

MODÉLISATION FONCTIONNELLE ET LISSAGE SOUS CONTRAINTES D'UN PROFIL SPATIAL DE VITESSE

Cindie Andrieu ¹, Guillaume Saint Pierre ¹ et Xavier Bressaud ²

¹ *IFSTTAR, IM, LIVIC, 14, route de la Minière, 78000 Versailles, France.*

E-mail : cindie.andrieu@ifsttar.fr et guillaume.saintpierre@ifsttar.fr

² *Université Paul Sabatier, Institut de Mathématiques de Toulouse, F-31062 Toulouse Cedex 9, France. E-mail : bressaud@math.univ-toulouse.fr*

Résumé. L'objet de cette étude est de proposer une modélisation fonctionnelle et une méthode d'estimation d'un profil spatial de vitesse (vitesse en fonction de la position du véhicule). Les mesures de vitesse et de position issues de capteurs du véhicule étant bruitées, on se ramène à un modèle de régression non paramétrique et on utilise une méthode de lissage sous contraintes utilisant les fonctions splines afin d'obtenir une estimation du vrai profil de vitesse.

Mots-clés. Profil spatial de vitesse, données fonctionnelles, régression non paramétrique, splines de lissage, lissage sous contrainte, monotonie.

Abstract. The purpose of this study is to propose a functional modelisation and an estimation method of a space-speed profile (i.e. speed depending on the position of the vehicle). Due to uncertainties in the measurements of speed and position from sensors, the problem is equivalent to a nonparametric regression model and we use a smoothing method under constraints using spline functions to obtain an estimate of the real speed profile.

Keywords. Space-speed profile, functional data, nonparametric regression, smoothing spline, constrained smoothing, monotonicity.

1 Introduction

La vitesse est un paramètre déterminant du risque routier. Ainsi, la vitesse pratiquée est caractéristique de l'usage réel du réseau routier (Ericsson (2000) [5]), et sa connaissance permet de détecter les points noirs de circulation, d'améliorer la prédiction des temps de parcours, ou encore d'évaluer les effets d'une modification de l'infrastructure (ajout de ralentisseurs ou de ronds-points). En parallèle, la généralisation des aides à la conduite embarquées, en particulier celles destinées à agir sur la vitesse comme le régulateur et le limiteur de vitesse, manuel ou intelligent (Ehrlich (2009) [4]), conduit à s'interroger sur la consigne de vitesse la plus pertinente à adopter, en tenant compte au mieux des spécificités de l'infrastructure. Les limitations de vitesse ne constituent qu'une limite et

non pas une consigne, et sont souvent inadaptées au contexte. De même, les limitations en vigueur ne tiennent pas compte de la météorologie, ni de la visibilité. La production d'une consigne de vitesse pertinente selon l'infrastructure devient donc un enjeu majeur du trafic routier, et de nombreux travaux s'intéressent aux modifications de cette consigne selon le contexte de conduite (trafic, météo, visibilité (ex. Gallen (2010) [7])).

L'information des vitesses réellement pratiquées devient accessible avec la généralisation de "véhicules traceurs" parcourant le réseau routier. Ces véhicules équipés d'un GPS peuvent être vus comme des capteurs mobiles explorant le réseau routier en continu et sont capables de transmettre leur position et leur vitesse. La majorité des véhicules traceurs actuels sont des flottes privées (taxis par exemple), mais la généralisation des smartphones équipés de GPS contribue à faire croître le nombre de "traceurs" susceptibles de renvoyer leur information de position en temps réel. Nous proposons dans ce papier les premiers éléments d'une méthodologie statistique permettant d'utiliser la connaissance des vitesses pratiquées le long d'un axe par un ensemble de véhicules "traceurs", pour en déduire un ou plusieurs profils de vitesse de référence.

2 Modélisation fonctionnelle des profils spatiaux de vitesse

Un profil de vitesse réel est constitué par une succession de mesures horodatées de position et de vitesse et peut donc être manipulé dans l'espace vitesse \times temps, vitesse \times position, ou position \times temps comme l'illustre la figure 1. La collecte de ces informations conduit à l'obtention de mesures bruitées, dont il est difficile d'assurer la cohérence. Un premier travail (non détaillé dans ce papier) a consisté à améliorer la qualité des mesures brutes en construisant un estimateur sans biais de la position du véhicule. Cet estimateur qui utilise le GPS et l'odomètre comme deux sources différentes d'information de position, permet d'obtenir une estimation de la position à environ 3m près.

Le second travail qui fait l'objet de cette communication consiste à proposer un lissage d'un profil de vitesse préservant la cohérence physique entre la vitesse et la position. L'originalité de notre approche consiste non plus à traiter les profils de vitesse comme des vecteurs de \mathbb{R}^n , mais comme des fonctions de l'espace vitesse \times position. En effet, l'analyse de données fonctionnelles (Ramsay et Silverman (2005) [11]) s'est considérablement développée ces dernières années, et constitue le fondement théorique de notre modélisation. Cela nécessite d'explicitier mathématiquement l'espace des profils de vitesse spatiaux (i.e. l'espace vitesse \times position). Nous proposons la définition suivante :

Definition 1. Soit $x_f \in \mathbb{R}^+$. Alors l'espace des profils de vitesse spatiaux est défini de la façon suivante :

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

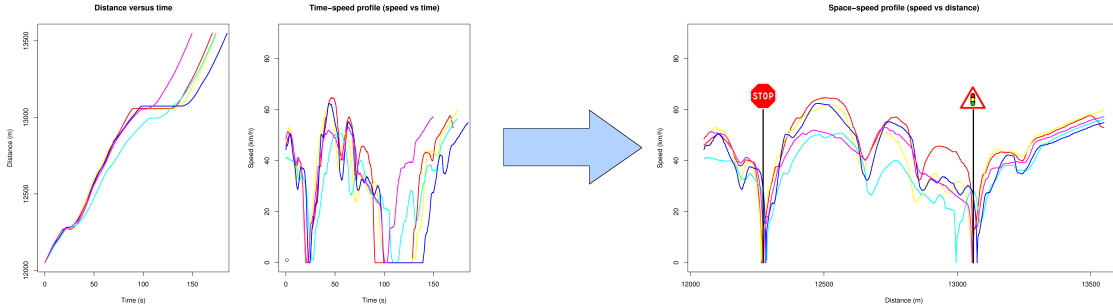


FIGURE 1 – Lien entre les trois espaces : $[position \times temps, vitesse \times temps]$ et $[vitesse \times position]$.

Les mesures issues des véhicules traceurs étant bruitées, l'enjeu est de trouver comment obtenir la fonction $v_S \in \mathcal{E}_{SSP}$ qui les représentent au mieux par utilisation d'une méthode de lissage appropriée. Cependant nous avons étudié certaines propriétés de l'espace \mathcal{E}_{SSP} et nous avons démontré que les fonctions appartenant à cet espace n'étaient pas dérivables aux points pour lesquelles elles s'annulent (correspondant aux arrêts du véhicule). Cette contrainte étant difficile à prendre en compte dans le lissage, nous avons choisi de nous placer dans un premier temps dans l'espace $position \times temps$ et d'estimer la fonction F représentant la distance parcourue au cours du temps en effectuant un lissage sous la seule contrainte de monotonie. On peut ensuite en déduire aisément un profil spatial de vitesse estimé \hat{v}_S par la transformation $\hat{v}_S(x) = \hat{F}'(\hat{F}^{-1}(x))$ donnée à la définition 1. Enfin notons également que le fait d'effectuer le lissage dans l'espace $position \times temps$ permet de préserver la cohérence entre les trois espaces illustrés à la figure 1.

3 Lissage sous contrainte de monotonie

On cherche ici à estimer la fonction F représentant la distance parcourue en fonction du temps à partir des données bruitées de position. On est alors ramené au modèle de régression non paramétrique suivant :

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

où les y_i sont les "observations" bruitées de la distance parcourue, les ε_i sont les erreurs aléatoires i.i.d. de moyenne 0 et de variance σ^2 , et où $F(t)$ est la fonction de régression que l'on cherche à estimer, plus couramment appelée fonction de lissage. Les méthodes de lissage dites locales comme l'estimateur à noyau ou l'estimation par polynômes locaux sont celles habituellement utilisées dans les problématiques d'estimation de profils de vitesse ou d'accélération (Jun et al. (2006) [8] et Bratt et Ericsson (1999) [3]). Cependant, afin

de garder le caractère fonctionnelle de nos données, nous avons choisi d'utiliser les splines de lissage. Cette méthode consiste à estimer F par minimisation du critère des moindres carrés pénalisés suivant dans l'ensemble des fonctions appartenant à l'espace de Sobolev $W^m[a, b]$:

$$\sum_{i=1}^n \{(y_i - F(t_i))^2\} + \lambda \int (F^{(m)}(t))^2 dt \quad (2)$$

où m est fixé et λ est le paramètre de lissage déterminant le compromis entre la fidélité aux données et la régularité de la solution. Un résultat connu est que l'unique fonction minimisant ce critère est une spline naturelle d'ordre $2m$ ayant pour noeuds les points d'observations t_1, \dots, t_n (démonstration dans Eubank (1999) [6]).

Cependant la fonction F que l'on cherche à estimer représentant la distance parcourue du véhicule au cours du temps, notre estimateur \hat{F} doit être une fonction croissante. Diverses méthodes de lissage sous contrainte de monotonie ont été traitées dans la littérature : régression isotonique (Barlow et al. (1972) [1]), splines monotones (Ramsay (1988) [10]), projection d'un estimateur non contraint dans un sous-espace de fonctions monotones (Mammen et al. (2001) [9]). Notre approche se situe dans cette dernière catégorie de méthodes qui présente l'avantage de s'adapter à tout type d'estimateur non contraint et qui se ramène en fait à deux étapes :

1. effectuer un lissage sans contraintes,
2. monotoniser l'estimateur obtenu à l'étape 1.

Cependant ce type de méthode présente deux inconvénients majeurs : la plupart de ces estimateurs monotones apparaissent moins lisses que l'estimateur sans contraintes initial, et l'étape de projection est difficile à mettre en oeuvre numériquement. Afin de palier ces deux difficultés, nous avons utilisé une méthode développée par Bigot et Gadat (2010) [2] et fondée sur des techniques développées dans le contexte de la déformation d'images pour la construction de difféomorphismes en deux et trois dimensions. En effet, si on se ramène en 1D, une fonction continue et strictement monotone est un homéomorphisme (\mathcal{C}^0 difféomorphisme). Cette méthode consiste à définir toute fonction monotone comme la solution d'une équation différentielle ordinaire (ODE) ayant pour conditions initiales l'identité et régie par un champ de vecteurs approprié dépendant du temps. Ce champ de vecteurs est estimé en utilisant les techniques classiques d'estimation par splines de lissage. L'estimateur monotone ainsi obtenu hérite des propriétés de convergence de l'estimateur non contraint utilisé.

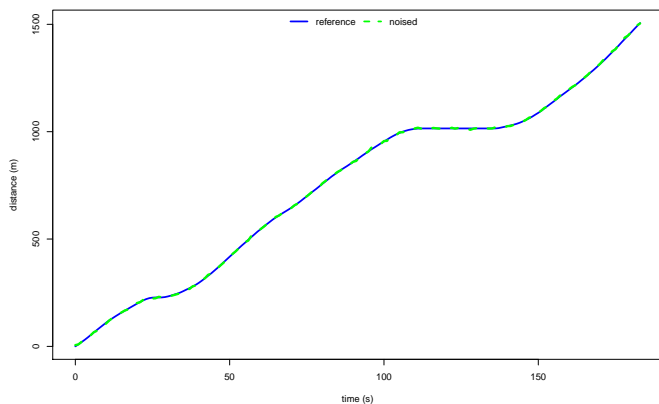
4 Application et discussion

On utilise des mesures réelles de la distance parcourue d'un véhicule sur un trajet de 1500m comprenant un stop (avec un arrêt court du véhicule) et un feu rouge (avec un arrêt long du véhicule) et avec une fréquence d'échantillonnage de 1Hz (soit 1 mesure/sec).

Afin d'évaluer les performances de la méthode de lissage décrite à la section précédente, on considère ces mesures comme étant la "vraie" distance parcourue par le véhicule. On introduit alors un bruit gaussien de variance 3^2 afin de simuler la distance estimée obtenue à partir de notre estimateur de position. Ces données de référence (dans l'espace $position \times temps$) sont illustrées figure 2 (a). On effectue alors les deux étapes décrites précédemment avec les paramètres suivants :

1. un lissage par spline d'ordre 6 (afin de pouvoir également étudier le profil d'accélération si besoin) et en choisissant la valeur du paramètre de lissage λ à l'aide du critère GCV ($\lambda = 19$) (Eubank (1999) [6]). Cette étape a été mise en oeuvre en utilisant le "package fda" du logiciel R et en prenant une base de B-splines.
2. une "monotonisation" de l'estimateur obtenu à l'étape 1 en choisissant un noyau gaussien, un λ très faible (10^{-6}) et un nombre de pas de temps de 300 pour la résolution de l'ODE.

(a) Distance parcourue en fonction du temps



(b) Profil de vitesse spatial après lissage monotone

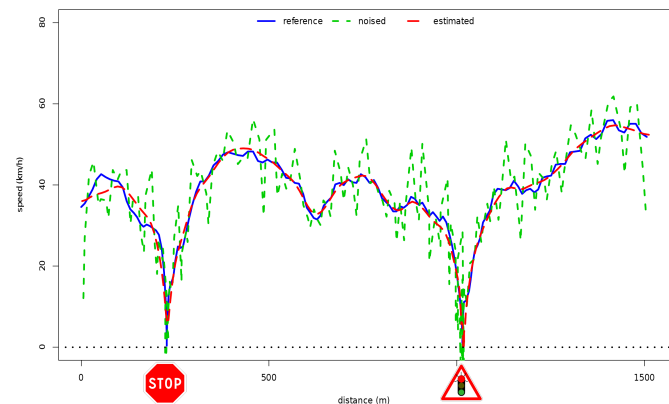


FIGURE 2 – Lissage monotone d'un profil de vitesse

La figure 2 (b) montre le profil de vitesse estimé (en rouge) obtenu après passage de l'espace $position \times temps$ dans l'espace $vitesse \times position$ ainsi que le vrai profil (en bleu) et le profil bruité (en pointillés verts). On observe que notre estimateur reproduit bien la forme générale du "vrai" profil de vitesse. Notons également que l'arrêt au feu est bien identifié (vitesse nulle de l'estimateur) mais que l'on observe uniquement un fort ralentissement au niveau du stop (vitesse estimée de 6km/h).

On peut donc en conclure que notre estimateur est bien adapté pour estimer un profil spatial de vitesse à partir de données bruitées. De plus, le lissage étant réalisé dans l'espace $distance \times position$, on dispose également d'un bon estimateur de la distance parcourue en fonction du temps (la racine carrée de l'erreur quadratique moyenne (RMSE) est de 3.17m) et par dérivation du profil temporel de vitesse (RMSE=1.80km/h). Une difficulté

reste la détection des arrêts courts ($< 5s$) à partir de données bruitées. Une solution envisagée est de déterminer un seuil en dessous duquel la vitesse est considérée comme nulle puis d'imposer une vitesse nulle aux instants d'arrêt en rajoutant un vecteur de poids dans le critère de lissage par fonctions splines.

Enfin, la suite des travaux sera consacrée à l'introduction des données de vitesse mesurées par le GPS (mesurées par effet Doppler et donc indépendantes de la position du GPS) dans notre modèle de régression afin d'améliorer les performances de notre estimateur, ainsi qu'à la mise en oeuvre de notre estimateur sur des trajets plus longs (de l'ordre de 15km par exemple).

Bibliographie

- [1] R.E. BARLOW, D.J. BARTHOLOMEW, JM BREMNER et HD BRUNK : *Statistical inference under order restrictions : the theory and application of isotonic regression*. J. Wiley, 1972.
- [2] J. BIGOT et S. GADAT : Smoothing under diffeomorphic constraints with homeomorphic splines. *SIAM Journal on Numerical Analysis*, 48(1):224–243, 2010.
- [3] H. BRATT et E. ERICSSON : Estimating speed and acceleration profiles from measured data. In *Conference Proceedings of the 8th International Symposium "Transport and Air Pollution"*, Graz, Austria, 1999.
- [4] J. EHRLICH : Towards isa deployment in europe : state of the art, main obstacles and initiatives to go forward. In *Proceedings from the 2009 Intelligent Speed Adaptation Conference*, Sidney, 2009.
- [5] E. ERICSSON : Variability in urban driving patterns. *Transportation Research Part D*, Volume 5:337–354, September 2000.
- [6] R.L. EUBANK : *Nonparametric regression and spline smoothing*. Statistics, textbooks and monographs. Marcel Dekker, 1999.
- [7] R. GALLEN, N. HAUTIERE et S. GLASER : Advisory speed for intelligent speed adaptation in adverse conditions. In *IEEE Intelligent Vehicles Symposium (IV'10)*, San Diego, California, USA, pages 107–114, 2010.
- [8] J. JUN, R. GUENSLER et J.H. OGLE : Smoothing methods to minimize impact of global positioning system random error on travel distance, speed, and acceleration profile estimates. *Transportation Research Record : Journal of the Transportation Research Board*, 1972:141–150, 2006.
- [9] E. MAMMEN, JS MARRON, BA TURLACH et MP WAND : A general projection framework for constrained smoothing. *Statistical Science*, pages 232–248, 2001.
- [10] J.O. RAMSAY : Monotone regression splines in action. *Statistical Science*, pages 425–441, 1988.
- [11] J.O. RAMSAY et B.W. SILVERMAN : *Functional Data Analysis, Second Edition*. 2005.