Statistical Analysis of Aircraft Trajectories in a Multivariate Functional Data Analysis Framework

Remi Perrichon, Xavier Gendre, Thierry Klein

Statistical Analysis of Aircraft Trajectories in a Multivariate Functional Data Analysis Framework

Rémi Perrichon

ENAC - Ecole Nationale de l'Aviation Civile. Universite de Toulouse, France.

E-mail: remi.perrichon@enac.fr

Xavier Gendre

ISAE SUPAERO. Université de Toulouse, France.

Institut de Mathématiques de Toulouse UMR5219. Université de Toulouse, France.

Thierry Klein

ENAC - Ecole Nationale de l'Aviation Civile. Universite de Toulouse, France.

Institut de Mathématiques de Toulouse UMR5219. Université de Toulouse, France.

Summary. While advanced inference and registration methods for functional data analysis have recently been developed in the literature, statistical analyses of aircraft trajectories have remained scarce, despite operational relevance. Using more than 3,000 trajectories matched with weather data, this paper explores how the experienced wind speed can be associated to en-route delays in a multivariate functional data analysis framework. The processing of aircraft trajectories is challenging as it requires constraint smoothing and rescaling. The paper emphasizes that the choice of the registration strategy influences further inference. Five scenarios are developed to compare registration strategies and find the most suited one for a pointwise functional two-sample test of means.

Keywords: Functional data analysis, curve registration, constraint smoothing, two-sample tests, aircraft trajectories

1. Introduction

To date, the statistical analysis of aircraft trajectories has been under-exploited in the Airspace Traffic Management (ATM) literature. This assessment was early made by

Puechmorel and Delahaye (2007), promoting the use of Functional Data Analysis (FDA) to study aircraft trajectories. Concurrently, FDA has experienced substantial growth and development in recent years. From the early work of Ramsay and Silverman (2005), some topics have gained visibility such as inference procedures developed by Horváth and Kokoszka (2012) and the elegant curve registration geometric framework of Srivastava and Klassen (2016). Statistical analyses of bird or hurricane trajectories exploiting these advanced frameworks have been made by Su et al. (2014). 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). The key point of this paper is to illustrate how advanced statistical frameworks are relevant to study operational topics. More specifically, the paper focuses on testing the association between en-route delays and wind speed, the working assumption being that delayed flights have experienced a stronger wind. 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, en-route and during taxi in. In order to provide some means of 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. Delay analysis is made complex by the fact that it is always cheaper to take delay on the ground than to impose a delay by speed control, by holding, or by re-clearing the aircraft to non-optimum flight levels. To put it differently, delays may be the result of a planned process. For instance, the Federal Aviation Administration (FAA) institutes Ground-Delay Programs (GDP) to delay flights before they depart from their originating airports. The effect of weather conditions on the characteristics of GDP events is investigated by Wang and Kulkarni (2011), and more recently by Liu et al. (2019). 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. For its part, atmospheric wind plays a mixed role in aircraft operations as it may have detrimental effects (delays, accidents) or beneficial ones. When detrimental, the magnitude of these effects has recently been quantified using econometric techniques by Borsky and Unterberger (2019). As Carvalho et al. (2020) shows, most contributions in the literature focus on flight delay prediction. Statistical studies of delays are often descriptive or are made at the airport level as in Mueller and Chatterji (2002), Pejovic et al. (2009) and Pérez–Rodríguez et al. (2017). To our knowledge, statistical inference has never been performed for en-route delays. Crucially, aircraft trajectories have never been the object of inference procedures, despite the great amount of information they contain about the true operating conditions. The aim of this work is to provide a framework in which:

(a) Aircraft trajectories can be associated to experienced weather conditions
(b) The multivariate FDA framework can be used to model aircraft trajectories
(c) Data transformations such as registration can be carefully done taking into account further inference

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. The resulting data set includes more than 3,000 flights departing from Toulouse-Blagnac and landing at Paris-Orly for the year 2015. Adding new airports or more recent years to the data set is almost effortless once the weather data have been downloaded for a given area, which in turn requires some storage capacity. Section 3 presents how the multivariate FDA framework can be used to model trajectories. Interestingly, this framework should take into account that some dimensions of a trajectory are physically constrained. For instance, it is expected from the smoothing step that altitude must be positive everywhere as well as null at the departure and arrival. This is why a focus is made on constrained smoothing in Section 3.4. Additionally, for a given trajectory, the longitude, the latitude, the altitude and the wind speed are not independent. For example, because of atmospheric layers, altitude and wind speed are associated. A cross-correlation analysis is detailed in Section 3.5. Because trajectories are never of same duration, a rescaling operation is necessary to compare trajectories. This basic homothety may have detrimental effects for further inference. Rationally, this problem is tackled as a registration issue. Within the framework developed in Srivastava and Klassen (2016), two strategies are compared on illustrative scenarios in Section 4. An original registration strategy seems to produce a desirable ground for the testing step. Finally, Section 5 presents the testing procedures and results.

2. Empirical background

To demonstrate the practical efficiency of a statistical analysis of aircraft trajectories, a data set is to be built. The latter has to rely the usual dimensions of an aircraft trajectory (longitude, latitude, altitude, time) - the so-called "4D trajectory" - to additional weather dimensions. To achieve so, two data sets are considered. Departing from the R&D data provided by Eurocontrol, weather dimensions are added. The two data sets are briefly introduced and the matching strategy is then presented.

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 four years of data, that is to say to more than 14 million flights as of April 2021. 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. It includes the last-filed flight plans and the actual route, the airspace and the route network that was in place at that time.

Only two data subsets are used in this work. The flights metadata and the actual flight points are merged thanks to a unique numeric identifier to produce a data set of raw trajectories. Each point of a trajectory is 4D as it includes the time, the crossed flight levels, the latitude and the longitude.

The duration of a flight is defined as being the time difference in seconds between the first actual point of the flight and the last one. The centered duration of a given flight is defined as being the centred duration (in seconds) with respect to the empirical mean duration.

Depending on the needs, two groups can easily be made using the centered duration: the group of delayed and on time flights, always relative to the observed mean duration of the year for the air link of interest. Figure 1 gives an example of the two groups made according to the 8^{th} decile.

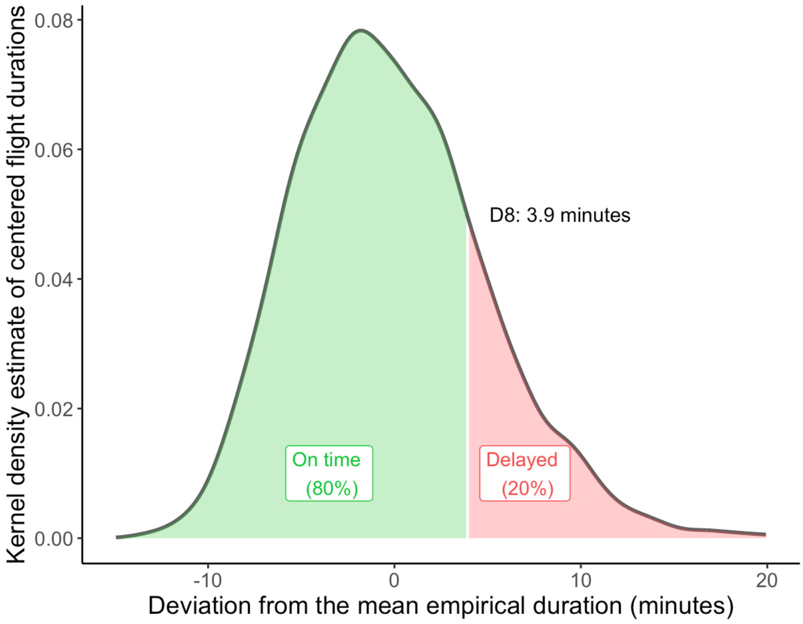

Fig. 1. Empirical density (Gaussian smoothing) of centered durations for the flights between Toulouse-Blagnac and Paris-Orly (2015). D8 indicates the value of the 8^{th} decile.

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. ERA5 provides hourly estimates for a large number of atmospheric, ocean-wave and land-surface quantities. Data has been regridded to a regular longitude-latitude grid of 0.25 degrees for the reanalysis. There are four main subsets: hourly and monthly products, both on pressure levels (upper air fields) and single levels (atmospheric, ocean-wave and land surface quantities). The subset of interest for trajectories is entitled "ERA5 hourly data on pressure levels from 1979 to present".

Among all weather variables, only two 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. Given the u and v components, the wind speed is computed as $\sqrt{u^2 + v^2}$.

2.3. Towards augmented trajectories

Departing from a 4D trajectory (longitude, latitude, altitude, time), the goal is to add weather dimensions to obtain a so-called augmented trajectory. To associate each point of a trajectory to a weather value, the matching should be done taking the closest altitude, the closest position (longitude, latitude) and the nearest time.

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 $288.15°K$. In what follows, the matching step is assumed to be perfect (no errors to be taken into account in the smoothing step). This can be refined in a future work.

2.4. Scope

Dealing with weather data ranging from -5 to 9 degrees in longitude, and 42 to 52 degrees in latitude, one is interested in studying flights departing from Toulouse-Blagnac and landing at Paris-Orly in 2015. Flights departing from Paris-Orly and landing at Toulouse-Blagnac in 2015 are not taken into account to avoid the implicit assumption that outbound flights are comparable with return flights in terms of trajectories. A summary of the data cleaning process is given in Section 5.2.

3. A multivariate functional data framework to model trajectories

3.1. The univariate functional data perspective

For $g = 1, 2$, let $X_g \equiv \{X_g(t), t \in [0, 1]\}$ be a L^2 stochastic process taking values in $\mathbb{R}$ with mean function $\mu_g(t)$ and covariance function $\gamma(s, t)$, both unknown in practice. It is assumed that X_1 and X_2 are independent. For aircraft data, the two groups are the on time and the delay groups presented above. Let $X_{g,i}$ be iid copies of X_g and $\varepsilon_{g,i}$ be iid copies of ε which is a stochastic process with a null mean and a covariance function given by $\gamma_\varepsilon(s, t) = \sigma^2(t)1_{\{s=t\}}$. For $g = 1, 2$, X_g and ε are assumed to be independent. Following Zhang and Chen (2007) and Wang (2021), a functional data set is modeled as independent realizations of an underlying stochastic process: $Y_{g,i}(t) = X_{g,i}(t)+\varepsilon_{g,i}(t)$. It is not fully observed in practice. Rather, for process i in group g, there are $M_{g,i}$ random observing times denoted $T_{g,i,1}, ..., T_{g,i,M_{g,i}}$. If the observing times are modeled as being deterministic, they may be regularly-spaced on $[0, 1]$ or not. Yet, if random, they are assumed to be at least iid with a bounded density function within [0,1]. Every process is observed at these random times such that $Y_{g,i}(T_{g,i,j}) = X_{g,i}(T_{g,i,j}) + \varepsilon_{g,i,j}(T_{g,i,j})$ where $j \in \{1, ..., M_{g,i}\}$. It is assumed that all is independent from $M_{g,i}$. In practice, a fixed design is used, meaning that all is done conditionally with respect to random observing times.

When observing times are not regularly-spaced but are numerous enough for each subject, a smoothing approach is often performed. The strategy is to reconstruct curves.

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. 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 all groups.

Historically, this approach has prevailed in the literature following the work of Ramsay and Silverman (2002) and Ramsay and Silverman (2005). It 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.

If 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 have recently been proposed by Wang (2021).

In this work, trajectories are assumed to be detailed enough for individual curves to be reconstructed. The average number of observed points in a trajectory is 14. The median is also 14.

3.2. A multivariate functional data perspective on augmented trajectories

By nature, an aircraft trajectory is a multivariate object. Let N_{Total} be the total number of flights that can be divided into N on time trajectories and M delayed trajectories such that $N_{\text{Total}} = N + M$.

For each flight $i = 1, ..., N_{\text{Total}}$ an augmented trajectory is

$$\left\{ \left(\boldsymbol{y_{g,i,j}}, t_{g,i,j} \right), j = \{1, ..., m_{g,i}\} \right\}$$

where $\boldsymbol{y_{g,i,j}}$ is a four-dimensional vector, $t_{g,i,j}$ a timestamp and $m_{g,i}$ the number of points in trajectory i in group g. The components of $y_{g,i,j}$ are denoted $y_{g,i,j,x}$, $y_{g,i,j,y}$, $y_{g,i,j,z}$, $y_{g,i,j,w}$, respectively referring to the longitude value for point j in trajectory i in group g, the latitude, the altitude and the wind speed. For two points in a given trajectory, the time gap between two points depends on the chosen points and on the trajectory itself. In other words, the sampling is not regular within a trajectory, nor between two trajectories. For all four dimensions of a given trajectory (longitude, latitude, altitude and wind speed), observing times are the same.

In the case of aircraft trajectories, the pilot follows a flight plan. Observed augmented trajectories hide a smooth underlying process. In this context, FDA offers a rigorous framework to perform inference. For each dimension (longitude, latitude, altitude, wind speed), observed augmented trajectories are modeled as independent realizations of an underlying stochastic process, as described above. In the multivariate setting, there are now four dimensions.

$$\boldsymbol{Y_{g,i}} : \begin{array}{ccc} [0,1] & \rightarrow & [-180,180] \times [-90,90] \times \mathbb{R}^+ \times \mathbb{R}^+ \\ t & \mapsto & (Y_{g,i,x}(t), Y_{g,i,y}(t), Y_{g,i,z}(t), Y_{g,i,w}(t))' \end{array}$$

3.3. Smoothing

Many smoothing methods are available to reconstruct individual trajectories. A review of four popular non-parametric techniques is given by Zhang (2013).

Whatever the method, the construction of the functional observations using the discrete data takes place separately for each flight and for each dimension of a flight: the longitude, the latitude, the altitude and the wind speed. As the regression B-splines strategy is adopted, the order and interior knots are chosen to be the same for all trajectories. Of course, choices can be different for each dimension. The B-spline basis system developed by de Boor (2001) is used.

For a given dimension (for instance the longitude), the latent function is estimated by some projection on a linear functional space spanned by K_x known basis functions $\phi_{x,1}, \ldots, \phi_{x,K_x}$

$$X_{g,i,x}(t) = \sum_{k=1}^{K_x} \theta_{g,i,x,k}\phi_{x,k}(t) = \boldsymbol{\theta'_{g,i,x}} \boldsymbol{\phi_x}(t)$$

where $X_{g,i,x}$ denotes the smooth function for the longitude dimension associated with flight i in group g and $\theta_{g,i,x,k}$ are some coefficients to estimate. In matrix form, $\boldsymbol{\theta_{g,i,x}}$ is a vector of length K_x that contains the coefficients and $\boldsymbol{\phi_x}(t)$ is a vector of length K_x containing the basis functions.

The choice of K_x is driven by a bias-variance trade-off. Note that the choice of the number of basis functions may change depending on the intrinsic variability: more basis functions may be needed to correctly smooth some dimensions.

Estimating the coefficients $\theta_{g,i,x,k}$ can be viewed as a least squares problem, for each flight and each dimension. In the least squares approach, the number of observing times should higher than the number of coefficients. Yet, a high number of B-spline functions may be needed for the smoothing to be satisfactory. In this case, the roughness penalty approach is relevant to avoid singularity problems on the computational side. This approach allows to use a large number of basis functions (bigger than the number of observing times) and to penalize some measure of function complexity. This is what is called regularization. If one focuses on the observed longitude point $y_{g,i,j,x}$ for every flight i in group g, the following optimization problem is considered

$$\underset{\theta_{g,i,x,k}}{\operatorname{argmin}} \sum_{j=1}^{m_{g,i}} \left[y_{g,i,j,x} - \sum_{k=1}^{K_z} \theta_{g,i,x,k} \phi_{x,k}(t_{g,i,j}) \right]^2 + \lambda \int_0^1 \left[D^2 \left(\boldsymbol{\theta'_{g,i,x}} \boldsymbol{\phi_x}(t) \right) \right]^2 dt$$

where D^2 is the second derivative and λ a smoothing parameter. Once the coefficients are estimated, the irregularity of the sampling is no longer a problem. A resampling on a regular grid can be done easily: this is the main advantage of the FDA framework.

3.4. Constrained smoothing

Regarding the altitude and the wind speed, the smoothing problem should be constrained: data are collected on functions that are strictly positive. A positive smoothing function s^+ can always be defined as the exponential of an unconstrained function s: $s^+(t) = exp\{s(t)\}$. The optimization problem for the altitude is then

$$\underset{\theta_{g,i,z,k}}{\operatorname{argmin}} \sum_{j=1}^{m_{g,i}} \left[y_{g,i,j,z} - exp \left\{ \sum_{k=1}^{K_z} \theta_{g,i,z,k} \phi_{z,k}(t_{g,i,j}) \right\} \right]^2 + \lambda \int_0^1 \left[D^2 \left(\boldsymbol{\theta'_{g,i,z}} \boldsymbol{\phi_z}(t) \right) \right]^2 dt.$$

Numerical methods are used to minimize the criterion. As the altitude is zero at both the departure and the arrival, that is to say at boundaries, the values in that region are poorly defined large negative numbers.

The above strategy does not ensure that the altitude is going to be zero at the departure ($t = 0$) and at the arrival ($t = 1$), even if one knows that it must be the case for an aircraft. Additionally, one wishes to smooth the longitude and the latitude such that for all trajectories, the first longitude-latitude values are the one of the departure airport. Similarly, the last longitude-latitude must be the one of the arrival airport.

These boundary constrains have not received much attention in the literature. They can be taken into account proceedings as follows. Let the Left Trapeze Function (LTF) be

$$LTF_\varepsilon(t) \equiv \begin{cases} \frac{t}{\varepsilon} & t \in [0, \varepsilon[\\ 1 & t \in [\varepsilon, 1] \end{cases}$$

for $t \in [0, 1]$ and $\varepsilon \in]0, \frac{1}{2}]$; the Right Trapeze Function (RTF) be

$$RTF_\varepsilon(t) \equiv \begin{cases} 1 & t \in [0, 1-\varepsilon] \\ \frac{-t}{\varepsilon} + \frac{1}{\varepsilon} & t \in]1-\varepsilon, 1] \end{cases}$$

for $t \in [0, 1]$ and $\varepsilon \in]0, \frac{1}{2}]$; and the Trapeze Function (TF) be

$$TF_\varepsilon(t) \equiv \inf\{LTF(t), RTF(t)\}$$

For a given trajectory i, the reconstructed functions are then estimated.

$$\widehat{\boldsymbol{Y_{g,i}}}(t) = \begin{pmatrix} \widehat{Y_{g,i,x}}(t) \\ \widehat{Y_{g,i,y}}(t) \\ \widehat{Y_{g,i,z}}(t) \\ \widehat{Y_{g,i,w}}(t) \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{K_x} \hat{\theta}_{g,i,x,k}\phi_{x,k}(t) \times TF_\varepsilon(t) + LTF_\varepsilon(t) \times \mathrm{lon}^D + RTF_\varepsilon(t) \times \mathrm{lon}^A \\ \sum_{k=1}^{K_y} \hat{\theta}_{g,i,y,k}\phi_{y,k}(t) \times TF_\varepsilon(t) + LTF_\varepsilon(t) \times \mathrm{lat}^D + RTF_\varepsilon(t) \times \mathrm{lat}^A \\ exp\left\{\sum_{k=1}^{K_z} \hat{\theta}_{g,i,z,k}\phi_{z,k}(t)\right\} \times TF_\varepsilon(t) \\ exp\left\{\sum_{k=1}^{K_w} \hat{\theta}_{g,i,w,k}\phi_{w,k}(t)\right\} \end{pmatrix}$$

where lon^D and lon^A are respectively denoting the longitude of the departure/arrival airport; lat^D and lat^A are respectively denoting the latitude of the departure/arrival airport. For a small value of ε, the idea is only to smoothly clip the longitude-latitude values to the airports' coordinates. Likewise, the altitude is constrained to be null at the departure and at the arrival. The smoothing is shown on Figures 2 and 3. Now that the functional data framework is used, vizualisation is much more easier as one may note from Figures 4 and 5.

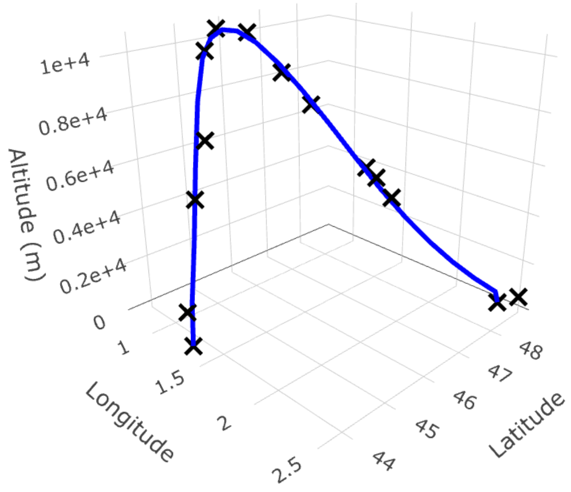

Fig. 2. 3D Visualisation of a flight between Toulouse-Blagnac and Paris-Orly in 2015

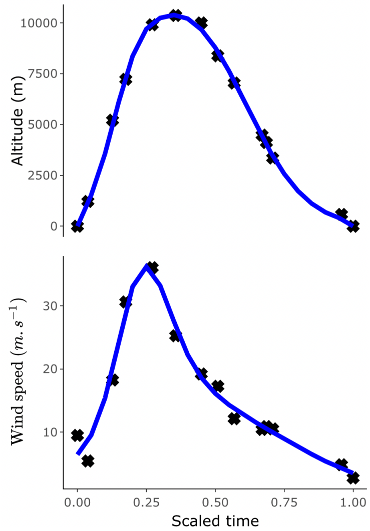

Fig. 3. Altitude (upper) and wind speed (lower) profiles of a flight between Toulouse-Blagnac and Paris-Orly in 2015

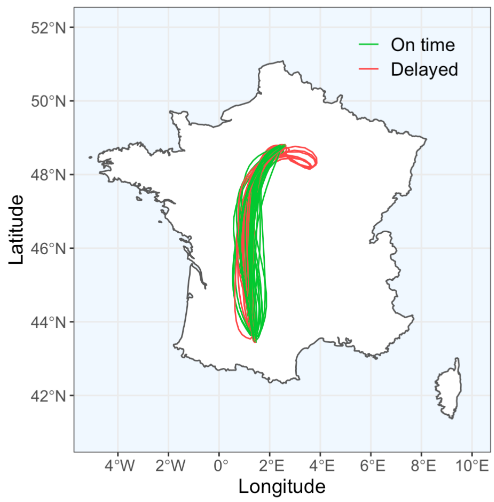

Fig. 4. Visualisation of all flights between Toulouse-Blagnac and Paris-Orly in 2015

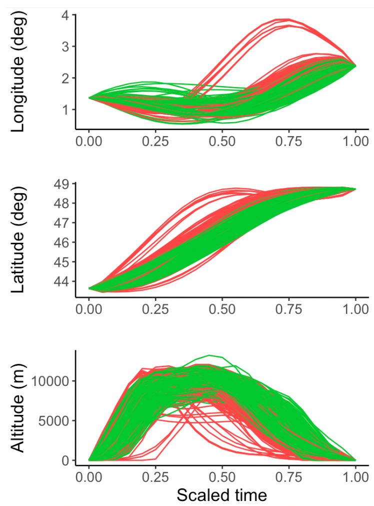

Fig. 5. Longitude (upper), latitude (middle) and altitude (lower) profiles of all flights between Toulouse-Blagnac and Paris-Orly in 2015

The map on Figure 4 clearly shows that some trajectories have a go-around pattern (loop before arriving at Paris-Orly). In aviation, a go-around is an aborted landing of an aircraft that is on final approach. A go-around can either be initiated by the pilot flying or requested by air traffic control for various reasons, such as an unstabilized approach or an obstruction on the runway. As this paper focuses on en-route delays, trajectories with a go-around pattern are excluded from the analysis because delays are clearly explained by the go-around. Some trajectories have an abnormal altitude profile with lot of redundant values. To avoid computational problems, there are also excluded from the analysis. Sophisticated approaches are available in the literature to detect such outliers. For instance, two adaptations of the Invariant Coordinate Selection (ICS) method from the multivariate to the multivariate functional case are proposed by Archimbaud et al. (2022).

3.5. Cross-correlation

Let Z and W be two square integrable random functions corresponding to the altitude and the wind speed. The way in which these random functions depend on one another can be quantified by the empirical cross-correlation function

$$\hat{\rho}_{Z,W}(t_1, t_2) = \frac{\widehat{\mathrm{Cov}}_{Z,W}(t_1, t_2)}{\sqrt{\widehat{\mathrm{Var}}_Z(t_1)\widehat{\mathrm{Var}}_W(t_2)}}$$

where the numerator is the empirical cross-covariance function, $\widehat{\mathrm{Var}}_Z$ is the empirical variance of the altitude and $\widehat{\mathrm{Var}}_W$ is the empirical variance of the wind speed.

Along the main diagonal of the cross-correlation plot (diagonal black dotted line on Figure 6), the value of the altitude (t_1) and the wind speed ($t_2 = t_1$) are correlated in the beginning of the flight and at the end of the flight. This is easily understandable: the wind speed increases (as well as the altitude) during the take-off phase, and the wind speed drops (as well as the altitude) when the aircraft is landing. It is a positive correlation. There is a low negative association between the altitude and the wind speed during the en-route phase of the flight. It is explained by atmospheric layers: above 10 km (depending on the latitude), the tropopause ends. It is within this layer that changes in the temperature values create a pressure variation at a given altitude. This change in

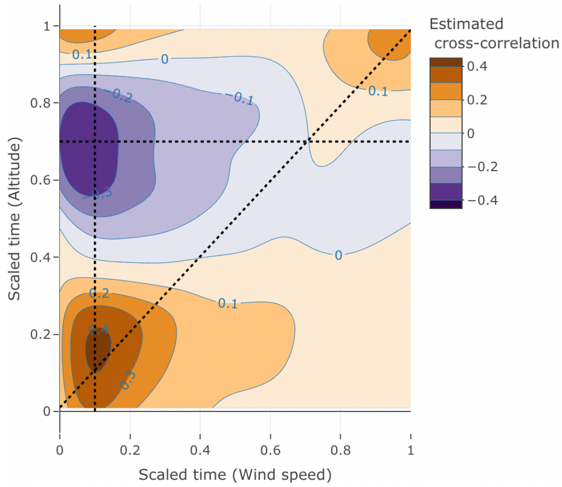

Fig. 6. Contour plot of the empirical cross-correlation function. In each panel t_1 is plotted on y-axis and t_2 on the x-axis ; the legends indicate which observations are being correlated against each other

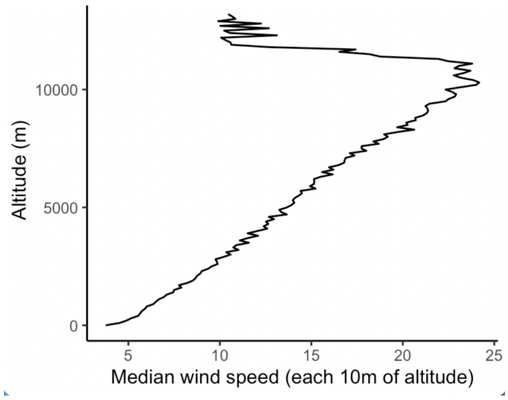

Fig. 7. Empirical median wind speed in $m.s^{-1}$ each 10m of altitude on all the flights between Toulouse-Blagnac and Paris-Orly in 2015

pressure increases with altitude within the tropopause and explains that wind speed is positively associated with altitude. The tropopause border demarcates the beginning of the temperature inversion. Above a given altitude, moving up to the stratosphere, the wind speed drops (Figure 7).

Along the horizontal black dotted line on Figure 6, the association between altitude and wind speed is strongly negative for early wind speed values and decreasing afterwards. An early strong wind speed is associated with a lower en-route altitude value (the highest one during a flight). The aircraft takes more time to reach its highest altitude. Of course, as time goes, the effect is vanishing: a strong wind at the end of the flight has no impact on the en-route altitude.

Along the vertical black dotted line on Figure 6, the association between altitude and wind speed is stronger when the aircraft experiences an early strong wind. At the end of the flight, the positive association between an early wind ($t_2 = 0.1$) and the altitude value ($t_1 = 0.9$) suggests that the wind values in Toulouse are related to the wind values in Paris (positive spatial correlation of the wind).

The lack of symmetry suggests that the cross-correlation between the altitude and the wind mostly goes one-way. The wind speed values follow the altitude profile. Yet, early wind values will strongly influence the timing at which future altitude values are going to be reached.

4. Elastic registration

4.1. The detrimental effect of rescaling to study the wind

Empirically, two flights never have the exact same duration. To compare two trajectories with different lengths, a scaling is made to the unified time interval [0,1]. This transformation is popular in FDA but may have an impact on statistical procedures. In this section, a small thought experiment is proposed to understand how this transformation may impact statistical inference. Five scenarios are presented and are named A, B, C, D, and E. In all of them, there are two theoretical trajectories having:

- A similar longitude-latitude profile (the path from the departure airport to the arrival airport is 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. In scenario E, the en-route altitude is higher for the green flight.
- A similar experienced wind speed, computed according to a stylised version of the empirical association that have been noticed in the previous section. The wind profile depends on the scenario.

Schematically, a scenario is made of 5 plots (see Figure 8): the two upper left plots are the original altitude and wind speed. The scaled versions are shown in the two lower left plots. The plot on the right gives the function used to compute the wind speed from the altitude profile. Let's focus on scenario A. In this scenario, the two altitude profiles are the same. There is no difference in the experienced wind speed. As one may note, comparing wind values for a given time is misleading due to the association with the altitude. At first glance, at each time, the wind was higher for the red trajectory.

This is obviously wrong from looking at the initial situation: each trajectory has been experiencing the same wind speed. The only difference was that the red trajectory had a longer en-route phase.

Intuitively, one wishes to compare wind values for similar flight phases. This well-known problem is called a registration problem in FDA.

4.2. Previous related works on registration

As put by Ramsay and Silverman (2005), registration starts with a key idea: 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 align wind speed values corresponding to the same flight phase.

In general, registration can be as simple as shifting the time or as tedious as applying a continuous transformation of the time scale. Methods for estimating the shift are traditionally based on Procrustes fitting following Silverman (1995). The main problem is the need for a landmark to perform the registration. Landmark registration has been studied by Kneip and Gasser (1992) and Gasser and Kneip (1995). For example, landmarks may be retrieved from the detection of flight phases based on some algorithms that exist in the literature as in Sun et al. (2016) or Liu et al. (2020). Yet, in general, a landmark may not be even available for every curve. If one does not want to rely on landmarks, a continuous registration may be used.

Continuous registration is about constructing an "optimal" transformation of time γ_i for each curve such that the registered curves are given by $x_i^*(t) = x_i\left[\gamma_i(t)\right]$. In the literature, γ is a time warping function. The family of boundary preserving diffeomorphisms $\Gamma_{[0,1]}$ is often considered. The pairwise alignment problem for registering f_2 to f_1, two functions in L^2, may be written

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

where E stands for an optimization criterion. E is often chosen to be a distance, e.g., the usual L^2 norm. Including a symmetry problem, using the L^2 norm has undesirable effects and a geometric framework has been proposed by Srivastava et al. (2011b) for separating the phase and the amplitude variability in functional data.

4.3. *A geometric framework*

The main idea is to take advantage of Square-Root Slope Function (SRSF) representation of functions, generalised to shape analysis in Srivastava et al. (2011a). Departing from a function f that is absolutely continuous on [0, 1], the map $Q : \mathbb{R} \to \mathbb{R}$ is defined as: $Q(x) \equiv \text{sign}(x)\sqrt{|x|}$. The square-root slope function associated to f is defined as $q : [0, 1] \to \mathbb{R}$, 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 square-root slope function is square integrable. For every $q \in L^2([0, 1], \mathbb{R})$ there exists a function f (unique up to a constant) such that the given q is the SRSF of that f. If the cost function for registering f_1 and f_2 is chosen to be $E[f_1, f_2] = \|q_1 - q_2\|$, where $\|.\|$ denotes the L^2 norm, finding the optimal registration for f_1 and f_2 amounts to solve

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

In practice, a dynamic programming algorithm (DPA) is used to compute the optimal time-warping function. Crucially, the registration of trajectories is not a pairwise alignment problem but a joint alignment problem, as there is no such thing as a reference flight. It requires to properly define the Karcher mean for which details can be found in Marron and Dryden (2021). The registration is performed using the R package developed by Tucker (2020).

4.4. *Adopted strategy*

The goal of the registration is align flight phases. It can be done in several ways. The most basic one would align wind profiles directly. Yet, as a trajectory is a multivariate functional object, other solutions may be preferable. To understand other strategies, one should distinguish between at least four main sources in wind speed variability:

- An altitude-related variability.

- A location-related variability (for instance, polar jet streams are known to be stronger than subtropical ones).
- A time-related variability (seasonal patterns).
- An intrinsic variability due to the state of the weather.

If air routes are roughly the same for all flights (delayed or not), the effects of the position (longitude, latitude) are negligible. It is especially the case for short-distance flights such as flights between Toulouse-Blagnac and Paris-Orly. Using the longitude and the latitude is not very relevant to register the wind values.

Two main strategies are compared: a registration based on the wind dimension and a registration based on the altitude. The former aligns the peaks and valleys of the wind speed, which is intuitive, as this is what would be done in a univariate functional data analysis framework. Doing so, the implicit working assumption is that flight phases are visible in the wind profile of the trajectory, local variations being a small noise for the registration. The latter strategy makes the most of the multivariate framework. It aligns altitude profiles and applies the optimal time warping function to the wind profiles. Doing so, one wishes to withdraw the altitude-related variability of the wind speed. The two approaches are compared on the illustrative scenarios (see Figure 9). A good registration should align the flight phases correctly for further inference. In scenario A (no difference in both altitude and wind speed), the two strategies give the same result. In scenario B (same altitude profile and a circumstantial higher wind above 10 km for the red trajectory), the registration is better with the altitude-based strategy as, indeed, the wind speed was the same before and after the en-route phase. No differences in amplitude are desirable outside the en-route phase. In scenario C (same altitude profile, gust of wind for the red trajectory during the en-route phase), the registration is better with the altitude-based strategy. The area between the two curves is coherent with the duration and the amplitude of the gust. In scenario D (same altitude profile, gust of wind for both the green and red trajectory during the en-route phase), the registration is better with the altitude-based strategy. The area between the curves is related to the relative duration of the two gusts. Both gusts have amplitude 10 $m.s^{-1}$ and duration 4 minutes, respectively representing 40% of the en-route phase duration (green trajectory) and 8% of the en-route phase duration (red trajectory). In scenario E (different altitude profile, same wind), both strategies give the same result. The altitude-based strategy seems to be more relevant to align wind speed values as any difference in wind speed values is attributable to the altitude-related variability of the wind or to an intrinsic variability (gusts of wind) even if it is impossible to distinguish between the two sources. Any pointwise test would be more accurate with the altitude-based strategy.

4.5. Results

Figure 10 shows that the elastic registration does not change the mean altitude profile very much. It can be explained by the fact that in the raw data, altitude profiles are already pretty smooth and en-route delays only differ by a few minutes. Yet, the effect of the registration on the pointwise standard deviation is important. Of course, as all flights start and end with a null altitude, and because constraint smoothing was used, the variance is null at the boundaries. From the raw data, it was already clear that the wind speed was very similar to the altitude profile. Using the altitude to register the wind profile does not change the mean and the standard deviation by a large amount (Figure 11). Registering the wind directly may provide a better insight on the true mean, yet, it is less adapted to perform inference.

5. Testing for a different experienced wind speed

5.1. Statistical framework

Following Zhang (2013), several families of inference techniques can be used for functional data: pointwise, 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. A L^2-norm-based test was introduced and a bootstrap approach allowed to approximate the null distribution of interest. The L^2-norm-based test is extended in Zhang and Chen (2007). The F-type test was first studied by Shen and Faraway (2004) and other contributions are listed in Zhang (2013). If groups are independent but some temporal correlation should be taken into account, a two-sample test for functional time series is proposed by Horváth et al. (2013).

In what follows, one is interested in proving that the experienced wind speed is greater for delayed trajectories. The focus is made on a pointwise test as an overall testing method (probably more powerful) is by nature not very appropriate to emphasize where the differences are located.

Let $\widetilde{Y_{g,i,w}}$ be the registered reconstructed wind curve for trajectory i in group g. It is assumed that the N_{Total} registered reconstructed wind curves are independent and identically distributed as $\widetilde{Y_{g,w}}$, a L^2 stochastic process with mean function $\mu_g(t)$ and covariance function $\gamma(s,t)$. A general one-sided pointwise two-sample test for functional data with a common covariance function can be formulated as follows

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

for a fixed $t \in [0,1]$. Let $\overline{\widetilde{Y_{1,w}}}(t) = \frac{1}{N}\sum_{i=1}^{N} \widetilde{Y_{1,i,w}}(t)$ be the usual estimator of the mean function $\mu_1(t)$, $\overline{\widetilde{Y_{2,w}}}(t) = \frac{1}{M}\sum_{i=1}^{M} \widetilde{Y_{2,i,w}}(t)$ be the usual estimator of the mean function $\mu_2(t)$ and $\widehat{\gamma}(s,t)$ be the pooled estimator of the common covariance function $\gamma(s,t)$. The pivotal test statistic is given by $z(t) = \frac{\overline{\widetilde{Y_{2,w}}}(t) - \overline{\widetilde{Y_{1,w}}}(t)}{\sqrt{(\frac{1}{N}+\frac{1}{M})\widehat{\gamma}(t,t)}}$. Under the assumptions of Zhang (2013), under the null, as $min(N,M) \to \infty$, the pivotal test statistic is asymptotically Gaussian for any fixed $t \in [0,1]$. Pointwise p-values can be computed.

Equivalently in a regression framework, this would translate into a pointwise z-test in the spirit of the pointwise test developed by Ramsay and Silverman (2005). To be more precise, this framework is the one of the linear model with functional response and a group covariate being time-independent. The model for the i^{th} trajectory in the g^{th} group (N_{Total} trajectories in total) is $\widetilde{Y_{g,i,w}}(t) = \mu(t) + x_{g,i}\alpha(t) + \eta_{g,i,w}(t)$ for $t \in [0,1]$, where μ denotes a grand mean function that indicates the average wind speed profile of all trajectories, α a specific effect, $\eta_{g,i,w}(t) \overset{i.i.d}{\sim} SP(0,\gamma)$, SP standing for stochastic process. The values of $x_{g,i}$ are either 0 (the trajectory is in group $g = 1$) or 1 (the trajectory is in group $g = 2$). Hypotheses go

$$H_{0,t} : \alpha(t) = 0$$
$$\text{against}$$
$$H_{1,t} : \alpha(t) > 0$$

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

This pointwise test may not be 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. It is the reason why, in the analysis of variance framework (ANOVA), global test statistics were investigated by Cuevas et al. (2004) using a L^2-norm-based test obtained by integrating the numerator of the pointwise F-test statistic over time. A global test whose test statistic is the supremum of the pointwise F-test statistic over time is studied by Zhang et al. (2019). A comparison of tests for the one-way ANOVA problem for functional data is provided by Górecki and Smaga (2015). Second, the same sample of trajectories is used to perform the test at each $t \in \{\tau_1, ..., \tau_m\}$, the set of evaluation points. 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 $\mathbb{P}\left(\text{reject} H_{0,\tau_k}, \forall \tau_k \in C_m\right)$ 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 $\frac{\alpha}{m} \underset{m \to \infty}{\longrightarrow} 0$. It is an obvious problem for functional data as m can be very large. Another famous correction is given by Holm (1979). Yet, both are relevant when rejection regions are disjoint. However, with functional data, rejection probabilities are likely to be correlated for two consecutive evaluation points. 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.

Given the estimated pointwise standard deviation of the wind speed, it seems that a group-specific covariance function would be more realistic. It is a Behrens-Fisher (BF) problem for functional data that have been investigated by Zhang et al. (2011). With a Gaussian assumption, a pointwise test would rely, for instance, on the famous approximation proposed by Welch (1947). Without a Gaussian assumption and if $min(N, M) \to \infty$, $z(t) = \frac{\widetilde{Y_{2,w}}(t) - \widetilde{Y_{1,w}}(t)}{\sqrt{\frac{\widehat{\gamma_2}(t,t)}{M} + \frac{\widehat{\gamma_1}(t,t)}{N}}}$ is asymptotically Gaussian for any fixed $t \in [0, 1]$. Equivalently in a regression framework, a heteroscedastic model can be used. Estimation would be done with Weighted Least Squares (WLS). Some correction is also needed under the heteroscedastic assumption. As said, one way to have a global test would be to use a L^2-norm test. Following Horváth et al. (2013), the null hypothesis would be $H_0 : \mu_1 = \mu_2 \ in \ L^2([0, 1])$ against the alternative that H_0 is false.

5.2. Results

There are 3,161 flights in the original data. Twenty of them are removed because the position of the aircraft is known for less than ten timestamps. Four of them are removed because there is at least one time gap between two known positions that is above thirty minutes. There are seven trajectories with a go-around that have been removed. In total, the test is performed on 3,130 trajectories, departing from Toulouse-Blagnac and landing at Paris-Orly in 2015. $N = 2,510$ trajectories are labeled to be on time and $M = 620$ are labeled to be delayed.

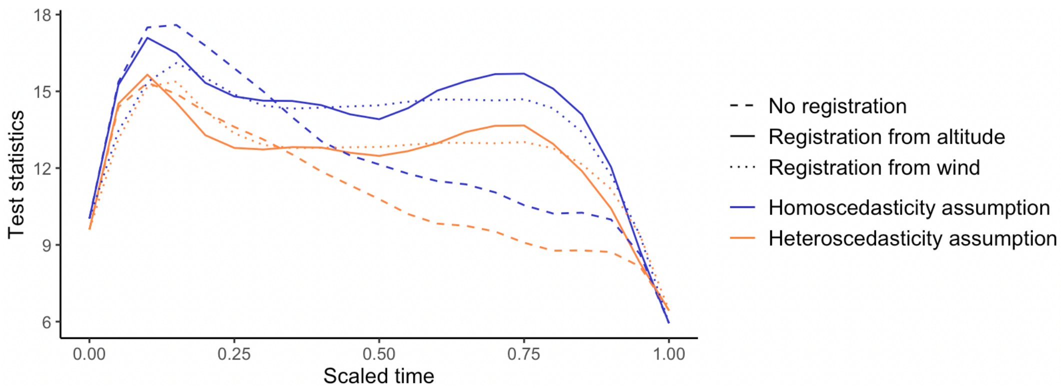

Fig. 12. Pointwise z statistics with different registration strategies

Regarding at the test statistics on Figure 12, there is a clear difference in performing the test with or without registration. Results are also different when comparing the two registration strategies. From the above illustrative scenarios, it would be relevant to only focus on the results obtained when wind profiles are aligned using the altitude optimal time warping. For both assumptions on the covariance functions, p-values are lower during the climb and the descent. P-values remain very low at the departure suggesting that delayed flights experience a stronger wind from the beginning of the trajectory. P-values are higher for the heteroscedastic case as it takes into account a higher variance in the middle of the flight for delayed trajectories. All p-values are below the Bonferroni correction value for $\alpha = 0.01$.

6. Discussion and future works

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 positive. Longitude and latitude profiles are coherent with the true locations of departure and arrival airports. Altitude drops to zero at the beginning and at the end of the flight.

To compare the experienced wind speed of trajectories having different lengths (which is always the case in practice), registration has to be performed. The optimal altitude time warping function has been used to register the wind speed profile. Few scenarios illustrate that it seems to be more efficient for further inference. In this context, efficiency means that differences in wind speed values are not coming from the fact that trajectories are of different lengths. Doing the test without registration will only prove that trajectories are of different lengths, which is a no-brainer.

Notably, the test shows that, in 2015, for flights departing from Toulouse-Blagnac and landing at Paris-Orly, the wind profiles were different for delayed and on time trajectories with a very high level of confidence. Incorporating a wind speed dimension in models is likely to enrich the prediction and classification of delays.

Several exciting aspects fall outside the scope of this paper. First, the paper focuses on a specific year and a specific air link. Given the above methodology, it is effortless to extend the analysis to other flights and other years. To answer operational concerns, it may be of interest to prove that the wind speed impact is positively associated with the travelled distance and to quantify this association.

Adding international flights require to adapt the registration strategy as the longitude and the latitude become of key importance. For instance, registering flights between Paris and New York City must take North Atlantic Tracks (NAT) into account. Adding new weather dimensions, for example the wind direction, calls for a complete study of a possible cross-correlation with the usual dimensions of the trajectory. In addition, testing for a seasonal impact of the wind on delays would require to complement trajectory data, for instance with Automatic Dependent Surveillance–Broadcast (ADS–B) data. Indeed, as mentioned earlier, R&D data are only available for four months each year.

Second, the independence assumption that is used in Section 5 could be relaxed. It is likely that wind speed profiles are temporally correlated. Under necessary assumptions, the testing procedure proposed in Horváth et al. (2013) can be generalized to functional time series. Yet, as the functional time series of aircraft trajectories is highly irregular, some work has to be done. Indeed, departures are not scheduled on a regular time scale.

Finally, the statistical analysis of aircraft trajectories offers a unique opportunity to make the most of advanced FDA frameworks and to bring together applied statistics and ATM.

References

Archimbaud, A., Boulfani, F., Gendre, X., Nordhausen, K., Ruiz-Gazen, A. and Virta, J. (2022) ICS for multivariate functional anomaly detection with applications to predictive maintenance and quality control. *Econometrics and Statistics*.

de Boor, C. (2001) *A Practical Guide to Spline*. Springer-Verlag New York Inc.

Borsky, S. and Unterberger, C. (2019) Bad weather and flight delays: The impact of sudden and slow onset weather events. *Economics of Transportation*, **18**, 10–26.

Carvalho, L., Sternberg, A., Gonçalves, L., Cruz, A., Soares, J., Brandão, D., Carvalho, D. and Ogasawara, E. (2020) On the relevance of data science for flight delay research: a systematic review. *Transport Reviews*, **41**, 499–528.

Cox, D. D. and Lee, J. S. (2008) Pointwise Testing with Functional Data Using the Westfall-Young Randomization Method. *Biometrika*, **95**, 621–634.

Cuevas, A., Febrero, M. and Fraiman, R. (2004) An anova test for functional data. *Computational Statistics & Data Analysis*, **47**, 111–122.

Faraway, J. (1997) Regression analysis for a functional response. *Technometrics*, **39**, 254–261.

Gasser, T. and Kneip, A. (1995) Searching for Structure in Curve Samples. *Journal of the American Statistical Association*, **90**, 1179–1188.

Górecki, T. and Smaga, (2015) A comparison of tests for the one-way ANOVA problem for functional data. *Computational Statistics*, **30**, 987–1010.

Hall, P. and Van Keilegom, I. (2007) Two-Sample Tests in Functional Data Analysis Starting From Discrete Data. *Statistica Sinica*, **17**, 1511–1531.

Holm, S. (1979) A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*, **6**, 65–70.

Horváth, L. and Kokoszka, P. (2012) *Inference for Functional Data with Applications*, vol. 200 of *Springer Series in Statistics*. Springer-Verlag New York Inc.

Horváth, L., Kokoszka, P. and 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**, 103–122.

Jarry, G., Delahaye, D., Nicol, F. and Feron, E. (2020) Aircraft atypical approach detection using functional principal component analysis. *Journal of Air Transport Management*, **84**, 101787.

Kneip, A. and Gasser, T. (1992) Statistical Tools to Analyze Data Representing a Sample of Curves. *The Annals of Statistics*, **20**, 1266–1305.

Kokoszka, P. and Reimherr, M. (2017) *Introduction to Functional Data Analysis*. Chapman and Hall/CRC.

Liu, D., Xiao, N., Zhang, Y. and Peng, X. (2020) Unsupervised Flight Phase Recognition with Flight Data Clustering based on GMM. In *2020 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, 1–6.

Liu, Y., Liu, Y., Hansen, M., Pozdnukhov, A. and Zhang, D. (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.

Marron, J. S. and Dryden, I. L. (2021) *Object Oriented Data Analysis.* Chapman and Hall/CRC.

Mueller, E. and Chatterji, G. (2002) Analysis of aircraft arrival and departure delay characteristics. Los Angeles, California. AIAA Aviation Technology, Integration, and Operations (ATIO) Technical, Los Angeles, California.

Nicol, F. (2017) Statistical Analysis of Aircraft Trajectories: a Functional Data Analysis Approach. pp.51. Alldata 2017, The Third International Conference on Big Data, Small Data, Linked Data and Open Data.

Pejovic, T., Williams, V. A., Noland, R. B. and Toumi, R. (2009) Factors Affecting the Frequency and Severity of Airport Weather Delays and the Implications of Climate Change for Future Delays. *Transportation Research Record*, **2139**, 97–106.

Puechmorel, S. and Delahaye, D. (2007) 4D trajectories : a functional data perspective. pp 1.C.6. IEEE.

Pérez–Rodríguez, J., Pérez–Sánchez, J. and Gómez–Déniz, E. (2017) Modelling the asymmetric probabilistic delay of aircraft arrival. *Journal of Air Transport Management*, **62**, 90–98.

Ramsay, J. O. and Silverman, B. W. (2002) *Applied Functional Data Analysis: Methods and Case Studies.* Springer-Verlag New York Inc.

— (2005) *Functional data analysis.* Springer series in statistics. Springer-Verlag New York Inc., 2nd edn.

Shen, Q. and Faraway, J. (2004) A F Test For Linear Models With Functional Responses. *Statistica Sinica*, **14**, 1239–1257.

Silverman, B. W. (1995) Incorporating Parametric Effects into Functional Principal Components Analysis. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **57**, 673–689.

Srivastava, A., Klassen, E., Joshi, S. H. and Jermyn, I. H. (2011a) Shape Analysis of Elastic Curves in Euclidean Spaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **33**, 1415–1428.

Srivastava, A. and Klassen, E. P. (2016) *Functional and Shape Data Analysis*. Springer-Verlag New York Inc., 1st edn.

Srivastava, A., Wu, W., Kurtek, S., Klassen, E. and Marron, J. S. (2011b) Registration of Functional Data Using Fisher-Rao Metric. URL: `http://arxiv.org/abs/1103.3817`. ArXiv: 1103.3817.

Su, J., Kurtek, S., Klassen, E. and Srivastava, A. (2014) Statistical analysis of trajectories on Riemannian manifolds: Bird migration, hurricane tracking and video surveillance. *The Annals of Applied Statistics*, **8**.

Sun, J., Ellerbroek, J. and Hoekstra, J. M. (2016) Large-Scale Flight Phase Identification from ADS-B Data Using Machine Learning Methods. *7th International Conference on Research in Air Transportation*.

Tucker, J. D. (2020) fdasrvf: Elastic Functional Data Analysis. URL: `https://CRAN.R-project. org/package=fdasrvf`. R package version 1.9.7.

Wang, Q. (2021) Two-sample inference for sparse functional data. *Electronic Journal of Statistics*, **15**.

Wang, Y. and Kulkarni, D. (2011) Modeling Weather Impact on Ground Delay Programs. *SAE International Journal of Aerospace*, **4**, 1207–1215.

Welch, B. L. (1947) The Generalization of 'Student's' Problem when Several Different Population Variances are Involved. *Biometrika*, **34**, 28–35.

Zhang, J.-T. (2013) *Analysis of Variance for Functional Data*. Chapman and Hall/CRC, 1st edn.

Zhang, J.-T. and Chen, J. (2007) Statistical inferences for functional data. *The Annals of Statistics*, **35**.

Zhang, J.-T., Cheng, M.-Y., Wu, H.-T. and Zhou, B. (2019) A new test for functional one-way ANOVA with applications to ischemic heart screening. *Computational Statistics & Data Analysis*, **132**, 3–17.

Zhang, J.-T., Liang, X. and Xiao, S. (2011) On the Two-Sample Behrens-Fisher Problem for Functional Data. *Journal of Statistical Theory and Practice*, **4**.

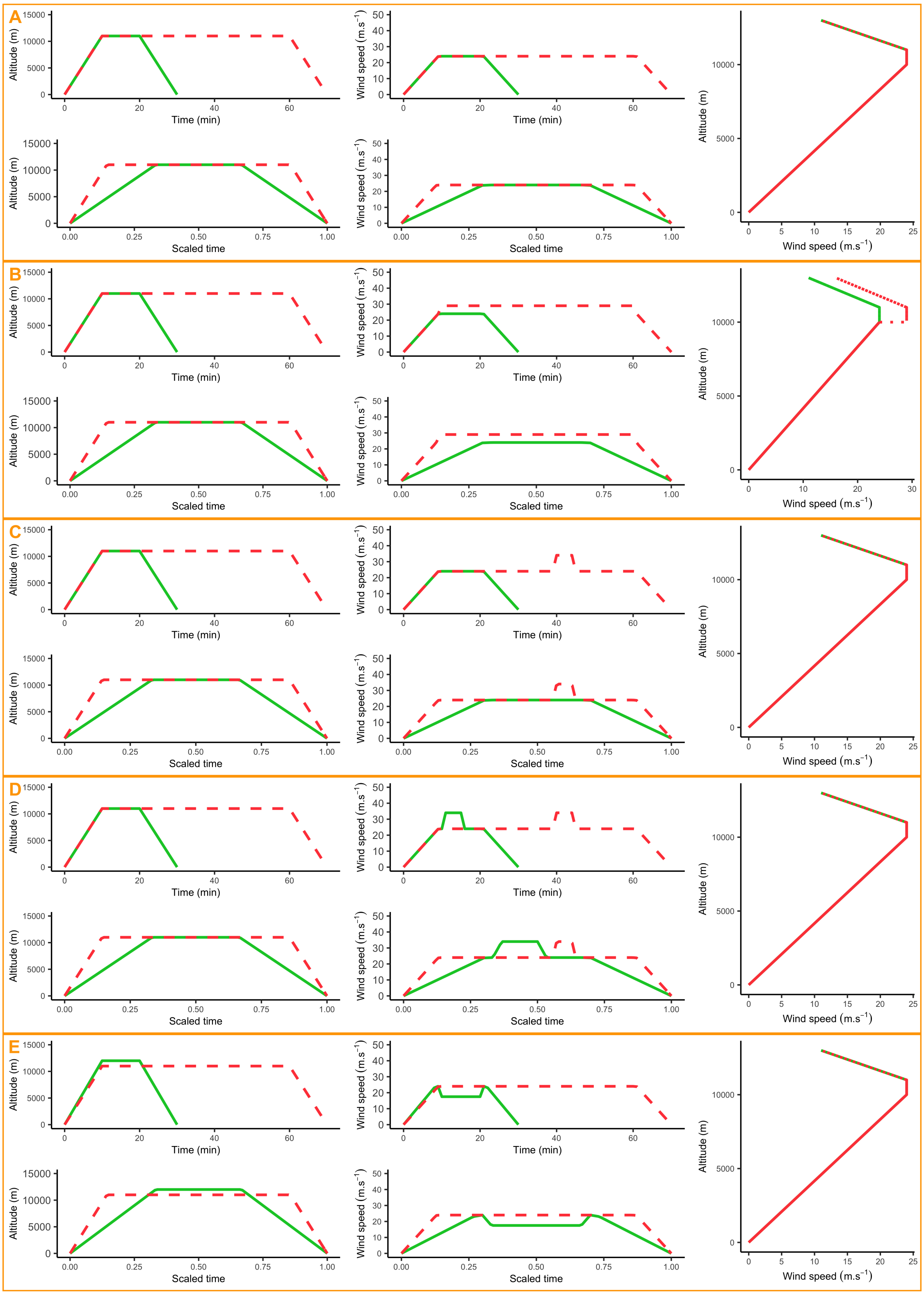

Fig. 8. Illustration of the five scenarios (A, B, C, D, E)

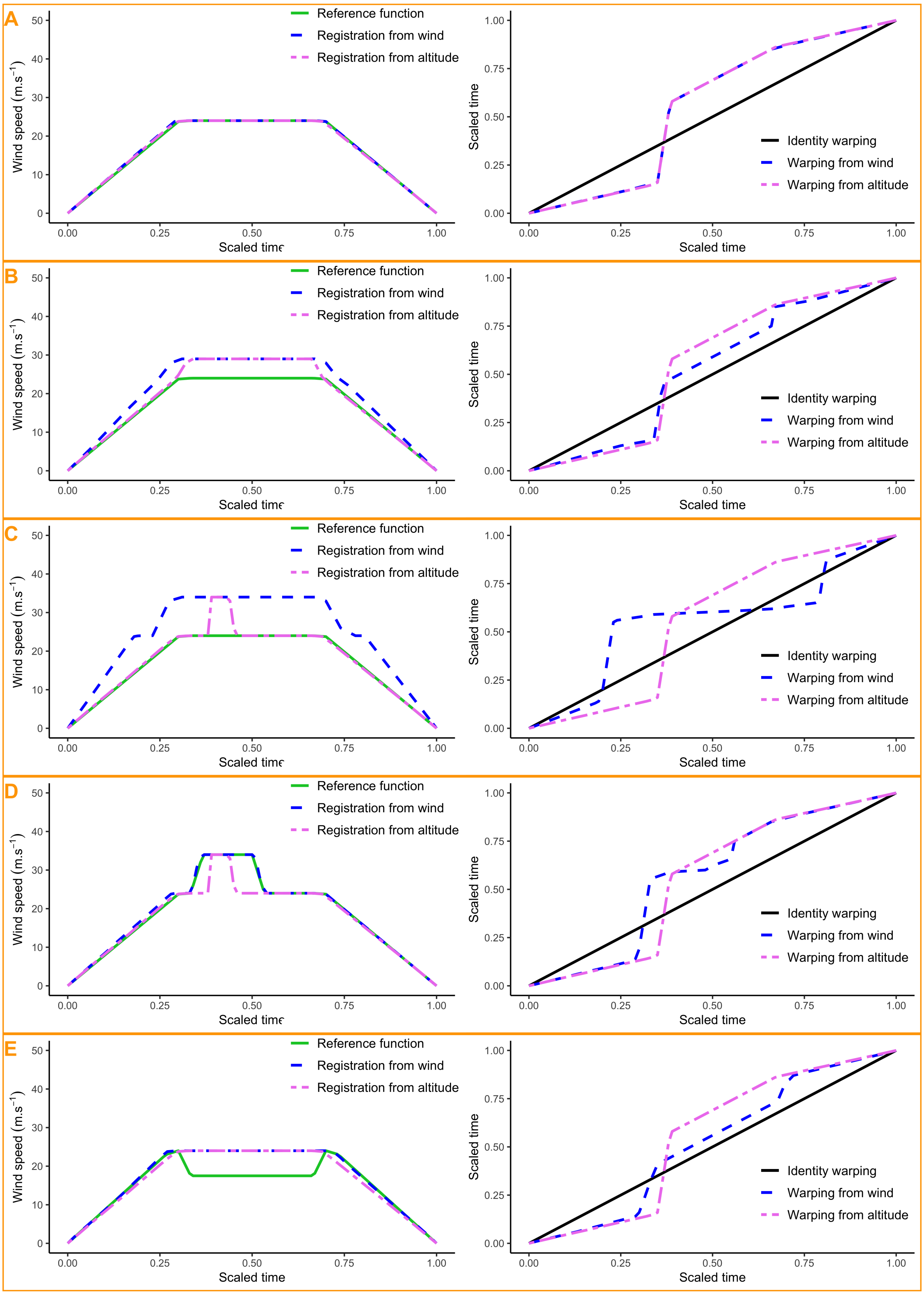

Fig. 9. Elastic registration strategies on the five scenarios

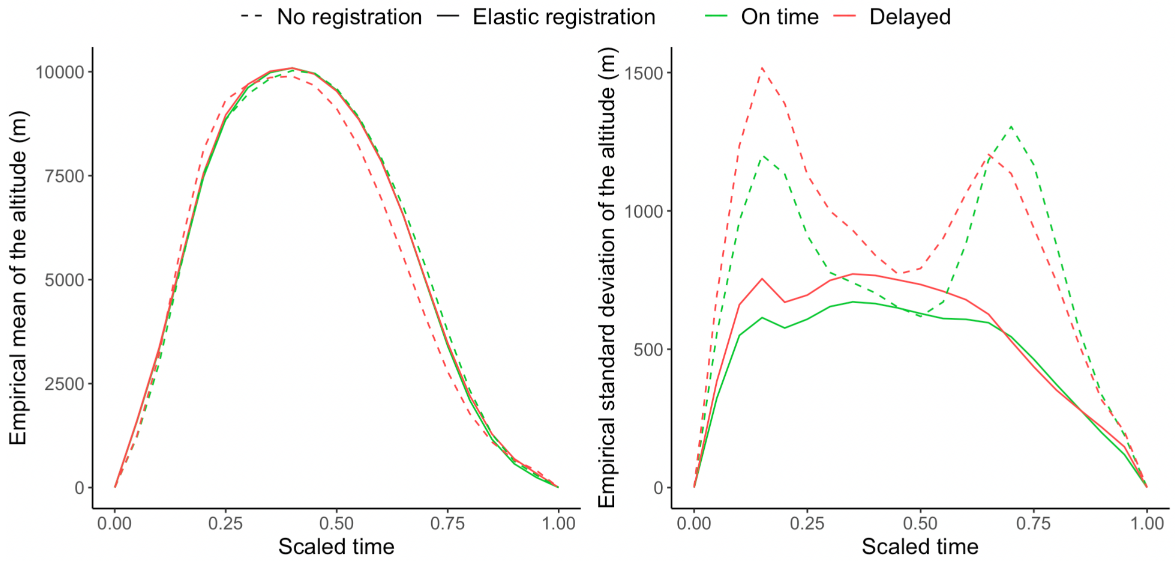

Fig. 10. Registered and raw altitude profiles by group for all flights between Toulouse-Blagnac and Paris-Orly in 2015

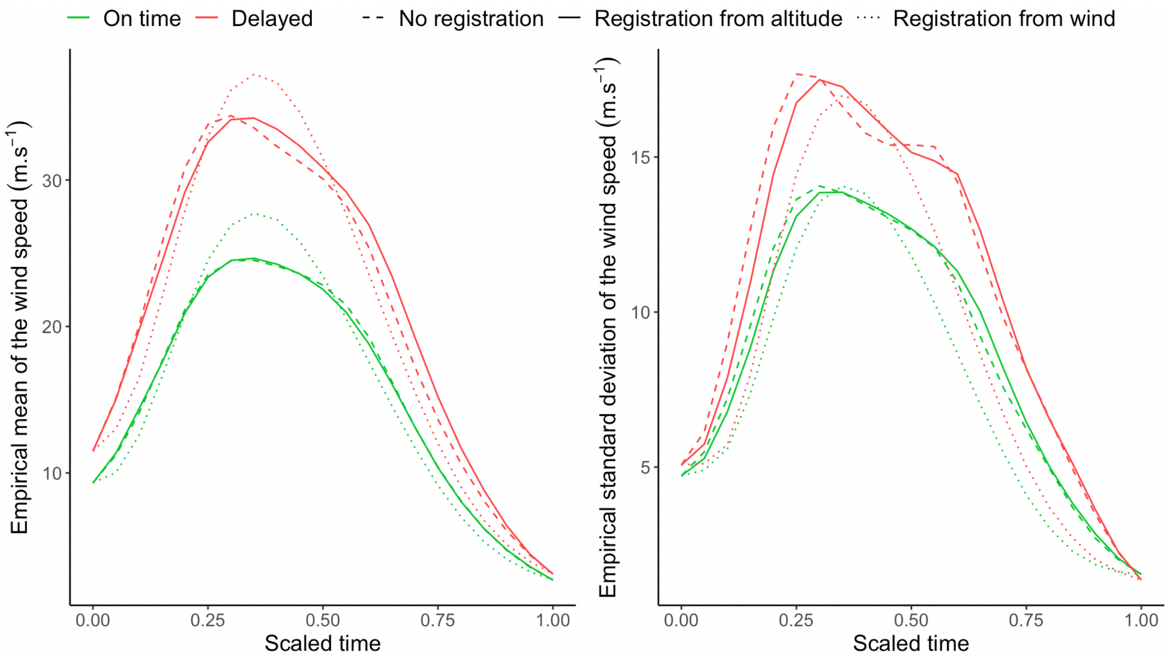

Fig. 11. Registered wind speed profiles by group for all flights between Toulouse-Blagnac and Paris-Orly in 2015