



**HAL**  
open science

## Espaces de courbes pour l'analyse de données aeronautiques

Cindie Andrieu, Baptiste Gregorutti, Florence Nicol, Stéphane Puechmorel

► **To cite this version:**

Cindie Andrieu, Baptiste Gregorutti, Florence Nicol, Stéphane Puechmorel. Espaces de courbes pour l'analyse de données aeronautiques . JDS 2016, 48e Journées de Statistiques de la Société française de statistiques, May 2016, Montpellier France. hal-01330957

**HAL Id: hal-01330957**

**<https://enac.hal.science/hal-01330957>**

Submitted on 13 Jun 2016

**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.

# ESPACES DE COURBES POUR L'ANALYSE DE DONNÉES AÉRONAUTIQUES

Cindie Andrieu<sup>1</sup> & Baptiste Gregorutti<sup>1</sup> & Florence Nicol<sup>2</sup> & Stéphane Puechmorel<sup>2</sup>

<sup>1</sup> *Safety Line*  
130 rue de Lourmel, 75015 Paris  
*prenom.nom@safety-line.fr*

<sup>2</sup> *MAIAA, École Nationale de l'Aviation Civile*  
7 Avenue Édouard Belin, 31000 Toulouse  
*prenom.nom@enac.fr*

**Résumé.** La sécurité des vols est une préoccupation constante dans le domaine de la gestion du trafic aérien. L'accroissement prévisible du nombre de vols et l'introduction progressive de nouveaux concepts de contrôle demandent une détection précoce des situations potentiellement dangereuses. Par ailleurs, de grandes masses de données sont collectées par les différents acteurs du trafic aérien, mais restent peu exploitées. En particulier, les données des radars de surveillance des aéroports peuvent être utilisées pour évaluer l'état des pistes et ainsi anticiper des phénomènes de dérapage. L'étude proposée ici vise à fournir des outils méthodologiques issus de la géométrie différentielle pour la détection d'un mauvais état de la piste d'atterrissage à partir de trajectoires mesurées par les radars. Faisant suite à un travail déjà réalisé par la société Safety Line, des outils nouveaux de calcul de similarité entre courbes sont introduits.

**Mots-clés.** Similarité entre courbes, variétés différentielles, détection de trajectoires atypiques, sécurité du transport aérien.

**Abstract.** Flight safety is a concern within the frame of air traffic management. The predicted increase in the number of flights and the gradual shift to new control procedures require an early detection of potentially hazardous situations. On the other hand, huge amounts of data are collected by the air transportation stakeholders, most of them remaining mainly unexploited. In particular, radar data can be used to measure runway conditions and to detect skid phenomenon. The study presented here aims at providing methodological tools from differential geometry for the detection of bad runway conditions based on radar data. Extending a work by the SafetyLine company, new tools for computing the similarity between trajectories are introduced.

**Keywords.** Curves similarity, curves manifolds, atypical trajectories mining, air transportations safety.

# 1 Introduction

Les grands projets fédérateurs SESAR en Europe et Nextgen aux Etats-Unis ont été initiés afin de répondre à une augmentation conséquente du trafic aérien mondial jusqu'à l'horizon 2030. Le rapport d'Airbus Industries (Perspectives 2030) table sur un accroissement du trafic de passagers dans les vingt prochaines années (4,2 % de 1986 à 2010 - 4,8 % de 2010 à 2030), ce qui aboutirait à un doublement du nombre de passagers en 2030 par rapport à 2010.

De ce fait, la gestion et l'analyse des flux de trafic sur une plateforme aéroportuaire devient primordiale dans l'objectif d'en améliorer la capacité et la sécurité. En effet, afin d'éviter la saturation d'un aéroport, il est nécessaire de déterminer les zones de congestion et d'agir en conséquence. Les trajectoires des objets mobiles évoluant sur un aéroport peuvent être reconstruites à partir des données issues des radars de surveillance des aéroports. Les traces radar permettent ainsi de localiser précisément tout mobile sur l'aéroport, aussi bien sur les taxiways que sur les pistes. Les données radar contiennent les coordonnées GPS et la vitesse de chaque objet mobile se déplaçant sur la plateforme. D'autres caractéristiques telles que le type d'avion peuvent également être enregistrées.

L'état des pistes sur un aéroport est une donnée critique à connaître. Lorsque celle-ci est supposée perdre de son adhérence, la procédure actuellement appliquée consiste à inspecter la piste ce qui implique un arrêt du trafic d'environ 20 minutes. Cette interruption provoque alors des perturbations très importantes (mises en attente, retards, etc.).

La société Safety Line propose ainsi à ses clients d'utiliser les traces radar pour optimiser le déclenchement des interventions sur les pistes et les voies de circulation (dénouement, traitement des pistes, etc.) et ainsi de limiter le temps d'interruption du trafic. L'approche retenue consiste à détecter les trajectoires des avions à l'atterrissage qui s'écartent notablement d'une trajectoire nominale (ex : trajectoire moyenne, centre de la piste). Lors du dégagement de piste, un tel écart peut être caractéristique d'un phénomène de dérapage de l'avion dû à une mauvaise adhérence de la piste (cf. Figure 1). Les indicateurs ainsi définis permettent d'estimer en temps réel les performances au freinage à chaque atterrissage.

Si les outils existants permettent de tenir compte de la nature fonctionnelle des données radar (voir par exemple [6]), la prise en compte de l'information à la fois temporelle et spatiale est un problème difficile. Il est donc nécessaire de proposer des outils plus sophistiqués pour mesurer l'écart entre différentes trajectoires sans perte d'information spatiale. L'approche retenue dans ce travail vise à appliquer des techniques issues de la géométrie différentielle à la construction d'une mesure de similarité entre trajectoires. Dans le cas particulier de l'étude des atterrissages, il est de plus nécessaire de prendre en compte le profil de décélération en tant que variable supplémentaire, ceci afin de confirmer que les écarts à la courbe nominale sont dûs à l'état de la piste.

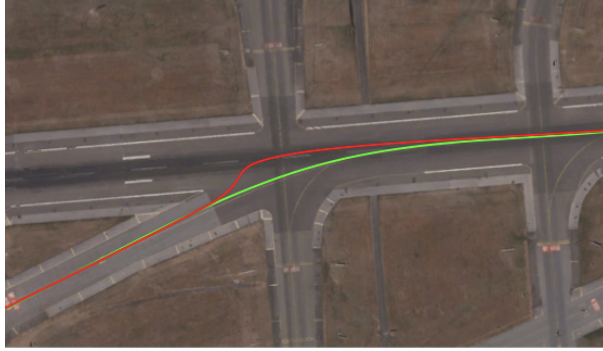


FIGURE 1 – Exemple de trajectoire avec un écart latéral (en rouge) par rapport à la trajectoire nominale (en vert).

## 2 Proximité dans les espaces de courbes

Une question fondamentale lorsque l'on travaille avec des données représentées sous la forme de courbes est la définition d'une notion de similarité qui capture les traits les plus caractéristiques des objets traités.

Nous proposons ici d'utiliser le cadre des variétés de courbes pour définir une métrique entre trajectoires comme une alternative aux distances  $L^p$  largement utilisées en pratique, en particulier dans des problèmes de recalage [6]. Dans ce cadre de travail, il devient possible de prendre en compte directement des caractéristiques géométriques jugées comme étant discriminantes pour l'application visée.

Afin de ne conserver dans l'étude que la forme des trajectoires et non la loi des vitesses, on choisit comme espace de représentation :

$$\mathcal{E} = \mathbf{Imm}([0, 1], \mathbb{R}^N) / \mathbf{Diff}^+([0, 1]),$$

où  $\mathbf{Imm}([0, 1], \mathbb{R}^N)$  est l'ensemble des immersions de l'intervalle  $[0, 1]$  dans  $\mathbb{R}^N$ , dont les points sont des trajectoires lisses dont la vitesse ne s'annule jamais et  $\mathbf{Diff}^+([0, 1])$  est le groupe des difféomorphismes croissants de  $[0, 1]$  dans lui-même. Ce dernier intervient comme quotient dans la définition de l'espace  $\mathcal{E}$  et permet de s'affranchir de la loi des vitesses. Si  $\gamma$  est une trajectoire de  $\mathbf{Imm}([0, 1], \mathbb{R}^N)$  et si  $\phi$  est un élément quelconque de  $\mathbf{Diff}^+([0, 1])$ , alors  $\gamma$  et  $\gamma \circ \phi$  seront associées au même point de  $\mathcal{E}$ . L'étude présente ne considère que le premier cas, avec  $N = 2$  qui est la dimension naturelle pour des trajectoires d'atterrissage.

On peut montrer [5] que  $\mathcal{E}$  possède une structure de variété lisse, de dimension infinie. En effet, si  $\gamma$  est un élément de  $\mathbf{Imm}([0, 1], \mathbb{R}^N)$ , il en sera de même pour  $\gamma + \epsilon$  si  $\epsilon$  est une application lisse de  $[0, 1]$  dans  $\mathbb{R}^N$  avec  $\sup_{t \in [0, 1]} \|\epsilon(t)\|$  suffisamment petit. On peut en déduire facilement l'espace tangent qui est formé de couples  $(\gamma, u)$  où  $\gamma$  est élément de  $\mathbf{Imm}([0, 1], \mathbb{R}^N)$  et  $u$  est une application de  $C^\infty([0, 1], \mathbb{R}^N)$  qui s'interprète comme

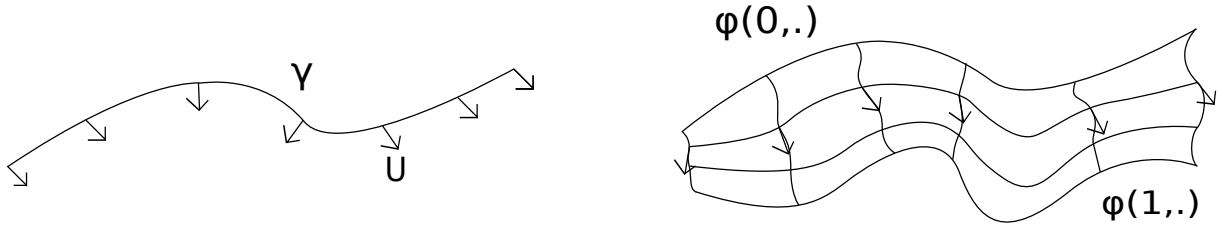


FIGURE 2 – Vecteur tangent et chemin entre courbes

un déplacement infinitésimal le long de  $\gamma$ . La construction passe directement à l'espace quotient  $\mathcal{E}$ .

Dans cet espace, chaque trajectoire est vue comme un point. Un chemin entre deux courbes  $\gamma_1, \gamma_2$  de  $\mathcal{E}$  peut se représenter par une application continue  $\phi: [0, 1] \times [0, 1] \rightarrow \mathbb{R}^N$  telle que  $\phi(0, \bullet) = \gamma_1(\bullet), \phi(1, \bullet) = \gamma_2(\bullet)$  avec  $D_t\phi = \frac{\partial\phi(s,t)}{\partial t}$  non nulle sur tout le domaine. On impose ici une régularité  $C^\infty$  à  $\phi$ . Dans ce cadre, on remarquera que si  $s \in [0, 1]$  est fixé, le champ de vecteurs  $D_s\phi$  définit un vecteur tangent à la courbe  $\phi(s, \bullet)$ .

Les métriques riemanniennes applicables à  $\mathcal{E}$  ont été étudiées en détail par plusieurs auteurs [1, 2, 4]. En particulier, [4] introduit la métrique locale en un point  $\gamma$  de  $\mathcal{E}$  :

$$g_\gamma(u, v) = \int_0^1 \langle u(\gamma(t)), v(\gamma(t)) \rangle (1 + \kappa_\gamma^2(t)) \|\gamma'(t)\| dt, \quad (1)$$

où  $(\gamma, u), (\gamma, v)$  est un couple de vecteurs tangents et  $\kappa_\gamma(t)$  est un terme de courbure de  $\gamma$  en  $t$ , nécessaire pour éviter une dégénérescence de la métrique. Pour un vecteur tangent  $(\gamma, u)$ ,  $g_\gamma(u, u)$  mesure l'énergie de déformation associée au déplacement  $u$ . Rappelons que l'objectif initial est de mesurer l'écart entre une courbe  $\gamma$  et une courbe de référence que l'on note  $\gamma_{\text{ref}}$ . Sur  $\mathcal{E}$ , cela revient à déterminer un chemin  $\phi: [0, 1] \times [0, 1] \rightarrow \mathbb{R}^N$  tel que  $\phi(0, \bullet) = \gamma(\bullet), \phi(1, \bullet) = \gamma_{\text{ref}}(\bullet)$  et dont l'énergie de déformation  $E(\phi)$  soit minimale, celle-ci étant calculée par :

$$E(\phi) = \int_0^1 \int_0^1 \langle D_s\phi(s, t), D_s\phi(s, t) \rangle \|D_t\phi(s, t)\| (1 + \kappa_{\phi(s,t)}^2(t)) dt ds.$$

On remarquera qu'il s'agit de la métrique (1) appliquée au vecteur tangent  $D_s\phi$  et intégrée le long du chemin  $\phi$ . Le terme  $\|D_t\phi\|$  garantit l'indépendance vis à vis du représentant dans  $\mathbf{Imm}([0, 1], \mathbb{R}^N)$ . La longueur du chemin optimal  $\phi_0$  est :

$$\ell(\phi_0) = \int_0^1 \int_0^1 \|D_s\phi_0(s, t)\| \|D_t\phi_0(s, t)\| (1 + \kappa_{\phi_0(s,t)}^2(t)) dt ds.$$

On peut montrer que si  $\phi_0$  minimise  $E$  alors il minimise également  $\ell$ .  $\ell(\phi_0)$  donne donc la distance minimale, ou géodésique, entre les deux courbes. On la notera  $d(\gamma, \gamma_{\text{ref}})$ .

En pratique,  $\phi$  est décrite par ses valeurs  $\phi_{ij}$  sur une grille régulière de  $[0, 1] \times [0, 1]$ , les dérivées étant approchées par différences finies. Le problème de détermination du

chemin géodésique  $\phi$  se ramène à un programme d’optimisation non linéaire de variables de décision les  $\phi_{ij}$ . Un algorithme de type LM-BFGS [3] permet une résolution efficace pour des programmes de grande taille. Dans l’étude présentée, les grilles retenues sont de taille  $10 \times 10$  à  $20 \times 20$ , valeurs suffisantes pour obtenir une estimation cohérente de la distance géodésique.

L’utilisation du terme de courbure induit une non-linéarité assez difficile à traiter numériquement. Il est préférable de supprimer ce terme en ayant recours à une métrique légèrement différente, qui fait intervenir des dérivées d’ordre supérieur :

$$g_\gamma(u, v) = \int_0^1 \langle Lu(\gamma(t)), Lv(\gamma(t)) \rangle \|\gamma'(t)\| dt, \quad (2)$$

où  $L$  est un opérateur différentiel elliptique d’ordre 2 au moins. Un choix approprié de  $L$  permet de prendre en compte dans l’application visée une dynamique avion simplifiée. Les exemples qui seront présentés ont été obtenus avec la métrique (2).

### 3 Application à la détection d’un écart latéral lors d’un atterrissage

Dans cette Section, nous décrivons la façon dont la métrique définie précédemment est utilisée pour mesurer des dégradations éventuelles de pistes d’atterrissage. Un indicateur de “glissance” est ainsi défini à partir de la distance “géométrique” proposée dans la Section précédente. Pour cela, nous observons un échantillon de  $n$  trajectoires  $X_1, \dots, X_n$  et supposons qu’elles sont des variables aléatoires à valeurs dans l’espace  $L^2([0, 1], \mathbb{R}^2)$ . Pour appliquer les résultats précédents, il est nécessaire d’avoir l’appartenance à  $\mathbf{Imm}([0, 1], \mathbb{R}^2)$ , qui est strictement inclus dans l’espace précédent, mais qui sort du cadre classique en statistique fonctionnelle. Dans la présente étude, une approche simplifiée a été retenue : on utilise une approximation régulière des trajectoires à l’aide d’une décomposition sur une base polynomiale locale, ce qui permet de remplacer les dérivées par des différences finies et ne nécessite que des valeurs échantillonnées, comme indiqué plus haut.

Une première étape consiste à construire une trajectoire de référence notée  $X_{\text{ref}}$  (en vert sur la Figure 1). Par exemple, il est possible de définir celle-ci par l’axe de la piste ou bien par la courbe médiane. En complément, nous définissons les variables aléatoires réelles  $Z_1, \dots, Z_n$  indépendantes et identiquement distribuées par

$$Z_i = d(X_i, X_{\text{ref}}),$$

où  $d$  représente la distance définie plus haut. Ces variables correspondent aux écarts à la courbe de référence. L’indicateur de glissance est basé sur l’estimation de la distribution

des écarts. Pour cela, soit  $F$ , la fonction de répartition des écarts et  $F_n$  la fonction de répartition empirique. Ainsi, pour une nouvelle trajectoire  $X_{n+1}$ ,

$$F_n(Z_{n+1}) = F_n(d(X_{n+1}, X_{\text{ref}}))$$

mesure sa déviation à la courbe de référence. L'application de la fonction de répartition aux données radar, nous fournit alors un indicateur naturel de l'état des pistes. Plus cet indice est proche de 1, plus l'état de la piste est estimé dégradé.

## Références

- [1] M. Bauer, M. Bruveris, S. Marsland, and P. W. Michor. Constructing reparametrization invariant metrics on spaces of plane curves. *ArXiv e-prints*, July 2012.
- [2] M. Bauer, M. Bruveris, and P. W. Michor. Overview of the geometries of shape spaces and diffeomorphism groups. *J. Math. Imaging Vis.*, 50(1-2) :60–97, September 2014.
- [3] Dong C. Liu and Jorge Nocedal. On the limited memory bfgs method for large scale optimization. *Mathematical Programming*, 45(1) :503–528.
- [4] Peter W Michor and David Mumford. Riemannian geometries on spaces of plane curves. *J. Eur. Math. Soc. (JEMS)*, 8 :1–48, 2006.
- [5] Peter W Michor and David Mumford. An overview of the riemannian metrics on spaces of curves using the hamiltonian approach. *Applied and Computational Harmonic Analysis*, 23(1) :74–113, 2007.
- [6] J. Ramsay and B.W. Silverman. *Functional Data Analysis*. Springer Series in Statistics. Springer New York, 2006.