



A Multivariate Functional Data Analysis of Aircraft Trajectories

Remi Perrichon, Xavier Gendre, Thierry Klein

► To cite this version:

Remi Perrichon, Xavier Gendre, Thierry Klein. A Multivariate Functional Data Analysis of Aircraft Trajectories. 2023. hal-03667323v2

HAL Id: hal-03667323

<https://enac.hal.science/hal-03667323v2>

Preprint submitted on 26 Jun 2023

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.

A Multivariate Functional Data Analysis of Aircraft Trajectories

Rémi Perrichon^{1*}, Xavier Gendre^{2,3} and Thierry Klein^{1,2}

^{1*}École Nationale de l'Aviation Civile (ENAC), Université de Toulouse, 7 avenue Edouard Belin, Toulouse, 31400, France.

²Institut de Mathématiques de Toulouse (IMT), Université de Toulouse, 1 R.3, Université Paul Sabatier, 118 route de Narbonne, Toulouse, 31400, France.

³Pathway.com, Paris, France.

*Corresponding author(s). E-mail(s): remi.perrichon@enac.fr;

Contributing authors: xavier.gendre@math.univ-toulouse.fr;
thierry01.klein@enac.fr;

Abstract

While advanced methods for functional data analysis have recently been developed in the literature, applications to aircraft trajectories have remained scarce, despite operational relevance. One reason is the practical difficulties affiliated with the multivariate nature of trajectories and associated physical constraints. Indeed, an aircraft trajectory usually involves three dimensions in space (longitude, latitude, altitude) but also weather values (say wind speed and direction), each dimension having its specificities. To name a few, smoothing altitude values requires to ensure both non-negativity and boundary constraints. Wind directions have support on the unit circle. Additional to constrained smoothing challenges, phase variations are to be taken into account as flights are never of the same duration. To tackle these issues, two smoothing methods respectively based on constrained splines and asymmetric kernels are implemented on real data. For each approach, two strategies to handle the circular nature of wind directions are compared. Registration is performed. A joint pointwise test is proposed to demonstrate that delayed flights have experienced less favorable wind conditions.

Keywords: Functional Data Analysis, Circular Data, Trajectory, Aircraft

1 Introduction

Roughly speaking, delay is the time lapse that occurs when a planned event happens after the planned time. In aviation, delays are easily measurable before departure, during taxi-out (the aircraft begins moving forward until it reaches the takeoff position), en-route and during taxi-in. In order to provide some means of international comparison, a standard set of delay definitions was introduced by the International Air Transport Association (IATA) Airport Services Committee. Regarding high-level groupings, one may name airline-related delays, airport-related delays, en-route delays and weather delays.

As they induce a bad customer perception, direct financial costs, a lack of efficiency and environmental issues, delays are subject to a quarterly report drafted by the European Organisation for the Safety of Air Navigation, commonly known as Eurocontrol. Crucially in Europe, the Central Office of Delay Analysis (CODA) is in charge of collecting operational data and drafting reports.

The transportation literature has mainly revolved around some major topics in analyzing delays:

- **Statistical modeling of delays at the airport level.** [Mueller and Chatterji \(2002\)](#) analyze delay characteristics for ten major airports in the United States using departure and arrival data. Focusing on London's Heathrow airport, [Pejovic et al \(2009\)](#) study weather-related delays.
- **Delay prediction.** [Pérez-Rodríguez et al \(2017\)](#) use an asymmetric logit probability model to estimate and predict the daily probabilities of delay in aircraft arrivals. [Yu et al \(2019\)](#) adopt a deep learning approach using data from Beijing International Airport. [Carvalho et al \(2020\)](#) present a literature review of data science techniques used in flight delay prediction.
- **Impact of Ground Delay Programs (GDP).** The effect of weather conditions on the characteristics of GDP events was investigated by [Wang and Kulkarni \(2011\)](#), and more recently by [Liu et al \(2019\)](#).
- **Delay propagation.** [Wang et al \(2022\)](#) have recently proposed an algorithm to estimate statistically significant time lags between airport delays from noisy, aggregate operational data.

Few studies are focusing on en-route delays on the scale of aircraft trajectories. One reason is the irrelevance of the usual statistical frameworks to model an aircraft trajectory. Indeed, techniques from multivariate statistics suffer from the high correlation in time that exists between two consecutive points of a trajectory. More worryingly, when observation times are more numerous than the number of trajectories, the so-called *curse of dimensionality* happens. The usual time series approach is not ideal either as trajectories are multidimensional time series with both irregular sampling and different durations.

An adapted and well-known framework to tackle the aforementioned issues is the one of Functional Data Analysis (FDA), popularised by [Ramsay and](#)

Silverman (2005) and Ferraty and Vieu (2006). FDA is an active field of statistics as highlighted by the development of inference procedures by Horváth and Kokoszka (2012), or more recently, by the blooming of the curve registration geometric framework of Srivastava and Klassen (2016). A review is proposed by Wang et al (2016), focusing on FDA concepts.

FDA had an early interest in applications as shown by Ramsay and Silverman (2002) and Valderrama (2007). The framework has been extensively used in many fields such as chemometrics, e-commerce, econometrics, management science and medicine, as recently highlighted by Aneiros et al (2019a) and Aneiros et al (2019b). Ullah and Finch (2013) pinpoint that applications of FDA in biomedicine are especially numerous.

The promotion of FDA to study aircraft trajectories was early made by Puechmorel and Delahaye (2007). Yet, the statistical analysis of aircraft trajectories has been very limited. Currently, the statistical literature focusing on aircraft trajectories revolves around Functional Principal Component Analysis (FPCA) carried out by Nicol (2017) and is applied to the detection of atypical energy behaviours by Jarry et al (2020).

In a recent contribution, Dai and Müller (2018), have studied trajectories of 969 commercial flights from Hong Kong to London (long-haul flights). They specifically take into account the directional nature of the longitude and the latitude. Spherical Principal Component Analysis (SFPCA) is performed.

An aircraft trajectory is multivariate statistical object as it involves three dimensions in space (longitude, latitude, altitude). Sometimes, weather values are also available (say wind speed and direction).

The ultimate goal of this work is to answer the following question from a statistical point of view:

Given two groups of trajectories, is it possible to state that, on average, the group with higher en-route delays has experienced less favorable wind conditions?

In response, a test is proposed in Section 5. Unfortunately, raw trajectory data must be preprocessed before performing any inference. These preprocessing steps are particularly important and sometimes subtle. They ensure that conclusions drawn from the test are not fallacious.

Two preprocessing steps are extensively developed.

The first one is the smoothing step that solves the irregular sampling problem. This smoothing step must account for specific physical constraints. To name a few:

- Altitude values are non-negative and must be null at takeoff and landing.
- Wind directions have support on the unit circle. It requires specific smoothing techniques.

If constraints are ignored, the mean trajectories of the two groups will not be correct and variances will be artificially inflated.

The second one is the registration step that accounts for the fact that trajectories are of different durations. If this registration step is poorly done,

the test statistic will basically compare weather values for flight phases that are very different (say, comparing the en-route wind values with landing wind values).

Note that taking the functional nature of trajectories is becoming essential with the generalization of Automatic Dependent Surveillance-Broadcast (ADS-B), soon mandated for all aircraft in Europe and the United States. As put in [Sun \(2021\)](#), ADS-B is a surveillance technology designed to allow aircraft to broadcast their flight state periodically without the need for interrogation. The position of an equipped aircraft is known almost every second.

The paper is organized as follows. Section 2 describes the two data sources that have been used as well as the method by which weather and raw trajectories are matched. Upon request, both data sources are in free access for researchers. Section 3 presents how the multivariate FDA framework can be used to model trajectories. Section 4 details the registration strategy. Section 5 presents the testing procedures and results.

2 Empirical background

As an empirical background, Subsection 2.1 introduces the data source for aircraft trajectories. The wind data source is presented in Subsection 2.2. The procedure to associate each point of a flight to a wind value is detailed in Subsection 2.3.

2.1 R&D data from Eurocontrol (trajectories)

Eurocontrol is an international organisation working to achieve safe and seamless air traffic management across Europe. Since 2020, Eurocontrol has given access to a R&D data archive containing more than six years of data, that is to say to more than 18 million flights as of January 2023. The data are collected from all commercial flights operating in and over Europe. To be more specific, Eurocontrol receives flight plans for all Instrument Flight Rules (IFR) flights. These flight plans are activated and updated based on live data from air navigation service providers. Data are available for 4 months each year: March, June, September and December. About 2 to 3 million flights are thus available each year. For each flight, files include the last-filed flight plan, the actual route, the airspace and the route network. Because the airspace and the route network are not of interest here, only two data subsets are used in this work.

- **Flight metadata.** Metadata include (but are not restricted to) a numeric identifier for each flight, the International Civil Aviation Organization (ICAO) airport code for the departure airport of the flight and associated spatial coordinates, the ICAO airport code for the destination airport of the flight and associated spatial coordinates, filed and actual off-block times and the filed and actual arrival times.
- **Actual flight points.** This data set includes a numeric identifier for each flight (allowing a matching with flight metadata), the time (UTC) at which

the point was actually crossed and corresponding altitude, longitude and latitude.

Raw files for the 6 years (2015-2020) amount to almost 38 Go. The adopted scope for the analysis is as follows:

- **Temporal scope.** We focus on the year 2015. All four months are considered (March, June, September, December).
- **Origin-Destination (OD) scope.** We focus on flights departing from Toulouse-Blagnac (LFBO) and landing at Paris-Orly (LFPO). This OD pair is the most flown over the year 2015. Table 1 gives the number of flights each year.

Table 1 Number of flights departing from Toulouse-Blagnac (LFBO) and landing at Paris-Orly (LFPO) over the time period.

2015	2016	2017	2018	2019	2020
3,161	3,137	3,130	2,972	2,824	991

Note: Only four months are taken into account each year, as explained in Subsection 2.1.

Flights with less than 12 known positions are not taken into account in this analysis because there are not reliable enough. The resulting sample consists in 3,024 trajectories.

2.2 ERA5 hourly data on pressure levels from 1979 to present (weather)

ERA5 is the fifth generation European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis for the global climate and weather for the past four to seven decades. Reanalysis combines model data with observations from across the world into a globally complete and consistent data set using the laws of physics. The data set of interest for trajectories is entitled “ERA5 hourly data on pressure levels from 1979 to present”. In this data set, several weather variables are available on an hourly basis for 37 pressure levels on a $0.25^\circ \times 0.25^\circ$ longitude-latitude grid. Note that the longitude-latitude grid is particularly fine. The time resolution is also very detailed as compared to data commonly used in the transportation literature. Relative to the work of [Liu et al \(2021\)](#), our longitude-latitude grid is ten times finer (from 2.5 degrees to 0.25 degrees), weather values are available every hour (not every six hours) for the 23 pressure levels that are kept.

In this work, only two weather variables are chosen: the eastward component of the wind and the northward component of the wind. The former, the u -component of wind, is the horizontal speed of air moving towards the east for which a negative sign indicates air moving towards the west. The latter,

the v -component of wind, is the horizontal speed of air moving towards the north for which a negative sign indicates air moving towards the south. Both are expressed in $m.s^{-1}$.

Combined together, they give the speed of the horizontal wind on each point of the longitude-latitude grid mentioned above. Let u and v be the two components of the wind.

- **Wind speed.** The wind speed ws is computed as $ws = \sqrt{u^2 + v^2}$ and is expressed in $m.s^{-1}$.
- **Wind direction.** The wind direction wd is computed as $wd = \text{atan2}(v, u)$ and is expressed in radians. The wind direction is valued in $]-\pi, \pi]$. In this paper, the wind direction is mapped to $]0, 2\pi]$. When the wind direction is π , air is moving towards the west (east wind); $\frac{\pi}{2}$ means that the air is moving towards the north (south wind); $\frac{3\pi}{2}$ means that the air is moving towards the south (north wind); 0 or 2π translate into the air moving towards the east (west wind).

Note that, *per se*, the wind speed and direction are not informative to unambiguously define favorable/unfavorable wind conditions. To this end, a following subsection introduces the definition of the crosswind and tailwind components of the wind.

2.3 Towards augmented trajectories

To associate each point of a trajectory to a weather value, a naive matching may be done taking the closest altitude, the closest position (longitude, latitude) and the nearest time. A trajectory with weather dimensions is called an *augmented trajectory*.

Regarding the altitude, weather data are coming by pressure levels whereas trajectory data are given in flight levels (in feet). A flight level is an aircraft's altitude at standard air pressure. To calculate a standard pressure p at a given altitude h in meters, the temperature is assumed to be standard and the air is assumed to be a perfect gas. The air pressure is computed assuming an International Standard Atmosphere pressure of $p_0 = 1013.25$ hPa at sea level.

$$p = p_0 \left(1 - \frac{0.0065 \times h}{T_0} \right)^{5.2561}$$

where T_0 is a baseline temperature equal to $288.15^\circ K$.

Approximations are inevitable: the aircraft's positions are not exactly on the weather grid. For a given dimension (say time, altitude, longitude, latitude), the magnitude of the approximation is defined as the Euclidean distance between the query point (from the trajectory) and the closest weather value (on the grid).

The matching strategy is deterministic (1-nearest neighbor). It does not involve any statistical model. To ensure that this naive strategy is still acceptable, we may want to check that the average matching approximation is half

the value of the worst one. For instance, regarding the time dimension, the worst approximation is 30 minutes as we deal with hourly data. Hence, on average, we would hope that the closest available weather value in time is within a 15-minute distance. It seems to be the case looking at the summary statistics given in Table 2.

Table 2 Summary statistics for matching approximations for the flights departing from Toulouse-Blagnac (LFBO) and landing at Paris-Orly (LFPO) in 2015.

	Time (min)	Longitude (deg)	Latitude (deg)	Pressure (hPa)
Mean	15.05	0.06	0.06	11.5
Standard error ¹	8.65	0.04	0.04	7.52

¹Denominator is $n - 1$; n being equal to 3,024.

Heading, bearing, crosswind, tailwind

To properly define what are favorable and unfavorable wind conditions, one may rely on the concept of *crosswinds* and *tailwinds*. A crosswind is any wind that has a perpendicular component to the line or direction of travel. It affects the aerodynamics. A tailwind has a parallel component to the line or direction of travel. Note that the direction of travel should be properly defined here.

As explained in the Pilot's Handbook of Aeronautical Knowledge written by the Federal Aviation Administration (FAA), for an aircraft, the *intended track* (or *route*) is typically a set of straight-line segments between waypoints. The pilot determines the *bearing* (the compass direction from the aircraft's current position) of the next waypoint. In navigation, because wind can cause an aircraft to drift, the pilot sets a *course to steer* that compensates for drift. The pilot points the aircraft on a *heading* that corresponds to the course to steer. If the predicted drift is correct, then the craft's track will correspond to the planned course to the next waypoint.

To summarize, the *heading* would be the direction of travel. Unfortunately, it is not available in the R&D data from Eurocontrol. When the wind speed is small, the bearing and the heading almost coincide. As a first approximation in this work, the bearing is taken as the direction of travel.

To compute the bearing β valued in $]-\pi, \pi]$ from a starting position (λ_1, φ_1) in lon-lat (radians) to a final position (λ_2, φ_2) , the following formula is used:

$$\beta = \text{atan2}(\sin(\lambda_2 - \lambda_1)\cos(\varphi_2), \cos(\varphi_1)\sin(\varphi_2) - \sin(\varphi_1)\cos(\varphi_2)\cos(\lambda_2 - \lambda_1)).$$

The result is mapped to $]0, 2\pi]$. With a bearing β , a wind speed ws , and a wind direction wd as defined in Subsection 2.2, it is possible to define the crosswind and tailwind components.

- **Crosswind component.** The crosswind component w_{cross} is computed as $w_{cross} = ws \sin(wd - \beta)$. A positive value means that the wind blows

westward, perpendicular to the direction of travel. A negative value means that the wind blows eastward, perpendicular to the direction of travel.

- **Tailwind component.** The tailwind component w_{tail} is computed as $w_{tail} = ws \cos(wd - \beta)$. A positive value means that wind blows to the direction of travel.

3 A functional framework to model trajectories

Now that all data sources have been presented, the functional framework is introduced. First, Subsection 3.1 highlights that raw augmented data are discrete. To retrieve a functional sample, the so-called *smoothing* step is presented in Subsections 3.2 and 3.3. Smoothing methods in the literature are reviewed in Subsection 3.3.1. Constrained smoothing is developed in Subsection 3.3.2. Smoothing circular data is detailed in Subsection 3.3.3. The two adopted strategies, respectively based on asymmetric kernels and constrained splines, are to be found in Subsection 3.3.4.

3.1 Raw augmented data

Raw augmented data are stored in a usual matrix format. A raw trajectory i , denoted tra_i , is a set of m_i pairs:

$$\text{tra}_i = \{(\mathbf{y}_{i,j}, t_{i,j}), j = \{1, \dots, m_i\}\}$$

m_i being the number of observation times associated to flight i , $\mathbf{y}_{i,j}$ being, a seven-dimensional vector (longitude, latitude, altitude, u-component of the wind, v-component of the wind, wind speed, wind direction), $t_{i,j}$ being a timestamp. In what follows, the components of $\mathbf{y}_{i,j}$ are respectively denoted $y_{i,j,x}$, $y_{i,j,y}$, $y_{i,j,z}$, $y_{i,j,u}$, $y_{i,j,v}$, $y_{i,j,ws}$, $y_{i,j,wd}$.

At first sight, keeping the two horizontal wind components as well as the speed and direction is useless. This choice is motivated in Subsection 3.3.3.

A few preliminary remarks should be made. First, the duration changes from one flight to another. As a consequence, the number of observation points is not the same from a trajectory to another. Second, for a given trajectory, observation times are not equally spaced. Third, between two trajectories, the spacing is irregular and is never the same. For a given trajectory, observation times are the same for all components of $\mathbf{y}_{i,j}$.

Given that most methods in time series and multivariate statistics are inappropriate in this context, a relevant statistical framework is to study trajectories from a FDA perspective. The main assumption in FDA is the existence of an underlying function that has given birth to a set of values the statistician observes on a discrete grid.

3.2 The functional perspective

As pinpointed by [Hsing and Eubank \(2015\)](#), there are two approaches to build the mathematical basis of FDA: the *stochastic process perspective* and the *random element perspective*.

In the *stochastic process perspective*, a functional data (a curve), is viewed as a *sample trajectory* (also called a *sample path*). Let Y be an indexed collection of random variables defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For a fixed $\omega \in \Omega$, the mapping $Y(., \omega) : E \rightarrow \mathbb{R}$ is called a *trajectory*. E is the index set, often taken as a compact metric space in the stochastic process literature. In FDA, E is some (closed) interval of the real line. Very often indeed, E is chosen to be $[0, 1]$. The stochastic process is assumed to be a second-order process, so that the mean function of the process Y and its covariance function are well-defined.

In the sequel, Y is taken to be a mean-square continuous stochastic process, that is to say with well-defined continuous mean and covariance functions. The *mean* of the process is given by:

$$\mu(t) = \mathbb{E}[Y(t)], \quad t \in [0, 1].$$

The *variance* is denoted:

$$\sigma^2(t) = \mathbb{E}[Y(t)^2] - \mu(t)^2, \quad t \in [0, 1].$$

As a trajectory has several dimensions, the multivariate FDA framework is relevant here. A so-called *functional random vector* of dimension p is denoted $\mathbf{Y} \equiv (Y_1, Y_2, \dots, Y_p)^\top$. Note that $\forall k \in \{1, \dots, p\}$, $Y_k \equiv (Y_{k,t})_{t \in [0,1]}$. The mean vector of $\mathbf{Y}$ is given by:

$$\mu_{\mathbf{Y}}(t) = \mathbb{E}(\mathbf{Y}_t) = \begin{pmatrix} \mathbb{E}[Y_1(t)] \\ \vdots \\ \mathbb{E}[Y_p(t)] \end{pmatrix}, \quad t \in [0, 1].$$

The covariance of $\mathbf{Y}$ involves the auto-covariance and the cross-covariance functions of the components. An augmented trajectory is modeled as the realisation of a functional random vector of dimension $p = 7$.

The first step of most studies in FDA is to reconstruct the continuous realisations from the raw discrete data. It is the *smoothing* step introduced in Subsection 3.3.

3.3 The smoothing step

The main assumption in FDA is the relevance of the following equation:

$$\mathbf{y}_{i,j} = \mathbf{f}_i(t_{i,j})$$

$\mathbf{f}_i$ being an underlying function. Two main approaches should be conceptually distinguished. When the statistician assumes that observations are recorded

with negligible errors, $\mathbf{f}_i$ can be retrieved by *interpolation*. More frequently, observations $\mathbf{y}_{i,j}$ are assumed to be noisy which can be written

$$\mathbf{y}_{i,j} = \mathbf{f}_i(t_{i,j}) + \varepsilon_{i,j}$$

where $\varepsilon_{i,j}$ are unobserved errors. In this second case, $\mathbf{f}_i$ can be approximated using smoothing techniques.

The smoothing approach has prevailed in the literature following the work of Ramsay and Silverman (2002) and Ramsay and Silverman (2005) and was almost viewed as a pre-processing step. This approach has been coined the “smoothing first, then estimation” strategy by Zhang and Chen (2007) who carefully investigated the effect of the substitution of the underlying individual functions by reconstructed functions obtained by Local Polynomial Kernel (LPK) smoothing.

Yet, as pinpointed by Hall and Van Keilegom (2007), there is rightly a debate as to whether statistical smoothing should be used at all, in a conventional sense, when constructing two-sample hypothesis tests (which is the ultimate objective of the paper). Smoothing is relevant if the new observed times are not too numerous as compared to the original ones and if the same tuning parameters are used to produce each curve in each group.

When functions are sparsely observed, smoothing should not be applied to individual sparse trajectories. As put by Kokoszka and Reimherr (2017), imputed smooth trajectories can be obtained only after information from the whole sample has been suitably combined. In this case, a two-sample test has recently been proposed by Wang (2021).

The underlying distinction is the one between the *sparse* and *dense* frameworks in FDA. Although no official definition exists, generally, if all m_i (number of points) are larger than some order of n (number of individuals), then the functional data are referred to as “dense” data. If all m_i are bounded, the data are commonly considered as “sparse”. Zhang and Wang (2016) have shown that the two frameworks lead to distinct asymptotic properties and have an impact on the choice of estimation procedures for the mean and covariance functions.

In this work, trajectories are assumed to be detailed enough for individual functions to be reconstructed. In the next section, smoothing methods are presented for a single dimension of a trajectory. The index i is then dropped to avoid complicated notations. The number of points in trajectory i is then simply $m = m_i$. Indications are given when a specific dimension is of interest (for instance, the altitude).

3.3.1 Smoothing methods

Many smoothing methods are available to reconstruct individual trajectories. The *basis expansion approach* is probably the most popular one and is at the heart of the monograph written by Ramsay and Silverman (2005). The

basic idea is to approximate a functional data by a linear combination of basis functions.

The basis choice is usually made *a priori*, motivated by the nature of the phenomenon under study and the choice to estimate derivatives or not. The Fourier basis and the B-spline basis are extensively used in the literature. Yet, wavelets also have a long history in the FDA literature as testified by the early work of [Antoniadis et al \(1994\)](#). Their use in FDA is not restricted to the curve estimation problem as illustrated by [Morettin et al \(2017\)](#).

The chosen number of basis functions involves a usual bias/variance trade-off. Algorithms for choosing the number of functions benefit from the vast literature on variable selection in multiple regression (e.g. stepwise variable selection or variable-pruning). Coefficients involved in the linear combination are often estimated to minimize a least square criterion, possibly penalized. Adding a penalization may conceptually solve the problem of choosing the number of basis functions. Indeed, one may depart from a generous choice of basis functions and exploit a theoretical continuum of smoothing levels. Yet, the roughness parameter is still to be chosen. In practice, a cross-validation rule is used.

When the regression B-spline strategy is adopted, the B-spline basis system developed by [de Boor \(2001\)](#) is commonly used. Good performances require relevant knot locations as well as a good choice of the number of knots. In the knots selection literature, the possible knots come from a predetermined set such as the design points or grid points in the range.

Another approach is based on *kernel smoothing* which is reviewed by [Zhang \(2013\)](#) as well as other popular non-parametric techniques used in FDA. Among the family of Local Polynomial Kernel (LPK) smoothers, the local constant smoother (known as the Nadaraya-Watson estimator) may suffer from boundary effects problem. This problem is key in FDA as the time interval is often taken to be $[0, 1]$. In practice, the local linear smoother is then preferable as it is known as a smoother with a free boundary effect, as shown by [Cheng et al \(1997\)](#). Yet, it is not robust against the presence of outliers.

In the multivariate context, two approaches are possible. A simple one is to obtain a p -dimensional curve by juxtaposition of p separate smoothing of the p functional coordinates of the curve (*basis expansion approach*, *kernel smoothing* or a mix of the two). A more ambitious approach is to go for an estimate that is itself a proper p -dimensional curve. In this regard, [Sangalli et al \(2009\)](#) smooth three-dimensional curves and their derivatives by free-knot regression splines with an application to the analysis of inner carotid artery centrelines. [Pigoli and Sangalli \(2012\)](#) develop a wavelet-based estimation of multidimensional curves and their derivatives with an application to the modeling of electrocardiogram records.

Yet, the later p -dimensional method may be hard to implement when specific constraints must be verified for some dimensions as it is typically the case for augmented aircraft trajectories.

3.3.2 Constrained smoothing methods

Smoothed augmented aircraft trajectories must verify some physical constraints. These are mainly of two types: *non-negativity constraints* and *boundary constraints*.

- **Non-negativity constraints.** These constraints refer to the fact that altitude and wind speed values must be non-negative once smoothed. There is an early literature in statistics on spline and kernel regression under shape restrictions. An excellent literature review is given in Chapter 5 of [Schimek \(2013\)](#). Usual constrained smoothing problems involving splines may be expressed as follows. Given data $\{(y_j, t_j)\}_{j=1}^m$ with $t_j \in [0, 1]$, the problem is:

$$\operatorname{argmin}_f \sum_{j=1}^m [y_j - f(t_j)]^2 + \lambda \int_0^1 [f^{(\ell)}(u)]^2 du$$

under the constraint that $f^{(r)}(t) \geq 0, \forall t \in [0, 1]$. Usual constraints are then non-negativity ($r = 0$), monotony ($r = 1$) and convexity ($r = 2$). Monotony and convexity are the two restrictions that have been the most studied.

A common practice is to go for a discretized version of the restrictions, obtained by choosing a finite number of t_j . This results in a finite number of linear constraints that can be included in the minimization problem. Yet, it is then impossible to ensure that the constraint is globally satisfied. Reversely, the drawback of most global methods is the lack of a computing algorithm. In the FDA literature, the non-negativity constraint in the *basis expansion approach* is handled following the idea of [Ramsay and Silverman \(2005\)](#) (Subsection 6.2.1). To ensure that $f(t) \geq 0 \forall t \in [0, 1]$, $f(t)$ is defined as the exponential of an unconstrained function. This unconstrained function is expanded in terms of a set of basis functions. Because the criterion is then not linear in terms of the coefficients, numerical methods are used.

When several constraints must be met simultaneously, [Turlach \(2005\)](#) adopted another point of view. He focused on splines of order 4, with knots located at observation points. Instead of choosing a suitable B-spline basis and identifying the necessary constraints on the coefficients of the basis functions, an unconstrained smoothing spline is first fitted. If there are any violations, constraints are added. The process of verifying-and-adding-new-constraints is iterated until there are no violations anymore.

On the *kernel smoothing* side, LPK smoothers may not meet the non-negativity condition. We may have negative predictions with a local linear smoother even if $y_j \geq 0, \forall j$. Yet, the Nadaraya-Watson smoother ensures non-negativity when $y_j \geq 0, \forall j$.

- **Boundary constraints.** These constraints ensure that the altitude must be zero at the departure ($t = 0$) and at the arrival ($t = 1$). Additionally,

the first longitude-latitude values must be the ones of the departure airport. Respectively, the last longitude-latitude values must be the ones of the arrival airport.

To ensure that boundary constraints are verified, several approaches are possible. In the *basis expansion approach* framework, one may drop some functions of the basis to ensure that the altitude would be zero at departure and arrival. Yet, it is not clear how to ensure some general boundary value constraints in the *kernel smoothing* approach.

3.3.3 Functional circular data

There is more to handle than the constraints detailed in Subsection 3.3.2. Strictly speaking, longitude, latitude and wind directions, are angular measurements.

As flights from Toulouse-Blagnac (LFBO) and Paris-Orly (LFPO) are taking place over Metropolitan France, the angular nature of the longitude and the latitude is not of paramount importance, as opposed to long-haul flights studied by Dai and Müller (2018). Yet, the direction of the wind should be handled carefully.

Wind directions are famous examples of circular data, that is to say with support on the unit circle. As put by Ley and Verdebout (2018), directional statistics probably started in the 1950s thanks to the seminal paper of Fisher (1953), studying the historic and prehistoric remanent magnetism of Icelandic lava flows. A famous monograph in directional statistics is the one written by Jupp and Mardia (1999). Some case studies are developed by Ley and Verdebout (2018).

When time is involved, *circular time series* have been early studied as testifies Fisher and Lee (1994). Unsurprisingly, circular time series of wind directions are an early application as highlighted by Breckling et al (1989). A recent literature review on circular time series is provided in Ugwuowo and Udokang (2022) as well as an application on hourly measurements of wind direction taken over a period of time at the Energy research Centre of the University of Nigeria, Nsukka.

When wind directions are not evenly spaced in time (which is the case in this work), it is natural to consider the framework of *functional circular data* that has been less common in the literature.

Smoothing functional circular data can be seen as a non-parametric regression task involving a circular variable as a response. The later has been developed by Di Marzio et al (2013) both for circular and real-line predictors. We consider the $[0, 1] \times \mathbb{T}$ -valued random vector (Δ, Φ) , where $\mathbb{T}$ is the circle with unit radius, Φ the response and Δ the predictor. We have a random sample $\{(\Delta_j, \Phi_j), j = 1, \dots, m\}$ of independent elements from (Δ, Φ) .

The regression model is given by:

$$\Phi_j = [f(\Delta_j) + \varepsilon_j] \pmod{2\pi}, \quad j = 1, \dots, m$$

where random error angles ε_j have zero mean direction (the mean direction of the resultant vector is null), finite concentration (the mean resultant length is finite) and are independent of the predictors. The estimator for the regression function at $\delta \in [0, 1]$ is:

$$\hat{f}(\delta) = \text{atan2}(\hat{g}_1(\delta), \hat{g}_2(\delta))$$

where $\hat{g}_1$ and $\hat{g}_2$ are respectively given by:

$$\hat{g}_1(\delta) = \frac{1}{m} \sum_{j=1}^m \sin(\Phi_j) W(\Delta_j - \delta)$$

and

$$\hat{g}_2(\delta) = \frac{1}{m} \sum_{j=1}^n \cos(\Phi_j) W(\Delta_j - \delta)$$

with W being a local weight.

Because we have redundant information for the wind values, two approaches are possible.

- **Component approach.** We avoid the directional aspect. First, one smooths the two (unconstrained) components of the wind. Second, one computes a smoothed wind speed and direction.
- **Directional approach.** One smooths the constrained wind speed and the directional wind direction straightaway.

Results may differ.

3.3.4 Adopted strategies and results for the smoothing step

Because raw observation times are not evenly spaced points (and m is not equal to 2^J for some integer $J \in \mathbb{N}$), using wavelets to smooth trajectories is not straightforward.

Yet, a *basis expansion* approach and a *kernel smoothing* one are both interesting. As physical constraints must be verified for some dimensions, the construction of the functional observations using the discrete data takes place separately for each flight and for each dimension of a flight. For each strategy, the component approach and direction approach are implemented.

- **Asymmetric kernel smoothing (first strategy).**

To ensure that the non-negativity constraint as well as the boundary constraints simultaneously hold, asymmetric kernels are used. These are recently reviewed by [Hirukawa \(2018\)](#). Asymmetric kernels emerged as a way to tackle the boundary bias problem. [Brown and Chen \(1999\)](#) approximated regression curves with support on the unit interval using the Bernstein polynomials smoothed by a family of beta densities. As an extension, [Chen \(1999\)](#) proposed to use a family of beta densities as kernels for density estimation. [Chen \(2000\)](#) used a beta kernel for the regression task. A weight function

$K_{t,b}(\cdot)$ is said to be an asymmetric kernel if it possesses the following two basic properties stated in [Hirukawa \(2018\)](#):

1. The weight function is a probability density function (pdf) with support either on the unit interval $[0, 1]$ or on the positive half-line $\mathbb{R}^+$.
2. Both the location and shape parameters in the weight function are functions of the design point x where smoothing is made and the smoothing parameter b .

Nonparametric estimators involving asymmetric kernels are free of boundary bias. Contrary to usual kernels, the shape of a given asymmetric kernel varies across design points. As a consequence, the amount of smoothing changes in an adaptive manner. The functional form of the beta kernel introduced in [Chen \(1999\)](#) is given by the probability density function of the beta distribution $B\left(\frac{t}{b} + 1, \frac{1-t}{b} + 1\right)$. Figure 1 shows the shapes of the beta kernel for two different values of smoothing.

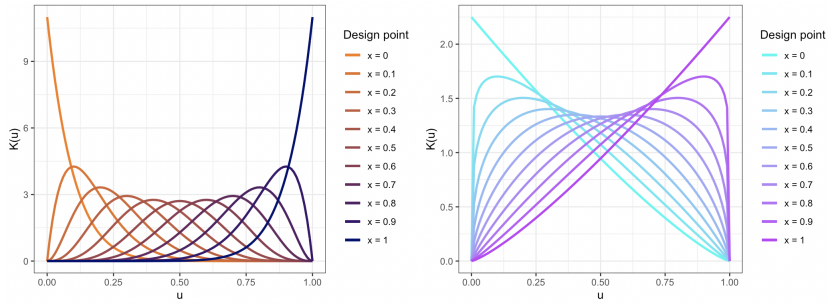


Fig. 1 Beta kernels for $b = 0.1$ [left] and $b = 0.8$ [right].

The Nadaraya-Watson regression smoother at a given design point t using the beta kernel (*fixed design*), is defined as:

$$\hat{f}_B^{NW}(t) = \frac{\sum_{j=1}^m Y_j K_{B(t,b)}(t_j)}{\sum_{j=1}^m K_{B(t,b)}(t_j)}.$$

Note that our asymmetric kernel approach is not the one developed by [Ferry and Vieu \(2006\)](#) or [Benhenni et al \(2007\)](#). These works explore the regression estimation problem when the explanatory data are curves and the response is real while we use asymmetric kernels to smooth functional data. Bandwidth selection is performed using a Generalized Cross-Validation (GCV). This criterion was introduced by [Craven and Wahba \(1978\)](#) to choose the correct degree of smoothing using spline functions and has been defined by [Shi and Song \(2016\)](#) for the Nadaraya-Watson regression

smoother using the gamma kernel. For the beta kernel, the GCV is written:

$$\text{GCV}(b) = \frac{m \sum_{j=1}^m (Y_j - \sum_{k=1}^m w_{j,k} Y_k)^2}{\left(m - \sum_{j=1}^m w_{jj}\right)^2}$$

where

$$w_{j,k} = \frac{K_{B(t_j,b)}(t_k)}{\sum_{\ell=1}^m K_{B(t_j,b)}(t_\ell)}, \quad j, k = 1, \dots, m.$$

- **Constrained B-spline smoothing (second strategy).** We consider the following minimization problem for the altitude which is a modified version of the one proposed by [Ramsay and Silverman \(2005\)](#) (Subsection 6.2.1):

$$\underset{\alpha_1, \dots, \alpha_K}{\text{argmin}} \sum_{j=1}^m \left[y_j - \exp \left\{ \sum_{k=1}^K \alpha_k \phi_k(t_j) \right\} t_j(1-t_j) \right]^2 + \lambda \int_0^1 \left[\sum_{k=1}^K \alpha_k \phi_k^{(2)}(u) \right]^2 du.$$

Note that the function $g : t \mapsto t(1-t)$ imposes the exact value constraint. This function is non-negative on $[0, 1]$, reaches a maximum value of $\frac{1}{4}$ when $t = \frac{1}{2}$. In practice, a Gauss-Newton algorithm is used to estimate the parameters. Figure 2 shows that the two smoothing strategies produce acceptable results. Constraints are verified. Smoothed trajectories are bumpier when using the asymmetric kernel smoother associated to the GCV criterion. In any case, the mean is very similar for the two smoothing strategies.

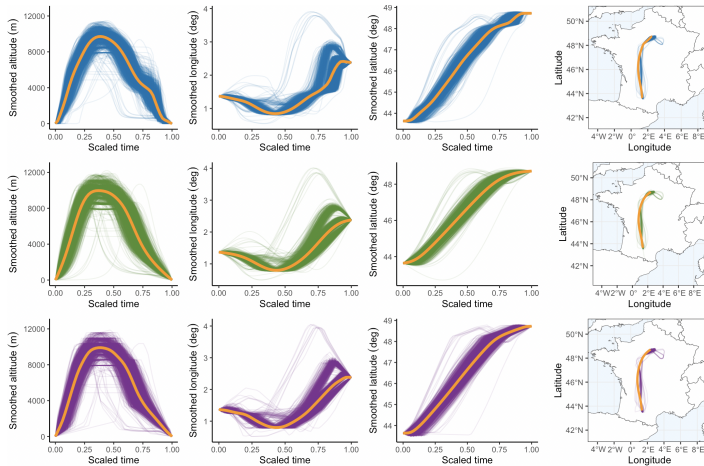


Fig. 2 Smoothing results for the first three dimensions using an asymmetric beta kernel [top row] and constrained smoothing B-splines [middle row]. Linear interpolation is also provided [bottom row]. pointwise means are indicated in orange.

- **Why is the circular nature of the wind direction important?**

If one decides to work with the wind direction directly, a non-parametric regression for a circular response will give accurate results. The regression has been implemented in R thanks to the `NPCirc` package for which details can be found in Oliveira et al (2014). Figure 3 shows, for a given flight, that taking the circular nature of the wind direction is key to smooth values and account for the discontinuity between 0 and 2π .

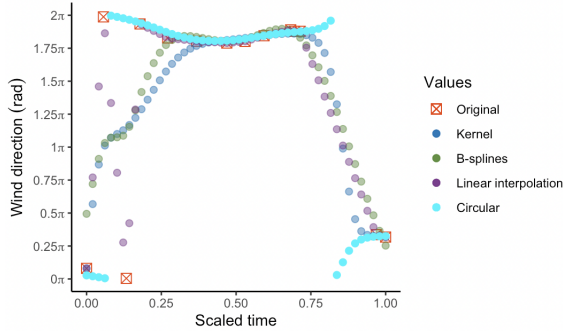


Fig. 3 Smoothing wind directions for a specific flight with several methods. Circular values refer to smoothed values obtained from a non-parametric regression for circular response and real-line predictor.

4 Elastic registration

This section details the phase problem associated to the fact that trajectories are all of different durations.

4.1 The detrimental effect of rescaling

Empirically, two flights never have the exact same duration. To compare trajectories with different durations, a scaling is made to the unified time interval $[0,1]$. This transformation is popular in FDA but has a well-known impact on summary statistics. To better understand, Figure 4 shows two theoretical trajectories having:

- A similar longitude-latitude profile (the path from the departure airport to the arrival airport is assumed to be the same)
- A similar trapezoid altitude profile except for the en-route phase that is assumed to be longer for one of the two flights. Specifically, the climb and the descent are the same.

Comparing altitude values for a given time is obviously misleading. From the scaled time version, at each time, the altitude seems higher for the red trajectory which is obviously wrong from looking at the initial situation. Intuitively, one wishes to compare trajectories for similar flight phases. This well-known problem is called a *registration* problem in FDA.

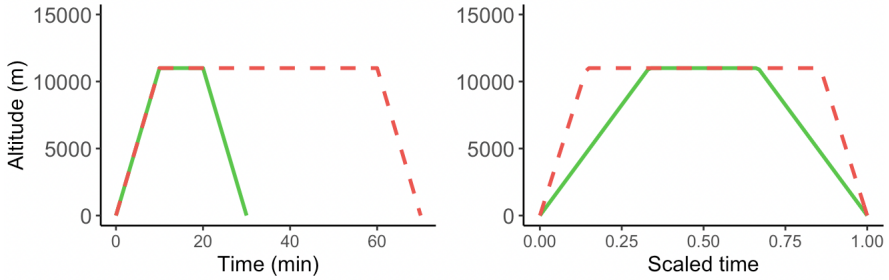


Fig. 4 Two theoretical altitude profiles (green and red). Time rescaling [right] imposes a registration step to avoid misguided interpretations.

4.2 Previous related works on registration - univariate framework

As put in Chapter 7 of [Ramsay and Silverman \(2005\)](#), registration starts with the key idea that the rigid metric of physical time may not be directly relevant to the internal dynamics of many real-life systems. To avoid fallacious comparisons, time should be stretched or shrunk for shapes to be aligned. For trajectories, one wishes to compare trajectories for the same flight phases.

In general, registration can be as simple as shifting time or as tedious as finding a continuous transformation of the time scale. One is typically more interested in creating a continuous transformation of the time scale that preserves boundaries.

Most registration approaches boil down to the resolution of an optimisation problem and is traditionally based on Procrustes fitting following [Silverman \(1995\)](#). An famous exception is landmark registration that has been studied by [Kneip and Gasser \(1992\)](#) and [Gasser and Kneip \(1995\)](#).

Trajectories are not already labeled: flight phases are not known in the raw data. A first option is to create landmarks from the detection of flight phases based on some algorithms that exist in the literature. Two examples are the ones of [Sun et al \(2016\)](#) and [Liu et al \(2020\)](#).

Yet, continuous registration is preferred as it is more flexible and does not require an additional pre-processing step. Conceptually, the registration can be done *pairwise* or *groupwise*.

The *pairwise alignment problem* for registering f_2 to f_1 (reference), f_2 and f_1 being two functions in $\mathbb{L}^2$, may be written as:

$$\gamma^* = \operatorname{argmin}_{\gamma \in \Gamma_{[0,1]}} E[f_1, f_2 \circ \gamma]$$

where E stands for an optimization criterion and $\Gamma_{[0,1]}$ for the set of boundary-preserving diffeomorphisms on $[0, 1]$. E is often chosen to be a distance, e.g., the usual $\mathbb{L}^2$ norm.

The *groupwise alignment problem* involves a set of functions $\{f_i, i = 1, \dots, n\}$ and aims at finding a set of warping functions $\{\gamma_i, i = 1, \dots, n\}$

such that, for any $t \in [0, 1]$, the values $f_i(\gamma_i(t))$ are said to be registered with each other.

Including a symmetry problem (registering f_2 to f_1 is not the same thing as doing the opposite), using the $\mathbb{L}^2$ norm has several undesirable effects. A famous one is the *pinching* effect: using the $\mathbb{L}^2$ allows for squeezing or pinching a large part of the candidate function to make the cost function arbitrarily close to zero. In practice, a regularization term is used to constrain the roughness of the warping function. An more elegant solution has been proposed by [Srivastava et al \(2011b\)](#) for separating the phase and the amplitude variability in functional data.

4.3 A geometric framework for functions

The main idea of the geometric approach is to take advantage of Square-Root Slope Function (SRSF) representation of functions. This representation allows to find the optimal warping function efficiently. Departing from a function f that is absolutely continuous on $[0, 1]$, the map $Q : \mathbb{R} \rightarrow \mathbb{R}$ is defined as

$$Q(x) \equiv \text{sign}(x)\sqrt{|x|}.$$

The SRSF associated to f is defined as $q : [0, 1] \rightarrow \mathbb{R}$, where $q(t) \equiv Q(\dot{f}(t))$. [Srivastava and Klassen \(2016\)](#) have shown that if the function f is absolutely continuous, then the resulting SRSF is square integrable. For every $q \in \mathbb{L}^2([0, 1])$ there exists a function f (unique up to an additive constant) such that the given q is the SRSF of that f . If the cost function for registering f_2 to f_1 is chosen to be $E[f_1, f_2] = \|q_1 - q_2\|_{\mathbb{L}^2}$, finding the optimal registration amounts to solve:

$$\underset{\gamma \in \Gamma_{[0,1]}}{\text{argmin}} E[f_1, f_2 \circ \gamma] = \underset{\gamma \in \Gamma_{[0,1]}}{\text{argmin}} \left\| q_1 - (q_2 \circ \gamma)\sqrt{\dot{\gamma}} \right\|_{\mathbb{L}^2}.$$

In practice, a Dynamic Programming Algorithm (DPA) is used to compute the optimal time-warping function. A common practice is rather to solve:

$$\underset{\gamma \in \Gamma_{[0,1]}}{\text{argmin}} \left\| q_1 - (q_2 \circ \gamma)\sqrt{\dot{\gamma}} \right\|_{\mathbb{L}^2}^2.$$

Regarding the *groupwise alignment problem*, a template function is required. Ideally, the template function should not be based on a mean built upon the $\mathbb{L}^2$ norm because it does not verify an isometry property under warping. Building a valid template relies on the Karcher mean, for which details can be found in [Marron and Dryden \(2021\)](#). In practice, the mean, computed iteratively, may serve as a reference function to which are registered all functions of the set.

4.4 Previous related works on registration - multivariate framework

When it comes to multivariate functional data, the issues of phase and amplitude variations still exist. Yet, phase variation for multivariate functional data is twofold. The so-called *observation phase* is the phase variation occurring across all dimensions of a given trajectory. *Component phase* is the phase variation occurring for a given dimension across all trajectories. Because component-wise registration is tempting, it is often assumed that cross-component variations are uncorrelated. Under this assumption, the registration of multivariate functional data consists in several univariate registrations: one should find several warping functions for each trajectory. When it is assumed that there are no component phase variations, registration is done to account for the observation phase. In this case, only one warping function is to be found for each trajectory. For example, [Park and Ahn \(2017\)](#) estimated a conditional subject-specific warping to compare growth patterns across individuals. [Carroll et al \(2021\)](#) accounted for both observation and component phases. In their work, observation phase variations are taken into account thanks to the estimation of time shifts, in line with a well-known pattern in pubertal spurts.

To register trajectories, one should take into account the correlation that exists between all the dimensions: component-wise registration is not an option. For instance, it is known that the wind speed will change with altitude because of atmospheric layers. Because the dimensions of a trajectory are the ones of an underlying coordinate system, an universal warping is to be found for all components. Luckily, there is no observation phase for trajectories because of the data acquisition process.

As stated by [Delahaye et al \(2014\)](#) an aircraft trajectory is naturally modeled as an open parameterised curve of $\mathbb{R}^p$. More rigorously, a trajectory $\beta : [0, 1] \rightarrow \mathbb{R}^p$ has an associated curve $\beta : (0, 1) \rightarrow \mathbb{R}^p$ (that has no end-points). While [Delahaye et al \(2014\)](#) focused on the case $p = 3$ relying on arc length parametrization, we propose a more flexible representation of aircraft trajectories.

4.5 A geometric framework for curves

A natural extension of the SRSF to curves has been proposed by [Srivastava et al \(2011a\)](#). This extension is the Square-Root Velocity Function (SRVF) and relates to absolutely continuous curves. Let $F : \mathbb{R}^p \rightarrow \mathbb{R}^p$ be a mapping given by:

$$F(v) = \begin{cases} \frac{v}{\sqrt{|v|}} & \text{if } |v| \neq 0 \\ 0 & \text{if } |v| = 0 \end{cases}$$

where $|\cdot|$ notes the Euclidean norm of $\mathbb{R}^p$. The SRVF is defined to be

$$q(t) \equiv F(\dot{\beta}(t)).$$

The curve β can be reconstructed from q up to a translation. Pairwise registration of curves in $\mathbb{R}^n$ is essentially an extension of the problem studied above. Likewise, underlying concepts and even algorithms remain mostly the same for the registration of multiple curves.

4.6 Registration of augmented aircraft trajectories

The goal of the registration is to align flight phases. As flight phases are determined by the longitude, the latitude and mostly by the altitude, these dimensions are used to compute the optimal time-warping functions. Results are shown in Figure 5.

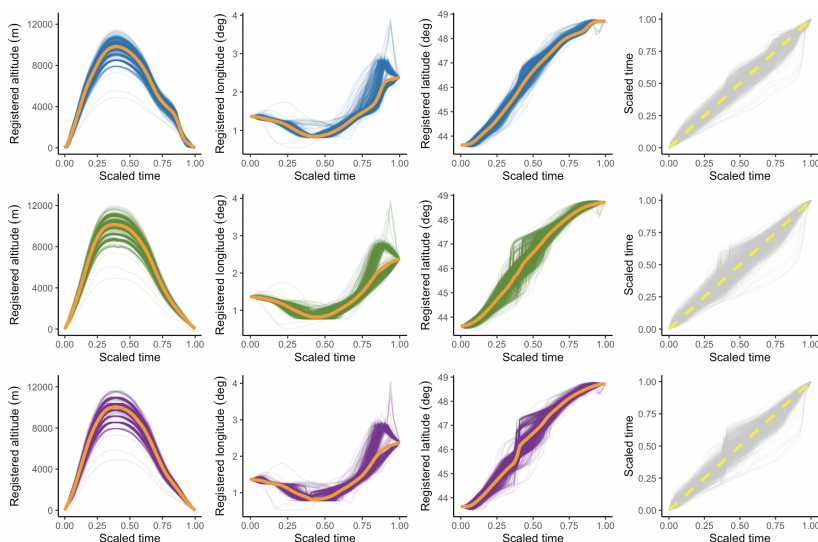


Fig. 5 Registration results for the first three dimensions when smoothed curves come from the asymmetric beta kernel [top row] and smoothing B-splines [middle row]. Linear interpolation is also provided [bottom row]. Pointwise means are indicated in orange. Optimal time-warping functions are given in grey. The dashed yellow line is the diagonal (no warping).

4.7 Descriptive statistics

Now that augmented aircraft trajectories have been registered, summary statistics can be computed. The crosswind and tailwind components of the wind are computed at each moment of the flight.

As one notices from Figure 6, the latest trajectories have greater tailwind values for most part of the flight. Reversely, earliest trajectories have highly negative tailwind values. Regarding the crosswind component, earliest trajectories have experienced smaller values. The next section develops a rigorous testing framework to prove it.

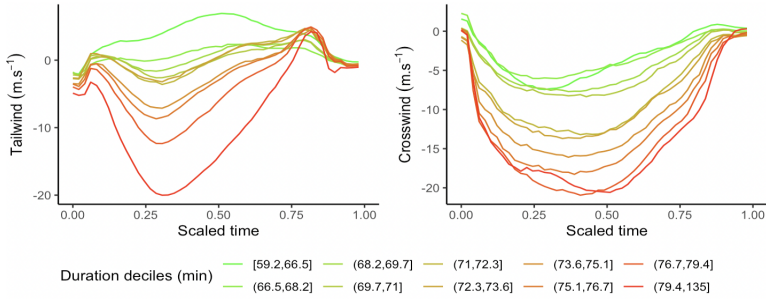


Fig. 6 Mean tailwind values by duration deciles computed from the asymmetric kernel approach [left] and mean crosswind values by duration deciles [right]. The directional approach was used.

5 Testing for a different experienced wind

Now that trajectories are smoothed taking constraints into account (Section 3) and that the registration has been done (Section 4), we wish to prove that delayed flights have experienced less favorable wind conditions. We define the on time and delayed flights in Subsection 5.1. Subsection 5.2 presents some inference procedures for functional data. Subsection 5.3 motivates the use of a pointwise test, implements and comments on the results of several procedures.

5.1 Definition of delayed and on time flights

Two groups can easily be made based on the duration of the flights. Note that the duration of a given flight is computed as the elapsed time between takeoff and landing. In what follows, the groups of delayed and on time flights are made according to the 8th decile of flight durations as shown in Figure 7.

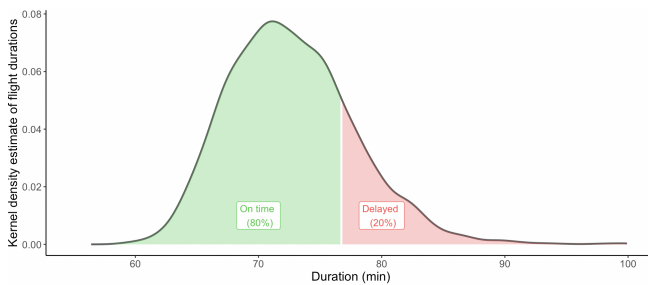


Fig. 7 Empirical density (Gaussian kernel smoothing) of all flight durations.

The goal of this section is to present the statistical framework to test for a different experienced wind between the two groups.

5.2 Statistical framework

Following Zhang (2013), several families of inference techniques can be used for functional data: pointwise, $\mathbb{L}^2$ norm-based, F -type, and bootstrap tests. One of the earliest contribution in the literature is probably the one of Faraway (1997) showing that a usual multivariate likelihood ratio test is not very adapted to compare nested linear models in the functional framework. Faraway (1997) motivates a $\mathbb{L}^2$ norm-based test and a bootstrap approach to approximate the null distribution of interest. The $\mathbb{L}^2$ norm-based test was extended in Zhang and Chen (2007). The F -type test was first studied by Shen and Faraway (2004). Other contributions are listed in Zhang (2013). If groups are independent but some temporal correlation should be taken into account in each group, a two-sample test for functional time series was proposed by Horváth et al (2013).

In what follows, one is interested in proving that the experienced wind was less favorable for delayed trajectories.

5.3 Pointwise tests for the crosswind and tailwind components

We chose to define a favorable/unfavorable experienced wind relying on the tailwind and crosswind components of the wind. Our working assumptions, motivated by summary statistics, are as follows:

- At each moment of the flight, tailwind values are higher, on average, for the on time group.
- At each moment of the flight, crosswind values are closer to zero, on average, for the on time group.

In this subsection, the focus is made on pointwise tests. An overall testing method (say, based on the $\mathbb{L}^2$ norm) is by nature not very appropriate to emphasize at which moment of the flight the differences between the two groups are important. Bigger differences in tailwind values are likely to happen in the middle of the flight, at high altitudes.

We have a sample of $n_1 = 2,419$ on time trajectories and a sample of $n_2 = 605$ delayed trajectories.

The tailwind component

We have two functional samples of tailwind values. The on time sample $\{Y_{1,i}^{\text{tail}}\}_{i=1}^{n_1}$ is made of independent stochastic processes identically distributed with mean function $\mu_1^{\text{tail}}(t)$ and covariance $\gamma^{\text{tail}}(s, t)$. The sample for the delayed group $\{Y_{2,i}^{\text{tail}}\}_{i=1}^{n_2}$ is made of independent stochastic processes identically distributed with mean function $\mu_2^{\text{tail}}(t)$ and covariance $\gamma^{\text{tail}}(s, t)$. By assumption, the covariance is the same for the two groups.

A general one-sided pointwise two-sample test for functional data with a common covariance function can be formulated as follows:

$$\begin{aligned} H_{0,t} : \mu_1^{\text{tail}}(t) &= \mu_2^{\text{tail}}(t) \\ &\text{against} \\ H_{1,t} : \mu_1^{\text{tail}}(t) &> \mu_2^{\text{tail}}(t) \end{aligned}$$

for a fixed $t \in [0, 1]$. Let $\widehat{\mu}_1^{\text{tail}}(t)$ be the usual estimator of the mean function $\mu_1^{\text{tail}}(t)$, $\widehat{\mu}_2^{\text{tail}}(t)$ be the usual estimator of the mean function $\mu_2^{\text{tail}}(t)$ and $\widehat{\gamma}^{\text{tail}}(s, t)$ be the pooled estimator of the common covariance function. The pivotal test statistic for each $t \in [0, 1]$ is given by:

$$z^{\text{tail}}(t) = \frac{\widehat{\mu}_1^{\text{tail}}(t) - \widehat{\mu}_2^{\text{tail}}(t)}{\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \widehat{\gamma}^{\text{tail}}(t, t)}}.$$

Let's make additional assumptions:

- $\mu_1^{\text{tail}}(t), \mu_2^{\text{tail}}(t) \in \mathbb{L}^2([0, 1])$
- $\int_0^1 \gamma^{\text{tail}}(t, t) dt < \infty$
- $\min(n_1, n_2) \rightarrow \infty$, n_1 and n_2 go to infinity proportionally
- The two groups are independent

If these assumptions hold, [Zhang \(2013\)](#), shows (Theorem 5.2, page 134), under the null, that the pivotal test statistic is asymptotically Gaussian for any fixed $t \in [0, 1]$. Pointwise p-values can be computed.

Yet, a pointwise test is not entirely satisfactory for at least two reasons. First, even if all the z -tests are significant at a given level, there is no guarantee for an overall significance. This caveat was early mentioned in [Ramsay and Silverman \(2005\)](#) speaking about the misleading interpretation of pointwise confidence intervals. Second, the same sample of trajectories is used to perform the test at each $t \in \{\tau_1, \dots, \tau_m\}$, the set of evaluation points for which curves have been sampled. In this context, a correction for the multiple comparisons should be found. Let $C_m = \{\tau_k : H_{0,\tau_k} \text{ is true}, 1 \leq k \leq m\}$ be the set of grid points for which H_{0,τ_k} is true and $\mathbb{P}(\text{reject } H_{0,\tau_k}, \forall \tau_k \in C_m)$ be the family-wise error rate. A correction would ensure that the family-wise error rate is less than or equal to α , where α is fixed, no matter what is the set C_m of true null hypotheses. Yet, as put by [Cox and Lee \(2008\)](#), the family-wise error rate depends on the set of grid points. The famous Bonferroni correction would require to test each individual hypothesis at a significance level of $\frac{\alpha}{m}$. Yet, there is an obvious problem for functional data when m is very large: $\frac{\alpha}{m} \xrightarrow{m \rightarrow \infty} 0$. Another famous correction is given by [Holm \(1979\)](#). Both corrections are relevant when rejection regions are disjoint. However, with functional data, rejection probabilities are likely to be correlated for two consecutive evaluation points in $\{\tau_1, \dots, \tau_m\}$. For functional data, [Cox and Lee \(2008\)](#) showed

that the Westfall-Young randomization method is appropriate if, among other hypotheses, a permutation pivotality condition holds.

More realistically, a group-specific covariance function is needed. Let $\gamma_1^{\text{tail}}(s, t)$ be the covariance function of the first group and $\gamma_2^{\text{tail}}(s, t)$ the covariance function of the second group. The pivotal test statistic, for each $t \in [0, 1]$ is now:

$$z_*^{\text{tail}}(t) = \frac{\widehat{\mu}_1^{\text{tail}}(t) - \widehat{\mu}_2^{\text{tail}}(t)}{\sqrt{\frac{\widehat{\gamma}_1^{\text{tail}}(t, t)}{n_1} + \frac{\widehat{\gamma}_2^{\text{tail}}(t, t)}{n_2}}}.$$

With similar assumptions, the pivotal test statistic is asymptotically Gaussian for any fixed $t \in [0, 1]$.

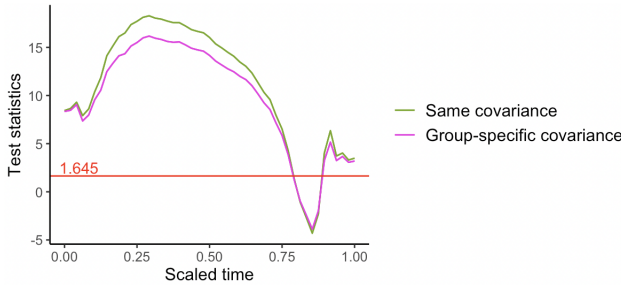


Fig. 8 Pointwise values of the test statistic for 50 values of $t \in [0, 1]$. The horizontal red line is the quantile that should be used to reject the null hypothesis at 5% level when no correction for multiple comparisons is taken into account. The test statistics are computed using asymmetric kernel smoothing with the directional approach.

Figure 8 shows that the on time group has an average tailwind that is greater than the delay group for most parts of the flight. Assuming a group-specific covariance lowers the test statistics. When asymmetric kernel smoothing is used and one uses the u and v component of the wind (component approach), results are very similar. Differences in test statistics are not very visible to the naked eye. If one looks at the test statistics obtained with the constrained splines approach, the global shape is the same. Yet, it is not possible to reject the null at 5% close to the landing phase.

The crosswind component

Note that any crosswind value that is too large in absolute value would affect the aerodynamics of the aircraft. The on time sample of absolute crosswind values $\{Y_{1,i}^{\text{cross}}\}_{i=1}^{n_1}$ is assumed to be made of independent stochastic processes identically distributed with mean function $\mu_1^{\text{cross}}(t)$ and covariance $\gamma_1^{\text{cross}}(s, t)$. The sample of absolute crosswind values for the delayed group $\{Y_{2,i}^{\text{cross}}\}_{i=1}^{n_2}$ is assumed to be made of independent stochastic processes identically distributed with mean function $\mu_2^{\text{cross}}(t)$ and covariance $\gamma_2^{\text{cross}}(s, t)$. Note a group-specific covariance is assumed. Regarding the crosswind component in absolute value,

a pointwise two-sample test should be written:

$$\begin{aligned} H_{0,t} : \mu_1^{\text{cross}}(t) &= \mu_2^{\text{cross}}(t) \\ &\text{against} \\ H_{1,t} : \mu_1^{\text{cross}}(t) &< \mu_2^{\text{cross}}(t) \end{aligned}$$

for a fixed $t \in [0, 1]$.

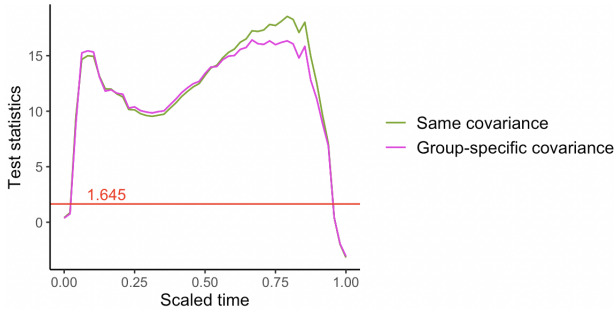


Fig. 9 Pointwise values of the test statistic for 50 values of $t \in [0, 1]$. The horizontal red line is the quantile that should be used to reject the null hypothesis at 5% level when no correction for multiple comparisons is taken into account. The test statistics are computed using asymmetric kernel smoothing with the directional approach.

Figure 9 shows that the delay group has an average absolute value of the crosswind component that is greater than the on time group except during landing and takeoff. Assuming a group-specific covariance lowers the test statistics at high altitudes. Again, results are similar to the component approach. If one looks at the test statistics obtained with constrained splines approach, the global shape is the same. Yet, it worth mentioning that one may reject the null at 5% for the takeoff and landing in this case. The splines and kernel approaches may lead to different conclusions at the boundaries.

A joint approach

We have two bivariate functional samples of wind component values. The one-time sample $\{(Y_{1,i}^{\text{tail}} \ Y_{1,i}^{\text{cross}})\}_{i=1}^{n_1}$ is assumed to be made of independent bivariate stochastic processes identically distributed with mean vector function $\mu_1(t)$ and covariance matrix function $\Gamma(s, t)$. The sample for the delayed group $\{(Y_{2,j}^{\text{tail}} \ Y_{2,j}^{\text{cross}})\}_{j=1}^{n_2}$ is made of independent bivariate stochastic processes identically distributed with mean vector function $\mu_2(t)$ and covariance matrix function $\Gamma(s, t)$. The pointwise two-sample joint test can be formulated as follows:

$$\begin{aligned} H_{0,t} : \mu_1(t) &= \mu_2(t) \\ &\text{against} \\ H_{1,t} : \mu_1^{\text{tail}}(t) &> \mu_2^{\text{tail}}(t) \text{ and } \mu_2^{\text{cross}}(t) > \mu_1^{\text{cross}}(t) \end{aligned}$$

A pivotal test statistic, for each $t \in [0, 1]$ is

$$z^{\text{joint}}(t) = \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \hat{\Psi}^{-\frac{1}{2}} \left[\begin{pmatrix} \hat{\mu}_1^{\text{tail}} - \hat{\mu}_2^{\text{tail}} \\ \hat{\mu}_2^{\text{cross}} - \hat{\mu}_1^{\text{cross}} \end{pmatrix} - \begin{pmatrix} \mu_1^{\text{tail}} - \mu_2^{\text{tail}} \\ \mu_2^{\text{cross}} - \mu_1^{\text{cross}} \end{pmatrix} \right]$$

where $\hat{\Psi}$ is an estimator of $\Psi = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \odot \Gamma(t, t)$, $\odot$ being the Hadamard product. Under $H_{0,t}$, with similar assumptions as above, the test statistic is asymptotically bivariate Gaussian for any fixed $t \in [0, 1]$. Details can be found in Appendix A. In practice, we choose, for each $t \in [0, 1]$, under the null:

$$\|z^{\text{joint}}(t)\|_2^2 \xrightarrow[\min(n_1, n_2) \rightarrow \infty]{d} \chi_2^2.$$

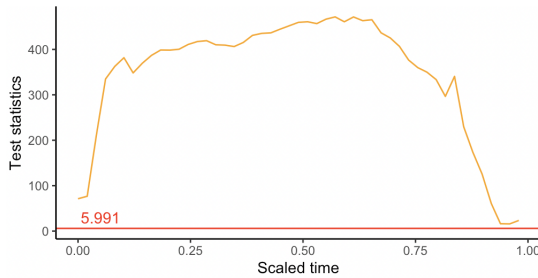


Fig. 10 Pointwise values of the test statistic for 50 values of $t \in [0, 1]$. The horizontal red line is the quantile that should be used to reject the null hypothesis at 5% level when no correction for multiple comparisons is taken into account. The test statistics are computed using asymmetric kernel smoothing with the directional approach.

Figure 10 shows that the delay group has, on average, less favorable wind conditions. Again, results are very similar with component approach. Constrained splines smoothing does not change the results.

6 Conclusion

In this paper, a multivariate functional data framework is used to model aircraft trajectories. Thanks to the constraint smoothing strategies, altitude and wind speed are guaranteed to be non-negative. Smoothed altitude drops to zero at the beginning and at the end of the flight.

Provided a good choice of bandwidth and roughness parameters, asymmetric kernel smoothing and constrained B-spline smoothing give similar results from a visual perspective.

It happens that the circular nature of the wind direction values must be handled carefully. Fortunately, it is also possible to depart from smoothed wind components to get similar results. Note that wind components are not always available. In ADS-B data, wind speed and direction are directly given.

Notably, a pointwise test allows to identify the time of the trajectory at which the wind was significantly less favorable for the delay group.

Promising results may stem from the developed multivariate functional data framework. For example, one may associate fuel consumption or noise emission values to each point of the flight instead of wind values. Doing so, it would be possible to check that one group has consumed more fuel.

Several exciting aspects fall outside the scope of this paper. A monthly study may help to detect any seasonal pattern. From a statistical point of view, the independence assumption that is used for the pointwise tests could be relaxed. It is likely that wind profiles are temporally correlated based on the takeoff time. One would deal with functional time series. Because departures are not scheduled on a regular time scale, relaxing the independence assumption requires some additional work.

Appendix A Asymptotic distribution

From the Central Limit Theorem, for each $t \in [0, 1]$, we have:

$$\sqrt{n_1} \left[\left(\frac{\frac{1}{n_1} \sum_{i=1}^{n_1} Y_{1,i}^{\text{tail}}(t)}{\frac{1}{n_1} \sum_{i=1}^{n_1} Y_{1,i}^{\text{cross}}(t)} \right) - \left(\frac{\mu_1^{\text{tail}}(t)}{\mu_1^{\text{cross}}(t)} \right) \right] \xrightarrow[n_1 \rightarrow \infty]{d} \mathcal{N}_2(0, \Gamma(t, t))$$

and

$$\sqrt{n_2} \left[\left(\frac{\frac{1}{n_2} \sum_{j=1}^{n_2} Y_{2,j}^{\text{tail}}(t)}{\frac{1}{n_2} \sum_{j=1}^{n_2} Y_{2,j}^{\text{cross}}(t)} \right) - \left(\frac{\mu_2^{\text{tail}}(t)}{\mu_2^{\text{cross}}(t)} \right) \right] \xrightarrow[n_2 \rightarrow \infty]{d} \mathcal{N}_2(0, \Gamma(t, t)).$$

With the assumptions of [Zhang \(2013\)](#) (Theorem 5.2, page 134), we have:

$$\sqrt{\frac{n_1 n_2}{n_1 + n_2}} \left(\frac{\overline{Y_1^{\text{tail}}(t)} - \overline{Y_2^{\text{tail}}(t)}}{\overline{Y_1^{\text{cross}}(t)} - \overline{Y_2^{\text{cross}}(t)}} \right) \xrightarrow[\min(n_1, n_2) \rightarrow \infty]{d} \mathcal{N}_2(\mu_1(t) - \mu_2(t), \Gamma(t, t)).$$

Let $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be the function $f(x, y) = (x, -y)$ with Jacobian matrix $J = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. For $\delta(t) = \mu_1(t) - \mu_2(t)$, let $J_{\delta(t)}$ denotes the differential of f at $\delta(t) \in \mathbb{R}^2$. As $J_{\delta(t)} \neq 0$, the Multivariate Delta Method gives:

$$\sqrt{\frac{n_1 n_2}{n_1 + n_2}} \Psi^{-\frac{1}{2}} \left[\begin{pmatrix} \hat{\mu}_1^{\text{tail}} - \hat{\mu}_2^{\text{tail}} \\ \hat{\mu}_2^{\text{cross}} - \hat{\mu}_1^{\text{cross}} \end{pmatrix} - \begin{pmatrix} \mu_1^{\text{tail}} - \mu_2^{\text{tail}} \\ \mu_2^{\text{cross}} - \mu_1^{\text{cross}} \end{pmatrix} \right] \xrightarrow[\min(n_1, n_2) \rightarrow \infty]{d} \mathcal{N}_2(0, I_2)$$

where $\Psi \equiv \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \odot \Gamma(t, t)$. Let $\hat{\Psi}$ be the estimator of Ψ involving the pooled covariance matrix $\hat{\Gamma}(t, t) = \frac{n_1 \hat{\Gamma}_1(t, t) + n_2 \hat{\Gamma}_2(t, t)}{n_1 + n_2}$. Using the law of large numbers and Slutsky's theorem, one shows that pooled covariance matrix converges to

the covariance matrix $\Gamma(t, t)$ in probability. Again using Slutsky's theorem,

$$\sqrt{\frac{n_1 n_2}{n_1 + n_2}} \hat{\Psi}^{-\frac{1}{2}} \left[\begin{pmatrix} \hat{\mu}_1^{\text{tail}} - \hat{\mu}_2^{\text{tail}} \\ \hat{\mu}_2^{\text{cross}} - \hat{\mu}_1^{\text{cross}} \end{pmatrix} - \begin{pmatrix} \mu_1^{\text{tail}} - \mu_2^{\text{tail}} \\ \mu_2^{\text{cross}} - \mu_1^{\text{cross}} \end{pmatrix} \right] \xrightarrow[\min(n_1, n_2) \rightarrow \infty]{d} \mathcal{N}_2(0, I_2).$$

References

- Aneiros G, Cao R, Fraiman R, et al (2019a) Recent advances in functional data analysis and high-dimensional statistics. *Journal of Multivariate Analysis* 170:3–9. <https://doi.org/10.1016/j.jmva.2018.11.007>
- Aneiros G, Cao R, Vieu P (2019b) Editorial on the special issue on Functional Data Analysis and Related Topics. *Computational Statistics* 34(2):447–450. <https://doi.org/10.1007/s00180-019-00892-0>
- Antoniadis A, Gregoire G, McKeague IW (1994) Wavelet Methods for Curve Estimation. *Journal of the American Statistical Association* 89(428):1340–1353. <https://doi.org/10.2307/2290996>
- Benhenni K, Ferraty F, Rachdi M, et al (2007) Local smoothing regression with functional data. *Computational Statistics* 22(3):353–369. <https://doi.org/https://doi.org/10.1007/s00180-007-0045-0>
- de Boor C (2001) *A Practical Guide to Spline*. Springer-Verlag New York Inc.
- Breckling J, Berger J, Fienberg S, et al (eds) (1989) *The Analysis of Directional Time Series: Applications to Wind Speed and Direction*, Lecture Notes in Statistics, vol 61. Springer, New York, NY
- Brown BM, Chen SX (1999) Beta-Bernstein Smoothing for Regression Curves with Compact Support. *Scandinavian Journal of Statistics* 26(1):47–59. URL <https://www.jstor.org/stable/4616540>
- Carroll C, Müller HG, Kneip A (2021) Cross-component registration for multivariate functional data, with application to growth curves. *Biometrics* 77(3):839–851. <https://doi.org/10.1111/biom.13340>
- Carvalho L, Sternberg A, Gonçalves L, et al (2020) On the relevance of data science for flight delay research: a systematic review. *Transport Reviews* 41(4):499–528. <https://doi.org/https://doi.org/10.1080/01441647.2020.1861123>
- Chen SX (1999) Beta kernel estimators for density functions. *Computational Statistics & Data Analysis* 31(2):131–145. [https://doi.org/https://doi.org/10.1016/S0167-9473\(99\)00010-9](https://doi.org/https://doi.org/10.1016/S0167-9473(99)00010-9)

- Chen SX (2000) Beta kernel smoothers for regression curves. *Statistica Sinica* 10(1):73–91
- Cheng MY, Fan J, Marron JS (1997) On Automatic Boundary Corrections. *The Annals of Statistics* 25(4):1691–1708. <https://doi.org/https://doi.org/10.1214/aos/1031594737>
- Cox DD, Lee JS (2008) Pointwise Testing with Functional Data Using the Westfall-Young Randomization Method. *Biometrika* 95(3):621–634. <https://doi.org/https://doi.org/10.1093/biomet/asn021>
- Craven P, Wahba G (1978) Smoothing noisy data with spline functions. *Numerische Mathematik* 31(4):377–403. <https://doi.org/10.1007/BF01404567>
- Dai X, Müller HG (2018) Principal component analysis for functional data on Riemannian manifolds and spheres. *The Annals of Statistics* 46(6B):3334–3361. <https://doi.org/10.1214/17-AOS1660>
- Delahaye D, Puechmorel S, Tsiotras P, et al (2014) Mathematical Models for Aircraft Trajectory Design: A Survey. In: *Air Traffic Management and Systems*. Springer Japan, Tokyo, *Lecture Notes in Electrical Engineering*, pp 205–247, https://doi.org/10.1007/978-4-431-54475-3_12
- Di Marzio M, Panzera A, Taylor CC (2013) Non-parametric Regression for Circular Responses. *Scandinavian Journal of Statistics* 40(2):238–255. <https://doi.org/https://doi.org/10.1111/j.1467-9469.2012.00809.x>
- Faraway J (1997) Regression analysis for a functional response. *Technometrics* 39(3):254–261. <https://doi.org/https://doi.org/10.2307/1271130>
- Ferraty F, Vieu P (2006) *Nonparametric Functional Data Analysis: Theory and Practice*. Springer Series in Statistics
- Fisher NI, Lee AJ (1994) Time Series Analysis of Circular Data. *Journal of the Royal Statistical Society Series B (Methodological)* 56(2):327–339. <https://doi.org/https://doi.org/10.1111/j.2517-6161.1994.tb01981.x>
- Fisher R (1953) Dispersion on a Sphere. *Proceedings of the Royal Society of London Series A* 217:295–305. <https://doi.org/10.1098/rspa.1953.0064>
- Gasser T, Kneip A (1995) Searching for Structure in Curve Samples. *Journal of the American Statistical Association* 90(432):1179–1188. <https://doi.org/https://doi.org/10.2307/2291510>
- Hall P, Van Keilegom I (2007) Two-Sample Tests in Functional Data Analysis Starting From Discrete Data. *Statistica Sinica* 17(4):1511–1531. URL <https://doi.org/10.1214/07-SS117>

[//www.jstor.org/stable/24307686](https://www.jstor.org/stable/24307686)

- Hirukawa M (2018) Asymmetric kernel smoothing: theory and applications in economics and finance. SpringerBriefs in statistics, JSS research series in statistics, Springer, Singapore
- Holm S (1979) A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics* 6(2):65–70. <https://doi.org/https://www.jstor.org/stable/4615733>
- Horváth L, Kokoszka P (2012) Inference for Functional Data with Applications, Springer Series in Statistics, vol 200. Springer-Verlag New York Inc.
- Horváth L, Kokoszka P, Reeder R (2013) Estimation of the mean of functional time series and a two-sample problem. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 75(1):103–122. <https://doi.org/10.1111/j.1467-9868.2012.01032.x>
- Hsing T, Eubank R (2015) Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators, John Wiley & Sons edn. Wiley Series in Probability and Statistics
- Jarry G, Delahaye D, Nicol F, et al (2020) Aircraft atypical approach detection using functional principal component analysis. *Journal of Air Transport Management* 84:101,787. <https://doi.org/10.1016/j.jairtraman.2020.101787>
- Jupp PE, Mardia KV (1999) Directional Statistics, 1st edn. Wiley, Chichester
- Kneip A, Gasser T (1992) Statistical Tools to Analyze Data Representing a Sample of Curves. *The Annals of Statistics* 20(3):1266–1305. <https://doi.org/https://doi.org/10.1214/aos/1176348769>
- Kokoszka P, Reimherr M (2017) Introduction to Functional Data Analysis. Chapman and Hall/CRC
- Ley C, Verdebout T (2018) Applied Directional Statistics: Modern Methods and Case Studies, 1st edn. Chapman and Hall/CRC, Boca Raton London New York
- Liu D, Xiao N, Zhang Y, et al (2020) Unsupervised Flight Phase Recognition with Flight Data Clustering based on GMM. In: 2020 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), pp 1–6, <https://doi.org/10.1109/I2MTC43012.2020.9128596>

- Liu Y, Liu Y, Hansen M, et al (2019) Using machine learning to analyze air traffic management actions: Ground delay program case study. *Transportation Research Part E: Logistics and Transportation Review* 131:80–95. <https://doi.org/10.1016/j.tre.2019.09.012>
- Liu Y, Hansen M, Ball MO, et al (2021) Causal analysis of flight en route inefficiency. *Transportation Research Part B: Methodological* 151:91–115. <https://doi.org/10.1016/j.trb.2021.07.003>
- Marron JS, Dryden IL (2021) *Object Oriented Data Analysis*. Chapman and Hall/CRC
- Morettin PA, Pinheiro A, Vidakovic B (2017) *Wavelets in Functional Data Analysis*, 1st edn. Springer
- Mueller E, Chatterji G (2002) Analysis of aircraft arrival and departure delay characteristics. Los Angeles, California, <https://doi.org/10.2514/6.2002-5866>, aIAA Aviation Technology, Integration, and Operations (ATIO) Technical, Los Angeles, California.
- Nicol F (2017) Statistical Analysis of Aircraft Trajectories: a Functional Data Analysis Approach. *Alldata 2017, The Third International Conference on Big Data, Small Data, Linked Data and Open Data*, p pp.51
- Oliveira M, Crujeiras RM, Rodríguez-Casal A (2014) NPCirc: An R Package for Nonparametric Circular Methods. *Journal of Statistical Software* 61:1–26. <https://doi.org/10.18637/jss.v061.i09>
- Park J, Ahn J (2017) Clustering multivariate functional data with phase variation. *Biometrics* 73(1):324–333. <https://doi.org/10.1111/biom.12546>
- Pejovic T, Williams VA, Noland RB, et al (2009) Factors Affecting the Frequency and Severity of Airport Weather Delays and the Implications of Climate Change for Future Delays. *Transportation Research Record* 2139(1):97–106. <https://doi.org/10.3141/2139-12>
- Pigoli D, Sangalli LM (2012) Wavelets in functional data analysis: Estimation of multidimensional curves and their derivatives. *Computational Statistics & Data Analysis* 56(6):1482–1498. <https://doi.org/10.1016/j.csda.2011.12.016>
- Puechmorel S, Delahaye D (2007) 4D trajectories : a functional data perspective. *IEEE*, p pp 1.C.6, <https://doi.org/10.1109/DASC.2007.4391832>
- Pérez-Rodríguez J, Pérez-Sánchez J, Gómez-Déniz E (2017) Modelling the asymmetric probabilistic delay of aircraft arrival. *Journal of Air Transport Management* 62:90–98. <https://doi.org/10.1016/j.jairtraman.2017.03.001>

- Ramsay JO, Silverman BW (2002) *Applied Functional Data Analysis: Methods and Case Studies*. Springer-Verlag New York Inc.
- Ramsay JO, Silverman BW (2005) *Functional data analysis*, 2nd edn. Springer series in statistics, Springer-Verlag New York Inc.
- Sangalli LM, Secchi P, Vantini S, et al (2009) Efficient estimation of three-dimensional curves and their derivatives by free-knot regression splines, applied to the analysis of inner carotid artery centrelines. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 58(3):285–306. <https://doi.org/10.1111/j.1467-9876.2008.00653.x>
- Schimek MG (2013) *Smoothing and Regression: Approaches, Computation, and Application*, 1st edn. Wiley
- Shen Q, Faraway J (2004) A F Test For Linear Models With Functional Responses. *Statistica Sinica* 14(4):1239–1257. URL <https://www.jstor.org/stable/24307230>
- Shi J, Song W (2016) Asymptotic results in gamma kernel regression. *Communications in Statistics - Theory and Methods* 45(12):3489–3509. <https://doi.org/10.1080/03610926.2014.890225>
- Silverman BW (1995) Incorporating Parametric Effects into Functional Principal Components Analysis. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 57(4):673–689. <https://doi.org/10.1111/j.2517-6161.1995.tb02055.x>
- Srivastava A, Klassen EP (2016) *Functional and Shape Data Analysis*, 1st edn. Springer-Verlag New York Inc.
- Srivastava A, Klassen E, Joshi SH, et al (2011a) Shape Analysis of Elastic Curves in Euclidean Spaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33(7):1415–1428. <https://doi.org/10.1109/TPAMI.2010.184>
- Srivastava A, Wu W, Kurtek S, et al (2011b) Registration of Functional Data Using Fisher-Rao Metric ArXiv: 1103.3817
- Sun J (2021) *The 1090 Megahertz Riddle: A Guide to Decoding Mode S and ADS-B Signals*. TU Delft OPEN
- Sun J, Ellerbroek J, Hoekstra JM (2016) Large-Scale Flight Phase Identification from ADS-B Data Using Machine Learning Methods. 7th International Conference on Research in Air Transportation

- Turlach BA (2005) Shape constrained smoothing using smoothing splines. *Computational Statistics* 20(1):81–104. <https://doi.org/https://doi.org/10.1007/BF02736124>
- Ugwuowo FI, Udokang AE (2022) Modelling Circular Time Series with Applications. In: SenGupta A, Arnold BC (eds) *Directional Statistics for Innovative Applications: A Bicentennial Tribute to Florence Nightingale*. Forum for Interdisciplinary Mathematics, p 407–424
- Ullah S, Finch CF (2013) Applications of functional data analysis: A systematic review. *BMC Medical Research Methodology* 13:43. <https://doi.org/10.1186/1471-2288-13-43>
- Valderrama MJ (2007) An overview to modelling functional data. *Computational Statistics* 22(3):331–334. <https://doi.org/10.1007/s00180-007-0043-2>
- Wang JL, Chiou JM, Müller HG (2016) Functional Data Analysis. *Annual Review of Statistics and Its Application* 3(1):257–295. <https://doi.org/10.1146/annurev-statistics-041715-033624>
- Wang Q (2021) Two-sample inference for sparse functional data. *Electronic Journal of Statistics* 15(1). <https://doi.org/https://doi.org/10.1214/21-EJS1802>
- Wang Y, Kulkarni D (2011) Modeling Weather Impact on Ground Delay Programs. *SAE International Journal of Aerospace* 4:1207–1215. <https://doi.org/https://doi.org/10.4271/2011-01-2680>
- Wang Y, Li MZ, Gopalakrishnan K, et al (2022) Timescales of delay propagation in airport networks. *Transportation Research Part E: Logistics and Transportation Review* 161:102,687. <https://doi.org/10.1016/j.tre.2022.102687>
- Yu B, Guo Z, Asian S, et al (2019) Flight delay prediction for commercial air transport: A deep learning approach. *Transportation Research Part E: Logistics and Transportation Review* 125:203–221. <https://doi.org/10.1016/j.tre.2019.03.013>
- Zhang JT (2013) *Analysis of Variance for Functional Data*, 1st edn. Chapman and Hall/CRC
- Zhang JT, Chen J (2007) Statistical inferences for functional data. *The Annals of Statistics* 35(3). <https://doi.org/https://doi.org/10.1214/009053606000001505>
- Zhang X, Wang JL (2016) From sparse to dense functional data and beyond. *The Annals of Statistics* 44(5):2281–2321. <https://doi.org/10.1214/>

