau 3 : Scénarios pour le filtrage et l'estimation au niveau chaînon..... 36

Tableau 4 - Variables utilisés dans les modules de prédiction..... 40

Tableau 5 : Structure des noms de fichier de sauvegarde..... 44

Tableau 6: Description du système de recueil de données des réseaux expérimentaux..... 62

Tableau 7 : Evaluation de la qualité des résultats ..... 62

Tableau 8 : Evaluation du module de qualification des données ..... 63

Tableau 9 : Moyennes des écarts des temps de parcours estimés et simulés pour un écoulement fluide suivant le taux d'équipement de GSM (cas d'un axe périphérique seul)..... 67

Tableau 10: Moyennes des écarts des temps de parcours estimés et simulés pour un écoulement lent suivant le taux d'équipement de GSM (cas d'un axe périphérique seul)..... 67

Tableau 11 : Moyennes des écarts des temps de parcours estimés et simulés suivant le taux d'équipement de GSM (cas d'un axe périphérique et d'une route adjacente)..... 67

Tableau 12: Modèles utilisés en urbain et périurbain en présence de données postes de mesure ..... 72

Tableau 13 : Variables et critères à minimiser pour chaque scénario de disponibilité de mesures ..... 73

Tableau 14: Evaluation du module de prédiction - Erreur moyenne relative (%) des indicateurs globaux ..... 80

Tableau 15 : Evaluation du module de prédiction - Erreur moyenne relative (%) des indicateurs individuels ..... 81

Tableau 16 : Caractéristiques des champs du fichier des postes de mesure ..... 82

Tableau 17 :Caractéristiques des champs du fichier des carrefours à feux..... 82

Tableau 18 : Caractéristiques des champs du fichier des chaînons..... 82

Tableau 19 : Caractéristiques des champs du fichier des nœuds..... 83

Tableau 20 : Caractéristiques des champs du fichier des Voiries..... 83

Tableau 21 : Caractéristiques des champs du fichier des Interdictions..... 83

Tableau 22 : Caractéristiques des champs du fichier des trajets..... 83

Tableau 23 : Caractéristiques des champs du fichier des pôles..... 84

Tableau 24 : Caractéristiques d'un profil historique d'une variable X (débit, taux d'occupation ou vitesse)..... 84

Tableau 25 : Caractéristiques d'un profil historique d'une variable TP par trajet..... 84

Tableau 26 : Caractéristiques des données de comptage (débit, taux et vitesse)..... 85

Tableau 27 : Caractéristiques des temps de parcours mesurés par chaînon ..... 85

Tableau 28: Caractéristiques du fichier circulaire des données de comptage corrigées..... 85

Tableau 29 : Caractéristiques du fichier circulaire des estimations relatives aux X chaînons..... 86

Tableau 30 : Caractéristiques du fichier circulaire des estimations relatives aux X trajets..... 86

Tableau 31 : Structure des fichiers de sauvegarde (tableau en ligne) ..... 86

Tableau 32 : Structure des fichiers de sauvegarde (tableau en colonne)..... 87