



HAL
open science

Short-term multi-step ahead traffic forecasting

Luis Leon Ojeda

► **To cite this version:**

Luis Leon Ojeda. Short-term multi-step ahead traffic forecasting. Automatic. Université de Grenoble, 2014. English. NNT: 2014GRENT081 . tel-01225248

HAL Id: tel-01225248

<https://theses.hal.science/tel-01225248>

Submitted on 5 Nov 2015

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

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

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Automatique-Productique**

Arrêté ministériel : 7 août 2006

Présentée par

Luis LEON OJEDA

Thèse dirigée par **M. Carlos CANUDAS DE WIT**

et codirigée par **M. Alain KIBANGOU**

préparée au sein **GIPSA-Lab, Département Automatique**

et de **Électronique, Électrotechnique, Automatique, Traitement du Signal**

Short-term multi-step ahead traffic forecasting

Thèse soutenue publiquement le **03 Juillet 2014**,

devant le jury composé de :

M, Habib HAJ-SALEM

Directeur de Recherche IFSTARR (Marne-la-vallée, France), Rapporteur

M, Hassane ABOUAISSA

Maître de conférences HDR, Université d'Artois (Béthune, France), Rapporteur

M, Christophe BERENQUER

Professeur Grenoble-INP, (Grenoble, France), Examineur

M, Gildas BESANÇON

Professeur Grenoble-INP, (Grenoble, France), Président

M, Fabien MOUTARDE

Maître de conférences HDR, Ecole de Mines, (Paris, France), Examineur

M, Alain KIBANGOU

Maître de conférences, Chaire CNRS, Université Joseph Fourier, (Grenoble, France), Examineur



Acknowledgements

My sincere gratitude to everyone who stand by me during these last three years. My supervisors, Carlos Canudas de Wit and Alain Kibangou, everyone in NeCS team, colleagues, friends, and family.

Special gratitude to my thesis committee, whose comments greatly improved the content of this document.

Muchisimas gracias a todos.

Abstract

This dissertation falls within the domain of the Intelligent Transportation Systems (ITS). In particular, it is concerned with the design of a methodology for the real-time multi-step ahead travel time forecasting using flow and speed measurements from a instrumented freeway. To achieve this objective this thesis develops two main methodologies.

The first one, a signal-based, uses only speed measurements collected from the freeway, where a mean speed is assumed between two consecutive collection points. The travel time is forecasted using a noise Adaptive Kalman Filter (AKF) approach. The process noise statistics are computed using an online unbiased estimator, while the observations and their noise statistics are computed using the clustered historical traffic data. Forecasting problems are reformulated as filtering ones through the use of pseudo-observations built from historical data.

The second one, a model-based, uses mainly traffic flow measurements. Its main appealing is the use of a mathematical model in order to reconstruct the internal state (density) in small road portions, and consequently exploits the relation between density and speed to forecast the travel time. The methodology uses only boundary conditions as inputs to a switched Luenberger state observer, based on the “Cell Transmission Model” (CTM), to estimate the road initial states. The boundary conditions are then forecasted using the AKF developed above. Consequently, the CTM model is run using the initial conditions and the forecasted boundaries in order to obtain the future evolution of densities, speeds, and finally travel time. The added innovation in this approach is the space discretization achieved: indeed, portions of the road, called “cells”, can be chosen as small as desired and thus allow obtaining a finer tracking of speed variations.

In order to validate experimentally the developed methodologies, this thesis uses as study case the Grenoble South Ring. This freeway, enclosing the southern part of the city from A41 to A480, consists of two carriageways with two lanes. For this study only the direction east-west was considered. With a length of about 10.5 km, this direction has 10 on-ramps, 7 off-ramps, and is monitored through the Grenoble Traffic Lab (GTL) that is able to provide reliable traffic data every 15 s, which makes it possible for the forecasting strategies to be validated in

real-time.

The results show that both methods present strong capabilities for travel time forecasting: considering the entire freeway, in 90% of the cases it was obtained a maximum forecasting error of 21% up to a forecasting horizon of 45 min. Furthermore, both methods perform as good as, or better than, the average historical. In particular, it is obtained that for horizons larger than 45 min, the forecasting depend exclusively on the historical data. For the dataset considered, the assessment study also show that the model-based approach is more suitable for horizons shorter than 30 min.

Table of contents

Acknowledgements	iii
Abstract	v
1 Introduction	1
1.1 Context of the thesis	3
1.1.1 Case study-Grenoble south ring	3
1.1.2 Research Objectives	6
1.2 Literature review	7
1.3 Main contributions of the thesis	12
1.4 Dissertation outline	14
2 Travel time review	17
2.1 Overview	17
2.2 Travel time synopsis	17
2.2.1 Travel time computation from direct measurements	19
2.2.2 Travel time computation from indirect measurements	21
2.3 Derivation of the forecasted travel time	24
2.4 Conclusions	27
3 Traffic data collection and pre-processing	29
3.1 Overview	29
3.2 Grenoble Traffic Lab (GTL)	29
3.3 Data description	32
3.4 Data pre-processing	35
3.4.1 Data cleaning	36
3.4.2 Data imputation	37
3.4.3 Data aggregation	41
3.5 Conclusions	43
4 Short-term multiple step ahead travel time forecasting: signal-based approach	45
4.1 Overview	45
4.2 Data description	45
4.3 Statement of the problem	46
4.4 Noise adaptive Kalman filter	46
4.4.1 State-space model	47
4.4.2 Pseudo-observations	48

4.4.3	Pseudo-observation noise statistics	49
4.4.4	Process noise statistics	50
4.4.5	Forecasting scheme and algorithm	51
4.4.6	Study of historical data	51
4.4.7	Real-time classification of the current data	59
4.5	Experimental results: travel time forecasting	62
4.6	Conclusions	67
5	Short-term multiple ahead travel time forecasting: model-based approach	69
5.1	Overview	69
5.2	Data description	70
5.3	Statement of the problem	70
5.4	Density reconstruction	71
5.4.1	LWR model	71
5.4.2	CTM model	74
5.4.3	Traffic density estimation review	78
5.4.4	Graph constrained-CTM observer	80
5.4.5	Experimental results: state observer	89
5.5	Boundary conditions, input, and output flows forecasting	95
5.5.1	Noise adaptive Kalman filter approach (AKF)	95
5.5.2	Experimental results: input flow forecasting	97
5.6	Density and travel time forecasting	99
5.6.1	Density forecasting	99
5.6.2	Speed and travel time forecasting	100
5.7	Experimental results: travel time forecasting	101
5.8	Conclusions	106
6	Conclusions and future work	107
6.1	Comparison between the proposed forecasting approaches	107
6.2	Main contributions	109
6.3	Future research efforts	112
A	Resumé en français	113
A.1	Introduction	113
A.1.1	Cas d'étude	114
A.1.2	Objectifs généraux	115
A.1.3	Contributions de la thèse	115
A.1.4	Revue de temps de parcours	115
A.2	Calcul du temps de parcours prédit	119
A.3	Collecte et pré-traitement de données de trafic	121
A.3.1	Nettoyage de données	121
A.3.2	Imputation de données	122
A.4	Prédiction à court terme et à pas multiples de temps de parcours : approche orientée signal	124
A.4.1	Description des données	124
A.4.2	Formulation du problème	124
A.4.3	Filtre de Kalman adaptatif	124
A.4.4	Résultats expérimentaux: prédiction du temps de parcours	129

A.5	Prédiction à court terme et à pas multiples de temps de parcours : approche orientée model	132
A.5.1	Description des données	133
A.5.2	Formulation du problème	133
A.5.3	Estimation d'état	134
A.5.4	Observateur CTM à contraintes graphiques	140
A.5.5	Résultats expérimentaux: estimation d'état	143
A.5.6	Prediction de conditions aux limites et débits des rampes d'accès et de sortie	144
A.5.7	Prédiction de densité	146
A.5.8	Prédiction du temps de parcours	147
A.5.9	Résultats expérimentaux: prédiction du temps de parcours.	147
A.6	Conclusions et travaux à venir	149
A.6.1	Comparaison entre les deux méthodes de prédiction proposées	149
A.6.2	Résumé des contributions et conclusions	151
A.6.3	Travaux en cours et à venir	151

Bibliography	153
---------------------	------------

List of Tables

3.1	Available individual traffic parameters. Each time an event (detection of vehicle) occurs, this information is generated.	32
3.2	Choices of the aggregation times given by <i>Sensys</i> and available aggregate traffic parameters.	33
3.3	Vehicle detector stations installed in Grenoble south ring. “ID” is a hexadecimal serial number associated to groups of magnetometers. The communication is via fiber optics, “f”, or GPRS, “g”.	35
3.4	Data cleaning rules for the 15 s aggregate data.	36
3.5	Assessment of the data imputation algorithms.	39
4.1	Maximum error observed for 90% of the realizations for different forecasting horizons.	66
5.1	Behavior of CTM model through different interfaces (Homogeneous section).	77
5.2	Behavior of CTM model through different interfaces (Not homogeneous section).	78
5.3	Operating mode table	81
5.4	Observability for different SMM modes	84
5.5	Transition conditions for a 3-cell example section	87
5.6	Boundary status written in terms of measured flow or speed for the ideal case	88
5.7	Boundary status written in terms of measured flow or speed for the not ideal case	89
5.8	Transition conditions for a 3-cell example section using the boundary speeds	89
5.9	MAPE(%) for two scenarios considered.	99
5.10	MAPE(%) for four scenarios considered.	103
5.11	Maximum error found for 90% of the realizations for different forecasting horizons.	105
A.1	règles de nettoyage de données pour les données agrégées en 15 s.	122
A.2	Débit traversant pour une section homogène	139
A.3	Débit traversant pour une section non homogène.	139
A.4	Table avec les modes du système	141
A.5	Observabilité du modèle SMM	142

A.6 Transitions en utilisant données des vitesses pour un exemple de 3 cellules . . . 143

List of Figures

1.1	Deployment of magnetic sensors on a freeway. The circles in the pavement indicate where the sensors are located. (source www.sensysnetworks.com)	2
1.2	Stream of information for the ITS	2
1.3	panoramic view of the city Grenoble.	4
1.4	<i>The south ring of Grenoble.</i> (a) View of a stretch of the south ring, direction north-east towards Meylan (image courtesy: DIR-CE); (b) Aerial view of the interchange “Rondeau” at the west end of the south ring (left, center): this site experiences heavy traffic congestion during the morning and afternoon rush hours, which may propagate upstream for several kilometers (source: Google Maps/Satellite).	5
1.5	<i>Traveling the south ring of Grenoble.</i> (a) Spatial trajectory, and (b) time evolution of the position of a car in the south ring on Thursday, December 5, 2013 (between 19:48:00 and 19:56:26), recorded with the GPS-based smart-phone application “My Tracks”. The average speed of the car is 75 km/h.	5
1.6	Markers indicating the location of collection points in Grenoble south ring(a) and (b), Traffic behavior on September 18, 2013 expressed in a speed contour plot.	6
1.7	Sensor placement in the ramps (red disks).	6
1.8	The new link division for the model-based approach	13
2.1	Trajectory of a vehicle. Departure point $[x_1, t_1]$ and arrival point $[x_2, t_2]$	18
2.2	Temporal and spatial illustration of a section travel time	18
2.3	Different travel time collection techniques	20
2.4	Link of freeway.	22
2.5	Illustration of the types of speed. TMS=mean of speed of circles in red, SMS=mean of speed of circles in blue.	23
2.6	Scenarios considered in order to obtain a formulation of the forecasted travel time. a) Vehicle trajectory from x_p to x_0 , at the passing times t_p and t_0 respectively. b) Space discretization in volumes Δx_i . c) Scenario in which we aim to forecast \hat{t} , which is the future time the vehicle would reach the point x_1	25

2.7	Travel time forecasting scheme. A vehicle enters the road at point x_0 and time k_0 , and we are interesting to forecast the travel time between $[x_0, x_n]$, departing at multi-step ahead in the future.	27
3.1	Grenoble south ring division	30
3.2	Three-level architecture of GTL	31
3.3	File for the individual raw traffic data(a) and File for the aggregate raw traffic data(b).	33
3.4	GTL's interface used to download traffic data.	34
3.5	GTL data management levels (data collection, storage, and supply).	34
3.6	Data pre-processing strategy for offline and real-time data.	41
3.7	Example of a 24 h profiles of speed while changing the aggregation time. a)15 s raw data b), 1 min b), 5 min b), 15 min.	42
3.8	Example of a 24 h profiles of vehicle count while changing the aggregation time. a)15 s raw data b), 1 min b), 5 min b), 15 min.	42
4.1	Abstraction of a freeway section. The grey stripes represent detector stations in the mainstream.	46
4.2	Derivation of the observation noise covariance. At each time k , the variance given in the historical values and increments are computed. These variances will be considered to defined the observation noise covariance.	50
4.3	Travel time forecasting scheme.	51
4.4	15 working days travel time for the second link of Grenoble south ring.	53
4.5	Variation with respect of the number of clusters of a)RMSSTD, and b) RS. The goal is to estimate the number of clusters for the travel time dataset in Fig. 4.4.	55
4.6	15 working days travel time for the second link of Grenoble south ring.	56
4.7	Clusters for the time zone between 00:00 to 07:00 of link 2.	57
4.8	Clusters for the time zone between 07:00 to 10:00 of link 2.	58
4.9	Clusters for the time zone between 10:00 to 16:00 of link 2.	58
4.10	Clusters for the time zone between 16:00 to 19:00 of link 2.	59
4.11	Clusters for the time zone between 19:00 to 24:00 of link 2.	59
4.12	Temporal assessment of the real-time classification of the current data from different values of N.	61
4.13	Temporal assessment of the real-time classification of the current data for N=10.	61
4.14	Freeway chosen path for the numerical tests, from start: Meylan to end: Rondeau.	62
4.15	Estimated travel time experienced by a driver that traverses the entire Grenoble south ring at different departing times for 15 different working days in September 2013. From the 2nd to the 20th of September.	63

4.16	Illustrative results of the proposed forecasting approach. a) multi-step ahead forecasts at $t_0 = 08:45$. b) forecasted and measured vehicle trajectory upon speed contour at $t_0 = 08:45$. c) multi-step ahead forecasts at $t_0 = 17:15$. d) forecasted and measured vehicle trajectory upon speed contour at $t_0 = 17:15$	64
4.17	Mean and standard deviation the APE (%) for a moving window from 06:00 am to 10:00 pm for different departure times. Every bar represents the mean or the standard deviation of the realizations computed for the fifteen days. a)Mean, b)Standard deviation.	65
4.18	Empirical cumulative distribution function evaluated with the APE at different forecasting horizons. a) $\Delta = 0$ (Current time), b) $\Delta = 15'$, c) $\Delta = 30'$, d) $\Delta = 45'$	67
5.1	Schematic representation of the model-based method proposed for travel time forecasting.	70
5.2	Abstraction of a freeway section. The grey stripes represent detector stations.	71
5.3	(a) Sub-division of the spatial domain in three cells, we are interested in computing the density values at the interfaces of the cell i , (b) Concave flow function $\Phi(\rho)$	74
5.4	The fundamental diagram. The shape is determined by the parameters: φ_m - maximum capacity, v - free flow velocity, w - congestion wave speed.	74
5.5	Representation of a freeway using the CTM model. (a) Homogeneous section divided into n cells of length L_i and densities $\rho_i, i = 1, \dots, n$, (b) Not homogeneous section divided into n cells of length L_i and densities $\rho_i, i = 1, \dots, n$. Each cell is accompanied with at most one on-ramp and one off-ramp.	75
5.6	Demand and Supply functions. The intersection of both functions characterize the triangular fundamental diagram.	76
5.7	Illustration of a homogeneous freeway section. This example is used to derive the system dynamics equations.	80
5.8	Illustration of the Graph \mathcal{G} associated to the constrained-CTM model, for an example of 3-Cells. The left Figure show the general case with all possible modes. The center figure show the graph path for when congestion moves upwards, and the right figure show the graph path when the congestion moves downwards.	81
5.9	Graphical representation of the CTM model calibration algorithm. a) compute the dividing point $(\bar{\rho}, \bar{\varphi})$, b) compute the free flow speed v , c) compute the congestion speed w , and d) with the intersection of the two lines compute the critical density and capacity (ρ_c, φ_m)	86

5.10 Relationship between the free flow speed v (mean-space speed), and the time-mean speeds measured at the cell's boundaries. The flow velocity $v(x,t)$ describes the velocity evolution of the cell. This evolution is modelled as constant v in the CTM model.	88
5.11 Freeway link considered for the observer validation. This link has a length of 2.13 km.	90
5.12 Results of the estimation when increasing the number of cells in the link. Each plot compares the measured density with the reconstructed one inside the cell that encloses the mainstream sensors. The number of cells chosen are: a)1, b)2, c)5, and d)10.	91
5.13 Empirical CDF showing the absolute error obtained, in the congested area, when dividing the considered link in an increasing number of cells.	92
5.14 Freeway division considering discretization in cells.	93
5.15 Contours depicting the 24 h estimation results for different days of the dataset: a),b) Thursday: 27th of February 2014, c),d) Friday: 28th of February 2014, e),f) Friday: 7th of March 2014, g),h) Friday: 14th of March 2014. Measured density on the left column, reconstructed ones on the right. Colors denote density in <i>veh/km</i>	94
5.16 AKF scheme.	95
5.17 5 min traffic information collected from Grenoble south ring from the 27th of February to the 15th of March of 2014. They represent: a) input flows from on-ramp 6, b) output flows from off-ramp 2, c) mainstream flow from upstream boundary of link 12, d) mainstream speed from upstream boundary of link 12. See Fig. A.19 to visualize the positions. This plot aims to highlight the high correlation within the flow profiles compared to those of speed.	97
5.18 Results of the knee criteria and clustering of a input flow dataset. a)RS, b) RMSSTD, c) Cluster 1, d) Cluster 2.	98
5.19 Results of the knee criteria and clustering of a input flow dataset. a)RS, b) RMSSTD, c) Cluster 1, d) Cluster 2.	99
5.20 Illustration of a not homogeneous freeway section. This example is used to present the equations for the density and interface flow forecasting.	100
5.21 Freeway section considered for the validation of travel time forecasting. This section has a length of 3.3 km.	101
5.22 Estimated travel time experienced by a driver that travels from Gabriel Peri (entrance 2) to SMH Centre (exit) at different departing times. From 27th of February to the 15th of March of 2014.	102

5.23	Illustrative results of the proposed forecasting approach. a) multi-step ahead forecasts at $t_0 = 16:00$. b) forecasted and measured vehicle trajectory upon speed contour at $t_0 = 16:00$. c) multi-step ahead forecasts at $t_0 = 18:30$. d) forecasted and measured vehicle trajectory upon speed contour at $t_0 = 18:30$. c) multi-step ahead forecasts at $t_0 = 19:30$. d) forecasted and measured vehicle trajectory upon speed contour at $t_0 = 19:30$	104
5.24	Cumulative distribution function evaluated with the APE at different forecasting horizons. a) $\Delta = 0$ (Current time), b) $\Delta = 15'$ 2, c) $\Delta = 30'$ 3, d) $\Delta = 45'$	105
6.1	Estimated travel time experienced by a driver that travels from Gabriel Peri (entrance 2) to SMH Centre (exit) at different departing times. From 27th of February to the 15th of March of 2014. The days pointed with arrows are the days that will be considered for the numerical comparison between the two proposed forecasting methods.	107
6.2	Cumulative distribution function evaluated with the APE at different forecasting horizons. a) $\Delta = 0$ (Current time), b) $\Delta = 15'$ 2, c) $\Delta = 30'$ 3, d) $\Delta = 45'$	108
A.1	Flux d'information pour les ITS	113
A.2	Différentes techniques pour la collecte de données de temps de parcours.	116
A.3	Lien d'autoroute.	118
A.4	scénarios envisagés afin d'obtenir une formulation du temps de parcours prédit. a) la trajectoire du véhicule à partir de x_p à x_0 , aux moments de t_p et t_0 respectivement. b) l'espace discrétisation en volumes Δx_i . c) Scénario dans lequel nous cherchons à calculer \hat{t}	120
A.5	Abstraction d'une section de l'autoroute. Les bandes grises représentent les stations de détection dans les voies principales.	124
A.6	Le calcul de la covariance du bruit d'observation. A chaque pas k , la variance des données historiques est calculées. Ces écarts seront considérés comme la variance du bruit d'observation.	127
A.7	Schème AKF pour la prediction du temps de parcours.	128
A.8	Chemin choisi pour les tests numériques, du début: Meylan à la fin: Rondeau.	130
A.9	Temps de parcours expérimenté par un conducteur qui traverse la Rocade sud de Grenoble aux moments différents de départ pour 15 jours ouvrables différents du 2 Septembre 2013 au 20 Septembre.	130
A.10	Résultats de prédiction pour la méthode proposée pour différents scenarios. a) $t_0 = 08h45$. b) trajectoires prédites et mesurées sur le contour de vitesse à $t_0 = 08h45$. c) $t_0 = 17h15$. d) trajectoires prédites et mesurées sur le contour de vitesse à $t_0 = 17h15$	132
A.11	Schéma de l'algorithme de prédiction orienté modèle.	133

A.12	Abstraction d'une section de l'autoroute. Les bandes grises représentent les stations de détection.	134
A.13	(a) Sous-division du domaine spatial en trois cellules, nous sommes intéressés dans le calcul des valeurs de densité au niveau des interfaces de la cellule i , (b) Fonction de débit concave $\Phi(\rho)$	135
A.14	Le diagramme fondamental. La forme est déterminée par les paramètres: φ_m - capacité maximum, v - vitesse libre, w - vitesse de congestion.	136
A.15	Représentation d'une autoroute en utilisant le modèle CTM. (a) section homogène divisée en n cellules de longueur L_i et densités ρ_i , $i = 1, \dots, n$, (b) section homogène divisée en n cellules de longueur L_i et densités ρ_i , $i = 1, \dots, n$. Chaque cellule est accompagnée d' au plus une rampe d'accès et une rampe de sortie.	136
A.16	Les fonctions de la demande et de l'offre.	138
A.17	Section considérée afin de dériver les équations dynamiques du système.	141
A.18	Illustration des règles de transition associées au graphe \mathcal{G} pour l'observateur CTM à contraintes graphiques. Exemple de 3 cellules.	142
A.19	Emplacement expérimental considéré pour la reconstruction d'état.	144
A.20	Résultats de l'observateur de densité pour différents jours: a),b) Vendredi: 28 Février 2014, c),d) Vendredi: 7 Mars 2014. Densité mesurée à gauche, densité reconstruite à droite.	145
A.21	schéma AKF pour prédiction de débit.	145
A.22	Profils de débit normalisés et agrégés à 5 min recueillies de la Rocade Sud de Grenoble du 27 Février au 15 Mars de 2014: a) Rampes d'accès 6, b) voie principal.	146
A.23	Illustration d'une section d'autoroute non homogène. Cet exemple est utilisé pour présenter les équations pour la prédiction des densités et des débits à l'interface.	146
A.24	Section utilisée pour la validation de l'approche de prédiction du temps de parcours approche orienté modèle.	147
A.25	Temps de parcours entre Gabriel Péri (entrée 2) et SMH Centre (sortie) pour différentes heures de départ. Du 27 Février au 15 Mars de 2014.	148
A.26	Résultats de l'approche de prédiction proposée. a) Prédiction à pas multiple à $t_0 = 18:30$. b) trajectoires prédites et mesurées sur contour de vitesse à $t_0 = 18:30$	149
A.27	Temps de parcours estimée entre Gabriel Péri (entrée 2) et Centre SMH (sortie) pour différentes heures de départ. Du 27 Février au 15 Mars 2014. Les jours de pointes de flèches sont les jours qui seront pris en considération pour la comparaison numérique.	149
A.28	Fonction de distribution cumulative évaluée avec les APE à différents horizons de prédiction. a) $\Delta =$, b) $\Delta = 15'$ 2,c) $\Delta = 30'$ 3, d) $\Delta = 45'$	150

Chapter 1

Introduction

Time is the longest distance between two places.

Tennessee Williams (1911-1983),
American playwright.

Rankings of the world's most liveable cities are released every year by several research institutions. The classification are based upon indexes referring to living conditions, which themselves depend on factors such as transportation services, urban design, economic development, and environmental conditions, among others. In particular, traffic congestions can have dramatic negative impact on a city such as productivity, energy waste, and pollution.

The evident way of solving this problem is to build new transportation infrastructure. Unfortunately, that is infeasible in most cities due to high costs and space limitations. Therefore, science needs to play a key role in the improvement of life conditions. In view of this, *Intelligent Transportation Systems* (ITS) has emerged in the 80's in order to increase the efficiency of transportation systems through the use of advanced technological and communication systems, and robust mathematical models.

Over the past decade, major progresses have taken place in the field of ITS. Massive traffic problems pushed public authorities to multiply the initiatives dealing with intelligent freeway systems that provide a continuous flow of information about the time evolution of traffic conditions (see as an example Fig. 1.1, where magnetic sensors are embedded in the pavement of a highway in order to collect traffic data). With the progress of traffic collection technologies, ITS applications such as collision avoidance, variable speed limit, and traffic forecasting are nowadays used and upgraded.

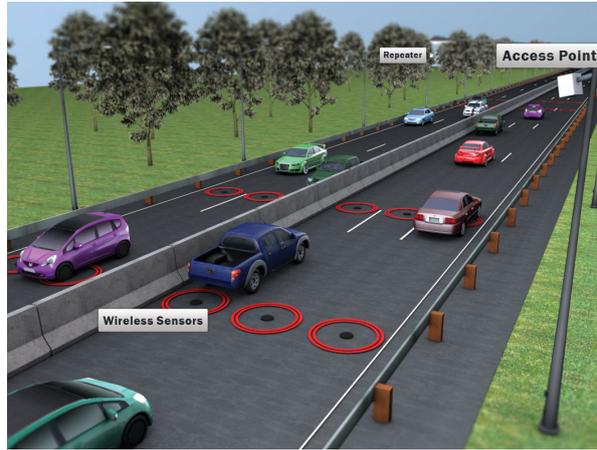


Figure 1.1: Deployment of magnetic sensors on a freeway. The circles in the pavement indicate where the sensors are located. (source www.sensysnetworks.com)

From a higher level, the typical stream of information for the ITS (see Fig. A.1) is the following:

1. Data collection: it embraces all existing data collection technologies.
2. Communication: it corresponds to the step where the collected traffic data are sent through a communication channel that varies according to the available technology.
3. Processing: in this step the algorithms developed by the researchers come into play using real traffic data for control, estimation, monitoring, and forecasting.
4. Serving: It consists in applying in real situations the technological solutions, either on *Advanced Traffic Management Systems* (ATMS) or on *Advanced Traveller's Information Systems* (ATIS).

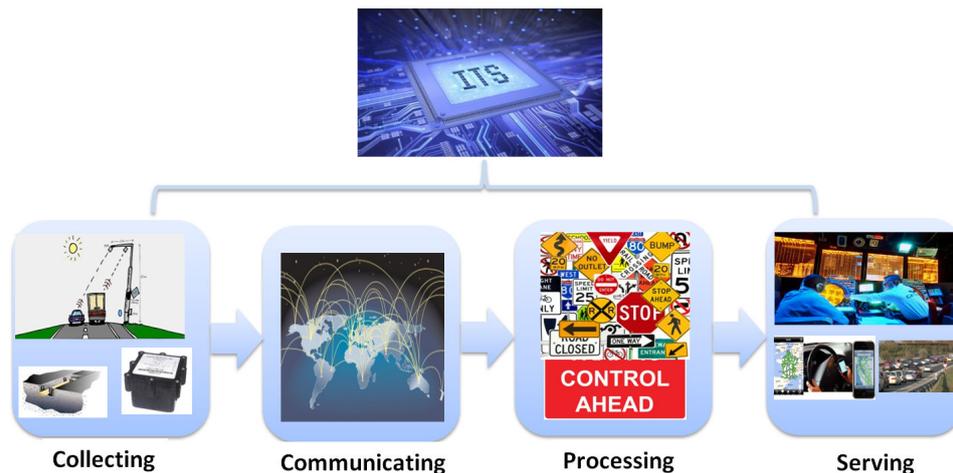


Figure 1.2: Stream of information for the ITS

Within this framework, this dissertation deals with the forecasting of freeway travel time by making use of flow and speed measurements collected by magnetic sensors. The aim is to

design short-term forecasting strategies to provide in real-time predictive travel times for several candidate routes.

1.1 Context of the thesis

This thesis is part of two research projects: The French project MOCoPo¹ and the European project Hycon2². Both projects deal with the estimation and forecasting of traffic where the Grenoble south ring constitutes the use case.

MOCoPo: project that expanded from January 2011 to December 2013. It was funded by the *French Ministry in charge of transport*, through the *Programme de recherche et d'innovation dans les transports terrestres (PREDIT)*. MOCoPo specifically deals with traffic and pollution data and modeling. The task involving the Research Institute INRIA³ via the NeCS team⁴, where this thesis has been prepared, is the forecasting of the travel time between an input and an output of the freeway.

Hycon2: European network of excellence coordinated by CNRS⁵ where the Grenoble south ring is one of the use-cases. It expands from September 2010 to September 2014, involving several academic partners. The objective of this project is the development of algorithms to attain the traffic control and travel time forecasting. This project also aims at building a common testing platform, upon which all partners can test their technological developments. The NeCS team is in charge of providing a robust and “standardized” testbed for the validation of all the partners’ theoretical work.

1.1.1 Case study-Grenoble south ring

Grenoble covers an area of 18.44 km² and with its 157424 inhabitants (in 2011) is the 16th largest city in France. The city is relatively flat, with an average elevation of 221 meters (see Fig. 1.3).

¹Measuring and modelling traffic COngestion and POLLution-<http://mocopo.ifsttar.fr/>

²Highly-complex and networked control systems-<http://www.hycon2.eu/>

³Institut National de Recherche en Informatique et en Automatique

⁴<http://necs.inrialpes.fr/>

⁵French National Center of Scientific Research



Figure 1.3: panoramic view of the city Grenoble.

The surface circulation is made difficult by the presence of mountains enclosing the city in the north, west, and south-east sides, and by the confluence of rivers Isère and Drac in the north-west side in the direction of Lyon. These natural boundaries have prevented the construction of a freeway surrounding the overall city until today, thus it makes vehicle circulation problematic especially during the peak hours. The *south ring of Grenoble* (a.k.a. “route nationale 87”) is a highway enclosing the southern part of the city from A41 to A480 (see Fig. 1.4). It consists of two carriageways with two lanes. For this study, only the direction east-west was considered (the carriageway on the left in Fig. 1.4-a). In this direction, there are 10 on-ramps and 7 off-ramps, and it stretches between the commune of Meylan (45.20531° N, 5.78353° E), and the interchange Le Rondeau (45.15864° N, 5.70384° E), for an overall length of about 10.5 km, (cf. Fig. 1.5). The south ring is a crucial transportation corridor for Grenoble: around 90000 vehicles (cars, vans, trucks, buses, etc.), with peaks of 110000, drive across it every day in both directions. The highway is operated by the DIR-CE (“Direction Interdépartementale des Routes Centre-Est”, www.dir-centre-est.fr), and the speed limit ranges between 70 and 90 km/h. The travel time from Meylan to Rondeau can vary in general from 7 minutes to 45 minutes.

The data collection is achieved through detector stations installed in the freeway Fig. 1.6-a. The direction chosen presents more interesting conditions as the topology of its last portion, Le Rondeau (cf. Fig. 1.5-b), makes drivers to behave in such way that strong congestions are created. Specially during morning and afternoon rush hours. As an example let us consider Fig. 1.6-b. The figure reports the speed contour plot of the south ring for the fast and slow lanes



Figure 1.4: *The south ring of Grenoble.* (a) View of a stretch of the south ring, direction north-east towards Meylan (image courtesy: DIR-CE); (b) Aerial view of the interchange “Rondeau” at the west end of the south ring (left, center): this site experiences heavy traffic congestion during the morning and afternoon rush hours, which may propagate upstream for several kilometers (source: Google Maps/Satellite).

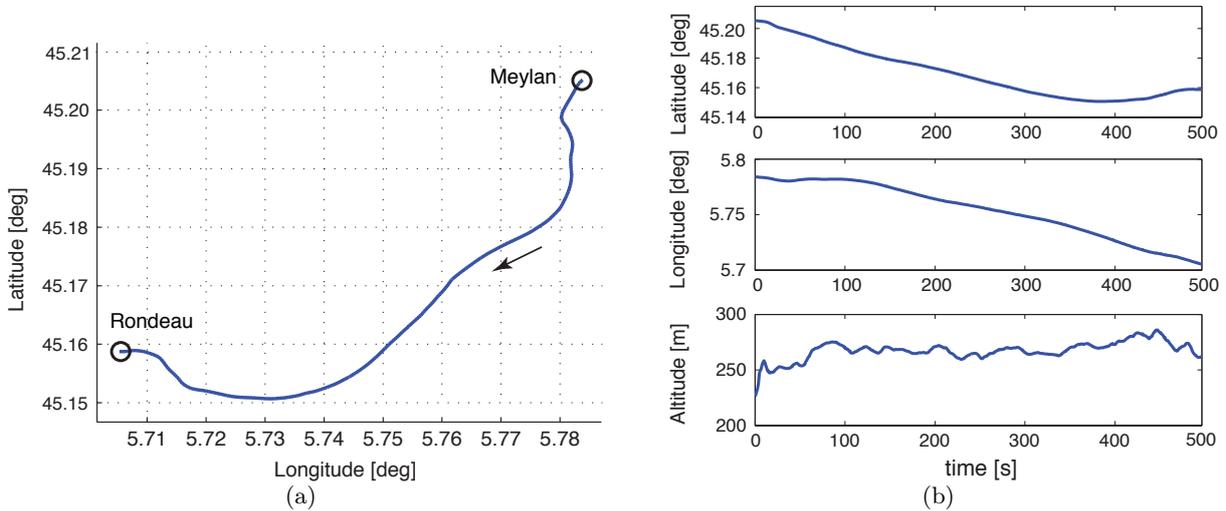


Figure 1.5: *Traveling the south ring of Grenoble.* (a) Spatial trajectory, and (b) time evolution of the position of a car in the south ring on Thursday, December 5, 2013 (between 19:48:00 and 19:56:26), recorded with the GPS-based smart-phone application “My Tracks”. The average speed of the car is 75 km/h.

for Wednesday, September 18, 2013: as it is evident in the figure (horizontal red stripes). Heavy congestions originating from Le Rondeau interchange are experienced during the morning and afternoon peak times.

The freeway sensing and actuation equipment includes: 130 wireless magnetic sensors, 4 junctions with in-ramp queue measurements, and 7 variable speed limit electronic panels (70-90 km/h).

The sensing instrumentation of the freeway was done from October 2012 to May 2013. The sensors are *Sensys Networks* VDS240 wireless magneto-resistive embedded in the pavement along the fast lanes, the slow lanes, and on/off ramps. The stations are constituted by three

- dynamic estimation of traffic density;
- dynamic forecasting of flows, speed, and travel time using historical and current data;
- Assessment of the estimation and forecasting algorithms with real traffic data.

1.2 Literature review

Nowadays as larger and larger amounts of “real-time” data are available the list of forecasting approaches has enriched substantially and it would be a big challenge to cite all works. Therefore, the author of this thesis will shorten the cited contributions to the most significant ones according to the objectives and the context of this thesis.

For this, the methodologies used for traffic forecasting will be detailed. By traffic forecasting we refer to the forecasting of any of the traffic parameters treated so far: travel time, speed, flow, density, and occupancy.

This discussion includes items that will be cited frequently in the document. These are listed below:

- **Sampling period:** the period with which the data are collected by the sensors.
- **Latency time:** the time for a data to arrive to the traffic center.
- **Raw data:** is the data collected from the sensor that has not been subjected to processing or any other manipulation.
- **Individual data:** information characterizing one single vehicle in the traffic stream.
- **Aggregate data:** individual data combined for a period of time.
- **Aggregation time:** the period of time upon which the raw data is assembled.
- **Forecasting step:** the time interval upon which the forecasts are made.
- **Forecasting horizon:** the extent of time ahead to which the forecast is referring.
- **Type of input:** the traffic parameter upon which the forecast is based (travel time, flow, speed, density, occupancy).
- **Same day data:** traffic data of the current day.
- **Current data:** data of the current day at the current time.
- **Same day past data:** same day data excluding current data.
- **Historical data:** traffic data of preceding days.

- **Forecast algorithm:** the methodology for forecasting.
- **Real-time algorithm:** algorithm able to provide forecasts when every new current data becomes available.

A major part of the literature reviewed focuses on the traffic flow forecasting problem. Although it is not clear why. We suggest that is because: first, authors tend to conclude that the forecasts based on flows are more stable [26, 38, 46] and second, traffic flow is more valuable in traffic control and ATMS applications. Nevertheless, we will cite the works regardless the type of forecasting input, keeping in mind that it can be applied also to travel time.

In the sequel, we will divide the methodologies according to the type of information used. First, those that use only same day data. Second, those using only historical data, and at last, the ones using a mix of both.

The use of historical information takes advantage of repetitive patterns in traffic, allowing forecasting at longer horizons. However, the exclusive use of this information does not allow the algorithms to react to current conditions. The use of same day data, on the other hand, overcome this problem, but at the cost of being limited for shorter forecasting horizons. When the mix of the information is used, the algorithms can capture current dynamics as well as enlarging their forecasting horizons.

Methodologies using exclusively historical data:

- Auto-regressive integrated moving average (ARIMA): it is an auto-regressive model used to better understand the data or forecast future points of the time series, it was first introduced in [3, 46]. These models handle time-correlated forecasts by exploiting the autocorrelation structures in time series data. The basic notation for the ARIMA model is ARIMA(p,d,q) and the underlying regression equation is as follows (Eq. 1.2.1):

$$(1 - B)^d \Upsilon_p(B) Z_t = \theta_q(B) a_t \tag{1.2.1}$$

where:

B = backshift operator defined by $B^j Z_t = Z_{t-j}$,

$\Upsilon_p(B) = 1 - \Upsilon_1 B - \dots - \Upsilon_p B^p$ = autoregressive (AR) polynomial of order p ,

$\theta_q(B) = 1 - \theta_1 B - \dots - \theta_q B^q$ = moving average polynomial of order q ,

$(1 - B)^d$ = order for the differentiation of Z_t , $((1 - B)Z_t = Z_t - Z_{t-1})$

Z_t = value of the series at time t ,

a_t = white noise with zero mean and variance σ^2 .

Some shortcomings are worth pointing out when using this approach. First, it needs an important historical database for parameters fitting. The fitting is not always straightforward when many parameters are involved. Second, it has problems for capturing extreme

values, i.e. it concentrates on the mean value of the time series, thus it can filter out some important characteristics. When a good model is fitted it presents satisfactory results. This has been mostly used to forecast flow and speed [23, 36, 82].

- Seasonal Auto-regressive integrated moving average (SARIMA): it is a further expansion from the ARIMA model when seasonal or cyclic components exist in the time series. The basic notation for a seasonal model with period S is ARIMA(p,d,q)(P,D,Q) and the equation is:

$$(1 - B)^d \Upsilon_p(B) (1 - B^S)^D \Psi_P(B^S) Z_t = \theta_q(B) \Theta_Q(B^S) a_t \quad (1.2.2)$$

where the additional terms represent:

$\Psi_P(B^S) = (1 - \Psi_1 B^S - \dots - \Psi_p B^{pS})$ = seasonal autoregressive polynomial of order P,
 $\Theta_Q(B^S) = (1 - \Theta_1 B^S - \dots - \Theta_Q B^{QS})$ = seasonal moving average polynomial of order Q,
 $(1 - B^S)^D$ = polynomials for seasonal differencing.

In traffic flow, for instance, the notion that seasonality could be achieved by a weekly seasonal difference was first recognized by [66]. The major drawback in this approach is that it can require an important amount of time and effort to fit and maintain a seasonal ARIMA in a real ITS system. The advantage is that once the model is calibrated it can produce accurate multi-step ahead forecasts. This is a widely used method for flow forecasting [81, 93].

Methodologies using exclusively same day data:

- Smoothing techniques: they are applied to time series under the assumption that the current data depends on the same day past data. These strategies are usually classified in two groups: moving average and exponential smoothing. The former averages the past observations. The latter assigns exponentially decreasing weights over time to the past observations. Since its forecast pattern is linear, this approach tends to not perform well for multiple-step ahead forecasts. However, they are very simple to implement and deterministic. This technique has been used to forecast flow, occupancy, travel time, and speed [10, 52, 71, 78, 94]. To cite an example, in [10] it was used a Holt's exponential smoothing implemented by using the following formulation:

$$\begin{aligned} S(k) &= \alpha Y(k) + (1 - \alpha)(S(k - 1) + b(k - 1)) \\ b(k) &= \beta(S(k) - S(k - 1)) + (1 - \beta)b(k - 1) \\ F(k + \Delta) &= S(k) + b(k)\Delta, \end{aligned}$$

where α and β are the smoothing constants, $S(k)$ is the smoothed value at the end of period k , $b(k)$ is the smoothed trend in period k , and Δ is the forecasting horizon.

- Kalman approach: related to a state-space model. It was gradually adopted by statisticians as a tool of time series modelling. In the forecasting framework, the limitations in this approach are the dependency of the noises statistics knowledge and the inability to forecast more than one-step ahead accurately. Nevertheless, if the correction step is possible, usually by assuming the observation as the last estimated value under a stationary hypothesis, it can be easily implementable, allowing multi-step ahead forecasts. Furthermore it is able to handle multiple parameters (as having a multivariate nature). This approach has been used to forecast, at multi-step ahead several, traffic parameters, flow, occupancy, travel time, and speed [13, 14, 66, 82, 96].

Methodologies using current day and historical information:

- Adaptive seasonal ARIMA: The bottom line is that a seasonal ARIMA model is fitted to a set of data, then some parameters of the model are changed in real-time if the historical and same day data draw away. This framework is of great interest given the inadequacy of the SARIMA to capture difference between same day and historical traffic conditions. The contributions have been presented using different directions. In [11] the expectation maximization (EM) and cumulative sum algorithms were used to monitor real-time shifts in the process mean for each new current data. The statics of the process noises are then adapted according to eventual detected changes. In [95] on the other hand, the authors used an adaptive predictor based on Kalman approach to adjust the forecasting error every new current data. The drawback in these approaches are the need to have fitted a seasonal ARIMA to a set of data, nevertheless once done, they are more reactive to current conditions with a high accuracy.
- Data-driven methodologies: They are based on the principle of pattern recognition, with the main purpose of identifying clusters of data with behavior similar to current traffic state [88]. They cover a large part of the literature. The most common strategies found are: nearest neighbor approaches and neural networks. Their limitation are the extensive and the good quality of the historical dataset needed to achieve a good forecasting, nevertheless if it is available, they can produce accurate multiple step-ahead forecasts with relatively low effort. Data-driven strategies are used to forecast several traffic parameters, flow, occupancy, speed, and travel time [1, 10, 14, 16, 17, 26, 33, 34, 43, 80, 81].
- Hybrid methods: They are based on the combination of two methods to achieve the forecast. In general a first method is used to reduce the problem's dimensionality, by applying clustering techniques, whereas a second one forecasts the time series using a known approach on the reduced data. The most common approaches are: ATHENA model (where a layered statistical approach clusters the data, and a linear regression model is applied to each cluster), and KARIMA model (Kohonen self-organizing map is

used for clustering and an ARIMA model is fitted then to each cluster). The most important drawback seen in these approaches is that the final forecast carries the inaccuracy of each individual method. Nevertheless, it takes the best of each one. These approaches have been adopted for flow forecasting [22, 24].

- Other methods: This category includes some methods that are not clearly classified into the previous categories:

- Gaussian Maximum Likelihood (GML): approach proposed by Lin in [51, 54]. It makes use of both historical and current data in an integrated way by using two key variables level and increment. GML was originally developed for flow and occupancy forecasting. Its major drawback is the inability to capture eventual dissimilarity between the historical and the current data, however it is quite effective and accurate for real-time implementations with multi-step ahead forecasting capabilities. The predictor is as follows.

Let $\varphi^h(k)$, $h = \{1, 2, \dots, D-1\}$ be consecutive observations of the traffic flow obtained at time k of day h . Let $\Phi^h(k) = \varphi^h(k) - \varphi^h(k-1)$ be the flow increment. Assuming that these two variables are normally distributed, an estimate for the flow in the next period k of the current day D was derived by maximizing the product of the two probability density functions of $\varphi^h(k)$ and $\Phi^h(k)$, yielding the predictor:

$$\hat{\varphi}^D(k) = \frac{\sigma_\varphi^2(k)(\mu_\Phi(k) + \hat{\varphi}^D(k-1)) + \sigma_\Phi^2(k)\mu_\varphi(k)}{\sigma_\varphi^2(k) + \sigma_\Phi^2(k)},$$

where $\mu_\varphi(k)$ and $\sigma_\varphi^2(k)$ are the mean and variance of $\varphi^h(k)$, respectively, and $\mu_\Phi(k)$ and $\sigma_\Phi^2(k)$ are the mean and variance of $\Phi^h(k)$, respectively. These means and variances are estimated from the historical data.

- Decay factor predictor: it was initially proposed for flow forecasting in [15]. This deterministic methodology is based on a combination between the current data, historical data, and a forgetting factor. The forecasted flow begins from the current flow data, the consecutive forecasted values tend to the average historical using a forgetting factor. This factor acts on a determined forecasting horizon. This approach succeeded to capture important difference between the historical and same day data. The predictor is as follows.

Let $\varphi^h(k)$ be the flow historical data for $h = \{1, 2, \dots, D-1\}$, D the current day, and k_0 the current time, then:

$$\hat{\varphi}^D(k) = \bar{\varphi}^h(k) + K(\Delta\varphi^D(k_0)),$$

with

$$\Delta\varphi^D(k_0) = \varphi^D(k_0) - \bar{\varphi}^h(k_0),$$

$$\text{and } K = \begin{cases} \eta(1 - \frac{\tau}{\Delta k_{\max}}), & \text{if } 0 < \tau \leq \Delta k_{\max} \\ 0, & \text{if } \tau > \Delta k_{\max} \end{cases},$$

$\bar{\varphi}^h(k)$ being the flow historical average, η being the weight of the mismatch between the current and the historical data, and Δk_{\max} the maximum horizon for the decay factor.

Although an important part of the literature has tried to compare the performance of the different forecasting approaches, there is still not a clear idea which one gives the best solution in terms of accuracy. Indeed, many variables are involved in the process, and not all works make use of the same basis of comparison; statistical and qualitative measures for comparing the algorithms being generally different. However, after its publication, the GML approach has been compared with other predictors, such as regression models, ARIMA, neural networks, smoothing techniques, concluding that GML presented the best forecasting performance, under typical traffic conditions [10, 39, 86].

1.3 Main contributions of the thesis

This dissertation contains contributions to the problem of travel time forecasting, traffic density estimation, and treatment and clustering of traffic data.

- **Travel time forecasting:** This dissertation proposes two approaches to solve the **short-term multi-step ahead travel time forecasting** problem, using in an integrated way the historical and the same day data.
 - Signal-based: This approach is based on a noise adaptive Kalman filter (AKF) scheme and uses only speed measurements. It transforms the forecasting problem to a filtering one, by using as system's observations a suitable combination of the historical data.
 - Model-based: This approach is based on the macroscopic traffic model Cell Transmission Model (CTM) and uses speed and flow measurements. This methodology exploits the relationship between traffic density and speed. It divides the section of interest in smaller road portions than the links, called cells (cf. Fig. 1.8). With this new division it aims to track with higher precision the traffic conditions of the section.

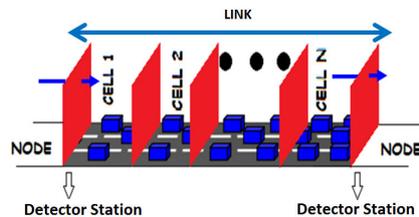


Figure 1.8: The new link division for the model-based approach

- **Traffic data treatment:** given that the traffic data is corrupted by erroneous values and missing data, we propose an algorithm, applicable also in real-time, for extracting the false samples and impute all missing values.
- **Traffic data clustering:** the travel time profiles computed from speed data present a strong variability within a dataset. To overcome this issue and have a consistently clustered historical dataset, we propose a new clustering approach. This approach first groups profiles into working and not working days, then for each group the following is applied: divide all the profiles in 5 time zones, chosen according to traffic conditions, then apply the k-means algorithm to each of the zones.

The forecasting schemes meet the following properties:

1. Reliability: they give consistent and realistic multi-step ahead forecasting results.
2. Transferability: they are able to be implemented across various locations, they are “plug-and-play” in nature.
3. Computation efficiency: they guarantee a fast response with a low computational burden and limited memory.
4. Adaptability: they are self-tuning to the incoming data stream.
5. Robustness: they handle incoming erroneous and missing data, caused by the unavoidable errors in the traffic data collection system.

The main contributions are summarized in the following publications:

Proceedings of peer-reviewed international conferences

- C. Canudas de Wit, L. Leon Ojeda, and A. Kibangou, **Graph constrained-CTM observer design for the Grenoble south ring**. In proceedings of the 13-th IFAC Symposium on Control in Transportation Systems, Sofia, Bulgaria, September 2012.
- L. Leon Ojeda, A. Kibangou, and C. Canudas de Wit, **Adaptive Kalman Filtering for Multi-Step ahead Traffic Flow**. In proceedings of the 2013 American Control Conference, Washington, DC, USA, June 2013.

- L. Leon Ojeda, A. Kibangou, and C. Canudas de Wit, **Online Dynamic Travel Time Prediction Using Speed and Flow Measurements**. In proceedings of the 2013 European Control Conference, Zurich, Switzerland, July 2013.
- F. Morbidi, L. Leon Ojeda, C. Canudas de Wit, I. Bellicot, **Robust mode selection for accurate traffic density estimation**, To appear in proceedings of the 2014 European Control Conference. Strasbourg, France.

Technical reports

- L. Leon Ojeda, D. Pisarski, and C. Canudas de Wit, **Preparation of the Traffic Modeling, Estimation and Control show case application** , D5.1.1 Deliverable HYCON2 project, 24 August 2011.

1.4 Dissertation outline

The rest of the dissertation is organized as follows:

Chapter 2

This chapter exposes the study of the traffic data collected from GTL. It aims to present first, the GTL platform, second, what type of data is collected and managed by GTL. Third, it presents the proposed strategies to tackle problems such as, data missing, erroneous data, data imputation, and data aggregation.

Chapter 3

This chapter aims to give first an overview of the technologies to collect travel time data, and second the scheme of how the forecasted travel time in a section is computed. It also aims to identify what type of travel time will be considered in this thesis.

Chapter 4

In this chapter, we present the signal-based method for the travel time forecasting problem. First, we describe the data considered for this study. Second, we derive the complete structure of the AKF scheme. Third, we present and validate the clustering of the travel time profiles. Then, we explain how the current data will be assign to a specific cluster in order to extract all the statistic information. And finally, we address the experimental validation of the signal-based method.

Chapter 5

The development of the model-based method is addressed in this chapter. Here, we describe the data considered for this study. Second, we present and validate the density estimation problem under the graph constrained-CTM observer. Third, we address the speed and flow forecasting using the same AKF scheme approach, numerical results will also be provided. Then, we show how the future evolution of densities and travel times in the cells are obtained from the CTM model. Finally, we provide experimental validation of this method.

Chapter 6

This chapter summarizes the results of the work. The first part will provide a numerical comparison between the forecasting methodologies developed in this thesis. The second part will summarize the main contributions. Finally, the last part will present some possible future perspectives.

Chapter 2

Travel time review

2.1 Overview

The main objective of this chapter is to give a general view of the traffic parameter: travel time. For this, in Sec. 2.2, we first present the general concept of the travel time in a road section, and then we highlight the numerous means to obtain experimental travel time data. In Sec. 2.3, we derive how the computation of the forecasted travel time at multiple steps ahead is achieved. In particular, we aim to establish a clear difference in the type of travel time that is used in this dissertation.

2.2 Travel time synopsis

Travel time is defined as the time needed to traverse a road between two points of interest. The interest in travel time can be traced back to several decades, but in the early 90's it has emerged as a key part in ITS systems. It is one of the most understood, clear, and communicated traffic measure indices, used by a wide variety of users including transportation engineers, planners, and consumers. It gives a clear idea of the road state and congestion conditions.

In order to present how the travel time for a road section during a given time interval is computed, let us take a measured vehicle trajectory as shown in Fig. 2.1.

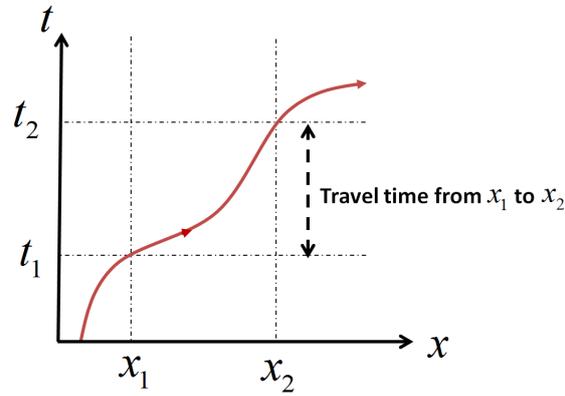


Figure 2.1: Trajectory of a vehicle. Departure point $[x_1, t_1]$ and arrival point $[x_2, t_2]$.

Clearly, the travel time experienced by that specific driver in the space interval $[x_1, x_2]$ is the difference between the times t_2 and t_1 . Now, if in the space and time intervals, $[x_1, x_2]$ and $[t_1, t_2]$ respectively, several vehicles passed, see Fig. 2.2, the travel time that characterizes the road section is computed by averaging the travel time of the vehicles that travelled the designated road during the time period of interest [75], (solid lines within the are $ABCD$).

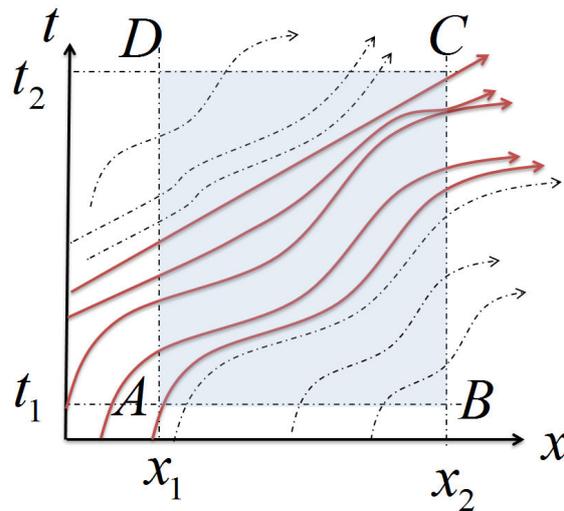


Figure 2.2: Temporal and spatial illustration of a section travel time

At present, there exist several ways to collect or estimate travel time data. Fig. A.2 summarizes the principal techniques existing nowadays. They can be divided into two groups: direct measurements, and indirect measurements. The main difference between the two groups is that, while direct measurements give the travel time of individual vehicles and thus the section travel time is given as the average of these, the indirect measurements characterize directly the road section as provide the information of a set of the vehicles. In the following a general overview for each is presented.

2.2.1 Travel time computation from direct measurements

1. **Test Vehicle:** this technique, introduced in the 1920's, was the most common one in early research. It basically consists of a vehicle that is specifically dispatched to drive with the traffic stream for the express purpose of data collection. The technique depends upon the instrumentation used for the collection. The used instrumentation are:

Manual: it requires a driver, that operates the test vehicle, and a passenger, that manually records the data at predefined checkpoints. The manual collection is generally performed by using pen and paper, tape recorder, or portable computer. Given its low implementation cost it is very popular.

Distance measuring instrument (DMI): it consists of an electronic DMI attached to the test vehicle's transmission and coupled to a portable computer. The travel time along a corridor is determined based upon speed and distance information provided by the DMI.

Global positioning system (GPS): a GPS receiver is connected to a portable computer and collects the latitude and longitude information that enables the tracking of the test vehicle in the traffic stream. With the development of the technology it has become the most used.

2. **License Plate Matching:** this technology, used since the beginning of the 1950's, consists of vehicle license plate number collection at specific checkpoints. The travel time is computed from difference of arrival times. There exist four basic methods of collecting and processing the license plates:

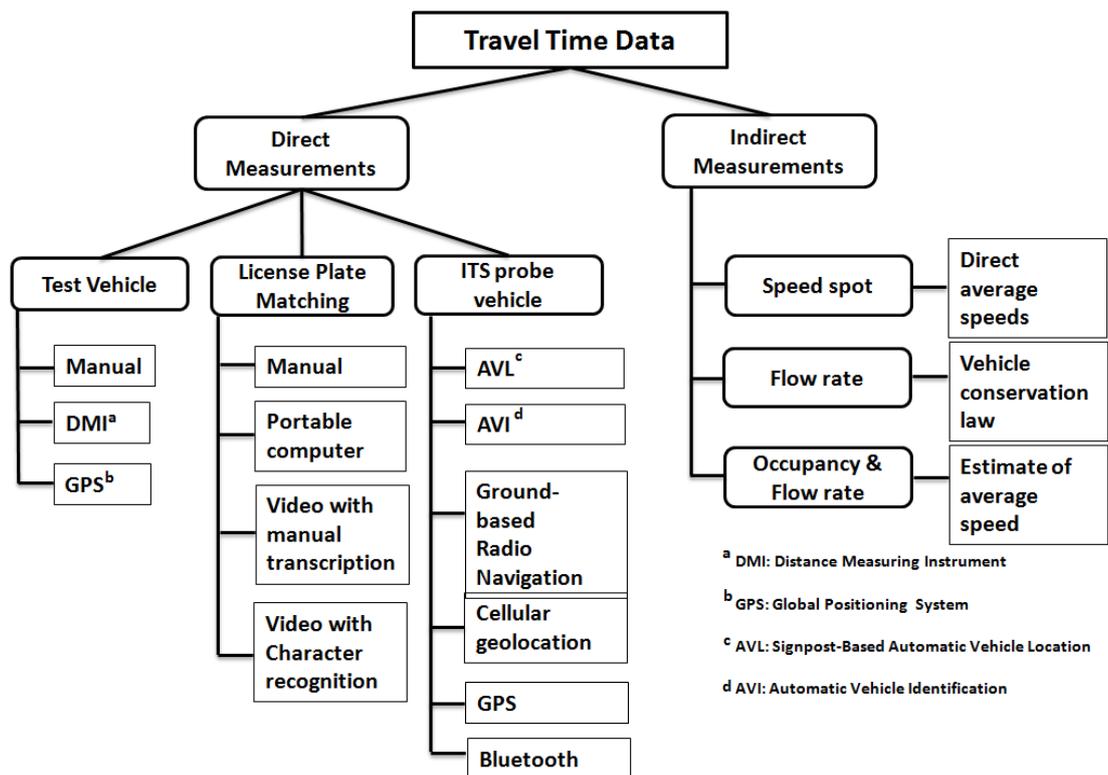


Figure 2.3: Different travel time collection techniques

- **Manual:** it collects license plates using pen and paper or using a recorder, all the information is then entered into a computer.
 - **Portable computer:** it collects license plates in the field using portable computers that automatically provide an arrival time stamp.
 - **Video with manual transcription:** license plates numbers are collected using video cameras and manually transcribing the information using human observers.
 - **Video with Character Recognition:** it collects license plates in the field using video, then automatically transcribes license plates and arrival times into a computer using computerized license plate character recognition.
3. **ITS probe vehicle:** it is a technique initially developed for real-time monitoring of traffic conditions, rather than for specific collection of travel time data, however it has been proved to efficiently collect the latter as well. As the probe vehicles are in the traffic streams for other purposes than traffic data collection, for instance taxis equipped with devices, they are often referred as “passive” test vehicles. These systems typically have a high implementation cost and are suited for large-scale data collection efforts. However, they allow for continuous data collection and for minimal human interaction. The travel time data collection using probe vehicles can be classified in five groups:

- **Signpost-Based Automatic Vehicle Location (AVL):** probe vehicles communicate with transmitters mounted on existing signpost structures. This technique has been widely used by transportation agencies to capture the position of buses.
- **Automatic Vehicle Identification (AVI):** probe vehicles are equipped with electronic tags. These tags communicate with roadside transceivers to identify the vehicle and collect travel times between transceivers.
- **Ground-Based Radio Navigation:** the data are collected by communication between probe vehicles and radio towers equipped with antennas. At every requested time, the vehicle communicates its unique ID and the time stamp to multiple towers.
- **Cellular Geo-location:** technology that collects travel time data by discretely tracking cellular telephone call transmissions.
- **Global Positioning System (GPS):** probe vehicles are equipped with GPS receivers that allow to receive signals from earth-orbiting satellites and then compute location and speed of the vehicle. The positional information determined from the GPS signals is transmitted to a control center to display real-time position of the probe vehicles.
- **Bluetooth:** it is a growing technology during the last years. The travel time is determined by re-identifying bluetooth devices in the probe vehicles such as cell phones between multiple sites. A single bluetooth reader mounted on the side of the road can be used to determine the travel time.

2.2.2 Travel time computation from indirect measurements

The travel time can be estimated from the macroscopic traffic parameters: flow (number of vehicles passing a point per unit of time), speed (distance travelled per unit of time), density (number of vehicles per unit of distance), or occupancy (percentage of time over a fixed time interval in which a sensor is occupied). These informations can be collected making use of several types of technologies that provide the traffic data at discrete time $t_k = k\Delta T$, with ΔT the sampling time. Note that while flow, speed, and occupancy are point-wise measurements, the density is a distributed one.

Speed is the traffic variable most closely related to the concept of travel time. The travel time is inverse of the speed per travelled space. Thus, when direct measurements of speed are available the travel time can be easily computed, otherwise it is usually inferred from the other traffic variables.

Before going further, there are two types of speeds that should be distinguished. For this, let us consider a link as shown in Fig. A.3, and assume that the link boundaries are equipped with sensors able to provide speed measurements, then:

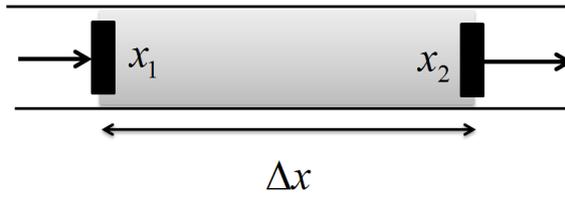


Figure 2.4: Link of freeway.

- Time-mean speed (TMS): it is the point-wise information given by the sensors. It is defined as the arithmetic average speed of all vehicles for a specified period of time at a specific point, Eq. A.1.1.

$$v_x(k) = \frac{1}{N_k} \sum_{i=1}^{N_k} v_{x,i}, \quad (2.2.1)$$

where $v_{x,i}$ is the speed of the vehicle i at the point x , and N_k is the number of vehicles that passed through x during $[k - 1, k]$.

- Space-mean speed (SMS): it is the average speed of the vehicles travelling in the link at a specified time.

$$v_{\Delta x}(k) = \frac{1}{N_{\Delta x}} \sum_{j=1}^{N_{\Delta x}} v_{\Delta x,j}(k), \quad (2.2.2)$$

where $v_{\Delta x,j}(k)$ is the speed of the vehicle j at time k that is inside the link Δx , and $N_{\Delta x}$ is the number of vehicles in the link at time k .

Fig. 2.5 illustrates the two types of speed. In particular, it evidences that TMS is associated with a single point along the link over time, whereas the SMS is associated with the specified length of the link.

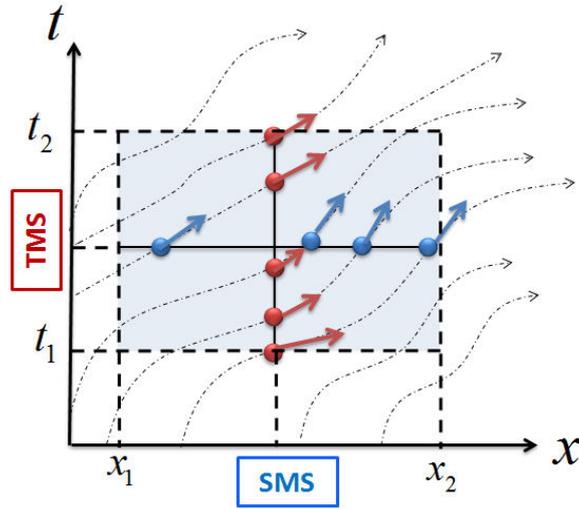


Figure 2.5: Illustration of the types of speed. TMS=mean of speed of circles in red, SMS=mean of speed of circles in blue.

SMS should be used for travel time computation in a road section [9, 75, 95]. However, as only TMS is the available measurement, we should derive a relation between these two speeds. It was demonstrated in [55] that when the time-mean speeds over fixed time intervals are stationary, the mean Eq. A.1.3 is equal to the space-mean speed.

$$v_{\Delta x}(k) = 2\left(\frac{1}{v_{x_1}(k)} + \frac{1}{v_{x_2}(k)}\right)^{-1}. \quad (2.2.3)$$

In the following, we will refer to Fig. A.3 to describe the methods to obtain the travel time using direct and indirect measurements of speed.

1. Measuring speed:

The estimated travel time in the link can be computed using the following equation:

$$tt(k) = \frac{\Delta x}{v_{\Delta x}(k)}, \quad (2.2.4)$$

where $v_{\Delta x}(k)$ is given by Eq. A.1.3.

2. Occupancy and flow:

When the technology at hand only allows the collection of flow and occupancy. The conventional method is first to estimate the TMS and then SMS:

$$\text{TMS} = \frac{\text{flow}}{\text{occupancy} \times g} \rightarrow v_x(k) = \frac{\phi_x(k)}{o_x(k)g_x(k)}, \quad (2.2.5)$$

where ϕ_x is the flow, o_x the occupancy, and g is the mean effective vehicle length (MEVL) through the detector located at position x during the period of time considered. In the

majority of real implementations, g is not known. Thus, several approaches have been proposed in order to infer its value. The simplest one considers a constant value over time [21, 53, 61]. While other assumes a different value for uncongested and congested traffic conditions [18, 31, 71].

Then the link speed and travel time can be deduced using Eq. A.1.3 and Eq. A.1.4 respectively.

3. flow and density:

Recalling the fluid mechanics equation stating that the flow is equal to the product of the density and the speed of a fluid: $\varphi = \rho v$. The speed of the fluid v is written as:

$$v = \frac{\varphi}{\rho}. \quad (2.2.6)$$

Under the assumptions that all the vehicles in the link travel at speed v , Eq. A.1.6 can be applicable to traffic (this will be deeper explained in the following chapters). Then, v corresponds to SMS, ρ to traffic density, and φ to traffic flow. Given that the measurement of flow data is point-wise, the following average value is usually adopted:

$$\varphi_{\Delta x}(k) = \frac{\varphi_{x_1}(k) + \varphi_{x_2}(k)}{2}.$$

Then:

$$v_{\Delta x}(k) = \Delta x \frac{\rho_{\Delta x}(k)}{\varphi_{\Delta x}(k)}, \quad (2.2.7)$$

where $\rho_{\Delta x}$ is the density in the link. The link travel time is deduced from Eq. A.1.4.

2.3 Derivation of the forecasted travel time

Once different methods for collecting travel time data have been presented, this section aims to derive how the travel time is forecasted. For this, let us imagine a map of a vehicle trajectory traversing through a route limited by the space interval $[x_p, x_0]$ and time interval $[t_p, t_0]$, as shown in Fig. A.4-a. Here, we would be interested in deriving a mathematical formula for the exit point of the vehicle $[x_0, t_0]$, given the entry point $[x_p, t_p]$. This formula will be particularly helpful in the forecasted travel time formulation.

If the velocity distribution $v(x, t)$ in $[x_p, x_0]$ is known, the vehicle travel time at infinitesimal space intervals is given by $dt \triangleq \frac{dx}{v(x, t)}$, which yields the following integral equation, that gives, in continuous time, the exact value of t_0 :

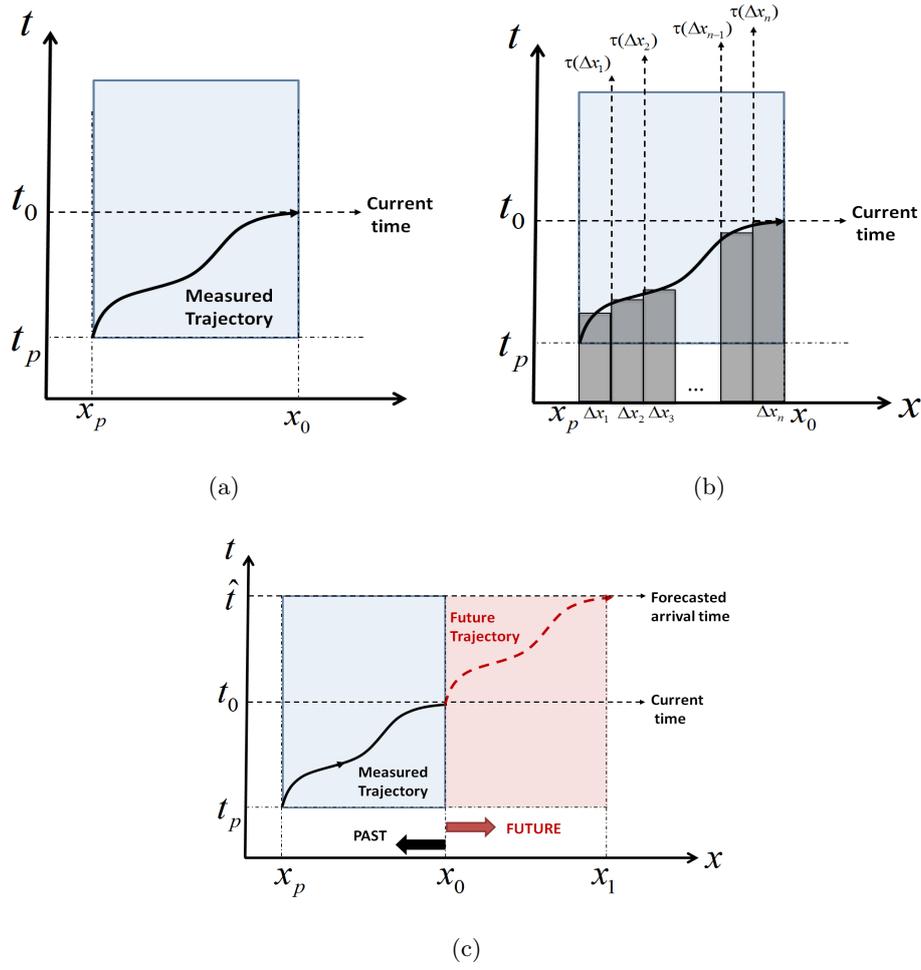


Figure 2.6: Scenarios considered in order to obtain a formulation of the forecasted travel time. a) Vehicle trajectory from x_p to x_0 , at the passing times t_p and t_0 respectively. b) Space discretization in volumes Δx_i . c) Scenario in which we aim to forecast \hat{t} , which is the future time the vehicle would reach the point x_1 .

$$t_0 - t_p = \int_{x_p}^{x_0} \frac{dx}{v(x, t)}. \quad (2.3.1)$$

Normally, the continuous distribution $v(x, t)$ is unknown. In general however, points of measures are placed within roads. These points allow us to approximate $v(x, t)$ by discretizing the space interval and making suitable assumptions within the intervals.

Making then a spatial discretization of Eq. A.2.1, we divide the space region into a set of n finite volumes of length Δx_i , where the speed in each volume is assumed constant, see Fig. A.4-b. Therefore, $v(x, t)$ is approximated with a piece-wise linear function of constant values $v_i(t)$. Now, the discrete version of Eq. A.2.1 yields the exit time:

$$t_0 = t_p + \sum_{i=1}^n \frac{\Delta x_i}{v_i(\tau(\Delta x_i))}, \quad (2.3.2)$$

where v_i is the volume's space-mean speed, and $\tau(\Delta x_i)$ the time in which a vehicle reaches the

upstream boundary of the volume i

Important attention must be given to $\tau(\Delta x_i)$. Its value is given by:

$$\tau(\Delta x_i) = t_p + \sum_{j=1}^{i-1} \frac{\Delta x_i}{v_j(\tau(\Delta x_j))} \quad (2.3.3)$$

Eq. A.2.3 considers the effect of traffic progression along the road. It states that the vehicle's arrival time at the upstream of the volume i is a function of the time spent in the volume $i - 1$. Eq. A.2.2 and Eq. A.2.3 define the *progressive travel time* (PTT). In literature, we can also find the so-called Instantaneous Travel Time (ITT). This differs from PTT as it assumes that the conditions at each volume are frozen as they were at the entry time t_p , i.e.

$$t_0 - t_p = \sum_{i=1}^n \frac{\Delta x_i}{v_i(t_p)} \leftarrow ITT \quad (2.3.4)$$

Clearly, it is seen then that while PTT is consistent with the traffic conditions experienced by a driver along the road, ITT is based on a very strong and not always realistic assumption, which becomes more critical as the length of space interval increases. Therefore, PTT provides a more rational formulation in order to compute t_0 .

In this dissertation, we will only be interested in PTT. This will be referred to, for the rest of the document, as just travel time.

Now, with the formulation of exit time, we can devise the forecasted travel time. For this, let us consider the scenario depicted in Fig. A.4-c. The objective is to forecast the arrival time \hat{t} at the point x_1 , given the entry point $[x_0, t_0]$. Thus, the problem reduces to the same scenario as before. The forecasted travel time is then:

$$\begin{aligned} \hat{t} - t_0 &= \sum_{i=1}^n \frac{\Delta x_i}{\hat{v}_i(\hat{\tau}(\Delta x_i))} \\ \hat{\tau}(\Delta x_i) &= t_0 + \sum_{j=1}^{i-1} \frac{\Delta x_i}{\hat{v}_j(\hat{\tau}(\Delta x_j))} \end{aligned} \quad (2.3.5)$$

If we simplify the notation by considering:

- $\hat{t}t_i(k)$: as the forecasted travel time of volume i at a discrete time k .
- $\hat{T}T_i(k)$: as the forecasted travel time from an entry point $[x_0, k]$ to the downstream of volume i .

The multi-step ahead forecasted travel time between the interval $[x_0, x_1]$ starting at the discrete current time k_0 , writes:

$$\begin{aligned} \hat{T}T_i(k) &= \hat{T}T_{i-1}(k) + \hat{t}t_i(k + \hat{T}T_{i-1}(k)) \\ \hat{t}t_i(k) &= \frac{\Delta x_i}{\hat{v}_i(\hat{T}T_{i-1}(k))}, \end{aligned} \quad (2.3.6)$$

where $i = \{1, \dots, n\}$ and $k = \{k_0 + 1, \dots, k_0 + \Delta\}$, with Δ as the forecasting horizon.

Note that the forecasted mean speed/travel time within the volumes, $(\hat{v}_i/\hat{t}t_i$ respectively), are the key to solve Eq. A.2.4.

The graphical concept of forecasted travel time is explained in Fig. 2.7, where the road will be divided in n links, and we are interested to know the time that the vehicle shall get to x_n , $\widehat{TT}_n(k)$.

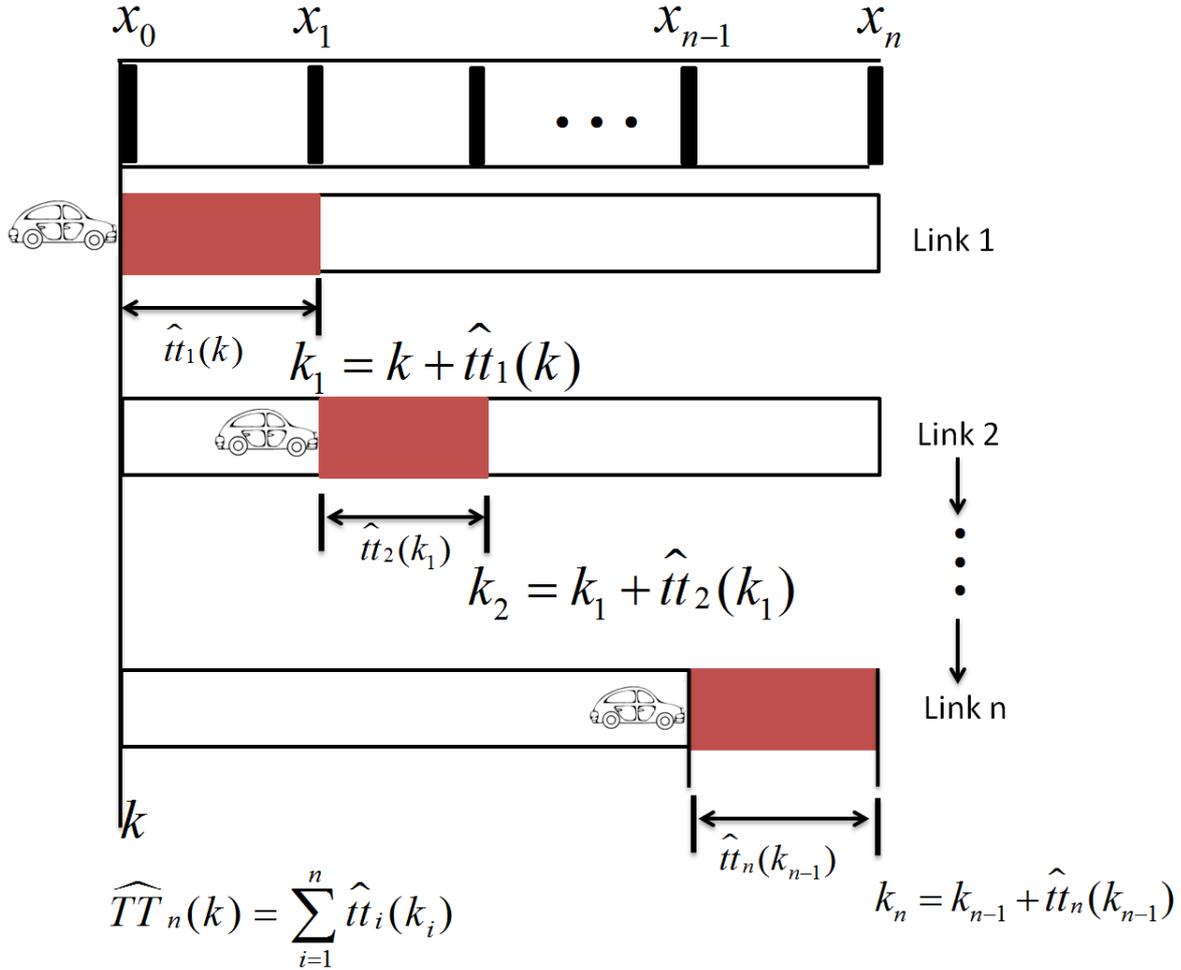


Figure 2.7: Travel time forecasting scheme. A vehicle enters the road at point x_0 and time k_0 , and we are interesting to forecast the travel time between $[x_0, x_n]$, departing at multi-step ahead in the future.

2.4 Conclusions

The main contributions in this chapter are:

- **Classification of the travel time data.** We put into light different means to obtain travel time data. These are classified in direct or indirect measurements. In the former we find the methods that allow the tracking of trajectory of the vehicles in a traffic stream, and therefore a map space-time can be known. The latter refers to the case where travel

time data is inferred from other traffic parameters such as, speed [km/h], flow [veh/h], or occupancy [%].

- **Mathematical derivation of the forecasted travel time.** For a discretized freeway, we derived a formulation to forecast the travel time. The computation took into account space and time progression of a vehicle in the freeway. The main idea was to evaluate the forecasted travel time in the links at the right time, so changes in the traffic conditions were eventually acknowledged.

Chapter 3

Traffic data collection and pre-processing

3.1 Overview

The main objective in this chapter is the study of GTL and the traffic data collected from GTL.

This chapter is organized as follows. In Sec. 3.2, we describe the GTL and its architecture. In Sec. 3.3, we present the type of traffic parameters provided by the platform, as well as the data format. In Sec. 3.4, we provide the data pre-processing steps, including data cleaning and imputation. We also give results regarding the suitable choice for the aggregation time of the traffic data.

3.2 Grenoble Traffic Lab (GTL)

As a part of the City Labs initiative, an experimental platform, Grenoble Traffic Lab, has been launched by INRIA/Gipsa-lab team NeCS in collaboration with Karrus-ITS (an Inria spin-off), and local traffic authorities (DIR-CE).

In the Fig. 3.2 we can find the different functional levels of GTL. In the following, these layers are explained.

1. Level 1: Physical layer

The south ring has been equipped with *Sensys Networks* VDS240 3-axis wireless magneto-resistive sensors embedded in the pavement. The installation and sensor-configuration process took approximately one year (the field installation is limited during the winter months). The sensors have a sampling rate of 128 Hz and are powered with non-rechargeable primary Lithium Thionyl Chloride (Li-SOCl₂) 3.6V, 7.2Ah batteries which guarantee 10 years of autonomy and up to 300 millions vehicle detections. The magne-

tometers use a ultra-low power 2.4 GHz TDMA protocol to communicate with an access point (configured and remotely operated with Sensys software “TrafficDOT2”), which sends the data to a server located at the DIR-CE via fiber optics, or via a wireless GPRS connection (20 out of 22 stations are connected via fiber optics). If the magnetometer is outside a radius of 45 meters from the access point, a repeater (mounted on the vertical signage: road lamps, sign or camera poles) is utilized to relay the signal to it. Overall, 20 access points and 21 repeaters are active in the south ring. The traffic data is monitored and stored in a database (or DB in short) at DIR-CE, where every 15 seconds a data exporter pushes it to a server located at Inria Grenoble Rhône-Alpes via an FTP connection (see Fig. 3.2). Note that the communication between DIR-CE and Inria is *unidirectional*: this means that GTL has not direct access to the data stored in the DIR-CE server or to the current health status of each magnetometer.

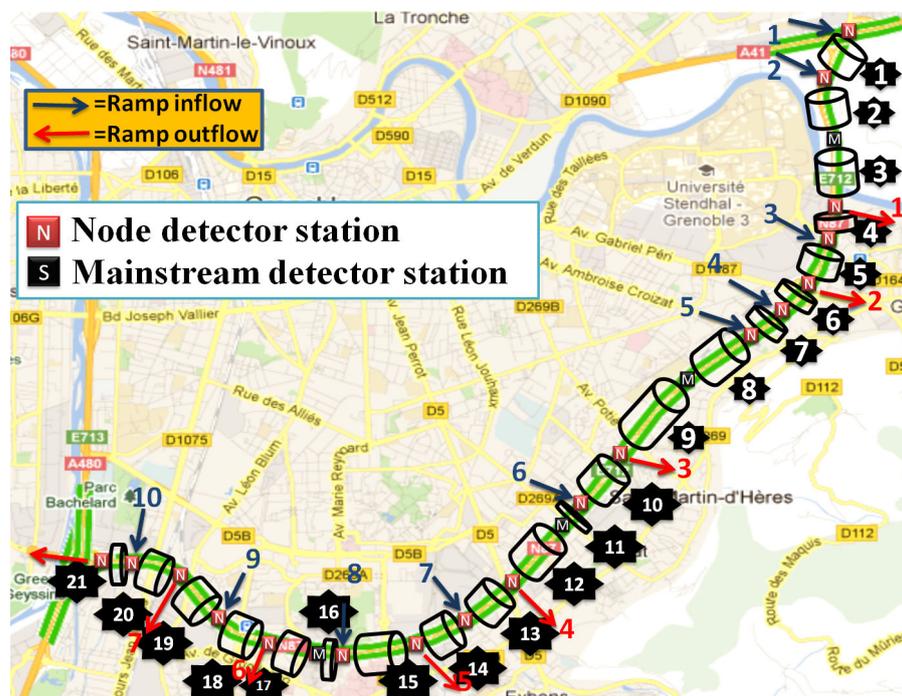


Figure 3.1: Grenoble south ring division

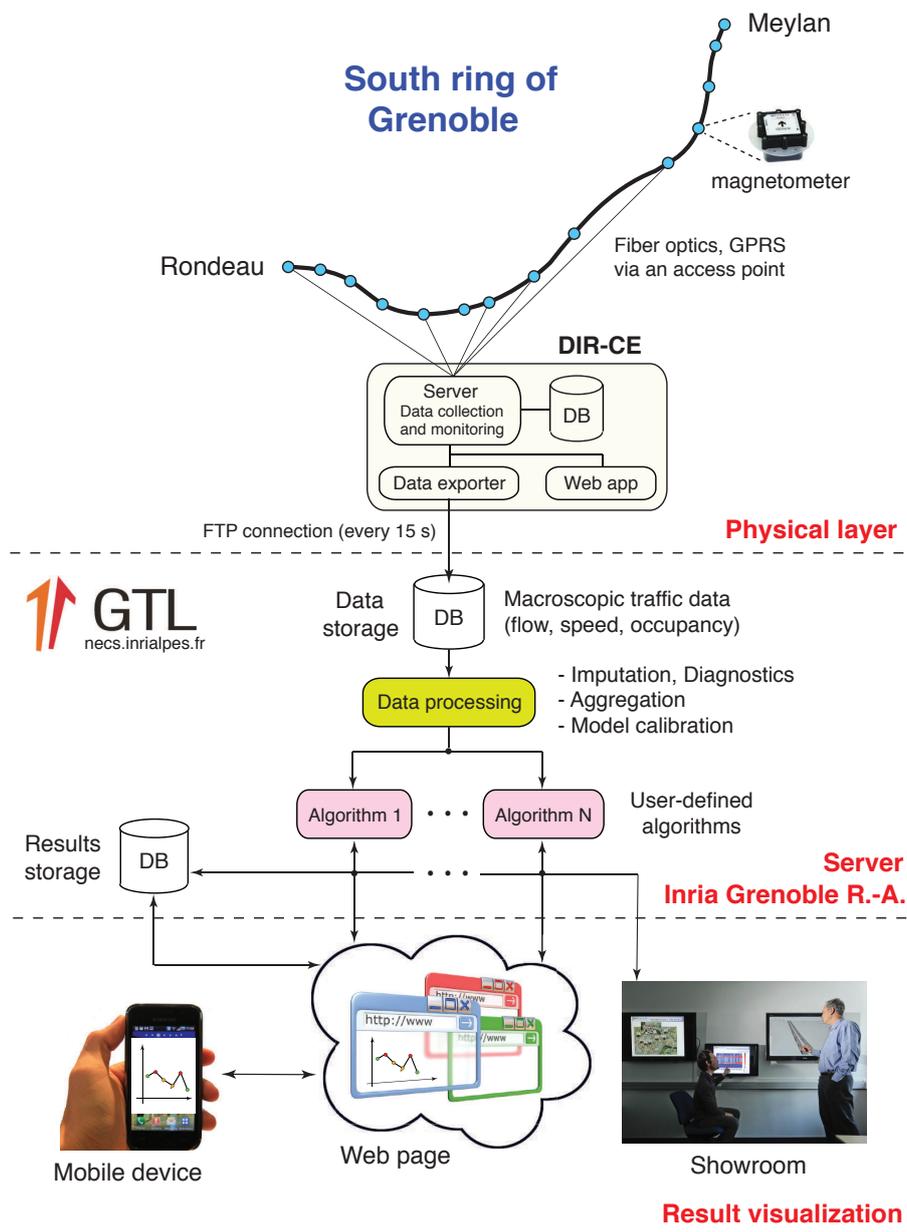


Figure 3.2: Three-level architecture of GTL

2. Level 2: Server Inria Grenoble Rhône-Alpes

Level 2 consists of an *upper* and *lower* layer, which are described in full detail below.

- **Lower layer:** the raw macroscopic traffic data coming every 15 seconds from the Sensys magnetometers (see Level 1) is stored in a database and then passed through a suite of signal-processing algorithms (green box in Fig. 3.2), which perform the data pre-processing.
- **Upper layer:** this layer is the repository of *user-defined* traffic-management algorithms. The results yielded by these algorithms are stored in a database either for later use or for visualization purposes (see Level 3 below).

3. Level 3: Result visualization

The results produced by the user-defined algorithms in Level 2 can be visualized using different media, including a:

- **Web page:** this page allows to visualize the collection points in the south ring and the color-coded level of traffic congestion in the intermediate links. The web page, that will be presented in detail in the next chapter, has been built upon the Google Maps interface.
- **Mobile device:** an Android smart-phone application called “GTLmobile” has been developed to display salient traffic information (travel time, energy consumption) to the private users in real time.
- **Showroom:** the content of the web page, plus additional diagnostic information about the prevailing traffic conditions, accident hot spots and quality of data (for internal use, only), is displayed 24 hours a day, 7 days a week, in seven monitors in a dedicated room at Inria Grenoble Rhône-Alpes.

3.3 Data description

The sensors embedded in the Grenoble south ring belong to *Sensys Networks* technology. This technology allows the collection of several traffic information. This information can be either individual or aggregate. For each type, *Sensys* sensors is able to provide several useful information. Recalling that in order to provide speed measurements one pair of sensors are placed per lane, Table 3.1 and Table 3.2 summarize the traffic parameters that can be collected, as well as the pre-defined aggregation times.

Table 3.1: Available individual traffic parameters. Each time an event (detection of vehicle) occurs, this information is generated.

Parameters (Individual data)	Description
Timestamp	The time the vehicle was detected by the first of the two sensors
Access Point Lane ID	Identifier of the Sensor
Speed [km/h]	Speed of the vehicle
Length [m]	Length of the vehicle
Gap [s]	Time elapsed since detection of previous vehicle.

GTL database is updated automatically at every time new raw traffic data is pushed from the data server at DIR-CE. These data arrive in compressed comma-delimited ASCII format and

contain, in a single file, individual or aggregate data for all detectors. Each file is identified by a file name, for example: “5001.txt.2014-01-20_01-00-07” (“identifier. file extension.date-collection time”). The GTL has chosen an aggregation time of 15 s.

Fig. 3.3 depicts a sample of the individual and aggregate raw ASCII data that come into the server every 15 s. Fig. 3.3-a contains only the sensors where events were detected and the information related to it (Table 3.1). Whereas Fig. 3.3-b presents the aggregate information corresponding to all the sensors in the freeway (Table 3.2).

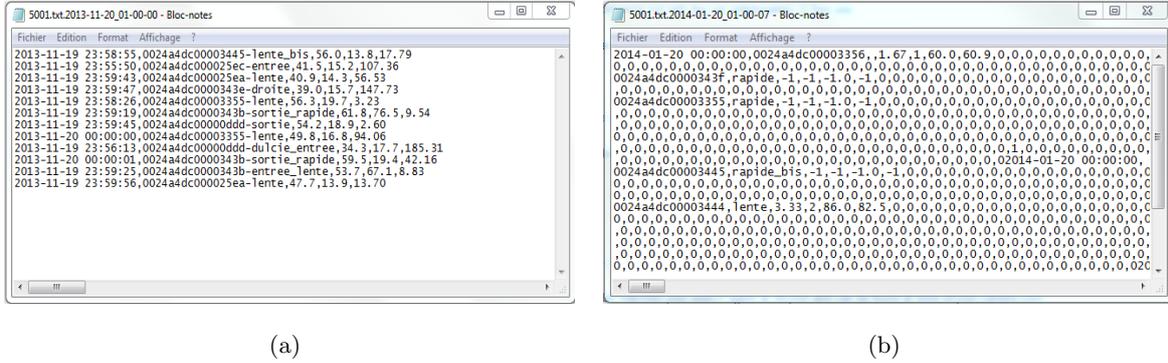


Figure 3.3: File for the individual raw traffic data(a) and File for the aggregate raw traffic data(b).

Once in the database, the raw data is subsequently disjoined by files. One file describes the traffic information of the lanes and the ramp present in the detector station. These data can be downloaded from GTL’s interface also in comma-delimited ASCII files. This interface provide same day and historical information, for both individual and aggregate data. Fig. 3.4 depicts the GTL’s interface.

Table 3.2: Choices of the aggregation times given by *Sensys* and available aggregate traffic parameters.

Aggregation times	Parameters (Aggregate data)	Description (by time interval)
10 s	Timestamp	End of the interval
15 s	Access Point Lane ID	Identifier of the Sensor
30 s	Speed [km/h]	Average speed of the vehicles
1 min	Counting	Number of vehicles
5 min	Occupancy [%]	Percentage of time the sensor is occupied
10 min	Speed histogram	Distribution vehicles speed
15 min	Length histogram	Distribution vehicles length

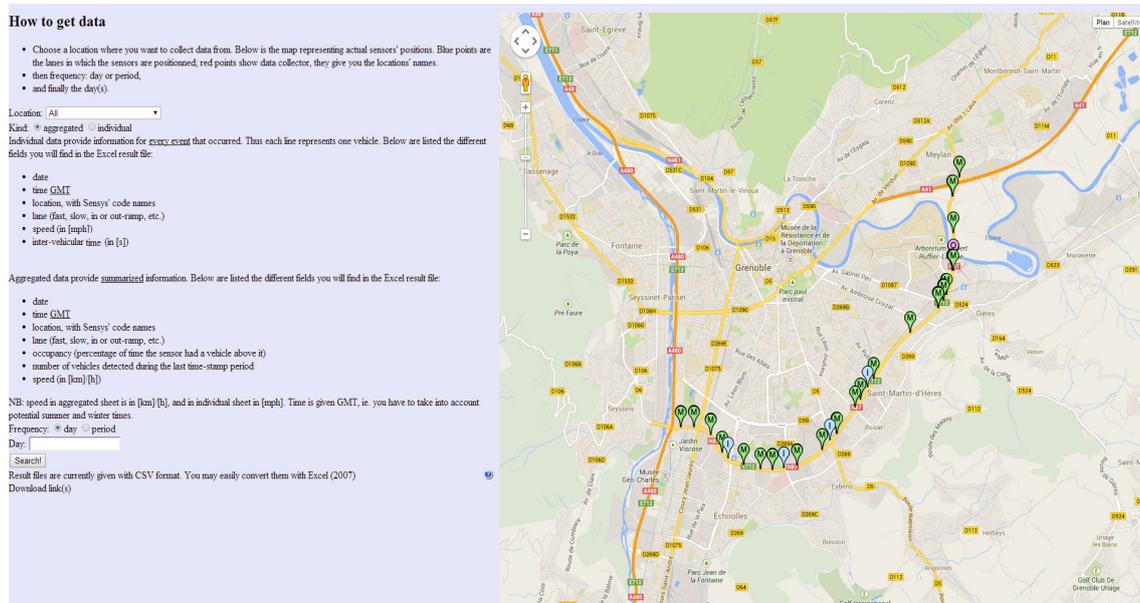


Figure 3.4: GTL's interface used to download traffic data.

Each file downloaded from the interface is also identified by a file name, for instance: “2014-01-01_0024a4dc0000343e_aggregate”. The first field represents the date “2014-01-01”, the second the access point related to the station “0024a4dc0000343e”, and the last, the type of data “aggregate” or “individual”.

Tab. 3.3 exhibits the 22 detector stations installed in Grenoble south ring and managed by GTL. It is important to notice that every station is associated to an access point. In some cases two stations can be connected to the same access point. In this case, their ID are different by adding a literal name (see D and E for instance).

Fig. 3.5 illustrates the steps executed by the GTL platform for the data collection, storage, and supply.

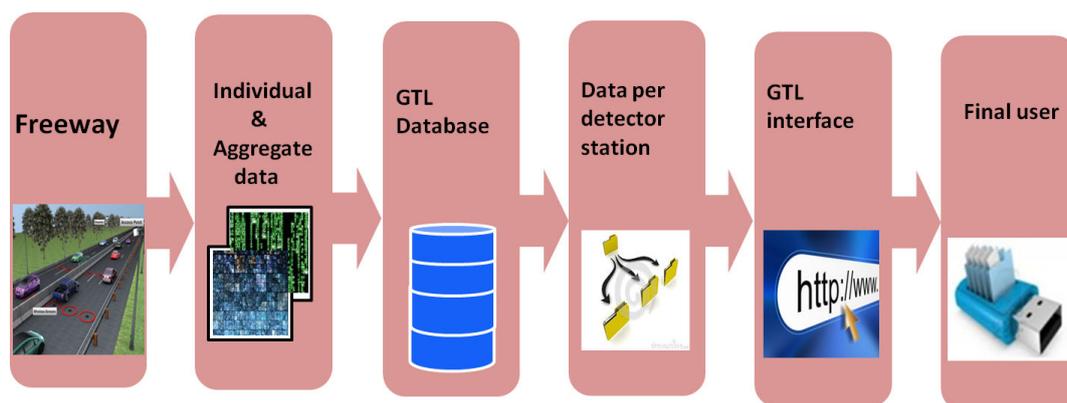


Figure 3.5: GTL data management levels (data collection, storage, and supply).

Table 3.3: Vehicle detector stations installed in Grenoble south ring. “ID” is a hexadecimal serial number associated to groups of magnetometers. The communication is via fiber optics, “f”, or GPRS, “g”.

	Location (type)	Lanes	ID, Com.
A	Meylan (entrance)	Slow, Fast, On-ramp	3356, f
B	A41 Grenoble (entrance)	Slow, Fast, On-ramp	3354, f
C	Carronnerie (mainstream)	Slow, Fast	343c, f
D	Domaine Univer. (exit)	Slow, Fast, Off-ramp	343b_ entry, f
E	Domaine Univer. (entrance)	Slow, Fast, On-ramp	343b_ exit, f
F	Gabriel Péri (exit)	Slow, Fast, Off-ramp	3445, f
G	Gabriel Péri (entrance 1)	Slow, Fast, On-ramp	3445_ bis, f
H	Gabriel Péri (entrance 2)	Slow, Middle, Fast, On-ramp	1b67, g
I	SMH (mainstream)	Slow, Fast	3357, f
J	SMH Centre (exit)	Slow, Fast	0ddd, f
K	SMH Centre (entrance)	Slow, Fast, On-ramp	3355, f
L	SMH Centre (mainstream)	Slow, Fast	3355_ entry, f
M	Eybens (exit)	Slow, Fast, Off-ramp	21d1, f
N	Eybens (entrance)	Slow, Fast, On-ramp	343f, f
O	Échirolles (exit)	Slow, Fast, Off-ramp	1b5c, g
P	Échirolles (entrance)	Slow, Fast, On-ramp	25eb, f
Q	Échirolles (mainstream)	Slow, Fast	25eb_ bis, f
R	États Généraux (exit)	Slow, Fast, Off-ramp	25ea, f
S	États Généraux (entrance)	Slow, Fast, On-ramp	13c6, f
T	Libération (exit)	Slow, Fast, Off-ramp	3444, f
U	Libération (entrance)	Slow, Fast, On-ramp	25ec, f
V	Rondeau (mainstream)	Left, Middle, Right	343e, f

3.4 Data pre-processing

Due to malfunctioning of the data collection, relay mechanism, and real system intrinsic noise, it is well known that real-time traffic data is always distorted by noise, and usually include false and missing data. For these reasons, the quality of the individual and 15 s aggregate measurements need to be checked. Specially, corrective actions need to be taken in order to assure that the estimation and forecasting algorithms are as reliable as possible.

In the sequel we present the steps for the data treatment. In order, these steps are: data cleaning (identification and elimination of erroneous data), and data repair (replacement or imputation of missing and erroneous data).

3.4.1 Data cleaning

For both individual and 15 s aggregate data, a series of rules have been worked out to discard false values. Based on analysis of the *Sensys* data, we will consider that the vehicle count has the highest priority, i.e. the system is always able to count the number of vehicles, or an event, but not the speed.

This can be explained firstly by recalling that, to measure speed it is necessary to have a pair of consecutive sensors and that secondly, and that *Sensys* sensors cover only a relative small area of the lane. It was found that when the lanes are too wide or the vehicle changes lanes drastically, the second sensor cannot detect the vehicle's passing time and consequently the speed is not measured. When the speed measurement is not available, the value received is -1. Table A.1 shows the combination rules for the 15 s aggregate data cleaning. These rules are applied to every sample. If a sample of vehicle count or speed does not pass the combination rule, its value is set to null. A null value indicates missing value.

Table 3.4: Data cleaning rules for the 15 s aggregate data.

Combination rules			Action		Description
Count	Speed	Occupancy	Count	Speed	
-	>150	-	-	Discard	Test 1: Maximum limit of speed set to 150 km/h
>0	>0	-	Accept	Accept	Test 2: When vehicle and speed are counted, this is assumed to be the correct functioning.
=0	-1	-	Accept	Discard	Test 3: No vehicle present during 15 s. As we are not able to distinguish if the road is in free flow or congestion, the speed measured is assumed missing.
>0	-1	-	Accept	Discard	Test 4: Vehicles were counted, but the sensors were not able to measure the speed.
=0	>0	>0	Discard	Accept	Test 5: Vehicle count will be assumed erroneous if both speed and occupancy are measured

For the individual data, every event represents a vehicle detected, and the speed of this vehicle may or may not have been measured. Hence, every vehicle detection is considered a right event, however the speed is assumed correct when it has been measured and with a value lower than 150 km/h. Otherwise it is set to null.

3.4.2 Data imputation

A crucial step for the traffic data repair is the imputation of missing or erroneous values. As such, several authors have worked out different strategies to fill the missing data and therefore improve its reliability, see [4, 12, 28, 79] and the references therein. The techniques can be generally classified in three groups: nearest neighbor, historical classification, and regression analysis.

In [12] in particular, the authors presented a set of imputation algorithms. The most efficient ones, in terms of accuracy and complexity, were the following:

1. Historical average (HistAv): it uses the average of the historical data to replace the missing data.
2. Time neighbors interpolation (TimeN): the missing sample is imputed with the average of the preceding and succeeding samples in time at the same detector station.
3. Spatial neighbors interpolation (SpatialN): the missing sample is imputed using the average value between the downstream and upstream detector stations.
4. Hybrid (historical and series analysis): it uses the historical data and time series analysis for the data imputation.

The assessment of the algorithms in [12] was done by imputing a percentage of randomly extracted samples in a dataset, and computing the error between the imputed and the real value of the sample. Defining applicability as the possibility to apply an algorithm to a missing sample, in this study the results showed that:

- The best results were found by replacing the missing data by the time neighbors interpolation. This algorithm however, can be applied only when these temporal neighbors are available. Thus, making it applicable for a certain percentage of missing data.
- Promising results were also found using the spatial neighbors interpolation. Nevertheless as TimeN, its applicability is also compromised by several factors: dependency on the distance between adjacent detector stations, presence of inputs of outputs between them, and existence of the neighbors.
- The hybrid strategy performed also well in terms of accuracy, although some worries arised when this algorithm needed to be applicable in real-time. The algorithm used in this work was a triple exponential smoothing model integrated with historical average.
- Compared to the other algorithms, the HistAv did not present notable results in terms of accuracy. However, its applicability was rather promising given that it could reconstruct almost 100% of the missing data. When there is a lack of historical values, the imputation must be done by addition of values created by some other means.

Taking into account these findings in [12] and the traffic data used in this thesis, the methods: time neighbors interpolation (TimeN), historical average (HistAv), and series analysis will be adopted. The strategy SpatialN was discarded due to the majority of the detector stations are located at node levels, hence the correlation between same-lane neighbors is weakened.

The series analysis is performed by the simple moving average (Eq. 3.4.1):

$$\hat{x}(t) = \frac{1}{N} \sum_{i=1}^N x(t-i), \quad (3.4.1)$$

where $\hat{x}(t)$ is the imputed data at the current time t . N is the window length and is chosen considering its smoothing effect on the time series. In this work, we use a value of $N = 4$ which takes into account traffic conditions for the last minute.

In the following we aim to assess the performance of each individual imputation algorithm. For this, we will make use of the same framework proposed in [12]. Thus, we present first the detector station selected for the test, as well as the description of the validation and training dataset to be used for the cross-validation assessment. Second, the performance metric chosen. Third, we give the selection of the percentages of samples to be randomly extracted from the dataset. And at last, we give numerical results of the performance for the different percentages of samples extracted. Notice that the historical dataset considered is not treated, thus there may be samples in which HistAv is not applicable.

Detector station

For a dataset constituted with working days from the 2nd to the 20th of September 2013, we will take the first half of profiles as a training set and the second half as validation, from the 2nd to the 11th and from the 12th to the 20th respectively.

The station chosen for the performance evaluation corresponds to C:“Carronnerie” station (ID:343c Table 3.3). This station was chosen given that for the period, it worked well (no major interruption) and more than 95% of the data passed the data cleaning rules in Table A.1.

Both flow and speed profiles can be imputed. For illustrative reasons, we will only present the results for the speed imputation.

Performance metric

To evaluate the performance the Mean Absolute Percentage Error (MAPE) will be chosen as metric. It is computed as follows 3.4.2:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|x_i - \hat{x}_i|}{x_i} \times 100, \quad (3.4.2)$$

where, n is the number of imputed samples, x_i is the real value of the sample, and \hat{x}_i is the imputed one.

This metric represents the average percentage of the deviation between the real and imputed n values.

Missing data percentage

After a study of the different stations. The percentage of missing data varied generally from 3% to 28%. To account for this variation, the metric MAPE was used to evaluate the algorithm performance assuming a percentage of missing samples of 10%, 20%, and 30% per day in the validation dataset.

Numerical assessment

In the sequel we will aim to present the statistical study, in terms of mean and standard deviation, of the MAPE for every percentage of missing samples. Each algorithm will be run separately with the objective to impute the majority of samples possible. Therefore, since their applicability can vary, we will compute their performance based on the effective samples imputed.

Table 3.5 exhibits the performance for each imputation algorithm.

Table 3.5: Assessment of the data imputation algorithms.

Algorithm	10% Miss- ing samples (MAPE)	20% Miss- ing samples (MAPE)	30% Miss- ing samples (MAPE)	Applicability for 30% Missing sam- ples
TimeN	5.17 ± 2.23	5.22 ± 3.75	6.27 ± 2.67	39.20%
HistAv	9.24 ± 3.17	9.02 ± 3.85	12.00 ± 3.44	91.17%
Moving average	7.34 ± 5.03	8.10 ± 4.02	8.11 ± 3.84	100%

The first to be noticed is that TimeN was more accurate, outperforming the other algorithms for both mean and standard deviation. Coming in second the Moving average and last HistAv. However in terms of standard deviation the latter presented better results. In average the accuracy of the algorithms decreased 15% when increasing the percentage of missing value from 10% to 30%.

When evaluating the applicability at 30% of missing samples, we found that the TimeN algorithm was able to impute only a 39.20% of these samples, whereas HistAv and Moving average took care of 91.17% and 100% respectively. Note that HistAv is applicable to all missing values if the historical dataset is complete.

According to the results in Table 3.5. We conclude that TimeN algorithm should have the highest priority when imputing data. However, since not all of the missing samples can be imputed, HistAv should take over for the remaining ones. We choose this one over Moving

average because we believe it is more efficient, for no computation is needed, as it only picks up the right historical average sample.

It is worth mentioning that in a real-time framework the problem is slightly changed, for the time neighbors interpolation cannot be applied. The priority sequence for an offline implementation is presented in Alg. 5, while for the real-time is given in Alg. 6.

Algorithm 1 Data imputation procedure (OFFLINE)

For each missing sample**Begin:****if TimeN** applicable **then**Missing value \leftarrow Average of temporal neighbors.**else****if HistAv** applicable **then**Missing value \leftarrow Historical average.**else**Missing value \leftarrow Moving average of the last 4 samples.**end if****end if**

Algorithm 2 Data imputation procedure (REAL-TIME)

For each missing sample**Begin:****if HistAv** applicable **then**Missing value \leftarrow Historical average.**else**Missing value \leftarrow Moving average of the last 4 samples.**end if**

Fig. 3.6 summarizes the data pre-processing strategy for online and offline data. The only difference is in the imputation step.

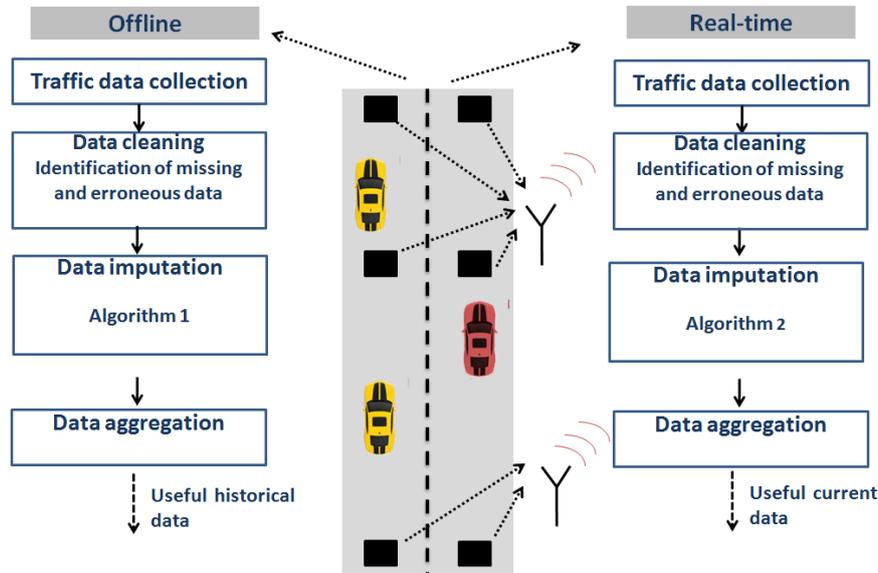


Figure 3.6: Data pre-processing strategy for offline and real-time data.

3.4.3 Data aggregation

Given that the accuracy of the forecasting suffers greatly from the strong variability present in the traffic data. Specially when the data is collected in very short intervals, 15 s in the GTL system. The majority of the forecasting strategies use as input aggregate information usually in 5 min interval or more [38, 69, 76].

However, even if larger levels of aggregation reduces noisy fluctuations, it may result in the loss of valuable information. Thus, the aggregation time is to be chosen considering a trade-off between error forecasting and strategy efficiency. In [1] the authors suggested that the higher aggregation time, the better the forecasting accuracy. In addition, they recommended to use a high aggregation when aiming at long forecasting horizons (30 min or more).

In the reviewed literature it was found a lack of consensus in the optimal choice of the aggregation time. In general this choice depends upon factors such as, the type of application, quality of data collected, and methodology used. Nevertheless the general agreement was the recommendation of using the aggregation time equal to the forecasting step.

In this work, the choice of this time for both speed and flow forecasting was of 5 min. The selection was based on the data treatment experience, as well as on the forecasting research reviewed involving California Performance Measurement System (PeMS) as test bed.

In the following, several figures contrasting 24 h profiles of speed and vehicle count at different aggregation times are presented. The data was acquired from the fast lane at I:“SMH” station (ID:357), see Tab. 3.3, in September 12th of 2013. Fig.3.7-a),b),c), and d) illustrates the speed for the 15 s raw data, and profiles aggregated at 1 min, 5 min, and 15 min, respectively. Fig.3.8-a),b),c), and d) depicts the same scenario using vehicle count data. The first to notice is the strong fluctuations in the 15 s data, Fig.3.7-a and Fig.3.8-a. While aggregating at 1 min these

fluctuations tend to cease, the 1 min aggregate data is still very noisy. While increasing the aggregation time to 5 min, Fig.3.7-c and Fig.3.8-c, a better compromise between information loss and fluctuations is observed. Although 15 min aggregate data present a quite smooth trend, too much information may be lost.

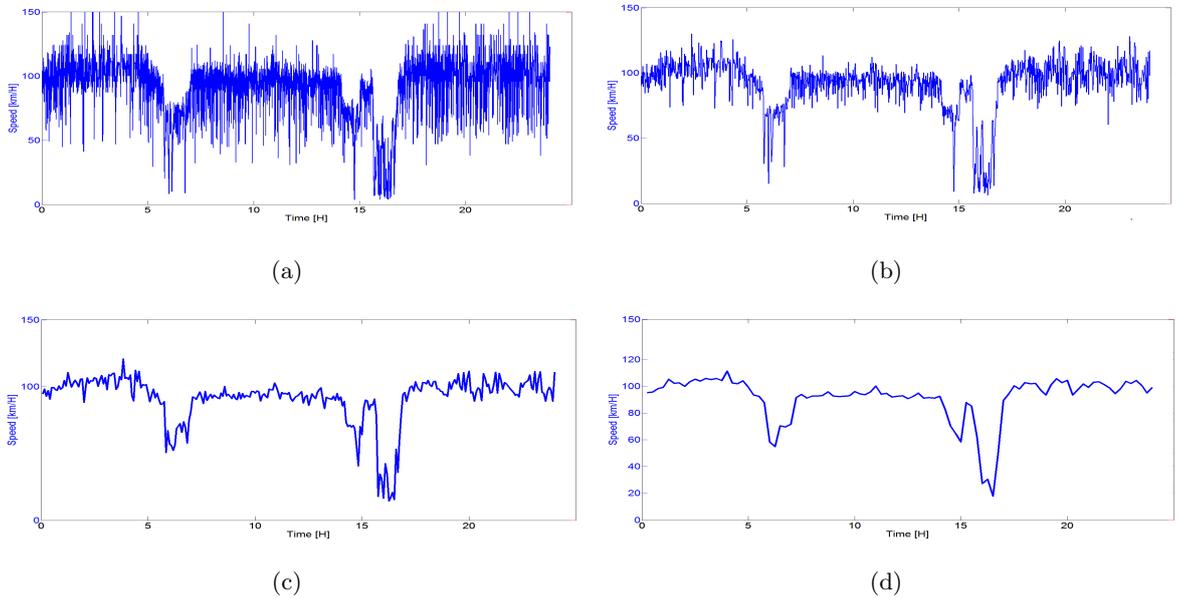


Figure 3.7: Example of a 24 h profiles of speed while changing the aggregation time. a)15 s raw data b), 1 min b), 5 min b), 15 min.

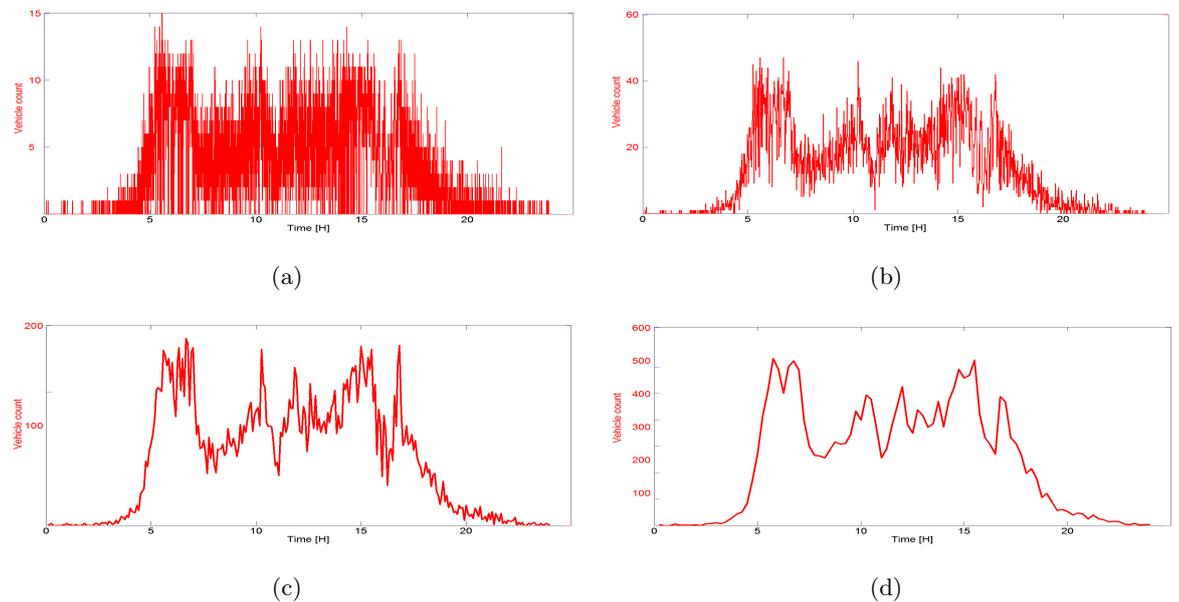


Figure 3.8: Example of a 24 h profiles of vehicle count while changing the aggregation time. a)15 s raw data b), 1 min b), 5 min b), 15 min.

3.5 Conclusions

The main contributions in this chapter are:

- **Identification of erroneous data.** We propose a series of rules to identify whether a sample is considered to be erroneous. These rules can be applied to the individual or the aggregate data of any of the traffic parameters collected, for either the historical or the current data.
- **Imputation of missing and erroneous samples.** Based on results from literature, we propose a series of algorithms for imputing the samples that are missing or considered erroneous. The combinations of the algorithms are proposed for both, an offline and a real-time framework. The algorithms were average of the time neighbors (TimeN), historical average (HistAv), and estimation using a moving average window.

The experimental assessment showed that imputing samples using TimeN presented the best results, however it was only implementable when the time neighbors were available. Coming in second the moving average window and finally the HistAv. This result was consistent with the findings from literature.

Chapter 4

Short-term multiple step ahead travel time forecasting: signal-based approach

4.1 Overview

Motivated by the introduction of a signal-based approach using only speed measurements, this chapter focuses on the development of an algorithm that can forecast, at multi-steps ahead in the future, the travel time between two points of interest of a freeway.

This methodology will be based only on mainstream speed measurements, and it will be particularly focused on forecasting the travel time inside the links enclosed in the section of interest, for it was shown in Chap. 2 that once each link's travel time is forecasted, the forecasted travel time in the section is easily derived. In that, it is worth to notice that each forecasting problem is decentralized at the link level. The space discretization discussed in Chap. 2 will be then fixed by the location of the detector stations.

This chapter is organized as follows. First, we describe the dataset considered for the validation of the method. Second, we derive the complete structure of the AKF scheme. Third, we present and validate the approach for clustering the historical dataset. Then, we explain and assess the algorithm that will assign a current data to a specific cluster. And finally, we address the experimental validation of the travel time forecasting using the signal-based method.

4.2 Data description

One of the core components in this chapter is the validation of the algorithms with real traffic data. The dataset consists of 15 working days, from the 2nd to 20th of September 2013.

The experimental raw data, downloaded from GTL, is pre-processed using the strategies given in Chapter 3.

4.3 Statement of the problem

Let us consider a freeway section divided in n links as shown in Fig. A.5. Given the mainstream speed measurements up to the current time k_0 , $k = \{1, \dots, k_0\}$, of the current day D , $v_i^D(k)$, $i = \{1, \dots, n\}$, and historical information of these data from preceding days $h = 1, 2, \dots, D - 1$, $v_i^h(k)$, the aim is to forecast, using an online multi-step ahead strategy, the section's travel time $\hat{T}_n^D(k) \forall k \in [k_0 + 1, \dots, k_0 + \Delta]$.

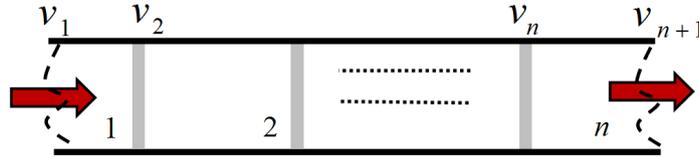


Figure 4.1: Abstraction of a freeway section. The grey stripes represent detector stations in the mainstream.

4.4 Noise adaptive Kalman filter

Chapter 1 presented a survey of different methodologies to forecast macroscopic traffic parameters. These methodologies were classified according to the type of data used. Their advantages are summarized considering that, by using only current data, the algorithms are more reactive to changes in the traffic conditions, but the forecasting horizon are usually very limited. On the other hand, when using only historical data, the forecasting horizon can be larger and the multi-step ahead forecasting is easier; however, in general, these strategies are unable to capture the current conditions. When a combination of both information is used, the methodologies are quite reactive to the present conditions and able to forecast at larger horizons, however they may not be always implementable in real time since the computational burden is typically important (large historical databases in general).

As historical traffic patterns generally provide valuable information, this thesis develops a forecasting methodology that makes use of the current day and historical data.

The methodology relies on the *Kalman filter* (KF) theory. This approach was selected for the following reasons: it runs successfully in real time applications and more important, it represented a novel approach in the multi-step ahead forecasting problem when fusing same day and historical data.

In order for the KF to provide updated state estimates at every time step along the considered time horizon, three key information are needed: noisy observations of the real system, statistics of the observations noise, and statistics of the process model noise. In the forecasting framework, these information, in particular the observations, are unknown. However we propose to circumvent this shortcoming by considering that: the unknown observations can be replaced by a suitable combination of the same day and historical data, referred to as pseudo-observations,

that their noise statistics can be captured in the dispersion of the historical data, and finally that the noise statistics of the process model is stationary and can be estimated online through an unbiased estimator proposed in literature. The first two assumptions are based on the hypothesis that the same day and historical data belong to the same probability distribution function.

In other words, this idea considers the historical data as noisy observations of the same day data, therefore as such, the Kalman filter operates recursively on streams of these noisy data to produce a statistically optimal estimate of the system state (travel time). Consequently, the forecasting problem is turned into a filtering one, by using these pseudo-observations.

This path presents a great improvement in the online traffic forecasting via multi-step ahead Kalman filtering, where it gives a novel work with the historical and same day data fusion concept.

In the following section, we aim to develop the structure of the AKF. For this, we first present the state-space model considered. Afterwards, we describe how the pseudo-observations are computed. Thereafter we expose how the noise statistics of the observations and the process model are obtained. At last, the final AKF algorithm and its graphical scheme are given.

It is worth to recall that in the following we will be referring to a generic link's travel time as the time series to forecast.

4.4.1 State-space model

The evolution of the travel time in a link can be modelled considering several approaches. The difference between these being the time necessary for their calibration as well as the amount of data needed to obtain a well-fitted model.

In particular, the authors in [66] used an n -order autoregressive model to describe the dynamics of traffic flow. The model parameters were fitted using n -past same day data such that the one-step ahead forecasting error was minimized. In [42], the authors compared two models, the one proposed in [66] and a random walk model, concluding that the latter presented better results in terms of flow forecasting accuracy, for their specific scenario.

Motivated by the findings in [42] and its straightforward structure, the travel time evolution is modelled by a simple random walk model. The process model is then written as:

$$tt^D(k) = tt^D(k-1) + w^D(k), \quad (4.4.1)$$

$w^D(k)$ is considered a realization of a white Gaussian random process $w^D(k) \sim \mathcal{N}(q^D(k), Q^D(k))$. Because we are assuming that the same day data are the same as the historical ones with the presence of noise and a scale factor, we define the scalar vector $H \in \mathbb{R}^{m \times 1}$ that maps these two informations, where m represents the number of pseudo-observations extracted from the historical data. Therefore, the observation model can be derived by:

$$y(k) = Htt^D(k) + \varepsilon^D(k), \quad (4.4.2)$$

where $\varepsilon^D(k) \sim \mathcal{N}(0, R^D(k))$ is the observation noise, and $y(k) \in \mathbb{R}^{m \times 1}$ the pseudo-observations.

Proposition 1.4.1 (Multi-step ahead Kalman predictor)

Given the state-space model:

$$\begin{aligned} tt^D(k) &= tt^D(k-1) + w^D(k) \\ y(k) &= Htt^D(k) + \varepsilon^D(k), \end{aligned} \quad (4.4.3)$$

where $w^D(k) \sim \mathcal{N}(q^D(k), Q^D(k))$, $\varepsilon^D(k) \sim \mathcal{N}(0, R^D(k))$, and there are available measurements $y(k)$, the multi-step ahead Kalman predictor is given by:

$$\begin{aligned} \hat{tt}^D(k_0 + \tau) &= tt^D(k_0) + \sum_{i=1}^{\tau} q^D(k_0 + i) + \sum_{i=1}^{\tau} K(k_0 + i)(y(k_0 + i) - H\hat{tt}^D(k_0 + i - 1)) - \\ &\quad \sum_{i=1}^{\tau} K(k_0 + i)Hq^D(k_0 + i), \end{aligned} \quad (4.4.4)$$

with K : Kalman gain.

Proof: From the assumptions, we can apply the known solution of the Kalman filter recursively in order to derive Eq. 4.4.1. ■

4.4.2 Pseudo-observations

A natural question that arises is how to combine the historical data in order to generate our pseudo-observations.

Our work in this respect was inspired by the approach proposed by Lin in [51] for flow forecasting. Lin derived a predictor that made use of the same day and historical information in an integrated way by defining two key variables: historical value and historical trend. Lin's approach presented high forecasting accuracy, which supported the idea that the level and the trend was a valuable information to extract from a historical dataset.

The pseudo-observations will be then built upon these findings. The first uses historical average only, while the second uses current data and the historical increment. Both are formulated as follows.

Let:

- $tt^D(k_0)$ be the current travel time value,
- $tt^h(k)$ and $x^h(k) = tt^h(k) - tt^h(k-1)$ be the historical: travel time values and travel time increments respectively of days $h = 1, 2, \dots, D-1$. Where $tt^h(k) \sim \mathcal{N}(\mu_{tt^h}(k), \sigma_{tt^h}^2(k))$ and $x^h(k) \sim \mathcal{N}(\mu_{x^h}(k), \sigma_{x^h}^2(k))$.

Under the assumptions that:

- For a given $k \in [k_0+1, \dots, k_0+\Delta]$, the values $tt^h(k)$ and $tt^D(k)$ are independent realizations of the same stochastic process.
- For a given $k \in [k_0+1, \dots, k_0+\Delta]$, the increments $x^h(k)$ and $x^D(k)$ are independent realizations of the same stochastic process.

The pseudo-observations,

$$y(k) = \begin{pmatrix} y^1(k) \\ y^2(k) \end{pmatrix},$$

can be defined as:

$$\begin{aligned} y^1(k) &= \mu_{tt^h}(k) \\ y^2(k) &= \mu_{x^h}(k) + y^2(k-1), y^2(k_0) = tt^D(k_0). \end{aligned} \tag{4.4.5}$$

As $\mu_{tt^h}(k)$ and $\mu_{x^h}(k)$ are available for $k > k_0$, the computation of the Kalman gain at every step is feasible.

4.4.3 Pseudo-observation noise statistics

The noise $\varepsilon^D(k)$ in Eq. A.4.2, is assumed to be a noise drawn from a zero mean multivariate normal distribution with covariance $R^D(k)$, i.e. $\varepsilon^D(t) \sim \mathcal{N}(0, R^D(t))$.

Moreover, this noise will be assumed to be given by the dispersion present in the historical data, as illustrated in Fig. A.6. Here, at each time instant k , the variance of the realizations $\{tt^h(k)\}$ and $\{x^h(k)\}$, historical values and historical increments respectively, are computed. These values will enter directly into the computation of the Kalman gain as the noise covariance. Therefore, the pseudo-observation whose variance is historical dispersion is smaller, will be considered the most trustworthy.

Hence, the time varying covariance matrix $R(k)$ is defined as:

$$R^D(k) = \begin{pmatrix} \sigma_{tt^h}^2(k) & 0 \\ 0 & \sigma_{x^h}^2(k) \end{pmatrix} \tag{4.4.6}$$

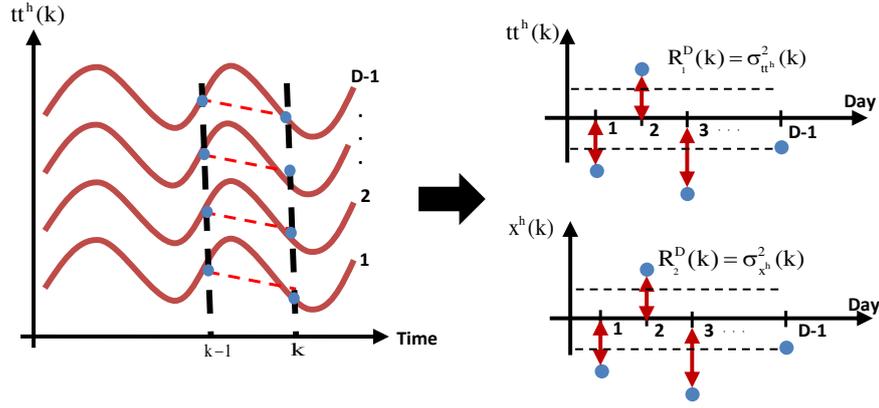


Figure 4.2: Derivation of the observation noise covariance. At each time k , the variance given in the historical values and increments are computed. These variances will be considered to defined the observation noise covariance.

4.4.4 Process noise statistics

In the majority of the real-world applications the statistics of the process noise used in the Kalman filter are not known. However, they can be usually identified using last estimated values of the states and of estimation error covariance of the KF. This is the so-called adaptive Kalman filter (AKF) problem.

Some works have been proposed in literature regarding the adaptive identification of these statistics. Here we can cite the works of [73],[58],[74],[64].

Given its ability to handle both systematic and random errors, as well as being well suited in a real-time framework, this study relies on the strategy proposed in [64] in order to estimate the first and second order statistics of $w^D(k) \sim \mathcal{N}(q^D(k), Q^D(k))$. This approach lies on the innovations property of the filter, and is derived by testing the whiteness of the residual. More details can be found in [64, 83].

In a moving time window of length N defined in $\{k - (N + 1), \dots, k\}$, an approximation of the process noise mean $q^D(k)$ is given by

$$\hat{q}(k) = \frac{1}{N} \sum_{l=1}^N \tilde{q}_l, \quad (4.4.7)$$

where \tilde{q}_l , is the component l of the vector $\tilde{q} \in \mathbb{R}^{N \times 1}$ computed from last estimated state values as: $\tilde{q}_l = \hat{tt}(k - l + 1) - \hat{tt}(k - l)$, $l = \{1, \dots, N\}$.

The estimation of noise covariance is given by:

$$\hat{Q}(k) = \frac{1}{N-1} \sum_{l=1}^N ((\tilde{q}_l - \hat{q}(k))(\tilde{q}_l - \hat{q}(k))^T - \frac{(N-1)}{N}(P_{l-1} - P_l)), \quad (4.4.8)$$

where P_l is the component l of the updated estimation error covariance matrix P .

The choice of N is related to the consistency of the unbiased estimator and the whiteness of the filter. The choice of this parameter will be explained later, and it will be done adequately according to the maximum forecasting horizon and forecasting step considered in this study.

4.4.5 Forecasting scheme and algorithm

In Fig. A.11 is depicted the proposed scheme for travel time forecasting strategy. It is important to notice the use we make of the same day information. The current data (at time k_0) and the same day past data are actively involved in the forecasting process, thus, we have the guarantee that at least for a certain future interval we can capture the current traffic conditions.

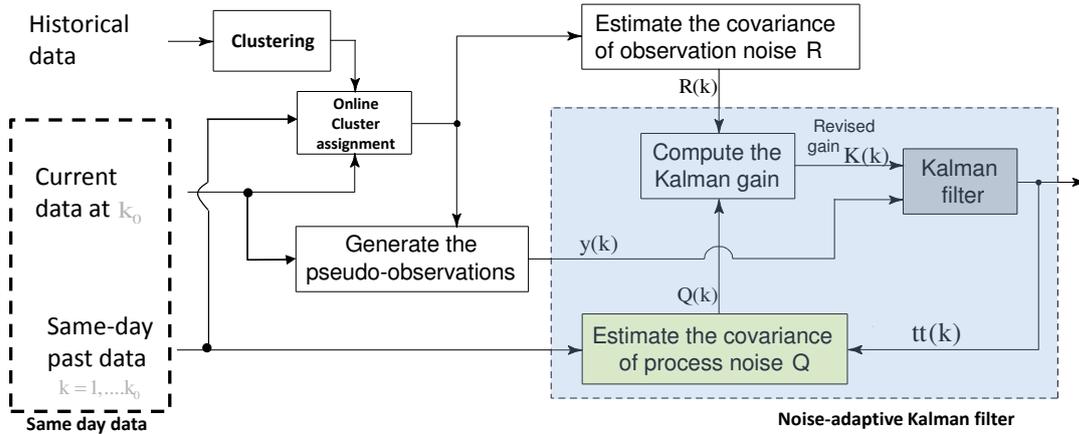


Figure 4.3: Travel time forecasting scheme.

The Adaptive Kalman algorithm is presented in Alg. 3.

Clearly, the proposed algorithm depends on the historical data in order to compute pseudo-observations and their statistics. Therefore, in order for the method to be provided with a “good” dataset, the following sections aim to develop first, an approach for clustering the historical information and second, assign to every new current data a consistent cluster according to the current conditions.

4.4.6 Study of historical data

Motivated by the problem of clustering 24 h travel time profiles within a dataset, this section focuses on the development of a methodology for data clustering, where the similarities between the profiles within limited time periods are indeed difficult to extract.

The main goal of analysing historical traffic information is to identify repetitive patterns in the underlying data. In other words, detect whether the traffic conditions on any given day have similarities with traffic conditions on other days. This study is carried out by using suitable clustering techniques. These techniques organize data into homogeneous groups where the within-group-object similarity is minimized and the between-group-object dissimilarity is

Algorithm 3 AKF algorithm for travel time forecasting, at current time k_0

System:

$$tt(k) = tt(k-1) + w(k), w(k) \sim \mathcal{N}(q(k), Q(k))$$

$$y(k) = Htt(k) + v(k), v(k) \sim \mathcal{N}(0, R(k)), H = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Input:

$$R(k), y(k): \forall k \in [k_0 + 1, \dots, k_0 + \Delta]$$

$$tt(k): \forall k \in [k_0 - N, \dots, k_0]$$

$$\hat{tt}(k_0) \leftarrow tt(k_0), P(k_0)$$

Output: $\hat{tt}(k)$

Process noise initial conditions: $\tilde{q}_l = tt(k_0 - l + 1) - tt(k_0 - l), l = \{1, \dots, N\}$

$$\hat{q}(k_0) = \frac{1}{N} \sum_{l=1}^N \tilde{q}_l$$

$$\hat{Q}(k_0) = \frac{1}{N-1} \sum_{l=1}^N (\tilde{q}_l - \hat{q}(k_0))^2$$

Begin

for $k = k_0 + 1$ to $k_0 + \Delta$ **do**

Prediction step:

$$\hat{tt}^-(k) \leftarrow \hat{tt}(k-1) + \hat{q}(k-1)$$

$$P^-(k) \leftarrow P(k-1) + \hat{Q}(k-1)$$

Update step:

$$K(k) \leftarrow P^-(k)H^T(HP^-(k)H^T + R(k))^{-1}$$

$$\hat{tt}(k) \leftarrow \hat{tt}^-(k) + K(k)(y(k) - H\hat{tt}^-(k))$$

$$P(k) \leftarrow (I - K(k)H)P^-(k).$$

Process noise estimation:

$$\tilde{q}_{k-k_0+N} \leftarrow \hat{tt}(k) - \hat{tt}(k-1)$$

if $(k - k_0) < N$ **then**

$$\hat{q}(k) \leftarrow \hat{q}(k_0)$$

$$\hat{Q}(k) \leftarrow \hat{Q}(k_0)$$

else

$$\hat{q}(k) \leftarrow \frac{1}{N} \sum_{l=k-N}^k \tilde{q}_l$$

$$\hat{Q}(k) \leftarrow \frac{1}{N-1} \sum_{l=k-N}^k \left[(\tilde{q}_l - \hat{q}(k))^2 + \frac{N-1}{N} (P_l - P_{l-1}) \right]$$

end if

end for

maximized.

At early stages, we studied different intuitive classification approaches for the travel time data. The first was to consider profiles of the same day of the week. While the second was by creating three groups: working days, week-ends, and special event days (days before holidays or accidents). However, these initial results were not satisfactory. This was due to high variability that presents the speed data, specially in the congested periods, which makes the clustering process tougher.

Fig. 4.4 shows the time evolution of travel time for 15 working days in the second link of Grenoble south ring. 5 time zones can be easily noticed. The first, in the early morning from 00:00 to 07:00, the second, in the morning congestion jams from 07:00 to 10:00, the third, in the afternoon from 10:00 to 16:00, the fourth, in the afternoon congestion jams from 16:00 to 19:00, and the last, in the night from 19:00 to 24:00. Clearly, we are more interested in assessing the forecasting in the second and fourth area. Therefore, we would aim to group profiles that behave similarly here. This is rather challenging if we perform clustering on the 24 h data. For in general, the techniques put together patterns with majority of “similar” attributes, thus in short time intervals where the forecasting is more crucial, important characteristics may be overlooked.

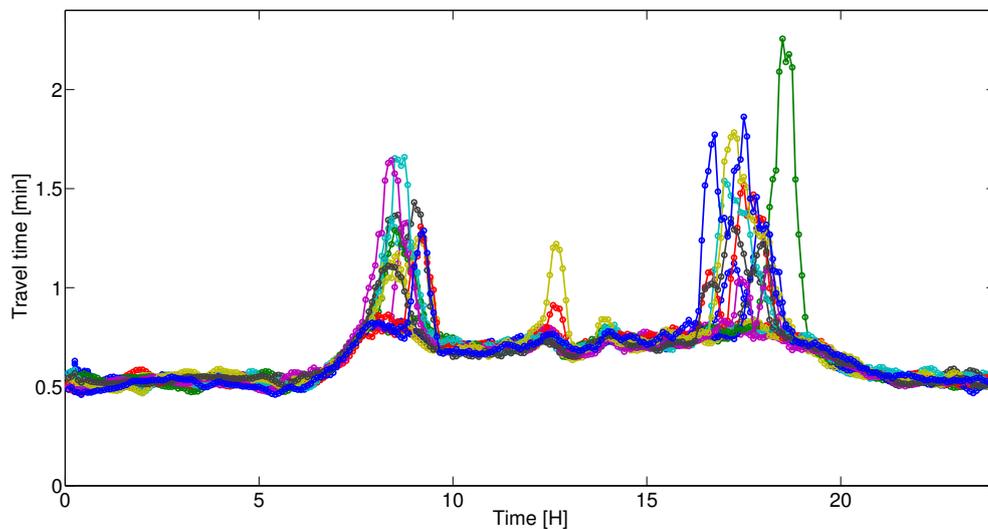


Figure 4.4: 15 working days travel time for the second link of Grenoble south ring.

In the interest of improving the clustering of the historical data, this thesis proposes a different clustering approach. In this, each profile will be divided according to the time zones established as follows.

1. **Time zone 1:** from 00:00 am to 07:00 am.
2. **Time zone 2:** from 07:00 am to 10:00 am.

3. **Time zone 3:** from 10:00 am to 04:00 pm.
4. **Time zone 4:** from 04:00 pm to 07:00 pm.
5. **Time zone 1:** from 07:00 pm to 00:00 am.

This division seems to be reasonable with the dynamics of traffic in the real world. After applying this division, the k-Means algorithm will be applied to the profiles. This technique aims at dividing the dataset in k groups according to some metric, with any concern in the dimensionality of the data. In the following a short review of the method is provided.

4.4.6.1 k-Means algorithm

K-means is a distance-based clustering algorithm that partitions the data into a predefined number of k clusters. This is one of the most commonly used algorithm in its category. The goal is to optimize the objective function:

$$J = \sum_{i=1}^k \sum_{x \in C_i} d(x, \bar{x}_i), \quad (4.4.9)$$

where \bar{x}_i is the center of the cluster C_i , while $d(x, \bar{x}_i)$ is the distance (Euclidean, Cosine, L_1 , etc.) between a point x and \bar{x}_i , Euclidean metric is used in this thesis. Hence the objective is to minimize the distance of each point from the center of the cluster to which the point belongs to. The algorithm is initialized by randomly generating k cluster centers. Then, it assigns each pattern of the dataset to the cluster whose mean is the nearest, and recomputes the new mean. The process continues until convergence is reached. At the end of computation, the algorithm outputs the cluster identifier for each profile and the cluster means.

K-means has several settings, here, the most relevant is how to set the a-priori known number of clusters k . In the following, we will describe a common approach to choose a suitable choice of k in a dataset.

Estimation of number of clusters

It is well known that a drawback of the partition algorithms is that an a-priori knowledge of the number of clusters that better fits a dataset is required. The choice of this number goes beyond visual inspection and in real applications can be estimated experimentally using some metrics. Several works dealing with validation of number of clusters are available in literature, see [30, 43, 92]. The most representative metrics are:

- **Root-mean-square standard deviation (RMSSTD):** it is defined as the square root of the pooled variance of all the profiles. It measures the homogeneity of the formed clusters.

$$RMSSTD = \sqrt{\frac{\sum_{i=1}^k \sum_{j=1}^{n_{C_i}} (x_j - \bar{x}_i)(x_j - \bar{x}_i)^T}{\sum_{i=1}^k (n_{C_i} - 1)}}, \quad (4.4.10)$$

where k is the number of clusters and n_{C_i} the number of objects inside the cluster C_i .

- **R-squared (RS)**: it measures the degree of difference between clusters.

$$RS = \frac{\sum_{j=1}^n (x_j - \bar{x})(x_j - \bar{x})^T - \sum_{i=1}^k \sum_{j=1}^{n_{C_i}} (x_j - \bar{x}_i)(x_j - \bar{x}_i)^T}{\sum_{j=1}^n (x_j - \bar{x})(x_j - \bar{x})^T} \times 100\%, \quad (4.4.11)$$

where n is the number of objects in the dataset and \bar{x} is its mean.

These indexes help to determine the number of potential clusters within a dataset. A typical strategy is the so-called “knee criterion”, which consists of plotting RMSSTD and RS for an increasing number of clusters and of identifying the first steepest knees in the curves, i.e. the first large jump. The knee in the RMSSTD graph indicates that the introduction of another cluster does not add new useful information, while the knee in the RS shows that the data is already divided into well-separated groups.

Only for illustration purpose, let us see how these metrics behave in a experimental setup. For this, let us determine a suitable number of clusters for the 24 h travel time profiles of Fig. 4.4. For this dataset, we have computed the RMSSTD and the RS for an increasing values of k , and the following curves are obtained:

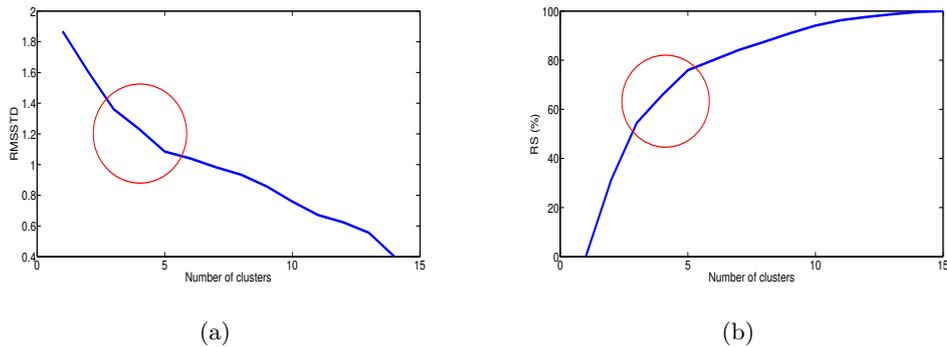


Figure 4.5: Variation with respect of the number of clusters of a)RMSSTD, and b) RS. The goal is to estimate the number of clusters for the travel time dataset in Fig. 4.4.

In Fig. 4.5-a the first break is found at 3 clusters, followed by another one at 5 clusters. This suggests that the optimal number of clusters is to be found in this range. While studying Fig. 4.5-b, the same behavior is observed, the first strongest knees are located at 3 and 5

clusters. As with 5 clusters 76 % of the variance is explained an estimated optimal number of clusters of 5 is a reasonable assumption.

4.4.6.2 Assessment of the proposed approach for historical data clustering

Here, the goal is to assess the proposed clustering methodology using experimental data. For this, we will consider the example in Fig. 4.4 and aim to cluster the dataset. This study will be done in two steps. First, we propose a number of clusters k for each time zone. Second, we provide the clustering results for each one.

Fig. 4.6 depicts the dataset divided in the time zones.

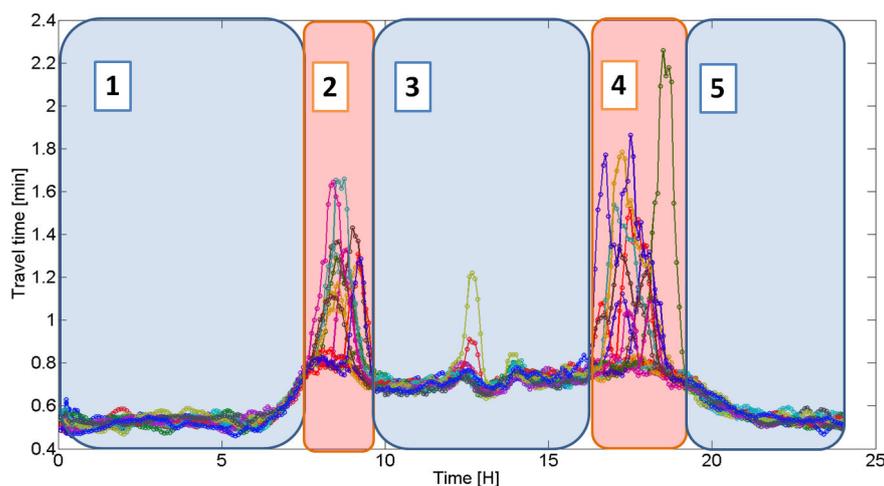


Figure 4.6: 15 working days travel time for the second link of Grenoble south ring.

Number of clusters

Certainly, the traffic conditions vary according to the time zone. Therefore, it would not be computational efficient to assume the same number of clusters in all of them. Moreover, since there are several links in the freeway, we want to generalize the choice of the clusters for each time zone. For this, we look for the worst case of each zone along all links in the freeway, and applying the knee-criteria to this case. The results for each zone, regardless the link considered, are as follows.

- Time zone 1: 2 clusters
- Time zone 2: 4 clusters
- Time zone 3: 4 clusters
- Time zone 4: 4 clusters
- Time zone 5: 2 clusters

Clustering results

For the considered dataset, clustering results are listed in the following plots.

Fig. 4.7 depicts the result for the first zone, where all traffic patterns are rather homogeneous. Fig. 4.8 shows a more challenging case. Here, the objective of capturing and putting together patterns with similar structure is reached: we can notice that in cluster 1 the patterns are homogeneous with consonant congestion build-up and phase-out times. Whereas in cluster 3 the patterns that not presented congestion are grouped together. The main difference between the patterns in the clusters 1, 2, and 4 are the level of congestion as well as congestion peak times.

Fig. 4.9 illustrates how the patterns are concentrated in the clusters 1 and 4, where high homogeneity is achieved. The clusters 2 and 4 on the other hand contain only one element, that by hypothesis will be considered as a non-recurrent event and thus discarded. The non-recurrent event is seen around 12:30 where a peak is observed for both the patterns: the k-means algorithm is able to single out these special patterns. Fig. 4.10 shows the results during the afternoon, from 16:00 to 19:00. Although smaller correlation is observed in the patterns inside the clusters compared to the morning area, specially in the cluster 2, still a high correlation is seen between elements. 2 outliers were singled out, in clusters 1 and 4.

In zone 5, see Fig. 4.11, we notice that the majority of profiles are located in the cluster 2. The algorithm picks up an outlier in cluster 1, which is consistent since it is the continuation of the profile in cluster 4 in Fig. 4.10.

In conclusion, by reducing the problem dimensionality, we achieve a finer clustering, which is important specially in the congestion areas, and second, special events, not considered recurrent, can be singled out.

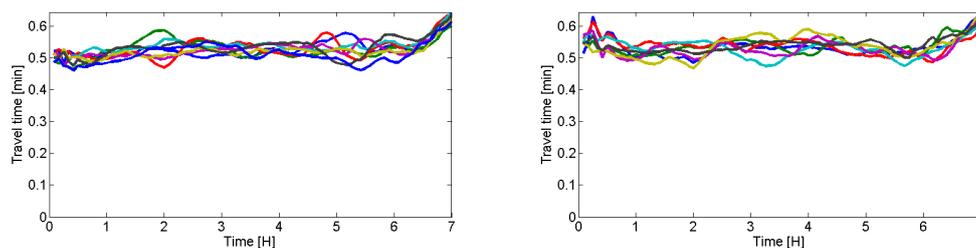


Figure 4.7: Clusters for the time zone between 00:00 to 07:00 of link 2.

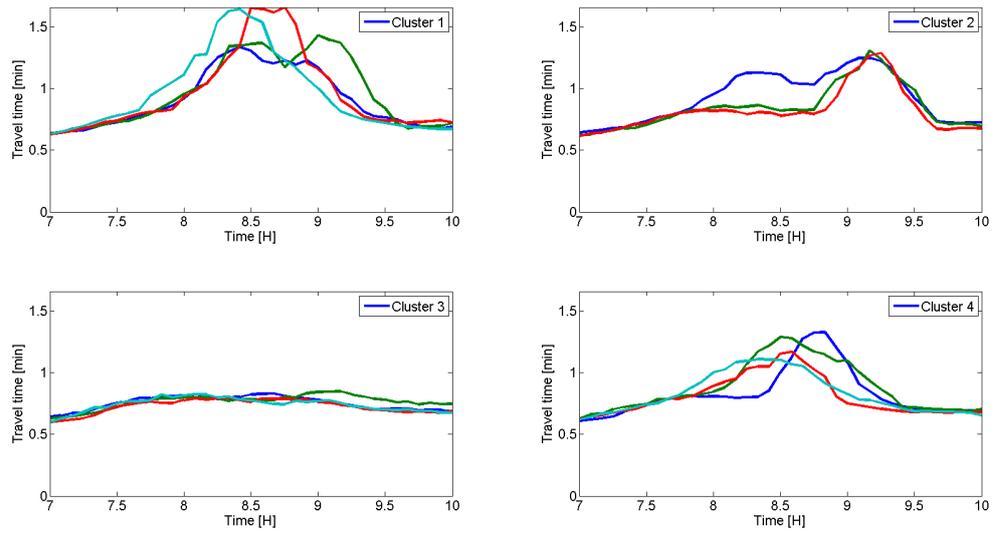


Figure 4.8: Clusters for the time zone between 07:00 to 10:00 of link 2.

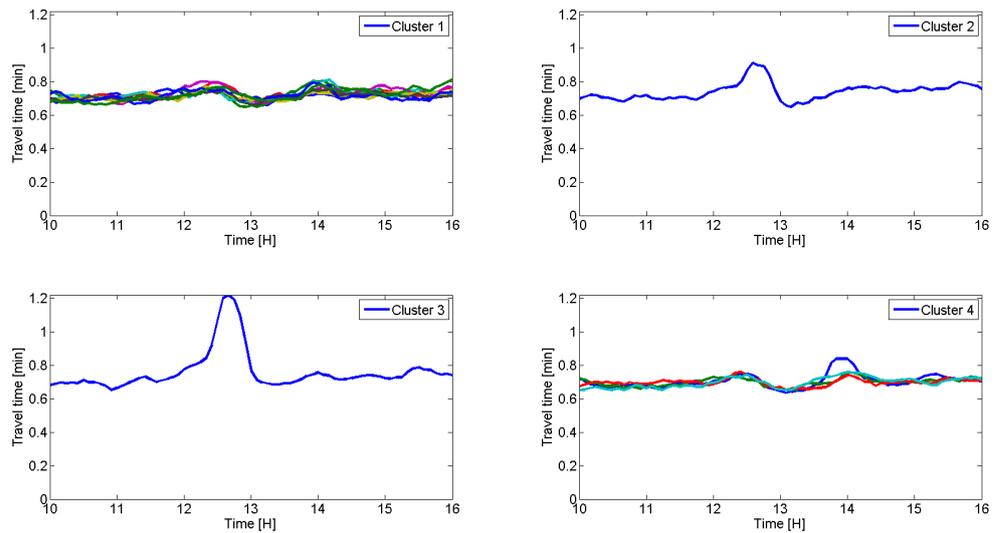


Figure 4.9: Clusters for the time zone between 10:00 to 16:00 of link 2.

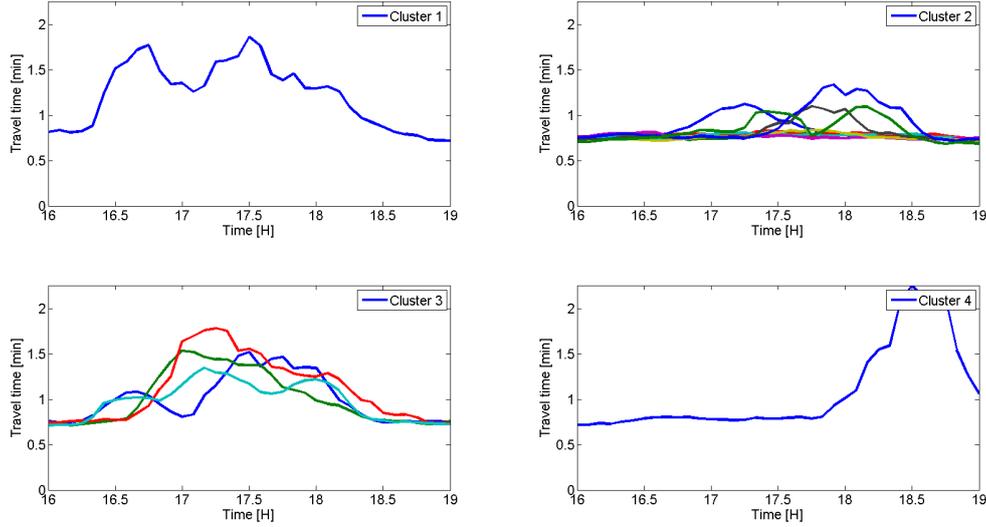


Figure 4.10: Clusters for the time zone between 16:00 to 19:00 of link 2.

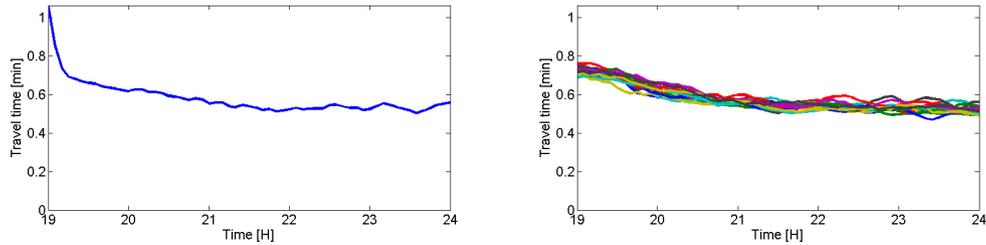


Figure 4.11: Clusters for the time zone between 19:00 to 24:00 of link 2.

4.4.7 Real-time classification of the current data

In the previous section, we have analysed differences in a dataset in order to place similar historical profiles in the same cluster. This section on the other hand, aims to assign a “new” data to a given cluster.

It was already established that clusters containing a single element would be discarded as they are likely to represent an exceptional event. Inside each non trivial cluster a center is located. In order to assign a cluster to the current data, a metric will be defined, and the cluster that minimizes the distance between the two will be chosen (and thus provide the pseudo-observations and model parameters).

The Mean Absolute Deviation (MAD) has been chosen as distance metric:

$$MAD_i = \frac{1}{N} \sum_{l=k_0-N}^{k_0} \left| tt^D(l) - \overline{tt}_i^h(l) \right|, i = \{1, \dots, k\} \quad (4.4.12)$$

whereby N , k_0 , $tt^D(l)$, and $\bar{tt}_i^h(l)$ stand for the number of time points preceding the current time, the current time, the current day data (travel time), and the mean of cluster i .

The choice of the cluster depends upon the distance metric in Eq. A.4.7. As such, we cannot rule out the possibility that the distance between the current data and the means is so small that the cluster discrimination becomes questionable. This case may be encountered especially at the limits of the time intervals. Considering this, two possible scenarios may occur:

- It is difficult to discern the cluster at the beginning of one time interval. The pseudo-observations will be computed using all the non trivial clusters. The advantage of doing this is that all the stable conditions are considered, the means are computed within homogeneous clusters and special events are discarded.
- It is difficult to discern the cluster in the middle of one time interval. The forecasting will be done by using the previous cluster until the selection is clear enough.

In the following the online cluster assignment will be validated. The goal is to study what is the success rate that Eq. A.4.7 can guess the right cluster. In particular, to estimate what is a good choice for the length of the time window, N .

Assessment

The analysis will proceed as follows: starting at the first freeway link and first time zone with $N = 0$, we take one profile, whose cluster is known a priori, and scan at every time step the closest cluster mean according to Eq. A.4.7, at the end of the time zone we then compute the percentage of times the right cluster was chosen and repeat the process using a *leave-one-out cross-validation* technique for the rest of the profiles, storing then all the percentages. The same is done for the subsequent time zones and consequently the subsequent links. After the computations are done for $N = 0$ we repeat the process for $N = \{1, \dots, 19\}$. An “optimal” value of N is finally chosen and a numerical quantification of the online cluster assignment is provided.

Fig. 4.12 shows the mean of the percentage of right cluster assignment, for the five time zones, when N varies between 0 and 19. The first notice an improvement on the right matching while increasing N . That is due to the fact that at the current time the cluster decision is taken by averaging current and past distances. Second, as expected the results for each time zone are different: in fact, the first, third, and fourth ones have similar performances in terms of percentage of right matching and the variability while changing N . Clearly, the negative impact of choosing a wrong cluster in the first and fifth time zone, on the forecasting accuracy can be almost neglected.

As for the second interval, it is observed the higher improvement when increasing N , whereas the fourth presents results rather consistent with respect to N . For $N > 6$ the second interval

has higher odds to capture the right cluster.

It can be also noticed that as N increases, the lines tend to become closer, which suggests that adding more same day past data at some point stops being relevant. The mean percentage of right assignments goes from 67% to 90% while varying N from 0 to 19.

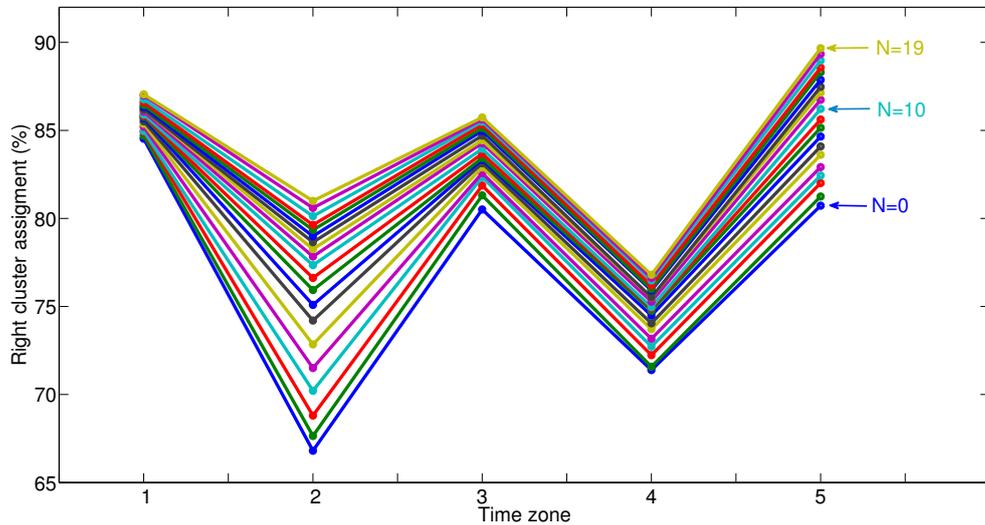


Figure 4.12: Temporal assessment of the real-time classification of the current data from different values of N .

For the final implementation a value of $N = 10$ was chosen: it represented a good trade-off between performance and manageable window length.

Fig. 4.13 exhibits the assessment of the method for the different time intervals with $N = 10$. The results are shown using the mean value and standard deviation of the percentage for all links at the specific interval. In the worst cases, second and fourth time zone, we obtain in terms of mean and standard deviation a success of right matching of 76% and 75% respectively.

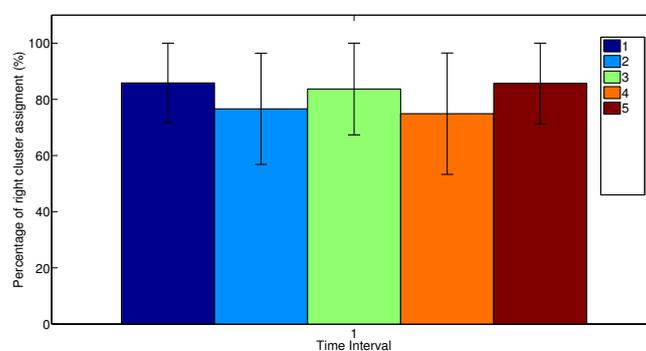


Figure 4.13: Temporal assessment of the real-time classification of the current data for $N=10$.

4.5 Experimental results: travel time forecasting

In the sequel, we address the assessment of the travel time forecasting methodology presented in this chapter. The experiments are organized as follows. First, we present the experimental location chosen. Second, the test scenario will be given. Finally, we generalize the forecasting assessment using a cross-validation technique for the available dataset.

To evaluate the forecasting performance, we use the absolute percentage error (APE) defined as follows:

$$APE(k) = \frac{|tt(k) - \hat{tt}(k)|}{tt(k)} \times 100\% \quad (4.5.1)$$

where $tt(k)$ and $\hat{tt}(k)$ are the true and forecasted value of the travel time at time k , respectively. The forecasting step and aggregation time will be chosen equal to 5 min. The length of the time window N for the process noise estimator will be then chosen equal to 5. In order to track the travel time in the entire freeway up to an horizon of 45 min, the travel at each link is forecasted up to 1h and 30min in the future.

1. Experimental location:

Fig. A.8 shows the path considered. The scenario chosen is a vehicle that wants to traverse the entire freeway.

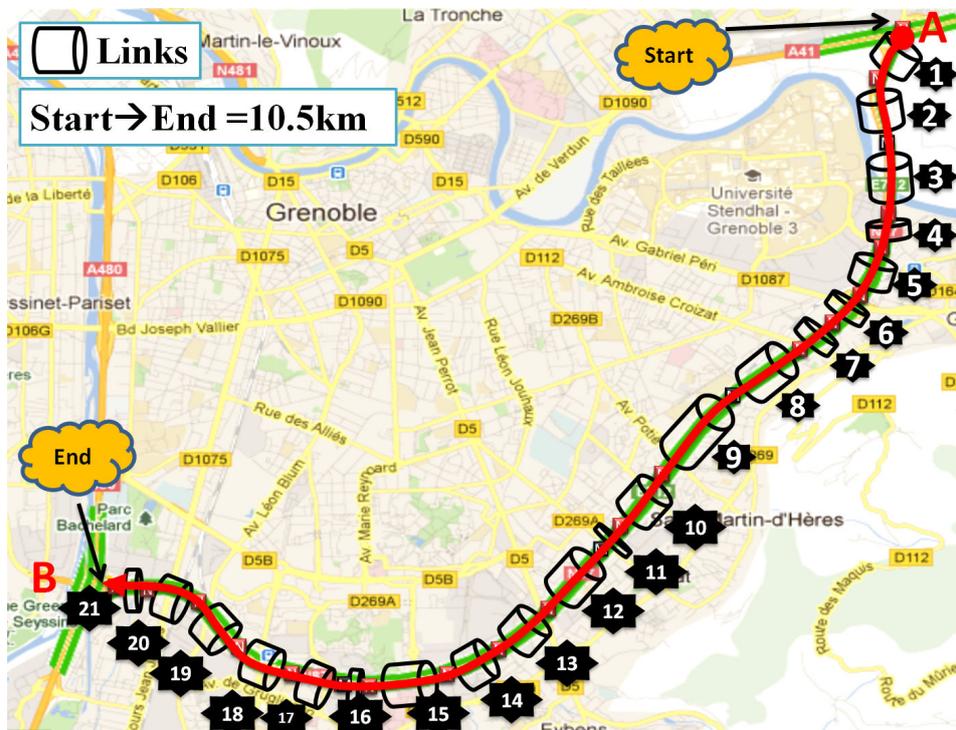


Figure 4.14: Freeway chosen path for the numerical tests, from start: Meylan to end: Rondeau.

For the available dataset and the selected path. Fig. A.9 shows the travel times computed.

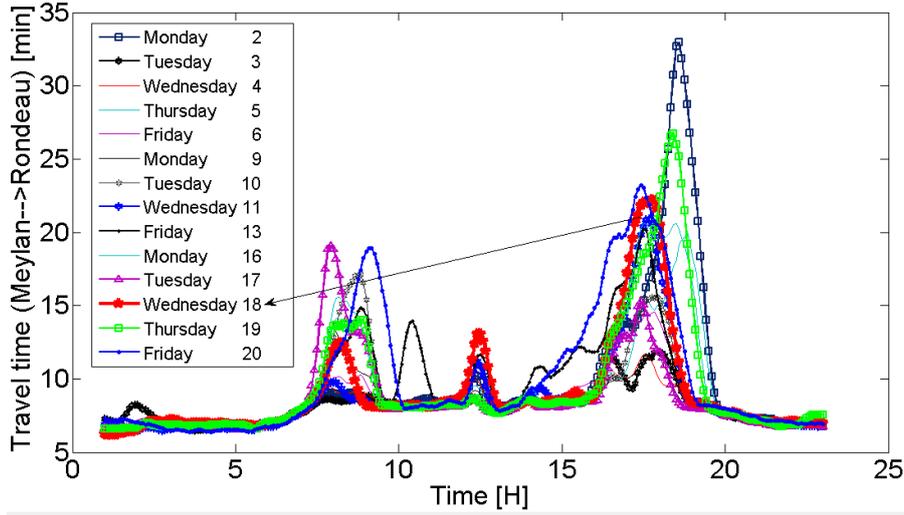


Figure 4.15: Estimated travel time experienced by a driver that traverses the entire Grenoble south ring at different departing times for 15 different working days in September 2013. From the 2nd to the 20th of September.

Analysing the figure, we can notice several recurrent patterns:

- Morning peak: between 8:00-9:00.
- Noon peak: between 12:00-13:00.
- Afternoon peak: between 16:30-18:30.
- The longest time to traverse the freeway was approximately 33 min. That corresponds to an average speed of 19 km/h.
- The shortest time observed was around 7 min. That corresponds with a average speed of 90 km/h.

2. Scenarios:

In order to illustrate the results, we chose the Wednesday 18th of September (cf. Fig. A.9). For this day, two forecasting scenarios will be considered. The first in the morning at 08:45 and the second at 17:15. These hours were considered because of the level of congestions. In particular, at 08:45 we can study the congestion phase-out time, and at 17:15 the congested state. For each one, the multi-step forecasts and the vehicle trajectory upon speed contour are provided.

The proposed approach will be also compared with the average of the historical data in the dataset defined as:

$$\hat{\varphi}(k) = \mu_{tt}(k), \quad (4.5.2)$$

where $\mu_{tt}(k)$ is the average value of the historical profiles of travel time for $k = \{k_0 + 1, \dots, k_0 + \Delta\}$.

Fig. A.10 depicts the results. In Fig. A.10-a it is observed a high accuracy using the proposed methodology. The forecasting tracks down accurately the congestion phase-out time. In this scenario it is also observed a fairly good forecasting result using the average of the historical. This is consistent with Fig. A.10-b. The trajectory of the vehicle is forecasted accurately. Fig. A.10-c shows the result at a current time 17 : 15. Here, the proposed method succeeds to capture the travel time in congested conditions, however compared to the morning period less accurate results are found, which is normal given that the road is more congested. Fig. A.10-d shows the forecasted trajectory. Clearly, the traffic conditions affects the accuracy of the forecasting. We can observed here that the longer the section considered, the higher the forecasting error. This was foreseen, given that the forecasting error accumulates from link to link.

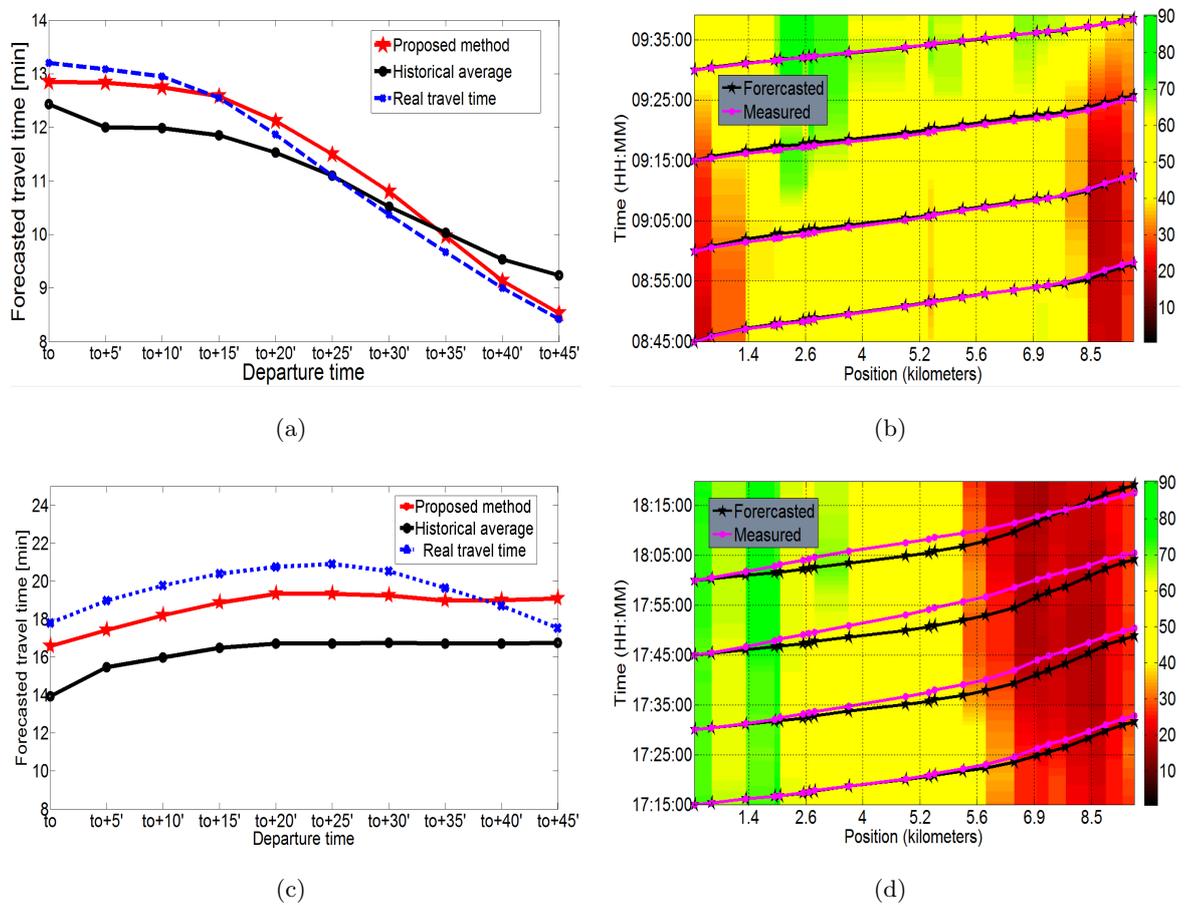


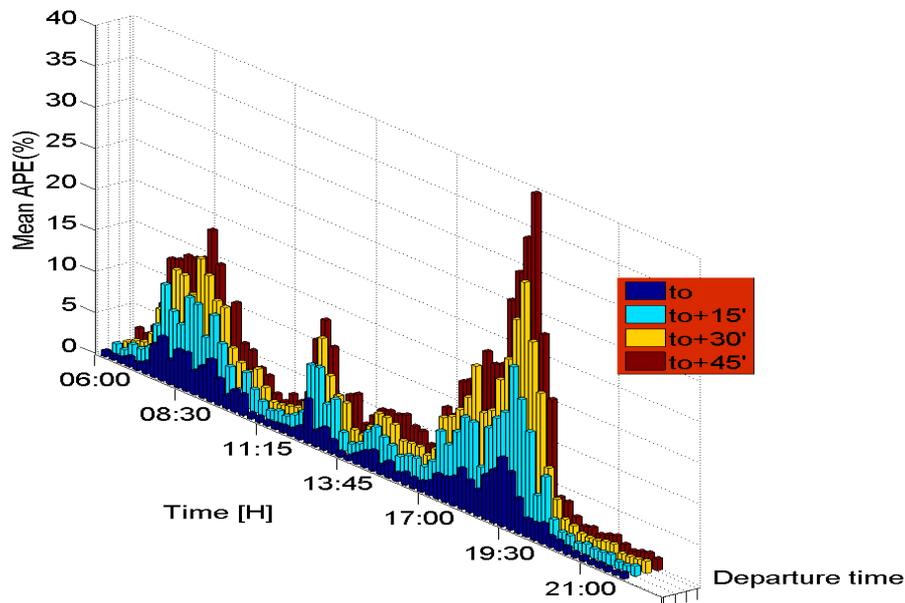
Figure 4.16: Illustrative results of the proposed forecasting approach. a) multi-step ahead forecasts at $t_0 = 08:45$. b) forecasted and measured vehicle trajectory upon speed contour at $t_0 = 08:45$. c) multi-step ahead forecasts at $t_0 = 17:15$. d) forecasted and measured vehicle trajectory upon speed contour at $t_0 = 17:15$.

3. Cross-validation:

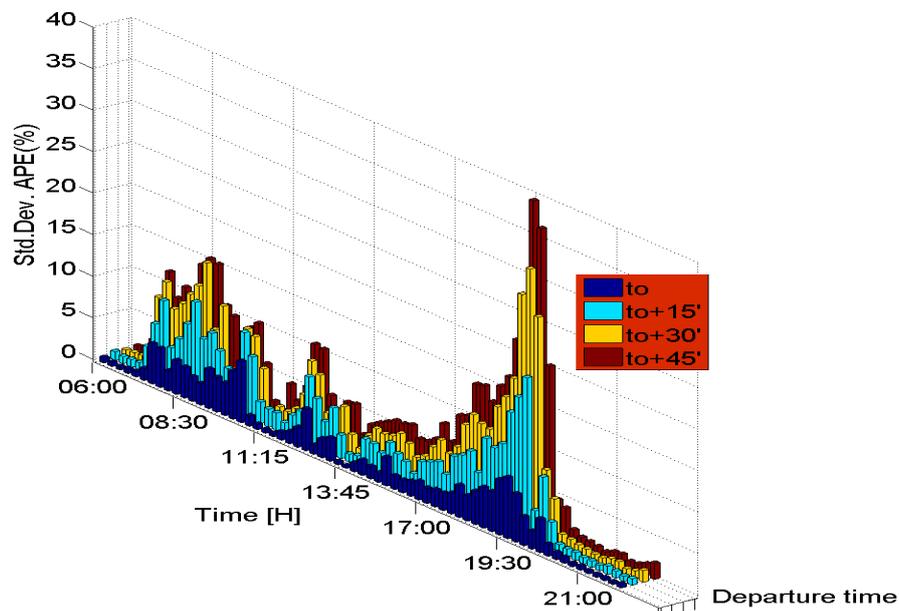
The tests will be carried out as follows: first, a leave-one-out cross-validation technique will be adopted, as such, one day at a time will be chosen as validation data and the other fourteen as training data. For the selected day, we forecast at different horizons

from 06:00 to 22:00 at steps of 5 min, for each horizon we compute the APE(%).

- We extract the mean and standard variation of the realizations in the Fig. 4.17-a and Fig. 4.17-b respectively.



(a)



(b)

Figure 4.17: Mean and standard deviation the APE (%) for a moving window from 06:00 am to 10:00 pm for different departure times. Every bar represents the mean or the standard deviation of the realizations computed for the fifteen days. a)Mean, b)Standard deviation.

Several points are worth mentioning:

- The forecasting error increases by increasing the departure time. This was

expected, as by enlarging the forecasting horizon we can not guarantee a good match between the forecasted and real traffic conditions.

- The highest peaks in the errors are found during three time intervals: morning, noon, and afternoon. The last being the most important (around 17:30). Hence, we conclude that in terms of forecasting accuracy, this is the less accurate period of the day, followed by the morning time around 08:30.
- Comparison between proposed methodology and average of the historical.

This test aims to compare the proposed method with the historical average. For all the realizations for which we computed the APE (%), we have evaluated the cumulative distribution function. Fig. 4.18 depicts the results for 4 different horizons. Experimentally, this function reveals the maximum forecasting error committed for a given percentage of the realizations.

Some conclusions can be drawn from the figures.

- As supported by Fig. 4.18, the smaller the forecasting horizon the higher the forecasting accuracy.
- The performance of the proposed methodology (dashed lines) is better or comparable than the achieved by historical average. As the horizon increases the difference between the two is less discernible. This is an interesting fact, as shows that the further we look into the future the more the proposed method depends only on the the historical average. Therefore, for long horizons the forecasting can be performed using directly the historical.
- Tab. 4.1 depicts the maximum forecasting error obtained for 90% of the realizations for different horizons. It is seen that at the current time ($\Delta = 0$), the proposed method is more accurate in a 133.33%. While the horizon increases, the error becomes more similar. For $\Delta = 45'$ the error difference decreases to 8%. In conclusion, for the scenario considered, it is observed that the proposed method was able to forecast 90% of the time with a maximum error of 23% up to an horizon of 45'.

Table 4.1: Maximum error observed for 90% of the realizations for different forecasting horizons.

Departure time	$\Delta = 0$	$\Delta = 15'$	$\Delta = 30'$	$\Delta = 45'$
Proposed method (%)	6	13.4	19	23
Historical average (%)	14	20	22	25

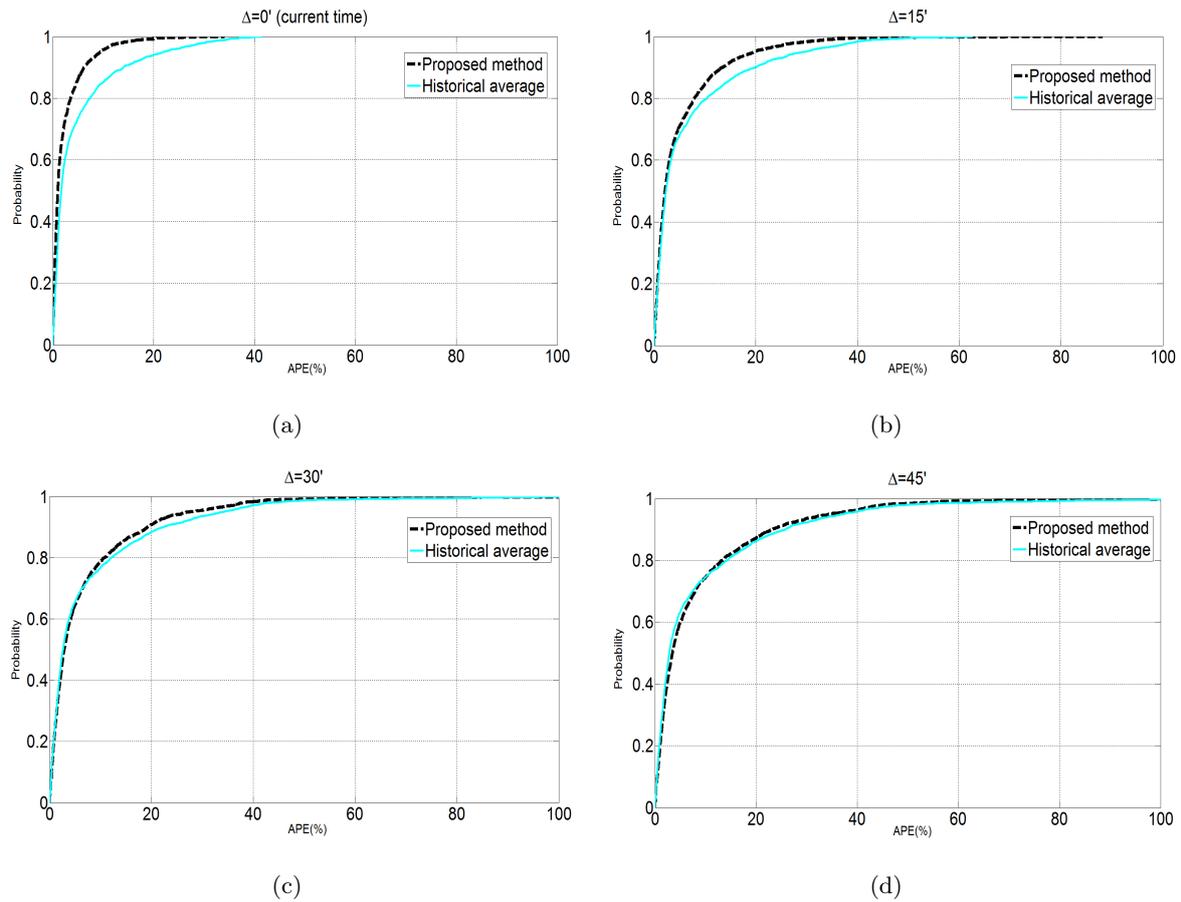


Figure 4.18: Empirical cumulative distribution function evaluated with the APE at different forecasting horizons. a) $\Delta = 0$ (Current time), b) $\Delta = 15'$, c) $\Delta = 30'$, d) $\Delta = 45'$.

4.6 Conclusions

In this chapter we developed a signal-based approach, based only on speed measurements, in order to tackle the multi-step ahead travel time forecasting. The choice of this approach was very pragmatic, and it was motivated by the need of developing a predictor able to work and achieve relatively good forecasts, even when the measurements were not 100% reliable.

The new contributions in this chapter were as follows.

- **Noise adaptive Kalman filter approach (AKF).** We proposed a novel approach for traffic forecasting. In particular, it addressed the forecasting of travel times inside the links enclosed by the section of interest. The approach was based on a noise adaptive Kalman filter, which used an adequate combination of historical and current data in order to replace the lack of system's observations. The statistics of the observation noise were computed from the historical data, while those of the process noise using an online unbiased estimator.
- **Approach for clustering a historical travel time dataset.** Given the high variability

in a dataset of travel time profiles, or equivalently of speed, we proposed a strategy to deal with the clustering of this dataset. In particular, it tackled the reduction of dimensionality of the clustering problem by dividing the time series in time zones, and applying the k-means algorithm to each one.

- **Real-time classification of the current data.** We proposed an algorithm that assigns a cluster to every new current data. This algorithm used a distance metric to compare a fixed length moving window composed of the current and same day past data, with the center of each cluster. We showed that with a window length equal to 10, in the worst cases we have a success rate of right classification of 75%.
- **Assessment of the forecasting methodology on experimental data.** For the considered dataset, we saw that the method proposed could forecast 90% of the time with a maximum error of 23% up to an horizon of 45'. Furthermore, we observed that the larger the forecasting horizon, the more the forecasts depend only on the historical data. Meaning that the current information loses importance when Δ increases, thus only the historical data is relevant.

Chapter 5

Short-term multiple ahead travel time forecasting: model-based approach

5.1 Overview

The goal of this chapter is to develop and validate a second methodology for the travel time forecasting problem. This methodology will be built upon a traffic macroscopic model, from which we will derive a state observer scheme. This new approach allows the discretization of the section of interest in smaller portions, referred later as cells, to later reconstruct and forecast the traffic conditions inside of them. The information used in this methodology are the section's boundary conditions, flow and speed, and the input and output flows enclosed by the section.

Noting that the number of detector stations is the same as in Chap. 4, we could look at the problem as if trying to segment trips in smaller pieces using “virtual” stations. Therefore, a finer tracking of the traffic dynamics can be achieved. This effect was experimentally remarked in [97]. In this work the authors noticed the beneficial effects of having halfway measurements in long sections when assessing the travel time forecasting.

Without getting into details, the complete scheme for travel time forecasting is as given in Fig. A.11, where all the signals are indexed by time k , k_0 represents the current time, and Δ the forecasting horizon.

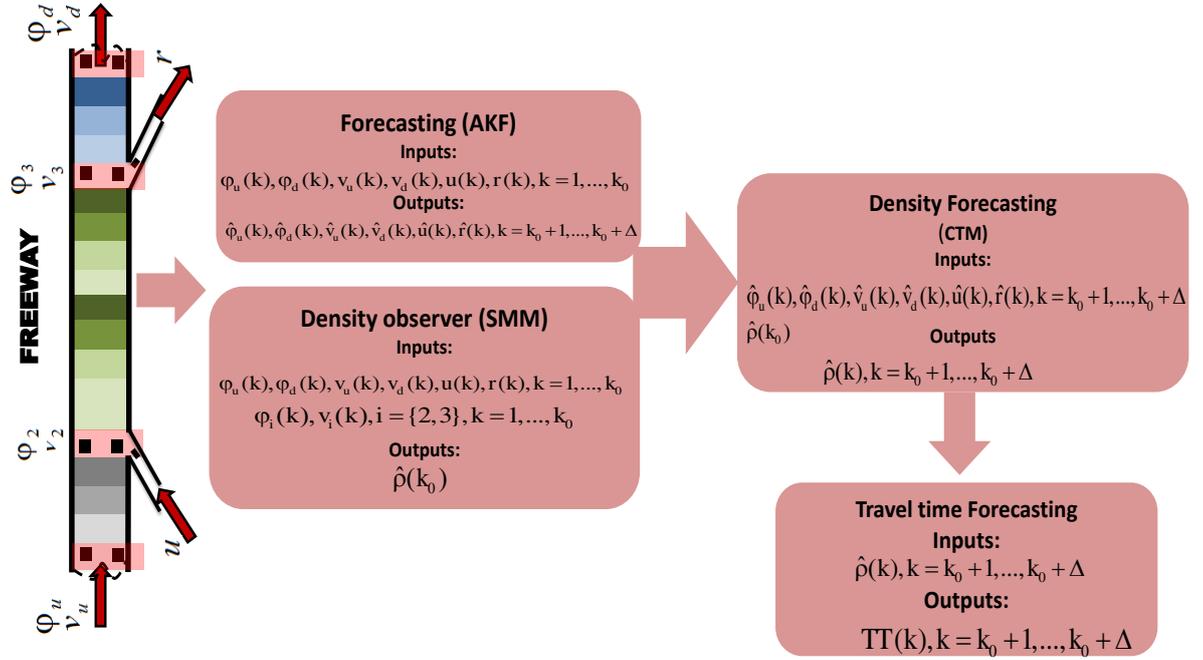


Figure 5.1: Schematic representation of the model-based method proposed for travel time forecasting.

Each of the four main blocks in the scheme is a key part for achieving the final objective. As such, experimental results will be given in each.

The rest of the chapter will be organized as follows. First, in Sec. 5.2 the description of the traffic data for the experimental tests will be given. Second, in Sec. A.5.2 the general problem statement will be presented. Then in Sec. 5.4 we derive the density observer based on SMM. Later in Sec. 5.5, we recall the AKF approach for multi-step traffic forecasting. In Sec. 5.6, we provide the relation between forecasted density and forecasted speed and travel time. Finally, in Sec. 5.7 the final experimental results for travel time forecasting are presented.

5.2 Data description

One of the core components in this chapter is the validation of the proposed methodologies with real traffic data. The dataset will be created from data collected in the period: 27th of February to the 15th of March of 2014. Therefore 12 days are available, without week-ends.

The experimental raw data, downloaded from GTL, is pre-processed using the strategies given in Chapter 3.

5.3 Statement of the problem

Considering a freeway section as shown in Fig. A.12. Given then mainstream measurements of flow and speed, and input and output flows up to the current time k_0 of the current day D , $\varphi_i^D(k), v_i^D(k), i = \{1, \dots, n + 1\}, u_l^D(k), l = \{1, \dots, m\},$ and $r_j^D(k), j = \{1, \dots, n\},$ respectively,

and historical data only of the section's boundaries and the input and output flows from preceding days $h = \{1, 2, \dots, D - 1\}$, $\varphi_1^h(k)$, $v_1^h(k)$, $\varphi_{n+1}^h(k)$, $v_{n+1}^h(k)$, $u_l^h(k)$, $r_j^h(k)$. The aim is to forecast, using an online multi-step ahead strategy, the section's travel time $T^{\hat{T}}(k) \forall k \in [k_0 + 1, \dots, k_0 + \Delta]$. Note that φ_1 and φ_{n+1} denote respectively the upstream flow φ_u and the downstream flow φ_d , analogously the speeds.

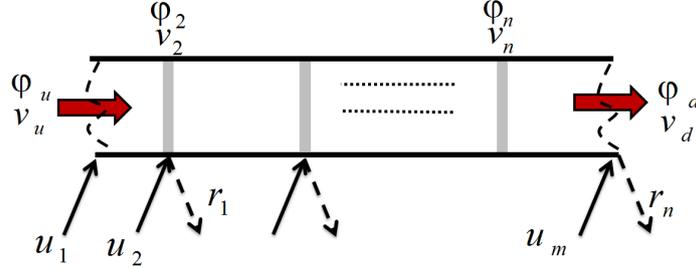


Figure 5.2: Abstraction of a freeway section. The grey stripes represent detector stations.

5.4 Density reconstruction

5.4.1 LWR model

The first and simplest continuous macroscopic freeway model, involving only density ρ and based on the car conservation law, is the first order Lighthill-Whitham-Richards (LWR) model initially proposed in [48, 72]. The LWR model can be easily derived in terms of a PDE, whose solutions both analytical and numerical can be also known. As this work is based on the numerical solution of the LWR model, we will focused on this as the final objective. Thus, we will first derive the LWR PDE model and thereafter present its numerical solution. In order to do so let us first state the conservation law.

For this, let us denote $\rho(x, t)$ and $\varphi(x, t)$ as the space-time distributions of density and flow respectively, the law states that the rate of change of the total number of vehicles in the section $[x_0, x_1]$ is given by difference in flows at x_0 and x_1 , i.e.:

$$\frac{d}{dt} \int_{x_0}^{x_1} \rho(x, t) dx = \varphi(x_0, t) - \varphi(x_1, t). \quad (5.4.1)$$

Assuming that ρ is derivable with respect to time, and φ with respect to space, Eq. A.5.1 is restated as:

$$\int_{x_0}^{x_1} \partial_t \rho(x, t) dx = \int_{x_0}^{x_1} \partial_x \varphi(x, t) dx. \quad (5.4.2)$$

Originating the differential form of the conservation law Eq. A.5.1:

$$\begin{aligned}\partial_t \rho(x, t) + \partial_x \varphi(x, t) &= 0 & (x, t) \in (-\infty, -\infty) \times (0, T) \\ \rho(x, 0) &= \rho_0(x) & x \in (-\infty, -\infty)\end{aligned}\tag{5.4.3}$$

In order to express the flow distribution in terms of velocity, we make use of the relationship first introduced in Chapter 1:

$$\varphi(x, t) = \rho(x, t)v(x, t).\tag{5.4.4}$$

Recalling now that the constitutive assumption of the LWR model, motivated by experimental data, is that the vehicles tend to travel at an equilibrium speed $v = V(\rho)$ for all locations at all times, Eq. 5.4.4 is written as:

$$\varphi(x, t) = \rho(x, t)V(\rho(x, t)) = \Phi(\rho(x, t)).\tag{5.4.5}$$

Then, the LWR PDE model can be expressed as follows:

$$\partial_t \rho + \partial_x \Phi(\rho) = 0.\tag{5.4.6}$$

In the transportation engineering community, the flow function $\Phi(\rho)$ is also known as the fundamental diagram, and is generally assumed to be concave and piecewise C^1 .

Analytically, solutions to the LWR PDE can be constructed through the method of characteristics, see [27, 44] for more details. However, even from smooth initial conditions, shocks may develop in finite time, and classical (smooth) solutions to the PDE may not exist. These shocks are the result of the encounter of two wave fronts. Numerically, one of most common methods for solving Eq. A.5.2 is the Godunov's discretization scheme [45]. Godunov's scheme is a conservative scheme based on Riemann problems.

The discrete model is obtained first by sub-dividing the spatial domain, x , into finite volumes or cells of length Δx_i with cell centres indexed as i . The function $\rho(x, t)$ is then approximated with a piecewise constant function $\rho_i(k)$, where $\rho_i(k)$ is constant in each cell. For a particular cell i , located between the coordinates $x_{i-1/2}$ and $x_{i+1/2}$ at time t , we can define its average density as:

$$\rho_i(t) = \int_{x_{i-1/2}}^{x_{i+1/2}} \rho(x, t) dx.\tag{5.4.7}$$

Integrating in space Eq. A.5.2, along the interval $[x_{i-1/2}, x_{i+1/2}]$, it is obtained:

$$\dot{\rho}_i(k) = -\frac{1}{\Delta x_i} (\Phi(\rho(x_{i-1/2}, t)) - \Phi(\rho(x_{i+1/2}, t))).\tag{5.4.8}$$

Exact time integration of Eq. 5.4.8, along $t_k = k\Delta t$ with the step Δt , yields the update formula:

$$\rho_i(k+1) = \rho_i(k) - \frac{1}{\Delta x_i} \int_{t_k}^{t_{k+1}} (\Phi(\rho(x_{i-1/2}, t)) - \Phi(\rho(x_{i+1/2}, t))) dt.\tag{5.4.9}$$

By applying the forward Euler method, the fully discrete model of the LWR PDE in Eq. A.5.2 writes:

$$\rho_i(k+1) = \rho_i(k) - \frac{\Delta t}{\Delta x_i} (\Phi(\rho(x_{i-1/2}, k)) - \Phi(\rho(x_{i+1/2}, k))). \quad (5.4.10)$$

As mentioned, the numerical values of $\rho(x_{i-1/2}, k)$ and $\rho(x_{i+1/2}, k)$ are computed solving a local Riemann problem. For this let us consider that (see Fig. A.13-a):

$$\rho(x, k) = \begin{cases} \rho_{i-1}(k) & \text{if } x < x_{i-1/2} \\ \rho_i(k) & \text{if } x > x_{i-1/2} \end{cases} \quad (5.4.11)$$

$$\rho(x, k) = \begin{cases} \rho_i(k) & \text{if } x < x_{i+1/2} \\ \rho_{i+1}(k) & \text{if } x > x_{i+1/2}. \end{cases}$$

Assuming a concave function for $\Phi(\rho)$, see as an example Fig. A.13-b, where ρ_c corresponds the point in which the function changes the sign of its derivative (called critical density), and ρ_m is the maximum density, also referred to as "jam density". The value of the density at the cell's interface is fixed by the direction in which the shock wave travels. The shock waves starting always at the discontinuity. By introducing the speed of this shock wave (known as the Rankine-Hugoniot relation) corresponding to the interface between the cells $i-1$ and i as:

$$z(k) = \frac{\Phi(\rho_{i-1}(k)) - \Phi(\rho_i(k))}{\rho_{i-1}(k) - \rho_i(k)}, \quad (5.4.12)$$

the solution at $x_{i-1/2}$ is then as follows:

$$\begin{aligned} \rho(x_{i-1/2}, t) = \rho_{i-1}(k) & \quad \text{if} \quad \partial_\rho \Phi(\rho_{i-1}(k)) \geq 0 \wedge \partial_\rho \Phi(\rho_i(k)) \geq 0 \\ \rho(x_{i-1/2}, t) = \rho_i(k) & \quad \text{if} \quad \partial_\rho \Phi(\rho_{i-1}(k)) < 0 \wedge \partial_\rho \Phi(\rho_i(k)) < 0 \\ \rho(x_{i-1/2}, t) = \rho_{i-1}(k)(z > 0) & \quad \text{or} \quad \text{if} \quad \partial_\rho \Phi(\rho_{i-1}(k)) \geq 0 \wedge \partial_\rho \Phi(\rho_i(k)) < 0 \\ \rho(x_{i-1/2}, t) = \rho_i(k)(z < 0) & \\ \rho(x_{i-1/2}, t) = \rho_c & \quad \text{if} \quad \partial_\rho \Phi(\rho_{i-1}(k)) < 0 \wedge \partial_\rho \Phi(\rho_i(k)) \geq 0 \end{aligned} \quad (5.4.13)$$

The last case refers to the transonic case [45]. The Riemann solution at interface $x_{i+1/2}$ is carried out in the same way.

In order to ensure numerical stability, the time and space steps are related by the Courant-Friedrichs-Lewy condition [45]: $v_{\max} \Delta t \leq \Delta x$, where v_{\max} denotes the maximal characteristic speed (referred to as free-flow speed). This conditions states that if a wave is moving across a discrete spatial grid and we want to compute its amplitude at discrete time steps of equal length, then this length must be less than the time for the wave to travel to adjacent grid points.

For the sake of clarity of the following, these definitions are in order:

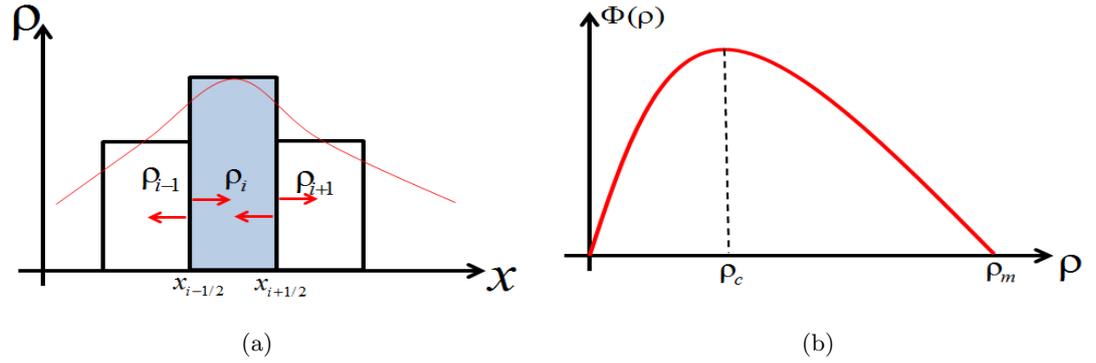


Figure 5.3: (a) Sub-division of the spatial domain in three cells, we are interested in computing the density values at the interfaces of the cell i , (b) Concave flow function $\Phi(\rho)$.

Definition 5.1. A cell i is considered free (F) if its associated density ρ_i is equal or below its critical density $\rho_{c,i}$, else, it is considered to be congested (C) ($\rho_{c,i} < \rho_i \leq \rho_{m,i}$).

Definition 5.2. A cell's interface, $x_{i-1/2}$ for instance, is considered free (F) if the shock wave starting at this point travels forward ($z > 0$), else is considered to be congested (C) ($z < 0$). In the former case $\varphi(x_{i-1/2}, t)$ is said to be diffusive (wave fronts moving downwards), while in the latter advective (wave fronts moving upwards), they will be denoted as $\vec{\varphi}(x_{i-1/2}, t)$ and $\overleftarrow{\varphi}(x_{i-1/2}, t)$ respectively.

5.4.2 CTM model

In the rest of this thesis we will work with the so-called Cell Transmission Model (CTM) proposed by Daganzo in [20]. This model is particularly well designed for model-based control and estimation, owing to its analytical simplicity and ability to reproduce congestion wave behavior. The CTM is indeed the first order discrete Godunov approximation of Eq. A.5.2 under the assumption that the fundamental diagram has a triangular form as given in Fig. A.14.

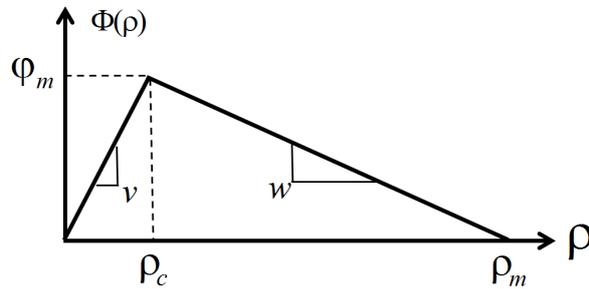


Figure 5.4: The fundamental diagram. The shape is determined by the parameters: φ_m - maximum capacity, v - free flow velocity, w - congestion wave speed.

CTM describes in discrete time the traffic evolution on a given freeway by evaluating the flow and density at finite number of intermediate points at different time steps. This model assumes a freeway representation of a sequence of n cells, where the cell's interface can be accompanied

or not by ramps, see Fig. A.15-a and Fig. A.15-b. In this context, Fig. A.15-a is called an homogeneous case, while Fig. A.15-b is called the non homogeneous one.

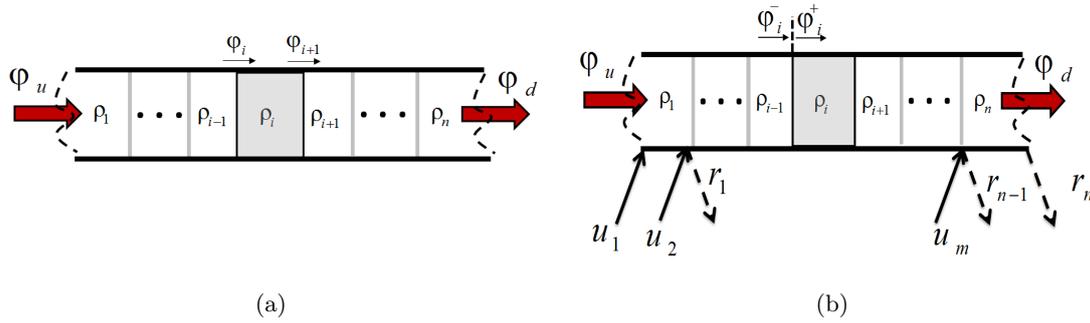


Figure 5.5: Representation of a freeway using the CTM model. (a) Homogeneous section divided into n cells of length L_i and densities ρ_i , $i = 1, \dots, n$, (b) Not homogeneous section divided into n cells of length L_i and densities ρ_i , $i = 1, \dots, n$. Each cell is accompanied with at most one on-ramp and one off-ramp.

Before presenting the governing model equations, for both the homogeneous and non homogeneous cases. We will introduce first the notation that will be used for the CTM model and second the Demand and Supply functions:

CTM notation

- ρ [veh/km]: density,
- L [km]: cell length,
- T [h]: discrete time,
- φ [veh/h]: mainstream flow,
- u [veh/h]: on-ramp flow,
- r [veh/h]: off-ramp flow,
- β : split ratio,
- v [km/h]: free-flow speed,
- w [km/h]: congestion wave speed,
- ρ_c [veh/km]: critical density,
- ρ_m [veh/km]: jam density,
- φ_m [veh/km]: mainstream flow capacity,

Demand/Supply paradigm

The structure for the Riemann solution can be interpreted using the Demand D_i and Supply S_i functions (see Fig. A.16):

$$D_i = \begin{cases} \Phi_i(\rho) & \text{if } \rho_i \leq \rho_{c,i} \\ \varphi_{m,i} & \text{if } \rho_i > \rho_{c,i} \end{cases} \quad (5.4.14)$$

$$S_i = \begin{cases} \varphi_{m,i} & \text{if } \rho_i \leq \rho_{c,i} \\ \Phi_i(\rho) & \text{if } \rho_i > \rho_{c,i} \end{cases}$$

The demand D_i explains how much traffic flow wants to enter the cell $i + 1$, while the supply S_i represents how much flow the cell i can receive.

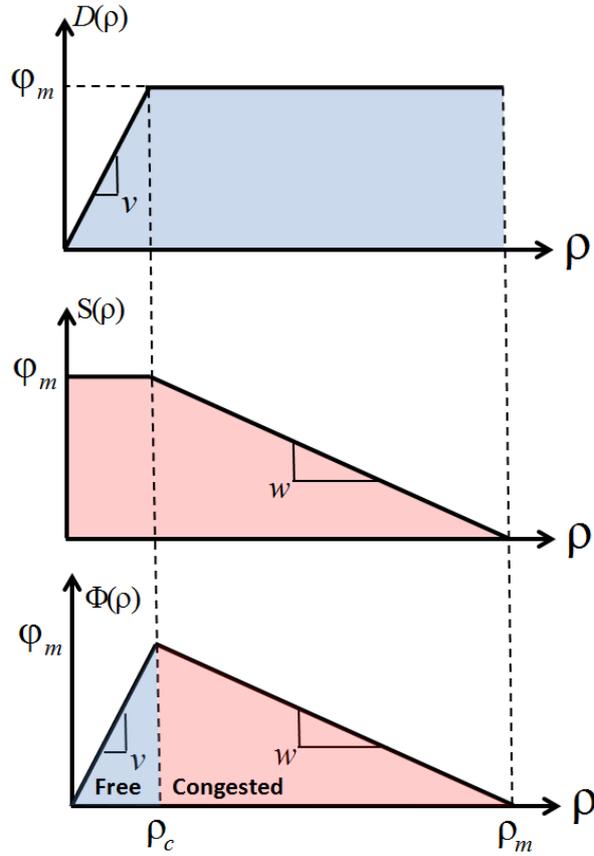


Figure 5.6: Demand and Supply functions. The intersection of both functions characterize the triangular fundamental diagram.

Governing equations

Homogeneous section

Considering a freeway representation as shown in Fig. A.15-a the CTM model is written as:

$$\rho_i(k + 1) = \rho_i(k) + \frac{T}{L_i} (\varphi_i(k) - \varphi_{i+1}(k)). \quad (5.4.15)$$

The interface flow φ_i between the cells $i - 1$ and i is given by:

$$\begin{aligned}\varphi_i &= \min\{D_{i-1}, S_i\} \\ D_{i-1} &= \min\{v_{i-1}\rho_{i-1}, \varphi_m\} \\ S_i &= \min\{\varphi_m, w_i(\rho_m - \rho_i)\}\end{aligned}\tag{5.4.16}$$

According to the status of the cells' interface, Tab. A.2 summarizes the expression of the through flow.

Table 5.1: Behavior of CTM model through different interfaces (Homogeneous section).

Through interface		
Condition	Mode	Interface flow
$v_{i-1}\rho_{i-1}(k) \leq w_i(\rho_m - \rho_i(k))$	Free	$\varphi_i(k) = v_{i-1}\rho_{i-1}(k)$
$v_{i-1}\rho_i(k) > w_i(\rho_m - \rho_i(k))$	Congested	$\varphi_i(k) = w_i(\rho_m - \rho_i(k))$

Non homogeneous section

CTM model can be also derived for a not homogeneous section. For this let us consider the representation in Fig. A.15-b. Each cell is accompanied with at most one on-ramp and one off-ramp. Let us denote φ_i^- the flow leaving the cell $i - 1$ and φ_i^+ the flow entering the cell i . The off-ramp flow r can also be expressed in terms of β , where $\beta \in [0, 1)$. β stands for the split ratio, and specifies the number of vehicles leaving through the off-ramp from the mainstream. The CTM model, under the presence of ramps, can be written as:

$$\rho_i(k + 1) = \rho_i(k) + \frac{T}{L_i}(\varphi_i^+(k) - \varphi_i^-(k))\tag{5.4.17}$$

The expressions of the interface flows, φ_i^- and φ_i^+ , depend on the presence of on/off ramps. These expressions are summarized in Tab. A.3.

These models are based under two working hypothesis. The first states that the direct sum of the mainstream flow $\varphi(k)$ and the input flow $u(k)$ is the merging output flow. The second states that the flow going through the off-ramp, $r(k)$, is a portion of the total flow leaving the section. The first assumption is the more critical one. However it is consistent with Daganzo's and Newel's merge connection model [20, 65], that state that if the rate of merging flows, $\varphi(k)$ and $u(k)$, under given demands, are determined, the merging output flow is $\varphi(k) + u(k)$. In our framework $\varphi(k)$ and $u(k)$ are known. More elaborate models can be applied for the merging areas, for instance in [40, 41], the authors presented a second order traffic flow modelling approach to capture the behaviour of the flow or vehicles at the intersections.

The CTM model presented can be captured by a finite set of linear systems with state-dependent switching function $s(k)$. The switching function depends upon the congestion status of the cells

Table 5.2: Behavior of CTM model through different interfaces (Not homogeneous section).

On-ramp interface		
Condition	Mode	Interface flow
$v_{i-1}\rho_{i-1}(k) + u(k) \leq w_i(\rho_m - \rho_i(k))$	Free	$\varphi_i^-(k) = v_{i-1}\rho_{i-1}(k)$ $\varphi_i^+(k) = v_{i-1}\rho_{i-1}(k) + u(k)$
$v_{i-1}\rho_{i-1}(k) + u(k) > w_i(\rho_m - \rho_i(k))$	Congested	$\varphi_i^-(k) = w_i(\rho_m - \rho_i(k)) - u(k)$ $\varphi_i^+(k) = w_i(\rho_m - \rho_i(k))$
Off-ramp interface (in terms of $r(k)$)		
Condition	Mode	Interface flow
$v_{i-1}\rho_{i-1}(k) - r(k) \leq w_i(\rho_m - \rho_i(k))$	Free	$\varphi_i^-(k) = v_{i-1}\rho_{i-1}(k)$ $\varphi_i^+(k) = v_{i-1}\rho_{i-1}(k) - r(k)$
$v_{i-1}\rho_{i-1}(k) - r(k) > w_i(\rho_m - \rho_i(k))$	Congested	$\varphi_i^-(k) = w_i(\rho_m - \rho_i(k)) + r(k)$ $\varphi_i^+(k) = w_i(\rho_m - \rho_i(k))$

and the direction of the wave fronts. This switching system, known as the Switching Mode Model (SMM), is a piecewise state dependent model. The key difference between the CTM and SMM models is that, with respect to density, the former is not linear, while each mode of the latter is linear.

In traffic, it is well known that an informative macroscopic parameter for describing the level of congestion and road conditions in a highway is the traffic density. Reconstructing densities in portions of the road not equipped with sensors constitutes an important task in traffic estimation, forecasting, and control. Since unfortunately, there do not exist, at present, practical and inexpensive ways to measure this parameter on the field, one should then resort to indirect measurements, such as flow, and mass conservation model-based approaches, such as the CTM and SMM models for instance.

In the following we present a short review of the most relevant works for traffic density estimation based on the SMM model. Furthermore, we will mention some works related to another macroscopic traffic model, METANET.

5.4.3 Traffic density estimation review

As mentioned, mass conservation model-based observers using flow measurements is one popular technique to build traffic densities. They can also be understood as “virtual sensors” deployed inside of the cells not equipped with “true sensors”. They are used to better track, in real-time, density variations with a fine degree of granularity in the space, as the cells can be selected as small as desired.

The most used approaches to tackle the density estimation problem can be roughly classified in

two main groups: linear approaches and non linear approaches. In the former we find stochastic strategies, based on the Kalman filter, and deterministic ones, based on Luenberger observer [56]. While in the latter, we find strategies based on extended Kalman filter, and others based on sliding modes [87].

The estimation approach is chosen according to the underlying macroscopic traffic model that is used. For instance, the extended Kalman filter approach is used when the so-called METANET model is considered [68, 90, 91]. METANET is a continuous non linear second order traffic model developed by Papageorgiou and his collaborators in the early 1980s [19, 67]. On the other hand, when using the SMM model, we generally find approaches based on Kalman filter [84, 85], Luenberger observer [63], and sliding modes [57].

Recently, the CTM has been used in state estimation problems using approaches outside the classification given above. For instance, particle filtering in [60], and numerical differentiation in [2]. In particular, the authors in [2] showed that an algebraic approach can be successfully applied in order to estimate traffic density for different freeway's modes. This approach was found interesting as does not use any stochastic nor asymptotic techniques.

The work proposed in this dissertation, for traffic density estimation, is based on a linear deterministic approach using a Luenberger observer and the SMM model. The benefits of the SMM model to tackle the estimation problem was first discovered in [63]. This estimator enabled the use of a set of linear equations to describe the state evolution for the distinct regimes on the freeway. This work proposed an observer based on two hypothesis. First, that the densities and flows at the section boundaries, as well as flows on all the on-ramps and off-ramps, are measured. Second, that there is at most one wave front in the section. In this work the authors also presented a study of observability and controllability of the switched observer. Using this approach we can find [29, 37, 84]. The weak point seen in this work was the lack of robustness in the switching function $s(k)$. Meaning that the system could jump into different modes when certain conditions were fulfilled. As we face with highly noisy input signals, this jump could happen in a rather high rate, degrading the quality of estimation.

In order to overcome this inconvenient, in a more recent contribution [62], the authors assumed an extra hypothesis. The congestions always appear at the last cell n and propagate upstream. The additional assumption reduced the number of possible modes with respect to [63] and forced the system to respect a certain transition order. However more robust, this constrained approach oversimplified the number of modes.

In the following we restate the idea presented in [62] with the correct number of modes and introduce the concept of graph constrained-CTM observer.

5.4.4 Graph constrained-CTM observer

In order to devise the density observer, let us consider a freeway section as in Fig. A.15-a. In fact, this part of the work focuses on the study of ring-road network designed as a sequence of nodes relid by links, having collection points located at the node level with our main concerning being the estimation of the density inside the links, each link being partitioned in several cells. In that, it is worth to notice that each observation problem is decentralized at the link level; the density of the cells belonging to a link is estimated using only the data provided by the sensors located at the link's boundaries.

Recalling that φ_1 and φ_{n+1} denote respectively the upstream flow φ_u and the downstream flow φ_d .

Making use of Def. 5.1 and Def. 5.2, and assuming that the congestions always appear at cell n and propagate upstream. The restricted number of modes can be evaluated. Indeed, under such a hypothesis only the following cell combinations do exist: FF , FC , CC . For each of these 3 combination we have:

- $F_{i-1}F_i$: $\varphi_i = \vec{\varphi}_i = D_{i-1}$ is diffusive
- $C_{i-1}C_i$: $\varphi_i = \overleftarrow{\varphi}_i = S_i$ is advective
- $F_{i-1}C_i$: φ_i is diffusive if $(D_{i-1} \leq S_i)$ or advective if $(D_{i-1} > S_i)$.

The cells' CTM parameters can be considered different between each other, but for simplicity we assume that they are the same within a link. Nevertheless, they are allowed to be different at different links. They will be identified from real data as it will be explained in latter sections. For illustration purposes, let us consider a section of 3-cells as shown in Fig. A.17 in order to derive the structure of the system dynamics. For it, we have the following possible cells status combinations: FFF , FFC , FCC , and CCC . Noticing that each interface of the type FC can have two possible modes, then, including the upstream and downstream interface flows, we can identify a total of 8 different modes, that we can classify as shown in Table A.4, where $\vec{\uparrow}$ and $\overleftarrow{\uparrow}$ indicate when the flow interface is diffusive or advective, respectively. The modes are indicated alternatively by the integer variable $s(k)$ or by the associate letter with the upper arrow indicating the direction of the waves. Using the letter notation will be useful to easily distinguish diffusive modes $\vec{(\cdot)}$ to the advective ones $\overleftarrow{(\cdot)}$.

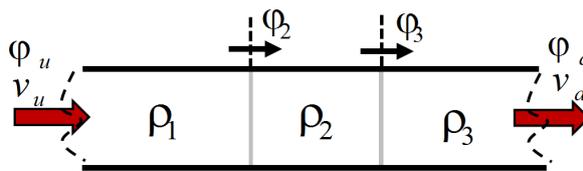


Figure 5.7: Illustration of a homogeneous freeway section. This example is used to derive the system dynamics equations.

Table 5.3: Operating mode table

Mode $s(k)$	Mode $\overleftrightarrow{(\cdot)}$	Cells configuration
1	\overrightarrow{a}	$FFF \overrightarrow{ }$
2	\overleftarrow{a}	$FFF \overleftarrow{ }$
3	\overrightarrow{b}	$FF \overrightarrow{ } C$
4	\overleftarrow{b}	$FF \overleftarrow{ } C$
5	\overrightarrow{c}	$F \overrightarrow{ } CC$
6	\overleftarrow{c}	$F \overleftarrow{ } CC$
7	\overrightarrow{d}	$\overrightarrow{ } CCC$
8	\overleftarrow{d}	$\overleftarrow{ } CCC$

The total number of modes in the constrained case, is then:

$$M = 2(n + 1).$$

These modes have an associated graph \mathcal{G} as the transition between modes are not arbitrary, but they follow the following rules:

- only one congestion wave may exist in the considered section, and that it will propagate upstream following the pattern: FFF, FFC, FCC, CCC ,
- Transitions $FF \rightarrow FC$ are only possible from an advective to a diffusive mode,
- Transitions $FC \rightarrow FF$ are only possible from a diffusive to a diffusive mode.

The last two properties can be derived from continuity of the fundamental diagram during the transient phases. Details of the graph are shown in Fig. A.18. The figure also illustrates the possible graph paths when the congestion moves upwards and downwards.

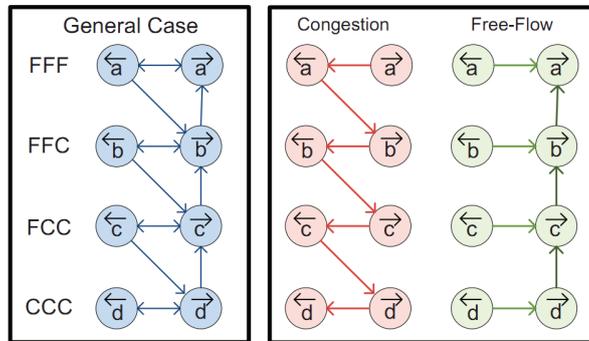


Figure 5.8: Illustration of the Graph \mathcal{G} associated to the constrained-CTM model, for an example of 3-Cells. The left Figure show the general case with all possible modes. The center figure show the graph path for when congestion moves upwards, and the right figure show the graph path when the congestion moves downwards.

Now, let us give the state-space representation of the system where the vector $\rho = (\rho_1, \dots, \rho_n)$ denotes the state of the system, the measured data used by the system being the upstream and downstream flows and speeds, (φ_u, φ_d) and (v_u, v_d) respectively. The index $s(k) \in \{1, 2, \dots, M\}$ is used in order to precise the mode of the entire freeway section.

With this notation the system dynamics are given by:

$$\begin{cases} \rho(k+1) = A_{s(k)}\rho(k) + B_{s(k)}\varphi(k) + \bar{B}_{s(k)}\rho_m \\ s(k) = \mathcal{G}(\rho(k), \varphi(k)) \\ y(k) = h(\rho(k), s(k)) \end{cases} \quad (5.4.18)$$

where $\varphi = (\varphi_u, \varphi_d)$ is the input,

$$h(\rho(k), s(k)) = \begin{cases} v_n \rho_n(k) & \text{if } s(k) = 1 \\ w_1(\rho_{m,1} - \rho_1(k)), & \text{if } s(k) = M \\ 0, & \text{otherwise} \end{cases} \quad (5.4.19)$$

Exact definition of the matrices $A_{s(k)} \in \mathbb{R}^{n \times n}$, $B_{s(k)} \in \mathbb{R}^{n \times 2}$, $\bar{B}_{s(k)} \in \mathbb{R}^{n \times 1}$, $\forall i \in \{1, 2, \dots, M\}$ can be found in [62], the main difference being the total number of modes M , and the associated Graph \mathcal{G} .

For the sake of illustration, we will report here the explicit expression for the system's matrices for the example $n = 3$ (thus $s(k) = 1, \dots, 8$), Fig. A.17.

$$\begin{aligned} s(k) = 1 \\ A_1 = \begin{bmatrix} 1 - \frac{T}{L_1} v_1 & 0 & 0 \\ \frac{T}{L_2} v_1 & 1 - \frac{T}{L_2} v_2 & 0 \\ 0 & \frac{T}{L_3} v_2 & 1 - \frac{T}{L_3} v_3 \end{bmatrix}, B_1 = \begin{bmatrix} \frac{T}{L_1} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \\ \bar{B}_1 = 0_{3 \times 3}. \end{aligned}$$

$$\begin{aligned} s(k) = 2 \\ A_2 = \begin{bmatrix} 1 - \frac{T}{L_1} v_1 & 0 & 0 \\ \frac{T}{L_2} v_1 & 1 - \frac{T}{L_2} v_2 & 0 \\ 0 & \frac{T}{L_3} v_2 & 1 \end{bmatrix}, B_2 = \begin{bmatrix} \frac{T}{L_1} & 0 \\ 0 & 0 \\ 0 & \frac{-T}{L_3} \end{bmatrix}, \\ \bar{B}_2 = 0_{3 \times 3}. \end{aligned}$$

$$\begin{aligned} s(k) = 3 \\ A_3 = A_2, B_3 = B_2, \bar{B}_3 = \bar{B}_2. \end{aligned}$$

$$\begin{aligned} s(k) = 4 \\ A_4 = \begin{bmatrix} 1 - \frac{T}{L_1} v_1 & 0 & 0 \\ \frac{T}{L_2} v_1 & 1 & \frac{T}{L_2} w_3 \\ 0 & 0 & 1 - \frac{T}{L_2} w_3 \end{bmatrix}, B_4 = \begin{bmatrix} \frac{T}{L_1} & 0 \\ 0 & 0 \\ 0 & \frac{-T}{L_3} \end{bmatrix}, \end{aligned}$$

$$\bar{B}_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -\frac{T}{L_2}w_3 \\ 0 & 0 & \frac{T}{L_3}w_3 \end{bmatrix}.$$

$$s(k) = 5$$

$$A_5 = A_4, B_5 = B_4, \bar{B}_5 = \bar{B}_4.$$

$$s(k) = 6$$

$$A_6 = \begin{bmatrix} 1 & \frac{T}{L_1}w_2 & 0 \\ 0 & 1 - \frac{T}{L_2}w_2 & \frac{T}{L_2}w_3 \\ 0 & 0 & 1 - \frac{T}{L_3}w_3 \end{bmatrix}, B_6 = \begin{bmatrix} \frac{T}{L_1} & 0 \\ 0 & 0 \\ 0 & \frac{-T}{L_3} \end{bmatrix},$$

$$\bar{B}_6 = \begin{bmatrix} 0 & -\frac{T}{L_1}w_2 & 0 \\ 0 & \frac{T}{L_2}w_2 & -\frac{T}{L_2}w_3 \\ 0 & 0 & \frac{T}{L_3}w_3 \end{bmatrix}.$$

$$s(k) = 7$$

$$A_7 = A_6, B_7 = B_6, \bar{B}_7 = \bar{B}_6.$$

$$s(k) = 8$$

$$A_8 = \begin{bmatrix} 1 - \frac{T}{L_1}w_1 & \frac{T}{L_1}w_2 & 0 \\ 0 & 1 - \frac{T}{L_2}w_2 & \frac{T}{L_2}w_3 \\ 0 & 0 & 1 - \frac{T}{L_3}w_3 \end{bmatrix}, B_8 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & \frac{-T}{L_3} \end{bmatrix},$$

$$\bar{B}_8 = \begin{bmatrix} \frac{T}{L_1}w_1 & -\frac{T}{L_1}w_2 & 0 \\ 0 & \frac{T}{L_2}w_2 & -\frac{T}{L_2}w_3 \\ 0 & 0 & \frac{T}{L_3}w_3 \end{bmatrix}.$$

$0_{3 \times 3}$ denotes 3×3 matrix of zeros.

5.4.4.1 Observability

The observability of the CTM model for different SMM modes has been studied elsewhere, see for example works in [63] and the references therein. The results are summarized in Table A.5.

In mode $s(k) = 1$, i.e. the $F \dots \vec{F}$ case, φ_u is not related to the density of the cell and it is seen as an external input to the system, whereas $\varphi_d = v\rho$ is state dependent, and hence can be associated to the output of the system, i.e.

$$y = \varphi_d = v_n \rho_n = C_1^T \rho, \quad C_1^T = [0, 0, \dots, 0, v_n]$$

the system is then *backward observable* (i.e. using downstream measurements).

Table 5.4: Observability for different SMM modes

Up. Cells	Down. Cells	Observable with
Free-flow	Free-flow	Downstream measurement
Congested	Congested	Upstream measurement
Congested	Free-flow	Up. and Down. measurement
Free-flow	Congested	Unobservable

In mode $s(k) = M$, i.e. the $\overleftarrow{C} \dots C$ case, φ_d is measured and $\varphi_u = w(\rho_m - \rho)$. Thus we have a reversed situation in which φ_d is seen as an input and φ_u is state dependent, and can be associated to the output of the system, i.e.

$$y = w_1(\rho_m - \rho_1) = C_M^T \rho + w_1 \rho_m, \quad C_M^T = [-w_1, 0, \dots, 0, 0]$$

In this case the system is *forward observable* (i.e. using upstream measurements).

Except for combinations of the type CF (ruled out here), all other modes are non observable because neither φ_u , nor φ_d depend on the density of the cell. The system is detectable, and as such, it allows for internal state forecasting (without state feedback).

5.4.4.2 Observer structure

A hybrid observer that uses only the section boundary conditions can be built from the basis of Eq. A.5.10 by reproducing these dynamics and activating a correcting term when the modes are observable.

$$\begin{cases} \hat{\rho}(k+1) &= A_{\hat{s}(k)} \hat{\rho}(k) + B_{\hat{s}(k)} \varphi(k) + \bar{B}_{\hat{s}(k)} \rho_m + \\ &+ K_{\hat{s}(k)} (y(k) - h(\hat{\rho}(k), \hat{s}(k))) \\ \hat{s}(k) &= \mathcal{G}(\hat{\rho}(k), \varphi(k)). \end{cases} \quad (5.4.20)$$

where $K_{\hat{s}(k)}$ is the observer gain. This gain depends on the estimated modes and is used only for the observable modes $\hat{s}(k) = 1$ and $\hat{s}(k) = M$.

Note that all the modes are stable and that the entire open-loop switched system is asymptotically stable for the arbitrary switching signal $s(k)$. This is rather positive as it is possible to get instability by switching among asymptotically stable modes. Hence, by introducing the observer correcting term we only aim to make the estimation error dynamics converge asymptotically faster to zero. Nevertheless as instability problems can be met, special attention must be given to the approach adopted.

In this work, we adopted a pole placement approach for the observer gain design. The observer gain design is thus as follows:

- spectral radius of $(A_1 - K_1 C_1)$ strictly lower than A_1 for $\hat{s} = 1$,

- spectral radius of $(A_M - K_M C_M)$ strictly lower than A_M for $\hat{s} = M$,
- $K_{\hat{s}} = 0$ for $\hat{s} = 2, 3, \dots, M - 1$.

In other words to improve the performance with respect to the open-loop estimator, the gains K_1 and K_M are selected such that the spectral radius of $(A_1 - K_1 C_1)$ (resp. $(A_M - K_M C_M)$) be upper-bounded by the lowest eigenvalue of \mathbf{A}_1 (resp. \mathbf{A}_M).

Based mainly on Lyapunov and LMIs theory, other approaches for the design of the gain for switched observers haven been proposed in literature. Interested readers can refer to [6, 5, 8, 7, 32, 35, 47, 49, 50, 59, 70, 77, 89].

5.4.4.3 CTM model calibration

The macroscopic model CTM makes use of the fundamental diagram Fig. A.14 to relate observed densities to observed flows at a particular point on the road. This fundamental diagram provides a direct mapping from density to flow. Thus, the selection of the parameters defining the curve is an important part of the model calibration procedure.

The first step of the calibration is the representation of the section of interest in the form of successive cells (space discretization). The section can be the entire freeway or a part of it. The cells must be longer than the free flow travel distance, i.e. $v_i T \leq L_i$. Each cell is assumed to be represented by a single fundamental diagram. These fundamental diagrams are subsequently calibrated on the data provided by the vehicle detector stations of the freeway.

From a top view, the freeway is divided into links. Thus, a fundamental diagram for each link can be computed. Given that a link is composed by several cells. We will assume that all cells inside one link share the same fundamental diagram as the link.

Since the fundamental diagrams are based on flow vs. density scatterplots, the measurement of these two quantities are crucial. The flows and speeds are measured directly from the magnetometers sensors, whereas the link's densities are determined indirectly by the following formula:

$$Density = \frac{Flow}{Speed} \Rightarrow \rho = \frac{(\varphi_u + \varphi_d)/2}{2(\frac{1}{v_u} + \frac{1}{v_d})^{-1}} \quad (5.4.21)$$

where φ_u , v_u , φ_d , and v_d , represent the link's boundary upstream flow, upstream speed, downstream flow, and downstream speed respectively. The calibration procedure is given in the following algorithm (Alg. 4) and Fig. 5.9.

The scatter plot showed in Fig. 5.9 was built with a 24 H data collected from the link of Grenoble south ring between the stations T:Libération (exit) and U:Libération (entrance), link 20, the 28th of February of 2014.

For this example the CTM parameters were: $v = 72$ km/h, $w = 13.31$ km/h, $\varphi_m = 2800$ veh/h, $\rho_m = 250$ veh/km, $\rho_c = 38.2$ veh/km.

Algorithm 4 CTM model calibration- Parameters per link

Input: Scatter plot of flow vs density.

Output: v, w, φ_m, ρ_c .

Begin

- 1) Determine the point of having maximum ordinate and split the data cloud using the value of its abscissa. Let us denote this point as $(\bar{\rho}, \bar{\varphi})$,
 - 2) The free flow speed, v , is estimated by using a first order model and least square approximation method to the data where $\rho \leq \bar{\rho}$,
 - 3) The jam density is fixed to 250 *veh/km* per a 2-lane link (average vehicle's length of 8 m),
 - 4) The congestion speed, w , is computed via a constrained least square (having a fixed point $(0, 250)$) to the data where $\rho > \bar{\rho}$,
 - 5) The maximum capacity and critical density are computed using the formula: $\rho_c = \frac{w}{v+w}\rho_m$, $\varphi_m = \rho_c \times v$.
-

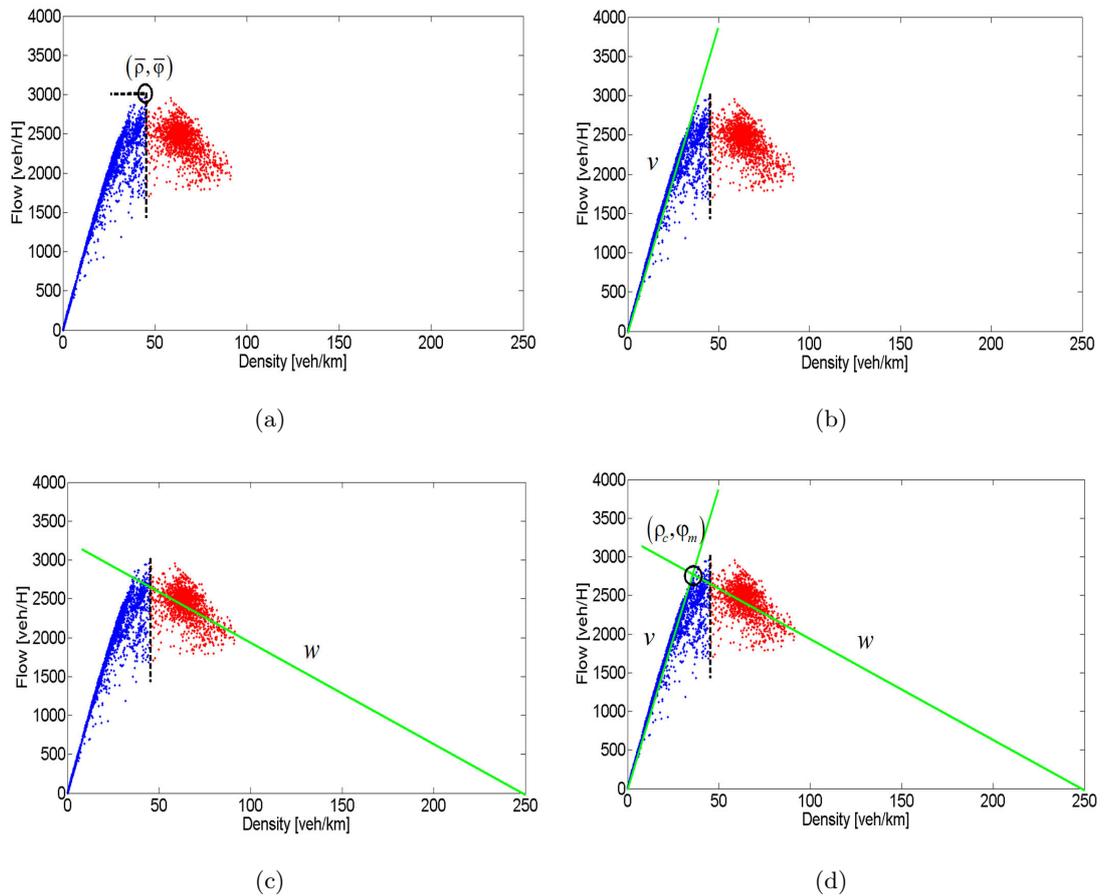


Figure 5.9: Graphical representation of the CTM model calibration algorithm.a) compute the dividing point $(\bar{\rho}, \bar{\varphi})$, b) compute the free flow speed v , c) compute the congestion speed w , and d) with the intersection of the two lines compute the critical density and capacity (ρ_c, φ_m) .

5.4.4.4 Transition function using speed measurements

Up to this point we have derived a systematic switched observer for traffic density estimation. As previously stated, the estimated transition signal $\hat{s}(k)$ depends on the estimation of both the state of the cells and the state of the cells' interfaces, the latter given by the direction of the wave front. This direction is computed using the Rankine-Hugoniot relation, that under the triangular hypothesis of the fundamental diagram reduces to Eq. A.5.8. Hence, $\hat{s}(k)$ can be fully determined by using only estimated flows.

It is known that making the mode selector $\hat{s}(k)$ dependent exclusively on estimations may not be convenient, for estimation errors affect directly the transitions. Therefore, our objective is to upgrade the density observer by making use of the information of the speed boundaries, in other words "robustify" the mode selection.

For illustration let us recall the 3-cell example in Fig. A.17. For this, the transition conditions using only estimated flows can be summarized in Tab. 5.5.

Table 5.5: Transition conditions for a 3-cell example section

Mode $s(k)$	Condition
1	$(\hat{\rho}_1 \leq \rho_{c,1}) \wedge (\hat{\rho}_2 \leq \rho_{c,2}) \wedge (\hat{\rho}_3 \leq \rho_{c,3}) \wedge (\varphi_d = \hat{D}_3)$
2	$(\hat{\rho}_1 \leq \rho_{c,1}) \wedge (\hat{\rho}_2 \leq \rho_{c,2}) \wedge (\hat{\rho}_3 \leq \rho_{c,3}) \wedge (\varphi_d \neq \hat{D}_3)$
3	$(\hat{\rho}_1 \leq \rho_{c,1}) \wedge (\hat{\rho}_2 \leq \rho_{c,2}) \wedge (\hat{\rho}_3 > \rho_{c,3}) \wedge (\hat{D}_2 \leq \hat{S}_3)$
4	$(\hat{\rho}_1 \leq \rho_{c,1}) \wedge (\hat{\rho}_2 \leq \rho_{c,2}) \wedge (\hat{\rho}_3 > \rho_{c,3}) \wedge (\hat{D}_2 > \hat{S}_3)$
5	$(\hat{\rho}_1 \leq \rho_{c,1}) \wedge (\hat{\rho}_2 > \rho_{c,2}) \wedge (\hat{\rho}_3 > \rho_{c,3}) \wedge (\hat{D}_1 \leq \hat{S}_2)$
6	$(\hat{\rho}_1 \leq \rho_{c,1}) \wedge (\hat{\rho}_2 > \rho_{c,2}) \wedge (\hat{\rho}_3 > \rho_{c,3}) \wedge (\hat{D}_1 > \hat{S}_2)$
7	$(\hat{\rho}_1 > \rho_{c,1}) \wedge (\hat{\rho}_2 > \rho_{c,2}) \wedge (\hat{\rho}_3 > \rho_{c,3}) \wedge (\varphi_u \neq \hat{S}_1)$
8	$(\hat{\rho}_1 > \rho_{c,1}) \wedge (\hat{\rho}_2 > \rho_{c,2}) \wedge (\hat{\rho}_3 > \rho_{c,3}) \wedge (\varphi_u = \hat{S}_1)$

The condition transition between the modes 1 and 2, and 7 and 8, are given by the local Riemann solution, where $\varphi_u = \min\{D, \hat{S}_1\}$ and $\varphi_d = \min\{\hat{D}_3, S\}$, D stands for the boundary demand and S for the boundary supply.

In order to transform these flows conditions in speed conditions, some observations derived from the CTM model are worth recalling. For this let us consider a cell between the space interval $[x_0, x_1]$, this cell being in free flow state for which the model calibration has been done and its speed is v , then:

1. The cell's flow velocity does not depend on the density and is always equal to v . As introduced in Chapter. 1, v represents the cell mean-space speed. In the calibration procedure the parameter v is fairly well estimated. It physically states that all vehicles that travel through the cell in free flow state will have a mean speed of v .

2. The measured boundary speeds, v_{x_0} and v_{x_1} , are the time-mean speeds. To this point we know that v is the harmonic mean of v_{x_0} and v_{x_1} . This relation is graphically represented in Fig. 5.10. If we neglect the dispersion present in the free flow part of the fundamental diagram (see Fig. 5.9), we will be assuming that regardless the values of v_{x_0} and v_{x_1} , its harmonic mean in free flow state will always be v . The time-mean speed can also be thought as the monitoring in time at a specific location of the flow velocity distribution.

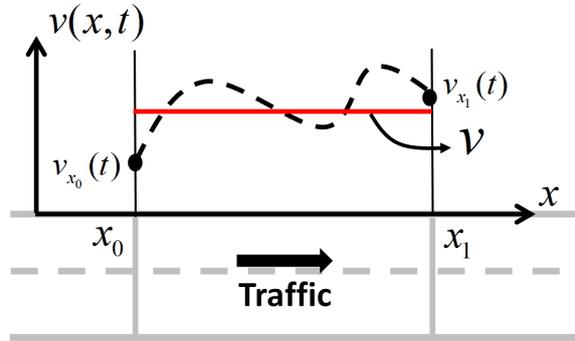


Figure 5.10: Relationship between the free flow speed v (mean-space speed), and the time-mean speeds measured at the cell's boundaries. The flow velocity $v(x, t)$ describes the velocity evolution of the cell. This evolution is modelled as constant v in the CTM model.

At this point the relation between the speeds has been presented. Now assuming that there are no congestion fronts in the network. The transition conditions in terms of speed for our original problem between the modes 1 and 2, and 7 and 8, should be divided in ideal case and non ideal case. The former assumes that the CTM model in free flow matches exactly the real velocity distribution, i.e. all the vehicles in the cell travel at the maximum speed (particularly $v_d = v_3$), the latter that there is a small dispersion in the speed drivers go in free flow. This dispersion can be characterized with experimental data.

For the **ideal case** the transitions then writes, Tab. 5.6:

Table 5.6: Boundary status written in terms of measured flow or speed for the ideal case

Downstream boundary status	Flow condition	Speed condition
Free flow	$\varphi_d = \hat{D}_3$	$v_d = v_3$
Congested	$\varphi_d \neq \hat{D}_3$	$v_d < v_3$
Upstream boundary status	Flow condition	Speed condition
Free flow	$\varphi_u \neq \hat{S}_1$	$v_u = v_1$
Congested	$\varphi_u = \hat{S}_1$	$v_u < v_1$

Remark: *The speed condition for the upstream boundary is true under the assumption that the left and right characteristics at the discontinuity travel at the same speed.*

For the **non ideal case** the speed transitions would write, Tab. 5.7:

Table 5.7: Boundary status written in terms of measured flow or speed for the not ideal case

Downstream boundary status	Flow condition	Speed condition
Free flow	$\varphi_d = \hat{D}_3$	$v_d = v_{\text{lim},3}$
Congested	$\varphi_d \neq \hat{D}_3$	$v_d < v_{\text{lim},3}$
Upstream boundary status	Flow condition	Speed condition
Free flow	$\varphi_u \neq \hat{S}_1$	$v_u = v_{\text{lim},1}$
Congested	$\varphi_u = \hat{S}_1$	$v_u < v_{\text{lim},1}$

Experimental results showed that in free flow state, drivers tend to travel not a lower speed than 90% of the maximum speed. This result was assessed by plotting the available time-speed measurements and comparing them to the maximum speed for the time intervals where the network was in free flow, for different days and links. Thus $v_{\text{lim},i} = 0.9v_i$. In particular $v_{\text{lim},3}$ can be also seen as the speed limit experienced by drivers when the congestion front reaches the downstream boundary and starts to propagates upwards.

The optimal choice of v_{lim} is still in progress. For instance [25], in accordance with the PeMS definition of congestion, assumed that a cell interface is deemed congested when the speed across the detector falls below 64.37 km/h for at least 5 minutes.

Finally the transition conditions using boundary speeds used in this thesis write, Tab. A.6:

Table 5.8: Transition conditions for a 3-cell example section using the boundary speeds

Mode $s(k)$	Condition
1	$(\hat{\rho}_1 \leq \rho_{c,1}) \wedge (\hat{\rho}_2 \leq \rho_{c,2}) \wedge (\hat{\rho}_3 \leq \rho_{c,3}) \wedge (v_d \geq 0.9v_3)$
2	$(\hat{\rho}_1 \leq \rho_{c,1}) \wedge (\hat{\rho}_2 \leq \rho_{c,2}) \wedge (\hat{\rho}_3 \leq \rho_{c,3}) \wedge (v_d < 0.9v_3)$
3	$(\hat{\rho}_1 \leq \rho_{c,1}) \wedge (\hat{\rho}_2 \leq \rho_{c,2}) \wedge (\hat{\rho}_3 > \rho_{c,3}) \wedge (\hat{D}_2 \leq \hat{S}_3)$
4	$(\hat{\rho}_1 \leq \rho_{c,1}) \wedge (\hat{\rho}_2 \leq \rho_{c,2}) \wedge (\hat{\rho}_3 > \rho_{c,3}) \wedge (\hat{D}_2 > \hat{S}_3)$
5	$(\hat{\rho}_1 \leq \rho_{c,1}) \wedge (\hat{\rho}_2 > \rho_{c,2}) \wedge (\hat{\rho}_3 > \rho_{c,3}) \wedge (\hat{D}_1 \leq \hat{S}_2)$
6	$(\hat{\rho}_1 \leq \rho_{c,1}) \wedge (\hat{\rho}_2 > \rho_{c,2}) \wedge (\hat{\rho}_3 > \rho_{c,3}) \wedge (\hat{D}_1 > \hat{S}_2)$
7	$(\hat{\rho}_1 > \rho_{c,1}) \wedge (\hat{\rho}_2 > \rho_{c,2}) \wedge (\hat{\rho}_3 > \rho_{c,3}) \wedge (v_u \geq 0.9v_1)$
8	$(\hat{\rho}_1 > \rho_{c,1}) \wedge (\hat{\rho}_2 > \rho_{c,2}) \wedge (\hat{\rho}_3 > \rho_{c,3}) \wedge (v_u < 0.9v_1)$

5.4.5 Experimental results: state observer

The core topic of this section is to assess the capabilities of the proposed graph-constrained CTM observer, using for this real world data collected from the Grenoble south ring.

The experiments will be divided in two sections. In the first we will validate the proposed observer considering a link between the stations H: Gabriel Péri (entrance 2) and J: SMH Centre (exit). This link has been chosen as a halfway mainstream sensor, I: SMH (mainstream), is available, and thus it will allow us to compare the reconstructed density with the measured one. For this the link will be discretized in an increasing number of cells and we will focus specially on quantifying the estimation accuracy in the congested areas.

After estimating a suitable number of cells, we will present a cell division scheme for the Grenoble south ring. Under this, space-time contours showing reconstructed and measured densities for different days will be given.

5.4.5.1 Observer assessment

Fig. 5.11 depicts the experiment location that will be considered. This link has a length of 2.13 km. Given both the level of congestion observed and the completeness of the data received, the day 14th of March of 2014 will be used for this study.

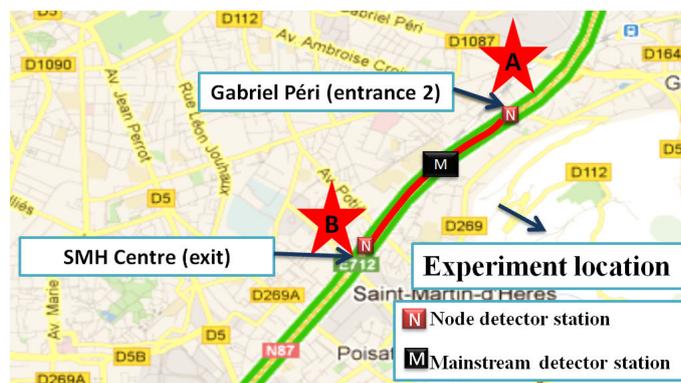


Figure 5.11: Freeway link considered for the observer validation. This link has a length of 2.13 km.

The idea here is to study the observer accuracy while increasing the space discretization. Despite increasing the number of cells, we will always maintain the poles of the system at the same position, so in all cases the system dynamics are invariant. Considering that all cells share the same CTM parameters, it is then sufficient to respect the relation $T_n = \frac{T_1}{n}$, where T_n is the new discrete time for a link divided in n cells, and T_1 the discrete time needed if $n = 1$. T_1 will be chosen to 50 s, in accordance with the numerical stability condition. Consequently, studying the assessment for values of n equal to: 1, 2, 5, and 10, T_n will take the values: 50 s, 25 s, 10 s, 5 s, respectively.

Reconstructed densities

Fig. 5.12 depicts the 24 h comparison result of the measured density (dashed line) with the reconstructed one (solid line) from the cell that encloses the mainstream station. In terms of performance it is seen, for all the tests, a fairly well tracking of the measured density in the free flow and congested areas. The density dynamics are consistent with the traffic behavior in

this freeway, for in general in the east-west direction the stronger congestions are observed in the afternoon time.

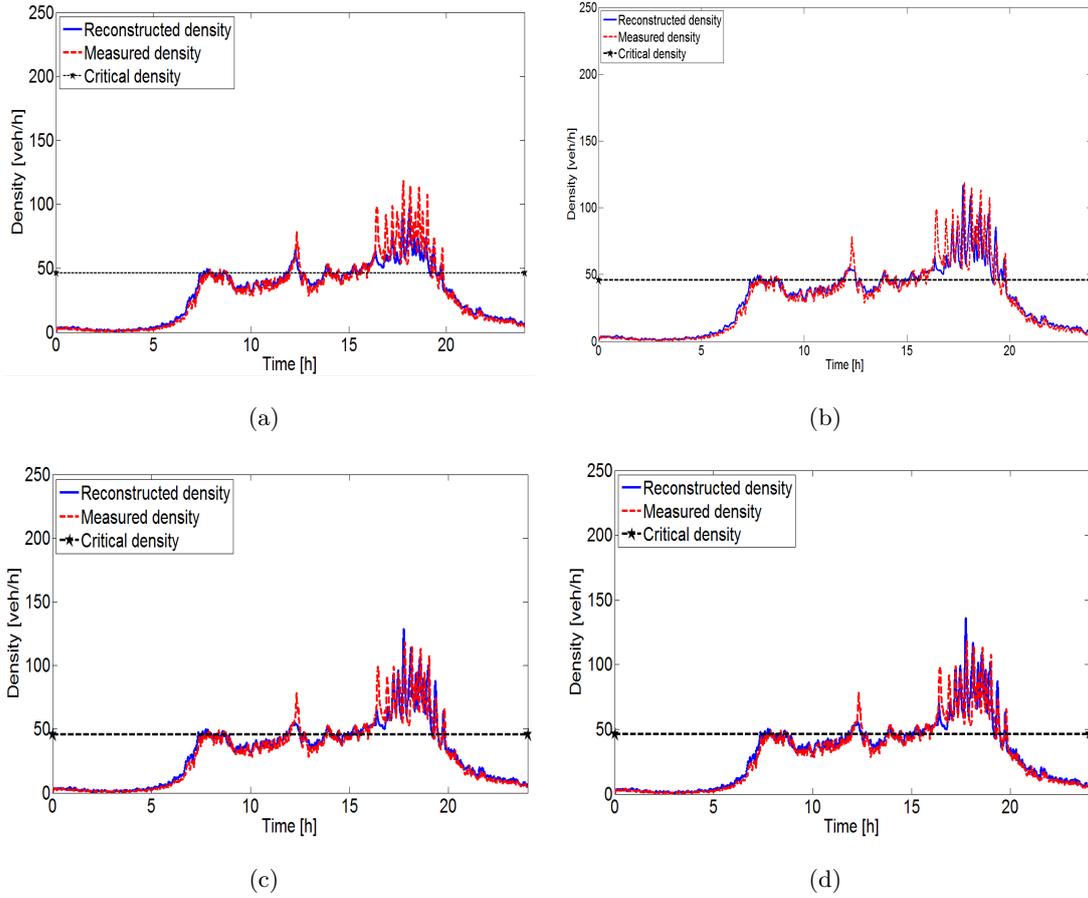


Figure 5.12: Results of the estimation when increasing the number of cells in the link. Each plot compares the measured density with the reconstructed one inside the cell that encloses the mainstream sensors. The number of cells chosen are: a)1, b)2, c)5, and d)10.

In order to assess a numerical quantification, we will focus on the more congested period of the day. It corresponds from 16:00 to 19:30.

Performance quantification in the congested are

In this study we will make use of absolute error metric (AE) as performance index:

$$AE(k) = |\rho(k) - \hat{\rho}(k)|, \quad (5.4.22)$$

where $\rho(k)$ and $\hat{\rho}(k)$ denote respectively the measured and the estimated densities.

Using the estimation error for the time-window and number of cells n chosen, we have evaluated the cumulative distribution function, which is depicted in Fig. 5.13. This analysis shows an increasing improvement trend in the estimation when increasing the number of cells. For instance, we have an error in the density estimation equal or less than 25 veh/km for 76%, 80%, 82%, and 84% of the time, in the case where n is 1,2,5, and 10 respectively. This result was expected, and can be explained by remembering that by increasing the space discretization

the model is able to capture more accurately the congestion front, for the CTM averages the conditions in smaller space portions.

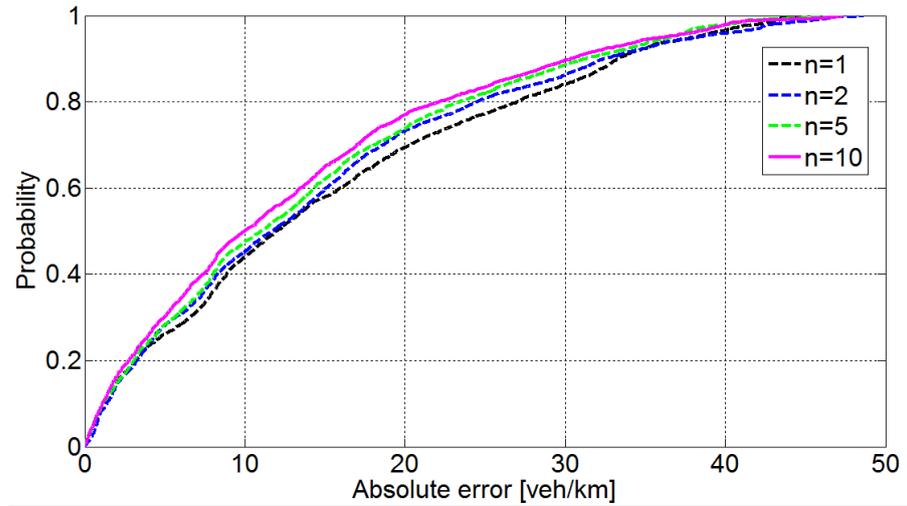


Figure 5.13: Empirical CDF showing the absolute error obtained, in the congested area, when dividing the considered link in an increasing number of cells.

Motivated by this result, we propose a freeway division using 48 cells. This was done by having two goals in mind. First, all the links were to be divided, when possible, at least with two cells. Second, the average length desired of the cells was around 200 m. Certainly, the cell's length cannot be lower than a threshold fixed by the numerical stability condition. Considering a maximum speed of 90 km/h and a discrete time of 5 s, this threshold was 125 m. Note that the inputs of the observer are re-sampled from 15 s to 5 s.

5.4.5.2 Density contours

Fig. A.19 shows the proposed freeway division.



Figure 5.14: Freeway division considering discretization in cells.

Note that here we consider 19 links, the freeway is originally partitioned in 21. This is because two mainstream collection points were bypassed (L: SMH Centre and Q: Échirolles), these correspond to the ones used for modelling purpose in the merging area. For each of the 19 links, we collect the boundary conditions, speed and flow, and run the designed observer.

Fig. A.20 shows, with contour plots, the comparison between the measured densities and the reconstructed ones for the entire freeway for different days. As measured density we will refer to the link’s density computed using Eq. 5.4.21. Thus, even if it is an approximation of the density field, gives us an good flavour of its evolution along the freeway. The days selected for illustration were chosen among the days that presented the strongest traffic jams. These days were February 27th and 28th, and March 7th and 14th of 2014. In particular the 27th and 28th of February, respectively a Thursday and a Friday, were interesting as Grenoble school holidays were in the following week and therefore the freeway was heavily congested due to travellers. One point worth mentioning is that the gray stripes on these days are due to communication failures from the collection points: A: Meylan (on-ramp 1), G: Gabriel Péri (on-ramp 4), and J: SMH Centre (off-ramp 3), see Fig. A.19). Thus any traffic data was received.

In terms of accuracy it is seen a well reconstruction of the freeway conditions, even under heavy traffic jams, using real traffic data. As above mentioned and evidenced here, the congestions are stronger in the afternoon period. Propagating backwards from the end of the freeway. In general, the more interesting part of the freeway, in terms of congestion, is the second half (from entrance 5 in Fig. A.19). Therefore, in order to limit our following forecasting study, we will concentrate in this sector.

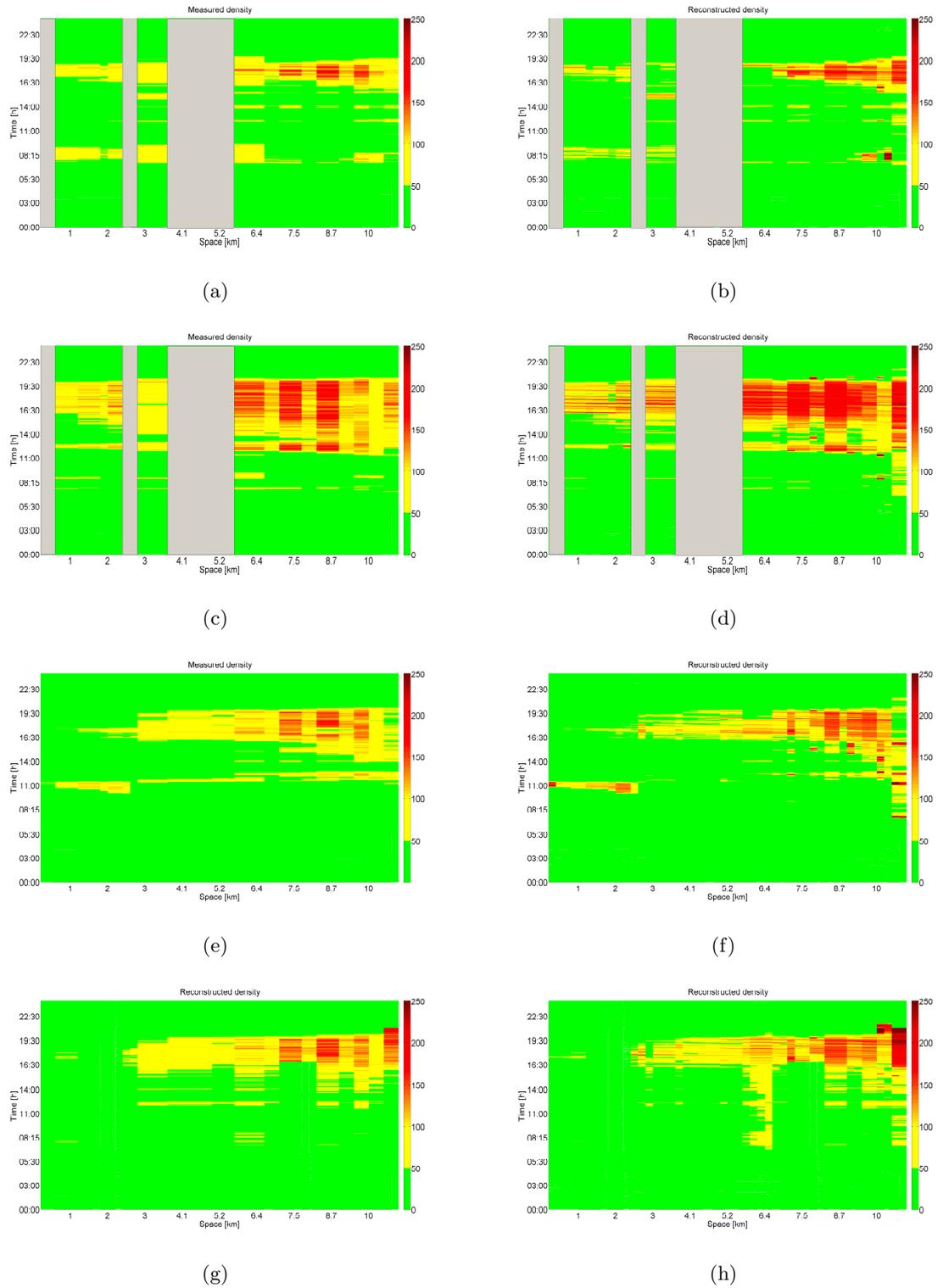


Figure 5.15: Contours depicting the 24 h estimation results for different days of the dataset: a),b) Thursday: 27th of February 2014, c),d) Friday: 28th of February 2014, e),f) Friday: 7th of March 2014, g),h) Friday: 14th of March 2014. Measured density on the left column, reconstructed ones on the right. Colors denote density in *veh/km*

5.5 Boundary conditions, input, and output flows forecasting

In this section, we describe how the boundary conditions, inputs, and outputs of a section of interest are forecasted. This step is rather important given that it will provide the future signals to our CTM-based model.

In order to present the specific information that will be treated here, let us recall the freeway representation of Fig. A.15. Now, we are interested to forecast:

- Boundary conditions: flows, φ_u, φ_d ; speeds, v_u, v_d .
- Input and output flows: u, r .

Note that in contrast with the observation problem, this block will not use mainstream information other than the boundaries. Moreover for the forecasting problem, we will use the 5 min aggregate data discussed in Chapter 3. The forecasting step chosen at 5 min, and forecasting horizon of 45 min.

The methodology adopted will be the noise adaptive Kalman filter approach (AKF) introduced in Chapter 4. Therefore, we will only provide the scheme of the algorithm, an overview of the treatment of the historical data for flow forecasting, and some experimental results revolving only the input flows.

5.5.1 Noise adaptive Kalman filter approach (AKF)

Fig. A.21 depicts the scheme of the multi-step ahead forecasting approach proposed in the preceding chapter. This forecasting strategy can be applied to any time series under the assumption that historical data are available. For specific details readers are referred to Chapter 2.

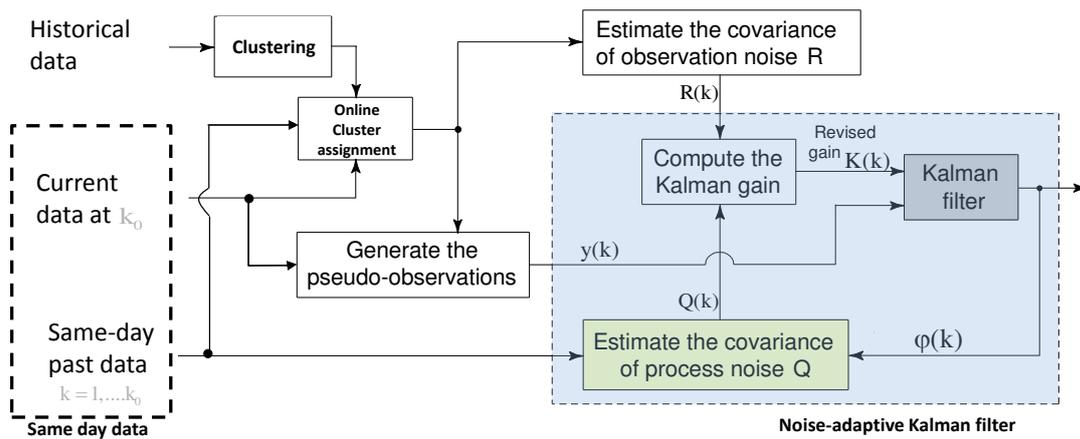


Figure 5.16: AKF scheme.

One of the key parts in this approach was indeed the treatment of the historical data (off-line clustering block). In Chapter 2, given that the travel time signals, directly proportional to

speed, available in the same dataset presented a rather strong variability between each other, we proceeded with a clustering approach by time zones. The K-means algorithm was then applied to each zone, and the optimal number of clusters was estimated using the validity indexes, Root-mean-square standard deviation (RMSSTD) and R-squared (RS). Note that this historical dataset was composed of only working day profiles.

Here, the off-line clustering for boundary speeds will follow the same approach. For flows, however, it was experimentally seen that their profiles are simpler to cluster than speeds. This is in accordance with the findings highlighted in [88]. Fig. A.22 shows an example of working day profiles for different types of flows (input, output, and mainstream) and speed, collected from Grenoble south ring. As the level of the flows depend on where they are collected, they have been normalized by using the maximum value in the dataset in order to focus of their cohesion. Recalling Fig. A.19: the locations chosen were, on-ramp 6, off-ramp 2, for the input (Fig. A.22-a) and output (Fig. A.22-b) flows respectively, and the upstream boundary of link 12 for the mainstream flows (Fig. A.22-c) and speeds (Fig. A.22-d). In the figure it is seen a high correlation between flow profiles, which is not necessarily the case with the ones of speeds. By looking at the afternoon period from 15:00 to 20:00, it can be noticed that in terms of flows, the variability between the profiles increases in some degree, nevertheless it does not seem as important as the one in speed profiles. In the following the proposed off-line clustering for flows will be presented.

Off-line clustering for flow

The 24 h flow profiles will be clustered in two steps. The first will divide profiles into working and not working days. The former will be in general constituted by week days, except when holidays, the latter mainly of week-end days. Finally the second step applies the k-means algorithm to each of the group. The main concern when clustering 24 h traffic data was that important features at certain parts of the days would not be captured, for they would appear in a short period over the 24 h. Nevertheless, with flow profiles it is seen marked characteristics through the day that helps this clustering process. Hence, the second step is rather acceptable. In order to apply k-means, a-priori known number of clusters, N , is needed. For this, the knee criteria using the indexes RMSSTD and RS is applied. Experimental tests have shown that a suitable choice is $N=2$. For illustration, in Fig. 5.18 we present the results of the knee criteria applied to the input flows of Fig. A.22-a. Fig. 5.18-a and Fig. 5.18-b depict respectively the RS and the RMSSTD for the dataset when increasing the number of clusters. It is clear observed that the strongest break occurs for both at $N=2$. Fig. 5.18-c and Fig. 5.18-d on the other hand, show the clustering results. Interestingly, it is seen that Fridays are grouped together, this is rather repetitive when studying other flow data, which suggest that among the working days Fridays have their own particular characteristics. This can be explained considering that in

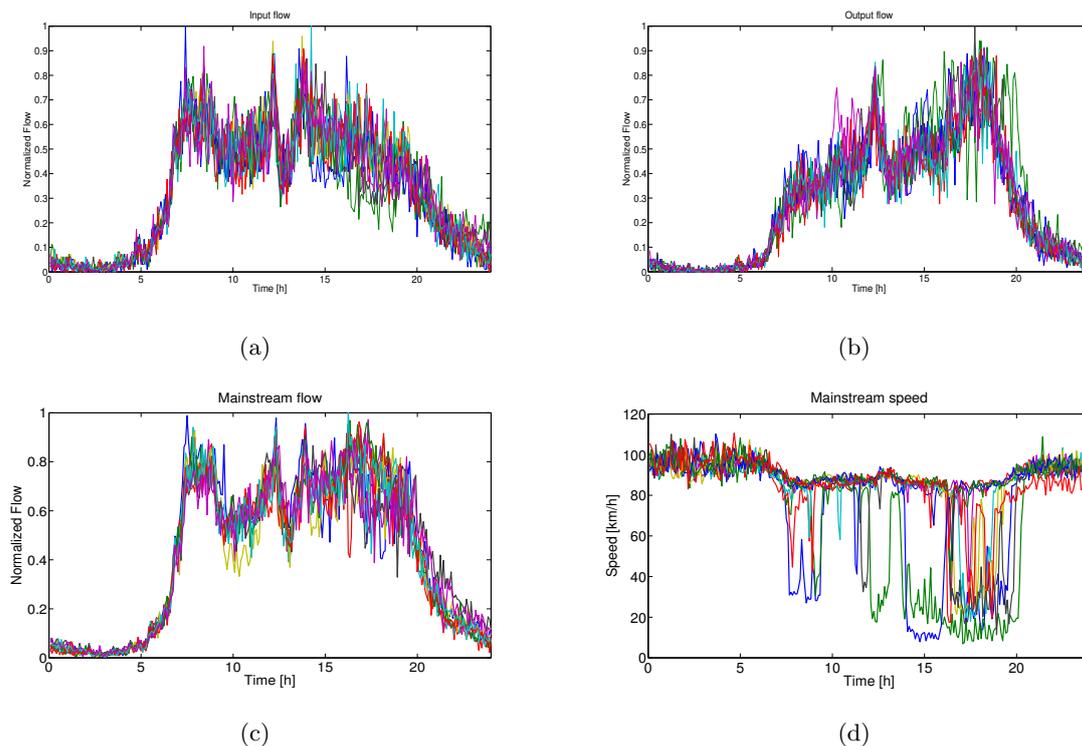


Figure 5.17: 5 min traffic information collected from Grenoble south ring from the 27th of February to the 15th of March of 2014. They represent: a) input flows from on-ramp 6, b) output flows from off-ramp 2, c) mainstream flow from upstream boundary of link 12, d) mainstream speed from upstream boundary of link 12. See Fig. A.19 to visualize the positions. This plot aims to highlight the high correlation within the flow profiles compared to those of speed.

Grenoble, a large portion of the population comes from a different city, Lyon for instance, thus on Fridays they travel to their family using the south ring, making then the freeway particularly different in terms of traffic volume.

Note that once the historical dataset is clustered, the profiles are categorized accordingly and centroids for each cluster is available, the online clustering block in Fig. A.21 gets into game. This block checks which cluster the current data is more likely to belong to.

5.5.2 Experimental results: input flow forecasting

This section aims to assess the AKF methodology, Fig. A.21, for the flow forecasting problem. In particular, only for illustration we restrict to input flow. For this study we will continue using the data collected from the on-ramp 6.

The experimental validation will aim to compare the proposed methodology with the real flow upon a forecasting horizon of 45 min. Here the historical average will also be given. For this two testing scenarios will be considered, the first at a current time at 08:00 and the second at 17:30. Both for Thursday 6th of March 2014. These periods were chosen as being the more interesting ones in terms of congestion and fluctuation levels.

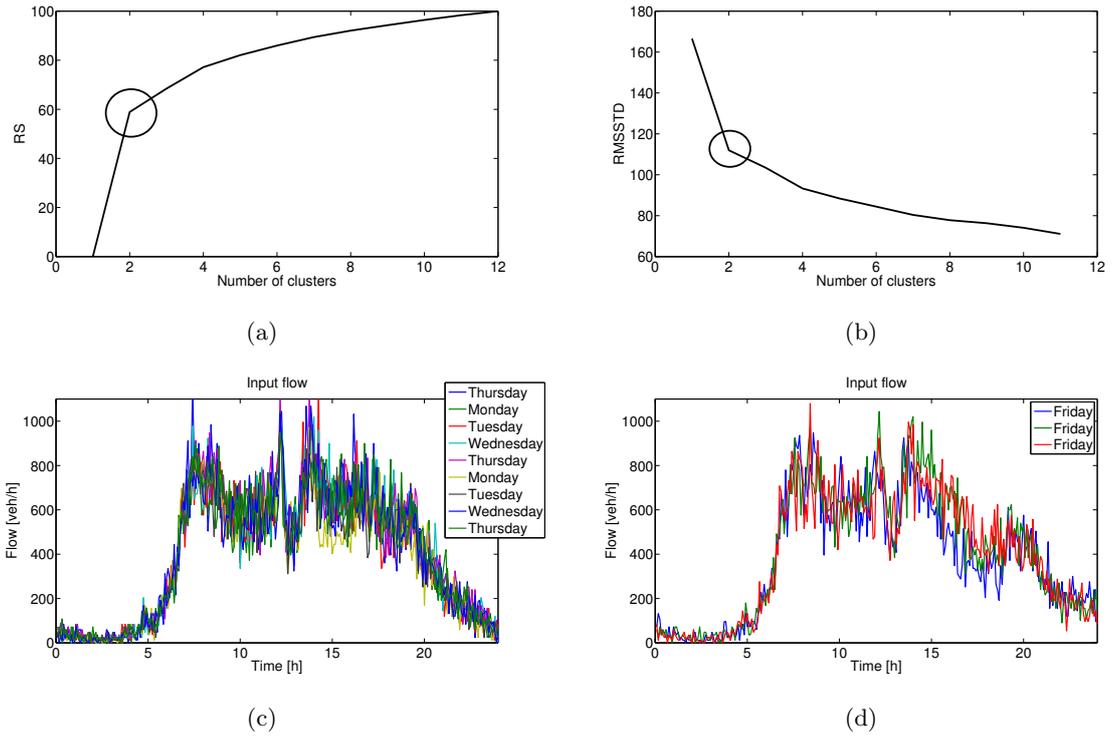


Figure 5.18: Results of the knee criteria and clustering of an input flow dataset. a)RS, b) RMSSTD, c) Cluster 1, d) Cluster 2.

To evaluate the forecasting performance, the mean absolute percentage error (MAPE) is adopted:

$$MAPE(\%) = \sum_{k=t_0+1}^{t_0+\Delta} \frac{|\varphi(k) - \hat{\varphi}(k)|}{\varphi(k)} \times 100\% \quad (5.5.1)$$

where $\varphi(t)$, $\hat{\varphi}(t)$, and Δ are the real and forecasted value of the input flow at time k , and the forecasting horizon respectively.

Fig. 5.19 shows the graphic forecasting results. Fig. 5.19-a depicts the morning period, while Fig. 5.19-b the afternoon. In the former, the real flow profile presents a rather smooth trend with a peak at t_0+15' . Our algorithm, AKF, succeeded to capture both the trend and the peak, outperforming the average historical. In particular this last failed to properly follow the peak. For the following steps, however, the historical average presented a fairly well performance. In the second scenario, Fig. 5.19-b, more oscillations are observed with respect to the first one, which is normal given that the road conditions are more critical. Here, the AKF performed well in terms of accuracy, outperforming again the historical average. Specially at t_0+20' and t_0+40' .

Tab. 5.9 depicts the forecasting numerical results for both the scenarios. Here, it is verified that the AKF approach outperforms the historical average. Specifically in 22.73% and 23.04% for the first and the second scenario respectively.

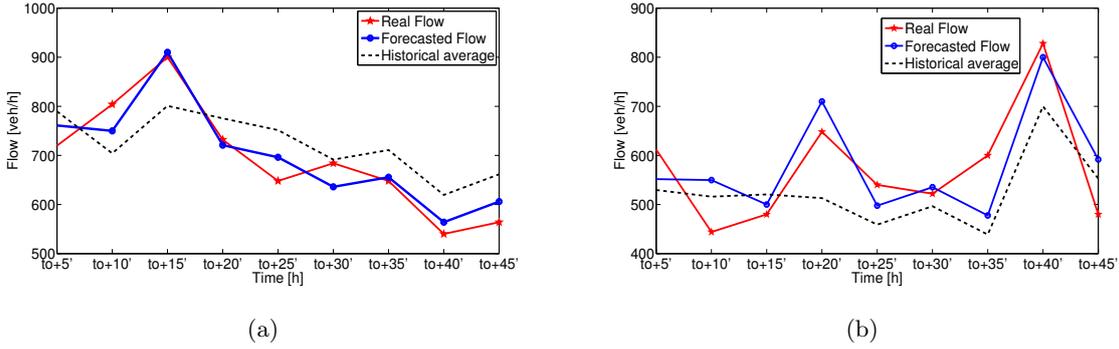


Figure 5.19: Results of the knee criteria and clustering of a input flow dataset. a)RS, b) RMSSTD, c) Cluster 1, d) Cluster 2.

Table 5.9: MAPE(%) for two scenarios considered.

Predictor	Scenario 1 @ 08:00 MAPE (%)	Scenario 2 @ 17:30 MAPE (%)
AKF	7.92	9.23
Historical average	10.49	13.63

These results show that the noise adaptive Kalman filter approach is also well suited for multi-step ahead traffic flow forecasting.

At this stage we were able to develop operational algorithms for density reconstruction and forecasting of boundary conditions, input, and output flows. As previously exposed, this information is the input for the block that will provide the forecasted densities and consequently travel times inside each cell. Therefore, in the following section we will aim to derive this last key part.

5.6 Density and travel time forecasting

5.6.1 Density forecasting

This last step will have the task of forecasting the densities inside each cell enclosed within the section of interest. For this we will make use, once more, of a model-based approach as for the density reconstruction problem in Sec. 5.4.4. Here, the key difference is that we will run the model using a simple “forward injection”, meaning open-loop model, whose inputs are the forecasted: boundary conditions, input and output flows (Sec. 5.5), i.e. without mainstream measurements. The clear advantage in this is that less information needs to be forecasted, thus it reduces the amount of computation and size of the historical database needed for the algorithm. For this we will make use of the CTM model derived in Sec. 5.4.2, whose density dynamics are given by Eq. A.5.9 and interface flows summarized in Tab. A.3.

For illustration purposes, let us consider a section as shown in Fig. A.23.

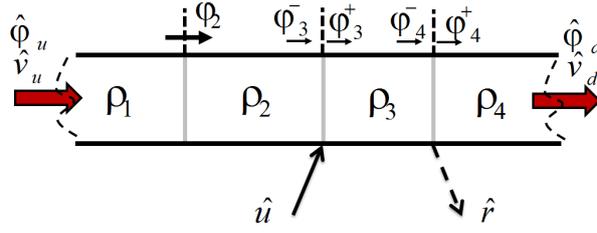


Figure 5.20: Illustration of a not homogeneous freeway section. This example is used to present the equations for the density and interface flow forecasting.

The forecasted densities and their respective interface flows, for the interval $k = t_0 + 1, \dots, t_0 + \Delta$, are computed as:

$$\begin{aligned}
 \hat{\rho}_1(k+1) &= \hat{\rho}_1(k) + \frac{T}{L_1}(\hat{\phi}_1(k) - \hat{\phi}_2(k)) \\
 \hat{\rho}_2(k+1) &= \hat{\rho}_2(k) + \frac{T}{L_2}(\hat{\phi}_2(k) - \hat{\phi}_3^-(k)) \\
 \hat{\rho}_3(k+1) &= \hat{\rho}_3(k) + \frac{T}{L_3}(\hat{\phi}_3^+(k) - \hat{\phi}_4^-(k)) \\
 \hat{\rho}_4(k+1) &= \hat{\rho}_4(k) + \frac{T}{L_4}(\hat{\phi}_4^+(k) - \hat{\phi}_5(k)) \\
 \hat{\phi}_2 &= \min\{v_1\hat{\rho}_1, w_2(\rho_{m,2} - \hat{\rho}_2)\} \\
 \hat{\phi}_3^- &= \min\{v_2\hat{\rho}_2, w_3(\rho_{m,3} - \hat{\rho}_3) - \hat{u}\} \\
 \hat{\phi}_3^+ &= \min\{v_2\hat{\rho}_2 + \hat{u}, w_3(\rho_{m,3} - \hat{\rho}_3)\} \\
 \hat{\phi}_4^- &= \min\{v_3\hat{\rho}_3, w_4(\rho_{m,4} - \hat{\rho}_4) + \hat{r}\} \\
 \hat{\phi}_4^+ &= \min\{v_3\hat{\rho}_3 - \hat{r}, w_4(\rho_{m,4} - \hat{\rho}_4)\} \\
 \hat{\phi}_1 &= \begin{cases} \hat{\phi}_u, & \text{if } \hat{v}_u \geq v_{\text{lim},1} \\ w_1(\rho_{m,1} - \hat{\rho}_1), & \text{otherwise} \end{cases} \\
 \hat{\phi}_5 &= \begin{cases} v_4\hat{\rho}_4, & \text{if } \hat{v}_d \geq v_{\text{lim},4} \\ \hat{\phi}_d, & \text{otherwise} \end{cases},
 \end{aligned} \tag{5.6.1}$$

where $\hat{\rho}(t_0)$ is estimated from the graph constrained-CTM observer.

Note that the input of this block, forecasted speeds and flows, need to be re-sampled in order to respect the numerical stability condition. The outputs of the AKF block are aggregated in 5 min, thus, considering the division in Fig. A.19 we re-sample into 5 s using a linear interpolation.

5.6.2 Speed and travel time forecasting

Now, each cell's travel time is forecasted using Eq. 5.4.21 as:

$$\begin{aligned}
 \hat{t}t_i &= \frac{L_i}{\hat{v}_i}, \\
 \text{where} & \\
 \hat{v}_i &= \frac{\hat{\phi}_i + \hat{\phi}_{i+1}}{2\hat{\rho}_i}.
 \end{aligned} \tag{5.6.2}$$

The forecasted progressive travel time in the section of n cells, $\hat{T}T_n$, is computed as presented in Chapter 2.

5.7 Experimental results: travel time forecasting

The goal of this section is to assess to what degree the introduction of a state observer framework can achieve an accurate travel time forecasting between two points of interest.

The experimental results are organized as follows. First, we describe the experimental location. Second, the test scenario will be given. Finally, we generalize the forecasting assessment using a cross-validation technique for the available dataset.

1. Experimental location:

The location considered for experimental validation is shown in Fig. A.24. This section, with a length of 3.3 km, was chosen given that the strongest jams are predominantly here.



Figure 5.21: Freeway section considered for the validation of travel time forecasting. This section has a length of 3.3 km.

For the available dataset, the progressive travel time between the on-ramp 6 and the off-ramp 7 (A to B) has been computed. Fig. A.25 shows these results. From the figure some comments are worth pointing out. First, in free flow conditions the time spent by a driver in this section is around 2.4 min, while in congestion goes up to 28 min. These values represent a section mean speed of 80km/h and 6.4km/h, respectively. The latter was a very extreme case, for it was the beginning of school holidays as mentioned before (see Fig. A.20-c). In more standard congestion conditions, the section travel time reaches 17 min, mean speed of 10 km/h. It is also seen that the congestions are restricted to the afternoon period, this can be explained as well because of the school holidays. For last, it is also observed that Fridays have particular conditions, that in general differs from the other days of the week. This is consistent with the findings in Sec. 5.5.

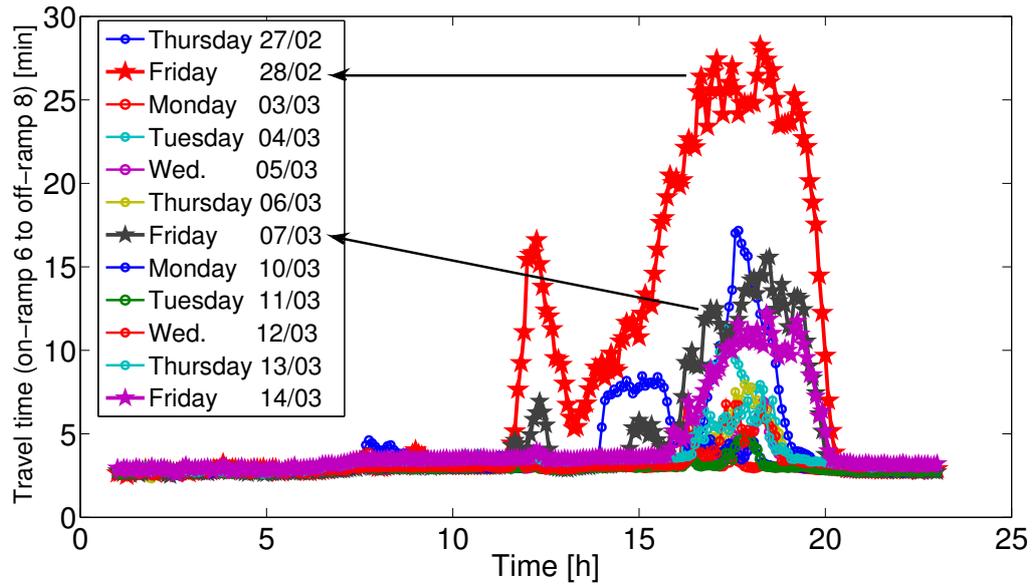


Figure 5.22: Estimated travel time experienced by a driver that travels from Gabriel Peri (entrance 2) to SMH Centre (exit) at different departing times. From 27th of February to the 15th of March of 2014.

2. Scenarios:

Illustrative results are presented. For this, we will choose Friday March 7 as testing day. Here, three scenarios will be presented. The difference between them being the current time assumed: $t_0 = 16:00$, $t_0 = 18:30$, and $t_0 = 19:30$. These specific times were selected because they allow to assess the forecasting approach in three different congestion states; congestion build-up time, congested condition, and congestion phase-out time, respectively.

For each of the scenario two results are given, the multi-step ahead forecasting and the vehicle trajectories upon the forecasted speed contour. In the former we also depict the historical average. Friday March 28 will not be considered in this study, as it will be used for comparison between the two forecasting approaches in the following chapter.

In Fig. A.26-a, it is seen how at the current time, $t_0 = 16:00$, the historical average and our algorithm are quite accurate. While the forecasting horizon increases, it is observed an increasing trend in the travel time. This is consistent with Fig. A.26-b, and indicates that the section starts to become congested over time. The proposed approach succeeds at getting this congestion build-up time, while the historical average does not seem to react accordingly. Fig. A.26-b shows the comparison between a vehicle's forecasted and measured trajectory when it departs at four different times. These times start at t_0 and move forward at steps of 15 min. The first to be noticed is that the forecasted speed contour seems consistent with the expected increasing congestions. Secondly, it is seen a good tracking of the measured vehicle's trajectory. The less accurate part is between the

kilometres 8 and 10 for the second and fourth trajectories. This difference suggests that the real conditions in this part are more congested than it was forecasted.

The scenario in Fig. A.26-c and Fig. A.26-d, computed at $t_0 = 18:30$, puts us into a context where most of the section is already congested. Here, Fig. A.26-c shows how the forecasted travel time follows fairly well the measured travel time with an maximum average percentage error of 30% at $t_0 + 10'$. On the other hand, the average historical does not see the congestion of the section. Fig. A.26-d shows how the forecasted speed contour is consistent with the expected traffic conditions. The measured trajectory is well tracked, and the biggest differences are observed between the kilometres 9 and 10 in the second and fourth trajectory. For the former, the results suggest that this part is less congested than it was forecasted, the opposite for the latter.

At $t_0 = 19:30$, the Fig. A.26-e and Fig. A.26-f show how the congestions start to fade away. Once more, in Fig. A.26-e is seen how the proposed method is able to capture the congestion phase-out time. Fig. A.26-f shows rather clear how the section starts to get in free flow state after 20:00. The measured trajectories are tracked down rather accurate. The difference between the second and third trajectory between the kilometres 6.5 and 8.2 indicate that the congestions stay longer in this part than it was forecasted.

Tab. 5.10 depicts the forecasting numerical results for the four scenarios. Here, at least for this specific example, it is verified how the model-based approach outperforms the historical average.

Table 5.10: MAPE(%) for four scenarios considered.

Predictor	Scenario 1 @ 16:00 MAPE (%)	Scenario 2 @ 17:30 MAPE (%)	Scenario 3 @ 19:30 MAPE (%)
Proposed method	10.88	11.53	5.20
Historical average	23.69	42.12	22.93

3. Cross-validation:

We will generalize the assessment of the method by considering all the days in the dataset and forecasting the travel time under the leave-one-out cross validation. The forecasting will be done for different horizons in the time window between 15:00 and 20:00 at steps of 1 min. At the specified horizons, we will compute the APE (%) and a CDF will be evaluated with all the realizations.

Fig. A.28 shows the cross-validation results for four different forecasting horizon. It is first seen that for all the horizons the historical average is outperformed by our method. This is not a general conclusion, given that we compute the direct average of days that are

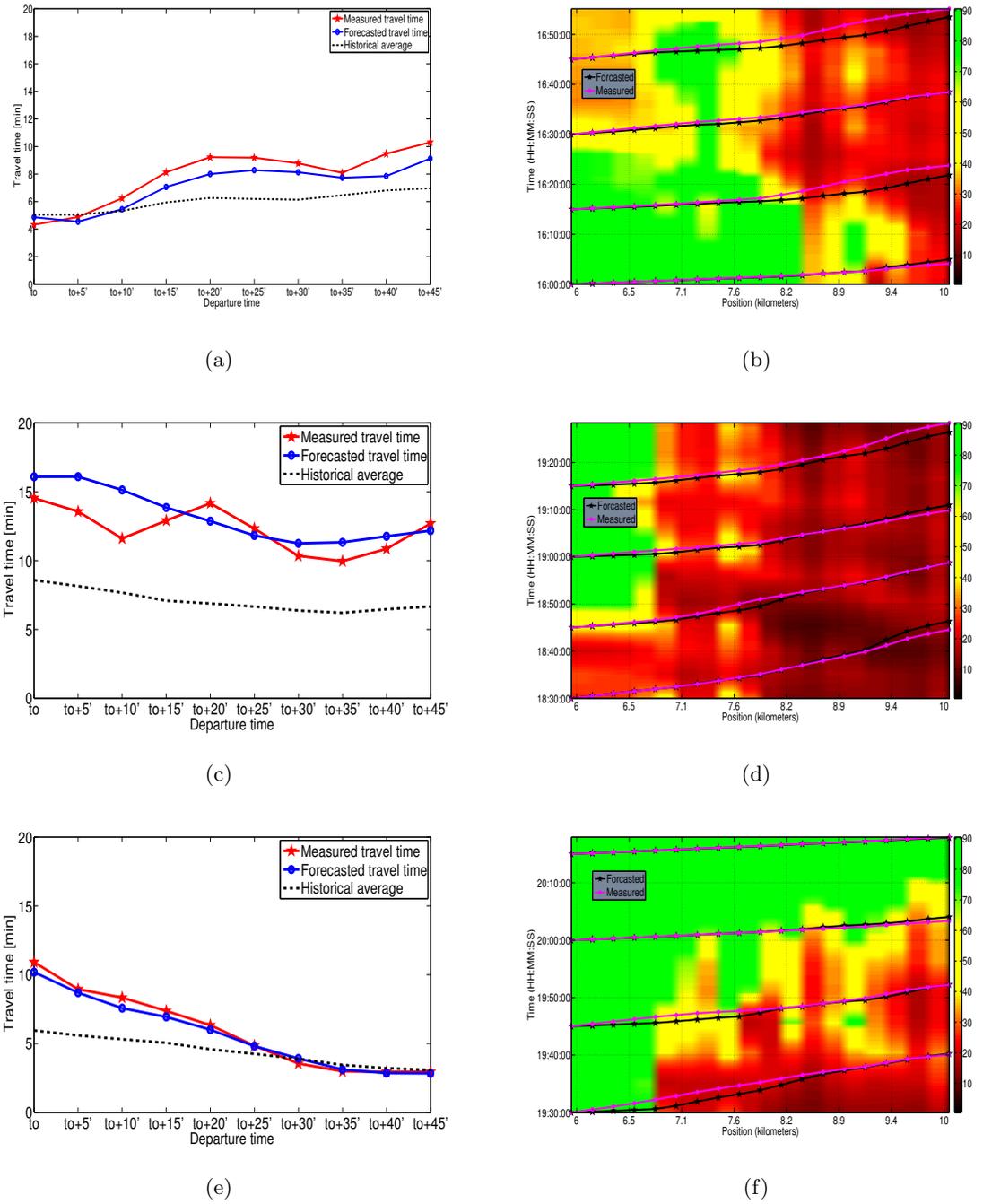


Figure 5.23: Illustrative results of the proposed forecasting approach. a) multi-step ahead forecasts at $t_0 = 16:00$. b) forecasted and measured vehicle trajectory upon speed contour at $t_0 = 16:00$. c) multi-step ahead forecasts at $t_0 = 18:30$. d) forecasted and measured vehicle trajectory upon speed contour at $t_0 = 18:30$. e) multi-step ahead forecasts at $t_0 = 19:30$. f) forecasted and measured vehicle trajectory upon speed contour at $t_0 = 19:30$.

somewhat special, as only a few presented strong congestions. It is also seen a positive result in terms of accuracy for the proposed approach, for all the horizons, 90% of the time we obtain an error below 20%. We also notice that the difference in the performance between the proposed approach and the historical average is smaller as we increase the

forecasting horizon. This is consistent with the findings in Chap. 4. This means that the larger this horizon, the more the forecasting depends only on the historical data.

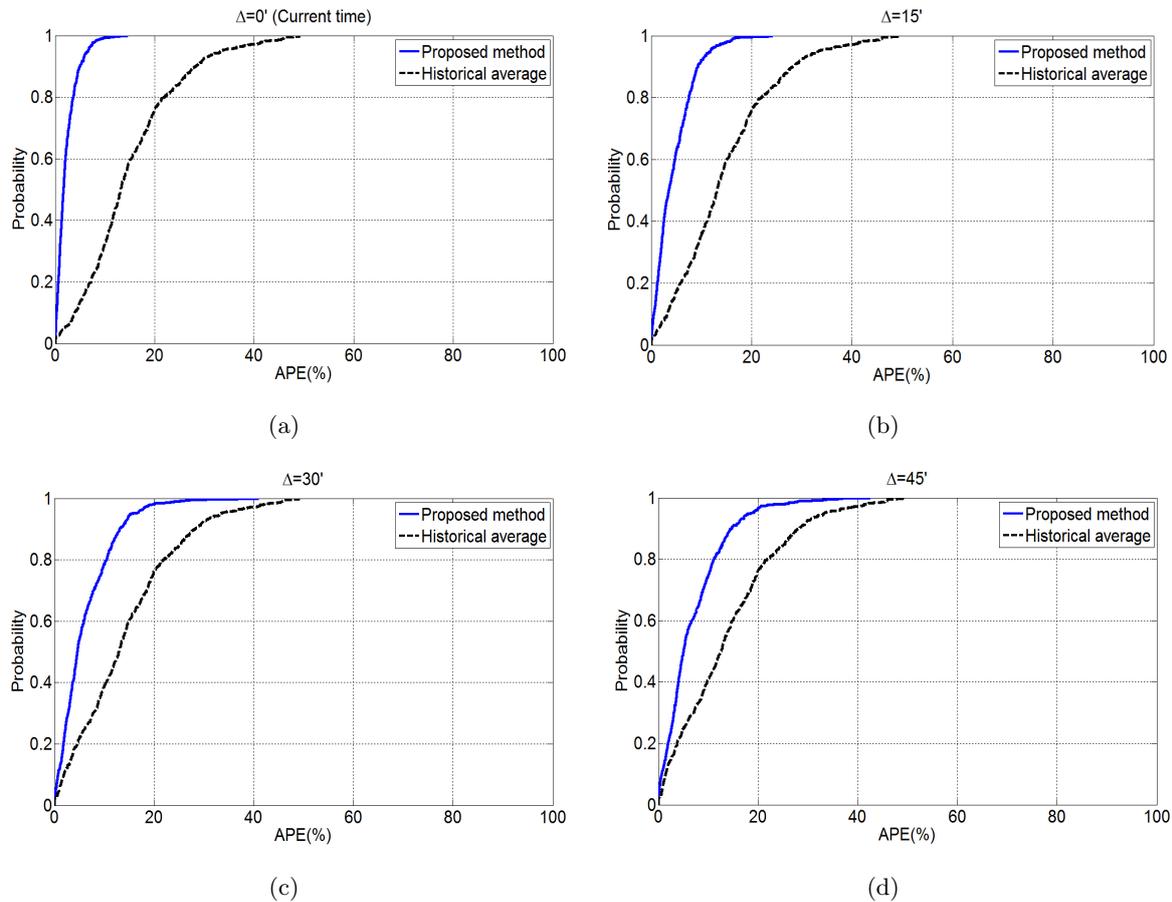


Figure 5.24: Cumulative distribution function evaluated with the APE at different forecasting horizons. a) $\Delta = 0$ (Current time), b) $\Delta = 15'$, c) $\Delta = 30'$, d) $\Delta = 45'$.

Tab. 5.11 displays the maximum error found for 90% of the realizations.

Table 5.11: Maximum error found for 90% of the realizations for different forecasting horizons.

Departure time	$\Delta = 0$	$\Delta = 15'$	$\Delta = 30'$	$\Delta = 45'$
Proposed method (%)	5.0	9.0	13.3	15.0
Historical average (%)	26	28.5	29.0	30.2

These results give a good idea of the capabilities of this second forecasting methodology. This suggests that in fact, the increase of the space discretization can be a powerful tool in order to obtain more accurate and consistent forecasting outcomes, specially at short forecasting horizons.

5.8 Conclusions

In this chapter we developed a model-based approach, based on flow and speed measurements, in order to tackle the multi-step ahead travel time forecasting. The choice of this approach was motivated by two points: the need of improving the tracking of traffic conditions by increasing the space discretization and the need of decreasing the number of signals to forecast. The main contributions in this chapter were:

- **Model-based density observer.** By measuring only boundary conditions of the links enclosed by the section of interest, we derived a discrete switching state observer, based on the “Switching Cell Transmission Model” (SMM) (linearization of the “Cell Transmission Model” (CTM)). This observer related measured flows to estimated densities and allowed the space discretization of the links in cells, reconstructing in time the densities inside of them.
- **Multi-step ahead traffic flow forecasting.** We showed that the noise adaptive Kalman filter (AKF) approach proposed in Chap. 2 was also suitable for traffic flow forecasting. Moreover we provided results in the context of traffic flow data clustering.
- **Multi-step ahead density forecasting.** We showed that by modelling the section with the traffic CTM model and running this model, whose inputs will be: forecasted boundary conditions, and forecasted input and output flows, we can obtain consistent forecasts of the cells’ densities.
- **Speed and travel time forecasting using forecasted densities.** By making use of forecasted densities, we used the relationship between the macroscopic traffic parameters, density and the speed, to accomplish the travel time forecasting.
- **Validation of results with real traffic data.** Each of these blocks are validated with real traffic data. The assessment showed that: while increasing the space discretization (number of cells), the performance of the proposed graph-constrained CTM observer was more accurate, the method proposed presented 90% of probability to forecast with an error smaller than 18% for a maximum horizon of 45 min, and the larger the forecasting horizon, the more the current data loses effect on the forecasts.

Chapter 6

Conclusions and future work

This chapter has the goal to summarize and highlight the main contributions in this thesis. It will be organized as follows. First, in Sec. 6.1 a final comparison between the two forecasting approaches will be provided. In Sec. 6.2 the main contributions of the thesis are briefly outlined. Finally, in Sec. 6.3 we focus on some future perspectives.

6.1 Comparison between the proposed forecasting approaches

This section aims to numerically compare the forecasting methods under the same scenario. For this, we will make use of the same dataset of Chap. 5: 27th of February to the 15th of March of 2014, without week-ends. As testing days, we chose the ones that presented more congestion, days with arrows in Fig. A.27.

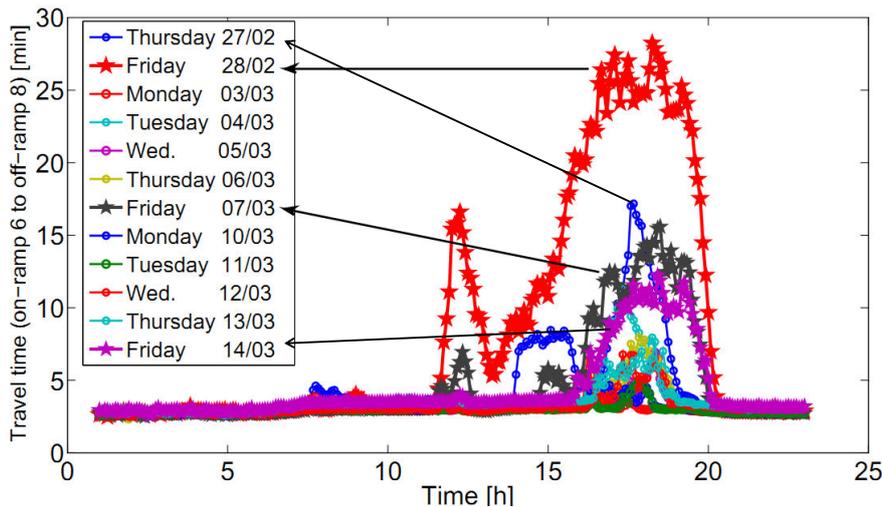


Figure 6.1: Estimated travel time experienced by a driver that travels from Gabriel Peri (entrance 2) to SMH Centre (exit) at different departing times. From 27th of February to the 15th of March of 2014. The days pointed with arrows are the days that will be considered for the numerical comparison between the two proposed forecasting methods.

The study will be performed as follows. The forecasting, for both methods for every day, will be done for different horizons in the time window between 15:00 and 20:00 at steps of 1 min. At the specified horizons, we will compute the APE (%) and a CDF will be evaluated with all the realizations.

Fig. A.28 shows the results. Several observations can be drawn from these results. First, it is observed a good accuracy for both the methods. For a maximum horizon of 45 min the methods have 90% of probability to forecast with an error lower than 21%. In particular it was found that for 90% of the realizations: at $\Delta = 0$ the model-based outperformed the signal-based in 37.1%, at $\Delta = 15$ the model-based outperformed the signal-based in 15.3%, at $\Delta = 30$ the model-based was outperformed by the signal-based in 5.3%, at $\Delta = 45$ the model-based was outperformed by the signal-based in 17.0%. This assessment suggests that for horizons shorter than 30 min, a model-based approach can be more suitable than the signal-based, the opposite for horizons larger than 30 min. This can be explained by recalling that the model-based uses a higher granularity of the space, thus it succeeds to better track the traffic conditions of the section especially at short forecasting horizons.

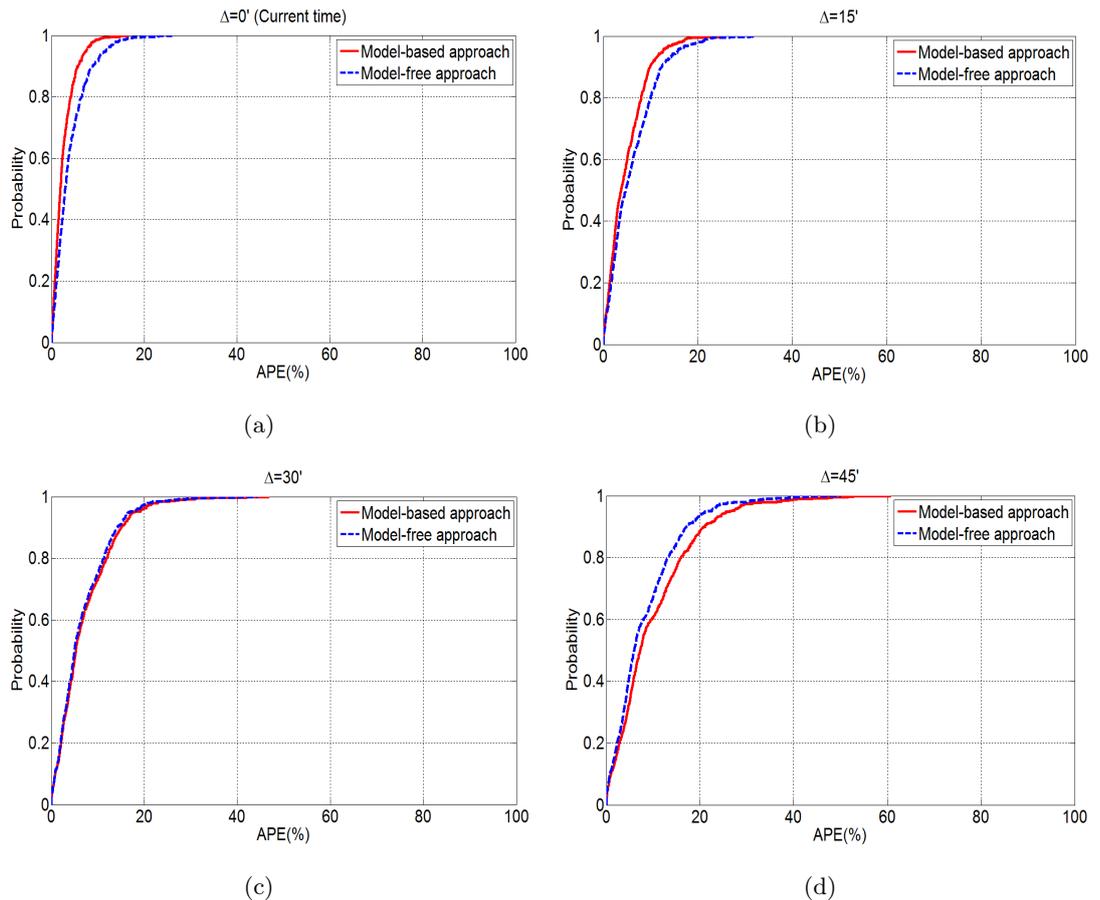


Figure 6.2: Cumulative distribution function evaluated with the APE at different forecasting horizons. a) $\Delta = 0$ (Current time), b) $\Delta = 15'$, c) $\Delta = 30'$, d) $\Delta = 45'$.

The final remarks are listed below:

- **Reliability:** both methods were found consistent and with a good performance. They were able to forecast with an error smaller than 21% at maximum horizons of 45 min. The results also suggest that the model-based is more apt for short-term forecasts, while the signal-based for larger horizons.
- **Transferability:** both approaches are able to be implemented across various locations.
- **Computation efficiency:** in order to estimate the current traffic state, the signal-based approach uses less information than the model-based method. The former uses only speed measurements from the mainstream, while the latter uses also flow data from mainstream and ramps. For forecasting on the other hand, the signal-based approach makes use of a larger historical dataset, while the model-based decreases the information needed by only using the section boundaries and the ramps. This last point is rather crucial, as the algorithms are to be running in real-time.
- **Adaptability:** they are self-tuning to the incoming data stream. Using the AKF approach for speed and flow forecasting, we make sure that the current traffic state is considered.
- **Robustness:** they both handle incoming erroneous and missing data. However, it is clear the advantage of the signal-based method over the model-based. The former can always forecast using speed measurements, for it is a signal-based approach. While the former needs more consistent data in order to respect the vehicle conservation law.

6.2 Main contributions

The aim of this dissertation was the design of a methodology for short-term multi-step ahead travel time forecasting using indirect freeway traffic measurements. The main contributions of this thesis are briefly listed below:

- **Procedure for traffic data pre-processing.** In Chap. 3, we thrive a procedure for the pre-processing of traffic data. This procedure allowed the analysis, cleaning, and imputation of the raw traffic data downloaded from GTL. Here it was shown that for the dataset considered, we can impute with a efficacy of 90% for a percentage missing data of 20%.
- **Development of a novel approach for multi-step ahead traffic forecasting based on Kalman filter.** In Chap. 4, we introduced a novel method for multi-step ahead traffic forecasting making use of an noise adaptive Kalman filter approach (AKF). This approach transformed the forecasting problem to a filtering one, by using as system's observations a suitable combination of the historical data, referred to as pseudo-observations. This step

made possible the computation of the Kalman gain, forcing then the multi-step ahead forecasts to follow a stable historical traffic pattern. The observation noise statistics were computed examining the dispersion presents in the historical data, i.e. the higher the dispersion, the less trustworthy the pseudo-observation. On the other hand, the statistics of the process noise were computed using an online unbiased estimator.

- **Speed data clustering.** In Chap. 4, we highlighted the issue of clustering speed profiles in a dataset. In particular, we showed that the variability present in profiles belonging to a working week, or even belonging to the same day of the week was too important. Consequently, to capture particular attributes in a particular time periods of the 24 h data, that, in addition made the difference between a satisfactory speed forecasting was quite tough. Therefore, we proposed to solve this problem by doing a clustering by time periods, i.e. reducing the problem dimensionality. The periods were chosen in order to present more marked traffic conditions between profiles, making then the clustering easier to achieve. In each cluster the k-means algorithm was used. Moreover, it was also shown that the k-means succeeded to capture outlier profiles in the dataset.
- **Flow data clustering.** In Chap. 5, it was shown that in fact, flow profiles (collected from mainstream, inputs, or outputs) had more cohesion between each other. Therefore, they could be clustered using a less elaborate approach. Here, we show that making use of the k-means upon the considered dataset, the working days could be classified into two main groups, one composed only of Fridays, and the other of the rest of the working days.
- **Density observer.** In Chap. 5, it was developed a deterministic discrete switching observer in order to aboard the traffic density estimation problem, from flow and speed measurements. This observer, referred to as graph constrained-CTM observer, was based on the Switching Mode Model (SMM), linearization of the Cell Transmission Model (CTM). Here we showed that with the assumption of the two hypotheses: only one congestion wave may exist in a freeway link, and that it will propagate upstream, we could reduce the number of affine modes and a constrained model could be devised. Furthermore the mode selector function depended on a diagraph derived from the continuity of the traffic fundamental diagram and the direction of the shock waves originated at space discontinuities. In order to robustify the mode selection, we also proposed to monitor the boundaries' state using speed measurements from the link's boundaries. The proposed observer was validated with real data; the results showed that method succeeded at capturing rather satisfactorily the different traffic conditions of the freeway. For these

experiments different days were considered, which varied in level of congestions. Moreover, this observer allowed us to propose an optimal partition of the entire freeway in 48 cells, with an average length of 220 m.

- **Freeway travel time forecasting based on mainstream speed measurements.** In Chap. 4, speed measurements from mainstream detector stations were used to assess the accuracy of the AKF approach for the purpose of freeway travel time forecasting. For this experiment, it was shown that forecasts obtained by feeding the AKF with clustered speed data, performed as well or as better than forecasts using historical average, in particular in atypical traffic conditions. While this result may vary with different datasets, is a strong indicator of the possibility for AKF to be used for forecasting purpose.
- **Freeway travel time forecasting based on a macroscopic traffic model, using flows and speed measurements.** As a second approach, in Chap. 5, we introduced another novel approach for travel time forecasting constructed upon the basis of the CTM model. This approach was built using several constitutive blocks. The first, used the Graph constrained-CTM observer to reconstruct the cells' densities at the current time, the second used the AKF and the respective clustering techniques to forecast boundary conditions, and input and output flows upon the forecasting horizon. The third, whose inputs were the forecasted signals and the density values at the current time, used the CTM model to compute the future evolution of densities. Finally the last block computed the freeway travel time from forecasted densities. This approach was validated with traffic data. The results showed a well performance in terms of accuracy. Again, the approach performed as well or as better than the historical average. While the size of the available dataset did not allow a deeper assessment study, the numerical results showed that this method has a lot of potential to become an efficient and systematic approach for travel time forecasting.

As a final remark we can say that, the fact that every contribution was, at a certain degree, validated with real traffic data, represented a great extra challenge and contribution per se. For several implementation issues, as well as techniques for data treatment, needed to be developed. Moreover, all the proposed algorithm: data pre-processing, data clustering (speed and flow), noise Adaptive Kalman filter (AKF), and graph constrained-CTM observer, have been implemented in a modular approach. As such, they plug and play in nature. At the time t they have been optimized and tailored in order to be implemented and utilized in real time by GTL.

Several other important findings have been uncovered through this work, which may be the focus of future research efforts.

6.3 Future research efforts

- **Graph constrained-CTM observer using only data from boundary and ramps.** A further effort should be considered in order to develop the density observer without using mainstream measurements, other than the section boundaries. This would make a clear impact on the amount of information the model-based approach would need.
- **Forecasting with confidence intervals.** As an important line of research to be followed after this thesis is the design of confidence levels in the forecasting. This would allow the results to be more flexible, and in some sense more realistic.
- **Robust Graph constrained-CTM observer.** A more robust version of the density observer developed in this thesis have been proposed. This version took into account the dispersion given in the congested are of the fundamental diagram. Nevertheless, a further study of this robust version should be considered, for instance by assuming probabilistic transitions for the system digraph.
- **Heterogeneous data fusion.** Nowadays, low-cost probe data have become more available. Thus, a likely scenario is one in which we use these data in order to improve the forecasting and estimation algorithms developed using data collected from dedicated sensors. In particular, estimation of traffic conditions using all available data is challenging due to the requirement of new models.

Appendix A

Resumé en français

A.1 Introduction

Il est de notoriété public que les problèmes de trafic ont un impact très négatif sur une ville. La façon évidente de résoudre ce problème est de construire de nouvelles infrastructures de transport. Malheureusement, c'est impossible dans la plupart des villes en raison des coûts élevés et des contraintes d'espace. Dans ce cadre, les systèmes de transport intelligents (ITS) ont émergé dans les années 80 afin d'augmenter l'efficacité des systèmes de transport par l'utilisation de nouvelles technologies d'information et de communication, et des modèles mathématiques complexes.

Au cours de la dernière décennie, des avancées majeures ont eu lieu dans le domaine des ITS. D'un point de vue globale, le flux typique d'information pour les ITS (voir Fig. A.1) est la suivante:

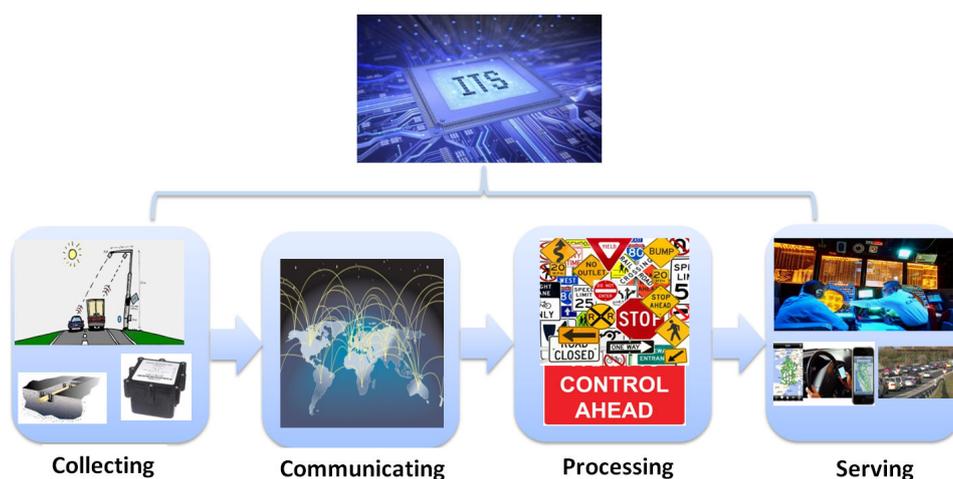


Figure A.1: Flux d'information pour les ITS

1. La collecte des données: elle concerne toutes les technologies de collecte de données existantes.

2. Communication: elle correspond à l'étape où les données de trafic collectées sont envoyées par l'intermédiaire d'un canal de communication qui varie en fonction de la technologie disponible.
3. Traitement: dans cette étape, les algorithmes développés par les chercheurs entrent en jeu. Algorithmes pour le contrôle, l'estimation, et la prédiction.
4. Service: elle consiste à appliquer dans des situations réelles, des solutions technologiques, soit sur les Advanced Traffic Management Systems (ATMS) soit sur les Advanced Traveller's Information Systems (ATIS).

Dans ce contexte, cette thèse aborde la prédiction du temps de parcours dans une autoroute, en faisant usage de mesures de débit et de vitesse recueillies par des capteurs magnétiques. L'objectif est de concevoir des stratégies de prédiction à court terme pour fournir des temps de déplacement prédictive en temps réel pour plusieurs routes candidates.

Cette thèse s'inscrit dans le cadre de deux projets de recherche: Le projet français MOCOpo¹ et le projet européen Hycon2². Les deux projets traitent de l'estimation et de prédiction de trafic, où La Rocade Sud de Grenoble constitue le cas d'étude.

- MOCOpo: aborde la modélisation et traitement des données de trafic et pollution. La tâche impliquant l'Institut de Recherche INRIA³ via l'équipe NeCS⁴, au sein de laquelle cette thèse a été faite, est la prédiction du temps de parcours entre une entrée et une sortie de l'autoroute.
- Hycon2: réseau d'excellence européen coordonné par CNRS où La Rocade Sud de Grenoble est un des cas d'étude. L'objectif de ce projet est le développement des algorithmes pour atteindre le contrôle du trafic et les prédictions du temps de parcours. Ce projet vise également à la construction d'une plateforme de test commune, sur laquelle tous les partenaires peuvent tester leurs développements technologiques.

A.1.1 Cas d'étude

La rocade sud de Grenoble, avec une longueur de 10km, comporte 11 entrées et 10 sorties. Dans cette thèse seulement le sens, Est/Ouest, (Meylan jusqu'à l'autoroute A480) a été considéré. La collecte des données est réalisée par l'intermédiaire des capteurs de type Sensys Networks VDS240 installés sur l'autoroute. Les stations sont constituées par trois paires de capteurs, deux pour la voie principale et une paire pour la rampe. Les capteurs constituant la paire sont placés à une distance de 4,5 m environ. Les données de trafic sont collectées toutes les 15

¹Measuring and mOdelling traffic COngestion and POllution-<http://mocopo.ifsttar.fr/>

²Highly-complex and networked control systems-<http://www.hycon2.eu/>

³Institut National de Recherche en Informatique et en Automatique

⁴<http://necs.inrialpes.fr/>

secondes et transférées à un serveur géré par GTL (Grenoble Traffic Lab).

A.1.2 Objectifs généraux

Cette thèse concerne la conception d'une méthodologie de prédiction efficace, en temps réel et pour différents horizons, du temps de parcours à partir des données de vitesse et de débit d'une route instrumentée. Par efficace, nous entendons une stratégie qui répond à deux exigences: être peu coûteuse et suffisamment précise en terme de des résultats.

A.1.3 Contributions de la thèse

Cette thèse contient des contributions au problème de la prédiction du temps de parcours, l'estimation de la densité du trafic, et le traitement et le groupement des données de trafic.

1. **Prédiction du temps de parcours:** cette thèse propose deux approches pour résoudre la prédiction à court terme et à pas multiples d'indicateurs de trafic routier, en utilisant de façon intégrée des données historiques et des données de la même journée.
 - Approche orientée signal: cette approche est basée sur un filtre de Kalman adaptatif (AKF) qui utilise uniquement des mesures de vitesse. Dans cet approche, le problème de prédiction est transformé en un problème de filtrage, en utilisant comme les observations du système une combinaison appropriée entre les données historiques et les données de la même journée.
 - Approche orientée model: cette approche est basée sur le modèle macroscopique de trafic Cell Transmission model (CTM) et elle utilise des mesures de vitesse et de débit. Cette méthodologie exploite la relation entre la densité et la vitesse. Elle divise la section d'intérêt en des portions de route plus petites, appelées cellules (cf. Fig. 1.8). Cette nouvelle division a pour objectif de suivre avec plus de précision les conditions du trafic dans la section.
2. **Clusterisation de données de trafic:** les profils de temps de parcours calculés à partir des données de vitesse présentent une forte dispersion. Pour surmonter ce problème, nous proposons une nouvelle approche d'agroupement. Cette approche groupe d'abord les profils en jours ouvrables et jours non-ouvrables, ensuite pour chaque groupe on divise tous les profils en cinq tranches horaires, choisies en fonction des conditions de circulation, auxquelles on applique l'algorithme des K plus proches voisins (K-Means).

A.1.4 Revue de temps de parcours

Le temps de parcours est défini comme le temps nécessaire pour parcourir une route entre deux points d'intérêt. Actuellement, il existe plusieurs façons de recueillir ou estimer les données

de temps de parcours. Fig. A.2 résume les principales techniques existantes aujourd’hui. Elles peuvent être divisées en deux groupes: les mesures directes et les mesures indirectes. La principale différence entre les deux groupes est que, tandis que les mesures directes donnent le temps de parcours des véhicules individuels et donc le temps de parcours effective est donné comme la moyenne de ceux-ci, les mesures indirectes caractérisent directement la section de route à partir de l’information obtenue d’un ensemble de les véhicules. Dans la suite, un aperçu général de chaque technique est présenté.

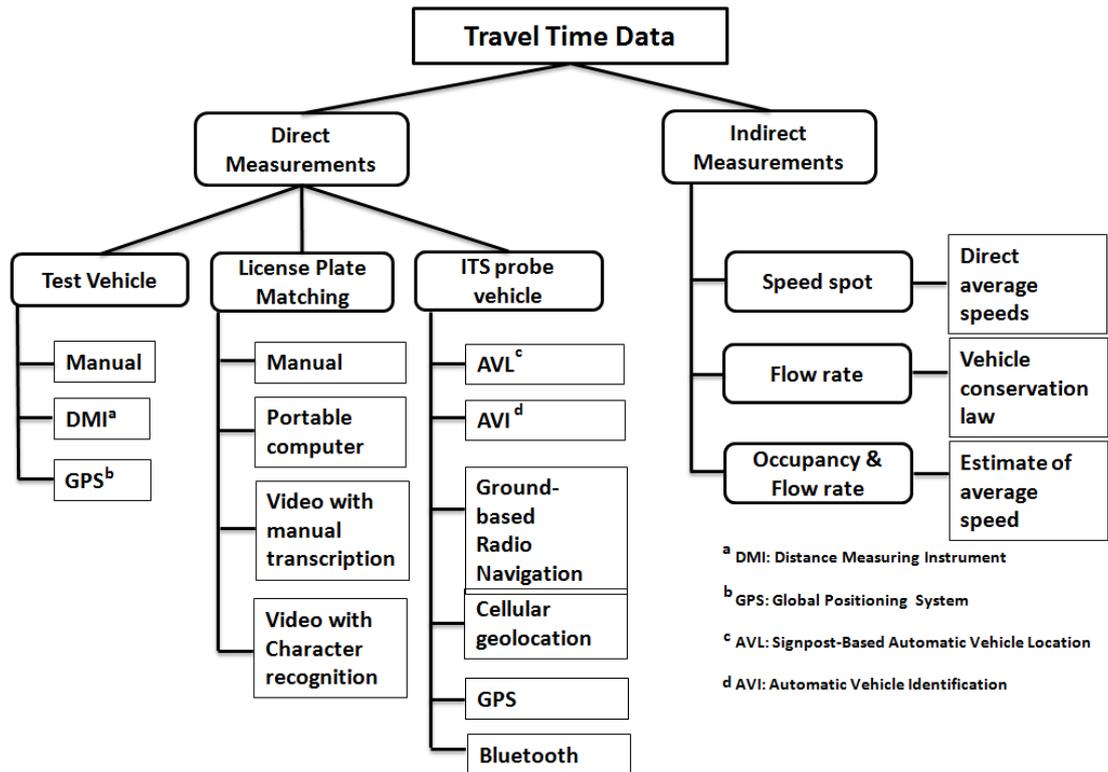


Figure A.2: Différentes techniques pour la collecte de données de temps de parcours.

Temps de parcours à partir de mesures directes

1. **Véhicule d’essai:** cette technique, introduite dans les années 1920, a été la plus commune dans les premières recherches. Il se compose essentiellement d’un véhicule qui est spécifiquement envoyé dans le flux dans le but exprès de la collecte des données. La technique dépend de l’instrumentation utilisée pour la collecte.
2. **Reconnaissance de plaque d’immatriculation:** cette technologie, utilisée depuis le début des années 1950, consiste dans l’identification du numéro de plaque à des véhicules dans points de contrôle spécifiques. Le temps de parcours est calculée à partir de la différence de temps d’arrivée.

3. **Véhicule de sonde ITS:** c'est une technique initialement développée pour la surveillance en temps réel des conditions de circulation plutôt que pour la collecte de temps de parcours. Comme les véhicules sont dans le trafic à d'autres fins que la collecte de données, ils sont souvent dénommé véhicules d'essai passifs. Ces systèmes ont généralement un coût élevé de mise en œuvre, cependant, ils permettent de collecter des données en forme continu et avec une interaction humaine minimale.

Temps de parcours à partir de mesures indirectes

Le temps de parcours peut être estimé à partir des paramètres de trafic macroscopiques: le débit (nombre de véhicules passant à un point par unité de temps), la vitesse (distance parcourue par unité de temps), la densité (nombre de véhicules par unité de distance), ou l'occupation (pourcentage de temps pendant un intervalle de temps fixe, dans lequel un capteur est occupé). Ces informations peuvent être collectées en faisant usage de plusieurs types de technologies qui fournissent les données à temps discret $t_k = k\Delta T$, avec Δt le temps d'échantillonnage.

La vitesse est la variable de la circulation plus étroitement lié à la notion de temps de parcours. Le temps de parcours est inverse de la vitesse par la distance parcourue. Ainsi, lorsque mesures directes de la vitesse sont disponibles le temps de parcours peut être facilement calculé, sinon il est généralement déduit des autres variables de trafic. Deux types de vitesses vont être considérées:

- Vitesse moyenne temporelle (TMS): c'est l'information donnée par les capteurs. Elle est définie comme la vitesse moyenne arithmétique de tous les véhicules pour une période de temps déterminée à un moment précis, l'Éq.A.1.1.

$$v_x(k) = \frac{1}{N_k} \sum_{i=1}^{N_k} v_{x,i}, \quad (\text{A.1.1})$$

où $v_{x,i}$ est la vitesse du véhicule i au point x , et N_k est le nombre de véhicules qui sont passés par x pendant $[k-1, k]$.

- Vitesse moyenne spatial (SMS): c'est la vitesse moyenne de tous les véhicules circulant dans un lien à un moment donné.

$$v_{\Delta x}(k) = \frac{1}{N_{\Delta x}} \sum_{j=1}^{N_{\Delta x}} v_{\Delta x,j}(k), \quad (\text{A.1.2})$$

où $v_{\Delta x,j}(k)$ est la vitesse du véhicule j au temps k qui est à l'intérieur du lien Δx et $N_{\Delta x}$ est le nombre de véhicules dans le lien au temps k . La relation entre vitesse moyenne temporelle et la vitesse moyenne spatial est décrite comme:

$$v_{\Delta x}(k) = 2\left(\frac{1}{v_{x_1}(k)} + \frac{1}{v_{x_2}(k)}\right)^{-1}. \quad (\text{A.1.3})$$

Dans ce qui suit, nous allons nous référer à la Fig. A.3 pour décrire les méthodes pour obtenir le temps de parcours à partir de mesures directes et indirectes de la vitesse.

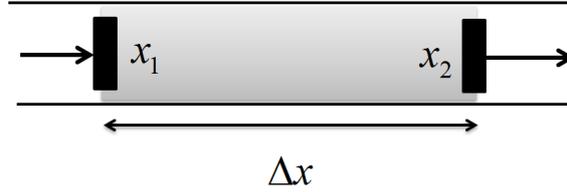


Figure A.3: Lien d'autoroute.

1. Mesure de la vitesse:

Le temps de parcours estimée dans la lien peut être calculée en utilisant l'équation suivante:

$$tt(k) = \frac{\Delta x}{v_{\Delta x}(k)}, \quad (\text{A.1.4})$$

où $v_{\Delta x}(k)$ est donnée par l'Éq. A.1.3.

2. Occupation et débit:

Lorsque la technologie ne permet que la collecte du débit et de l'occupation. La méthode classique est d'abord d'estimer la TMS et ensuite la SMS:

$$\text{TMS} = \frac{\text{flow}}{\text{occupancy} \times g} \rightarrow v_x(k) = \frac{\varphi_x(k)}{o_x(k)g_x(k)}, \quad (\text{A.1.5})$$

où ϕ_x est le débit, o_x l'occupation, et g est la longueur effective moyenne du véhicule (MEVL) évalués à la position x pendant la période de temps considérée. Dans la plupart des mises en œuvre réelle, g n'est pas connue. Ainsi, plusieurs approches ont été proposées dans le but de déduire de sa valeur. La plus simple considère une valeur constante dans le temps [21, 53, 61].

Ensuite, la vitesse du lien et le temps de parcours peuvent être déduites en utilisant l'Éq. A.1.3 et l'Éq. A.1.4 respectivement.

3. Débit et la densité:

Rappelant l'équation de la mécanique des fluides indiquant que le débit est égal au produit de la densité et de la vitesse d'un fluide: $\varphi = \rho v$. La vitesse du fluide v s'écrit:

$$v = \frac{\varphi}{\rho}. \quad (\text{A.1.6})$$

Sous les hypothèses que tous les véhicules parcourant dans un lien à la vitesse v , l'Éq. A.1.6 peut être applicable à la circulation. Ensuite, v correspond à SMS, ρ à la densité du trafic, et φ au débit. Donc, on obtient:

$$\varphi_{\Delta x}(k) = \frac{\varphi_{x_1}(k) + \varphi_{x_2}(k)}{2}.$$

puis:

$$v_{\Delta x}(k) = \Delta x \frac{\rho_{\Delta x}(k)}{\varphi_{\Delta x}(k)}, \quad (\text{A.1.7})$$

où $\rho_{\Delta x}$ est la densité dans le lien. Le temps de parcours dans le lien se déduit de l'Éq. A.1.4.

A.2 Calcul du temps de parcours prédit

Cette section vise à montrer la façon dont laquelle le temps de parcours est prédit. Pour cela, imaginons une carte d'une trajectoire d'un véhicule qui traverse une route limitée par l'intervalle spatial $[x_p, x_0]$ et l'intervalle temporelle $[t_p, t_0]$, comme le montre la Fig. A.4-a. Ici, nous serions intéressés à dériver une formule mathématique pour le point de sortie du véhicule $[x_0, t_0]$, étant donné le point d'entrée $[x_p, t_p]$. Cette formule sera particulièrement utile dans la formulation de temps de parcours prédit.

Si la distribution de vitesse $v(x, t)$ dans $[x_p, x_0]$ est connue, le temps de parcours du véhicule à des intervalles d'espace infinitésimales est donné par $dt \triangleq \frac{dx}{v(x, t)}$, ce qui donne l'équation intégrale suivante, qui donne, en temps continu, la valeur exacte de t_0 :

$$t_0 - t_p = \int_{x_p}^{x_0} \frac{dx}{v(x, t)}. \quad (\text{A.2.1})$$

Normalement, la distribution continue $v(x, t)$ est inconnue. Cependant, en général, les points de mesures sont placés dans les routes. Ces points nous permettent d'approcher $v(x, t)$ par discrétisation de l'intervalle de l'espace.

Avec une discrétisation spatiale de l'Éq. A.2.1, nous divisons la région de l'espace dans un ensemble de n volumes finis de longueur Δx_i , où la vitesse de chaque volume est supposée constante, voir Fig. A.4-b. Par conséquent, $v(x, t)$ est approchée par une fonction linéaire par morceaux de valeurs constantes $v_i(t)$. Maintenant, la version discrète de l'Éq. A.2.2 donne le temps de sortie:

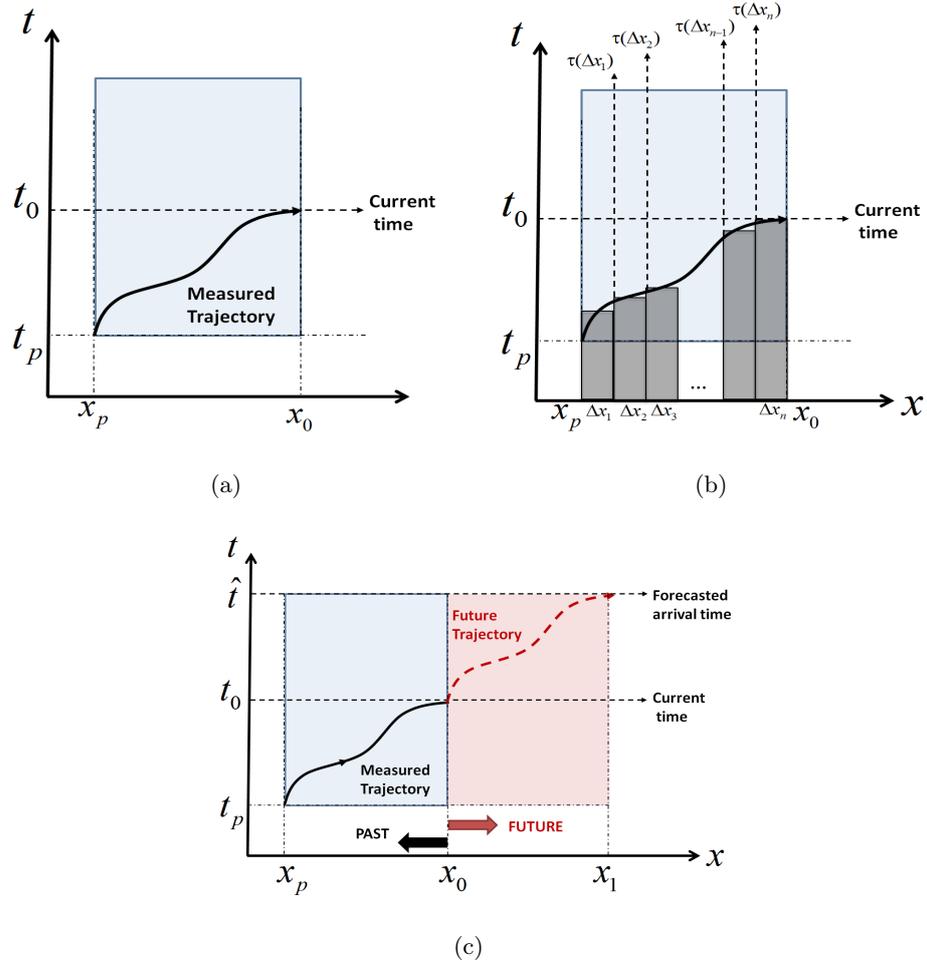


Figure A.4: scénarios envisagés afin d'obtenir une formulation du temps de parcours prédit. a) la trajectoire du véhicule à partir de x_p à x_0 , aux moments de t_p et t_0 respectivement. b) l'espace discrétisation en volumes Δx_i . c) Scénario dans lequel nous cherchons à calculer \hat{t} .

$$t_0 = t_p + \sum_{i=1}^n \frac{\Delta x_i}{v_i(\tau(\Delta x_i))}, \quad (\text{A.2.2})$$

où v_i est la vitesse spatiale moyenne du volume, et $\tau(\Delta x_i)$ est l'heure à laquelle un véhicule a atteint la limite amont du volume i .

Une attention importante doit être donnée à $\tau(\Delta x_i)$. Sa valeur est donnée par:

$$\tau(\Delta x_i) = t_p + \sum_{j=1}^{i-1} \frac{\Delta x_j}{v_j(\tau(\Delta x_j))} \quad (\text{A.2.3})$$

Éq. A.2.3 considère l'effet de la progression du trafic sur la route. Il indique que l'heure d'arrivée du véhicule à l'amont du volume i est une fonction du temps passé dans le volume $i - 1$.

Éq. A.2.2 et l'Éq. A.2.3 définissent le *temps de parcours progressif* (PTT). Dans cette thèse, nous ne serons intéressés que par le PTT.

Maintenant, avec la formulation de temps de sortie, nous pouvons calculer le temps parcours

prédit. Pour cela, considérons le scénario décrit dans la Fig. A.4-c. L'objectif est de prédire l'heure d'arrivée \hat{t} au point x_1 , compte tenu du point d'entrée $[x_0, t_0]$. Ainsi, le problème se réduit à la même situation que précédemment. Le temps de parcours prédit est alors:

$$\begin{aligned}\hat{t} - t_0 &= \sum_{i=1}^n \frac{\Delta x_i}{\hat{v}_i(\hat{\tau}(\Delta x_i))} \\ \hat{\tau}(\Delta x_i) &= t_0 + \sum_{j=1}^{i-1} \frac{\Delta x_j}{\hat{v}_j(\hat{\tau}(\Delta x_j))}.\end{aligned}\tag{A.2.4}$$

Si nous simplifions la notation en tenant compte:

- $\hat{t}t_i(k)$: comme le temps parcours prédit du volume i à un temps discret k .
- $\hat{T}T_i(k)$: comme le temps parcours prédit à partir d'un point d'entrée $[x_0, k]$ à l'aval du volume i .

La prédiction à court terme et à pas multiples du temps de parcours dans l'intervalle $[x_0, x_1]$ à partir du temps courant k_0 , est redéfini:

$$\begin{aligned}\hat{T}T_i(k) &= \hat{T}T_{i-1}(k) + \hat{t}t_i(k + \hat{T}T_{i-1}(k)) \\ \hat{t}t_i(k) &= \frac{\Delta x_i}{\hat{v}_i(\hat{T}T_{i-1}(k))},\end{aligned}\tag{A.2.5}$$

A.3 Collecte et pré-traitement de données de trafic

En raison d'un dysfonctionnement de la collecte des données, il est bien connu que les données de trafic sont toujours corrompues par le bruit, et elles contiennent habituellement des données manquantes et des faux données. Pour ces raisons, leurs qualité doivent être vérifiée. Spécialement, des mesures correctives doivent être prises afin d'assurer que les algorithmes d'estimation et de prédiction sont aussi fiables que possible.

Dans la suite, nous présentons les étapes pour le traitement de données. Dans l'ordre, ces étapes sont: nettoyage des données (l'identification et l'élimination des données erronées), et la réparation de données (remplacement ou imputation des données manquantes et erronées).

A.3.1 Nettoyage de données

Pour les données individuelles et les données agrégées en 15 s, des règles ont été élaborées à fin de rejeter les données erronées. Basé sur l'analyse des données *Sensys*, nous considérons que le nombre de véhicule a la priorité la plus élevée, c'est à dire le système est toujours en mesure de compter le nombre de véhicules, ou un événement, mais pas la vitesse. Lorsque la mesure de vitesse n'est pas disponible, la valeur reçue est égale à -1.

Tab. A.1 montre les règles de combinaison pour le nettoyage de données agrégées à 15 s. Ces règles sont appliquées à tous les échantillons. Si un échantillon de débit ou de vitesse ne

ne passe pas la règle de combinaison, sa valeur est définie null. Une valeur null indique valeur manquante.

Pour les données individuelles, chaque événement représente un véhicule détecté, et sa vitesse est supposée correcte lorsqu'il a été mesuré et d'une valeur inférieure à 150 km/h. Sinon, il est défini null.

A.3.2 Imputation de données

Une étape cruciale pour la réparation des données de trafic est l'imputation des valeurs manquantes ou erronées. Les techniques d'imputation utilisées dans cette thèse sont les suivantes:

1. Moyenne historique: il utilise la moyenne des données historiques pour remplacer les données manquantes.
2. Interpolation des voisins temporelles: l'échantillon manquant est imputé par la moyenne temporelle des échantillons voisins dans le temps à la même station de détection.
3. Hybride: il utilise une analyse des données historiques pour estimer la possible valeur manquante.

La technique d'imputation sont montrés dans Alg. 5 et Alg. 6.

Table A.1: règles de nettoyage de données pour les données agrégées en 15 s.

Règles			Action		Description
Débit	Vitesse	Occupation	Débit	Vitesse	
-	>150	-	-	Jetée	Test 1: Vitesse maximum de 150 km/h
>0	>0	-	Accepté	Acceptée	Test 2: Ceci est supposé être le fonctionnement correct.
=0	-1	-	Accepté	Jetée	Test 3: Pas de débit pendant 15 s. Vitesse est assumée nulle car on arrive pas à savoir si la route est en libre circulation ou congestionnée.
>0	-1	-	Accepté	Jetée	Test 4: Débit mesuré, mais le capteur n'a pas mesuré la vitesse.
=0	>0	>0	Jeté	Acceptée	Test 5: Si la vitesse et l'occupation sont été mesurées, alors le débit est assumé null.

Algorithm 5 Data imputation procedure (OFFLINE)

For each missing sample

Begin:

if TimeN applicable **then**

Missing value \leftarrow Average of temporal neighbors.

else

if HistAv applicable **then**

Missing value \leftarrow Historical average.

else

Missing value \leftarrow Moving average of the last 4 samples.

end if

end if

Algorithm 6 Data imputation procedure (REAL-TIME)

For each missing sample

Begin:

if HistAv applicable **then**

Missing value \leftarrow Historical average.

else

Missing value \leftarrow Moving average of the last 4 samples.

end if

A.4 Prédiction à court terme et à pas multiples de temps de parcours : approche orientée signal

Motivé par l'introduction d'une approche qui n'utilise que des mesures de vitesses, cette section se concentre sur le développement d'un algorithme qui peut prédire, à pas multiples, le temps de parcours entre deux points d'intérêt d'une autoroute.

Il est intéressant de remarquer que chaque problème de prédiction est décentralisé au niveau de chaque lien. La discrétisation de l'espace sera alors fixée par l'emplacement des stations de détection.

A.4.1 Description des données

L'une des composantes essentielles de cette section est la validation des algorithmes avec des données réelles de circulation. La base de données se compose de 15 jours ouvrables, du 2 au 20 Septembre 2013.

Les données brutes expérimentales, téléchargé à partir de GTL, sont pré-traitées en utilisant les stratégies énoncées à la section précédemment.

A.4.2 Formulation du problème

Considérons une section de l'autoroute divisée en n liens comme indiqué sur la Fig. A.5. Etant donné les mesures de vitesse de la voie principale jusqu'à au temps courant k_0 , $k = \{1, \dots, k_0\}$, du jour courant D , $v_i^D(k)$, $i = \{1, \dots, n\}$, et des informations historiques de ces données pour des jours précédents $h = 1, 2, \dots, D - 1$, $v_i^h(k)$, l'objectif est de prédire le temps de parcours dans la section $\hat{T}_n^D(k) \forall k \in [k_0 + 1, \dots, k_0 + \Delta]$.

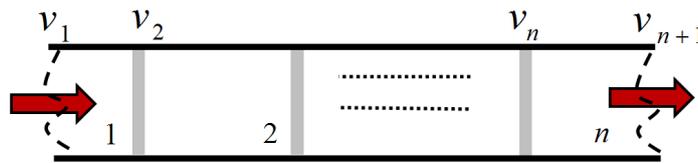


Figure A.5: Abstraction d'une section de l'autoroute. Les bandes grises représentent les stations de détection dans les voies principales.

A.4.3 Filtre de Kalman adaptatif

Les différents méthodes de prédiction du trafic peuvent être classés en fonction du type de données qu'ils utilisent. Les avantages de ces méthodes sont résumés considérant que, en utilisant seulement les données actuelles, les algorithmes sont plus réactifs à l'évolution des conditions de circulation, mais l'horizon de prédiction sont généralement très limitée. D'autre

part, lorsqu'on utilise uniquement des données historiques, l'horizon de prédiction peut être plus grand et la prédiction à plusieurs pas est plus facile; cependant, en général, ces stratégies sont incapables de capter les conditions actuelles. Quand une combinaison des deux informations sont utilisées, les méthodes sont très réactifs aux conditions actuelles et ils sont en mesure de prédire à plus grands horizons.

Cette thèse développe une méthode de prédiction qui fait usage des données de la journée en cours et des données historiques, et qui repose sur la théorie du filtre de Kalman (KF).

Le KF peut fournir des estimations actualisées de l'état à chaque pas si trois informations clés sont fournis: observations bruitées du système réel, les statistiques du bruit des observations et des statistiques sur le bruit du modèle de processus. Dans le cadre de la prédiction, ces informations, en particulier les observations, sont inconnues. Cependant, nous proposons de contourner cet inconvénient en tenant compte de ce qui suit: les observations inconnues peuvent être remplacées par une combinaison appropriée des données de la même journée et des données historiques, appelés pseudo-observations, que leurs statistiques de bruit peuvent être capturées dans la dispersion des données historiques, et enfin que les statistiques de bruit du modèle de processus est stationnaire et peuvent être estimées en ligne grâce à un estimateur sans biais proposé dans la littérature. Les deux premières hypothèses sont basées sur l'hypothèse que les données de la même journée et les données historiques appartiennent à la même fonction de distribution de probabilité.

Dans la section suivante, nous visons à développer la structure de l'AKF. Pour cela, nous présentons d'abord le modèle état-espace considéré. Ensuite, nous décrivons comment les pseudo-observations sont calculées. Par la suite, nous exposons comment sont obtenues les statistiques de bruit des observations et du modèle de processus. Enfin, la représentation finale de l'algorithme AKF est donnée.

Il convient de rappeler que, dans la suite, nous ferons référence au temps de parcours d'un lien générique comme la série temporelle à prédire.

Représentation d'état.

L'évolution du temps de parcours est modélisée par un modèle de marche aléatoire écrit comme:

$$tt^D(k) = tt^D(k-1) + w^D(k), \quad (\text{A.4.1})$$

$w^D(k)$ est considérée comme une réalisation d'un processus aléatoire gaussien blanc $w^D(k) \sim \mathcal{N}(q^D(k), Q^D(k))$.

Le modèle d'observation peut donc être obtenu par:

$$y(k) = Htt^D(k) + \varepsilon^D(k), \quad (\text{A.4.2})$$

où $H \in \mathbb{R}^{m \times 1}$, avec m le nombre de pseudo-observations extraites des données historiques, $\varepsilon^D(k) \sim \mathcal{N}(0, R^D(k))$ est le bruit d'observation, et $y(k) \in \mathbb{R}^{m \times 1}$ le vecteur des pseudo-observations.

Pseudo-observations.

Inspiré par l'approche proposée par Lin dans [51] pour la prédiction de débit, les pseudo-observations seront basées sur la tendance et le niveau des données historiques. Par conséquent, deux pseudo-observations peuvent être formulées comme suit :

- $tt^D(k_0)$: la valeur courante du temps de parcours,
- $tt^h(k)$ et $x^h(k) = tt^h(k) - tt^h(k-1)$: temps de parcours et variation de temps de parcours des jours précédents $h = 1, 2, \dots, D-1$. Où $tt^h(k) \sim \mathcal{N}(\mu_{tt^h}(k), \sigma_{tt^h}^2(k))$ et $x^h(k) \sim \mathcal{N}(\mu_{x^h}(k), \sigma_{x^h}^2(k))$.

Sous les hypothèses suivantes:

- Pour $k \in [k_0+1, \dots, k_0+\Delta]$, les valeurs $tt^h(k)$ et $tt^D(k)$ sont des réalisations indépendantes d'un même processus stochastique.
- Pour $k \in [k_0+1, \dots, k_0+\Delta]$, les incréments $x^h(k)$ et $x^D(k)$ sont des réalisations indépendantes d'un même processus stochastique.

Les pseudo-observations,

$$y(k) = \begin{pmatrix} y^1(k) \\ y^2(k) \end{pmatrix},$$

peuvent être définies comme suit:

$$\begin{aligned} y^1(k) &= \mu_{tt^h}(k) \\ y^2(k) &= \mu_{x^h}(k) + y^2(k-1), y^2(k_0) = tt^D(k_0). \end{aligned} \tag{A.4.3}$$

Comme $\mu_{tt^h}(k)$ et $\mu_{x^h}(k)$ sont disponibles pour $k > k_0$, le calcul du gain de Kalman à chaque pas est possible.

Statistiques du bruit des pseudo-observations.

Le bruit $\varepsilon^D(k)$ dans l'Éq. A.4.2, est supposé être un bruit tiré d'une distribution normale de moyenne nulle de variance $R^D(k)$, i.e. $\varepsilon^D(t) \sim \mathcal{N}(0, R^D(t))$.

De plus, ce bruit est supposé être donné par la dispersion dans les données historiques, comme illustré dans la Fig. A.6.

La matrice de covariance $R(k)$ est définie comme:

$$R^D(k) = \begin{pmatrix} \sigma_{tt^h}^2(k) & 0 \\ 0 & \sigma_{x^h}^2(k) \end{pmatrix} \quad (\text{A.4.4})$$

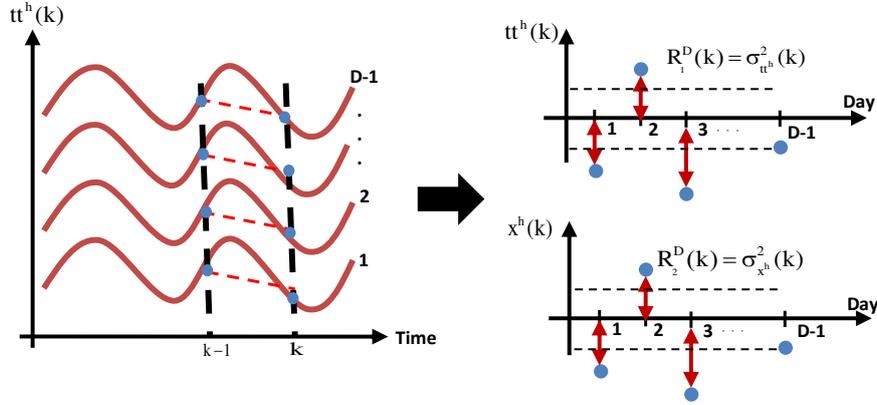


Figure A.6: Le calcul de la covariance du bruit d'observation. À chaque pas k , la variance des données historiques est calculée. Ces écarts seront considérés comme la variance du bruit d'observation.

Statistiques de bruit du processus.

Dans la majorité des applications dans le monde réel, les statistiques du bruit de processus utilisé dans le filtre de Kalman ne sont pas connues. Cependant, ils peuvent généralement être estimés à l'aide des dernières valeurs estimées des états et de l'estimation de la covariance de l'erreur du KF. Ceci est le problème du filtre de Kalman adaptatif (AKF).

Étant donné sa capacité à gérer à la fois les erreurs systématiques et aléatoires, ainsi que d'être bien adapté dans un cadre temps réel, cette étude s'appuie sur la stratégie proposée dans [64] afin d'estimer les statistiques du premier et deuxième ordre de $w^D(k) \sim \mathcal{N}(q^D(k), Q^D(k))$.

En utilisant une fenêtre glissante de longueur N définie par $\{k - (N + 1), \dots, k\}$, une approximation de la moyenne du bruit de processus $q^D(k)$, est donnée par

$$\hat{q}(k) = \frac{1}{N} \sum_{l=1}^N \tilde{q}_l, \quad (\text{A.4.5})$$

où \tilde{q}_l , est le composant l du vecteur $\tilde{q} \in \mathbb{R}^{N \times 1}$ calculé à partir des dernières valeurs d'état estimées: $\tilde{q}_l = \hat{tt}(k - l + 1) - \hat{tt}(k - l)$, $l = \{1, \dots, N\}$.

L'estimation de la variance du bruit est donnée par:

$$\hat{Q}(k) = \frac{1}{N-1} \sum_{l=1}^N ((\tilde{q}_l - \hat{q}(k))(\tilde{q}_l - \hat{q}(k))^T - \frac{(N-1)}{N}(P_{l-1} - P_l)), \quad (\text{A.4.6})$$

où P_L est le composant l de la variance de l'erreur d'estimation mise à jour P .

Schéma AKF.

Dans la Fig. A.11 est représenté le schéma proposé pour la stratégie de prédiction du temps de parcours. Il est important de noter l'usage que nous faisons de l'information de la même journée. Les données actuelles (au moment k_0) et les données passées du même jour sont activement impliquées dans le processus de prédiction, ainsi, nous avons la garantie que, au moins pour un certain intervalle à venir, nous pouvons capturer les conditions de trafic.

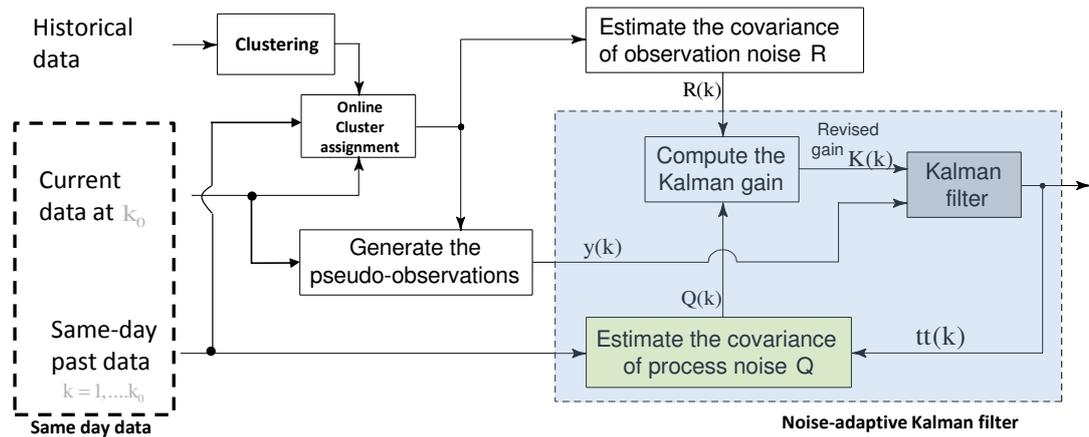


Figure A.7: Schéma AKF pour la prédiction du temps de parcours.

Il est clair que l'algorithme proposé utilise des données historiques pour calculer des pseudo-observations et leurs statistiques. Par conséquent, les sections suivantes visent à développer en premier lieu, une approche pour grouper l'information historique et d'autre part, attribuer à tous les nouvelles données actuelles, un groupe cohérent selon les conditions de trafic actuelles.

Étude des données historiques.

L'objectif principal de cette section est d'identifier et regrouper les caractéristiques répétitives dans les conditions de trafic sur différentes journées.

Dans le but d'améliorer le groupement des données historiques du temps de parcours, cette thèse propose une approche de classification différente. Pour ce faire, les différentes journées sont d'abord classées en jours ouvrables et non-ouvrables, puis différentes tranches horaires sont définies:

1. **Tranche horaire 1:** 00h00-07h00.
2. **Tranche horaire 2:** 07h00-10h00.
3. **Tranche horaire 3:** 10h00-16h00.
4. **Tranche horaire 4:** 16h00-19h00.

5. Tranche horaire 1: 19h00-00h00.

Cette division semble être raisonnable, avec la dynamique de la circulation. Après l'application de la division temporelle, l'algorithme K-Means sera appliquée sur les profils obtenus.

Les groupes avec un seul élément sont considéré comme groupes avec événements spéciaux, et par suite ils ne seront pas considérés comme des événements récurrents et ils seront jetés.

Classification en temps réel des données actuelles.

Dans la section précédente, nous avons analysé les différences dans un ensemble de données afin de placer les profils historiques similaires dans le même group. Cette section d'autre part, vise à attribuer une nouvelle donnée à un groupe donné.

Il a déjà été établi que les groupes contenant un seul élément seraient éliminés car ils sont susceptibles de représenter un événement exceptionnel. Dans l'intérieur de chaque groupe non trivial un centre est situé. Pour assigner un groupe aux données actuelles, une mesure sera définie, et le groupe qui contient le centre plus proche aux données actuelles sera choisi.

L'écart absolu moyen (MAD) a été choisi comme métrique de distance:

$$MAD_i = \frac{1}{N} \sum_{l=k_0-N}^{k_0} |tt^D(l) - \bar{tt}_i^h(l)|, i = \{1, \dots, k\} \quad (\text{A.4.7})$$

où N , k_0 , $tt^D(l)$, et $\bar{tt}_i^h(l)$ correspondent à la longueur de la fenêtre temporelle, l'heure actuelle, la donnée de la journée en cours, et la moyenne du groupe i respectivement.

Pour la mise en œuvre finale une valeur de $N = 10$ a été choisie: elle représente un bon compromis entre la performance et la longueur de la fenêtre.

A.4.4 Résultats expérimentaux: prédiction du temps de parcours

Dans la suite, nous abordons l'évaluation de la méthode de prédiction de temps de parcours présentée. Les expériences sont organisées comme suit. Tout d'abord, nous présentons l'emplacement expérimental choisi, en suite le scénario de test, et enfin les résultats seront donnés.

Pour évaluer la performance de prédiction, nous utilisons le pourcentage d'erreur absolue (APE) défini comme suit:

$$APE(k) = \frac{|tt(k) - \hat{tt}(k)|}{tt(k)} \times 100\% \quad (\text{A.4.8})$$

où $tt(k)$ et $\hat{tt}(k)$ sont les valeurs calculées et prédites du temps de parcours pour le temps k respectivement. Le pas et l'horizon de prédiction seront choisis à 5 min et 45 min respectivement.

La fenêtre temporelle pour l'estimation du bruit de processus sera 5.

Fig. A.8 montre le scénario choisi. Il correspond à un véhicule qui veut traverser toute l'autoroute.



Figure A.8: Chemin choisi pour les tests numériques, du début: Meylan à la fin: Rondeau.

Pour l'ensemble de données disponibles et le chemin sélectionné, Fig. A.9 indique les temps de parcours calculés.

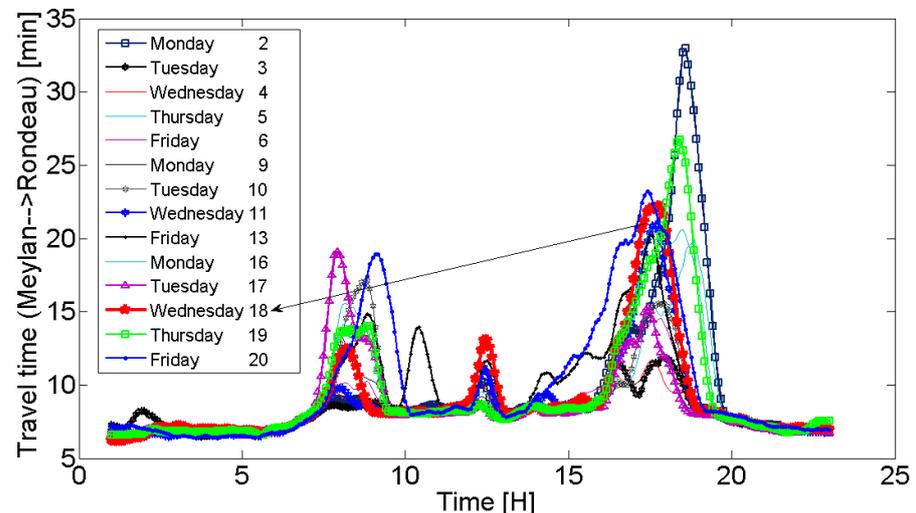


Figure A.9: Temps de parcours expérimenté par un conducteur qui traverse la Rocade sud de Grenoble aux moments différents de départ pour 15 jours ouvrables différents du 2 Septembre 2013 au 20 Septembre.

Afin d'illustrer les résultats, nous avons choisi le mercredi 18 Septembre (cf. Fig. A.9). Pour ce jour, deux scénarios de prédiction seront considérés. Le première correspond à 08h45 et le deuxième à 17h15. Ces heures ont été considérées en raison de leurs niveaux de congestion.

En particulier, à 08h45, nous pouvons étudier le temps d'élimination de congestion, et à 17h15 l'état de congestion. Pour chacun d'eux, les prédictions à pas multiple et la trajectoire prédite d'un véhicule sur le contour de vitesse sont fournis.

L'approche proposée sera également comparée à la moyenne des données historiques comme suit:

$$\hat{\varphi}(k) = \mu_{tt}(k), \quad (\text{A.4.9})$$

où $\mu_{tt}(k)$ est la valeur moyenne des profils de temps de parcours historiques pour $k = \{k_0 + 1, \dots, k_0 + \Delta\}$.

Fig. A.10 montre les résultats expérimentaux. Dans la Fig. A.10-a on observe une grande précision en utilisant la méthodologie proposée. La prédiction traque avec précision le temps d'élimination de la congestion. Dans ce scénario, ils sont également observés des résultats de prédiction assez précis en utilisant la moyenne de l'historique. Ces résultats sont consistents avec la Fig. A.10-b. Les trajectoires du véhicule pour différentes heures de départ sont prédites avec précision.

Fig. A.10-c montre le résultat à 17h15. Ici, la méthode proposée réussit à capturer le temps parcours dans une période de congestion, mais par rapport à la période du matin on retrouve des résultats moins précis, ce qui est normal étant donné que la route est plus congestionnée.

Fig. A.10-d montre de bons résultats pour les trajectoires prédites. De toute évidence, les conditions de circulation affecte la précision des résultats.

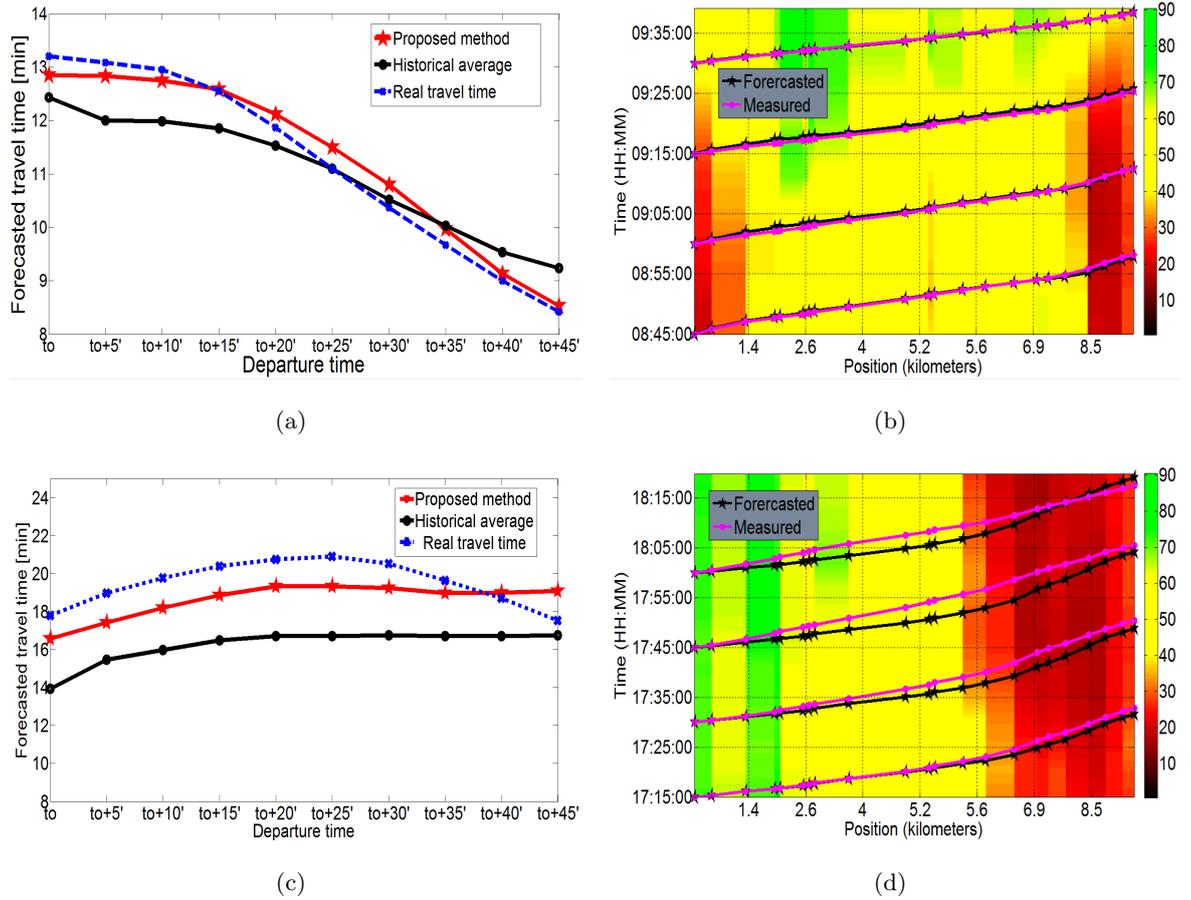


Figure A.10: Résultats de prédiction pour la méthode proposée pour différents scénarios. a) $t_0 = 08h45$. b) trajectoires prédites et mesurées sur le contour de vitesse à $t_0 = 08h45$. c) $t_0 = 17h15$. d) trajectoires prédites et mesurées sur le contour de vitesse à $t_0 = 17h15$.

A.5 Prédiction à court terme et à pas multiples de temps de parcours : approche orientée model

Le but de cette section est de développer et de valider une deuxième méthode pour le problème de prédiction de temps de parcours. Cette méthodologie sera construite sur un modèle macroscopique de trafic, dont un observateur d'état sera développé. Cette nouvelle approche permet la discrétisation de la section d'intérêt dans portions plus petites, appelées cellules. Les conditions aux limites (débit et vitesse de la section) ainsi que les débits des rampes d'entrée et de sortie constituent l'information exploitée dans cette méthode.

Sans entrer dans les détails, le schéma complet pour la prédiction du temps de parcours est donné dans la Fig. A.11, où tous les signaux sont indexés par le temps k , k_0 représentant l'heure actuelle, et Δ l'horizon de prédiction.

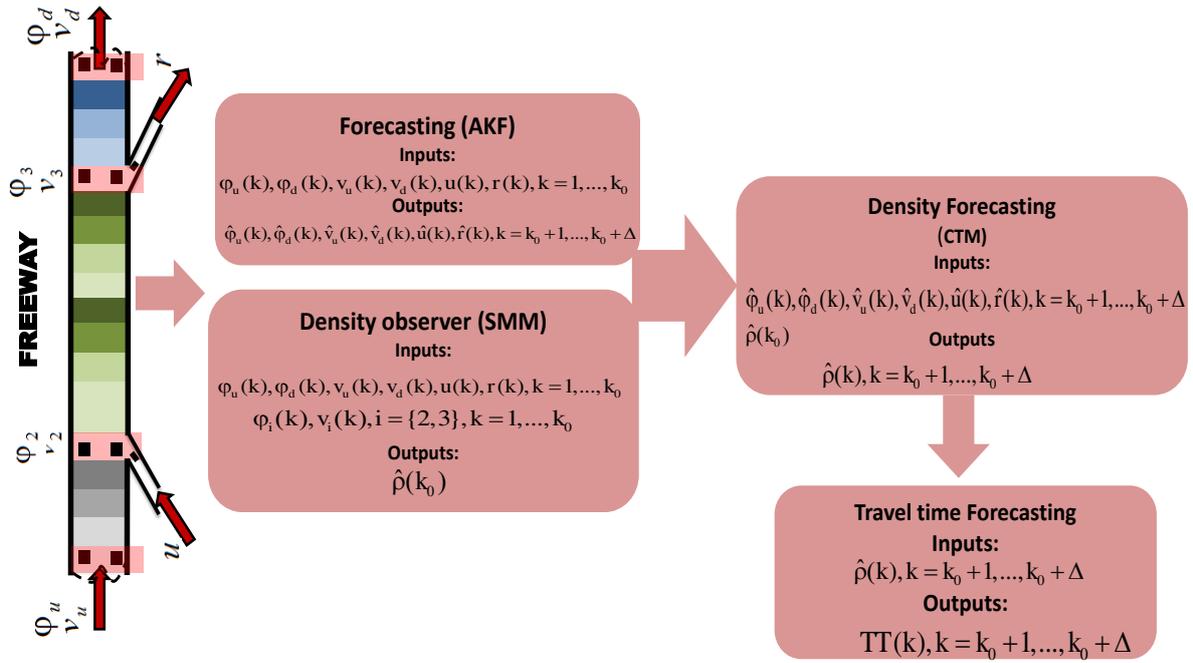


Figure A.11: Schéma de l'algorithme de prédiction orienté modèle.

A.5.1 Description des données

L'une des composantes essentielles de cette section est la validation des algorithmes avec des données réelles de circulation.

La base de données se compose de 12 jours ouvrables, du 27 Février au 15 Mars 2014.

A.5.2 Formulation du problème

Étant donné une section de l'autoroute comme le montre la Fig. A.12, avec des mesures de débit et de vitesse de la voie principale, et les débits des rampes d'entrées et de sorties, jusqu'au temps courant k_0 du jour actuel D , $\varphi_i^D(k)$, $v_i^D(k)$, $i = \{1, \dots, n + 1\}$, $u_l^D(k)$, $l = \{1, \dots, m\}$, et $r_j^D(k)$, $j = \{1, \dots, n\}$, respectivement, ainsi que les mesures de vitesse et de débit des entrées et sorties et des limites de section des jours précédents $h = \{1, 2, \dots, D - 1\}$, $\varphi_1^h(k)$, $v_1^h(k)$, $\varphi_{n+1}^h(k)$, $v_{n+1}^h(k)$, $u_l^h(k)$, $r_j^h(k)$. L'objectif est de prédire le temps de parcours $\hat{TT}(k) \forall k \in [k_0 + 1, \dots, k_0 + \Delta]$. Notez que φ_1 et φ_{n+1} désignent respectivement les débit en amont φ_u et en aval φ_d de la section, de façon analogue les vitesses.

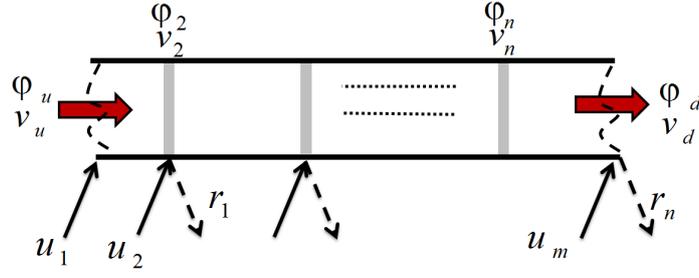


Figure A.12: Abstraction d'une section de l'autoroute. Les bandes grises représentent les stations de détection.

A.5.3 Estimation d'état

Le modèle en premier ordre Lighthill-Whitham-Richards (LWR) initialement proposé dans [48, 72] est le premier et le plus simple des modèles macroscopiques du trafic autoroutier. Il est basé sur la loi de conservation des véhicules.. Ce modèle LWR peut être facilement déduit en termes d'une équation aux dérivées partielles (EDP), dont les solutions analytique et numérique peuvent être connues.

$$\frac{d}{dt} \int_{x_0}^{x_1} \rho(x, t) dx = \varphi(x_0, t) - \varphi(x_1, t). \quad (\text{A.5.1})$$

Pour une section $[x_0, x_1]$, la loi de conservation de véhicules indique que le taux de variation du nombre total de véhicules dans la section est donnée par la différence des débits à x_0 et x_1 . Il est possible de récrire cette équation de la manière suivante::

$$\partial_t \rho + \partial_x \Phi(\rho) = 0. \quad (\text{A.5.2})$$

où $\rho(x, t)$ et $\varphi(x, t)$ représentent les distribution espace-temps de la densité et de la vitesse respectivement.

La fonction $\Phi(\rho)$ est également connu comme le diagramme fondamental, et elle est généralement supposé être concave et continue par morceaux.

En supposant que ρ est dérivable par rapport au temps et φ par rapport à l'espace, l'Éq. A.5.1 peut être discretisée dans l'espace et dans le temps, ce qui donne:

$$\rho_i(k+1) = \rho_i(k) - \frac{\Delta t}{\Delta x_i} (\Phi(\rho(x_{i-1/2}, k)) - \Phi(\rho(x_{i+1/2}, k))). \quad (\text{A.5.3})$$

Les valeurs numériques de $\rho(x_{i-1/2}, k)$ et $\rho(x_{i+1/2}, k)$ sont calculées en résolvant un problème de Riemann local. Pour ce, nous considérons que (see Fig. A.13-a):

$$\rho(x, k) = \begin{cases} \rho_{i-1}(k) & \text{if } x < x_{i-1/2} \\ \rho_i(k) & \text{if } x > x_{i-1/2} \\ \rho_i(k) & \text{if } x < x_{i+1/2} \\ \rho_{i+1}(k) & \text{if } x > x_{i+1/2}. \end{cases} \quad (\text{A.5.4})$$

Sous l'hypothèse que la fonction $\Phi(\rho)$ est concave (Fig. A.13-b), où ρ_c correspond au point où la fonction change le signe de sa dérivée (densité critique) et ρ_m est la densité maximale (jam density), la solution à $x_{i-1/2}$ est la suivante:

$$\begin{aligned}
 \rho(x_{i-1/2}, t) = \rho_{i-1}(k) & \quad \text{if} & \quad \partial_\rho \Phi(\rho_{i-1}(k)) \geq 0 \wedge \partial_\rho \Phi(\rho_i(k)) \geq 0 \\
 \rho(x_{i-1/2}, t) = \rho_i(k) & \quad \text{if} & \quad \partial_\rho \Phi(\rho_{i-1}(k)) < 0 \wedge \partial_\rho \Phi(\rho_i(k)) < 0 \\
 \rho(x_{i-1/2}, t) = \rho_{i-1}(k) (z > 0) & \quad \text{or} & \quad \text{if} & \quad \partial_\rho \Phi(\rho_{i-1}(k)) \geq 0 \wedge \partial_\rho \Phi(\rho_i(k)) < 0 \\
 \rho(x_{i-1/2}, t) = \rho_i(k) (z < 0) & & & \\
 \rho(x_{i-1/2}, t) = \rho_c & \quad \text{if} & \quad \partial_\rho \Phi(\rho_{i-1}(k)) < 0 \wedge \partial_\rho \Phi(\rho_i(k)) \geq 0
 \end{aligned} \tag{A.5.5}$$

La solution de Riemann à $x_{i+1/2}$ est effectuée de la même manière.

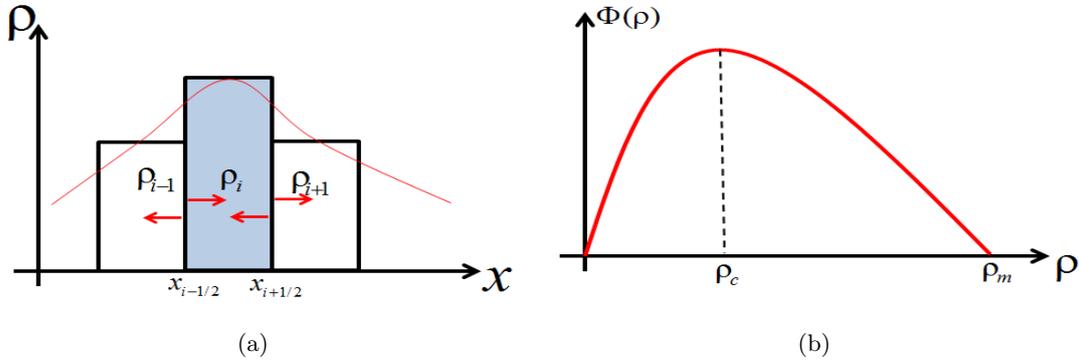


Figure A.13: (a) Sous-division du domaine spatial en trois cellules, nous sommes intéressés dans le calcul des valeurs de densité au niveau des interfaces de la cellule i , (b) Fonction de débit concave $\Phi(\rho)$.

Afin d'assurer la stabilité numérique, les pas de temps et d'espace sont liés par: $v_{\max} \Delta t \leq \Delta x$, où v_{\max} désigne la vitesse libre.

Pour des raisons de clarté des éléments suivants, ces définitions sont dans l'ordre:

Definition 1. Une cellule est considérée comme libre si sa densité associée ρ_i est inférieure ou égale à sa densité critique $\rho_{c,i}$, sinon, elle est considérée comme congestionnée (C) ($\rho_{c,i} < \rho_i \leq \rho_{m,i}$).

Definition 2. L'interface d'une cellule, $x_{i-1/2}$ par exemple, est considérée comme libre (F) si l'onde de choc à partir de ce point se déplace vers l'avant ($z > 0$), sinon, elle est considérée comme congestionnée (C) ($z < 0$). Dans le premier cas $\varphi(x_{i-1/2}, t)$ est dite onde de diffusion (fronts d'onde se déplacent vers le bas), tandis que dans le second cas il s'agit d'une onde d'advection (fronts d'onde se déplacent vers le haut), ils seront désignés comme $\vec{\varphi}(x_{i-1/2}, t)$ et $\overleftarrow{\varphi}(x_{i-1/2}, t)$ respectivement.

Cell Transmission Model (CTM)

CTM a été proposé par Daganzo dans [20]. Le CTM est en effet la discrétisation de premier ordre de l'équation de Godunov Éq. A.5.2 sous l'hypothèse que le diagramme fondamental a la forme triangulaire donnée par la Fig. A.14.

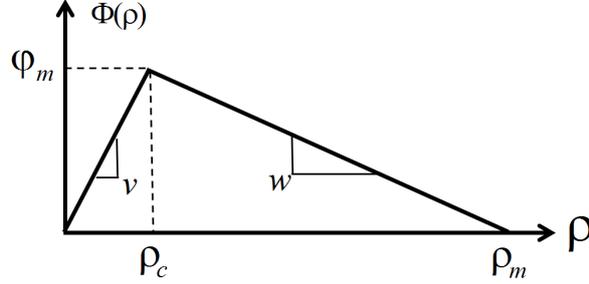


Figure A.14: Le diagramme fondamental. La forme est déterminée par les paramètres: φ_m - capacité maximum, v - vitesse libre, w - vitesse de congestion.

Ce modèle suppose une représentation de l'autoroute comme une séquence de n cellules, où l'interface de la cellule peut être accompagnée ou pas par des rampes, voir la Fig. A.15-a et la Fig. A.15 -b. Dans ce contexte, la Fig. A.15-a est appelée un cas homogène, tandis que la Fig. A.15-b est appelée un cas non homogène.

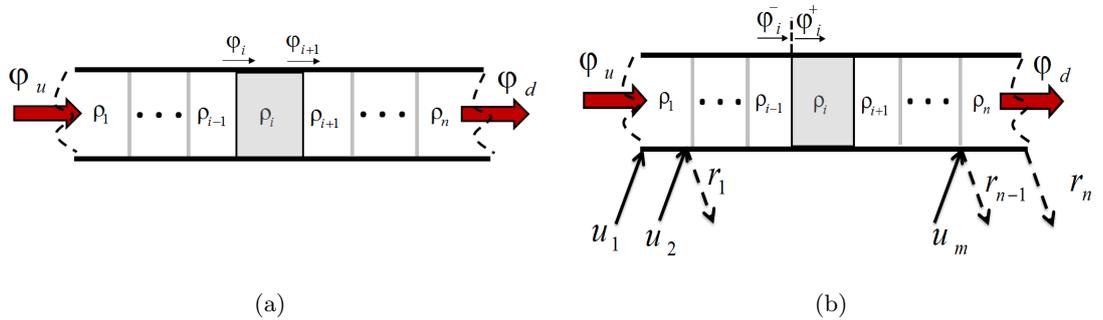


Figure A.15: Représentation d'une autoroute en utilisant le modèle CTM. (a) section homogène divisée en n cellules de longueur L_i et densités ρ_i , $i = 1, \dots, n$, (b) section homogène divisée en n cellules de longueur L_i et densités ρ_i , $i = 1, \dots, n$. Chaque cellule est accompagnée d' au plus une rampe d'accès et une rampe de sortie.

Avant de présenter les équations du modèle, pour le cas homogène et non homogène, nous allons présenter d'abord la notation qui sera utilisé pour le modèle CTM et suite les fonctions de la demande et l'offre:

Notation du modèle CTM

- ρ [veh/km]: densité,
- L [km]: longueur de la cellule,
- T [h]: temps discret,

- φ [veh/h]: débit de la voie principale,
- u [veh/h]: débit de la rampe d'accès,
- r [veh/h]: débit de la rampe de sortie,
- β : split ratio,
- v [km/h]: vitesse libre,
- w [km/h]: vitesse de congestion,
- ρ_c [veh/km]: densité critique,
- ρ_m [veh/km]: densité maximum,
- φ_m [veh/km]: capacité maximum,

Fonctions: Demande/Offre

La solution de Riemann peut être restructurée en utilisant la demande D_i et l'offre S_i (Fig. A.16):

$$\begin{aligned} D_i &= \begin{cases} \Phi_i(\rho) & \text{if } \rho_i \leq \rho_{c,i} \\ \varphi_{m,i} & \text{if } \rho_i > \rho_{c,i} \end{cases} \\ S_i &= \begin{cases} \varphi_{m,i} & \text{if } \rho_i \leq \rho_{c,i} \\ \Phi_i(\rho) & \text{if } \rho_i > \rho_{c,i} \end{cases} \end{aligned} \quad (\text{A.5.6})$$

La demande D_i constitue le débit pouvant être produit par la cellule i , tandis l'offre S_i est le débit pouvant être reçu dans la cellule $i + 1$.

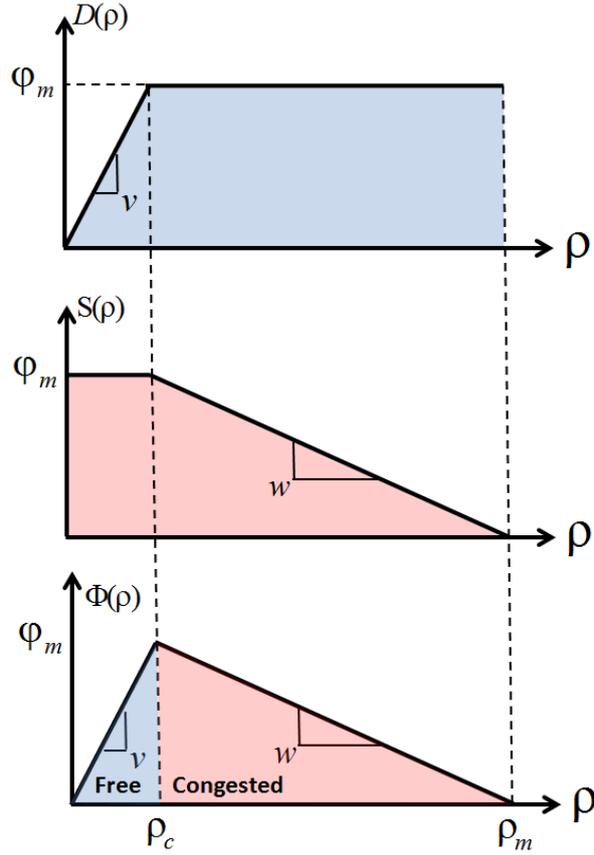


Figure A.16: Les fonctions de la demande et de l'offre.

Équations dynamiques: section homogène

En considérant un modèle d'autoroute comme celui décrit par la Fig. A.15-a, le modèle CTM est décrit comme:

$$\rho_i(k+1) = \rho_i(k) + \frac{T}{L_i}(\varphi_i(k) - \varphi_{i+1}(k)). \quad (\text{A.5.7})$$

Le débit d'interface φ_i entre les cellules $i-1$ et i est:

$$\begin{aligned} \varphi_i &= \min\{D_{i-1}, S_i\} \\ D_{i-1} &= \min\{v_{i-1}\rho_{i-1}, \varphi_m\} \\ S_i &= \min\{\varphi_m, w_i(\rho_{m,i} - \rho_i)\} \end{aligned} \quad (\text{A.5.8})$$

Selon l'état de l'interface des cellules, le tableau Tab. A.2 résume l'expression du débit traversant.

Section non-homogène

Pour la représentation de la Fig. A.15-b chaque cellule est accompagnée d'au plus une rampe d'accès et une rampe de sortie. En sachant que le débit de sortie r peut aussi être exprimé

Table A.2: Débit traversant pour une section homogène

Interface		
Condition	Mode	Interface
$v_{i-1}\rho_{i-1}(k) \leq w_i(\rho_m - \rho_i(k))$	Libre	$\varphi_i(k) = v_{i-1}\rho_{i-1}(k)$
$v_{i-1}\rho_i(k) > w_i(\rho_m - \rho_i(k))$	Congestionée	$\varphi_i(k) = w_i(\rho_m - \rho_i(k))$

en termes de β , où $\beta \in [0, 1)$. β représente le rapport de division, et indique le nombre de véhicules sortant par la rampe.

Le modèle CTM, en vertu de la présence de rampes, peut être écrit comme:

$$\rho_i(k+1) = \rho_i(k) + \frac{T}{L_i}(\varphi_i^+(k) - \varphi_i^-(k)) \quad (\text{A.5.9})$$

φ_i^- and φ_i^+ sont donnés dans Tab. A.3.

Table A.3: Débit traversant pour une section non homogène.

interface pour l'accès		
Condition	Mode	Interface
$v_{i-1}\rho_{i-1}(k) + u(k) \leq w_i(\rho_m - \rho_i(k))$	Libre	$\varphi_i^-(k) = v_{i-1}\rho_{i-1}(k)$ $\varphi_i^+(k) = v_{i-1}\rho_{i-1}(k) + u(k)$
$v_{i-1}\rho_{i-1}(k) + u(k) > w_i(\rho_m - \rho_i(k))$	Congestionée	$\varphi_i^-(k) = w_i(\rho_m - \rho_i(k)) - u(k)$ $\varphi_i^+(k) = w_i(\rho_m - \rho_i(k))$
Interface pour la sortie (en fonction de $r(k)$)		
Condition	Mode	Interface
$v_{i-1}\rho_{i-1}(k) - r(k) \leq w_i(\rho_m - \rho_i(k))$	Libre	$\varphi_i^-(k) = v_{i-1}\rho_{i-1}(k)$ $\varphi_i^+(k) = v_{i-1}\rho_{i-1}(k) - r(k)$
$v_{i-1}\rho_{i-1}(k) - r(k) > w_i(\rho_m - \rho_i(k))$	Congestionée	$\varphi_i^-(k) = w_i(\rho_m - \rho_i(k)) + r(k)$ $\varphi_i^+(k) = w_i(\rho_m - \rho_i(k))$

Le modèle CTM présenté peut être capturé par un ensemble fini de systèmes linéaires avec une fonction de transition. Cette fonction dépend de l'état de congestion des cellules et la direction des fronts d'ondes. Ce système à commutation est connu sous le nom de Switching Mode Model (SMM).

Avec l'utilisation de SMM, des observateurs de la densité peuvent être développés. Ce travail de thèse est basé sur ce type d'observateur. Les avantages du modèle SMM pour résoudre le problème d'estimation ont été mis en évidence pour la première fois dans [63]. Cet estimateur a permis l'utilisation d'un ensemble d'équations linéaires pour décrire l'évolution de l'état pour les différents régimes de l'autoroute. Ce travail a proposé un observateur basé sur deux

hypothèses. Tout d'abord, que les densités et les débits aux frontières de la section, ainsi que les débits sur toutes les rampes sont mesurés. D'autre part, qu'il y ait au plus un front d'onde dans la section. Dans ce travail, les auteurs ont également présenté une étude de l'observabilité du système sur la base du modèle SMM. Cependant la fonction de commutation est peu robuste aux bruits de mesure et aux erreurs d'estimation.

Afin de remédier à cet inconvénient, dans une contribution plus récente [62], les auteurs ont introduit une hypothèse supplémentaire. Les congestions apparaissent toujours à la dernière cellule et se propagent en amont. L'hypothèse supplémentaire réduit le nombre de modes possibles par rapport à [63] et contraint le système à respecter un certain ordre de transition. Dans ce qui suit, nous allons réaffirmer l'idée présentée dans [62] avec le bon nombre de modes et introduire le concept d'observateur CTM à contraintes graphiques.

A.5.4 Observateur CTM à contraintes graphiques

Cette partie du travail concerne l'étude d'un réseau conçue comme une séquence de nœuds liés par liens. Le but est l'estimation de la densité à l'intérieur des liens, chaque lien est partitionné dans plusieurs cellules (Fig. A.15-a). Des points de mesures sont positionnés aux limites de chaque lien.

En faisant usage de Def.1 et Def.2, et en supposant que les congestions apparaissent toujours à la cellule n et se propagent en amont. Le nombre de modes peut être évalué. En effet, dans une telle hypothèse, seules les combinaisons cellulaires suivantes existent: FF , FC , CC . Pour chacune de ces trois combinaison nous avons:

- $F_{i-1}F_i$: $\varphi_i = \overrightarrow{\varphi}_i = D_{i-1}$ est une onde de diffusion
- $C_{i-1}C_i$: $\varphi_i = \overleftarrow{\varphi}_i = S_i$ est une onde d'advection
- $F_{i-1}C_i$: φ_i est une onde de diffusion si $(D_{i-1} \leq S_i)$ ou d'advection si $(D_{i-1} > S_i)$.

Les paramètres du modèle CTM sont identifiés à partir de données réelles. À titre d'illustration, on considère une section de 3 cellules, comme indiqué dans Fig. A.17, afin d'obtenir la structure de la dynamique du système. Pour cela, nous avons les combinaisons suivantes: FFF , FFC , FCC et CCC . En remarquant que chaque interface de type FC peut avoir deux modes possibles, alors, y compris les débits d'interface amont et en aval, nous pouvons identifier un total de 8 modes différents qu'on peut classer comme indiqué dans la Table A.4, où $\overrightarrow{\cdot}$ et $\overleftarrow{\cdot}$ indiquent un débit d'interface de diffusion ou d'advection, respectivement. Les modes sont indiqués alternativement par la variable entière $s(k)$ ou par la lettre associée avec la flèche supérieure qui indique la direction des ondes. La notation de lettre sera utile afin de distinguer facilement les modes de diffusion $\overrightarrow{(\cdot)}$ de ceux d'advection $\overleftarrow{(\cdot)}$.

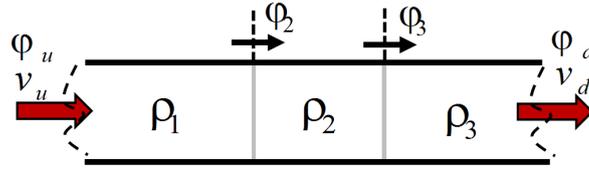


Figure A.17: Section considérée afin de dériver les équations dynamiques du système.

Table A.4: Table avec les modes du système

Mode $s(k)$	Mode $\overleftrightarrow{(\cdot)}$	Configuration des cellules
1	\overrightarrow{a}	$FFF \overrightarrow{ }$
2	\overleftarrow{a}	$FFF \overleftarrow{ }$
3	\overrightarrow{b}	$FF \overrightarrow{ } C$
4	\overleftarrow{b}	$FF \overleftarrow{ } C$
5	\overrightarrow{c}	$F \overrightarrow{ } CC$
6	\overleftarrow{c}	$F \overleftarrow{ } CC$
7	\overrightarrow{d}	$\overrightarrow{ } CCC$
8	\overleftarrow{d}	$\overleftarrow{ } CCC$

Le nombre total de modes dans le cas contraint, est alors:

$$M = 2(n + 1).$$

Ces modes ont des règles associé à un graphe \mathcal{G} car les transitions entre les modes ne sont pas arbitraires, elles suivent les règles suivantes:

- une seule onde de congestion peut exister dans une section et elle se propage suivant le modèle: FFF , FFC , FCC , CCC ,
- Les transitions $FF \rightarrow FC$ sont possibles d'un mode d'advection vers un mode de diffusion,
- Les transitions $FC \rightarrow FF$ sont possibles d'un mode de diffusion vers un mode d'advection.

Ces deux dernières propriétés peuvent être obtenues à partir de la continuité du diagramme fondamental pendant les phases transitoires. Les détails du graphe sont présentés dans la Fig. A.18.

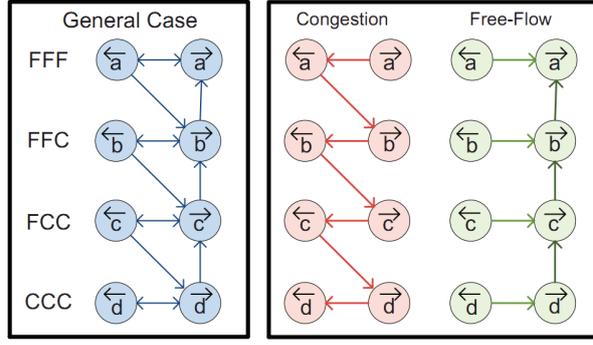


Figure A.18: Illustration des règles de transition associées au graphe \mathcal{G} pour l'observateur CTM à contraintes graphiques. Exemple de 3 cellules.

La représentation d'état, dont $\rho = (\rho_1, \dots, \rho_n)$ dénote l'état du système, est donnée par:

$$\begin{cases} \rho(k+1) = A_{s(k)}\rho(k) + B_{s(k)}\varphi(k) + \bar{B}_{s(k)}\rho_m \\ s(k) = \mathcal{G}(\rho(k), \varphi(k)) \\ y(k) = h(\rho(k), s(k)) \end{cases} \quad (\text{A.5.10})$$

$$h(\rho(k), s(k)) = \begin{cases} v_n \rho_n(k) & \text{if } s(k) = 1 \\ w_1(\rho_{m,1} - \rho_1(k)), & \text{if } s(k) = M \\ 0, & \text{sinon} \end{cases} \quad (\text{A.5.11})$$

où les données mesurées du système sont les débits et les vitesses en amont et aval, (φ_u, φ_d) et (v_u, v_d) respectivement, tandis que l'index $s(k) \in \{1, 2, \dots, M\}$ précise le mode de la section.

Les matrices $A_{s(k)} \in \mathbb{R}^{n \times n}$, $B_{s(k)} \in \mathbb{R}^{n \times 2}$, $\bar{B}_{s(k)} \in \mathbb{R}^{n \times 1}$, $\forall i \in \{1, 2, \dots, M\}$ peuvent être trouvées dans [62].

Observabilité

L'observabilité du modèle CTM pour différents modes de SMM a été étudiée ailleurs, voir par exemple [63] et les références qui y sont. Les résultats sont résumés dans la Table A.5.

Table A.5: Observabilité du modèle SMM

Cellule en amont	Cellule en aval	Observable avec
Libre	Libre	Mesures en aval
Congestionnée	Congestionnée	Mesures en amont
Congestionnée	Libre	Mesures en aval et amont
Libre	Congestionnée	Inobservable

Structure de l'observateur.

Un observateur hybride qui utilise uniquement les conditions aux limites de la section peut être construit à partir de Éq. A.5.10 en reproduisant ces dynamiques et l'activation d'un terme de correction lorsque les modes sont observables.

$$\begin{cases} \hat{\rho}(k+1) &= A_{\hat{s}(k)}\hat{\rho}(k) + B_{\hat{s}(k)}\varphi(k) + \bar{B}_{\hat{s}(k)}\rho_m + \\ &+ K_{\hat{s}(k)}(y(k) - h(\hat{\rho}(k), \hat{s}(k))) \\ \hat{s}(k) &= \mathcal{G}(\hat{\rho}(k), \varphi(k)). \end{cases} \quad (\text{A.5.12})$$

où $K_{\hat{s}(k)}$ le gain de l'observateur. Dans ce travail, nous avons adopté une approche de placement des pôles pour calculer $K_{\hat{s}(k)}$.

Fonction de transition en fonction des mesures de vitesse.

Afin d'améliorer la sélection des modes du système, on utilisera les données des vitesses aux limites de la section. À titre d'illustration rappelons l'exemple 3 cellules dans la Fig. A.17. Pour cela, les conditions de transition peuvent être résumées dans Tab. A.6.

Table A.6: Transitions en utilisant données des vitesses pour un exemple de 3 cellules

Mode $s(k)$	Condition
1	$(\hat{\rho}_1 \leq \rho_{c,1}) \wedge (\hat{\rho}_2 \leq \rho_{c,2}) \wedge (\hat{\rho}_3 \leq \rho_{c,3}) \wedge (v_d \geq 0.9v_3)$
2	$(\hat{\rho}_1 \leq \rho_{c,1}) \wedge (\hat{\rho}_2 \leq \rho_{c,2}) \wedge (\hat{\rho}_3 \leq \rho_{c,3}) \wedge (v_d < 0.9v_3)$
3	$(\hat{\rho}_1 \leq \rho_{c,1}) \wedge (\hat{\rho}_2 \leq \rho_{c,2}) \wedge (\hat{\rho}_3 > \rho_{c,3}) \wedge (\hat{D}_2 \leq \hat{S}_3)$
4	$(\hat{\rho}_1 \leq \rho_{c,1}) \wedge (\hat{\rho}_2 \leq \rho_{c,2}) \wedge (\hat{\rho}_3 > \rho_{c,3}) \wedge (\hat{D}_2 > \hat{S}_3)$
5	$(\hat{\rho}_1 \leq \rho_{c,1}) \wedge (\hat{\rho}_2 > \rho_{c,2}) \wedge (\hat{\rho}_3 > \rho_{c,3}) \wedge (\hat{D}_1 \leq \hat{S}_2)$
6	$(\hat{\rho}_1 \leq \rho_{c,1}) \wedge (\hat{\rho}_2 > \rho_{c,2}) \wedge (\hat{\rho}_3 > \rho_{c,3}) \wedge (\hat{D}_1 > \hat{S}_2)$
7	$(\hat{\rho}_1 > \rho_{c,1}) \wedge (\hat{\rho}_2 > \rho_{c,2}) \wedge (\hat{\rho}_3 > \rho_{c,3}) \wedge (v_u \geq 0.9v_1)$
8	$(\hat{\rho}_1 > \rho_{c,1}) \wedge (\hat{\rho}_2 > \rho_{c,2}) \wedge (\hat{\rho}_3 > \rho_{c,3}) \wedge (v_u < 0.9v_1)$

A.5.5 Résultats expérimentaux: estimation d'état

Fig. A.19 indique l'emplacement expérimentale qui sera utilisée dans cette section.



Figure A.19: Emplacement expérimental considéré pour la reconstruction d'état.

Pour chacun des 19 liens, nous recueillons les conditions aux limites, les vitesses, et les débits, et on estime la densité dans chaque cellule.

Fig. A.20 montre, avec des tracés de contours, la comparaison entre les densités mesurées et celles reconstruites pour deux jours différents.

Les jours sélectionnés à titre d'illustration ont été choisis parmi les jours qui ont présenté les plus forts embouteillages. Ces journées ont été le 28 Février et le 9 Mars 2014.

En termes de précision, il est perçu une bonne reconstruction des conditions du trafic, même dans les périodes de grands embouteillages. Comme mentionné ci-dessus, les congestions sont plus fortes dans la période de l'après-midi. Elles se propagent vers l'arrière à partir de la fin de l'autoroute. En général, la partie la plus intéressante de l'autoroute, en termes de congestion, est la seconde moitié (de l'entrée 5 dans la Fig. A.19).

A.5.6 Prediction de conditions aux limites et débits des rampes d'accès et de sortie

La méthodologie adoptée sera l'approche du filtre de Kalman adaptatif (AKF) présentée dans la section précédente. Par conséquent, nous ne fournirons que le schéma de l'algorithme et une vue du traitement de l'ensemble des données historiques.

AKF

Fig. A.21 décrit le schéma utilisé pour la prédiction des profils de débit. Cette stratégie

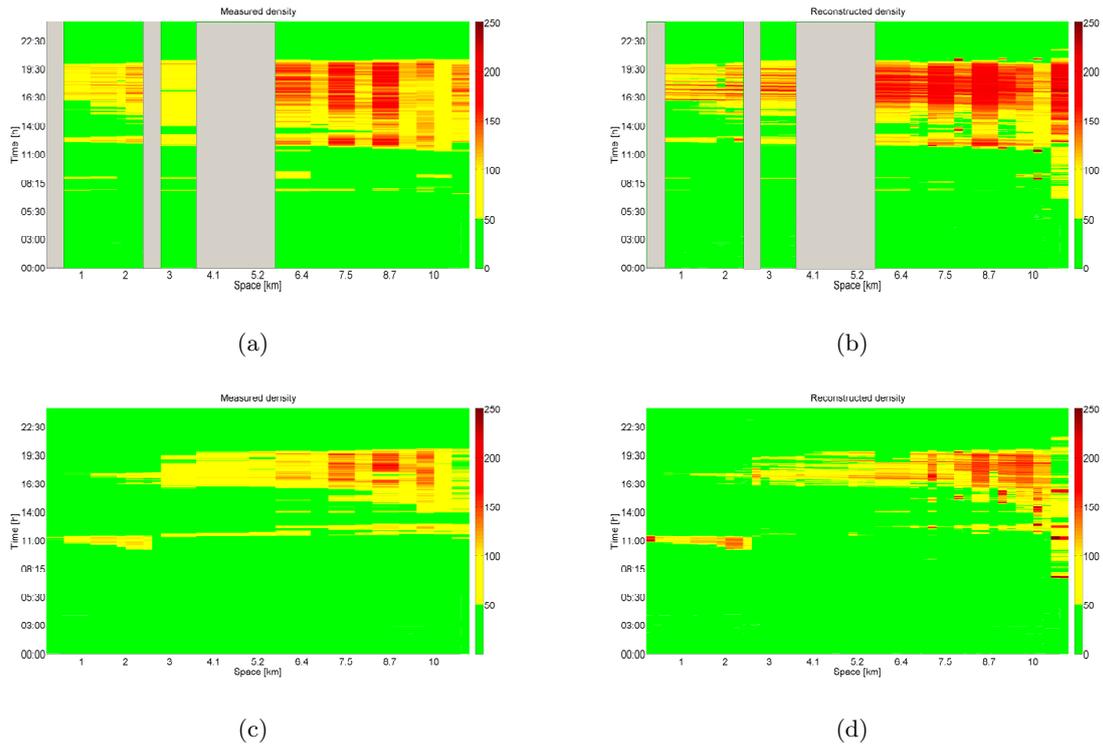


Figure A.20: Résultats de l'observateur de densité pour différents jours: a),b) Vendredi: 28 Février 2014, c),d) Vendredi: 7 Mars 2014. Densité mesurée à gauche, densité reconstruite à droite.

de prédiction peut être appliquée à n'importe quelle série temporelle sous l'hypothèse que les données historiques soient disponibles.

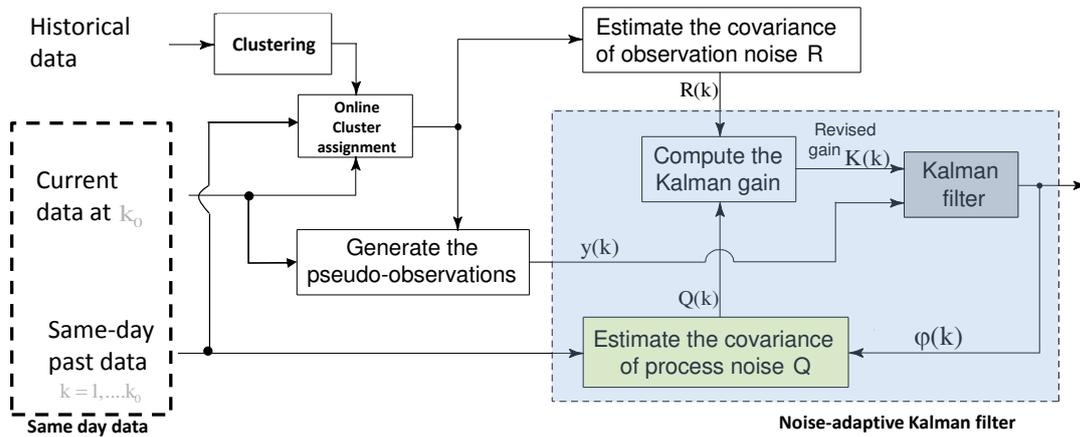


Figure A.21: schéma AKF pour prédiction de débit.

Un des éléments clés de cette approche était en effet le traitement des données historiques (bloc d'agroupement off-line).

Étant donné que les profils de flux de trafic ont plus de corrélation entre eux, voir la Fig. A.22. L'agroupement des profils se fera d'abord en séparant les profils dans deux groupes, jours ouvrables et jours non-ouvrables, et en suite en appliquant K-means pour chaque groupe.

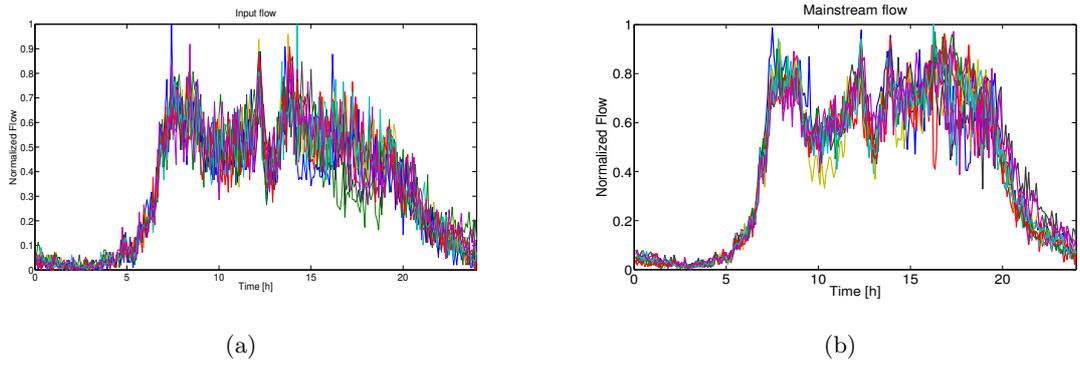


Figure A.22: Profils de débit normalisés et agrégés à 5 min recueillies de la Rocade Sud de Grenoble du 27 Février au 15 Mars de 2014: a) Rampes d'accès 6, b) voie principal.

À ce stade, nous avons développé des algorithmes opérationnels pour la reconstruction de la densité et de la prédiction des débits. Dans la suite, la stratégie pour la prédiction de la densité et temps de parcours sera présentée.

A.5.7 Prédiction de densité

Nous visons à prédire les densités à l'intérieur de chaque cellule de la section qui nous intéresse. Pour cela, nous allons utiliser, une fois de plus, l'approche basée sur un modèle comme pour le problème de la reconstruction de la densité. Ici, la principale différence est que nous allons utiliser le modèle CTM en boucle ouverte, dont les entrées sont les débits prédites: conditions aux limites, rampes d'entrées et rampes de sorties, c'est à dire sans les débits de la voie principale. L'avantage évident de cette approche est qu'elle réduit la quantité de calcul et la taille de la base de données historique de l'algorithme.

À titre d'illustration, considérons une section comme indiqué dans Fig. A.23.

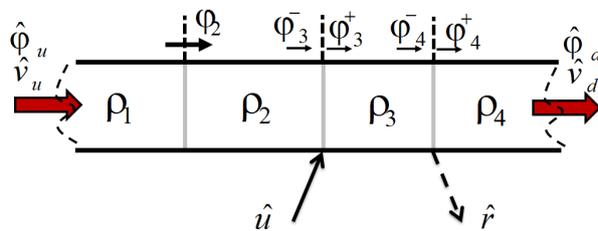


Figure A.23: Illustration d'une section d'autoroute non homogène. Cet exemple est utilisé pour présenter les équations pour la prédiction des densités et des débits à l'interface.

Les densités et débits prédits dans l'intervalle $k = t_0 + 1, \dots, t_0 + \Delta$, sont:

$$\begin{aligned}
 \hat{\rho}_1(k+1) &= \hat{\rho}_1(k) + \frac{T}{L_1}(\hat{\varphi}_1(k) - \hat{\varphi}_2(k)) \\
 \hat{\rho}_2(k+1) &= \hat{\rho}_2(k) + \frac{T}{L_2}(\hat{\varphi}_2(k) - \hat{\varphi}_3^-(k)) \\
 \hat{\rho}_3(k+1) &= \hat{\rho}_3(k) + \frac{T}{L_3}(\hat{\varphi}_3^+(k) - \hat{\varphi}_4^-(k)) \\
 \hat{\rho}_4(k+1) &= \hat{\rho}_4(k) + \frac{T}{L_4}(\hat{\varphi}_4^+(k) - \hat{\varphi}_5(k)) \\
 \hat{\varphi}_2 &= \min\{v_1\hat{\rho}_1, w_2(\rho_{m,2} - \hat{\rho}_2)\} \\
 \hat{\varphi}_3^- &= \min\{v_2\hat{\rho}_2, w_3(\rho_{m,3} - \hat{\rho}_3) - \hat{u}\} \\
 \hat{\varphi}_3^+ &= \min\{v_2\hat{\rho}_2 + \hat{u}, w_3(\rho_{m,3} - \hat{\rho}_3)\} \\
 \hat{\varphi}_4^- &= \min\{v_3\hat{\rho}_3, w_4(\rho_{m,4} - \hat{\rho}_4) + \hat{r}\} \\
 \hat{\varphi}_4^+ &= \min\{v_3\hat{\rho}_3 - \hat{r}, w_4(\rho_{m,4} - \hat{\rho}_4)\} \\
 \hat{\varphi}_1 &= \begin{cases} \hat{\varphi}_u, & \text{if } \hat{v}_u \geq v_{\text{lim},1} \\ w_1(\rho_{m,1} - \hat{\rho}_1), & \text{otherwise} \end{cases} \\
 \hat{\varphi}_5 &= \begin{cases} v_4\hat{\rho}_4, & \text{if } \hat{v}_d \geq v_{\text{lim},4} \\ \hat{\varphi}_d, & \text{otherwise} \end{cases},
 \end{aligned} \tag{A.5.13}$$

où $\hat{\rho}(t_0)$ est estimée à partir de l'observateur.

A.5.8 Prédiction du temps de parcours

La prédiction du temps de parcours dans chaque cellule est faite comme suit:

$$\begin{aligned}
 \hat{t}t_i &= \frac{L_i}{\hat{v}_i}, \\
 \hat{v}_i &= \frac{\hat{\varphi}_i + \hat{\varphi}_{i+1}}{2\hat{\rho}_i}.
 \end{aligned} \tag{A.5.14}$$

A.5.9 Résultats expérimentaux: prédiction du temps de parcours.

Le but de cette section est d'évaluer dans quelle mesure l'introduction d'un cadre d'observation de l'état peut améliorer la prédiction du temps de parcours.

L'emplacement pour la validation expérimentale est représentée sur la Fig. A.24. Cette section, d'une longueur de 3,3 km, a été choisie car les congestions les plus fortes sont principalement observées dans cette section.



Figure A.24: Section utilisée pour la validation de l'approche de prédiction du temps de parcours approche orienté modèle.

Pour l'ensemble de données disponibles, le temps de parcours progressif, entre la rampe d'accès 6 et la rampe de sortie 7 (A à B), a été calculé. Fig. A.25 montre ces résultats. On peut noter que le temps passé par un conducteur dans cette section est d'environ 2,4 min en mode fluide (mode libre), tandis que la congestion va jusqu'à 28 min. Soulignons que ce dernier était un cas extrême, car c'était le début des vacances scolaires.

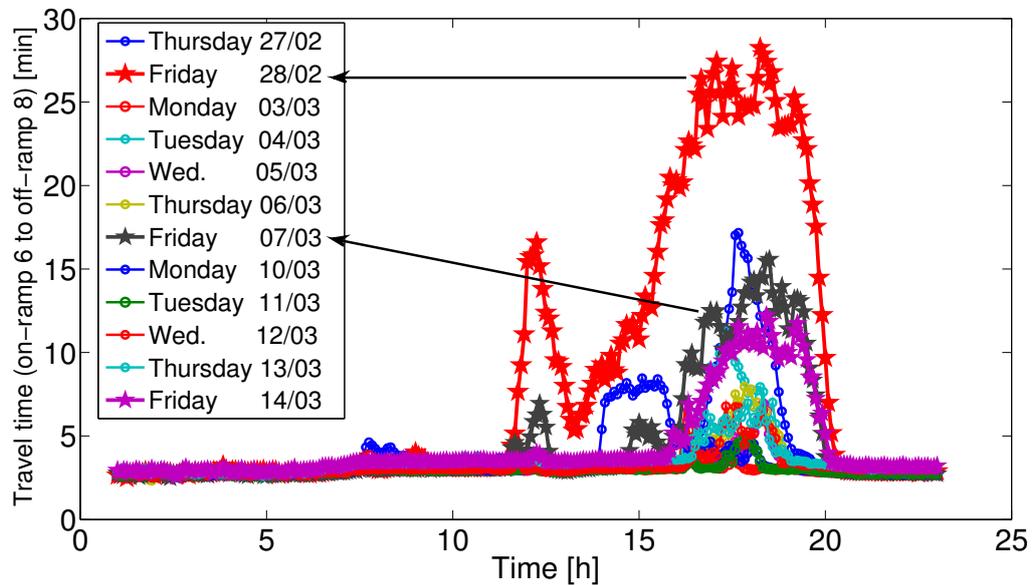


Figure A.25: Temps de parcours entre Gabriel Péri (entrée 2) et SMH Centre (sortie) pour différentes heures de départ. Du 27 Février au 15 Mars de 2014.

Le scénario considéré est un véhicule qui part à 18h30. Les résultats de la Fig. A.26-a et la Fig. A.26-b, Nous mettent dans un contexte où la plupart de la section est déjà congestionnée. Ici, la Fig. A.26-a montre comment le temps parcours prédit suit assez bien le temps de parcours mesuré avec un pourcentage d'erreur moyenne maximale de 30 % . D'autre part, la moyenne historique ne capture pas la congestion. La Fig. A.26-b montre comment le contour de la vitesse prédite est compatible avec les conditions de circulation attendues. La trajectoire mesurée est bien suivie, et les plus grandes différences sont observées entre les kilomètres 9 et 10 dans la deuxième et la quatrième trajectoire.

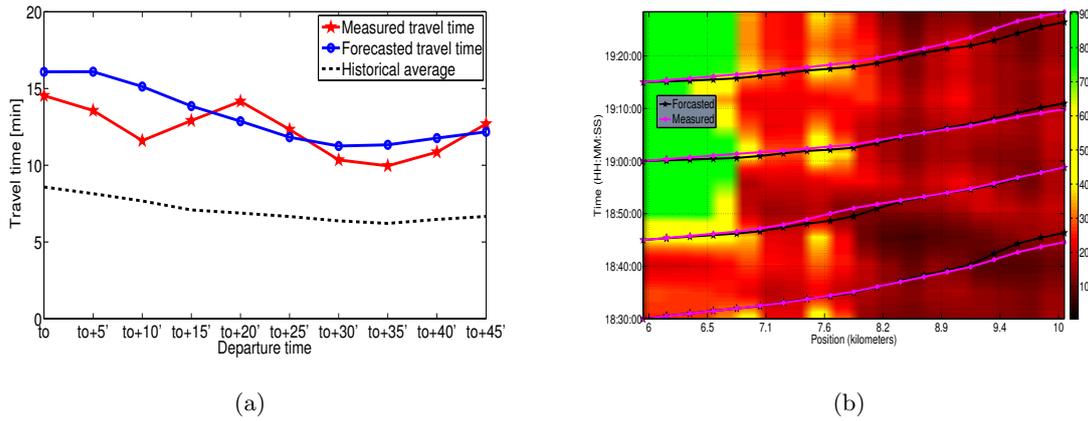


Figure A.26: Résultats de l’approche de prédiction proposée. a) Prédiction à pas multiple à $t_0 = 18:30$. b) trajectoires prédites et mesurées sur contour de vitesse à $t_0 = 18:30$.

A.6 Conclusions et travaux à venir

Le but de ce chapitre est de résumer les contributions présentées dans le manuscrit et d’introduire quelques perspectives de futures recherches pour compléter et améliorer ce travail.

A.6.1 Comparaison entre les deux méthodes de prédiction proposées

Cette section a pour but de comparer numériquement les méthodes de prédiction sous un même scénario. Pour cela, nous allons utiliser le même ensemble de données de la section précédente: du 27 Février au 15 Mars 2014, sans week-end. Dans ce test, nous avons choisi ceux qui ont présenté plus de congestion, les jours avec des flèches dans la Fig. A.27.

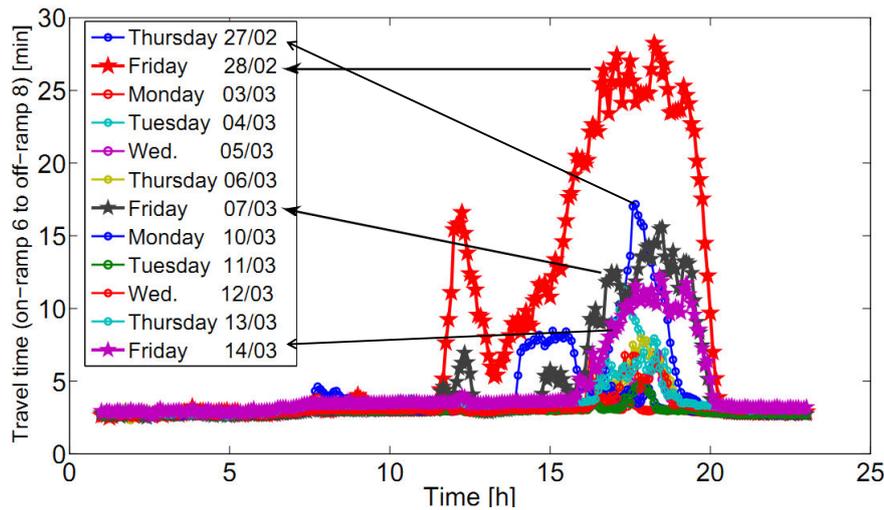


Figure A.27: Temps de parcours estimée entre Gabriel Péri (entrée 2) et Centre SMH (sortie) pour différentes heures de départ. Du 27 Février au 15 Mars 2014. Les jours de pointes de flèches sont les jours qui seront pris en considération pour la comparaison numérique.

Cette étude est réalisée comme suit. La prédiction sera faite dans la fenêtre temporelle de

15h00 à 20h00 à 1 min d'intervalle pour chaque jour de la base de données. À l'horizon spécifié, nous allons calculer l'APE (%) et des CDF seront évalués avec toutes les réalisations.

Plusieurs observations peuvent être tirées des résultats de la Fig. A.28. En premier lieu, on observe une bonne précision pour les deux méthodes. Pour un horizon maximum de 45 minutes, les méthodes ont 90 % de probabilité de prédire avec une erreur inférieure à 21 %.

En particulier, il a été constaté que pour 90 % des réalisations: à $\Delta = 0$ et à $\Delta = 15$ la méthode orientée modèle a présenté des meilleurs résultats (37.1% et 15.3% respectivement). Par contre, pour les horizons $\Delta = 30$ et $\Delta = 45$ les résultats sont inversés, la méthode orientée signal a présenté des meilleurs résultats (dans 5.3% et 17.0% respectivement). Cette évaluation suggère que pour des horizons plus courts de 30 minutes, l'approche orientée modèle peut être plus appropriée que celui orientée signal, à l'opposé des horizons de plus de 30 min. Ceci peut être expliqué en rappelant le modèle utilise une granularité plus fine, elle réussit donc mieux à suivre les conditions de trafic dans la route, notamment à court horizons de prédiction.

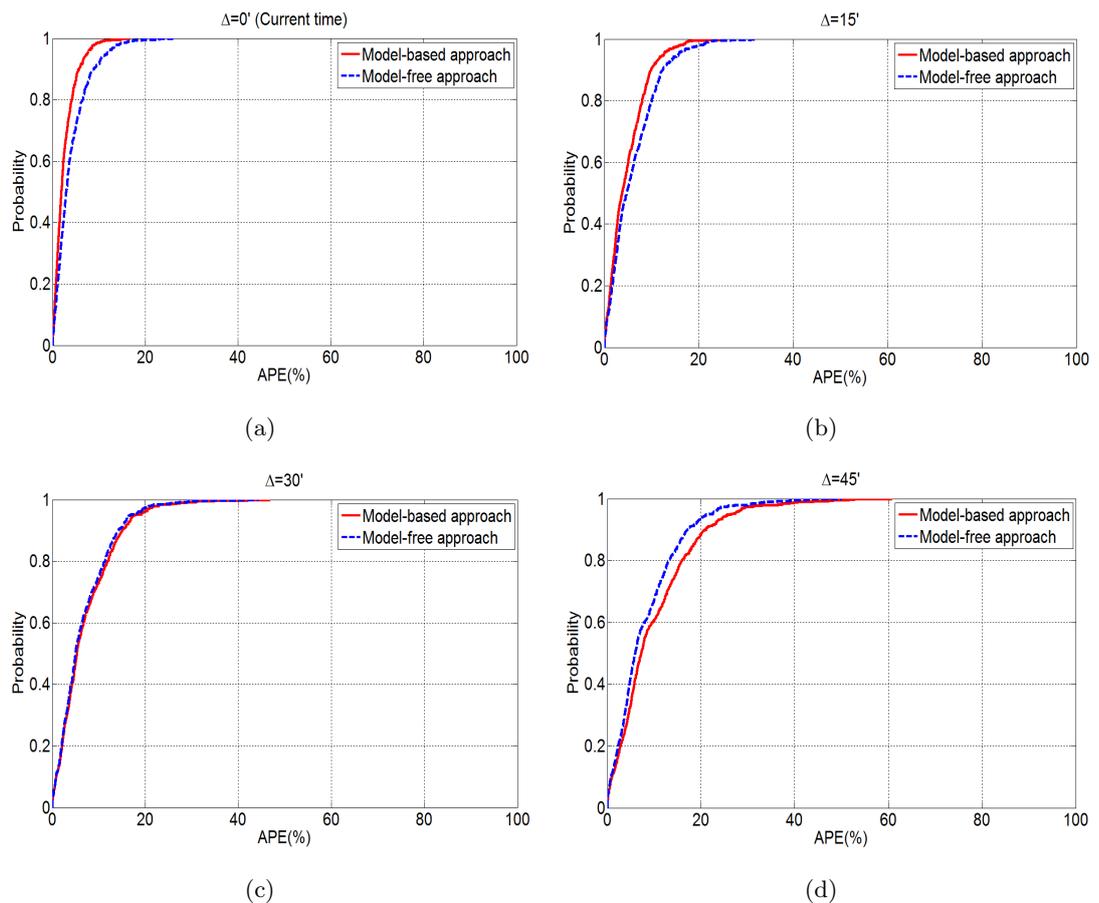


Figure A.28: Fonction de distribution cumulative évaluée avec les APE à différents horizons de prédiction. a) $\Delta = 0$, b) $\Delta = 15'$, c) $\Delta = 30'$, d) $\Delta = 45'$.

A.6.2 Résumé des contributions et conclusions

- **Procédure de pré-traitement de données de trafic:** nous avons travaillé sur une procédure de pré-traitement des données de trafic. Cette procédure a permis l'analyse, le nettoyage et l'imputation des données de trafic brut.
- **Développement d'une nouvelle approche pour la prédiction à court terme et à pas multiples de trafic approche basée sur le filtre de Kalman:** nous avons introduit une nouvelle méthode de prédiction en faisant usage d'une approche de filtre de Kalman adaptatif (AKF). Cette approche transforme le problème de prédiction en un problème de filtrage, en utilisant comme observations du système une combinaison appropriée des données historiques, appelés pseudo-observations. Les statistiques de bruit d'observation ont été calculés en examinant la dispersion dans les données historiques. D'autre part, les statistiques du bruit de processus ont été calculées en utilisant un estimateur sans biais en ligne.
- **Observateur de densité:** il a été développé un observateur déterministe afin de résoudre le problème d'estimation de la densité du trafic, à partir de mesures de débits et de vitesse. Cet observateur a été basé sur le modèle SMM, la linéarisation du Cell Transmission Model (CTM). Ici nous avons montré qu'avec les deux hypothèses: une seule onde de congestion peut exister dans un lien d'autoroute, et qu'il se propage en amont, nous avons pu réduire le nombre de modes. Un modèle contraint a été mis en œuvre. En outre, afin de rendre plus robuste la sélection des modes du système, nous avons proposé d'utiliser des mesures de vitesse aux limites de la section.
- **Prédiction de temps de parcours approche basée sur un modèle de trafic macroscopique, en utilisant les mesures de débits et de vitesses:** nous avons introduit une autre nouvelle approche pour la prédiction du temps de parcours construite sur la base du modèle CTM. Cette approche a été construite en utilisant plusieurs blocs constitutifs. Le premier a utilisé l'observateur de densité pour reconstruire les densités des cellules à l'instant courant, le deuxième a utilisé l'AKF et l'agroupement des données historiques pour prédire les conditions aux limites, les flux d'entrées et de sorties à l'horizon de prédiction. Le troisième, dont les entrées sont les signaux prédites et les valeurs de densité à l'instant courant, a utilisé le modèle CTM pour calculer l'évolution future des densités. Enfin, le dernier bloc a calculé le temps de parcours de l'autoroute à partir des densités prédites.

A.6.3 Travaux en cours et à venir

- **Observateur de densités en n'utilisant que des données aux limites et des rampes:** un effort supplémentaire doit être pris en compte afin de développer l'observateur

de densité sans utiliser des mesures de la voie principale. Ceci ferait un impact évident sur la quantité d'informations que l'approche orientée modèle aurait besoin.

- **Prédiction des intervalles de confiance:** une autre direction de recherche s'occupe de la conception des intervalles de confiance dans les prédictions. Cela permettrait de définir une marge d'erreur entre les résultats réels et prédits.
- **Observateur de densités robuste:** une version plus robuste de l'observateur de densité développée dans cette thèse a été proposée. Cette version prend en compte la dispersion dans la partie congestionnée du diagramme fondamental. Néanmoins, une étude plus approfondie de cette version robuste devrait être envisagée.

Bibliography

- [1] B. Abdulhai, H. Porwal, and W. Recker. Short term freeway traffic flow prediction using genetically-optimized time-delay-based neural networks. *Research Reports . Institute of Transportation Studies, University of California, Berkeley*, 1999.
- [2] H. Abouaissa and V. Iordanova. Algebraic methods for traffic flow densities estimation. *Journal of Cybernetics and Information Technologies, CIT*, 13:5–17, 2013.
- [3] M. Ahmed and A. Cook. Analysis of freeway traffic time-series data by using Box-Jenkins techniques. *Transportation Research Record: Journal of Transportation Research Board*, 722:1–9, 1979.
- [4] H. Al-Deek and C. Chandra. New algorithms for filtering and imputation of real-time and archived dual-loop detector data in i-4 data warehouse. *Transportation Research Record: Journal of Transportation Research Board*, 1867:116–126, 2004.
- [5] A. Alessandri, M. Baglietto, and G. Battistelli. Luenberger observers for switching discrete-time linear systems. In *Proc. of Conference on Decision and Control (CDC) and the European Control Conference (ECC)*, Seville, Spain, December 2005.
- [6] A. Alessandri, M. Baglietto, and G. Battistelli. Design of observers with commutation-dependent gains for linear switching systems. In *Proc. of American Control Conference*, New York City, USA, July 2007.
- [7] A. Alessandri and P. Coletta. *Design of Luenberger Observers for a Class of Hybrid Linear Systems*. Springer Berlin Heidelberg, Rome, Italy, 2001.
- [8] A. Alessandri and P. Coletta. Switching observers for continuous time and discrete-time linear systems. In *Proc. of American Control Conference*, Arlington, USA, June 2001.
- [9] C. Buisson and J.B. Lesort. *Comprendre le trafic routier-Méthodes et calculs*. Centre d'études sur les réseaux, les transports, l'urbanisme et les constructions publiques, 2010.
- [10] M. Castro-Neto, Y. Jeong, M. Jeong, and L. Hana. Online-SVR for short-term traffic flow prediction under typical and atypical traffic conditions. *Journal Expert Systems with Applications: An International Journal archive*, 36(3):6164–6173, April 2009.

- [11] M. Cetin and G. Comert. Short-term traffic flow prediction with regime switching models. *Transportation Research Record*, 1965:23–31, 2006.
- [12] C. Chen and J. Xia. Developing a strategy for imputing missing volume data. *85th Annual Meeting of Transportation Research Board*, 2006.
- [13] M. Chen and S. Chien. Dynamic freeway travel time prediction using probe vehicle data: link-based vs. path-based. *Transportation Research Board 80th Annual Meeting*, 2001.
- [14] S. Chien and C. Kuchipudi. Dynamic travel time prediction with real-time and historical data. *Transportation Research Board 81th Annual Meeting*, 2002.
- [15] R. Chrobok, J. Kaumann, and M. Schreckenberg. Different methods of traffic forecast based on real data. *European Journal of Operational Research*, 155:558–568, 2004.
- [16] S. Clark. Traffic prediction using multivariate non parametric regression. *Journal of Transportation Engineering*, 129(2):161–168, 2003.
- [17] S. Clark, M. Dougherty, and H. Kirby. The use of neural networks and time series models for short-term traffic forecasting: a comparative study. In *Proceedings of the 21st summer Annual Meeting*, 1993.
- [18] B. Coifman. A new methodology for smoothing freeway loop detector data: Introduction to digital filtering. *Transportation Research Record 1554, TRB, National Research Council*, pages 142–152, 1996.
- [19] M. Cremer and M. Papageorgiou. Parameter identification for a traffic flow model. *Automatica*, 16:837–847, 1981.
- [20] C.F. Daganzo. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B*, 28(4):269–287, 1994.
- [21] D. Dailey. Travel time estimates using a series of single loop volume and occupancy measurements. *Presented at the Transportation Research Board 76th Annual Meeting [CD-ROM]*, 1997.
- [22] M. Danech-Pajouh and M. Aron. ATHENA: a method for short-term inter-urban motorway traffic forecasting. *Recherche Transport Securite*, 6:11–16, 1991.
- [23] G. Davis and N. Niham. Adaptive forecasting of freeway traffic congestion. *Transportation Research Record*, 1287:29–33, 1991.
- [24] M. Der Voort, M. Dougherty, and S. Watson. Combining Kohonen maps with ARIMA time series models to forecast traffic flow. *Transportation Research Part C*, 4(5):307–318, 1996.

- [25] G. Dervisoglu, G. Gomes, J. Kwon, R. Horowitz, and P. Varaiya. Automatic calibration of the fundamental diagram and empirical observations on capacity. *Transportation Research Board 88th Annual Meeting*, (09-3159), 2009.
- [26] M. Dougherty and M. Cobbet. Short-term inter-urban traffic forecasts using neural networks. *International Journal of Forecasting*, 13(1):21–31, 1997.
- [27] L. Evans. *Partial Differential Equations*. American Mathematical Society, Providence, 1998.
- [28] D. Gold, S. Turner, B. Gajewski, and S. Spiegelman. Imputing missing values in its data archives for intervals under 5 minutes. *80th Annual Meeting of Transportation Research Board*, November 2011.
- [29] G. Gomes, R. Horowitz, A.A. Kurzhanskiy, J. Kwon, and P. Varaiya. Behavior of the cell transmission model and effectiveness of ramp metering. *Transportation Research, C*, 16(4):485–513, 2008.
- [30] M. Halkidi, Y. Batistakis, and M. Vazirgiannis. On clustering techniques. *Journal of Intelligent Information Systems*, 17:107–145, December 2001.
- [31] F. Hall and B. Persaud. Evaluation of speed estimates made with single-detector data from freeway traffic management systems. *Transportation Research Record 1232*, TRB, National Research Council, 1989.
- [32] A. Hassibi and S. Boyd. Quadratic stabilization and control of piecewise-linear systems. In *Proc. of American Control Conference*, Philadelphia, USA, June 1998.
- [33] S. Haykin. *NN: A comprehensive Foundation*. NY: Macmillan, 1994.
- [34] S. Ishak and C. Alecsandru. Optimizing traffic prediction performance of neural networks under various topological, input and traffic condition. In *Proceedings of the TRB 82nd Annual Meeting*, Washington, DC, USA, 2003.
- [35] M. Johansson and A. Rantzer. Computation of piecewise quadratic Lyapunov functions for hybrid systems. *IEEE Trans. on Automatic Control*, 43:555 – 559, April 1998.
- [36] H. Kirby, M. Dougherty, and S. Watson. Short-term prediction of traffic volume in urban arterials. *ASCE Journal of TransportationEngineering*, 121(3):249–254, 1995.
- [37] A.A. Kurzhanskiy. Set-valued estimation of freeway traffic density. In *Proc. of 12th IFAC Symposium on Control in Transportation Systems*, 2009.
- [38] Florio L. and Mussone L. Neural-network models for classification and forecasting of freeway traffic flow stability. *Control Engineering Practice*, 4:153–164, February 1996.

- [39] W. Lam, Y. Tang, K. Chan, and M. Tam. short-term traffic flow forecast using Hong Kong Annual Traffic Census. *Transportation*, 33:291–310, 2006.
- [40] J. Lebacque, S. Mammam, and H. Haj-Salem. Generic second order traffic flow modeling. In *Proc. of ISTTT (International Symposium on Transportation and Traffic Flow theory)*, London, UK, 2007.
- [41] J. Lebacque, S. Mammam, and H. Haj-Salem. An intersection model based on the gsom model. In *Proc. of the 17th IFAC World Congress*, Seoul, South Korea, Julliet 2008.
- [42] L. Leon, A. Kibangou, and C. Canudas-de Wit. Adaptive Kalman filtering for multi-step ahead traffic flow prediction. In *Proc. of American Control Conference*, Washington, DC, USA, June 2013.
- [43] B. Leva. Travel time forecasting. Master’s thesis, Technical University of Denmark, Kongens Lyngby (Denmark), 2007.
- [44] R.J. LeVeque. *Numerical methods for conservation laws*. Birkhäuser, 1992.
- [45] R.J. LeVeque. The Godunov scheme and what it means for first order traffic flow models. In *Internaional symposium on transportation and traffic theory*, pages 647–677, 1996.
- [46] M. Levin and Y. Tsao. On forecasting freeway occupancies and volumes. *Transportation Research Record: Journal of Transportation Research Board*, 773:47–49, 1980.
- [47] D. Liberzon and A. Morse. Basic problems in stability and design of switched systems. *IEEE Control Systems Magazine*, 19:59 – 70, October 1999.
- [48] M.J. Lighthill and G.B. Whitham. On kinematic waves ii: A theory of traffic flow on long crowded roads. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 229(1178):317–345, 1955.
- [49] H. Lin and P. Antsaklis. Stability and stabilizability of switched linear systems: A short survey of recent results. In *Proc. of IEEE International Symposium on Intelligent Control*, Limassol, Cyprus, June 2005.
- [50] H. Lin and P. Antsaklis. Stability and stabilizability of switched linear systems: A survey of recent results. *IEEE Trans. on Automatic Control*, 54:308 – 322, February 2009.
- [51] W. Lin. A Gaussian maximum likelihood formulation for short-term forecasting of traffic flow. In *IEEE Intelligent Transportation Systems Conference Proceedings*, Oakland, CA, USA, August 2001.
- [52] W. Lin, J. Dahlgram, and H. Huo. Short-term forecasts of freeway traffic volumes and lane occupancies, phase 1-volume IV. Technical report, Washington State Transportation Center (TRAC), University of Washington, Washington, USA, June 1993.

- [53] W. Lin, J. Dahlgram, and H. Huo. An enhancement to speed estimation with single loops. california path working paper. california path working paper, ucb-its-pwp-2003-14. Technical report, California, USA, March 2003.
- [54] W. Lin, Q. Lu, and J. Dalhgren. A dynamic procedure for short-term prediction of traffic conditions. *Transportation Research Record*, 1783:149–157, 2002.
- [55] C. Linveld and R. Thijs. Online travel time estimation using inductive loop. data: the effect of instrumentation peculiarities on travel time estimation quality. In *Proc. of the 6th ITS World Congress*, Toronto, Canada, 1999.
- [56] D. Luenberger. An introduction to observers. *IEEE Transactions on Automatic Control*, 16:596–602, 1971.
- [57] H. Majid, H. Abouaissa, D. Jolly, and G. Morvan. State reconstructor for real-time free-way ramp metering. In *Proc. of the 10th IEEE International Conference on Networking, Sensing and Control, ICNSC*, Paris, France, 2013.
- [58] R. Mehra. Adaptive filtering with unknown prior statistics. In *Proceedings of the joint Automatic Control Conference*, pages 760–769, Colorado, USA, August 1969.
- [59] D. Mignone, G. Ferrari-Trecati, and M. Morari. Stability and stabilization of piecewise affine and hybrid systems: an LMI approach. In *Proc. of Conference on Decision and Control*, Sydney, Australia, December 2000.
- [60] L. Mihaylova and R. Boel. A particle filter for freeway traffic estimation. In *Proc. of Conference on Decision and Control*, volume 2, pages 2106–2111, Bahamas, UK, 2004.
- [61] B. Mikhalkin, H. payne, and L. Isaksen. Estimation of speed from presence detectors. In *Proc. of Highway Research Record*, pages 73–83, Washington DC, USA, 1972.
- [62] I.-C. Morarescu and C. Canudas-de Wit. Highway traffic model-based density estimation. In *Proc. of American Control Conference*, San Francisco, CA, USA, June 2011.
- [63] L. Muñoz, X. Sun, R. Horowitz, and L. Alvarez. Traffic density estimation with the cell transmission model. In *Proc. of the American Control Conference*, pages 3750–3755, 2003.
- [64] K. Myers and B. Tapley. Adaptive sequential estimation with unknown noise statistics. *IEEE Transactions on Automatic Control*, 21(24):520–523, August 1976.
- [65] G. F. Newell. A simplified theory on kinematic waves in highway traffic part i, ii and iii. *Transportation Research*, 27:281–313, 1993.
- [66] I. Okutani and Y. Stephanedes. Dynamic prediction of traffic volume through Kalman filtering theory. *Transportation Research Part B*, 18(1):1–11, February 1984.

- [67] M. Papageorgiou. *Applications of Automatic Control Concepts to Traffic Flow Modeling and Control*. Springer-Verlag New York, Inc., Secaucus, 1983.
- [68] M. Papageorgiou, H. Haj-Salem, A. Tzimitsi, and J. Blosseville. ALINEA: A local feedback control law for on-ramp metering. *Transportation Research Record*, pages 58–64, 1991.
- [69] D. Park and L. Rilett. Forecasting multiple-period freeway link travel times using modular neural networks. *Transportation Research Record*, 1617:163–170, 1998.
- [70] A. Rantzer and M. Johansson. Piecewise linear quadratic optimal control. *IEEE Trans. on Automatic Control*, 45:629–637, April 2000.
- [71] J. Rice. A simple and effective method for predicting travel times on freeways. *IEEE Transactions on Intelligent Transportation System*, 5(3):200–207, September 2004.
- [72] P.I. Richards. Shock waves on the highway. *Operations Research*, 4(1):42–51, 1956.
- [73] A. Sage and G. Husa. Adaptive filtering with unknown prior statistics. In *Proceedings of the joint Automatic Control Conference*, pages 760–769, Colorado, USA, August 1969.
- [74] A. Sage and J. Melsa. *System identification*. Academic press, New York (USA), 1971.
- [75] M. Shawn, L. William, J Robert, and J. Douglas. Travel time data collection handbook. Technical report, Texas, USA, March 1998.
- [76] J.-B. Sheu. A stochastic modeling approach to dynamic prediction of section-wide inter-lane and intra-lane traffic variables using point detector data. *Transportation Research Part A*, 33(2):79–100, February 1999.
- [77] R. Shorten, F. Wirth, O. Mason, K. Wulff, and C. King. Stability criteria for switched and hybrid systems. *The Flagship Journal of the Society for Industrial and Applied Mathematics*, 49:545–592, February 2007.
- [78] B. Smith and M. Demetsky. Traffic flow forecasting: comparison of modelling approaches. *Journal of Transportation Engineering*, 123(4):261–266, 1997.
- [79] B. Smith, W. Scherer, and J. Conklin. Exploring imputation techniques for missing data in transportation management systems. *Transportation Research Record: Journal of Transportation Research Board*, 1836:132–142, 2003.
- [80] B. Smith, B. Williams, and K. Oswald. Parametric and nonparametric traffic volume forecasting. In *Proceedings of the 80th TRB Annual Meeting*, Washington, DC, USA, 2000.

- [81] B. Smith, B. Williams, and K. Oswald. Comparison of parametric and nonparametric models for traffic flow forecasting. *Transportation Research Part C: Emerging Technologies*, 10(4):303–321, 2002.
- [82] A. Stathopoulos and A. Karlaftis. A multivariate state-space approach for urban traffic flow modelling and prediction. *Transportation Research Part C: Emerging Technologies*, 11(2):121–135, April 2003.
- [83] R. Stengel. *Optimal Control and Estimation*. Dover publications, INC, New York (USA), 1994.
- [84] X. Sun, L. Muñoz, and R. Horowitz. Highway traffic state estimation using improved mixture Kalman filters for effective ramp metering control. In *Proc. of Conference on Decision and Control*, pages 6333–6338, 2003.
- [85] X. Sun, L. Muñoz, and R. Horowitz. Mixture kalman filter based highway congestion mode and vehicle density estimator and its application. In *Proc. of the American Control Conference*, pages 2098–2103, Boston, USA, 2004.
- [86] Y. Tang, W. Lam, and P. Ng. Comparison of four modeling techniques short-term AADTT forecasting in Hong Kong. *Journal of Transportation Engineering*, 129(3):271–277, 2003.
- [87] V. Utkin. *Sliding modes in control and optimization*. Springer, Berlin, Germany, 1992.
- [88] E. Vlahogiannia, J. Golias, and M. Karlaftisa. Short term traffic forecasting: overview of objectives and methods. *Transport Reviews*, 24(5):553–557, September 2004.
- [89] Chen W. and M. Saif. Observer design for linear switched control systems. In *Proc. of American Control Conference*, Boston, USA, June 2004.
- [90] Y. Wang and M. Papageorgiou. Real-time freeway traffic state estimation based on extended kalman filter: a general approach. *Transportation Research Part B*, 39(2):141–167, 2005.
- [91] Y. Wang, M. Papageorgiou, P. Coppola, A. Tzimitsi, and A. Nuzzolo. An adaptive freeway traffic state estimator. *Automatica*, 45:10–24, 2009.
- [92] T. Warren. Clustering of time series data—a survey. *Journal of Journal Pattern Recognition*, 38:1857–1874, November 2005.
- [93] B. Williams and L. Hoel. Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results. *ASCE Journal of Transportation Engineering*, 129(6):664–672, 2003.

- [94] B. Williams, K. Prya, and D. Brown. Urban freeway traffic flow prediction. application of seasonal autoregressive integrated moving average and exponential smoothing models. *Transportation Research Record*, 1644:179–188, 1998.
- [95] J. Xia. *Dynamic freeway travel time prediction using single loop detector and incident data*. University of Kentucky Doctoral Dissertations, Kentucky (USA), 2006.
- [96] X. Yuanchang, Z. Yunlong, and Y. Zhirui. Short-term traffic volume forecasting using Kalman filter with discrete wavelet decomposition. *Computer-Aided Civil and Infrastructure Engineering*, 22(5):326–334, July 2007.
- [97] X. Zhang and R. John. Short-term travel time prediction. *Transportation Research Part C*, 11:187–210, 2003.