



HAL
open science

Temps de parcours de référence, Phase A, Conception, réalisation et validation d'un moteur de calcul

Mehdi Danech-Pajouh, Pierrick Joubert, Florent Blondeau

► **To cite this version:**

Mehdi Danech-Pajouh, Pierrick Joubert, Florent Blondeau. Temps de parcours de référence, Phase A, Conception, réalisation et validation d'un moteur de calcul. 2007, 68p. hal-00545846

HAL Id: hal-00545846

<https://hal.science/hal-00545846v1>

Submitted on 13 Dec 2010

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

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



Institut national de recherche
sur les transports et leur sécurité

Mehdi DANECH-PAJOUH
Pierrick JOUBERT
Florent BLONDEAU

TEMPS DE PARCOURS DE RÉFÉRENCE

Phase A

Conception, réalisation et validation d'un moteur de calcul

Convention PREDIM/INRETS
C06-01
Juin 2007



Institut national de recherche
sur les transports et leur sécurité

Mehdi DANECH-PAJOUH

Pierrick JOUBERT

Florent BLONDEAU

TEMPS DE PARCOURS DE RÉFÉRENCE

Phase A

Conception, réalisation et validation d'un moteur de calcul

*Convention PREDIM//INRETS
C06-01
Juin 2007*

Les auteurs :

*Mehdi DANECH-PAJOUH, chargé de recherche à l'INRETS-GRETIA
courriel : mehdi.danech@inrets.fr*

*Pierrick JOUBERT, ingénieur statisticien
Florent BLONDEAU, ingénieur informaticien*

*L'unité de recherche :
Génie des Réseaux de Transport et Informatique Avancée (GRETIA)*

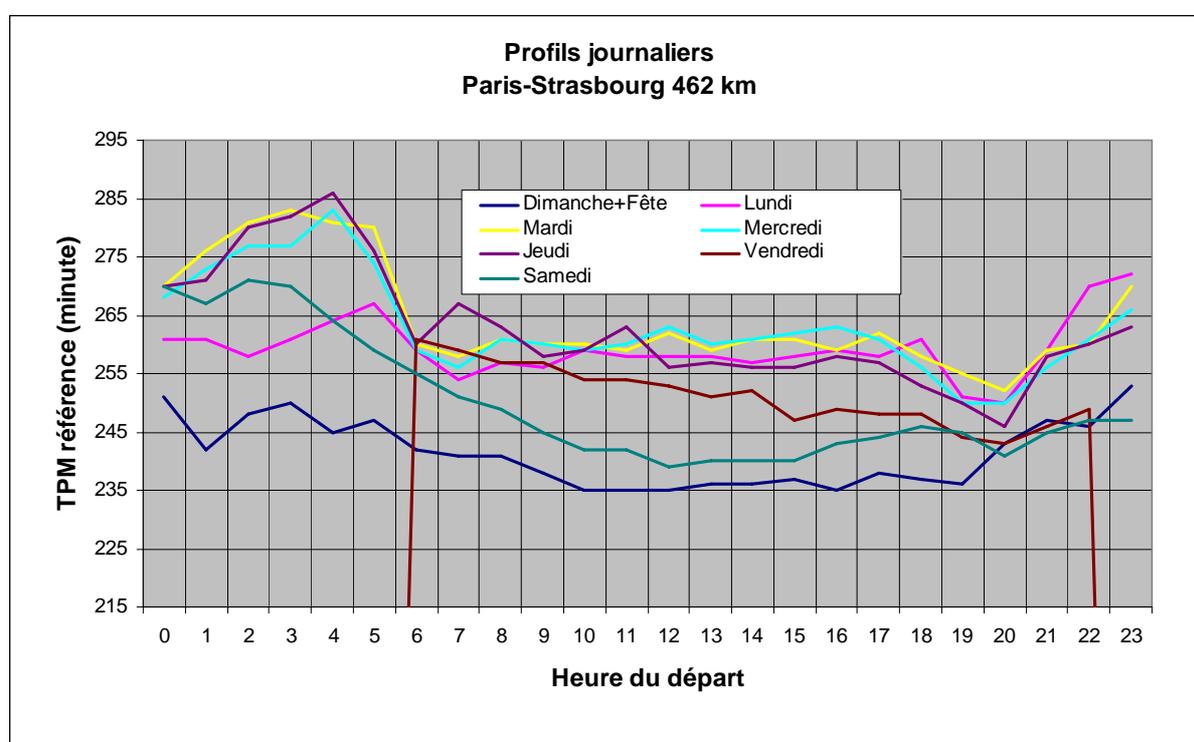
TABLE DES MATIERES

<u>SYNTHESE</u>	<u>4</u>
<u>1. INTRODUCTION.....</u>	<u>7</u>
1.1. LE CONTEXTE.....	7
1.2. LES OBJECTIFS.....	7
1.3. LES DIFFERENTES COMPOSANTES D'UN TEMPS DE PARCOURS.....	8
<u>2. ASPECTS THEORIQUES.....</u>	<u>9</u>
2.1. CONCEPTION THEORIQUE DE L'ALGORITHME	9
2.1.1. DEFINITION D'UNE SECTION DE ROUTE	9
2.1.2. VERSION NON PONDEREE	9
2.1.3. VERSION PONDEREE	11
2.2. RECHERCHE D'UN TEMPS DE PARCOURS DE REFERENCE	12
2.2.1. DEFINITION	12
2.2.2. L'AGREGATION DES \overline{TP} 6MINUTES EN HORAIRE	12
2.2.3. CALCULER UNE COURBE MEDIANE PAR TYPE DE JOUR	12
2.2.4. RECHERCHE UN JOUR MEDIAN PAR TYPE DE JOUR	13
2.3. CRITERE DE VALIDITE	13
2.4. COMMENT INTRODUIRE LES EFFETS D'UNE PERTURBATION ?.....	14
2.4.1. POUR LA VERSION NON PONDEREE.....	14
2.4.2. POUR LA VERSION PONDEREE	15
2.5. COMMENT EVALUER LE TP DES VEHICULES LEGERS (VL) ?	15
2.5.1. ESTIMATION DE LA VITESSE DES VL SUR UNE SECTION DE ROUTE.....	15
2.5.2. ESTIMATION DE TP DES VL SUR UN ITINERAIRE.....	16
<u>3. MISE EN ŒUVRE DE L'ALGORITHME.....</u>	<u>17</u>
3.1. SCHEMA RECAPITULATIF	17
3.2. PARTIE PREALABLE A L'ALGORITHME	18
3.2.1. BASE DE DONNEES	18
3.2.2. EPURATIONS PRIMAIRES.....	19
3.2.3. VERIFICATION DE COHERENCE TEMPORELLE ET SPATIALE.....	20
3.2.4. TRAITEMENT DES DONNEES MANQUANTES.....	21
3.2.5. CALCUL DES PARAMETRES CRITIQUES.....	23
3.3. DESCRIPTION PROCEDURALE DE L'ALGORITHME DE CALCUL DE TPM.....	26
3.3.1. CALCUL DES TEMPS DE PARCOURS 6 MINUTES	26
3.3.2. L'AGREGATION HORAIRE	26
3.3.3. LA NOTION DE PROFILS JOURNALIERS.....	27
3.3.4. PERTURBATION DES RESULTATS	27
3.4. FONCTIONNEMENT DE L'ALGORITHME	29
3.5. EXEMPLE DE L'APPLICATION	32

4. EVALUATION	34
4.1 PERIODE D'EVALUATION.....	35
4.2 LES TRAJETS CHOISIS	35
4.3 TRAJET COUTEVROULT SCHWINDRATZHEIM	36
4.3.1 ANALYSE DES TELEPEAGES.....	36
4.3.2 COMPARAISON AVEC LES TP ESTIMES.....	38
4.3.3 IDENTIFICATION DES JOURS DE REFERENCE	39
4.4 TRAJET TAISSY BEAUMONT	41
4.4.1 ANALYSE DES TELEPEAGES	41
4.4.2 COMPARAISON AVEC LES TP ESTIMES.....	43
4.4.3 IDENTIFICATION DES JOURS DE REFERENCE	44
4.5 TRAJET BEAUMONT ST AVOLD.....	46
4.5.1 ANALYSE DES TELEPEAGES	46
4.5.2 COMPARAISON AVEC LES TPM ESTIMES	48
4.5.3 IDENTIFICATION DES JOURS DE REFERENCE	49
4.6 TRAJET DORMANS TAISSY.....	51
4.6.1 ANALYSE DES TELEPEAGES	51
4.6.2 COMPARAISON AVEC LES TP ESTIMES.....	52
4.6.3 IDENTIFICATION DES JOURS DE REFERENCE	54
4.7 L'ITINERAIRE PARIS – STRASBOURG DANS LES DEUX SENS	55
5. CONCLUSIONS	57
BIBLIOGRAPHIE	59
TABLE DES ANNEXES.....	60

SYNTHESE

Le développement d'une information routière sur les temps de parcours correspond à une nouvelle exigence tant pour les exploitants afin d'améliorer la gestion de leur réseau que pour les usagers de la route afin de mieux appréhender leur déplacement à la fois avant et pendant leur voyage. La mise à disposition de cette information passe par la construction d'un temps de parcours moyen de référence (TPM) qui soit non seulement fonction de la demande de déplacements (type de jour, période horaire) et de l'itinéraire considéré, mais aussi sensible à la présence d'événements susceptibles d'affecter la capacité des routes (tels les chantiers, et les conditions météorologiques).



Un exemple de TPM de référence

Problématique

Les exploitants et les usagers sont tous convaincus que la connaissance des temps de parcours est utile, c'est pourquoi elle occupe une partie importante de l'information routière. Néanmoins il existe une différence importante entre deux notions : le temps de parcours moyen TPM qu'un exploitant peut fournir et le temps de parcours individuel TPI qu'un usager perçoit. Les exploitants qui gèrent les déplacements en masse n'ont pas accès au TPI qui dépend des facteurs individuels (le comportement des conducteurs et la puissance de leur véhicule). Contrairement à ceux qui existent actuellement sur le marché de l'information routière, le TPM n'est pas une notion statique. La problématique est de concevoir des TPM qui soient sensibles à la fois à la demande des déplacements (dans sa dimension temporelle) et à la capacité des routes (dans sa dimension événementielle).

Il existe, actuellement, un important potentiel de recueil de données du trafic sur les axes routiers via plusieurs milliers de capteurs appartenant au Ministère d'Équipement ou faisant partie des systèmes d'exploitations des sociétés d'autoroutes. Ces capteurs de la norme SIREDO fournissent toutes les 6 minutes des mesures agrégées comme le débit, la vitesse, le taux d'occupation et parfois le pourcentage de poids lourds. Le logiciel d'acquisition des données MI² et le réseau Intranet I² du Ministère permettent d'acheminer ces informations vers des serveurs dédiés. Les données recueillies, celles de temps réel et celles formant des historiques permettent de procéder à différentes études du trafic routier comme le calcul de temps de parcours par itinéraire.

Objectifs

Cette recherche appuyée par la PREDIM comporte un double objectif :

- A. Disposer d'un algorithme et d'un moteur de calcul de TPM par itinéraire en utilisant des données historiques des capteurs SIREDO déjà implantés sur les routes.
- B. Etudier la faisabilité d'un système d'information routière sur les temps de parcours moyen et des services associés.

Phase A : L'algorithme TPM

L'objectif principal de cette première partie consiste à fournir aux usagers des informations de temps de parcours par axe pour de grands itinéraires. Il s'agit de fournir des temps de parcours référentiels par itinéraire, par type de jour et par heure de la journée

De plus, le modèle saura estimer le temps de retard dû aux événements exogènes perturbant la capacité des voiries tout au long d'un itinéraire, comme certains chantiers ou les mauvaises conditions météorologiques.

Produits

- Des procédures de traitement des données manquantes et de cohérence des données ont été développées.

Ces fonctions se sont avérées indispensables pour pouvoir disposer d'un maximum de données, et pour que celles-ci puissent être utilisées de manière pertinente.

- Le moteur de calcul dispose actuellement d'un historique de données d'environ un an portant sur 36 capteurs de l'autoroute A4 reliant Paris et Strasbourg. Il comprend deux variantes de calcul, une pondérée et une non pondérée, fonctionnant dans les deux sens de parcours.

L'algorithme a été implémenté sous Java et peut être étendu à d'autres trajets que Paris-Strasbourg, à condition de lui fournir les données des capteurs correspondants au trajet que l'on souhaite étudier. Pour avoir des résultats fiables, il est nécessaire de disposer d'un historique de données assez étendu (1 an minimum).

Evaluation

Dans une première étape, en comparant les TPM obtenus par nos algorithmes et ceux disponibles sur Internet via les fournisseurs actuels d'information routière, on constate que

nous sommes en mesure de produire, avec au moins autant de précision, des temps de parcours référentiel par itinéraire qui de plus varient selon les conditions de circulation liées au type de jour et aux heures de départ. C'est-à-dire des informations semi dynamiques. Des informations pouvant être réajustées selon les perturbations connues de l'état de l'itinéraire. Ainsi un temps de retard est calculé selon la nature de ces perturbations.

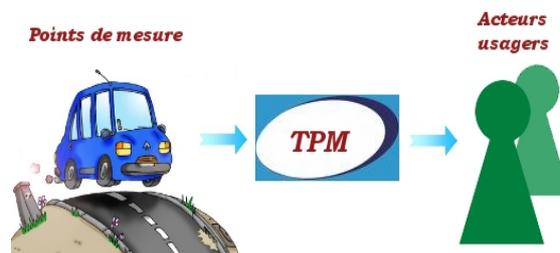
La deuxième facette de l'évaluation a été possible grâce aux données fournies par la SANEF. Il s'agit de données télépéages relatives à une période donnée. C'est-à-dire des temps de trajet individuel (TTI) réalisés par des clients, dans de vraies conditions de circulation, contrairement à nos résultats qui sont une estimation du temps de parcours à partir de données historiques.

Dans cette partie de l'évaluation après avoir tenté de reconstituer des conditions semblables de comparaison, nous avons constatés que les distributions statistiques des TTI et des TPM bâties par itinéraires sont relativement proches.

Les données de télépéages ne peuvent pas remplacer une vraie campagne de mesures de TP via les véhicules flottants mais elles nous ont aidé à confirmer la vraisemblance et la validité des TPM estimés par nos algorithmes.

Phase B : Modélisation du SI TPM

Disposer d'un outil technique de calcul de temps de parcours de référence est le premier pas pour pouvoir diffuser une information de TPM, mais la chaîne qui part du traitement des données brutes pour aller à l'information finale doit être explicitée.



C'est le rôle de la partie B du projet qui comporte plusieurs objectifs :

- Faire un état de l'art des systèmes d'information routière en France pour évaluer les acquis et définir les besoins.
- Définir les services associés à la nouvelle donnée TPM.
- Définir une architecture fonctionnelle cible du système permettant la création et la diffusion d'une information de temps de parcours de référence.
- Préparer la mise en place du système en dégagant des scénarii de déploiement.

La phase B fait l'objet d'un rapport distinct.

1. Introduction

1.1. *Le contexte*

Il existe, actuellement, un important potentiel de recueil de données du trafic sur les axes routiers via plusieurs milliers de capteurs appartenant au Ministère ou faisant partie des systèmes d'exploitations des sociétés d'autoroutes. Ces capteurs de la norme SIREDO fournissent toutes les 6 minutes des mesures agrégées comme le débit, la vitesse, le taux d'occupation et parfois le pourcentage de poids lourds. Le logiciel d'acquisition des données MI2 et le réseau Intranet du Ministère, I² permettent d'acheminer ces informations vers des serveurs dédiés (celui de l'INRETS est baptisé SI²ITAR). Les données recueillies, celles de temps réel et celles formant des historiques permettent de procéder à différentes études du trafic routier dont le calcul de temps de parcours par itinéraire.

Cartographie des capteurs SIREDO

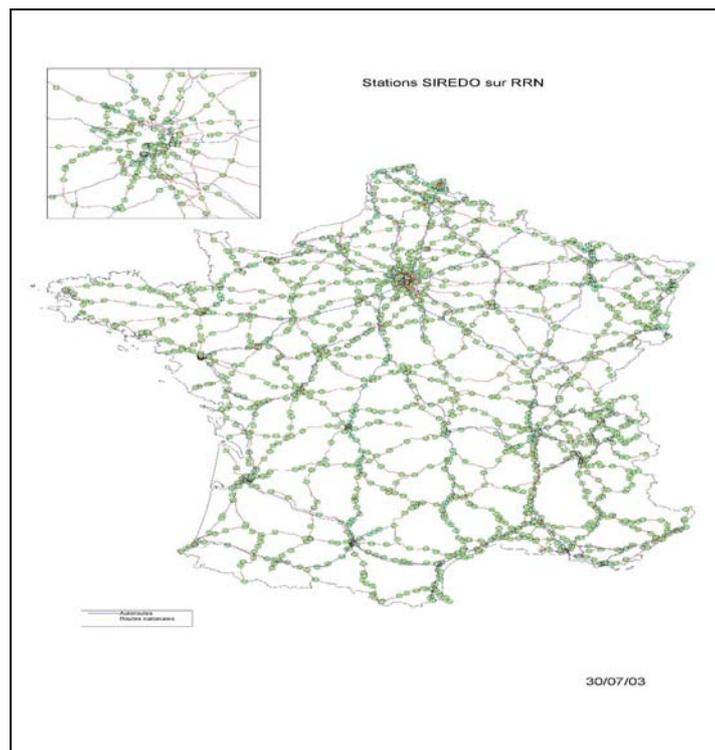


Figure 1 : Stations SIREDO en France

1.2. *Les objectifs*

Cette recherche appuyée par la PREDIM comporte un double objectif :

- A. Disposer d'un algorithme et d'un moteur de calcul de TPM par itinéraire par type de jour et par heure prêt à appliquer sur l'ensemble du réseau routier, partout où il existe suffisamment de stations de comptage.
- B. Etudier la faisabilité d'un système d'information routière sur les temps de parcours moyen et des services associés.

Ce rapport ne traite que nos démarches relatives au premier volet de ces objectifs.

Les exploitants routiers, la DSCR en tête, souhaitent fournir aux usagers des informations relatives aux temps de parcours par axe des itinéraires des grandes mailles. Il s'agit de fournir des temps de parcours référentiels :

- par itinéraire,
- par type de jour,
- et par heure de la journée

Il ne s'agit pas de calculer le temps de parcours moyen d'un itinéraire à partir des données statiques (longueur d'un itinéraire et vitesses autorisées tout au long du trajet), comme on peut en voir chez les différents opérateurs de diffusion d'information routière tels que MAPPY ou Via Michelin. Il s'agit d'un algorithme de calcul de temps de parcours moyen (TPM) sensible aux conditions de circulation. Notre démarche consiste à concevoir un algorithme produisant des temps de parcours de référence, qui ne soient perturbés par aucun événement exogène du type pannes de mesure, accidents, chantiers ou encore mauvaises conditions météorologiques. La particularité de cet algorithme résiderait dans l'introduction de paramètres autres que la vitesse, comme le nombre de voies, la vitesse critique et le débit de saturation par tronçon dans sa formulation. C'est cette particularité qui doit permettre, lors de l'application temps réel, d'actualiser (voire de pénaliser) un TPR relatif à un itinéraire selon les événements exogènes connus ou annoncés.

L'algorithme sera fondé sur les données recueillies à l'aide des capteurs SIREDO existant tout au long des itinéraires. Seules les routes nationales équipées de suffisamment de capteurs feront partie de ces itinéraires.

1.3. Les différentes composantes d'un temps de parcours

Chaque itinéraire « porte à porte » et interurbain se décompose en cinq parties :

- la voirie urbaine associée au point de départ,
- les voies rapides,
- la partie autoroutière de l'itinéraire,
- les voies rapides à l'approche du point de destination,
- et la voirie urbaine s'acheminant vers la destination finale.

La contrainte de l'homogénéité de calcul impose une seule et unique notion de temps de parcours. L'une des difficultés d'un tel calcul est due à ce respect de l'homogénéité imposée par des considérations mathématiques et statistiques. D'après nos connaissances, chaque exploitant routier (urbain ou interurbain) a actuellement sa propre procédure de calcul, et la notion d'interopérabilité n'existe pas encore.

2. Aspects théoriques

2.1. Conception théorique de l'algorithme

2.1.1. Définition d'une section de route

Prenons un itinéraire, situé sur un axe routier, séparant un point A d'un point B. Supposons qu'entre les deux points existent $m+1$ capteurs fournissant des mesures agrégées 6 minutes. Nous considérons comme une section de route l'espace séparant deux capteurs successifs. Le trajet entre A et B est donc composé de m sections. La longueur L_i de la section i s'obtient par la différence entre les PK des capteurs situés aux deux extrémités de cette section. La longueur totale du trajet est L :

$$L = \sum_{i=1}^m L_i \quad (2.1)$$

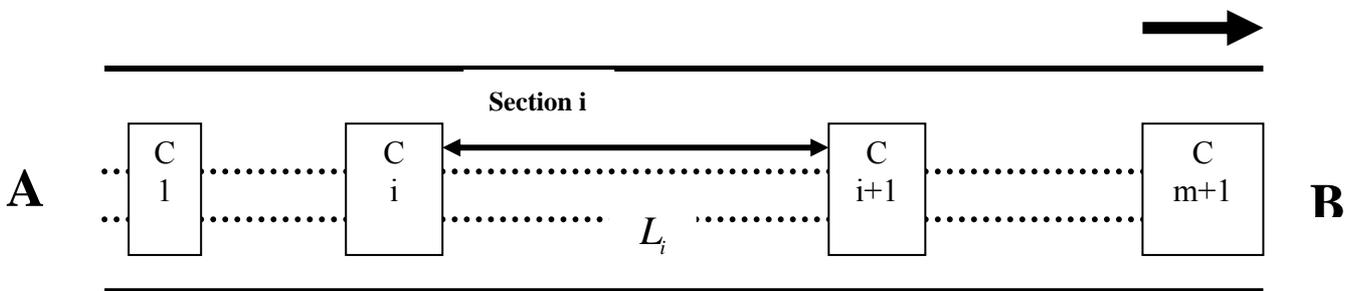


Figure 2 : Schéma d'un axe routier allant du point A vers le point B

Nous distinguons deux approches de calcul du temps de parcours moyen \overline{TP} : le cas le plus fréquent où l'unique variable mise en jeu est la vitesse (version non pondérée) et le cas où la variable du débit intervient comme un facteur de pondération (version pondérée). Dans la partie expérimentation nous débattons des avantages et des inconvénients de ces deux approches.

2.1.2. Version non pondérée

Le temps de parcours d'un itinéraire se déduit par :

$$\overline{TP} = L / \overline{V} \quad (2.2)$$

où \overline{V} est la vitesse moyenne à estimer. Cette estimation se fait à partir des données historiques. Elles sont auparavant classifiées en type de jour et suivant l'heure de la journée. Dans cette estimation pour refléter le dynamisme de l'écoulement du trafic nous discrétisons l'espace (décalage spatial) et les variables sont échantillonnées selon l'avancement des véhicules sur l'itinéraire (décalage temporel).

2.1.2.1 Décalage spatial

Les sections étant de longueurs différentes, nous subdivisons chaque section en sections élémentaires (tronçons) de longueurs ne dépassant pas 1 km.

$$L_i = \sum_{j=1}^{p_i} l_i^j \quad (2.3a)$$

$$\{l_i^j = 1km \quad j = 1, p_i - 1\} \quad (2.3b)$$

$$l_i^{p_i} \leq 1km \quad (2.3c)$$

Où p_i est le nombre de tronçons de la section i .

Pour le calcul de \bar{V} , nous faisons intervenir la vitesse, par section, moyennée sur 6 minutes sans tenir compte du nombre de véhicules (c'est-à-dire un poids unique pour les vitesses) :

$$\bar{V} = \frac{L}{\sum_{i=1}^n \sum_{j=1}^{p_i} l_i^j / \hat{v}_i^j(t_i^j)} \quad (2.4) \quad \overline{TP} = \sum_{i=1}^n \sum_{j=1}^{p_i} l_i^j / \hat{v}_i^j(t_i^j) \quad (2.4 \text{ bis})$$

Où $\hat{v}_i^j(t_i^j)$ est la vitesse estimée du tronçon j de la section i pour le pas de temps t_i^j où les véhicules sont susceptibles d'arriver sur ce tronçon.

En faisant une simple interpolation spatiale (à temps constant) les vitesses des tronçons sont estimées de la manière suivante :

$$\hat{v}_i^j(t_i^j) = V_i(t_i) + (j-1) \frac{V_{i+1}(t_i) - V_i(t_i)}{p_i} \quad (2.5)$$

Où t_i indique le pas de temps d'arrivée des véhicules sur la section i .

2.1.2.2 Décalage temporel

Pour tenir compte du décalage temporel pour chaque tronçon on calcule le temps de parcours partiel, entre le début de la section correspondante jusqu'au tronçon en cours :

$$Tp_i^j = \sum_{k=1}^{j-1} l_i^k / \hat{v}_i^k(t_i^k) \quad (2.6)$$

Si $Tp_i^j < 6$ alors le débit et la vitesse du tronçon en cours se calculent avec les relations (1.5) et (1.6).

Si non, l'indice du temps doit être incrémenté d'un pas de temps :

$$\hat{V}_i^j(t_i^j) = V_i(t_i + 1) + (j - 1) \frac{V_{i+1}(t_i + 1) - V_i(t_i + 1)}{p_i} \quad (2.5 \text{ bis})$$

Ce calcul se poursuit jusqu'à la fin de la dernière section élémentaire située à la fin du trajet.

Remarque : l'usage des données historiques a cet avantage sur les données temps réel que l'on peut échantillonner les variables en tenant compte du décalage temporel. Ainsi, l'échantillonnage des mesures sur chaque section se fait en fonction du temps que les véhicules ont mis pour parcourir la section précédente. C'est cette possibilité de décalage qui permet de calculer un temps de parcours qui ne s'écarte pas trop du réel.

2.1.3. Version pondérée

Ici, le raisonnement porte sur le fait qu'une vitesse mesurée sur un pas de temps est une moyenne des vitesses individuelles de l'ensemble des véhicules rencontrés sur ce même laps de temps, c'est-à-dire le débit. Ce qui distingue cette version de la première, c'est que nous attribuons à chaque mesure de vitesse un poids égal au débit relatif à cette vitesse. Il faut remarquer que si le débit mesuré avait été constant sur tout le parcours, les deux versions aboutiraient au même résultat. Le temps de parcours moyen d'un trajet se calcule toujours de la même manière que dans la première version $\overline{TP} = L / \overline{V}$ (2.2)

Mais, ici, pour calculer \overline{V} , nous faisons intervenir à la fois la vitesse par section moyennée sur 6 minutes et le débit. Il s'agit de calculer la distance totale parcourue par l'ensemble des véhicules observés sur le trajet et l'intégrale du temps passé par ces véhicules. La vitesse moyenne s'obtient par le rapport des ces deux entités.

$$\overline{V} = \frac{\sum_{i=1}^m \sum_{j=1}^{p_i} \hat{q}_i^j(t_i^j) l_i^j}{\sum_{i=1}^n \sum_{j=1}^{p_i} \hat{q}_i^j(t_i^j) l_i^j / \hat{v}_i^j(t_i^j)} \quad (2.7)$$

Le débit du tronçon j de la section i $\hat{q}_i^j(t_i^j)$ est estimé par une simple interpolation linéaire :

$$\hat{q}_i^j(t_i^j) = q_i(t_i) + (j - 1) \frac{q_{i+1}(t_i) - q_i(t_i)}{p_i} \quad (2.8)$$

Sa vitesse $\hat{v}_i^j(t_i^j)$ est estimée comme précédemment avec le même type d'interpolation :

$$\hat{v}_i^j(t_i^j) = V_i(t_i) + (j - 1) \frac{V_{i+1}(t_i) - V_i(t_i)}{p_i} \quad (2.5)$$

Comme dans la première version, afin de s'approcher du temps de parcours réel, nous devons tenir compte du décalage temporel dans le calcul des débits et vitesses des tronçons et des sections. Nous appliquons donc une procédure semblable à celle expliquée plus haut (2.1.2.2).

2.2. Recherche d'un temps de parcours de référence

2.2.1. Définition

Un temps de parcours de référence relatif à un itinéraire, un type de jour et une heure de la journée (l'heure de départ) est pris égale à la médiane de la distribution des TP calculés à partir de l'historique. La médiane par rapport à la moyenne a cette particularité d'écarter les valeurs extrêmes ; celles affectées par des événements exogènes (travaux, intempérie, accident ou incident).

Pour identifier des TP de référence, nous procédons d'abord à une agrégation des TP 6minutes en horaires et ensuite nous constituons pour chaque jour de l'historique une courbe horaire. Le jour type de référence est le jour dont sa courbe de TP est le plus près de la courbe médiane.

2.2.2. L'agrégation des \overline{TP} 6minutes en horaire

Les deux versions ci-dessus donnent des TP moyen \overline{TP} par pas de 6 minutes, ainsi \overline{TP}_j^{hm} est le \overline{TP} au pas m de l'heure h du jour j. L'agrégation horaire des \overline{TP} peut s'obtenir via le calcul de la médiane par tranche horaire.

$$\widetilde{TP}_j^h = \text{médiane}(TP_j^{hm} \mid m = 1, n_j^h) \quad (2.2)$$

Il est possible que les dix pas des \overline{TP} relatifs à une heure donnée ne soient pas tous disponibles, la médiane sera alors calculée selon le nombre de pas disponibles dans l'historique. Le paramètre n_j^h indiquant le nombre de pas de 6 minutes disponibles pour le jour j et l'heure h, varie entre 1 et 10 (nous écartons le cas où aucun \overline{TP} n'a pu être calculé).

Choix d'une période utile :

Nous nous intéressons à la période utile de la journée P_u pour laquelle la connaissance du temps de parcours est importante. Elle est délimitée par l_1 et l_2 : exemple 6h-21h

L'effectif des jours du type s dans la base est N_s où l'indice s variant de 1 à 7 désigne les sept jours de la semaine. Les jours fériés et les dimanches sont classés ensemble. Il est bien entendu que si l'historique est suffisamment important cette typologie peut être étendue à plus de classes (voir le dispositif Bison Futé). Les jours intégrés dans ce calcul sont ceux pour lesquels les \overline{TP} sont disponibles pendant toute la période P_u .

2.2.3. Calculer une courbe médiane par type de jour

La courbe médiane par type de jour \widetilde{M}_s se calcule de la façon suivante :

$$\widetilde{M}_s^h = \text{médiane}(\widetilde{TP}_j^h | j = 1, N_s) \quad (2.3)$$

Il est évident qu'une courbe médiane ainsi calculée ne peut être, forcément, un jour réel, or ce qui nous intéresse est le choix d'un jour médian, ceci permet d'avoir la trace de ses données élémentaires de la base. Nous cherchons, donc, à partir de l'ensemble des courbes journalières des jours du même type, le jour le plus proche de la courbe médiane.

2.2.4. Recherche un jour médian par type de jour

Nous définissons la distance entre la courbe médiane et la courbe d'un jour j de la manière suivante :

$$d_j^m = \text{Max} \left[\text{abs}(\widetilde{TP}_j^h - \widetilde{M}_s^h) | h = l_1 \dots l_2 \right] \quad (2.4)$$

Il s'agit de trouver pour chaque jour et à l'intérieur de la période utile la différence maximale (en valeur absolue) qui sépare la courbe du jour à la courbe médiane. Le jour médian (il y en a un par type de jour) est considéré comme le jour où cette différence est la plus petite :

$$d_s^m = \text{Min} \left[d_j^m | j = 1, N_s \right] \quad (2.5)$$

S'il y a plusieurs jours à d_s^m identique, le jour médian sera le jour où la quantité suivante calculée pour les périodes « hors utile » est minimale.

$$d_j^{+m} = \text{Max} \left[\text{abs}(\widetilde{TP}_j^h - \widetilde{M}_s^h) | h = 1 \dots l_1 \text{ et } l_2 \dots 24 \right] \quad (2.6)$$

Ainsi tandis que la courbe médiane est souvent un jour fictif, le jour médian est un jour de l'historique du type s le plus proche de sa courbe médiane. C'est ce jour qui sera considéré comme le jour de référence pour sa catégorie et le temps de parcours de chaque heure de ce jour est considéré comme une référence.

2.3. Critère de validité

Ce critère est construit par type de jour et par heure. C'est à l'aide du premier et du neuvième décile que nous évaluons la dispersion autour de la médiane (le cinquième décile).

Le premier décile :

$$D1_s^h = \text{premier_décile}(\widetilde{TP}_j^h | j = 1, N_s) \quad (2.7)$$

Le cinquième décile

$$D5_s^h = \widetilde{M}_s^h = \text{médiane}(\widetilde{TP}_j^h | j = 1, N_s) \quad (2.8)$$

Le neuvième décile :

$$D9_s^h = \text{neuvième_décile}(\widetilde{TP}_j^h | j = 1, N_s) \quad (2.9)$$

Si la population à étudier (ici les TP horaires d'un type de jour) suivait une distribution gaussienne dans ce cas les deux déciles $D1_s^h$ et $D9_s^h$ seraient à égale distance de la médiane. Nous n'espérons pas avoir une distribution normale mais les TP calculés seront qualifiés selon les deux critères suivants :

Critère un :

$E_5^1 = (D5 - D1) / D5$ Correspond à 40 % des cas ayant une valeur inférieure à la médiane

Critère deux :

$E_9^5 = (D9 - D5) / D5$ Correspond à 40 % des cas ayant une valeur supérieure à la médiane.

Etant donnée la précision des mesures faites par les capteurs, nous pouvons admettre que ces deux critères ne doivent pas dépasser 10%, sinon le TP calculé est invalide.

L'intervalle entre les deux critères couvre 80% des observations.

2.4. Comment introduire les effets d'une perturbation ?

Toute perturbation peut s'exprimer soit par l'augmentation de la demande (annonce via les médias du beau temps, grève du transport collectif, manifestation sportive ou politique...) soit par la réduction de la capacité (chantiers, météo défavorable...). Dans ce qui suit nous nous intéressons uniquement à ce deuxième type de perturbations. Dans cette partie de calcul, nous avons besoin de connaître les caractéristiques de chacune des sections. A savoir, le nombre de voies, la vitesse critique et la capacité.

On suppose pouvoir quantifier la capacité réduite d'une section de route suite à un évènement exogène comme un chantier. La capacité d'une section i est noté par Q_s^i , sa vitesse critique par V_c^i et sa capacité réduite par Q_r^i . De plus nous attribuons à chaque section une vitesse minimale (par exemple 5km/h) qui sera notée par V_m^i

La baisse de la capacité d'une section de route fait baisser la vitesse d'écoulement. L'estimation de la vitesse réduite de chaque tronçon dépend de la version de calcul du temps de parcours.

2.4.1. Pour la version non pondérée

Si $\hat{q}_i^j(t_i^j) < Q_r^i$ alors la vitesse estimée $\hat{v}_i^j(t_i^j)$ ne changera pas.

Si $\hat{q}_i^j(t_i^j) \geq Q_r^i$ on prévoit deux cas :

Premier cas $v_i^j(t_i^j) < v_c^i$.

Chaque tronçon sera scindé en deux parties avec deux vitesses différentes.

Sur la distance $\frac{Q_r^i}{\hat{q}_i^j(t_i^j)} l_i^j$ les véhicules rouleront avec la vitesse $\hat{v}_i^j(t_i^j)$.

Sur la distance restante $\frac{\hat{q}_i^j(t_i^j) - Q_r^i}{\hat{q}_i^j(t_i^j)} l_i^j$ les véhicules rouleront avec la vitesse V_m^i .

Deuxième cas $\hat{v}_i^j(t_i^j) \geq v_c^i$ idem, le tronçon concerné sera divisé en deux parties.

Sur la distance $\frac{Q_r^i}{\hat{q}_i^j(t_i^j)} l_i^j$ véhicules rouleront avec la vitesse V_c^i

Sur la distance restante $\frac{\hat{q}_i^j(t_i^j) - Q_r^i}{\hat{q}_i^j(t_i^j)} l_i^j$ véhicules rouleront avec la vitesse V_m^i .

2.4.2. Pour la version pondérée

Si $\hat{q}_i^j(t_i^j) < Q_r^i$ alors la vitesse estimée $\hat{v}_i^j(t_i^j)$ ne changera pas.

Si $\hat{q}_i^j(t_i^j) \geq Q_r^i$ on prévoit deux cas :

Premier cas $\hat{v}_i^j(t_i^j) < v_c^i$ alors Q_r^i véhicules rouleront avec la vitesse $\hat{v}_i^j(t_i^j)$
et $\hat{q}_i^j(t_i^j) - Q_r^i$ véhicules rouleront avec la vitesse V_m^i .

Deuxième cas $\hat{v}_i^j(t_i^j) \geq v_c^i$ alors Q_r^i véhicules rouleront avec la vitesse V_c^i
et $\hat{q}_i^j(t_i^j) - Q_r^i$ véhicules rouleront avec la vitesse V_m^i .

2.5. Comment évaluer le TP des véhicules légers (VL) ?

2.5.1. Estimation de la vitesse des VL sur une section de route

Les mesures enregistrées, toutes les 6 minutes, sur des stations SIREDO telles que nous les recevons via MI2 sont par sens de circulation et ne distinguent pas le type de véhicule. Par ailleurs, il existe des enregistrements horaires venant d'autres stations qui permettent d'accéder à des pourcentages horaires de poids lourds (PL). L'idée est de se servir de cette information quantitative pour ajuster le TP des véhicules légers VL.

Pour simplifier les calculs nous faisons l'hypothèse que V_i la vitesse attribuée à un pas de 6 minutes d'une section i est obtenue par la moyenne harmonique des vitesses individuelles de q_i véhicules (tout type confondu) observées pendant ce même laps de temps.

Soit α_i le pourcentage horaire de PL de la section i , la vitesse moyenne des $\alpha_i q_i$ PL rencontrés sur la section est désignée par v_i^{pl} .

$$1/v_i = \frac{(1 - \alpha_i)q_i(1/v_i^{vl}) + \alpha_i q_i(1/v_i^{pl})}{q_i} \quad (2.10)$$

Où v_i^{vl} est la vitesse moyenne des véhicules légers et on peut déduire que

$$1/v_i = (1 - \alpha_i)(1/v_i^{vl}) + \alpha_i(1/v_i^{pl}) \quad (2.11)$$

$$v_i^{vl} = \frac{1 - \alpha_i}{1/v_i - \alpha_i/v_i^{pl}}$$

Ainsi, en supposant que la vitesse moyenne des PL est connue (par exemple 80km/h), v_i^{vl} est estimée par la relation ci-dessus.

$$\text{Exemple : } v_i = 120 \text{ km/h} \quad v_i^{pl} = 80 \text{ km/h} \quad \alpha_i = 15\%$$

$$\hat{v}_i^{vl} = 132 \text{ km/h}$$

2.5.2. Estimation de TP des VL sur un itinéraire

La procédure la plus simple consiste à faire intervenir le pourcentage de PL à la fin de calculs de TP, ceci quelque soit la variante choisie. On fait l'hypothèse que le pourcentage de PL observé au début d'itinéraire et à l'heure de départ reste le même tout au long de l'itinéraire. Du TP calculé on déduit la vitesse moyenne sur l'itinéraire et ensuite on applique la relation (2.11) pour estimer la vitesse moyenne sur l'itinéraire et une estimation de TP des véhicules légers.

Remarque 1 : Pour introduire un peu plus de précision dans ce calcul, nous pouvons prendre comme pourcentage de PL sur l'itinéraire la moyenne des pourcentages de PL observés sur les sections de l'itinéraire pendant une durée correspondant à un TPM approximatif et statique.

Remarque 2 : Nous pensons que vu notre connaissance imparfaite de la vitesse moyenne des PL, son introduction dans le calcul détaillé de TP (décalage temporel, discrétisation) n'apporterait pas plus de précisions.

3. Mise en œuvre de l'algorithme

3.1. Schéma récapitulatif

Le schéma ci-dessous présente les principales étapes à effectuer, du recueil des données à l'obtention de résultats issus du calcul. Toutes ces étapes vont être expliquées dans ce chapitre.

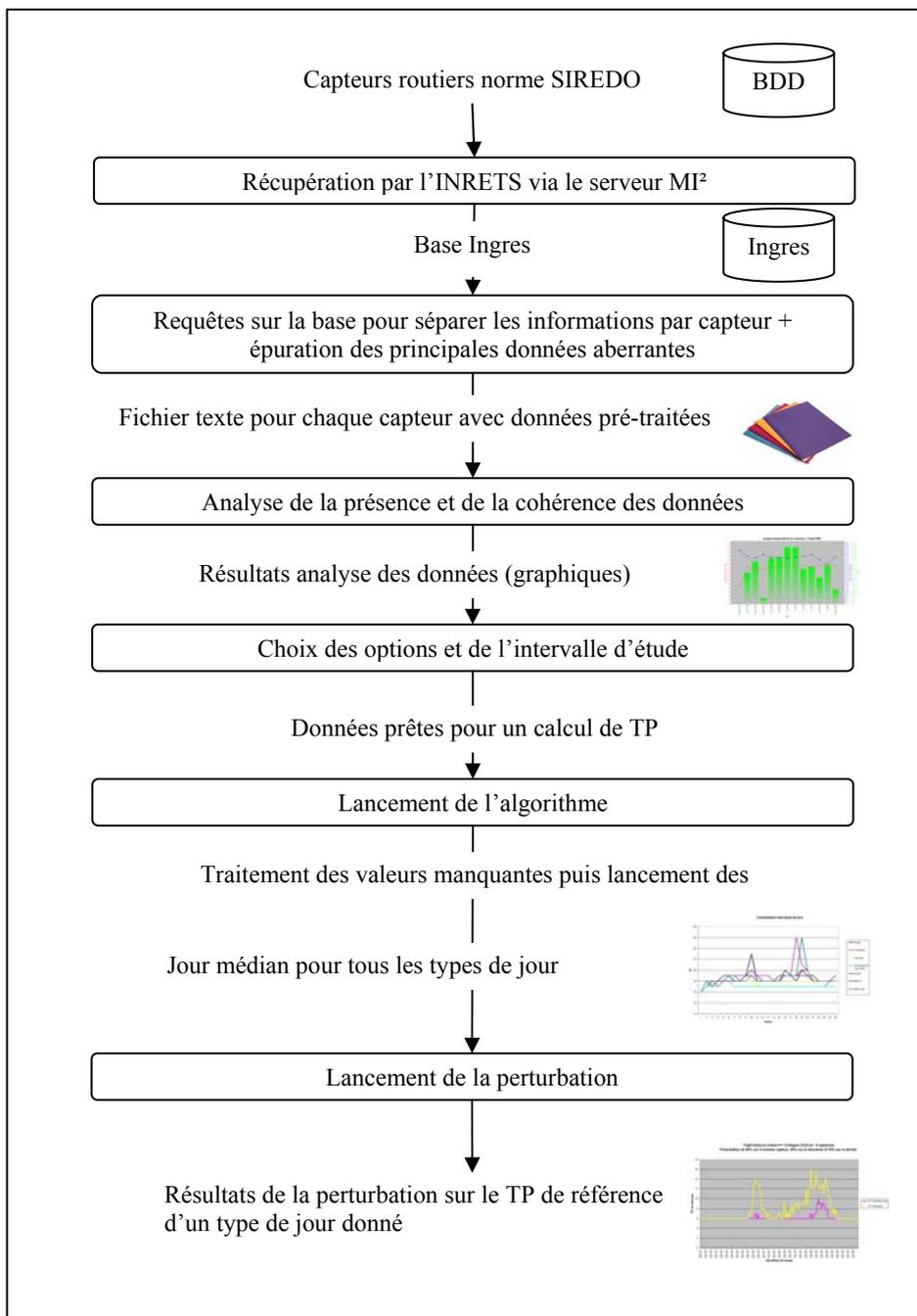


Fig. 3 : le schéma récapitulatif de l'algorithme TPM

3.2. *Partie préalable à l'algorithme*

3.2.1. Base de données

⇒ Quelle est la source des données ?

Les données recueillies sont issues de capteurs bordant les routes et autoroutes (pour les routes, ce sont surtout les grands axes qui disposent de tels dispositifs). Ces capteurs relèvent des informations comme la vitesse, le taux d'occupation et le débit, qui sont souvent agrégées en périodes de 1 ou 6 minutes. Ces données sont datées d'une manière précise (année, mois, jour, heure et pas de 6 minutes). On connaît également la longueur des sections composant le parcours. Nous avons définie (cf. 2.1.1) une section comme l'intervalle entre deux capteurs successifs.

Pour ce qui concerne notre exemple d'expérimentation (cf. 3.5) nous travaillerons avec 36 capteurs basés sur l'autoroute A4 (gérée par la société SANEF) reliant Paris à Strasbourg, qui effectuent des mesures toutes les 6 minutes. L'INRETS récupère les données via le logiciel MI2 et stocke le tout dans une base INGRES. Ainsi, on dispose de données s'étalant d'août 2005 à octobre 2006. Des mises à jour ont été effectuées au cours de l'étude afin d'intégrer de nouvelles données, et pourront l'être encore par la suite.

⇒ Récupération des données

Nous avons créé une macro d'extraction et de préparation des données sous le logiciel SASTM. Cette macro a pour but de stocker les informations dans des fichiers susceptibles d'être appelés ultérieurement par un programme, mais met également en œuvre une première étape de traitement des données.

Nous avons créé une variable au cours de cette macro, appelée *Identifiant*. Elle permet de résumer toutes les informations de date en une seule. Cette variable est numérique. Pour comprendre comment nous avons créé cette variable, nous donnons l'exemple suivant :

50930147 : le premier chiffre (5) correspond à l'année translatée de 2000. Ainsi 5 correspond à 2005. Les deux suivants (09) correspondent au mois, ici septembre. Ensuite on trouve le jour (30) et l'heure (14). Le dernier chiffre correspond au pas de 6 minutes. Ainsi 7 représente le 7^{ème} pas de 6 minutes, ce qui signifie que la mesure a été réalisée à $7*6 = 42$ minutes. Le code 50930147 correspond donc à une mesure effectuée le 30 septembre 2005 à 14h42.

Enfin, cette macro réalise également une première épuration de données. Nous expliciterons toutes les étapes préliminaires au lancement de l'algorithme de calcul de temps de parcours dans la suite.

⇒ Fichiers obtenus (données par capteur)

Le macro d'extraction nous a permis de créer pour chaque capteur et par sens de circulation (sens 1 pour Paris-Strasbourg et sens 2 pour le retour) un fichier contenant les informations relevées toutes les 6 minutes. On trouve donc dans un tel fichier autant de lignes

que d'observations durant la période souhaitée (si une donnée manque pour une certaine date, elle ne figure pas dans la table sous forme de ligne vide, mais est purement et simplement absente) et 5 colonnes comme suit :

- Première colonne : *débit* ou Q (en véhicules/h)
- Deuxième colonne : *taux d'occupation* ou TO
- Troisième colonne : *vitesse* (en km/h)
- Quatrième colonne : *Identifiant* (variable créée comme indiqué précédemment)
- Cinquième colonne : *Jour de la semaine* ou jds (variable créée pour identifier le type de jour)

⇒ Localisation des capteurs

En plus de ces données relevées par les capteurs, on dispose d'autres informations annexes, qui concernent l'itinéraire. Ainsi, on dispose des PK (Points Kilométriques) de chacun des capteurs, qui nous permettent de calculer la longueur de chaque section, ce qui nous est indispensable pour calculer les temps de parcours.

⇒ Etat des données

Nous avons évoqué brièvement que les données, après avoir été récupérées dans la base de données INGRES, avaient subi une première épuration. Cette étape a permis de supprimer les principales données aberrantes. En effet, les informations fournies par les capteurs ne sont pas fiables à 100% dans la mesure où des dysfonctionnements peuvent intervenir tels que des erreurs de mesure ou des arrêts inopinés pour cause de panne par exemple.

Comme nous cherchons à travailler sur un modèle non perturbé, un prétraitement des données s'impose. Pour cette raison, il nous faut préalablement vérifier la cohérence des données. Ainsi, si un capteur fournit des informations suspectes ou si on note une forte différence entre un capteur et ceux qui l'entourent, on enquêtera pour s'assurer de la pertinence de garder ces informations pour le calcul du temps de parcours. Les données conservées après cette étape devront être suffisamment justes pour être exploitées par la suite.

Une fois le champ d'étude (spatial et temporel) défini pour disposer de données non perturbées sur tout le trajet, une procédure d'estimation des valeurs manquantes sera mise en place, afin de pouvoir disposer du maximum d'informations possible.

3.2.2. Epurations primaires

La première opération se fait sur les données récupérées directement dans la base initiale. Elle consiste à éliminer 3 types de données aberrantes.

En premier lieu, on supprime de la base toutes les lignes de données où la somme des 3 variables (vitesse, débit et taux d'occupation) est strictement supérieure à 0, et dont le produit est inférieur ou égal à 0. En résumé, toutes les lignes où ces conditions sont vérifiées seront supprimées :

$$V + Q + TO > 0 \text{ et } V * Q * TO \leq 0$$

On élimine ensuite toutes lignes où tous les triplets où une des 3 variables est manquante. Une telle situation laisse présager un mauvais fonctionnement du capteur, il est en effet peu concevable qu'un capteur ait pu mesurer une ou deux variables sans réussir à mesurer la ou les autres.

Enfin, on élimine également les observations pour lesquelles les valeurs prises par une des trois variables semblent trop élevées. Ainsi, on supprime toutes les lignes suspectes répondant aux critères suivants :

$$V > 250 ; Q > 20000 ; TO > 99.$$

L'opération suivante consiste à supprimer un autre type de données aberrantes, au moyen d'une procédure un peu plus complexe, qui nécessite de tracer un diagramme vitesse*taux d'occupation. Cette étape permet de supprimer des données aberrantes comme une vitesse faible associée à un taux d'occupation très faible, situation difficilement envisageable en réalité. Cette procédure d'épuration est détaillée en annexe A.

Le dernier traitement appliqué aux données consiste en une élimination de doublons afin de s'assurer de l'unicité de chaque pas de temps récupéré.

3.2.3. Vérification de cohérence temporelle et spatiale

A l'issue des 3 étapes d'épuration primaires évoquées ci-dessus, on doit donc disposer de données corrigées des valeurs aberrantes. Il s'agit désormais de fournir à l'opérateur¹ les statistiques globales de disponibilité des données. Ainsi, on doit être en mesure de vérifier, pour chaque capteur, le pourcentage de données disponibles sur une période de temps donnée. Cette option doit amener l'opérateur à **choisir quels capteurs garder pour la suite du calcul**, ce choix devant tenir compte de la quantité d'information disponible pour chaque capteur.

Analyse de la cohérence des données

Afin de fournir un temps de parcours de référence, il nous faut disposer de données non perturbées. Il faut donc pouvoir se donner un moyen de vérifier que les observations à notre disposition sont conformes à notre attente. Ainsi, nous allons effectuer une analyse de la cohérence des données. Cette analyse aura lieu en deux temps. Tout d'abord, une analyse « temporelle » puis une analyse « spatiale ».

Analyse temporelle

Les capteurs à notre disposition fournissent des données de vitesse, débit et taux d'occupation toutes les 6 minutes. Sur une période de temps relativement grande (plus d'un an par exemple), la masse d'informations est trop importante pour une analyse pas à pas. Pour pouvoir se donner une idée des valeurs de vitesse et débit mesurées par les capteurs, nous avons décidé de faire, pour chaque capteur et sur la période utile de la journée, une moyenne par mois de ces deux variables. En représentant graphiquement ces moyennes, on est ainsi capable de repérer d'éventuelles perturbations pouvant affecter l'une ou l'autre des variables, comme une baisse de la vitesse ou du débit, consécutive à des travaux sur une chaussée ou à une réduction de la vitesse sur une portion de route. Cette procédure ne permet de repérer que

¹ Il s'agit de la personne qui fera tourner le programme et non pas de l'utilisateur final

des perturbations récurrentes, mais il serait difficile et très fastidieux de faire ce même travail sur les données 6 minutes.

Analyse spatiale

Cette deuxième analyse ressemble à la précédente mais cette fois le but est de comparer les vitesses et débits des capteurs au cours du trajet. On calculera cette fois les vitesses moyennes et débits moyens mensuels d'un capteur, et on représentera pour chaque mois ces moyennes sur un graphique « spatial » où l'axe des abscisses sera formé des capteurs successifs d'un trajet donné. Ce type de représentation permet de comparer les moyennes des vitesses et des débits entre les capteurs, afin de repérer des perturbations pouvant affecter une section. Cette analyse est complémentaire de la première car une perturbation récurrente sur un capteur au cours de toute la période étudiée se verrait ici, alors qu'elle aurait pu passer inaperçue avec l'analyse temporelle.

A l'issue de ces deux analyses, l'opérateur doit effectuer un nouveau choix quant aux capteurs à conserver ou à la période d'étude. S'il estime qu'un capteur a connu trop de perturbations ou de dysfonctionnements, il peut choisir de l'ôter de l'étude. De même, si une période de temps n'est pas compatible avec la mise en place du calcul, on peut décider de réduire l'intervalle de temps sur lequel on récupère les données. Bien sûr, il faudra veiller, lors de ces choix, à rester dans une configuration permettant d'avoir des résultats fiables, c'est-à-dire garder une période d'étude assez grande pour que le résultat fourni puisse servir de référence. L'historique de données en notre possession ne nous permet pas de fixer un seuil minimal pour ce point. Cependant, il nous semble qu'une période de un an soit un minimum pour avoir une base suffisamment étoffée.

3.2.4. Traitement des données manquantes

Les capteurs peuvent être sujets à des dysfonctionnements, ce qui entraîne des périodes d'absences de mesures. Ces pertes de données ont des conséquences sur le résultat fourni par le moteur de calcul car elles peuvent bloquer le calcul du temps de parcours pour un pas de temps (identifiant) de départ donné. Rappelons que le calcul du temps de parcours s'obtient par sommation sur les sections successives, utilisant les données de chacun des capteurs composant le parcours. Les premiers résultats ont montré la nécessité de traiter une partie des données manquantes. En effet, sans prétraitement approprié, le programme s'avère incapable de fournir suffisamment de données pour que le tout soit exploitable.

Pour essayer de limiter au maximum le nombre de fois où le calcul ne pourra pas aller à son terme, nous avons donc mis en place une procédure de traitement des données manquantes, qui vise à récupérer une information pertinente quand c'est possible.

Deux méthodes ont été choisies pour traiter ce problème. La première consiste en une interpolation temporelle, la seconde en une interpolation spatiale. L'interpolation temporelle se justifie par la corrélation existant entre les données d'un même capteur pour des pas de temps successifs allant jusqu'à 30 minutes². Comme nos données proviennent d'un historique, une valeur manquante pourra se voir attribuer une valeur de son passé ou de son futur, la plus proche possible. Dans le cas où on ne disposerait pas de l'information avec cette méthode, une seconde méthode serait appliquée, qui s'appuierait sur la corrélation spatiale des données.

² Voir DANECH-PAJOUH M., GILLAIZEAU F., PIERRELEE J-C., Amélioration du dispositif bison futé prévision de l'état du trafic à l'horizon d'une demi-heure, Convention DSCR/INRETS (CV04-0006) Rapport final, INRETS, Mars 2006

Cette méthode consiste à calculer un temps de parcours prenant en compte la vitesse de la section $i-1$ quand on a une donnée manquante à la section i . Toutefois, afin de ne pas aboutir à des calculs aberrants qui prendraient par exemple la même valeur pour toutes les sections, nous avons limité l'utilisation de cette interpolation à une distance donnée. Ainsi, la procédure retenue permet d'aller chercher l'information seulement pour le capteur précédent et si la distance ne dépasse pas 40 km (on considère qu'il est hasardeux de prendre l'information d'un capteur qui serait trop éloigné).

A l'issue de cette procédure de traitement des valeurs manquantes, on peut fournir des statistiques montrant le nombre de fois où elle a été utilisée. Ainsi, l'opérateur pourra décider ou non de garder un capteur en fonction du pourcentage de données présentes après le traitement.

Voici un schéma qui résume cette procédure de traitement des valeurs manquantes :

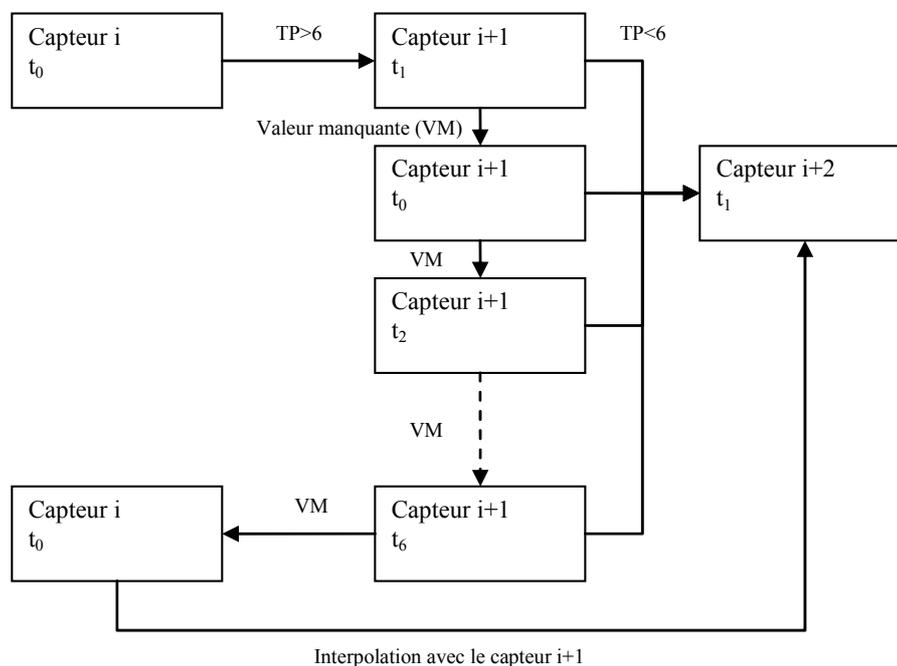


Fig. 4 : Schéma de traitement des valeurs manquantes

Le graphique ci-dessous permet de visualiser le taux de présence de l'information pour un capteur. On peut également constater le gain en pourcentage de données utilisables suite à la procédure de valeur manquante.

Résultats de l'analyse des données des capteurs - Projet TPM

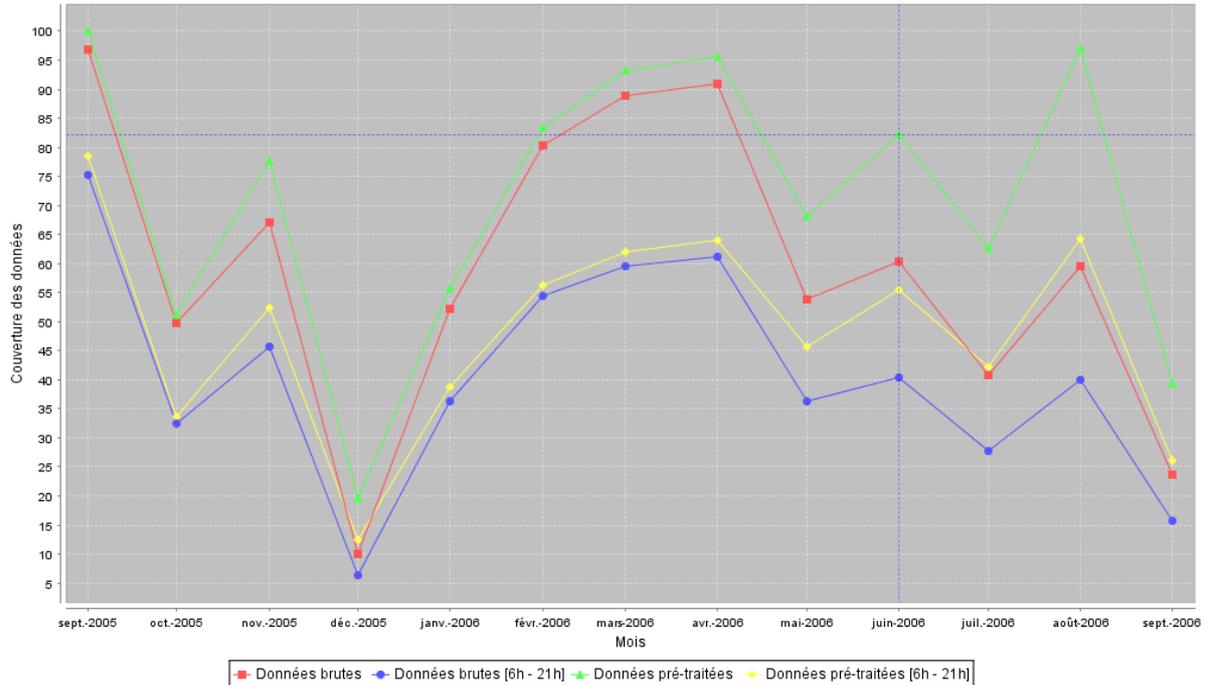


Fig. 5 : Pourcentage de présence des données au cours du temps pour le capteur MFS93A1

Ce graphique est le résultat d'une analyse de données effectuée par notre programme sur le capteur MFS93A1³. Les courbes représentent le pourcentage de données disponibles par mois. Ce pourcentage est calculé comme le nombre des données effectivement présentes par rapport au nombre de données qu'on aurait pu attendre, c'est-à-dire le nombre de pas de 6 minutes dans un mois.

On peut tout d'abord constater que les données brutes concernant la plage utile définie par nos soins comme la période horaire 6h-21h (courbe bleue) sont beaucoup moins nombreuses que celles concernant toute la journée (courbe rouge). Les deux autres courbes (vertes et jaunes) montrent respectivement le pourcentage de données présentes après le prétraitement pour toute la journée, et pour la période utile. On peut ainsi noter un gain de quelques points pour la plupart des mois, pouvant atteindre 25 points pour les meilleurs mois.

Cette procédure nous apparaît donc essentielle car elle permet de disposer d'informations a priori pertinentes mais surtout plus nombreuses. Toutefois, rappelons que l'intervention d'un opérateur est nécessaire à cette étape pour vérifier la cohérence des données recueillies (par exemple, des capteurs ayant fourni peu d'informations pourraient sembler fournir beaucoup d'informations grâce au traitement des valeurs manquantes, mais il convient de s'interroger sur le bien-fondé de garder un tel capteur pour la suite de l'étude)

3.2.5. Calcul des paramètres critiques

Pour pouvoir perturber le modèle, on a besoin de connaître les paramètres critiques des différentes sections composant un parcours. En effet, dans le cas d'une perturbation de type travaux ou la mauvaise météo, on pourra intervenir dans le calcul en modifiant les paramètres

³ Capteur de l'autoroute A4 le plus proche de Paris, situé à Noisy-le-Grand, pour le sens Paris-Strasbourg

caractéristiques d'une ou des sections concernées. Il s'agit de la vitesse critique et le débit de saturation les deux notions classiquement utilisées dans la théorie du trafic routier.

En observant les données à notre disposition dans cette étude, on remarque que l'on doit différencier plusieurs cas. En effet, les données relatives à une section de voirie rapide ne permettent pas systématiquement à une estimation cohérente de la vitesse critique. En particulier, on constate que les capteurs situés près des villes, et principalement ceux se trouvant proches de Paris, se distinguent des autres capteurs quand on trace leurs nuages de points comme par exemple le nuage débit/vitesse. Nous complétons donc la procédure classique d'ajustement d'une courbe de deuxième degré dans le plan débit-vitesse par les trois étapes préalables suivantes :

1. On calcule la variable $k = q/v$ comme un indicateur de la concentration.
2. On fait passer une droite de régression $v = \alpha k + \beta$ dans le plan concentration-vitesse.
3. On calcule le coefficient de détermination de cette droite : R^2

Deux cas se présentent alors :

- Si $R^2 > 0.60$, la **vitesse critique** peut être calculée par l'ajustement d'une courbe de deuxième degré dans le plan débit-vitesse (voir annexe B). Cette situation correspond à un trafic plutôt « urbain » où des congestions se forment assez régulièrement.
- Si le $R^2 \leq 0.60$, tout ajustement de courbe dans le plan débit-vitesse n'aura aucun sens. Nous sommes en présence d'une section sans aucun problème de congestion dans le sens courant du terme. Dans ce cas, à la place de la vitesse critique, nous définissons la notion de **vitesse de confort** qui est prise égale à la médiane de la fonction de répartition des vitesses enregistrées à tous les pas de 6 minutes de l'historique.

Dans les deux cas, le **débit de saturation** sera pris égal à la médiane de tous les débits tels que dans le plan débit-vitesse, les vitesses correspondantes ne s'éloignent pas trop de la vitesse critique.

Voici deux graphiques illustrant la nécessité de traiter différemment les capteurs :

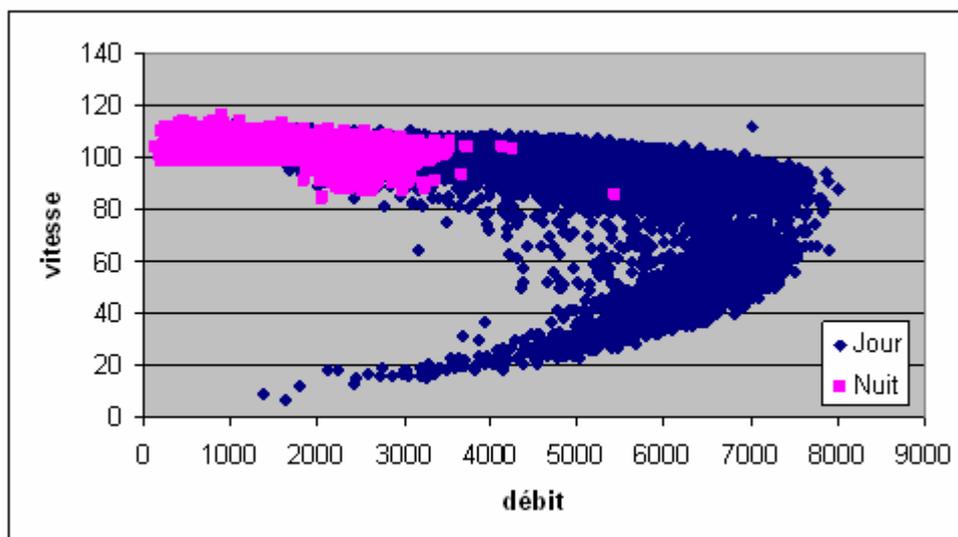


Fig. 6 : Diagramme débit/vitesse - Capteur MFS93A1

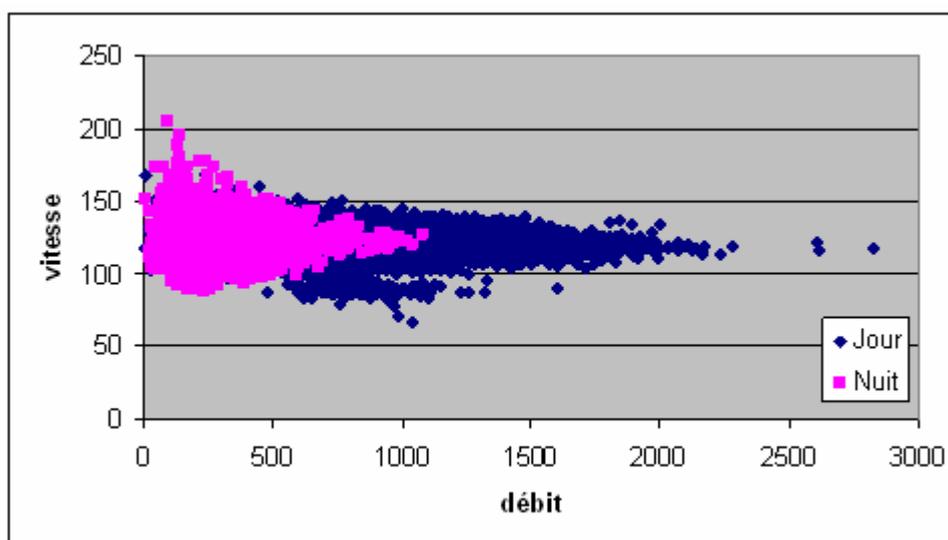


Fig. 7 : Diagramme débit/vitesse - Capteur MZo57.81

Ces deux graphiques sont des diagrammes débit/vitesse réalisés sur deux capteurs différents au cours de la même période de relevé des données. Les points roses représentent les données de nuit (ou période non utile, c'est-à-dire comprise entre 22h et 5h du matin) alors que les bleus correspondent à la tranche horaire utile 6h-21h.

Le premier graphique (MFS93.A1) concerne le capteur le plus proche de Paris à notre disposition. Il est situé en réalité à Noisy-le-Grand. Il présente toutes les caractéristiques d'un capteur « périurbain » avec une courbe bien parabolique permettant d'identifier relativement facilement le débit de saturation et la vitesse critique, même visuellement. En revanche, le deuxième, situé aux Etangs, en Moselle, ne présente pas ce profil classique du capteur urbain. On ne distingue pas de forme parabolique et les données sont beaucoup plus regroupées. Ce deuxième graphique montre bien l'impossibilité de déduire un débit de saturation et une vitesse critique par la procédure habituelle.

3.3. Description procédurale de l'algorithme de calcul de TPM

Avant de chercher un temps de parcours de référence, plusieurs étapes sont nécessaires. En effet, rappelons qu'un temps de parcours de référence n'est pas le résultat d'un calcul mais résulte d'une observation des temps de parcours calculés, puis d'un choix en adéquation avec la définition qu'on se donne d'une référence (voir 2.2.1).

La première étape consiste donc à calculer les temps de parcours moyens par pas de temps de 6 minutes. Il nous faut ensuite chercher la manière la plus pertinente de diffuser une information de temps de parcours. Nous avons choisi d'agréger les TPM ainsi obtenus en données horaires. En effet, la précision du calcul ne permet pas de différencier des résultats obtenus pour un ou deux pas de 6 minutes d'écart. L'agrégation horaire permet donc de résumer l'information en donnant une vision plus synthétique.

Enfin, il faut trouver des représentations pertinentes pour synthétiser l'information, analyser l'évolution du temps de parcours au cours du temps (pour un itinéraire donné) et repérer les intervalles de temps significativement différents. Pour cela, nous représentons les TPM par profils journaliers. L'analyse de ces profils et la possibilité de tracer ceux-ci suivant différents types de jour nous aident à choisir un temps de parcours de référence, qui serait le jour médian. Nous allons détailler toutes ces parties dans la suite.

3.3.1. Calcul des temps de parcours 6 minutes

On rappelle que les données d'origine sont celles relevées par les capteurs, et que chaque résultat de temps de parcours sera le calcul effectué pour un identifiant de départ 6 minutes attribué au premier capteur.

A cette étape du programme les données ont été corrigées. On peut donc entamer le calcul proprement dit. Nous avons défini dans la partie théorique les deux versions de calcul (pondérée et non pondérée). La procédure de calcul applique les formulations telles qu'elles ont été présentées auparavant (2.1.2 et 2.1.3). En revanche, nous précisons un peu plus la notion de décalage temporel évoquée brièvement plus haut.

L'idée de décalage temporel est de tenir compte de l'écoulement du temps au cours d'un parcours. Les conditions de trafic peuvent changer au cours du temps, ce qui fait qu'un départ à une heure donnée d'un point A peut correspondre à une arrivée au point B à une heure problématique. Par exemple, admettons que l'arrivée au point B à 9h du matin soit synonyme de bouchons ou de forts ralentissements. Une telle situation contribue à une hausse de temps de parcours. Un algorithme sans décalage temporel revient à l'échantillonnage simultané des données de tous les capteurs présents sur l'itinéraire. Ainsi, que l'on parte à 9h d'un point A situé à 10 km du point B ou d'un point C situé à 400 km du point B, on prendra pour tous les capteurs la mesure à l'instant $t=9h$. Or en pratique, on se rend bien compte que cette situation n'est pas réaliste pour un itinéraire relativement long (s'il dépasse 30 à 40 km). Cet exemple permet d'illustrer la nécessité de procéder à un décalage temporel dans l'échantillonnage des mesures.

3.3.2. L'agrégation horaire

Pour résumer les temps de parcours obtenus par l'algorithme toutes les 6 minutes, nous avons choisi d'agréger les données par heure. En effet, la précision des mesures et de

notre calcul n'est pas assez forte pour que les variations entre deux pas de six minutes puissent être exploitées.

Cette agrégation a dû s'accompagner de certains choix lors de la programmation. En effet, nous avons vu plus avant que nous ne pouvions pas calculer un temps de parcours pour chaque identifiant de départ en raison de données manquantes. Cela pose un problème lors de l'agrégation, car il faut que les données agrégées soient représentatives des données d'une heure et pas seulement d'un pas. Pour cette raison, nous avons choisi de limiter l'agrégation aux situations où l'on dispose de plus de la moitié des pas de temps dans une heure.

Pour un temps de parcours agrégé à l'heure nous avons choisi comme critère la médiane, plus représentative de l'ensemble des données en raison de la faible dispersion supposée des valeurs au cours d'une même heure.

3.3.3. La notion de profils journaliers

L'intérêt de disposer de résultats fondés sur un historique repose sur la comparaison que l'on peut effectuer entre les temps de parcours calculés à différentes heures et pour différents jours. La première de ces comparaisons qui nous semble la plus intuitive concerne ce qu'on appelle les profils journaliers. Ce terme est assez intuitif et indique que l'on cherche à se représenter l'évolution du temps de parcours d'un itinéraire pour une journée donnée. Une telle représentation peut permettre de ressortir les tranches horaires à problème. On pourra aussi l'utiliser pour comparer des journées entre elles. Par exemple, on peut envisager de représenter tous les profils journaliers d'un type de jour donné (lundi, mardi, ..., dimanche, jour férié) afin de voir si une tendance générale s'affirme.

Afin de disposer du maximum d'informations possible, nous avons restreint une journée à une « période utile », pour laquelle la connaissance du temps de parcours est importante. Nous l'avons délimitée par deux bornes horaires l_1 et l_2 . Pour la suite, nous avons choisi de prendre $l_1 = 6h$ et $l_2 = 21h$.

Pour la construction d'un profil, on collecte tout d'abord tous les jours qui correspondent au type demandé. Ensuite, pour chaque heure, on calcule la médiane correspondante. On aurait pu s'arrêter là et fournir comme profil journalier une courbe dont chacun point soit la médiane de l'heure correspondante. Pour des raisons évoquées en 2.2.2 chaque courbe médiane relative à un type de jour est remplacée par un jour de l'historique du même type dont son profil est le plus proche de la courbe médiane.

3.3.4. Perturbation des résultats

Après avoir lancé un calcul de temps de parcours à l'aide de l'algorithme défini plus haut, on dispose de temps de parcours correspondant à des situations normales de circulation, c'est-à-dire non perturbées. Cependant, on doit pouvoir perturber le système pour tenir compte d'informations supplémentaires dont on disposerait à propos d'un itinéraire donné. Par exemple, si on sait qu'un chantier sera mise en place sur une section particulière, bloquant plusieurs voies de circulation, ou si on prévoit une mauvaise condition météorologique sur une partie d'un itinéraire ; pouvoir intégrer ce type d'information on calcul un temps de retard par rapport à la situation non perturbée (cf. 2.4.)

Notre programme permet de prendre en compte ce type de perturbations. Cependant, il n'y a pas actuellement de prise en compte automatique de tels événements. Ainsi, pour utiliser cette fonctionnalité dans notre programme, il faudra choisir pour chaque capteur sélectionné un pourcentage de réduction de la capacité.

Explications	
Perturbations (les valeurs sont des % de pénalité) :	
MFS93.A1	0.1
MFS93.B1	0.5
MFM77.A1	0.3
MFM77.C1	0.2
MFM77.D1	0.0
MFM77.E1	0.0
MFM77.F1	0.0
MFM77.G1	0.0
MFM77.H1	0.0
MFS77.L1	0.0

Choisissez les perturbations pour chaque capteur de votre sélection.

Fig. 8 : Exemple de choix de perturbation

Cette perturbation s'inscrit dans un objectif d'information usager (qu'il soit utilisateur ou opérateur). Une telle demande doit donc répondre à un besoin précis. Pour cette raison, il nous a semblé logique de demander sur quel type de jour on souhaite voir s'appliquer la perturbation. Ainsi, si on sait qu'une perturbation aura lieu un vendredi, on choisira cette option et les informations délivrées par notre programme ne concerneront que ce type de jour. On pourra relancer un calcul si on souhaite finalement avoir la même information pour un autre type de jour.

Choisissez le type de jour

Dimanche & jour férié

OK Annuler

Fig. 9 : Choix du type de jour pour lancer un calcul perturbé

Nous devons préciser que le calcul en mode perturbé est effectué sur les données du jour de référence relatif au jour choisi. Ainsi, on ne relance pas tout l'algorithme pour fournir un temps de parcours de référence perturbé, mais on perturbe le jour de référence identifié précédemment. Outre l'avantage en terme de temps de calcul que cela comporte, on peut aussi ajouter que ce choix est totalement justifié par le fait que le jour de référence est justement censé faire abstraction des perturbations, et donc représenter au mieux les jours de ce type.

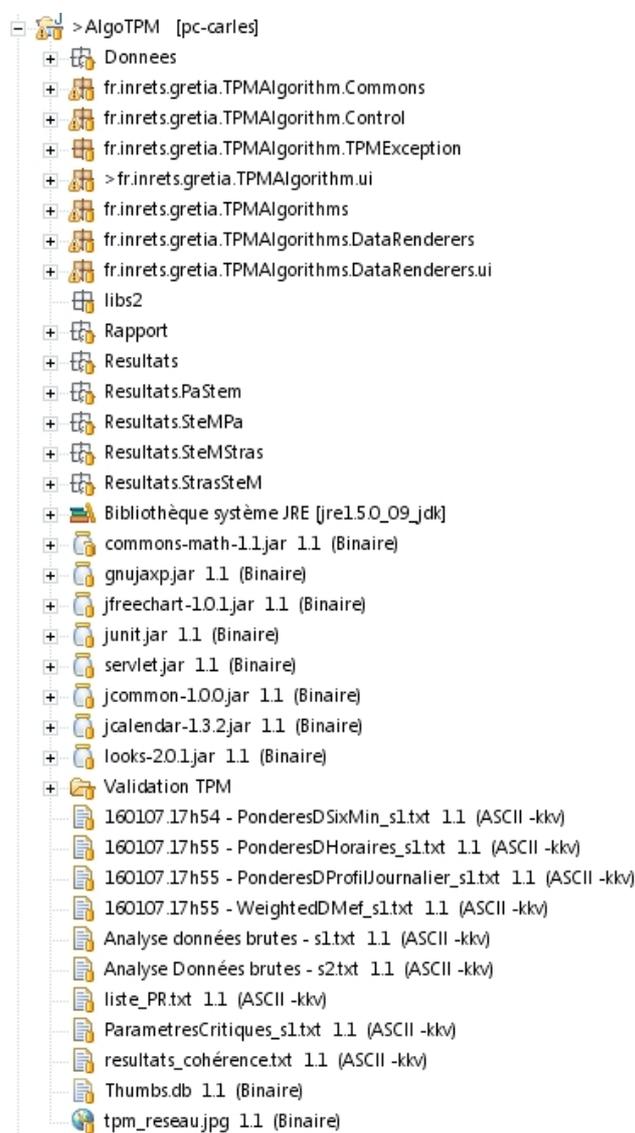
3.4. Fonctionnement de l'algorithme

Nous allons expliquer dans la suite le fonctionnement du prototype de l'application TPM. Cette application a été réalisée pour permettre un calcul de temps de parcours de référence, ainsi que pour permettre l'introduction de perturbations pour calculer un temps de retard par rapport au temps de parcours de référence.

Ce projet a nécessité la récolte de données autoroutières pour être mis en œuvre. Ce sont les données de l'autoroute A4, entre Paris et Strasbourg, qui ont été choisies. Le prototype avait cependant pour but d'être suffisamment générique pour permettre une réutilisation sur d'autres itinéraires. Ainsi, la programmation a été effectuée de telle sorte que les fonctions ne dépendent pas d'un parcours spécifique.

Le projet Eclipse de l'API est disponible sous CVS sur un serveur présent à l'INRETS.

Voici une vue du projet telle qu'on peut l'avoir en ouvrant Eclipse :



Il est composé de 7 packages :

- TPMAAlgorithm.Commons : ce package contient les classes communes aux algorithmes du moteur
- TPMAAlgorithm.Control : le contrôleur de l'interface graphique, là où se trouve la gestion des barres de menu
- TPMAAlgorithm.TPMException : les exceptions spécifiques au projet TPM
- TPMAAlgorithm.ui : les classes de gestion de l'interface graphique
- TPMAAlgorithms : les classes des algorithmes du moteur
- TPMAAlgorithms.DataRenderers : les classes de rendu des résultats des algorithmes
- TPMAAlgorithms.DataRenderers.ui : les classes graphiques du rendu des résultats des algorithmes

La fenêtre s'ouvrant au lancement de l'application ressemble à l'image suivante :

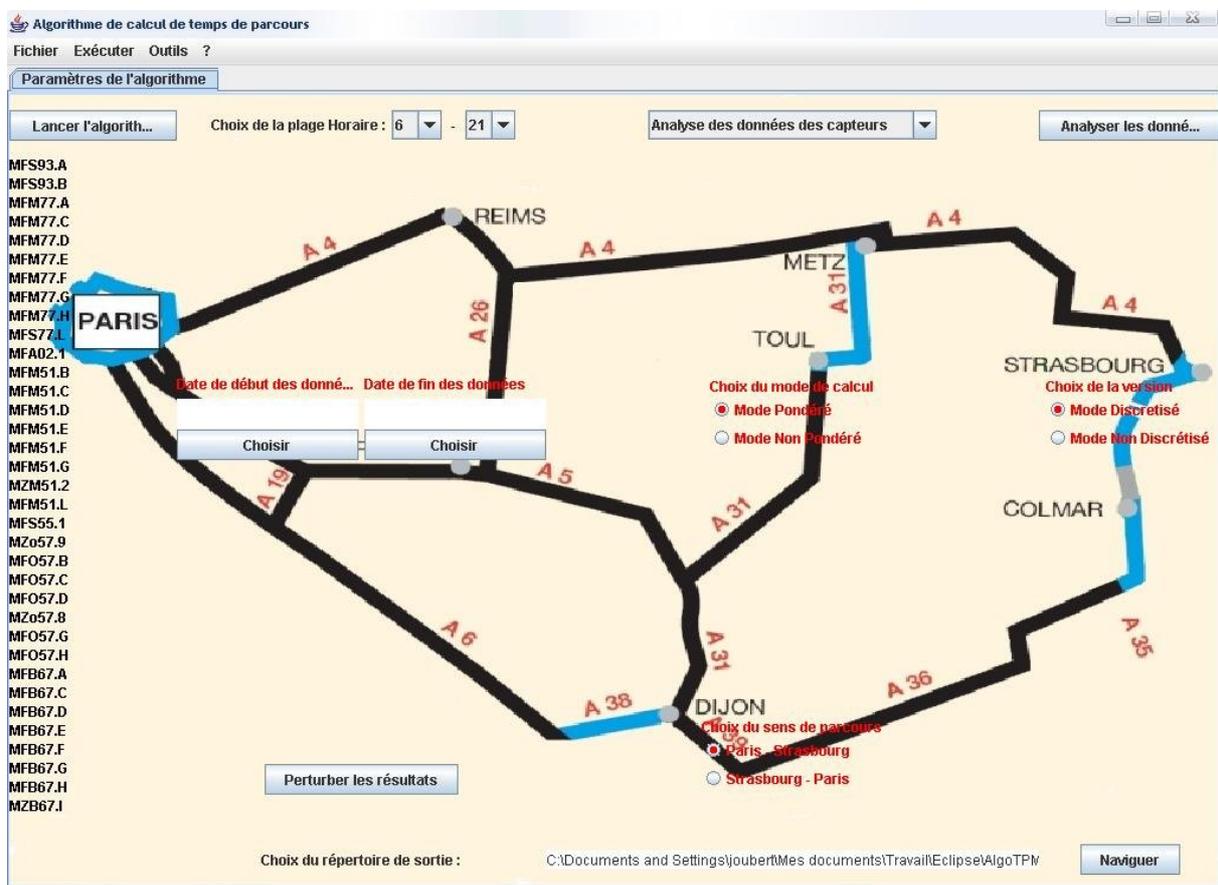


Fig. 10 : fenêtre de lancement de l'application

A partir de cette fenêtre, différentes manipulations peuvent être effectuées pour obtenir des résultats souhaités.

Sélection des capteurs

La première chose à faire est de choisir les capteurs sur lesquels vous voulez travailler. Les codes des capteurs sont situés à gauche (ex : MFS93.A). Pour savoir à quel lieu correspond un capteur, il faudra se référer à un fichier descriptif des capteurs le long de l'A4. La sélection multiple est possible, ainsi que la « suppression » d'un capteur au cours du trajet (dans ce cas le programme fonctionne comme si ce capteur n'existait pas).

Les capteurs sont ordonnés spatialement dans le sens du parcours. Ainsi ceux situés en haut sont les capteurs proches de Paris et ceux du bas correspondent aux capteurs situés à proximité de Strasbourg. Pour faire une simulation dans le sens Strasbourg-Paris, il faudra sélectionner le bouton-radio « *Strasbourg – Paris* » dans le groupe « *Choix du sens de parcours* ». Par défaut, le sens de parcours est réglé sur le sens Paris-Strasbourg.

Choix du répertoire de sortie

Tout en bas de la fenêtre, une barre indique le chemin de sortie par défaut des fichiers résultats créés par le programme. Ce chemin est paramétrable grâce au bouton Naviguer.

Analyse des données

La barre de menu déroulant située en haut à droite ainsi que le bouton « *Analyser les données* » correspondent à la partie d'analyse des données des capteurs. Deux possibilités sont offertes dans le menu déroulant. Cette analyse sera lancée sur les capteurs sélectionnés.

Options pour l'algorithme de calcul de temps de parcours

Enfin, avant de lancer l'algorithme de calcul de temps de parcours, il y a différentes options que l'on peut paramétrer. Tout d'abord, les dates de début et de fin des données. Par défaut, le calcul est réalisé sur toutes les données à disposition, mais un utilisateur peut souhaiter restreindre son étude à une période particulière.

Les options situées dans les groupes de radio-boutons sont les suivants :

- Choix du mode de calcul : permet de choisir la formule de calcul utilisée dans l'algorithme, pondérée ou non. Par défaut c'est la formule pondérée qui est choisie.
- Choix de la version (expérimental) : ce groupe permet de choisir le niveau de précision du calcul au niveau spatial. La version non discrétisée était une étape expérimentale afin de tester l'algorithme. Cette dernière n'est plus utilisée mais néanmoins conservée en cas de nécessité. Par défaut, la version discrétisée est donc sélectionnée.
- Choix du sens de parcours : expliquée auparavant, cette fonctionnalité permet de choisir si on veut que les capteurs choisis soient ceux du sens Paris-Strasbourg ou l'inverse.

Un bouton supplémentaire apparaît à gauche du groupe de bouton concernant le choix du sens de parcours. Ce bouton permet de lancer l'algorithme calculant un retard sur le temps de parcours de référence. Ce bouton ne doit être choisi que si l'on a déjà effectué un calcul de temps de parcours non perturbé (en ayant cliqué sur le bouton « lancer l'algorithme ») car les données prises en compte pour calculer un retard prennent en entrée des informations calculées auparavant. Cette fonctionnalité permet de choisir si on introduit ou non des perturbations dans le calcul (comme les chantiers, la météo...).

Note : actuellement (fin février 2007) cette fonctionnalité permet de choisir un pourcentage de perturbations pour chaque capteur composant le parcours sélectionné. Il faut donc un opérateur pour choisir les perturbations, il n'y a pas de lien avec une base de données événementielle.

Choix de la plage horaire

Les deux listes déroulantes associées à ce groupe permettent de changer la plage horaire « utile ». Cet intervalle de temps a été défini afin de prendre en compte principalement les données comprises dans la plage, horaires considérés comme les plus pertinents dans notre analyse. Cependant, l'utilisateur est libre de changer cet intervalle de temps. Par défaut, nous avons fixé celui-ci à 6h-21h.

3.5. Exemple de l'application

Après avoir fait une analyse de la cohérence de nos données, nous nous sommes aperçus que le capteur situé à Sainte-Ménéhould présentait un problème. En effet, au niveau de ce capteur, on a pu constater des vitesses relativement faibles (par rapport aux sections alentour). Après avoir contacté un expert de la SANEF, il nous est apparu que ces vitesses faibles étaient dues à un dysfonctionnement de la station de mesure correspondant à cette section. Comme nous souhaitons travailler dans un premier temps sur des données non perturbées, nous avons décidé de ne pas prendre en compte ce capteur. Nous lancerons donc notre programme pour les trajets Paris-Ste-Ménéhould et Ste-Ménéhould-Strasbourg.

On peut noter une différence assez importante entre les profils journaliers des lundis selon les 2 versions. La première, discrétisée mais non pondérée a des valeurs beaucoup plus dispersée que la seconde version, qui est discrétisée et pondérée. Toutefois, on constate que les courbes sont assez ressemblantes d'un point de vue global, ce qui est plutôt rassurant. Ainsi on identifie un pic aux alentours de 18h, ce qui peut correspondre à une sortie de Paris difficile. Un autre résultat important qui apparaît sur ce graphique concerne les jours atypiques. En effet, parmi tous les lundis à notre disposition se trouvent quelques jours fériés. Ces jours fériés ont été classés par la suite dans une autre catégorie (avec les dimanches), mais nous nous sommes limités dans un premier temps à une segmentation basique en jours de la semaine. Le fait de les voir ressortir sur ce graphique est très encourageant, car cela montre le bon fonctionnement du moteur de calcul. On peut ainsi repérer le 14 juillet 2006, le premier et le huit mai ainsi que le lundi de Pentecôte. Les profils de ces jours là ressemblent beaucoup plus à ceux des dimanches, ce qui ne semble pas très surprenant.

On peut ainsi analyser les résultats pour tous les types de jour et apercevoir des différences assez marquées entre les versions non pondérées et pondérées (que ça soit pour la version discrétisée ou non). On remarque aussi des pics communs à toutes les courbes, qu'il nous faut analyser plus précisément afin de cerner quelles peuvent en être les causes. Enfin, des pics pour certaines journées seulement mais ne reflétant pas une tendance générale se doivent aussi d'être inspectés.

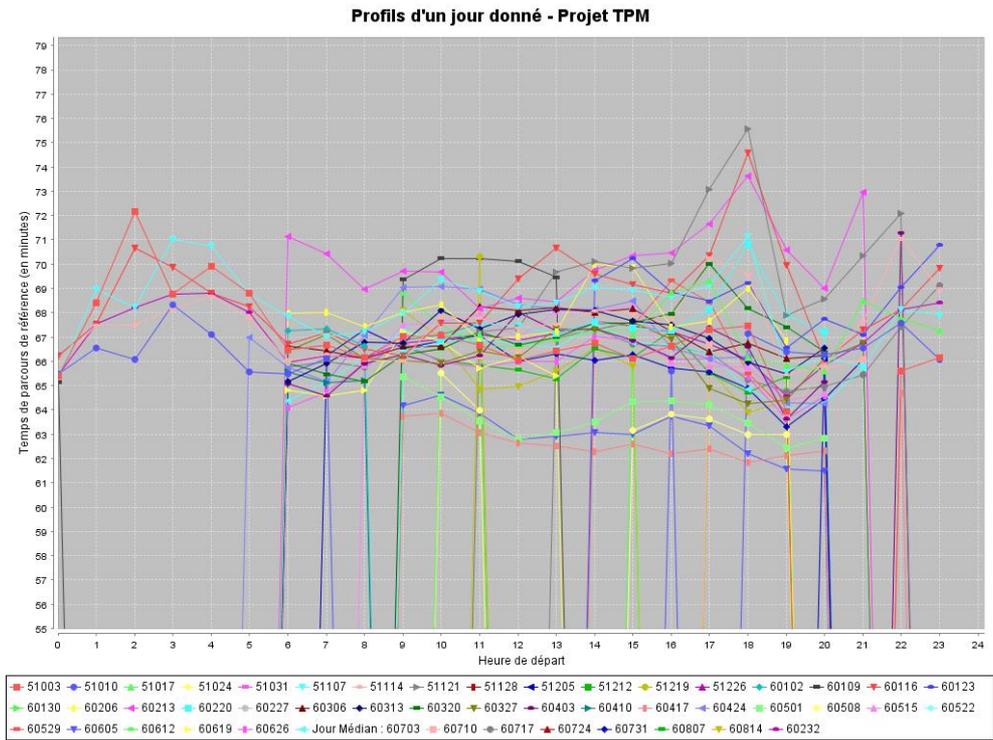


Fig. 11 : Version non pondérée - Discrétisée

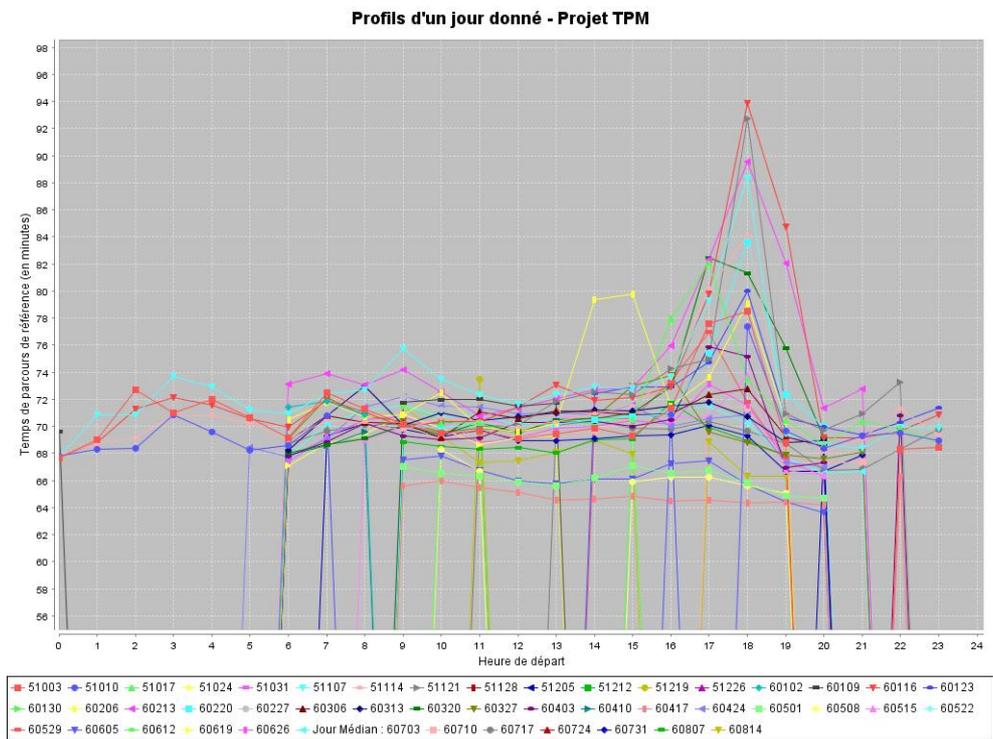


Fig. 12 : Version pondérée discrétisée

4. Evaluation

Cette évaluation consiste à comparer le temps de parcours estimé par moteur de calcul avec une entité proche de la réalité. En principe une telle comparaison se fait via une enquête de véhicules témoins ce qui incombe un coût financier relativement important. C'est pourquoi nous nous sommes adressés à la SANEF⁴ société qui gère l'autoroute A4, en particulier l'axe Paris Strasbourg notre champ d'expérimentation. Cette évaluation tente de comparer les temps de parcours estimés par moteur de calcul avec les temps de trajet découlant des données de transactions télépéages fournies par la SANEF après les avoir rendues anonymes.

Quelques notions importantes

Un temps de trajet individuel TTI est la différence entre l'instant d'arrivée et l'instant de départ. Il comprend, donc, le temps d'arrêt ou de ralentissement (volontaire ou subi). Un temps de parcours individuel TPI est le temps de trajet moins le temps d'arrêt. D'une transaction télépéage se déduit un TTI et non pas un TPI. Des transactions relatives à un trajet on tente de distinguer deux types d'utilisateurs : ceux qui ont été arrêtés ou ralentis sur leur trajet et les autres. Pour faire cette distinction nous faisons l'hypothèse que le temps de trajet sans aucun arrêt (retard) correspond au TTI le plus fréquent, c'est-à-dire au mode de sa distribution \widetilde{TTI} ⁵.

La notion de retard individuel RI intègre à la fois la somme des temps d'arrêts volontaires et le temps dépensé dans les ralentissements. Elle dépend bien sûr de la longueur du parcours mais elle dépend aussi des variables exogènes :

$$RI = f(TTI, \text{ jour, heure, météo, chantier....})$$

$$TPI = TTI - RI$$

L'expression exacte de cette fonction exige des études et des enquêtes appropriées. A défaut, nous sommes conduits à appliquer l'hypothèse ci-dessus donc à formuler une règle très simplifiée :

$$\widehat{RI} = \begin{cases} TTI - \widetilde{TTI} & \text{si } TTI > \widetilde{TTI} \\ 0 & \text{sinon} \end{cases}$$

Remarque importante : Il existe trois différences conceptuelles importantes entre TPM et TTI qui empêchent une comparaison directe entre ces deux quantités. Premièrement, tandis que les TTI déduits des télépéages ne concernent que les temps de trajet individuels des véhicules légers, les TPM estimés à partir des données des capteurs (moteur de calcul) englobent tout type de véhicule (VL+PL). Deuxièmement, le TTI inclut le RI éventuel ce qui est évidemment exclus dans les TPM. Troisièmement, les TPI sont des mesures individuelles et les TPM des mesures agrégées. C'est pourquoi les comparaisons les plus cohérentes se reporteront aux périodes où la circulation de PL est interdite (ou %PL très faible). Par ailleurs

⁴ Nous remercions vivement Monsieur Pierre Florentin de la Direction du développement des grands projets à la SANEF sans qui cette évaluation n'aurait pas été possible.

⁵ Cette hypothèse est suggérée par les exploitants de la SANEF.

il faut procéder à une estimation des TPI à partir des TTI en suivant la règle de calcul de retard mentionnée plus haut.

4.1 Période d'évaluation

La période d'analyse retenue comporte 16 jours consécutifs, du jeudi 3 au vendredi 18 novembre 2005. Elle comporte des typologies assez variées. Elle commence juste après la fin des vacances scolaires de Toussaint (mercredi 02/11/2005) et comporte un long week-end autour du 11 novembre. La semaine du lundi 14/11 au vendredi 18/11 a davantage un profil de jours ouvrés « ordinaires ».

4.2 Les trajets choisis

Dans cette évaluation les temps de parcours de plusieurs trajets ont été examinés, ici nous nous contentons d'en présenter quatre :

- Coutevroult Schwindratzheim (421 km), reliant Paris à Strasbourg
- Taissy Beaumont (151), un système fermé de péage
- Beaumont St Avold (66 km), contournement Nord-Est de Metz
- Dormans Taissy (33 km), traversée urbaine de Reims

Parmi ces quatre trajets Taissy Beaumont est le seul à être couvert par un système de télépéage fermé, ainsi les temps de parcours individuels découlent directement de l'identification des badges. Les autres cas nécessitent un repérage des passages en séquence d'un même numéro de badge.

Pour chacun de ces trajets on analyse d'abord les données de télépéage d'une manière indépendante et ensuite nous procédons à la comparaison entre les temps de parcours estimés et ceux déduits des télépéages.

4.3 Trajet Coutevroult Schwindratzheim

4.3.1 Analyse des télépéages

Pour ce trajet d'une longueur de 421km, les transactions télépéages nous ont permis d'énumérer 1923 VL. La médiane de la distribution des TTI est estimée à 216 minutes (3 heures et 36 minutes). Environ 1% des VL roulent relativement vite. La queue de la distribution côté gauche est courte $E_5^1 = 11\%$ (voir la définition en 2.3). Par contre coté droit cette queue est presque deux fois plus longue $E_9^5 = 19\%$, celle-ci est sans doute due aux arrêts éventuels des véhicules sur un trajet de plus de 400km.

Centile	V km/h	TTI min
1	143	176
10	131	193
25	124	203
Médiane	117	216
75	108	234
90	98	257
95	92	274

Tab. 1 : Distribution de TTI

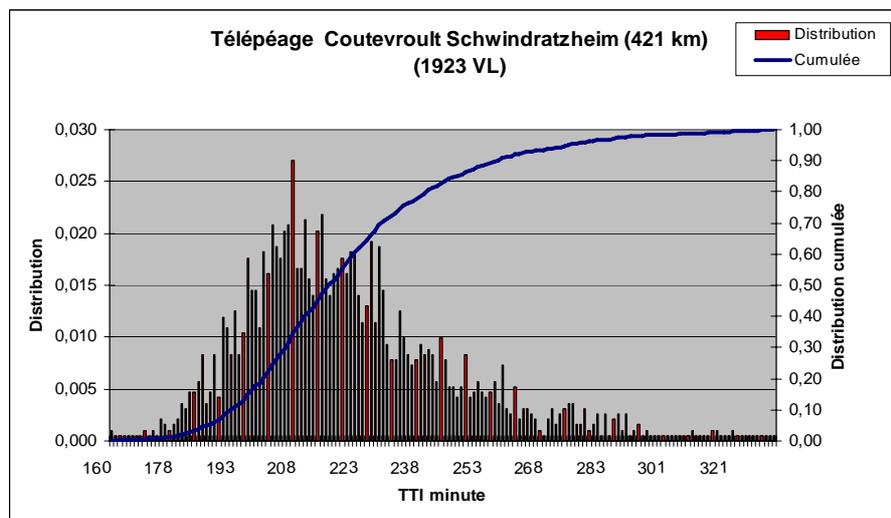


Fig. 13 : histogramme de distribution de TTI

Pour évaluer le temps d'arrêt nous appliquons la règle ci-dessus qui est basée sur l'écart au mode de la distribution, c'est-à-dire 207 minutes (ou 122 km/h). La courbe de pourcentage selon l'heure des VL avec arrêt montre qu'il varie peu et qu'il est presque indépendant de l'heure de départ. Le retard moyen dû à ces arrêts est évalué à 27 minutes. Ce délai nous paraît acceptable vu les conseils de la sécurité routière qui préconisent un arrêt toutes les deux heures.

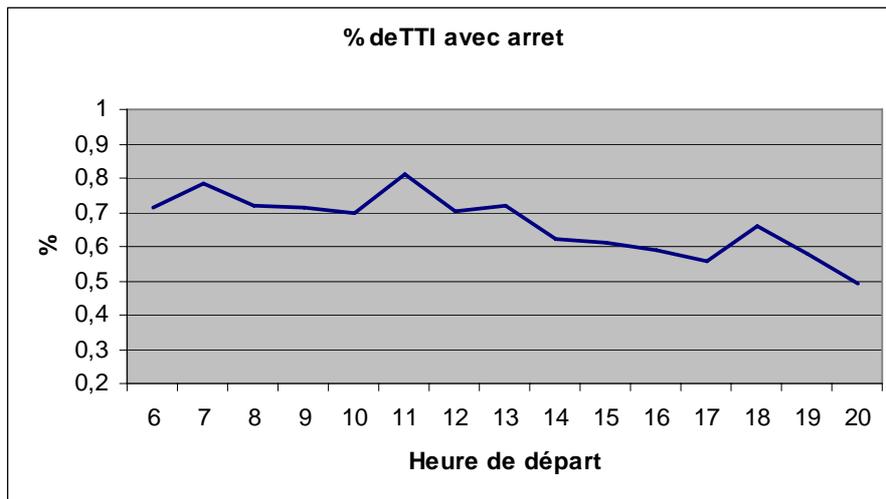


Fig. 14 : courbe de % d'arrêt

Nous faisons l'hypothèse que la durée de retard suit une distribution exponentielle $\Pr(RI = x) = \lambda e^{-\lambda x}$ de paramètre $\lambda = 1/27$. Le graphique suivant montre la différence de celle-ci avec la distribution empirique (le test du χ^2 est largement significatif). Cette modélisation de retard variant selon la longueur du trajet nous servira ultérieurement dans la conception de l'information routière fondée sur le TPM.

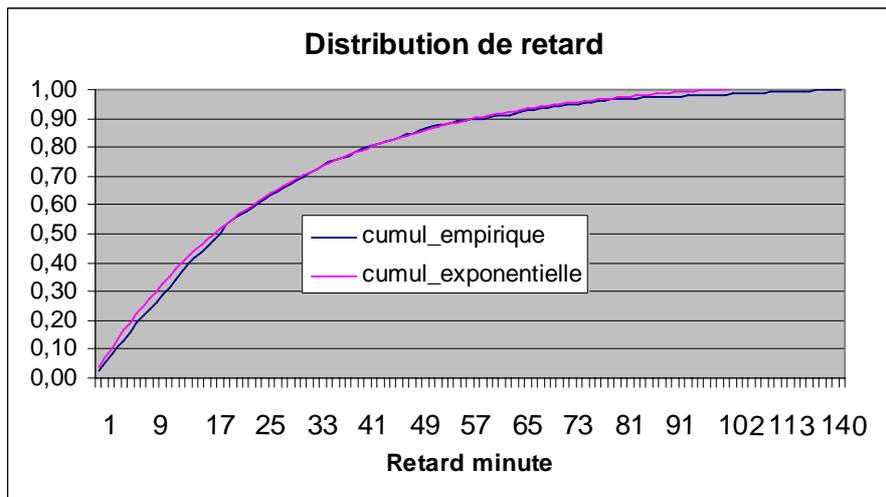


Fig. 15 : courbes de distribution de retard

Pour se positionner dans des situations comparables nous avons extrait de l'ensemble des 1923 véhicules décomptés ceux empruntant ce trajet les dimanches 6 et 13 novembre ainsi que le jour férié du 11 novembre, c'est à dire, les jours où la circulation des PL est interdite. Nous obtenons, ainsi, une autre distribution dont la médiane est toujours environ 216 minutes et son mode est 214. La queue droite de distribution est toujours relativement longue. La distribution des RI pour ce sous échantillon est presque la même que pour l'échantillon global. Sa moyenne est égale à 24 minutes (contre 27 minutes pour l'ensemble des 16 jours). Sa médiane est toujours 18 minutes, 69% des VL ont un RI inférieur à 30 minutes et dans 94% des cas le RI est inférieur à une heure.

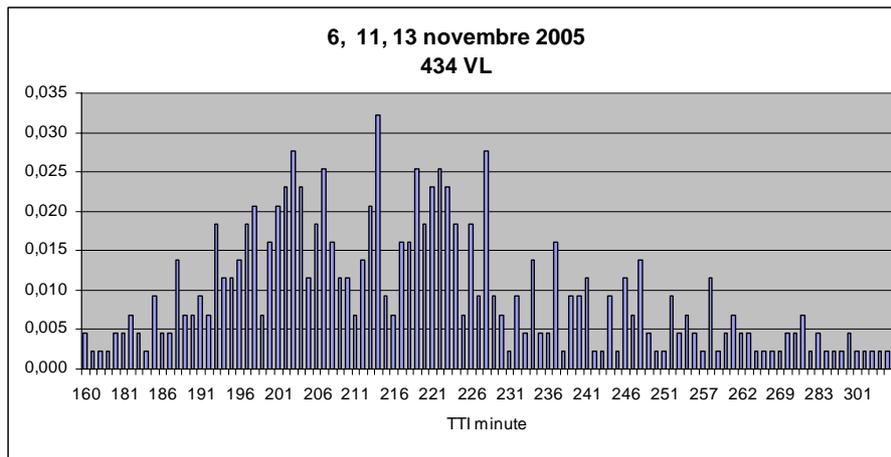


Fig. 16 : histogramme de TTI (dimanche et férié)

4.3.2 Comparaison avec les TP estimés

Pour ce trajet nous disposons via MI2 d'un an historique (octobre 2005 à septembre 2006) de 24 capteurs (cf. annexe C). La position des capteurs situés au début et à la fin d'un trajet ne coïncident pas toujours avec celle des gares de péages, c'est pourquoi il y a une différence de longueur de trajet selon que celui-ci a été traité via les télépéages ou le TPM. Dans le cas du trajet Coutevroult Schwindratzheim la différence est d'environ 500 mètres ce qui est négligeable par rapport à la totalité du trajet c'est-à-dire 421km. La distribution des TP estimés pour ce trajet par la méthode de calcul pondérée a la même allure que celle obtenue par la méthode non pondérée. Mais la médiane de la méthode pondérée est légèrement supérieure à celle de la non pondérée (225 minutes contre 222). Cette différence de quelques minutes qui existe aussi au niveau du neuvième décile (233 contre 229) vue la longueur du trajet est négligeable. Les deux histogrammes montrent que les deux distributions sont bimodales, dues à la composition du trafic mélangeant les PL et les VL. Les TPM (non pondérés) estimés pour les périodes avec interdiction de PL ont comme médiane 204 minutes ; et 90% des VL ont un TPM inférieur à 208 minutes. Les TP estimés pour les autres périodes forment une distribution distincte dont la médiane est environ 223 minutes.

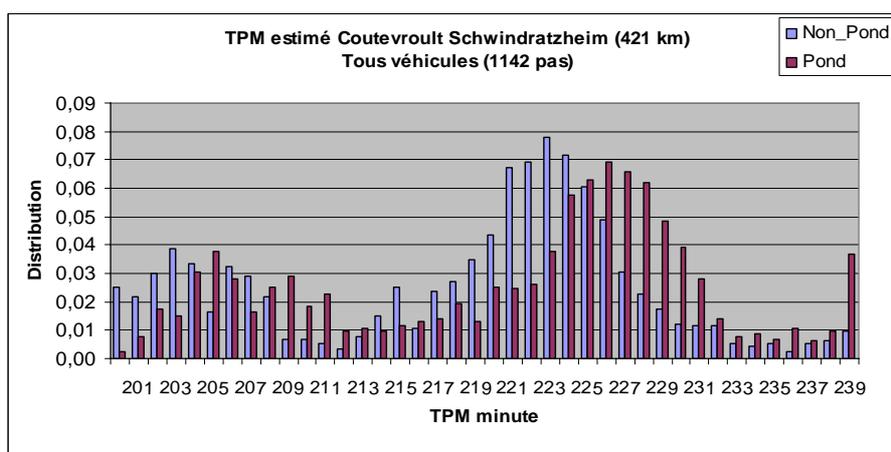


Fig. 17 : histogramme de TPM

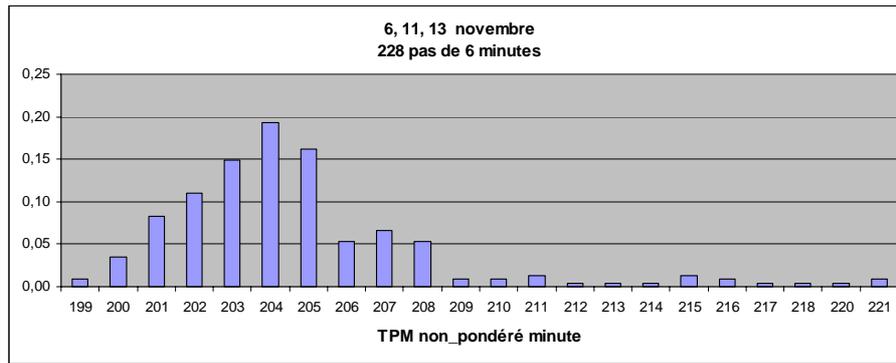


Fig. 18 : histogramme de TPM (dimanche et férié)

En se contentant des périodes d'interdiction de PL et en séparant les TTI supposés incluant des temps d'arrêt aux autres, la médiane des TTI et celle des TPM estimés ne s'écartent que de 3 minutes pour la version non pondérée (6 minutes pour la pondérée). Le tableau suivant donne quelques éléments pour cette comparaison. Tandis que 10% des TTI énumérés ont une vitesse qui peut dépasser 135km/h, pour la même portion des TPM la vitesse moyenne ne dépasse que de 126 km/h. Par ailleurs 90% des TTI sont effectués avec une vitesse supérieure ou égale à 118km/h (ce qui est envisageable pour les dimanche et fête), selon les TPM pour la même proportion la vitesse moyenne est au moins égale à 121km/h.

		Pondéré	Non pondéré	Télépéage sans RI	Télépéage total
Tous les jours	1 ^{er} décile	206	204	184	192
Tous les jours	Médiane	225	222	198	216
Tous les jours	9 ^{ème} décile	233	229	205	257
Dimanche Fête	1 ^{er} décile	202	201	186	192
Dimanche Fête	Médiane	207	204	201	217
Dimanche Fête	9 ^{ème} décile	209	208	213	254

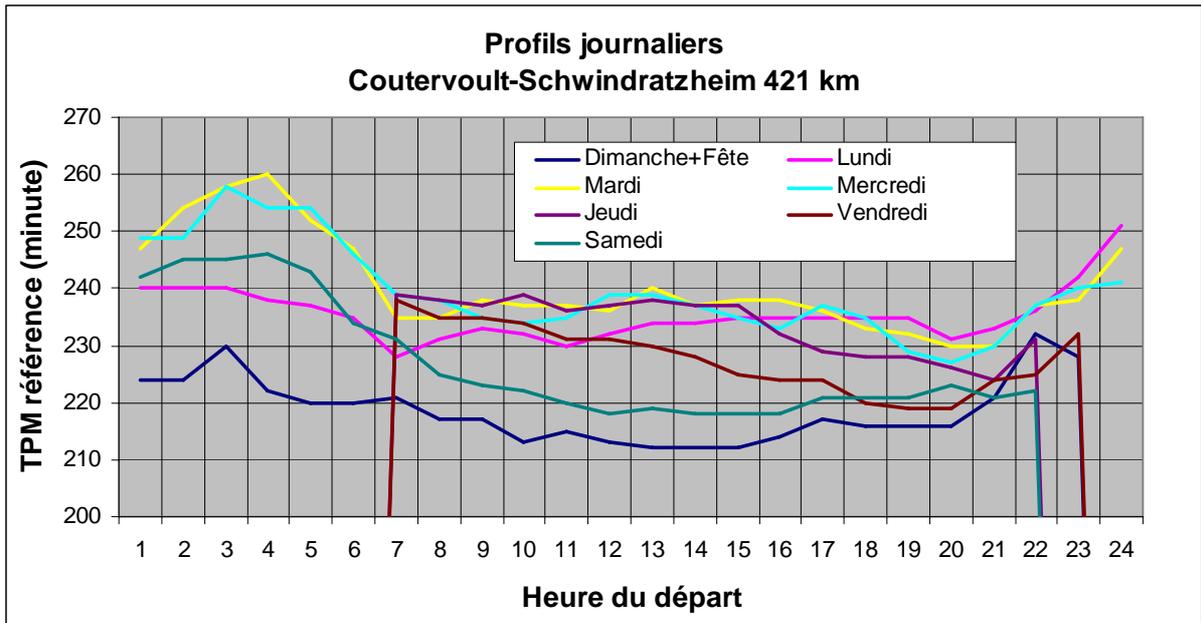
Tab. 2 : Comparaison TTI et TPM

4.3.3 Identification des jours de référence

Pour le trajet **Coutevroult Schwindratzheim** les jours médians (voir la définition en 2.2.4) pris comme référence sont les suivants :

- Dimanche : le 8 janvier 2006
- Lundi : le 14 novembre 2005
- Mardi : le 8 novembre 2005
- Mercredi : le 12 octobre 2005
- Jeudi : le 6 avril 2006
- Vendredi : le 7 avril 2006
- Samedi : le 4 février 2006

La période utile de la journée est prise entre 6h et 21h. Les TP nuls correspondent à l'impossibilité de calcul due à l'absence des données cohérentes.



Variation TPM (6h-21h)

Jour	Moyenne min	Ecart-type min	CV %	Min min	Max min	Etendu min
Dimanche+Fête	216,5	5,0	2,3	212	232	20
Lundi	233,1	2,2	1,0	228	236	8
Mardi	235,6	2,9	1,2	230	240	10
Mercredi	234,9	3,6	1,6	227	239	12
Jeudi	233,5	5,1	2,2	224	239	15
Vendredi	227,6	6,0	2,6	219	238	19
Samedi	221,3	3,3	1,5	218	231	13
Ensemble	228,9	8,1	3,5	212	240	28

4.4 Trajet Taissy Beaumont

4.4.1 Analyse des télépéages

Ce trajet de longueur de 151km correspond à un système fermé de péage. Le nombre de VL énumérés s'élève à 13.337. La médiane est estimée à 73 minutes, c'est-à-dire une vitesse de 124 km/h. Ici aussi uniquement 1% des VL roulent relativement vite. La queue de la distribution est étalée deux fois plus à droite qu'à gauche, $E_5^1 = 11\%$ et $E_9^5 = 25\%$.

Centile	V km/h	TTI min
1	153	59
10	139	65
25	131	69
Médiane	124	73
75	112	81
90	100	91
95	91	100

Tab. 3 : Distribution de TTI

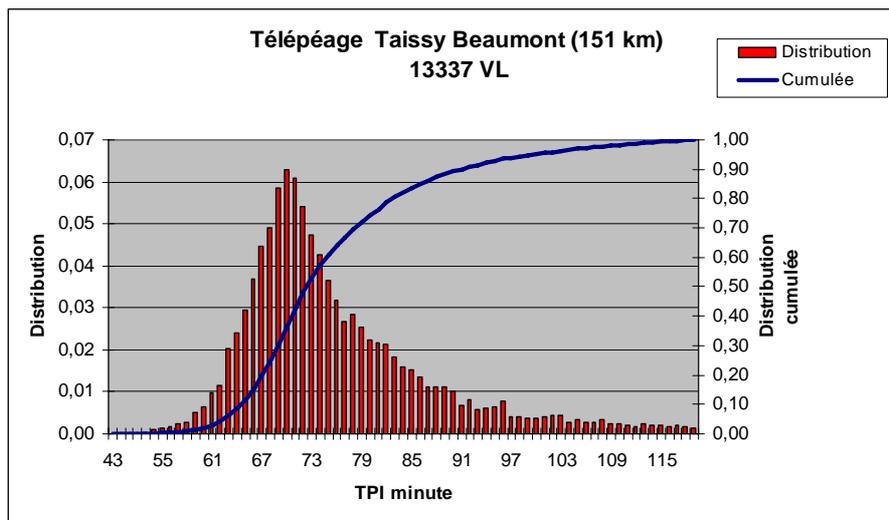


Fig. 19 : histogramme de distribution de TTI

Le mode de la distribution est 70 minutes ou 130Km/h. En appliquant la règle de l'estimation de retard on obtient une durée moyenne de 11 minutes. La courbe de pourcentage des VL avec arrêt montre qu'en dehors des heures nocturnes, ce pourcentage est très stable, il varie entre 60 et 70%. En effet étant donné le point de départ du trajet (Taissy situé à la sortie de la ville de Reims) et le sens de la circulation (Taissy Beaumont) les TTI sont moins affectés par les ralentissements que si l'on observait le TTI dans le sens inverse du trajet.

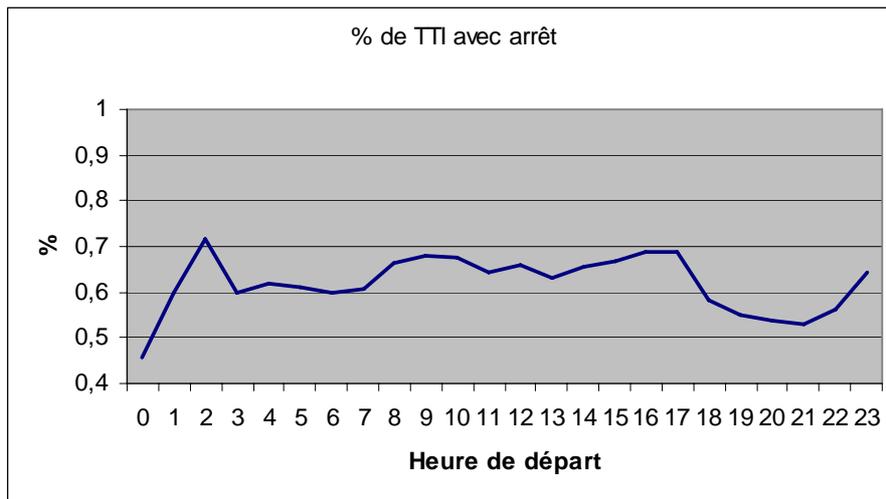


Fig. 20 : courbe de % d'arrêt

La moyenne de la durée des retards est évaluée à 12 minutes. La distribution de RI peut s'exprimer par un modèle exponentiel avec le paramètre $\lambda = 1/12$ (le test de χ^2 est assez significatif). La médiane de retard s'évalue à 8 minutes.

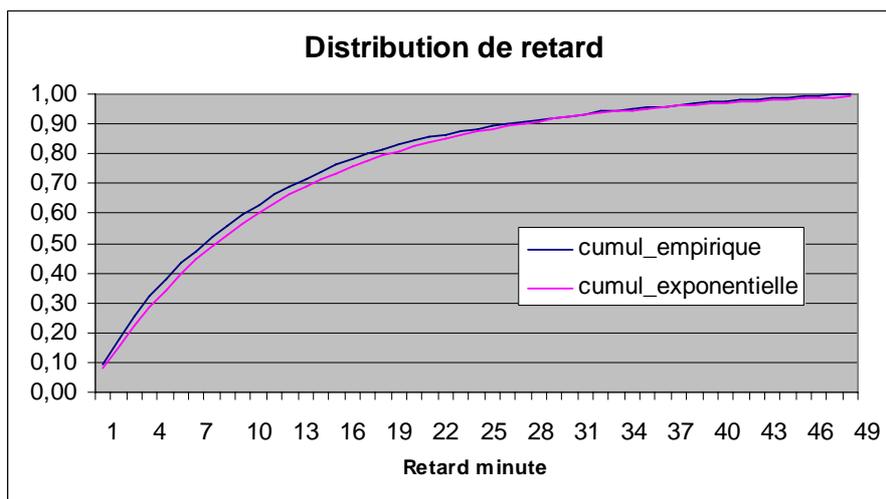


Fig. 21 : courbes de distribution de retard

La distribution des TTI sur les trois jours sans PL (6, 11 et 13 novembre) ressemble beaucoup à celle de l'ensemble des jours. Sa médiane est de 72 minutes, seulement 10% des VL ont un TTI dépassant 92 minutes. La médiane des RI est toujours 8 minutes et dans 10% des cas le RI dépasse 28 minutes.

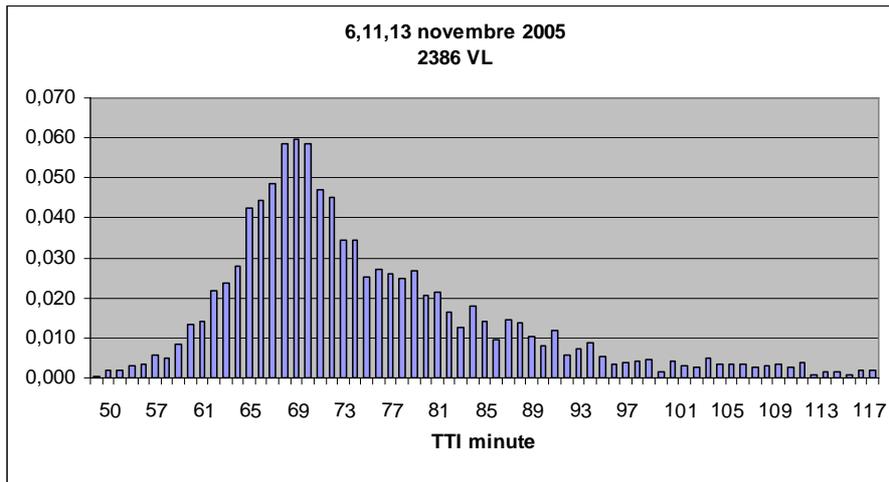


Fig. 22 : histogramme de TTI (dimanche et férié)

4.4.2 Comparaison avec les TP estimés

Pour ce trajet nous disposons d'uniquement quatre capteurs (cf. annexe C). La distance à parcourir entre les deux capteurs les plus proches des péages est 148 km, par conséquent les TTI seront majorés proportionnellement à l'écart de 3km (=158-148). Encore une fois cette différence est trop faible pour qu'elle fausse les comparaisons. Les deux distributions (non pondéré et pondéré) sont très semblables. Leur médiane est 76 minutes et leur dernier décile vaut 78 minutes. Chacune a une forme bimodale. Ceci s'explique par l'absence ou la présence des PL sur le trajet. La médiane de la distribution des TPM (non pondérés) relatifs aux trois jours sans PL est de l'ordre de 70 minutes.

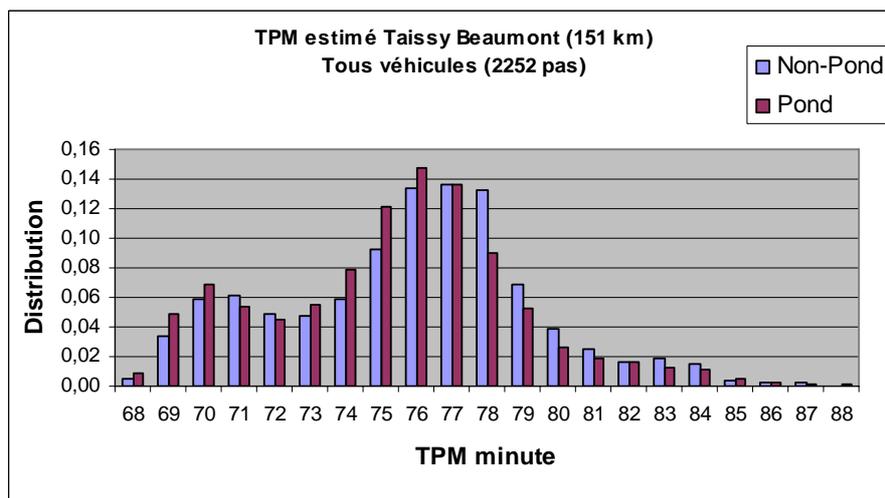


Fig. 23 : histogramme de TPM

En limitant la comparaison aux périodes sans PL et aux TTI supposés sans arrêt, on constate que la médiane des TPM s'écarte de 3 minutes de la médiane des TTI. Calculés dans les mêmes conditions les premiers déciles ont 8 minutes de différence et entre les 9^{ème} déciles la différence est de l'ordre de 12 minutes.

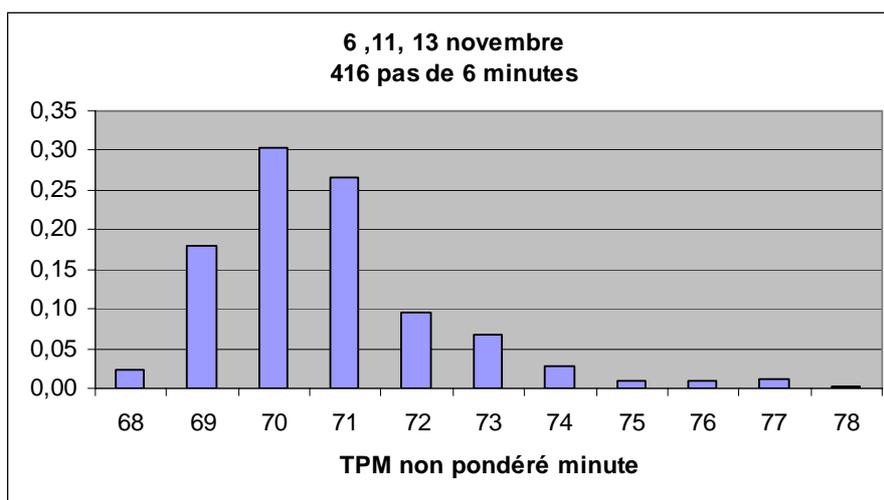


Fig. 24 : histogramme de TPM (dimanche et férié)

		Pondéré	Non pondéré	Télépéage sans RI	Télépéage total
Tous les jours	1 ^{er} décile	70	70	62	65
Tous les jours	Médiane	75	76	66	73
Tous les jours	9 ^{ème} décile	79	80	69	91
Dimanche Fête	1^{er} décile	69	69	61	64
Dimanche Fête	Médiane	70	70	67	72
Dimanche Fête	9^{ème} décile	73	73	69	92

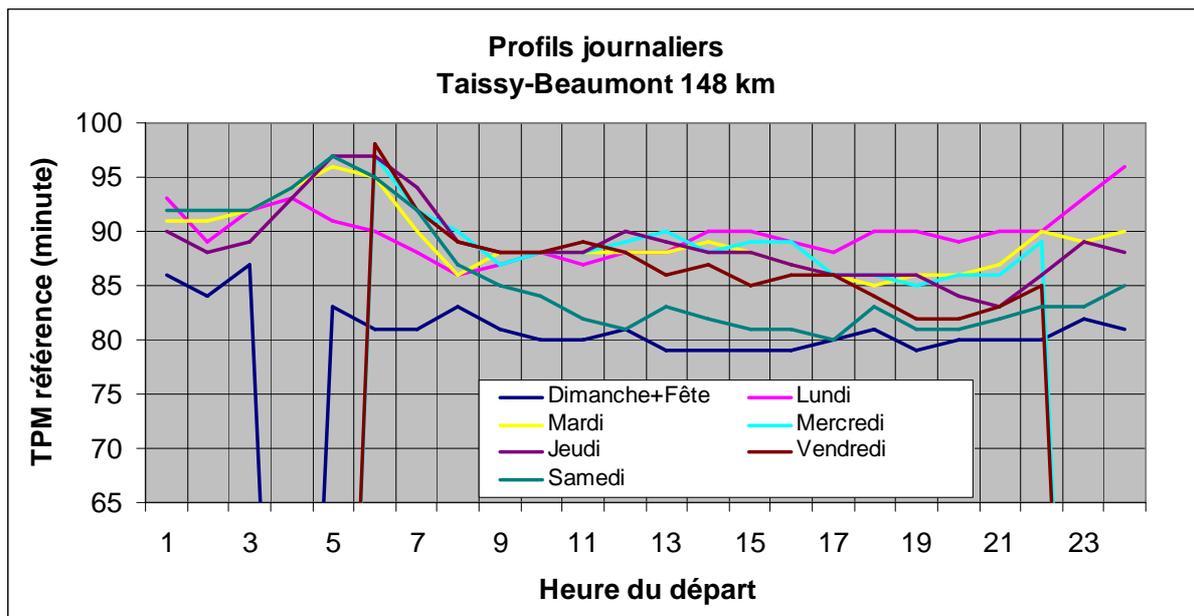
Tab. 4 : Comparaison TTI et TPM

4.4.3 Identification des jours de référence

Pour le trajet **Taissy Beaumont** les jours médians pris comme référence sont les suivants :

- Dimanche : le 8 janvier 2006
- Lundi : le 23 janvier 2006
- Mardi : le 4 octobre 2005
- Mercredi : le 26 avril 2006
- Jeudi : le 13 octobre 2005
- Vendredi : le 17 mars 2006
- Samedi : le 5 novembre 2005

La période utile de la journée est prise entre 6h et 21h. Les TP nuls correspondent à l'impossibilité de calcul due à l'absence des données cohérentes.



Variation TPM (6h-21h)

Jour	Moyenne min	Ecart-type min	CV %	Min min	Max min	Etendu min
Dimanche+Fête	80,1	1,1	1,4	79	83	4
Lundi	88,6	1,3	1,5	86	90	4
Mardi	87,5	1,5	1,7	85	90	5
Mercredi	88,0	1,9	2,2	85	92	7
Jeudi	87,5	2,5	2,9	83	94	11
Vendredi	86,3	2,7	3,2	82	92	10
Samedi	83,0	3,0	3,6	80	92	12
Ensemble	85,9	3,6	4,2	79	94	15

4.5 Trajet Beaumont St Avoild

4.5.1 Analyse des télépéages

Ce trajet de longueur 66km contourne le Nord-Est de Metz. Pendant la période d'évaluation les données des télépéages ont décompté 6347 VL. Tandis que seulement 1% des VL ont une vitesse relativement élevée, 10% roulent avec une vitesse relativement faible. Ceci est sans doute dû à la nature périurbaine du trafic et les ralentissements des heures de pointe. Ceci se constate aussi par l'étendue très importante de la queue de droite de la distribution. L'étendue de la queue de droite est 7 fois plus grande que celle de gauche $E_5^1 = 9\%$ et $E_9^5 = 69\%$. Dans le cas des trajets précédents, ce rapport était de l'ordre de 2.

Centile	V km/h	TTI min
1	152	26
10	136	29
25	128	31
Médiane	124	32
75	102	39
90	73	54
95	59	67

Tab. 5 : Distribution de TTI

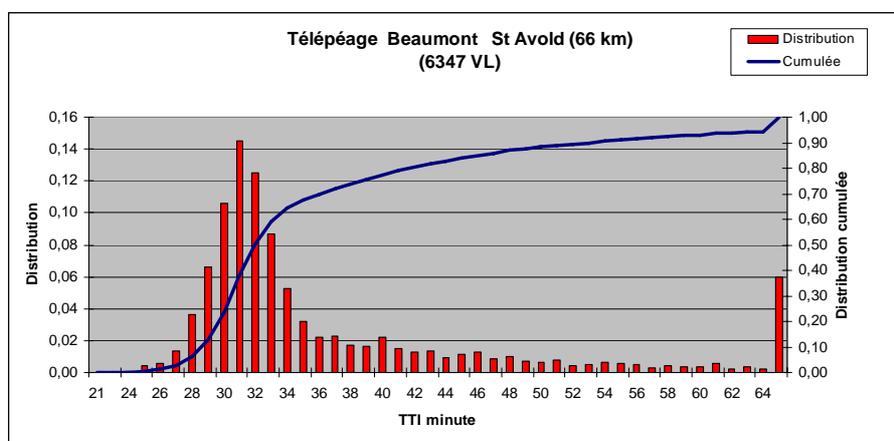


Fig. 25 : histogramme de distribution de TTI

La médiane de la distribution des TTI est évaluée à 32 minutes, ce qui correspond à une vitesse de 124 km/h. Le mode est de 31 minutes, pas trop loin de la médiane. En appliquant la règle de l'estimation de retard on obtient la durée moyenne de retard qui s'évalue à 16 minutes, c'est-à-dire la moitié de la médiane des TTI. La courbe de pourcentage des VL avec arrêt montre sa forte stabilité entre 5h et 19h où il varie entre 60 et 70%.

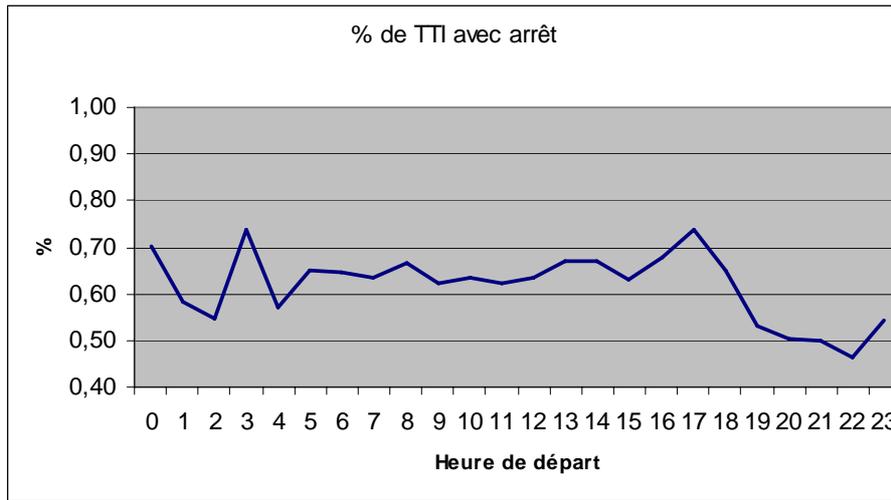


Fig. 26 : courbe de % d'arrêt

Cette fois la distribution théorique exprimée par un modèle exponentiel avec le paramètre $\lambda = 1/16$ ne colle pas assez bien avec la distribution empirique du retard (le test du χ^2 est peu significatif).

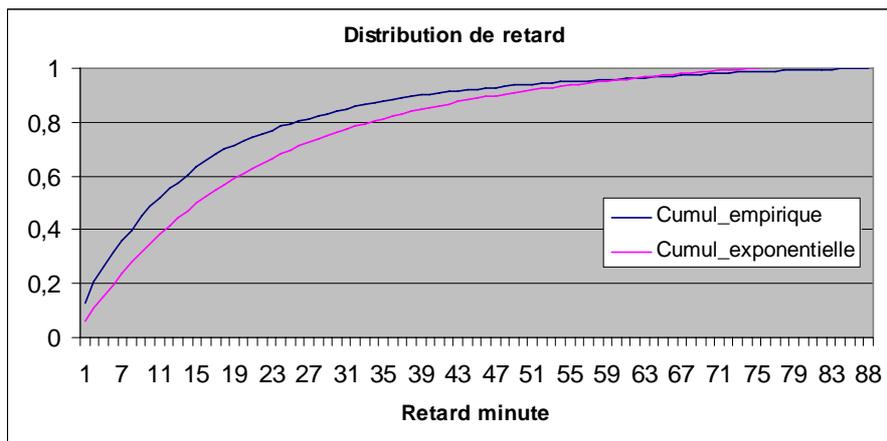


Fig. 27 : courbes de distribution de retard

La distribution des TTI sur les trois jours sans PL (6, 11 et 13 novembre) ressemble beaucoup à celle de l'ensemble des 16 jours. Son mode est 31 minutes et sa médiane 32. Seulement 7% des VL ont un TTI dépassant 60 minutes. La médiane des RI est toujours 7 minutes et dans 12% des cas le RI dépasse 30 minutes.

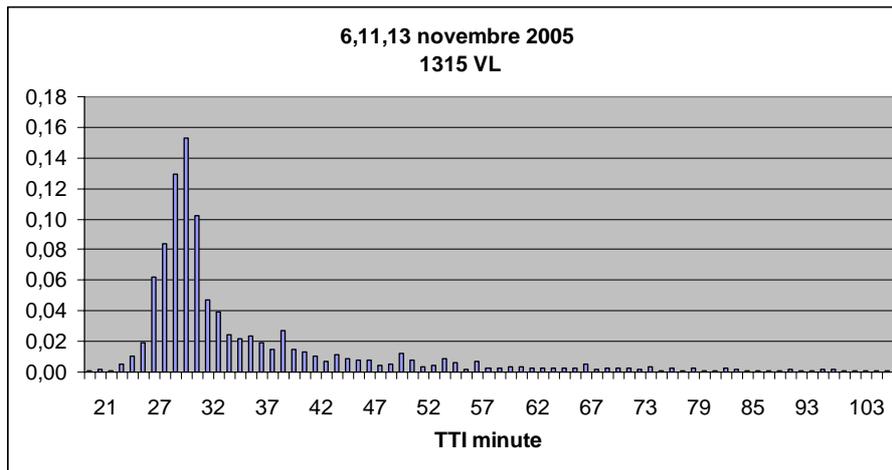


Fig. 28 : histogramme de TTI (dimanche et férié)

4.5.2 Comparaison avec les TPM estimés

Le nombre de capteurs disponibles pour ce trajet est six (cf. annexe C). La différence entre la distance à parcourir entre les deux péages (origine et destination) et celle des deux capteurs les plus près est de l'ordre de 5km. La distribution des TP estimés pour ce trajet par la méthode de calcul pondérée a la même allure que celle obtenue par la méthode non pondérée. En particulier la médiane de chacune des deux distributions est environ 32 minutes. La bimodalité des deux distributions due au mélange de VL et PL est moins prononcée que dans les cas précédents.

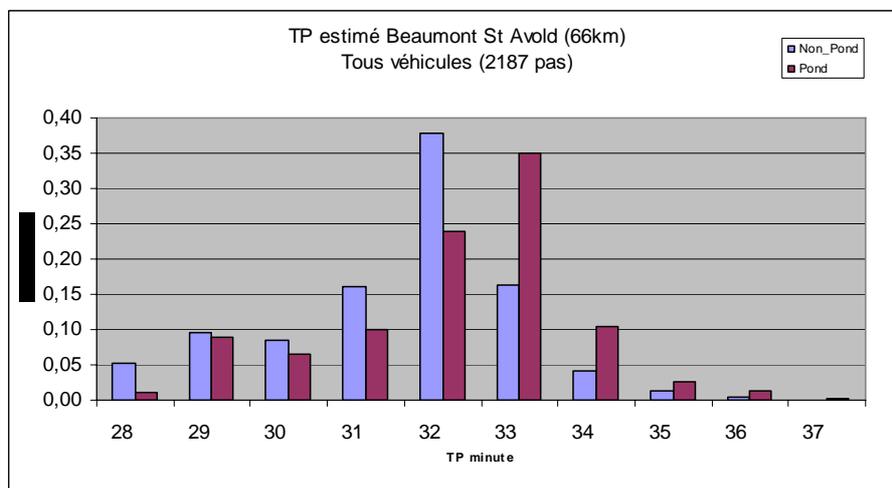


Fig. 29 : histogramme de TPM

L'étendue de la variation de TPM est de l'ordre de 11 minutes (entre 26 et 37 minutes) mais quant il s'agit des jours avec interdiction de PL cette étendue se réduit à 5 minutes (entre 28 et 33 minutes). Dans ce type de jour 75 % des TPM sont inférieurs à 30 minutes.

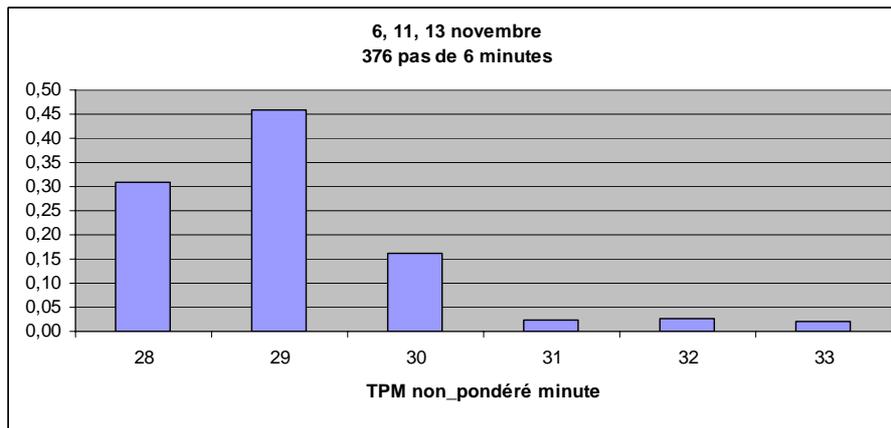


Fig. 30 : histogramme de TPM (dimanche et férié)

En limitant la comparaison aux périodes sans PL et aux TTI supposés sans arrêt, on constate que la médiane des TPM est estimée égale à la médiane des TTI. Calculés dans les mêmes conditions les premiers déciles ont 1 minute de différence et entre les 9^{ième} déciles la différence est nulle.

		Non pondéré	Télépage sans RI	Télépage total
Tous les jours	1 ^{er} décile	28	24	26
Tous les jours	Médiane	32	29	32
Tous les jours	9 ^{ème} décile	33	31	54
Dimanche Fête	1 ^{er} décile	28	27	28
Dimanche Fête	Médiane	29	29	31
Dimanche Fête	9 ^{ème} décile	30	30	53

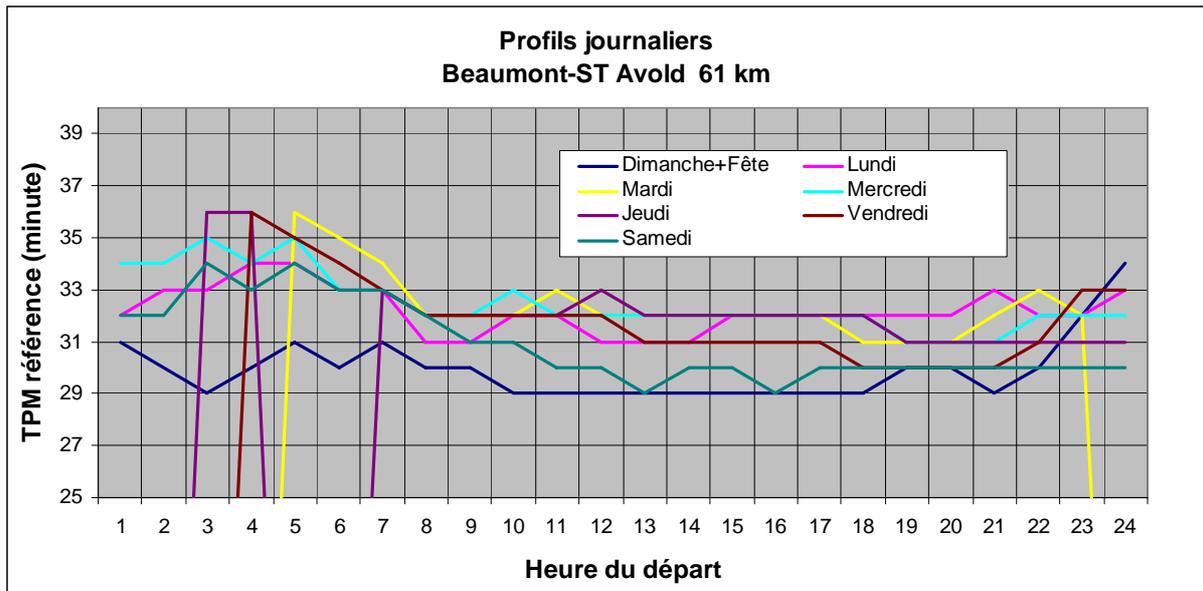
Tab. 6 : Comparaison TTI et TPM

4.5.3 Identification des jours de référence

Pour le trajet **Beaumont St Avoild** les jours médians pris comme référence sont les suivants :

- Dimanche : le 8 janvier 2006
- Lundi : le 14 novembre 2005
- Mardi : le 4 avril 2006
- Mercredi : le 12 octobre 2005
- Jeudi : le 6 avril 2006
- Vendredi : le 7 avril 2006
- Samedi : le 5 novembre 2005

La période utile de la journée est prise entre 6h et 21h. Les TP nuls correspondent à l'impossibilité de calcul due à l'absence des données cohérentes.



Variation TPM (6h-21h)

Jour	Moyenne min	Ecart-type min	CV %	Min min	Max min	Etendu min
Dimanche+Fête	29,4	0,6	2,1	29	31	2
Lundi	31,8	0,7	2,1	31	33	2
Mardi	32,1	0,8	2,4	31	34	3
Mercredi	31,9	0,6	1,8	31	33	2
Jeudi	31,9	0,6	1,9	31	33	2
Vendredi	31,2	0,9	2,9	30	33	3
Samedi	30,3	1,0	3,3	29	33	4
Ensemble	31,2	1,2	3,8	29	34	5

4.6 Trajet Dormans Taissy

4.6.1 Analyse des télépéages

D'après les données de télépéage (voir fig. ci dessous), pendant la période d'évaluation, 50% des 9237 individus escomptés ont mis au plus 16 minutes pour parcourir ce trajet avec un VL. Un peu plus d'un individu sur trois (31%) ont mis 16 minutes, ceci correspond au mode de la distribution. Uniquement 6% des individus ont un TTI qui dépasse 32 minutes (c'est-à-dire deux fois le TTI médiane). La grande dispersion entre les TTI est due, en partie, aux comportements individuels des conducteurs. L'étendue de la queue de droite est 5 fois plus grande que celle de gauche $E_5^1 = 12\%$ et $E_9^5 = 56\%$. Celle-ci est due aux ralentissements aux heures de pointes.

Centile	V km/h	TTI min
1	152	13
10	141	14
25	132	15
Médiane	124	16
75	116	17
90	79	25
95	56	35

Tab. 7 : Distribution de TTI

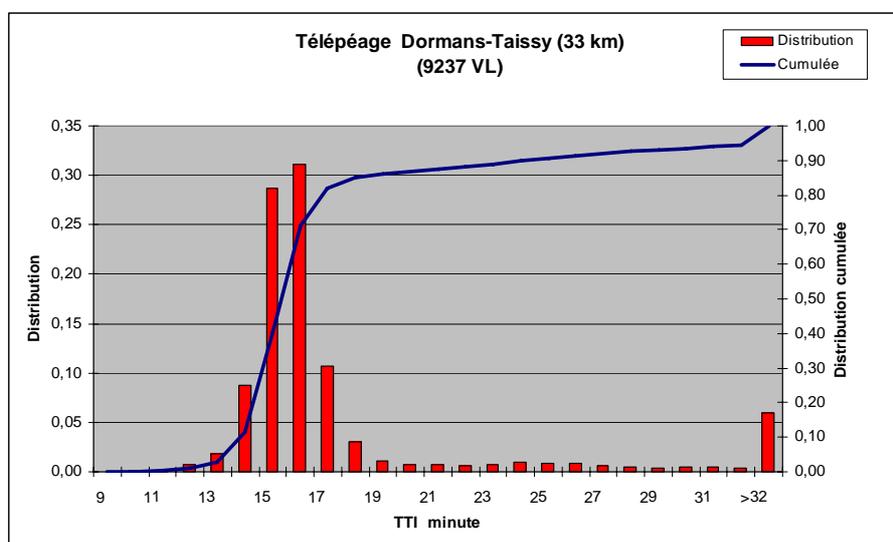


Fig. 31 : histogramme de distribution de TTI

En appliquant la règle de l'estimation de retard on obtient la durée moyenne de retard égale à 18 minutes, c'est-à-dire 2 minutes de plus que la médiane des TTI. La courbe de pourcentage des VL avec arrêt montre sa forte stabilité, il varie souvent entre 25 et 35%. Il est vraisemblable que les retards ainsi identifiés sont dus aux ralentissements et non pas aux arrêts dans les airs. Le modèle théorique d'estimation de temps de retard calculé pour les trois cas précédents (modèle exponentiel) ne s'applique pas bien ici.

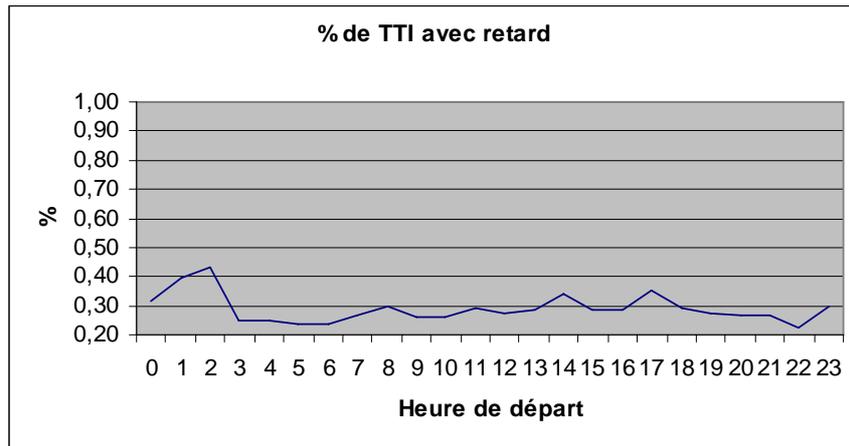


Fig. 32 : courbe de % d'arrêt

La distribution des TTI sur les trois jours sans PL (6, 11 et 13 novembre) ressemble beaucoup à celle de l'ensemble des 16 jours. Son mode est 15 minutes et sa médiane 16. Seulement 12% des VL ont un TTI dépassant 20 minutes. La médiane des RI est évaluée à 1 minute et dans 20% des cas le RI dépasse 8 minutes.

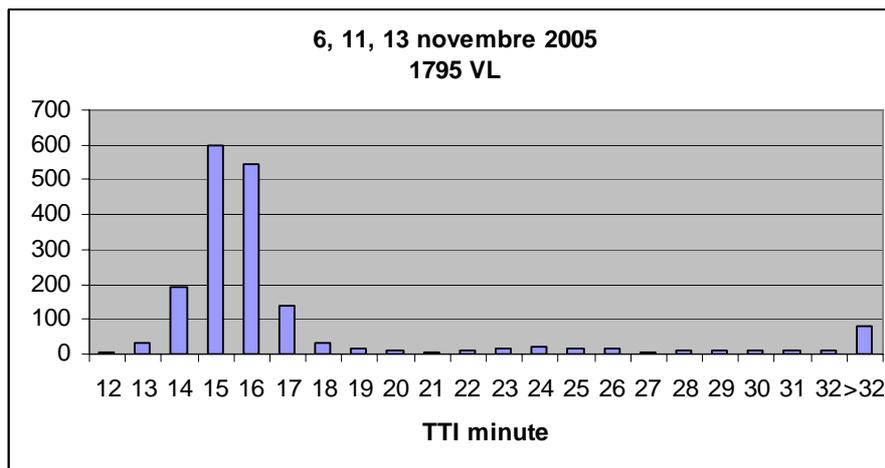


Fig. 33 : histogramme de TTI (dimanche et férié)

4.6.2 Comparaison avec les TP estimés

Le nombre de capteurs disponibles pour ce trajet est sept (cf. annexe C). Pour ce trajet la différence entre la distance à parcourir entre les deux péages (origine et destination) et celle des deux capteurs les plus près est de l'ordre de 1,4 km. La distribution des TP estimés pour ce trajet par la méthode de calcul pondérée a la même allure que celle obtenue par la méthode non pondérée. En particulier la médiane de chacune des deux distributions est environ 18 minutes. La bimodalité des deux distributions due au mélange de VL et PL est très faible.

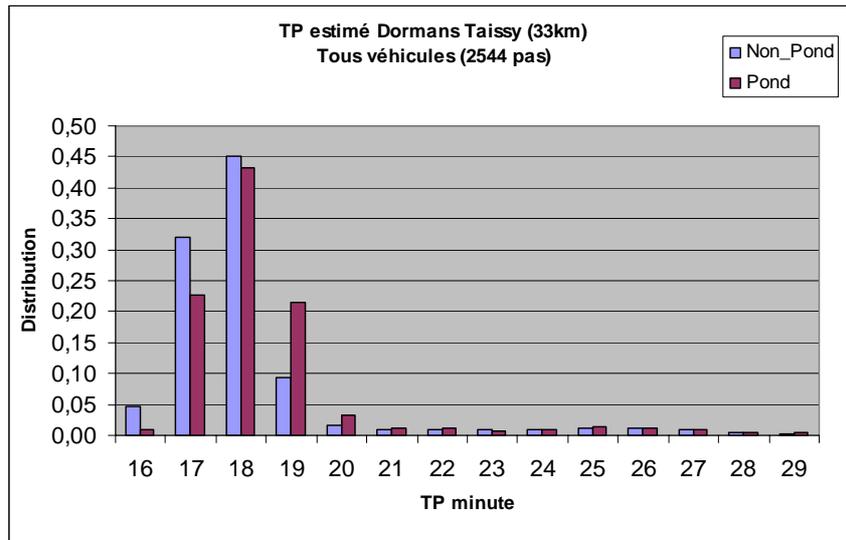


Fig. 34 : histogramme de TPM

L'étendue de la variation de TPM est de l'ordre de 13 minutes (entre 16 et 39 minutes) mais quant il s'agit des jours avec interdiction de PL cette étendue se réduit à 2 minutes (entre 16 et 18 minutes). Dans ce type de jour 10 % des TPM sont supérieurs à 17 minutes.

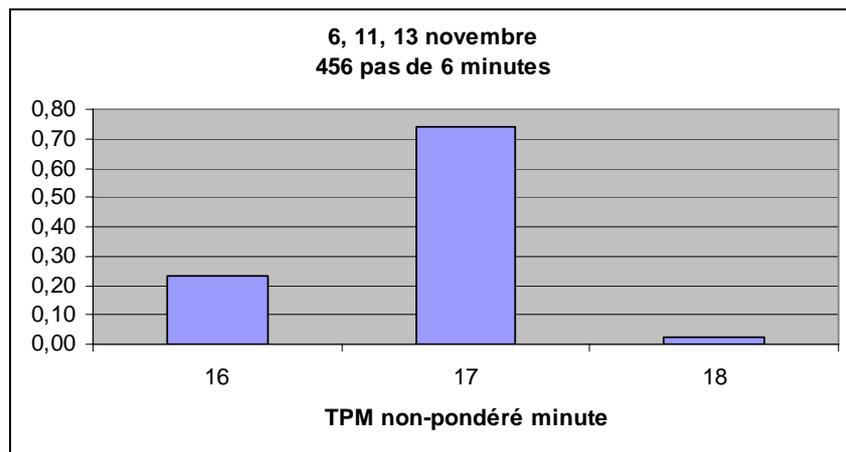


Fig. 35 : histogramme de TPM (dimanche et férié)

En limitant la comparaison aux périodes sans PL et aux TTI supposés sans arrêt, on constate que la médiane des TPM a 3 minutes de différence avec la médiane des TTI. Calculés dans les mêmes conditions les premiers déciles ont 2 minutes de différence et entre les 9^{èmes} déciles la différence est également 2 minutes.

		Non pondéré	Télépage sans RI	Télépage total
Tous les jours	1 ^{er} décile	17	13	14
Tous les jours	Médiane	18	14	16
Tous les jours	9 ^{ème} décile	19	15	25
Dimanche Fête	1 ^{er} décile	16	14	14
Dimanche Fête	Médiane	17	14	16
Dimanche Fête	9 ^{ème} décile	17	15	24

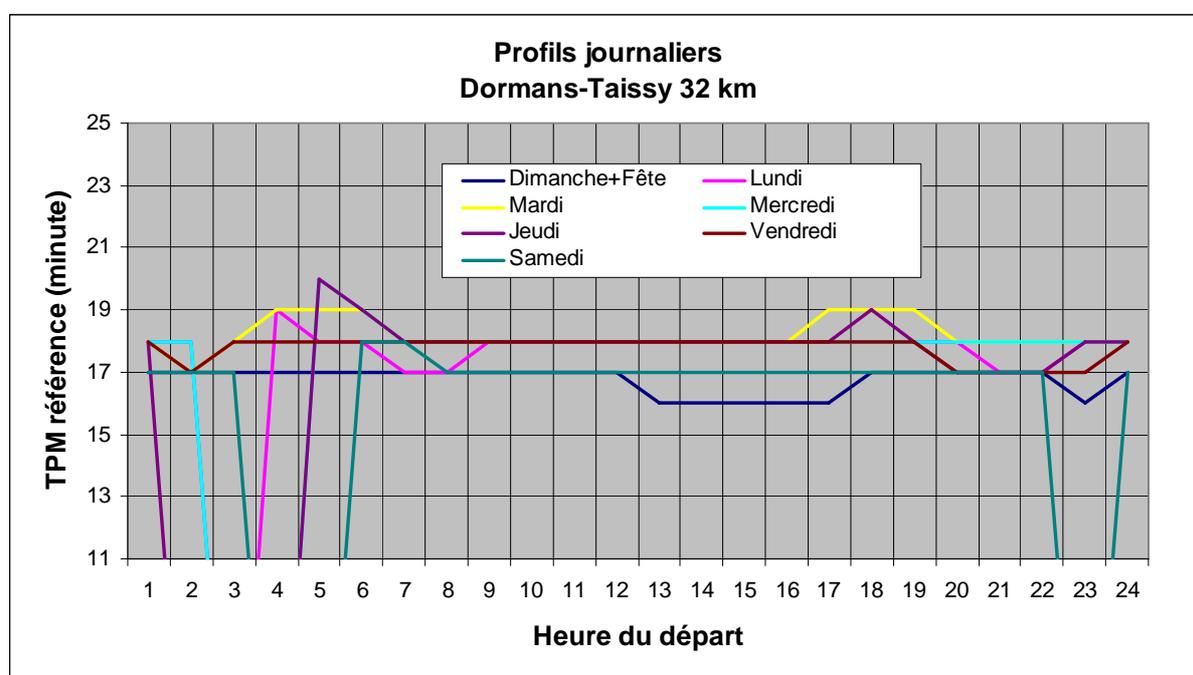
Tab. 6 : Comparaison TTI et TPM

4.6.3 Identification des jours de référence

Pour le trajet **Dormans Taissy** les jours médians pris comme référence sont les suivants :

- Dimanche : le 2 octobre 2005
- Lundi : le 13 mars 2006
- Mardi : le 8 novembre 2005
- Mercredi : le 22 février 2006
- Jeudi : le 9 février 2006
- Vendredi : le 27 janvier 2005
- Samedi : le 9 septembre 2006

La période utile de la journée est prise entre 6h et 21h. Les TP nuls correspondent à l'impossibilité de calcul due à l'absence des données cohérentes.



Variation TPM (6h-21h)

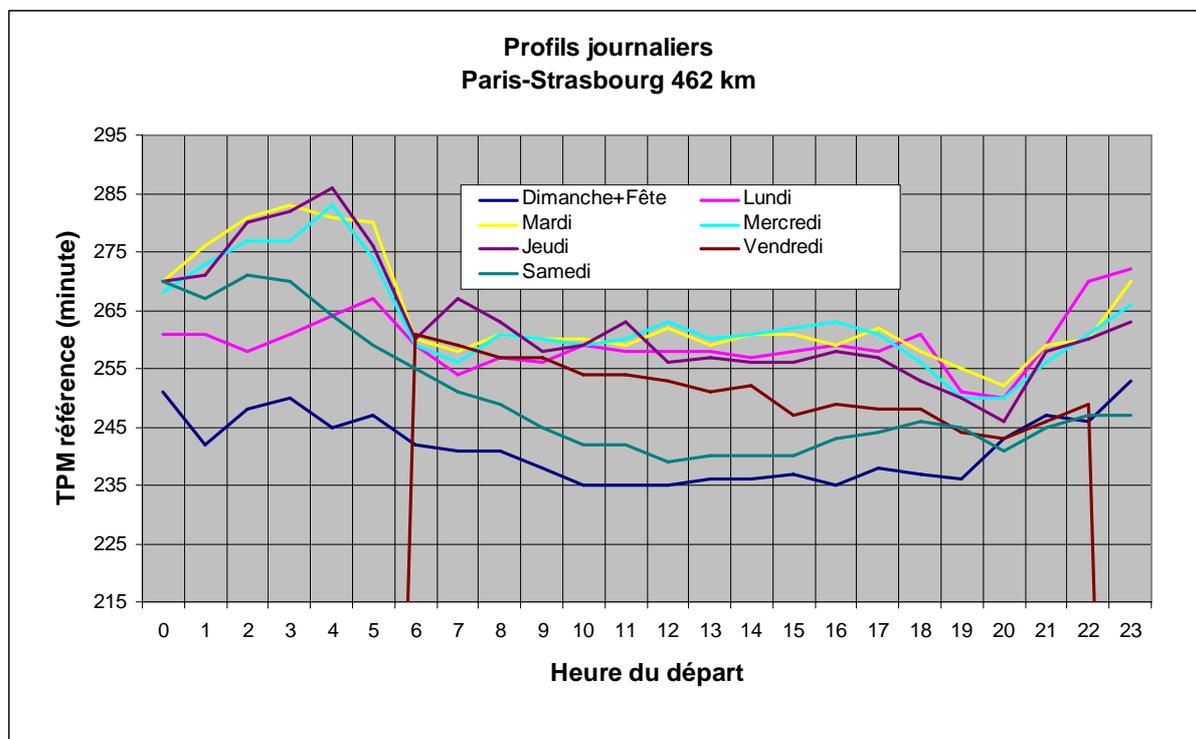
Jour	Moyenne min	Ecart-type min	CV %	Min min	Max min	Etendu min
Dimanche+Fête	16,7	0,5	2,9	16	17	1
Lundi	17,8	0,4	2,5	17	18	1
Mardi	18,2	0,4	2,2	18	19	1
Mercredi	18,0	0,0	0,0	18	18	0
Jeudi	17,9	0,5	2,8	17	19	2
Vendredi	17,8	0,4	2,3	17	18	1
Samedi	17,1	0,3	1,5	17	18	1
Ensemble	17,6	0,6	3,6	16	19	3

4.7 L'itinéraire Paris – Strasbourg dans les deux sens

Le trajet le plus long de notre terrain d'expérience est entre Paris et Strasbourg ou plus exactement entre le capteur situé au Noisy le Grand ouest (Pk 13,3) et celui près de Strasbourg à Reichstett (Pk 475,2). Ce trajet traverse 36 capteurs SIREDO (cf. annexe C). Les jours médians pris comme référence sont les suivants :

- Dimanche : le 2 octobre 2005
- Lundi : le 7 novembre 2005
- Mardi : le 8 novembre 2005
- Mercredi : le 9 novembre 2005
- Jeudi : le 17 novembre 2005
- Vendredi : le 28 avril 2006
- Samedi : le 7 janvier 2006

La période utile de la journée est prise entre 6h et 21h. Les TP nuls correspondent à l'impossibilité de calcul due à l'absence des données cohérentes.

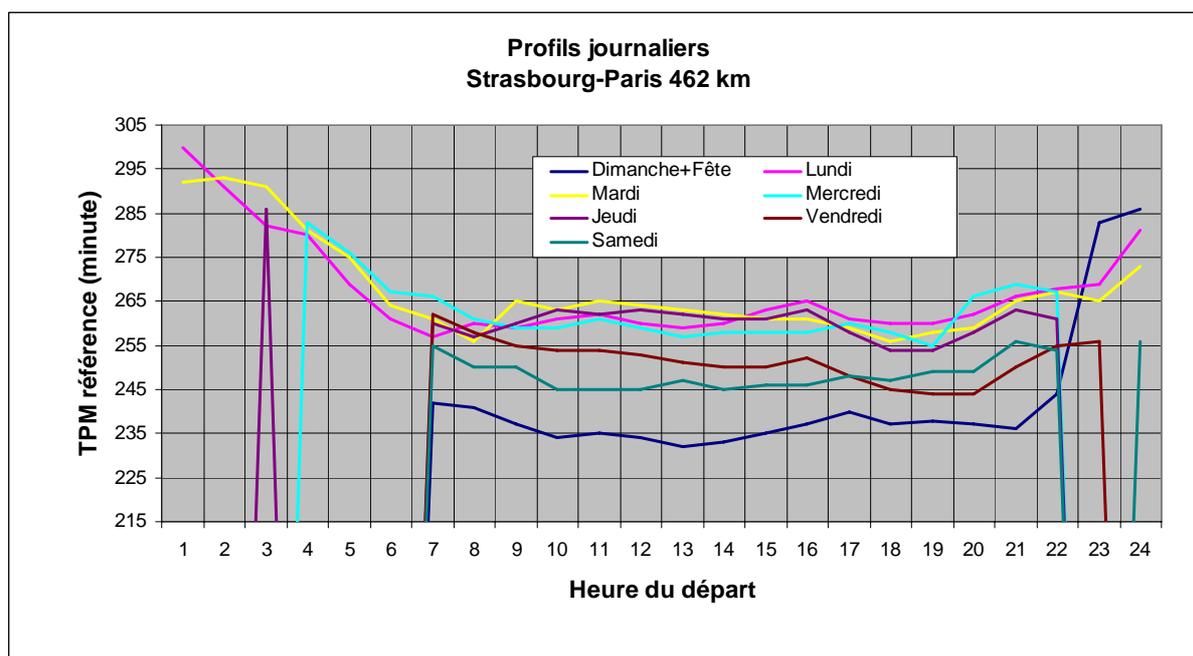


Variation TPM (6h-21h)

Jour	Moyenne min	Ecart-type min	CV %	Min min	Max min	Etendu min
Dimanche+Fête	238,3	3,5	1,5	235	247	12
Lundi	257,0	3,0	1,2	250	261	11
Mardi	259,1	2,6	1,0	252	262	10
Mercredi	258,6	4,0	1,6	250	263	13
Jeudi	257,3	5,0	1,9	246	267	21
Vendredi	251,4	5,4	2,1	243	261	18
Samedi	244,2	4,4	1,8	239	255	16
Ensemble	252,3	8,5	3,4	235	267	32

Ce même trajet dans le sens de retour (Strasbourg- Paris) a son propre comportement. Les jours médians pris comme référence sont les suivants :

- Dimanche : le 22 janvier 2006
- Lundi : le 7 novembre 2005
- Mardi : le 8 novembre 2005
- Mercredi : le 4 janvier 2006
- Jeudi : le 13 octobre 2005
- Vendredi : le 14 octobre 2005
- Samedi : le 8 octobre 2005



Variation TPM (6h-21h)

Jour	Moyenne min	Ecart-type min	CV %	Min min	Max min	Etendu min
Dimanche+Fête	237,0	3,3	1,4	232	244	12
Lundi	261,4	2,9	1,1	257	268	11
Mardi	261,6	3,3	1,3	256	267	11
Mercredi	260,7	4,1	1,6	255	269	14
Jeudi	260,0	3,0	1,2	254	263	9
Vendredi	251,6	4,9	2,0	244	262	18
Samedi	248,6	3,7	1,5	245	256	11
Ensemble	254,4	9,3	3,7	232	269	37

5. Conclusions

La problématique

Les exploitants et les usagers de la route sont tous convaincus que la connaissance des temps de parcours est utile. Néanmoins il existe une différence importante entre deux notions : le temps de parcours moyen TPM qu'un exploitant peut fournir et le temps de parcours individuel TPI qu'un usager perçoit. Les exploitants qui gèrent les déplacements en masse n'ont pas accès au TPI qui dépend des facteurs individuels (le comportement des conducteurs et la puissance de leur véhicule). La problématique est de concevoir des TPM qui soient sensibles à la fois à la demande des déplacements (dans sa dimension temporelle) et à la capacité des routes (dans sa dimension événementielle).

A la demande du Ministère des Transports, l'INRETS a élaboré un prototype de moteur de calcul permettant d'obtenir un TPM de référence qui varie selon la date et l'heure du départ, et des conditions de circulation (chantiers, conditions météorologiques). Ce temps que nous proposons, est un temps semi dynamique contrairement à celui, complètement statique, proposé par certains calculateurs d'itinéraire que l'on peut trouver sur le web. Il ne faut pas non plus le confondre avec celui fourni par les systèmes de navigation qui, eux, indique un temps réel. Le TPM est un temps de parcours différé qui permet de planifier un déplacement à l'avance. Il permet ainsi aux usagers de prévoir leur déplacement, et aux exploitants de mieux gérer leurs réseaux routiers.

Le terrain d'expérimentation : Paris-Strasbourg

Le moteur de calcul dispose actuellement d'un historique de données d'environ un an. Grâce aux 36 capteurs routiers disposés sur l'autoroute A4, reliant Paris et Strasbourg, des données tels que le débit, la vitesse, le taux d'occupation et parfois le pourcentage de poids lourds ont pu être recueillies et, ont surtout permis d'obtenir des TPM de référence.

La synthèse de l'évaluation

Dans une première étape, en comparant les TPM obtenus par nos algorithmes et ceux disponibles sur Internet via certains calculateurs d'itinéraire comme MAPY et Via Michelin, on s'aperçoit que nous sommes en mesure de produire, avec au moins autant de précision, des temps de parcours référentiel par itinéraire qui varient selon les conditions de circulation liées au type de jour aux heures de départ. C'est-à-dire des informations semi dynamiques. De plus ces informations peuvent être réajustées selon les perturbations connues de l'état de l'itinéraire. Ainsi un temps de retard est calculé selon la nature de ces perturbations.

Dans une deuxième étape nous avons cherché une méthode de comparaison plus fiable pour pouvoir valider nos résultats. Cette autre facette de l'évaluation a été possible grâce aux données fournies par la SANEF. Il s'agit de données télépéages relatives à une période donnée. C'est-à-dire des temps de trajet réalisés par des clients, dans de vraies conditions de circulation, contrairement à nos résultats qui sont une estimation du temps de parcours à partir de données historiques. Toutefois, la comparaison n'est quand même pas immédiate au moins pour trois raisons qui sont les suivantes :

- Les données télébadges ne concernent que les véhicules légers alors que les données des capteurs englobent tous les types de véhicules passant sur les sections de routes.

- Un TTI calculé à partir des transactions télépéages englobe le retard dû aux arrêts volontaires ou des ralentissements subis, alors qu'un TPM déduit des capteurs ne peut tenir compte des temps d'arrêts hors circulation, par exemple les arrêts dans les airs de stationnement.
- Les TTI à partir des télépéages sont des données individuelles alors que les données des capteurs sont agrégées toutes les six minutes.

Dans cette partie de l'évaluation après avoir tenté de reconstituer des conditions semblables de comparaison, nous avons constaté que les distributions statistiques des TTI et des TPM bâties par itinéraires sont relativement proches. Dans le sens où leurs caractéristiques de tendance centrale (moyenne et médiane par exemple) ainsi que leurs caractéristiques de dispersion (centiles par exemple) ne sont pas trop éloignées. Les données de télépéages ne peuvent pas remplacer une vraie campagne de mesures de TP via les véhicules flottants mais elles nous ont aidés à confirmer la vraisemblance et la validité des TPM estimés par nos algorithmes.

Les enjeux d'une meilleure appréhension des temps de parcours sont multiples. Cette recherche et l'évaluation des résultats de l'expérimentation mettent en exergue l'utilité et l'importance de l'information sur les temps de parcours interurbains dans les cas ci-dessous :

- 1) **La longueur d'un trajet interurbain dépasse une centaine de kilomètres,**
- 2) **Le calcul des TPM intègre l'effet des conditions de circulation (situations variées mais récurrentes) et il peut, aussi, tenir compte de l'impact des événements exogènes connus à l'avance (comme la météo, les chantiers),**
- 3) **La combinaison de ces deux cas.**

Bibliographie

- 1 F. Gillaizeau, M. Danech-Pajouh, JC, Pierrelée, Préviation qualitative du trafic routier par régression logistique, RTS 93, 2006
- 2 F.ROYER, *Une procédure d'épuration des données issues des capteurs SIREDO*, INRETS, 2004
- 3 F.ROYER, *Une procédure d'estimation des vitesses critiques sur un réseau autoroutier*, INRETS, 2004
- 4 M. Danech-Pajouh, janvier- mars 2003, Les modèles de prévision du dispositif Bison Futé et leur évolution, RTS n°78, 2003
- 5 M. Danech-Pajouh, novembre 2000, Le temps de parcours et ses deux variantes de calcul, la revue TEC n° 162
- 6 Dr.ir. H.J.M. van Grol, Hague Consulting Group, The Netherlands, dr. M. Danech-Pajouh, INRETS, France, dr. S. Manfredi, CSST, Italy, dr. J. Whittaker, University of Lancaster, United Kingdom, 1998, DACCORD : On-Line Travel Time Predictions 8th WCTR
- 7 Piet H.L. BOVY, Rammelt THIJS, Estimation of travel time, Delft University Press, 2000
- 8 CERTU, Le temps de parcours, dossiers 138, 2002
- 9 Will Recker and all. Considering Risk-Taking Behavior in travel time reliability, Institute of Transportation Studies, University of California, Irvine, janvier 2005

Adresses de quelques sites

<http://www.mapy.fr/>

<http://www.viamichelin.fr>

<http://www.mapquest.com>

<http://iww.web.de>

<http://www.sanef.com>

<http://www.transportdirect.info/> interurbain

<http://scottishexecutive.itisholdings.com/> interurbain

<http://www.sytadin.tm.fr/> urbain et suburbain

Table des annexes

Annexe A : Une procédure d'épuration des données issues des capteurs SIREDO

Annexe B : Une procédure d'estimation des vitesses critiques sur un réseau autoroutier

Annexe C : Liste des capteurs disponibles sur A4 via MI2

Annexe A : Une procédure d'épuration des données issues des capteurs SIREDO

I- Présentation des données

Les données fournies par le ministère de l'Équipement proviennent de capteurs de la norme SIREDO placés le long des axes routiers. Ces capteurs nous permettent d'accéder toutes les 6 minutes à des mesures agrégées comme le débit, la vitesse et le taux d'occupation et parfois le pourcentage de poids lourds. Ainsi pour chaque créneau horaire de 6 minutes et chaque capteur, on dispose d'une base de données [Date::Heure::Min ; vitesse brute ; débit brut (ramené à l'heure) ; taux d'occupation brut].

II- Épuration des données incomplètes

Dans la base de données ainsi constituée, chaque individu est un pas de 6 minutes identifié par sa date, l'heure et minute, il est considéré comme un point dans l'espace de trois dimensions débit, to , et vitesse.

On supprime d'office les observations qui ne disposent pas des 3 données débit, vitesse, taux d'occupation. En effet ces événements sont souvent consécutifs et ainsi laissent présager un mauvais fonctionnement du capteur. On supprimera également les individus correspondant à des vitesses ou des débits nuls. De cette nouvelle base on calcule U_{max} arrondie à la dizaine supérieure ainsi que TO_m : la plus grande valeur des taux d'occupation pour laquelle le débit correspondant est acceptable (équivalent à 50 véhicules à l'heure) ; nous assimilerons U_{max} à la vitesse libre : vitesse à laquelle un véhicule se déplacerait s'il ne subissait aucune contrainte (limitations de vitesse, congestion, ...).

III- Épuration des données aberrantes

Ensuite on observe le nuage de point vitesse*taux d'occupation.

Théoriquement il existe une relation linéaire entre le taux d'occupation et la vitesse :

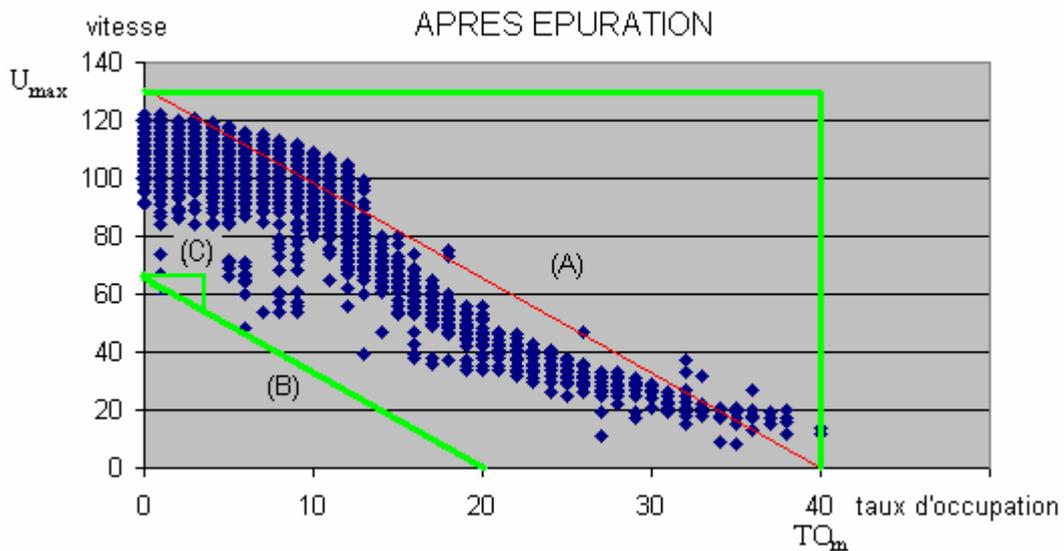
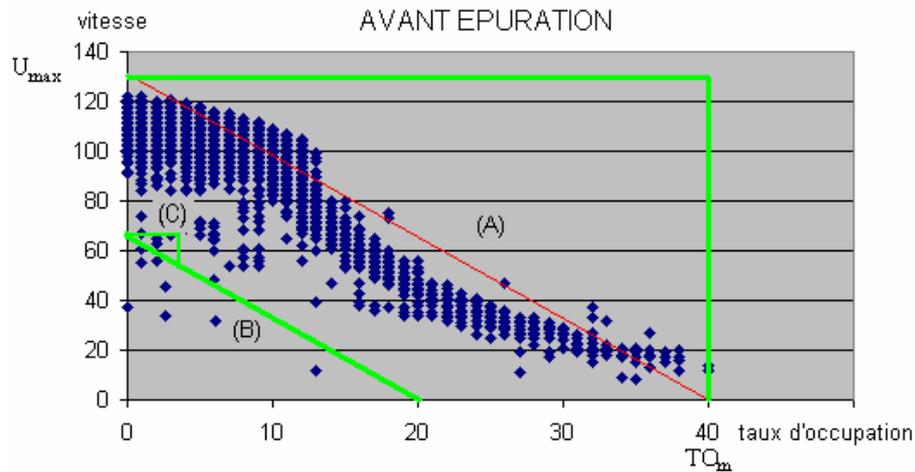
$$U = U_{max} \left(1 - \frac{TO}{TO_m}\right) \quad (A)$$

La procédure d'épuration se décrit de la façon suivante :

- On supprime de toute façon les observations extérieures à la **surface rectangulaire** délimitée par la vitesse libre U_{max} et TO_m (dans l'exemple ci-dessus cette situation ne se présente pas).
- On trace ensuite la droite (B) parallèle à (A) et ayant comme ordonnée à l'origine la moitié de la vitesse libre U_{max} . S'il existe moins de 5% de points sous cette droite (B), on les élimine et l'épuration est terminée ; dans le cas contraire, parallèlement à (A), on abaisse cette droite jusqu'à ce que 5% du nuage de points se trouvent sous cette dernière. **Au plus** 5% des points auront donc été supprimés après cette étape.
- On délimite également la région triangulaire (C) associée au 3 points de coordonnées :

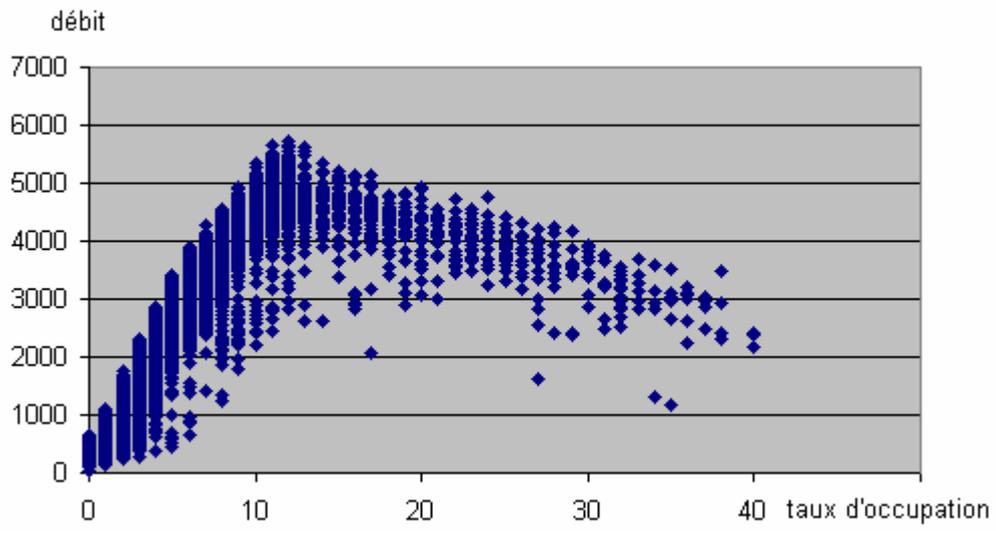
$(0 ; 0.5*TO_m)$, $(0.1*TO_m ; 0.5*TO_m)$ et $(0.1*TO_m ; 0.1*U_{max})$.

Les points sous la droite (B) ou dans le triangle (C) ont été éliminés car un taux d'occupation faible et une vitesse faible sont difficilement envisageables.



On évite ainsi des périodes de mauvais fonctionnement des capteurs et on conserve un comportement habituel de l'écoulement du trafic.

Nous aurions également pu regarder les nuages de points taux d'occupation*débit ou débit*vitesse mais il paraît plus difficile de se donner une règle de décision permettant de repérer des individus aberrants. En effet la forme des nuages obtenus est beaucoup moins évidente, théoriquement il s'agit d'une parabole.



Annexe B : Une procédure d'estimation des vitesses critiques sur un réseau autoroutier

I- Préliminaire

Définition de la capacité d'une route :

La capacité en un point d'une route (appelée également débit de saturation Q_s) est définie comme le nombre maximal de véhicules pouvant être écoulés en ce point, pendant un intervalle de temps de référence correspondant en général à l'heure. On comprend aisément qu'elle dépend des caractéristiques d'infrastructures et des habitudes de circulation constatées sur la section concernée.

Ces caractéristiques d'infrastructure désignent les facteurs géométriques, comme par exemple le nombre de voies, la largeur des voies, la présence de bandes d'arrêt d'urgence ou la pente du profil en long. L'environnement de la route précise son insertion - urbaine, périurbaine ou interurbaine -, les facteurs météorologiques ou encore les conditions de visibilité. Les caractéristiques de circulation concernent la composition du trafic par catégories de véhicules (voitures particulières, poids lourds, motos), la nature des déplacements (domicile travail, migrations de loisirs) et les conditions d'exploitation.

On définit ensuite la vitesse critique V_c comme la vitesse correspondant à des débits très proches de la capacité.

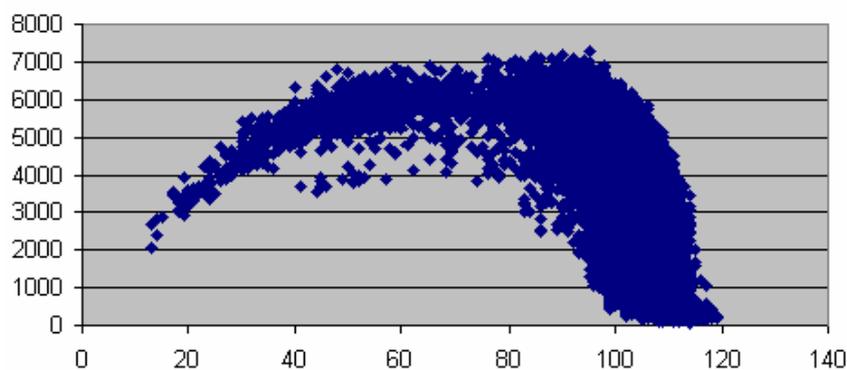
Ensuite le taux d'occupation critique TO_c s'obtient en prenant la médiane des taux observés pour des vitesses appartenant à l'intervalle $[V_c - 3\text{Km/h} ; V_c + 3\text{Km/h}]$.

Cette marge sur la vitesse se justifie par la précision des capteurs mesurant les données.

II- Estimation de la vitesse critique et de la capacité

On suppose qu'à ce niveau les données ont été pré-traitées grâce à une procédure d'épuration. Le nuage de points observé entre le débit et la vitesse a souvent l'allure globalement parabolique (cf. ci-dessous station MST67.G2 Bisheim sens 2).

Débit fonction de la vitesse



Annexe C : Liste des capteurs disponibles sur A4 via MI2

Code SIREDO	PK	Nom	Péage proche	PK
MFS93.A1	13300	Noisy le Grand Ouest		
MFS93.B1	14870	Noisy le Grand Est		
MFM77.A1	15900	Champs sur Marne		
MFM77.B1	18935	Noeud A4 RN104		
MFM77.C1	20600	Croissy Beaubourg		
MFM77.D1	23800	Collégien A4.A104		
MFM77.E1	25000	Ferrieres en Brie		
MFM77.F1	31300	Serris		
MFM77.G1	37675	Coutevroult	Coutevroult	35066
MFM77.H1	40350	Crecy la Chapelle		
MFS77.L1	54320	St Jean les Deux Jumeaux	Montreuil	64325
MFA02.11	91800	Dormans Ouest		
MFM51.B1	125980	Reims Gueux	Dormans	111302
MFM51.C1	133107	Reims Tinquieux		
MFM51.D1	136250	Reims Centre		
MFM51.E1	137700	Reims Cathédrale		
MFM51.F1	139150	Reims St Rémi		
MFM51.G1	140415	Reims Cormontreuil		
MZM51.21	157600	Châlon en Champagne	Taissy	143610
MFM51.L1	220000	Sainte Ménéhould		
MFS55.11	260720	Audiomont		
MZo57.91	305600	Saint Privat	Beaumont	294917
MFO57.B1	314600	Sémécourt		
MFO57.C1	317510	Argancy		
MFO57.D1	321092	Charly Oradour		
MZo57.81	337000	Les Etangs		
MFO57.G1	366440	Saint Avold	St Avold	361435
MFO57.H1	372450	Farébersviller		
MFB67.A1	395120	Sarre Union	Loupershouse	379700
MFB67.C1	434010	Saverne A4		
MFB67.D1	441850	Saverne Est		
MFB67.E1	459275	Brumath Nord	Schwindratzheim	456072
MFB67.F1	464740	Brumath		
MFB67.G1	469600	Brumath Sud		
MFB67.H1	472980	Hoerdt A4		
MZB67.I1	475200	Reichstett		

Les capteurs disponibles sur A4 via MI2 (sens Paris –Strasbourg)
 Les codes SIREDO des capteurs du sens opposé se terminent par le chiffre 2