

Modélisation fonctionnelle d'un profil de vitesse de référence adapté à l'infrastructure

Cindie Andrieu

Directeur de thèse : X. Bressaud (Pr. à l'université Paul Sabatier de Toulouse)

Encadrant IFSTTAR : G. Saint Pierre (CR au LIVIC)

Colloque "Jeunes Probabilistes et Statisticiens",
CIRM-Marseille, 16-20 Avril 2012



Plan

- 1 Introduction
 - Contexte
 - Objectifs de la thèse
- 2 Modélisation fonctionnelle des profils spatiaux de vitesse
 - Définition
 - Propriétés
- 3 Lissage sous contraintes
 - Introduction
 - Lissage sous contrainte de monotonie
 - Application
- 4 Perspectives

Gestion de la vitesse

La connaissance de la vitesse réelle pratiquée sur nos routes est essentielle pour :

- ▶ Situer les points noirs du réseau
- ▶ Améliorer la connaissance des temps de parcours
- ▶ Connaître les effets de modification de l'infrastructure (ajout de dos d'ânes, rond-point, ...)

Construction d'un profil de vitesse de référence

Pour qui ?

Construction d'un profil de vitesse de référence

Pour qui ?

- ▶ Tout le monde !
Développement des ISA (Intelligent Speed Adaptation).

Construction d'un profil de vitesse de référence

Pour qui ?

- ▶ Tout le monde !
Développement des ISA (Intelligent Speed Adaptation).
- ▶ Gestionnaire d'infrastructure.
Impact de la modification de l'infrastructure (ajout de rond-points, de ralentisseurs).

Construction d'un profil de vitesse de référence

Pour qui ?

- ▶ Tout le monde !
Développement des ISA (Intelligent Speed Adaptation).
- ▶ Gestionnaire d'infrastructure.
Impact de la modification de l'infrastructure (ajout de rond-points, de ralentisseurs).
- ▶ Véhicules "intelligents".
Vitesse de consigne.

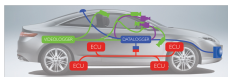
Véhicules traceurs

Mesures continues de la vitesse et de la position tout au long du trajet (capteurs mobiles).

Véhicules traceurs

Mesures continues de la vitesse et de la position tout au long du trajet (capteurs mobiles).

Actuellement : flottes privées (taxis, bus, ...). Non représentatif de l'usage réel du réseau. Instrumentation lourde.



Futur proche : Toute personne équipée d'un smartphone.
Instrumentation légère.



Objectifs de la thèse

- ▶ Utiliser les mesures répétées de vitesses réellement pratiquées issues de **véhicules traceurs**.

Objectifs de la thèse

- ▶ Utiliser les mesures répétées de vitesses réellement pratiquées issues de **véhicules traceurs**.
- ▶ Développer une **méthodologie statistique** permettant d'extraire **divers profils de vitesse de référence**, chacun étant adapté à une situation de conduite (feu rouge ou vert, trafic libre ou contraint, ...)

Objectifs de la thèse

- ▶ Utiliser les mesures répétées de vitesses réellement pratiquées issues de **véhicules traceurs**.
- ▶ Développer une **méthodologie statistique** permettant d'extraire **divers profils de vitesse de référence**, chacun étant adapté à une situation de conduite (feu rouge ou vert, trafic libre ou contraint, ...)

Originalité de notre approche : Traiter les profils de vitesse comme des fonctions et non comme des vecteurs de \mathbb{R}^n .

Objectifs de la thèse

- ▶ Utiliser les mesures répétées de vitesses réellement pratiquées issues de **véhicules traceurs**.
- ▶ Développer une **méthodologie statistique** permettant d'extraire **divers profils de vitesse de référence**, chacun étant adapté à une situation de conduite (feu rouge ou vert, trafic libre ou contraint, ...)

Originalité de notre approche : Traiter les profils de vitesse comme des fonctions et non comme des vecteurs de \mathbb{R}^n .

Remarque : On s'intéresse à la construction d'un profil de référence adapté à l'infrastructure, i.e. à l'étude des **profils spatiaux de vitesse** ("space-speed profiles") correspondant à la vitesse en fonction de la position du véhicule.

Plan

- 1 Introduction
 - Contexte
 - Objectifs de la thèse
- 2 Modélisation fonctionnelle des profils spatiaux de vitesse
 - Définition
 - Propriétés
- 3 Lissage sous contraintes
 - Introduction
 - Lissage sous contrainte de monotonie
 - Application
- 4 Perspectives

Lien entre vitesse, position et temps

En pratique : profil spatial de vitesse = succession de mesures horodatées de position et de vitesse.

Conséquence : Un profil spatial de vitesse peut être manipulé dans les trois espaces suivants :

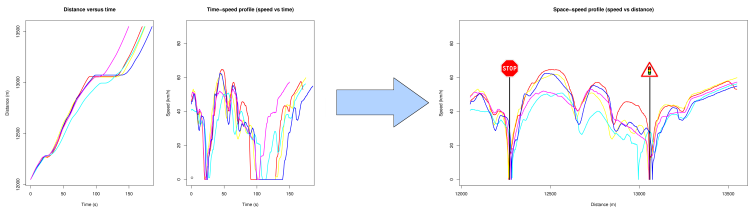


Figure: Lien entre les trois espaces : $[position \times temps]$, $[vitesse \times temps]$ et $[vitesse \times position]$.

Principale difficulté : Assurer la correspondance entre vitesse et position (et implicitement temps) afin de préserver les caractéristiques dynamiques d'un profil de vitesse valide.

Définition de l'espace \mathcal{E}_{SSP}

Définition

Soit $x_f \in \mathbb{R}^+$. Alors l'espace des profils de vitesse spatiaux est défini de la façon suivante :

$\mathcal{E}_{SSP} = \{v_S : [0, x_f] \rightarrow \mathbb{R}^+ \text{ tel que il existe un réel positif } T \text{ et une fonction } \mathcal{C}^2 \text{ croissante } F : [0, T] \rightarrow [0, x_f] \text{ avec } F(0) = 0 \text{ tels que } v_S(x) = F'(F^{-1}(x)), x \in [0, x_f] \text{ où } F^{-1} \text{ est l'inverse généralisée de } F \text{ défini par } F^{-1}(x) = \inf \{t \in [0, T], F(t) = x\}\}$

où $F : [0, T] \rightarrow [0, x_f]$ représente la distance parcourue au cours du temps (espace *position* \times *temps*),

$F' : [0, T] \rightarrow \mathbb{R}^+$ représente la vitesse au cours du temps (espace *vitesse* \times *temps*),

$v_S : [0, x_f] \rightarrow \mathbb{R}^+$ avec $v_S(x) = F'(F^{-1}(x))$ représente le profil spatial de vitesse (espace *vitesse* \times *position*).

Propriétés de l'espace \mathcal{E}_{SSP}

Propriétés

- ▶ Positivité
- ▶ Continuité
- ▶ Sous certaines hypothèses (faibles), non dérivabilité en tout point de $H_0 = \{x \in [0, x_f], v_S(x) = 0\}$.

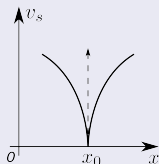


Figure: Point de rebroussement de 1^{ère} espèce.

- ▶ La somme de deux profils spatiaux de vitesse n'est pas toujours un profil spatial de vitesse (cas où l'un s'annule en x_0 mais pas l'autre).

Plan

- 1 Introduction
 - Contexte
 - Objectifs de la thèse
- 2 Modélisation fonctionnelle des profils spatiaux de vitesse
 - Définition
 - Propriétés
- 3 Lissage sous contraintes
 - Introduction
 - Lissage sous contrainte de monotonie
 - Application
- 4 Perspectives

Introduction

Données : mesures bruitées (y_i, x_i) , $i = 1, \dots, n$, de la position et de la vitesse du véhicule issues de capteurs.

Objectif : Estimer au mieux le vrai profil v_S par un estimateur \widehat{v}_S .

Modélisation : Modèle de régression non paramétrique :

$$y_i = m(x_i) + \varepsilon_i, \quad i = 1, \dots, n$$

où les ε_i sont les erreurs non corrélés entre elles, de moyenne 0 et de variance σ^2 et m est la fonction de régression que l'on cherche à estimer.

Contrainte : L'estimateur \widehat{m} de m doit appartenir à l'espace \mathcal{E}_{SSP} des profils spatiaux de vitesse \Rightarrow Non dérivabilité en tout point de $H_0 = \{x \in [0, x_f], \widehat{m}(x) = 0\}$

Choix de l'espace

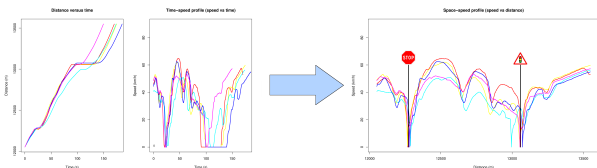


Figure: Lien entre les trois espaces : $[position \times temps]$, $[vitesse \times temps]$ et $[vitesse \times position]$.

Principe : On se place dans l'espace $[position \times temps]$ et l'on cherche à estimer la fonction F représentant la distance parcourue au cours du temps. Modèle de régression non paramétrique :

$$y_i = F(t_i) + \varepsilon_i, \quad i = 1, \dots, n$$

où les y_i sont les "observations" bruitées de la distance parcourue (obtenues avec l'un des deux estimateurs de position), les ε_i sont les erreurs non corrélés entre elles, de moyenne 0 et de variance σ^2 et $F(t)$ est la fonction de régression que l'on cherche à estimer.

Choix de l'espace

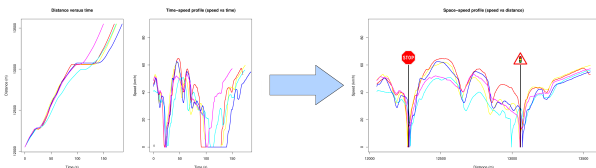


Figure: Lien entre les trois espaces : $[position \times temps]$, $[vitesse \times temps]$ et $[vitesse \times position]$.

Principe : On se place dans l'espace $[position \times temps]$ et l'on cherche à estimer la fonction F représentant la distance parcourue au cours du temps. Modèle de régression non paramétrique :

$$y_i = F(t_i) + \varepsilon_i, \quad i = 1, \dots, n$$

Avantage : La seule contrainte sur l'estimateur \hat{F} est une contrainte de monotonie.

Choix de l'espace

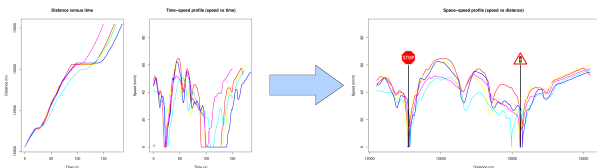


Figure: Lien entre les trois espaces : $[position \times temps]$, $[vitesse \times temps]$ et $[vitesse \times position]$.

Principe : On se place dans l'espace $[position \times temps]$ et l'on cherche à estimer la fonction F représentant la distance parcourue au cours du temps. Modèle de régression non paramétrique :

$$y_i = F(t_i) + \varepsilon_i, \quad i = 1, \dots, n$$

On en déduit une estimation du profil spatial de vitesse par :

$$\hat{v}_S(x) = \hat{F}'(\hat{F}^{-1}(x))$$

Revue des différentes méthodes

- ▶ Régression isotonique ;
 - *Brunk (1955)* : estimateur constant par morceaux ;
 - *Friedman & Tibshirani (1984)* : lissage de l'estimateur isotone par moyenne mobile ;
 - *Mukerjee (1988), Mammen (1991)* : lissage à noyau puis isotonisation ;

Revue des différentes méthodes

- ▶ Régression isotonique ;
 - *Brunk (1955)* : estimateur constant par morceaux ;
 - *Friedman & Tibshirani (1984)* : lissage de l'estimateur isotone par moyenne mobile ;
 - *Mukerjee (1988), Mammen (1991)* : lissage à noyau puis isotonisation ;
- ▶ Splines monotones ;
 - *Utretas (1980)* : caractérisation dans le cas où F est \mathcal{C}^1 ;
 - *Ramsay (1988), Kelly & Rice (1990)* : lissage sous contraintes aux points d'observations ($\widehat{F}(t_{i+1}) - \widehat{F}(t_i) \geq 0$) ;
 - *Ramsay (1998)* : définition d'une fonction monotone par $D^2f = wf$ et estimation sans contraintes de w ;

Revue des différentes méthodes

- ▶ Régression isotonique ;
 - *Brunk (1955)* : estimateur constant par morceaux ;
 - *Friedman & Tibshirani (1984)* : lissage de l'estimateur isotone par moyenne mobile ;
 - *Mukerjee (1988), Mammen (1991)* : lissage à noyau puis isotonisation ;
- ▶ Splines monotones ;
 - *Utretas (1980)* : caractérisation dans le cas où F est \mathcal{C}^1 ;
 - *Ramsay (1988), Kelly & Rice (1990)* : lissage sous contraintes aux points d'observations ($\widehat{F}(t_{i+1}) - \widehat{F}(t_i) \geq 0$) ;
 - *Ramsay (1998)* : définition d'une fonction monotone par $D^2f = wf$ et estimation sans contraintes de w ;
- ▶ Projection d'un estimateur non contraint dans un espace de fonctions monotones.
Mammen & Thomas-Agnan (1999), Mammen, Marron, Turlach & Wand (2001), Dette, Neumeyer & Pilz (2006).

Revue des différentes méthodes

- ▶ Régression isotonique ;
 - *Brunk (1955)* : estimateur constant par morceaux ;
 - *Friedman & Tibshirani (1984)* : lissage de l'estimateur isotone par moyenne mobile ;
 - *Mukerjee (1988), Mammen (1991)* : lissage à noyau puis isotonisation ;
- ▶ Splines monotones ;
 - *Utretas (1980)* : caractérisation dans le cas où F est \mathcal{C}^1 ;
 - *Ramsay (1988), Kelly & Rice (1990)* : lissage sous contraintes aux points d'observations ($\widehat{F}(t_{i+1}) - \widehat{F}(t_i) \geq 0$) ;
 - *Ramsay (1998)* : définition d'une fonction monotone par $D^2f = wf$ et estimation sans contraintes de w ;
- ▶ Projection d'un estimateur non contraint dans un espace de fonctions monotones.

Avantages de ces méthodes de projection : se ramène à deux étapes :

- 1 Lissage non contraint ;
- 2 Monotonisation.

Revue des différentes méthodes

- ▶ Régression isotonique ;
 - *Brunk (1955)* : estimateur constant par morceaux ;
 - *Friedman & Tibshirani (1984)* : lissage de l'estimateur isotone par moyenne mobile ;
 - *Mukerjee (1988), Mammen (1991)* : lissage à noyau puis isotonisation ;
- ▶ Splines monotones ;
 - *Utretas (1980)* : caractérisation dans le cas où F est \mathcal{C}^1 ;
 - *Ramsay (1988), Kelly & Rice (1990)* : lissage sous contraintes aux points d'observations ($\widehat{F}(t_{i+1}) - \widehat{F}(t_i) \geq 0$) ;
 - *Ramsay (1998)* : définition d'une fonction monotone par $D^2f = wf$ et estimation sans contraintes de w ;
- ▶ Projection d'un estimateur non contraint dans un espace de fonctions monotones.
Inconvénients de ces méthodes de projection :
 - Etape de projection difficile à mettre en oeuvre numériquement ;
 - La plupart des méthodes sont spécifiques à un estimateur non contraint ;

Choix de la méthode : splines homéomorphes (Bigot & Gadat (2010))

Principe : Adapter les outils développés dans le contexte de la déformation d'images au problème de régression monotone.

Choix de la méthode : splines homéomorphes (Bigot & Gadat (2010))

Principe : Adapter les outils développés dans le contexte de la déformation d'images au problème de régression monotone.

Déformation d'images :

Soient $\Omega \subset \mathbb{R}^2$ et deux images $I_1, I_2 : \Omega \rightarrow \mathbb{R}$.

On cherche $\phi : \Omega \rightarrow \Omega$ telle que :

$$I_1(x) = I_2(\phi(x)) \quad \text{et } \phi \text{ est bijective.}$$

Choix de la méthode : splines homéomorphes (Bigot & Gadat (2010))

Principe : Adapter les outils développés dans le contexte de la déformation d'images au problème de régression monotone.

Déformation d'images :

Soient $\Omega \subset \mathbb{R}^2$ et deux images $I_1, I_2 : \Omega \rightarrow \mathbb{R}$.

On cherche $\phi : \Omega \rightarrow \Omega$ telle que :

$$I_1(x) = I_2(\phi(x)) \quad \text{et } \phi \text{ est bijective.}$$

Trouvé, Younes et al. : Méthodologie pour construire des fonctions lisses et bijectives (difféomorphismes) dans plusieurs dimensions.

Choix de la méthode : splines homéomorphes (Bigot & Gadat (2010))

Principe : Adapter les outils développés dans le contexte de la déformation d'images au problème de régression monotone.

Déformation d'images :

Soient $\Omega \subset \mathbb{R}^2$ et deux images $I_1, I_2 : \Omega \rightarrow \mathbb{R}$.

On cherche $\phi : \Omega \rightarrow \Omega$ telle que :

$$I_1(x) = I_2(\phi(x)) \quad \text{et } \phi \text{ est bijective.}$$

Trouvé, Younes et al. : Méthodologie pour construire des fonctions lisses et bijectives (difféomorphismes) dans plusieurs dimensions.

Remarque : En 1D, une fonction strictement monotone et continue est un homéomorphisme (\mathcal{C}^0 difféomorphisme).

Principe de la méthode

- Soit f une fonction strictement monotone. On écrit f comme la solution au temps $t = 1$ d'une ODE avec comme condition initiale l'identité, i.e.

$$\phi_0(x) = x \quad \phi_1(x) = f(x) \quad \frac{\partial \phi_t}{\partial t} = v_t^f(\phi_t(x))$$

ou encore

$$f(x) = x + \int_0^1 v_t(\phi_t(x)) dt$$

Principe de la méthode

- Soit f une fonction strictement monotone. On écrit f comme la solution au temps $t = 1$ d'une ODE avec comme condition initiale l'identité, i.e.

$$\phi_0(x) = x \quad \phi_1(x) = f(x) \quad \frac{\partial \phi_t}{\partial t} = v_t^f(\phi_t(x))$$

ou encore

$$f(x) = x + \int_0^1 v_t(\phi_t(x)) dt$$

- On estime le champ de vecteurs $(v_t^f, t \in [0, 1])$ à partir des observations y_i en utilisant un estimateur non contraint \hat{f}^{nc} de f : utilisation de la méthode classique de lissage splines.

Principe de la méthode

- Soit f une fonction strictement monotone. On écrit f comme la solution au temps $t = 1$ d'une ODE avec comme condition initiale l'identité, i.e.

$$\phi_0(x) = x \quad \phi_1(x) = f(x) \quad \frac{\partial \phi_t}{\partial t} = v_t^f(\phi_t(x))$$

ou encore

$$f(x) = x + \int_0^1 v_t(\phi_t(x)) dt$$

- On estime le champ de vecteurs $(v_t^f, t \in [0, 1])$ à partir des observations y_i en utilisant un estimateur non contraint \hat{f}^{nc} de f : utilisation de la méthode classique de lissage splines.
- On calcule la solution au temps $t = 1$ de l'ODE :

$$\hat{f}^c(x) = \phi_1(x) \quad \text{avec} \quad \frac{\partial \phi_t}{\partial t} = \hat{v}_t^f(\phi_t(x)) \quad \text{et} \quad \phi_0(x) = x$$

Principe de la méthode

- Soit f une fonction strictement monotone. On écrit f comme la solution au temps $t = 1$ d'une ODE avec comme condition initiale l'identité, i.e.

$$\phi_0(x) = x \quad \phi_1(x) = f(x) \quad \frac{\partial \phi_t}{\partial t} = v_t^f(\phi_t(x))$$

ou encore

$$f(x) = x + \int_0^1 v_t(\phi_t(x)) dt$$

- On estime le champ de vecteurs $(v_t^f, t \in [0, 1])$ à partir des observations y_i en utilisant un estimateur non contraint \hat{f}^{nc} de f : utilisation de la méthode classique de lissage splines.
- On calcule la solution au temps $t = 1$ de l'ODE :

$$\hat{f}^c(x) = \phi_1(x) \quad \text{avec} \quad \frac{\partial \phi_t}{\partial t} = \hat{v}_t^f(\phi_t(x)) \quad \text{et} \quad \phi_0(x) = x$$

Remarque : \hat{f}^c hérite des propriétés de convergence de \hat{f}^{nc} .

Application : estimateur par splines de lissage

Rappel : Pour $t_1 < \dots < t_n$ n points d'un intervalle $[a, b]$, on appelle spline d'ordre m (m étant un entier ≥ 1) ayant pour noeuds simples les points t_1, \dots, t_n toute fonction f de $[a, b]$ dans \mathbb{R} telle que :

- ▶ f est continûment dérivable jusqu'à l'ordre $m - 2$ (si $m \geq 2$),
- ▶ sur chaque intervalle inter-noeuds $(a, t_1), \dots, (t_i, t_{i+1}), \dots, (t_n, b)$, f coïncide avec un polynôme de degré $m - 1$.

Application : estimateur par splines de lissage

Principe : Méthode résultant d'un problème d'optimisation.

Le problème suivant :

$$\min_{F \in \mathcal{H}^m([a,b])} \sum_{i=1}^n \{(y_i - F(t_i))^2\} + \lambda \int (F^{(m)}(t))^2 dt.$$

admet une unique solution qui est une spline naturelle d'ordre $2m$ (i.e. dont les dérivées d'ordre $2m - 2$ sont nulles aux frontières de $[a, b]$) ayant pour noeuds les points d'observations t_1, \dots, t_n dans $[a, b]$.

Cette solution peut s'écrire sous la forme :

$$\hat{F}(t) = \sum_{j=1}^n \hat{\beta}_j B_j \quad (1)$$

où les B_j forment une base de splines naturelles (en général, on prend la base des B-splines).

Ainsi trouver \hat{F} revient à déterminer les coefficients $\hat{\beta} = (\hat{\beta}_1, \dots, \hat{\beta}_n)$.

Application : estimateur par splines de lissage

Etape 1 : Lissage non contraint

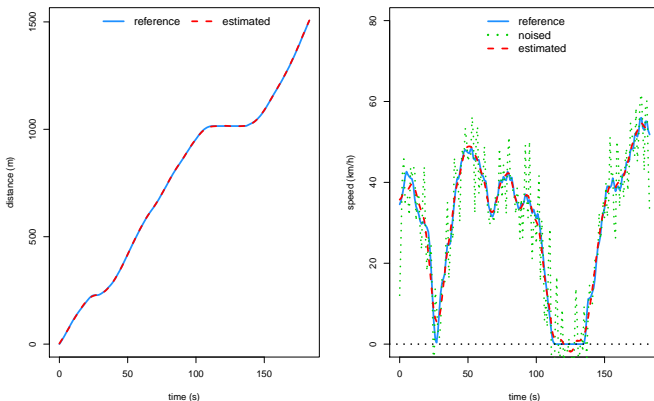


Figure: Estimateur par splines de lissage non contraint. Paramètres : ordre $m = 6$; $\lambda = 17.8$ (GCV).

Application : estimateur par splines de lissage

Etape 2 : Monotonisation

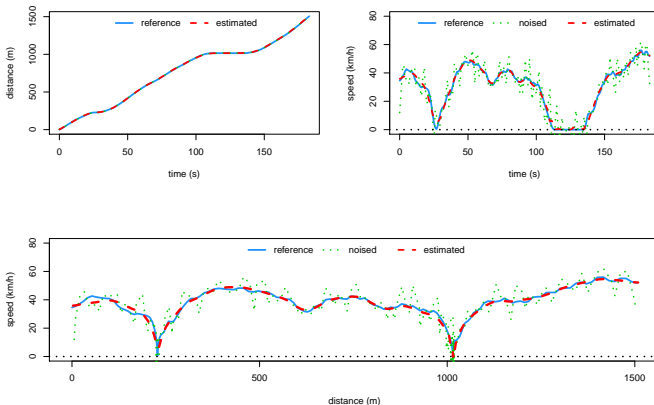


Figure: Estimateur par splines de lissage monotone. $RMSE(F(t)) = 2.11$; $RMSE(F'(t)) = 1.81$.

Suite des travaux

- ▶ Lissage de plusieurs profils : problème de la gestion des arrêts.
- ▶ Détermination de méthodes de classification fonctionnelle non supervisée afin de discriminer différentes situations de conduite (ex : feu rouge, feu vert).
- ▶ Etude de l'impact de certains éléments de l'infrastructure (rond-point, ralentisseur, ...) sur le profil de vitesse de référence.

Merci de votre attention