



HAL
open science

Semi Parametric Estimations of rotating and scaling parameters for aeronautic loads

Edouard Fournier, Stéphane Grihon, Thierry Klein

► **To cite this version:**

Edouard Fournier, Stéphane Grihon, Thierry Klein. Semi Parametric Estimations of rotating and scaling parameters for aeronautic loads. 2018. hal-01879135v1

HAL Id: hal-01879135

<https://hal.science/hal-01879135v1>

Preprint submitted on 22 Sep 2018 (v1), last revised 4 Apr 2019 (v3)

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.

Semi Parametric Estimations of rotating and scaling parameters for aeronautic loads

Edouard Fournier^{*1,2,3}, Stéphane Grihon^{†2}, and Thierry Klein^{‡ 1,3}

¹Institut de mathématique, UMR5219; Université de Toulouse;
CNRS, UPS IMT, F-31062 Toulouse Cedex 9, France

²Airbus France 316, Route de Bayonne, Toulouse France

³ENAC - Ecole Nationale de l'Aviation Civile, Université de
Toulouse, France

Abstract

In this paper, we perform registration of noisy curves. We provide an appropriate model in estimating the rotation and scaling parameters to adjust a set of curves through a M -estimation procedure. We prove the consistency and the asymptotic normality of our estimators. Numerical simulation and a real life aeronautic example are given to illustrate our methodology.

Keywords:

Semiparametric model, Registration of curves, Industrial monitoring

MSC Classification:

62F12, 62F30, 62P30

1 Introduction

An airframe structure is a complex system and its design is a complex task involving today many simulation activities generating massive amounts of data. This is, for example, the process of loads and stress computations of an aircraft. That is the computations of the forces and the mechanical strains suffered by the structure. The overall process exposed in Figure 1 is run to identify load cases (i.e aircraft mission and configurations: maneuvers, speed, loading, stiffness...), that are critical in terms of stress endured by the structure and, of course, the parameters which make them critical. The final aim is to size and design the structure (and potentially to reduce loads in order to reduce the weight of the

*edouard.fournier.grihon@airbus.com

†stephane.grihon@airbus.com

‡thierry.klein@math.univ-toulouse.fr or thierry01.klein@enac.fr

structure). Typically for an overall aircraft structure, millions of load cases can be generated and for each of these load cases millions of structural responses (i.e how structural elements react under such conditions) have to be computed. As a consequence, computational times can be significant.

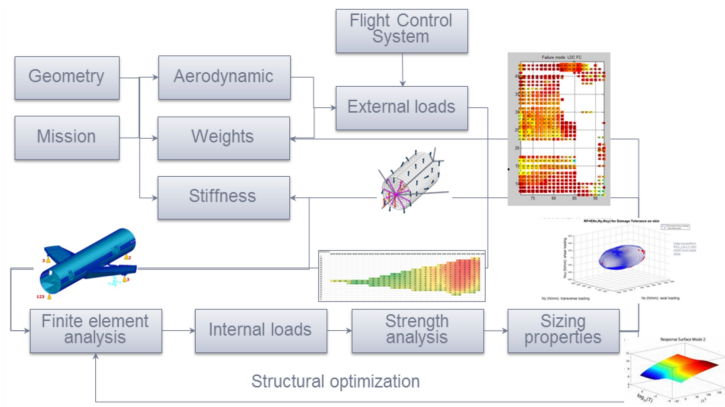


Figure 1: Flowchart for loads and stress analysis process

In an effort to continuously improve methods, tools and ways-of-working, Airbus has invested a lot in digital transformation and the development of infrastructures allowing to treat data (newly or already produced). The main industrial challenge for Airbus is to reduce lead time in the computation and preliminary sizing of an airframe as well as extracting value from already calculated loads.

In this paper, we focus on the external loads of a wing: for each load case are calculated the shear forces (transverse forces near to vertical arising from aerodynamic pressure and inertia) and bending moments (resulting from the shear forces, they represent the flexion of the wing) such as shown in Figure 2.

These external loads appeared to be extremely regular and one can legitimately suppose that it exists a link between all those curves. Indeed, it is natural to assume that it exists a reference bending moment (a reference curve) which can be morphed through a deformation model to give all the other curves. This problem of point sets, curves and images registration has been widely studied and extended to image and shape analysis [1, 8].

The first problem of registration is the identification of a reference function: which binds all the others and the second is to define the transformation model. Numerous papers refers to the identification of a proper reference template such as [3] who used a wavelet threshold to identify the mean pattern, or [5] who proposed to estimate a common template to through an approximated geodesic distance for manifold embedding. Nevertheless, the use of a template is not

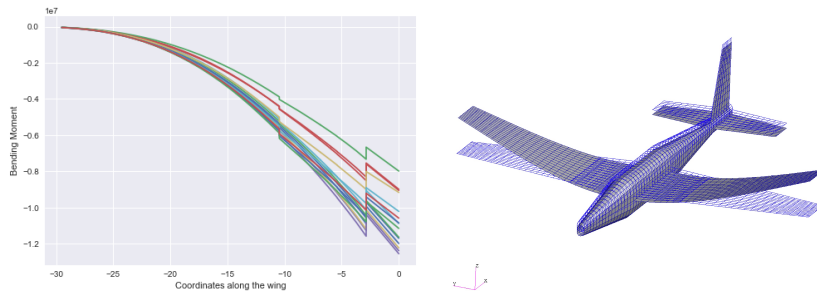


Figure 2: (a) Examples of bending moments along the wing for different load cases - (b) Finite element model of a generic aircraft representing the wing deformation [13]

always necessary as shown in [4]. In our work, we do not focus on the identification of the reference curve and use simply the mean curve as a template.

While the estimation of the rotation and scaling parameters have been widely studied for image registration such as shown in [4], [10], researches on curve registration have been focused on warping strategies and estimations of shifts such as shown in [12], [7]. However, to the best of our knowledge, none work has been realized on the estimation of rotation and scaling parameters when the transformation is applied on the couple of ordinate and abscissa of a curve. More precisely in our study, we observe a large number of functions differing from each other only by the non-linear combination of transformations: a rotation, a homothetic transformation and a dilation of the design space. The idea is then to use a transformation model in order to predict the deformation parameters to estimate loads for different loads configurations.

Our paper is organized as follow: in Section 1, we define the observations as well as the transformation model used and the non linear regression model. Estimation of the rotation and scaling parameters is made following the classical guideline of M -estimators and we provide M -estimators of the unknown parameters and we give their asymptotic behaviour. Then, in Section 3, we give a simulated example illustrating our methodology and discuss an application to the aeronautic problem. The proofs are postponed to the last section.

2 Framework, model and analytic results

In this section, we describe the statistical model studied and give the asymptotic behaviour of the M -estimators of the unknown parameters.

2.1 The observations

Notation 1 Let be $x \in [0, 1]$, and $f : [0, 1] \rightarrow \mathbb{R}^+$. We denote by C the curve $C := \left(\begin{array}{c} x \\ f(x) \end{array} \right)_{x \in [0, 1]}$.

In our framework, we have at hand $K + 1$ curves. \tilde{C} is the reference curve and we assume that the K other curves C_j , $j = 1, \dots, K$ are the images of \tilde{C} by the transformation model described bellow. The K curves are observed on the same random grid $\mathcal{D}_N := \{X_1, \dots, X_N\}$ where $(X_i)_{i=1, \dots, N}$ are iid random variables with uniform distribution. Hence we have at our disposal

$$C_j^N = (X_i, f_j(X_i))_{i=1, \dots, N}, j=1, \dots, K.$$

We will consider the following two cases

- i) \tilde{C} is known everywhere: $\tilde{C} = (x, \tilde{f}(x)), \forall x \in [0, 1]$,
- ii) \tilde{C} is only known on \mathcal{D}_N : $\tilde{C} = (x, \tilde{f}(x)), \forall x \in \mathcal{D}_N$.

2.2 Transformation model

Before defining the transformation model linking the K observed curves to the pattern \tilde{C} , one must ensure that the functions f_j , $j = 1, \dots, K$ and \tilde{f} are "admissible". This is the aim of the next definition:

Definition 2 Let $\theta_0 \in]0, \frac{\pi}{2}[$, \mathcal{F}_{θ_0} is the set of applications defined by:

$$\mathcal{F}_{\theta_0} := \{f : [0, 1] \rightarrow \mathbb{R}^+, f \text{ is differentiable on } [0, 1], \\ f(0) > 0, f(1) = 0, \\ f'(x) < 0, f'(1) = 0, f'(0) > -\cot \theta_0\}.$$

Let now define the parametric family of transformation that we will put in action.

Notation 3 For any function $T_\alpha : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ depending on a parameter $\alpha \in \Theta$, we will denote by T_α^1 and T_α^2 its two coordinates.

In the following, we define our transformation model.

Set $0 < \theta_0 < \theta_1 < \frac{\pi}{2}$, and $0 < \lambda_{min} < \lambda_{max}$. For any $\alpha := (\theta, \lambda) \in \Theta = [-\theta_0, \theta_1] \times [\lambda_{min}, \lambda_{max}]$, set

$$T_\alpha := H_\lambda \circ S_\theta \circ R_\theta$$

the composition of a rotation R_θ , a rescaling S_θ applied on the x -axis, and a homothetic transformation H_λ on the y -axis. More precisely

- R_θ is the rotation centered in $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and of angle θ . That is

$$R_\theta : [0, 1] \times \mathbb{R}^+ \rightarrow [-A_{min}, 1] \times \mathbb{R}$$

$$\begin{pmatrix} x \\ y \end{pmatrix} \rightarrow \begin{pmatrix} (x-1) \cos \theta - y \sin \theta + 1 \\ (x-1) \sin \theta + y \cos \theta \end{pmatrix},$$

with $A_{min} = \sqrt{1 + y(0)^2} - 1$.

- After rotating a curve, depending of the angle θ , the resulting design space of the curve is generally no more $[0,1]$. Hence, we define the following transformation to re-position the curve on $[0,1]$:

$$S_\theta : [-A_{min}, 1] \times \mathbb{R} \rightarrow [0, 1] \times \mathbb{R}$$

$$\begin{pmatrix} x \\ y \end{pmatrix} \rightarrow \begin{pmatrix} \frac{x-a}{1-a} \\ y \end{pmatrix},$$

with a is the real minimum value of the design space after rotating the curve.

- The scaling transformation of parameter $\lambda > 0$ acting on the second coordinate of the curve:

$$H_\theta : [0, 1] \times \mathbb{R}^+ \rightarrow [0, 1] \times \mathbb{R}^+$$

$$\begin{pmatrix} x \\ y \end{pmatrix} \rightarrow \begin{pmatrix} x \\ \lambda y \end{pmatrix}.$$

We now introduced the assumption that will be used:

(A1): The functions $(f_j)_{j=1, \dots, K}$, and \tilde{f} belong to \mathcal{F}_{θ_0} .

(A2): For any $\theta \in [-\theta_0, \theta_1]$, for any $x \in [0, 1]$ and any $f \in \mathcal{F}_{\theta_0}$, the second coordinate of R_θ is positive.

(A3): $\Theta = [-\theta_0, \theta_1] \times [\lambda_{min}, \lambda_{max}]$.

(A4): \tilde{C} is known on $[0, 1]$.

(A5): \tilde{C} is known on the grid \mathcal{D}_N .

Thus, the transformation model considered is as follows:

$$T_\alpha : [0, 1] \times \mathbb{R}^+ \rightarrow [0, 1] \times \mathbb{R}^+$$

$$\begin{pmatrix} x \\ f(x) \end{pmatrix} \rightarrow \begin{pmatrix} \frac{(x-1) \cos \theta - f(x) \sin \theta}{\cos \theta + f(0) \sin \theta} + 1 \\ \lambda((x-1) \sin \theta + f(x) \cos \theta) \end{pmatrix}.$$

2.3 Non-linear regression model

Recall that we wish to adjust the reference curve \tilde{C} on the other curves C_j ($j = 1, \dots, K$) by transformations defined in the previous subsection. Notice that these transformations act on both axis. For any α , we want to compare the value of the transformed curve $(T_\alpha \tilde{C})(X_i)$ with $f_j(X_i)$. Since the abscissa points are affected by the transformation, we denote by $X_i(\alpha)$ the point such that $T_\alpha^1(X_i(\alpha)) = X_i$. For that reason, we introduce the following definition:

Definition 4 We denote by $x(\alpha, g)$ the solution of the equation:

$$T_\alpha^1(u, g(u)) = x. \quad (1)$$

To ease the notation, we finally set $x(\alpha) := x(\alpha, g)$, so $X_i(\alpha) = X_i(\alpha, \tilde{f})$. We consider the non-linear parametric regression model:

$$f_j(X_i) = T_{\alpha_j^*}^2(X_i(\alpha_j^*), \tilde{f}(X_i(\alpha_j^*))) + \epsilon_{j,i}, (j = 1, \dots, K). \quad (2)$$

Where:

- (X_i) are iid with distribution \mathbb{P}_X . It is the design on which we observe the curves C_j ;
- $\alpha_j^* = (\theta_j^*, \lambda_j^*)$ is the couple of true parameters for each curve ($j = 1, \dots, K$);
- $\epsilon_{j,i}$, $\forall i = 1, \dots, N$, $\forall j = 1, \dots, K$ are iid $\mathcal{N}(0, \sigma^2)$ random variables. These variables are assumed to be independent of X_i .

2.4 Estimation

2.4.1 Estimation when \tilde{C} is known on $[0, 1]$

For the sake of simplicity, let us fix j . Relying on a classical M -estimation procedure, we consider a semi-parametric method to estimate the parameters and

define consequently the following empirical contrast function to fit the reference curve \tilde{C} to C_j ($j = 1, \dots, K$):

$$\begin{aligned} M_N^j(\alpha) &= \frac{1}{N} \sum_{i=1}^N (f_j(X_i) - T_\alpha^2(X_i(\alpha), \tilde{f}(X_i(\alpha))))^2 \\ &= \frac{1}{N} \sum_{i=1}^N m_\alpha^j(X_i). \end{aligned} \quad (3)$$

The random function M_N^j is non negative. Furthermore, intuitively, its minimum value should be reached close to the true parameter α_j^* . Indeed, the following theorem gives the consistency of the M -estimator, defined by :

$$\hat{\alpha}_N^j = \operatorname{argmin}_{\alpha \in \Theta} M_N^j(\alpha). \quad (4)$$

Recall that our empirical contrast function enters in the general theory of M -estimator. The Central Limit Theorem will be shown by using M -estimator arguments.

Theorem 5 *Assume that A1, A2, A3 and A4 are satisfied. Then*

$$i) \hat{\alpha}_N^j \xrightarrow[N \rightarrow +\infty]{\mathbb{P}} \alpha_j^*, \quad (5)$$

$$ii) \sqrt{N}(\hat{\alpha}_N^j - \alpha_j^*) \xrightarrow[N \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \Gamma_{\alpha_j^*}). \quad (6)$$

In particular, the covariance matrix has the following form

$$\Gamma_{\alpha_j^*} = V_{\alpha_j^*}^{-1} 2\sigma^2, \quad (7)$$

with $V_{\alpha_j^} = 2\mathbb{E}[\dot{T}_{\alpha_j^*}^2 \dot{T}_{\alpha_j^*}^{2T}]$, and $\dot{T}_{\alpha_j^*}^2$ is the vector of partial derivatives of $T_{\alpha_j}^2$ w.r.t elements of α taken at α_j^* .*

2.4.2 Estimation when \tilde{C} is observed on \mathcal{D}

In this section, we consider the case where the reference curve \tilde{C} is observed on the same grid $\mathcal{D} := (X_i)_{i=1, \dots, N}$ as the other curve C_j , i.e $\tilde{C} = \begin{pmatrix} x \\ \tilde{f}(x) \end{pmatrix}_{x \in \mathcal{D}}$. By

applying the transformation T_α to \tilde{C} , the transformed pattern $T_\alpha \tilde{C}$ is no longer observable on \mathcal{D} . As a consequence, one must make use of an approximation process over \tilde{f} . Let \tilde{f}_N be the linear interpolate of \tilde{f} , defined by:

$$\tilde{f}_N(x) = \sum_{i=1}^N \Delta_i(x) \mathbf{1}_{x \in [X_{(i)}, X_{(i+1)})}, \quad (8)$$

where

$$\Delta_i(x) = \frac{\tilde{f}(X_{(i+1)}) - \tilde{f}(X_{(i)})}{X_{(i+1)} - X_{(i)}} x + \tilde{f}(X_{(i)}) - \frac{\tilde{f}(X_{(i+1)}) - \tilde{f}(X_{(i)})}{X_{(i+1)} - X_{(i)}} X_{(i)}. \quad (9)$$

It is easy to see that \tilde{f}_N belongs also to \mathcal{F}_{θ_0} . Replacing \tilde{C} by \hat{C} in (3) we obtain

$$\hat{M}_N^j(\alpha) = \frac{1}{N} \sum_{i=1}^N (f_j(X_i) - T_\alpha^2(X_i(\alpha, N), \tilde{f}_N(X_i(\alpha, N))))^2,$$

where $X_i(\alpha, N)$ is the solution to the equation $T_\alpha^1(u, \tilde{f}_N(u)) = X_i$.

Using the linear interpolate defined by (8), we show the consistency and asymptotic normality of our M -estimator defined as follow:

$$\hat{\alpha}_N^j = \underset{\alpha \in \Theta}{\operatorname{argmin}} \hat{M}_N^j(\alpha). \quad (10)$$

Theorem 6 *Assume that A1, A2, A3 and A5 are satisfied. Let \tilde{f}_N be defined by (8) and assume that $\exists C > 0$ s.t $\forall x \in [0, 1]$, $\tilde{f}'_N(x) \leq C$, and $\tilde{f}'(x) \leq C$, then*

$$i) \hat{\alpha}_N^j \xrightarrow[N \rightarrow +\infty]{\mathbb{P}} \alpha_j^*, \quad (11)$$

$$ii) \sqrt{N}(\hat{\alpha}_N^j - \alpha_j^*) \xrightarrow[N \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \Gamma_{\alpha_j^*}) \quad (12)$$

with $\Gamma_{\alpha_j^*}$ such defined in (7).

3 Simulations and applications

In this section we illustrate the method on numerical applications. The first subsection is dedicated to some simulated toy example while the second to a real problem. The optimisation problems (4) and (10) will be numerically solved by using the BFGS algorithm [11].

3.1 Simulated toy example

We consider the following model:

$$f_{\lambda, \theta}(x) = \lambda[(x - 1) \sin \theta + g(x) \cos \theta],$$

with $g(x) = 2(\cos \pi x + 1)$. We observe $f_{\lambda_j, \theta_j}(x_i)$ for $i = 1, \dots, 100$, for $J = 25$ values of (λ, θ) with some iid errors ϵ_{ij} .

- The observations points x_i , $i = 1, \dots, 100$ are iid random variables with uniform distribution on $[0, 1]$.

- The parameters are chosen randomly with the following arbitrary distribution $(\lambda, \theta) \hookrightarrow \mathcal{U}([0, 15]) \times \mathcal{U}([-\frac{\pi}{3}, \frac{\pi}{30}])$.
- The errors are assumed to be $\mathcal{N}(0, 0.01)$.

Results are given in Figure 3. The simulated data are shown in Figure 3 (a). Each of these curves is rescaled to $[0, 1]$ to avoid numerical issues. The curve in blue is the reference curve. After the estimation of the parameters λ_j and θ_j , all curves can be rescaled back to their original space as shown in Figure 3 (b).

Rescaling the curves allows to easily choose the initial point of parameters for the optimization algorithm taking 1 for λ and 0 for θ .

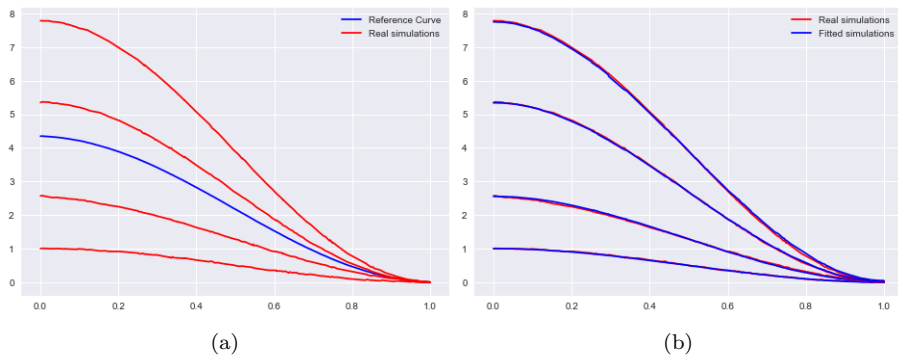


Figure 3: Rotating and scaling results for simulated data: (a) Simulated data; (b) Fitted data

3.2 Aeronautic loads

In [6], the authors present an aeronautic model that computes the loads (forces and moments) on the wing of some aircraft model denoted by *ACM1*. They present several statistical methods in order to study these data. In this section, we will compare the method used in [6] with the model presented in Section 2 for a new aircraft model called *ACM2*. The data at our disposal represents bending moments of a wing (representing its flexion) of an aircraft calculated for 1152 different configurations (load cases). Each configuration is defined by 28 features (speed of the aircraft, mass, altitude, quantity of fuel, etc.), leading to a bending moment calculated on 45 stations along the wing. In a more formal way, we observe the couple $(X_j, Y_j)_{j=1, \dots, 1152}$, where $X_j = (X_j^1, \dots, X_j^{28})$ are the features and $Y_j = (Y_j^1, \dots, Y_j^{45})$ is the bending moment. The idea is to predict the bending moment for different configurations. The data are represented in Figure 4.

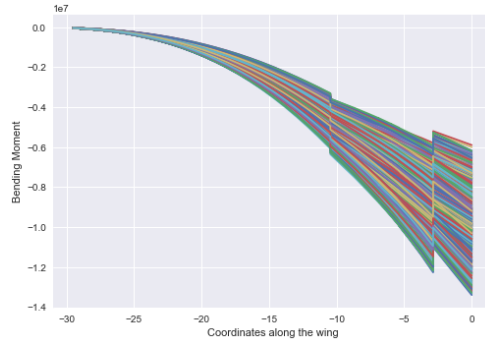


Figure 4: Representation of all the bending moments of a wing of our data base: the wing root is located at the null coordinate, where the strains are maximum when the wing bends.

Due to the discontinuities at the 3rd and 20th stations, we apply our methodology to each section independently. Then, each section can be represented by its minimum and maximum values, and by its rotation and scaling coefficients λ_j and θ_j . Figure 5 assess the quality of the matching process (the reference curve used is the average bending moment).

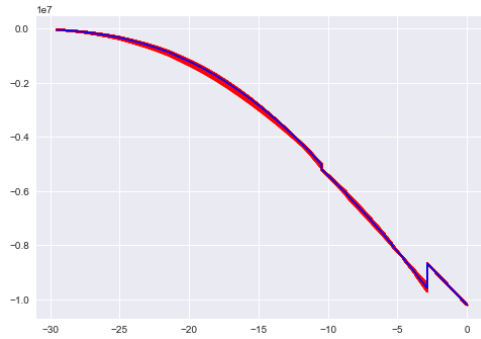


Figure 5: Results of the matching process

Thus the dimensional space of the outputs is reduced to 12 instead of 45. We compare our method to three other methods of [6] applied on the outputs: no transformation (we call it raw - we build 45 models, one per station); a PCA (the three first principal components represent 99,9% of the explained variance - 3 models instead of 45); a polynomial fitting per section (of degree 4 for the first section, of degree 2 for the second and of degree 1 for the third section) which leads to 10 models instead of 45. The Table 1 sums up the number of outputs to predict depending on the method used.

Table 1: Number of outputs to be predicted depending on the method used on the raw outputs: Raw, Deformation Model, PCA, polynomial fitting

| | Number of outputs | Names of outputs |
|--------------------|-------------------|---|
| Raw | 45 | Bending moment value at station 0 to 44 |
| Deformation Model | 12 | $\theta_1, \theta_2, \theta_3, \lambda_1, \lambda_2, \lambda_3, min_1, min_2, min_3, max_1, max_2, max_3$ |
| PCA | 3 | Principal components 1 to 3 |
| Polynomial fitting | 10 | Coefficients of polynomials |

The significant advantage of the reduction dimension techniques used is that the response of the model would have a physical form contrary to the simple linear models performed on the raw data. To build our models, we use the Orthogonal Greedy Algorithm (OGA) also known as the Matching Pursuit Algorithm. Detailed explanations can be found in [2], [14] and [9]. Roughly speaking, we consider the problem of approximating a function by a sparse linear combination of inputs.

To assess the goodness of fit of our models, we defined for a curve of bending moment j the error rate as follows:

$$error(j) = \sqrt{\frac{\sum_{i=1}^{45} (\hat{y}_j(x_i) - y_j(x_i))^2}{\sum_{i=1}^{45} y_j^2(x_i)}}, j = 1, \dots, n_{test},$$

where n_{test} is the size of the sample of test. We compute the error rates on (the sample of test is of 25% the size of the total database). It gives an idea of how far our predictions are. For this standpoint, we can easily compute the empirical cumulative distribution function (CDF): $\forall j = 1, \dots, n_{test}$, let $\alpha \in [0, 1]$. The empirical CDF is defined as:

$$\alpha \rightarrow G(\alpha) = \frac{1}{n} \sum_{j=1}^{n_{test}} \mathbb{1}_{(error(j) \leq \alpha)}$$

In Table 2, we give the values of $G(\alpha)$ for $\alpha = 1\%, 2\%, 5\%, 10\%$ and the mean error. In Figure 6 we give the plots of the function $G(\alpha)$ for the different methods.

Concerning the approximation and prediction of loads, our model is equivalent in average to other tested methods, there are just slightly more observations with an error below 1%. Nevertheless, in our case, the linear models built through the deformation model are sparser than the other. Indeed, in average, 11 variables are chosen as optimal parameter of the greedy algorithm by cross-validation for the deformation model, 13 for the polynomial fitting, 15 for the PCA and 14 for the raw outputs one.

Table 2: Average estimated $\mathbb{P}(\text{error} \leq 1\%)$, $\mathbb{P}(\text{error} \leq 2\%)$, $\mathbb{P}(\text{error} \leq 5\%)$, $\mathbb{P}(\text{error} \leq 10\%)$, $\mathbb{E}(\text{error})$ calculated on several random test data set (25% of the size of the total dataset)

| | Deformation Model | Polynomial Fitting | PCA | Raw |
|--------------------------------------|-------------------|--------------------|-----|------|
| $\mathbb{P}(\text{error} \leq 1\%)$ | 17% | 14% | 16% | 15% |
| $\mathbb{P}(\text{error} \leq 2\%)$ | 45% | 45% | 43% | 51% |
| $\mathbb{P}(\text{error} \leq 5\%)$ | 88% | 88% | 86% | 88% |
| $\mathbb{P}(\text{error} \leq 10\%)$ | 98% | 97% | 95% | 98% |
| $\mathbb{E}(\text{error})$ | 2.9% | 2.9% | 3% | 2.8% |

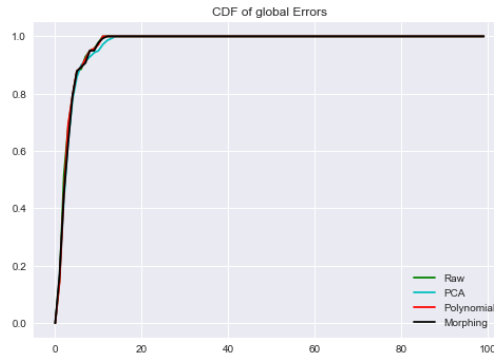


Figure 6: Empirical CDF of error rates ($\mathbb{P}(\text{error} \leq \alpha)$)

Even though the prediction of loads with the deformation model is so likely equivalent to none transformation, it obtains better results than the polynomial fitting and the PCA. Besides, it is important to notice that it gives to engineers a physical interpretation and idea of how react the wing to new constraints. Besides, using this deformation model gives a physical response contrary to a simple linear model per station whose response could be irregular.

4 Proofs and technical result

4.1 Technical result

This section is dedicated to the technical result used in the proof of Theorem 4.

Lemma 7 *Let X_1, \dots, X_N be N independent and identically distributed random variables with uniform distribution on $[0, 1]$ and let $X_{(1)} \leq \dots \leq X_{(N)}$ be the reordered sample. Let $a_N = O(\sqrt{N})$, then:*

$$\mathbb{P}(a_N \sup_j |X_{(j+1)} - X_{(j)}| \geq \epsilon) \xrightarrow{N \rightarrow +\infty} 0.$$

Proof of Lemma 1 Let Z_1, \dots, Z_{N+1} be N independent and identically distributed random variables with exponential distribution with parameter 1. It is a well known fact that

$$\left(\frac{Z_1}{\sum_{k=1}^{N+1} Z_k}, \frac{Z_1+Z_2}{\sum_{k=1}^{N+1} Z_k}, \dots, \frac{Z_1+\dots+Z_N}{\sum_{k=1}^{N+1} Z_k} \right) \stackrel{\mathcal{L}}{=} (X_{(1)}, \dots, X_{(N)}) \text{ and we have}$$

$$X_{(j+1)} - X_{(j)} \stackrel{\mathcal{L}}{=} \frac{Z_{j+1}}{\sum_k Z_k}. \quad (13)$$

Now, for $\epsilon > 0$

$$\begin{aligned} \mathbb{P}(\sup_j |X_{(j+1)} - X_{(j)}| \geq \epsilon) &\leq \sum_j \mathbb{P}(X_{(j+1)} - X_{(j)} \geq \epsilon) \\ &\leq N \max_j \mathbb{P}(X_{(j+1)} - X_{(j)} \geq \epsilon). \end{aligned}$$

By using (13), we have

$$\mathbb{P}(X_{(j+1)} - X_{(j)} \geq \epsilon) = (1 - \epsilon)^{N-1}.$$

Then,

$$\mathbb{P}(\sup_j |X_{(j+1)} - X_{(j)}| \geq \epsilon) \leq N(1 - \epsilon)^{N-1}.$$

The result follows replacing ϵ by $\frac{\epsilon}{a_N}$ and letting $N \rightarrow +\infty$. ■

4.2 Proofs of Theorems 3 and 4

Proof of Theorem 3 To ease the notation, we do not display the dependency in j .

i) By (3) it is easy to see that $M_N(\alpha)$ is an empirical mean of iid bounded random variables. Thus, by the Strong Law of Large Number (SLLN)

$$M_N(\alpha) \xrightarrow[N \rightarrow +\infty]{p.s} M(\alpha),$$

with $M(\alpha) = \mathbb{E}[\epsilon^2] + \mathbb{E}[(T_\alpha^2(X(\alpha), \tilde{f}(X(\alpha))) - T_{\alpha^*}^2(X(\alpha^*), \tilde{f}(X(\alpha^*))))^2]$.

$M(\alpha)$ is continuous and has an obvious unique minimum α^* . Since Θ is compact, this implies that $\inf_{\alpha: d(\alpha, \alpha^*) \geq \epsilon} M(\alpha) > M(\alpha^*)$ is satisfied (see Problem 27 p. 84 in [15]).

It remains to prove that $\{m_\alpha : \alpha \in \Theta\}$ is a Glivenko-Cantelli class. Thanks to the remark following the proof of Theorem 5.9 in [15], this is an easy consequence of the continuity of $\alpha \rightarrow m_\alpha$ and the fact that the function is bounded by a continuous and integrable function on $[0, 1]$. Indeed, it exists at least a function f^* in \mathcal{F}_{θ_0} which bounds every other functions, and two constants $K_1 > 0, K_2 > 0$ such that

$$m_\alpha(x) \leq K_1(f^*(x) + K_2)^2,$$

and

$$\sup_{\alpha \in \Theta} |M_N(\alpha) - M(\alpha)| \xrightarrow[N \rightarrow +\infty]{\mathbb{P}} 0. \quad (14)$$

The result follows from the Theorem 5.7 in [15].

ii) The Central Limit Theorem will be a consequence of Theorem 5.23 in [15]. Recall that

$$m_\alpha(x) = [f(x) - \lambda((x(\alpha) - 1) \sin \theta + \cos \theta \tilde{f}(x(\alpha)))]^2.$$

By the Implicit Function Theorem, that is easy to see that $\alpha \rightarrow x(\alpha)$ is C^1 on a compact set. This implies that the norm of the gradient of m_α is uniformly bounded in α . Hence $\exists \dot{\phi}(x) \in L^1$ such that $\|\nabla_\alpha m_\alpha(x)\| \leq \dot{\phi}(x)$ hence

$$|m_{\alpha_1}(x) - m_{\alpha_2}(x)| \leq \dot{\phi}(x) \times \|\alpha_1 - \alpha_2\|.$$

In order to give an explicit formula for the limit variance, we apply the results of Example 5.27 in [15] where f_θ becomes in our case T_α^2 and hence, we have

$$\sqrt{N}(\hat{\alpha}_N^j - \alpha_j^*) \xrightarrow[N \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \Gamma_{\alpha_j^*}),$$

with $\Gamma_{\alpha_j^*} = V_{\alpha_j^*}^{-1} 2\sigma^2$ and $V_{\alpha_j^*} = 2\mathbb{E}[\dot{T}_{\alpha_j^*}^2 \dot{T}_{\alpha_j^*}^{2\mathbf{T}}]$.

■

Proof of Theorem 4

i) To prove the consistency of $\hat{\alpha}_N$ we have to show that

$$\sup_{\alpha \in \Theta} |\hat{M}_N(\alpha) - M(\alpha)| \xrightarrow[N \rightarrow +\infty]{\mathbb{P}} 0.$$

We have,

$$\sup_{\alpha \in \Theta} |\hat{M}_N(\alpha) - M(\alpha)| \leq \sup_{\alpha \in \Theta} |\hat{M}_N(\alpha) - M_N(\alpha)| + \sup_{\alpha \in \Theta} |M_N(\alpha) - M(\alpha)|.$$

It has been shown in the proof of Theorem 3 that

$$\sup_{\alpha \in \Theta} |M_N(\alpha) - M(\alpha)| \xrightarrow[N \rightarrow +\infty]{\mathbb{P}} 0.$$

It remains to prove that

$$\sup_{\alpha \in \Theta} |\hat{M}_N(\alpha) - M_N(\alpha)| \xrightarrow[N \rightarrow +\infty]{\mathbb{P}} 0.$$

To ease the notation, we write $T_\alpha^2(i, N) = T_\alpha^2(X_i(\alpha, N), \tilde{f}_N(X_i(\alpha, N)))$, and $T_\alpha^2(i) = T_\alpha^2(X_i(\alpha), \tilde{f}(X_i(\alpha)))$. Set

$$\begin{aligned} D_N(\alpha) &= M_N(\alpha) - \hat{M}_N(\alpha) \\ &= \frac{1}{N} \sum_{i=1}^N 2y(X_i) [T_\alpha^2(i, N) - T_\alpha^2(i)] - \frac{1}{N} \sum_{i=1}^N [T_\alpha^2(i, N) - T_\alpha^2(i)] [T_\alpha^2(i, N) + T_\alpha^2(i)]. \end{aligned}$$

As f and T_α^2 are continuous and bounded on $\Theta \times [0, 1]$, this implies that:

$$\begin{aligned} |D_n(\alpha)| &\leq \left| \frac{1}{N} \sum_{i=1}^N 2f(X_i) [T_\alpha^2(i, N) - T_\alpha^2(i)] \right| + \left| \frac{1}{N} \sum_{i=1}^N [T_\alpha^2(i, N) - T_\alpha^2(i)] [T_\alpha^2(i, N) + T_\alpha^2(i)] \right| \\ &\leq K \left(\frac{1}{N} \sum_{i=1}^N [T_\alpha^2(i, N) - T_\alpha^2(i)]^2 \right)^{\frac{1}{2}} \\ &\leq K' \left(\frac{1}{N} \sum_{i=1}^N [(X_i(\alpha, N) - X_i(\alpha))(1 + C) + (\tilde{f}_N(X_i(\alpha)) - \tilde{f}(X_i(\alpha)))^2] \right)^{\frac{1}{2}}. \end{aligned}$$

By construction, there exists j such that $X_{(j)} \leq X_i(\alpha) \leq X_{(j+1)}$, and $X_{(j)} \leq X_i(\alpha, N) \leq X_{(j+1)}$ which leads to:

$$X_i(\alpha, N) - X_i(\alpha) = \gamma(X_{(j+1)} - X_{(j)}).$$

Besides, since there exists $\gamma' > 0$ such that $\tilde{f}_N(X_i(\alpha)) = \gamma' \tilde{f}(X_{(j+1)}) + (1 - \gamma') \tilde{f}(X_{(j)})$ we have

$$\begin{aligned} |\tilde{f}_N(X_i(\alpha)) - \tilde{f}(X_i(\alpha))| &\leq \gamma' |\tilde{f}(X_{(j+1)}) - \tilde{f}(X_{(j)})| \\ &\leq C\gamma' |X_{(j+1)} - X_{(j)}|, \end{aligned}$$

and

$$\begin{aligned} |D_n(\alpha)| &\leq K' \left(\frac{1}{N} \sum_{j=1}^N [X_{(j+1)} - X_{(j)}]^2 \right)^{\frac{1}{2}} \\ \sup_{\alpha \in \Theta} |D_n(\alpha)| &\leq K' \left(\frac{1}{N} \sum_{j=1}^N \sup_{\alpha \in \Theta} [X_{(j+1)} - X_{(j)}]^2 \right)^{\frac{1}{2}} \\ &\leq K' \sup_j |X_{(j+1)} - X_{(j)}|. \end{aligned}$$

By Lemma 1 $\mathbb{P}(K' \sup_j |X_{(j+1)} - X_{(j)}| \geq \epsilon) \xrightarrow{N \rightarrow +\infty} 0$. Hence D_N is bounded by an integrable and continuous function which goes to 0 in probability on Θ

$$\sup_{\alpha \in \Theta} |\hat{M}_N(\alpha) - M(\alpha)| \xrightarrow{N \rightarrow +\infty} 0.$$

So we may conclude.

ii) First, we use that $\sqrt{N}(\hat{\alpha}_N - \alpha^*) = \sqrt{N}(\hat{\alpha}_N - \hat{\alpha}_N) + \sqrt{N}(\hat{\alpha}_N - \alpha^*)$. By Theorem 3, $\sqrt{N}(\hat{\alpha}_N - \alpha^*) \xrightarrow{N \rightarrow +\infty} \mathcal{N}(0, \Gamma_{\alpha^*})$ with Γ_{α^*} defined in (6). It remains to prove that $\sqrt{N}(\hat{\alpha}_N - \hat{\alpha}) \xrightarrow{N \rightarrow +\infty} 0$.

Using the same arguments as in the proof i), we have

$$\mathbb{P}(\sqrt{N} \sup_{\alpha \in \Theta} |\hat{M}_N(\alpha) - M_N(\alpha)| \geq \epsilon) \leq \mathbb{P}(K \sqrt{N} \sup_{\alpha \in \Theta} |X_{(j+1)} - X_{(j)}| \geq \epsilon) \quad (15)$$

The right hand side of (15) converges to 0 by Lemma 1. This implies that $\sqrt{N}(\hat{\alpha}_N - \hat{\alpha}) \xrightarrow{N \rightarrow +\infty} 0$.

■

5 Perspectives and conclusion

One of the main quality of our approach is that it is easy to implement and execute. The cost function being simple, we use a BFGS algorithm to find the optimal parameters, and because of the regularity of curves we deal with, the initial points for optimization can be easily defined.

Furthermore, the search of the coordinate of the reference curve which is sent to the coordinate of the curve to fit can be easily implemented with a simple value search.

Besides, the deformation parameters can be exploited through an explainable model such as the linear model used in the real world problem.

It seems that the deformation model is robust if the noise is controlled. An interesting extension of this work would be to study what is going on when the reference curve is noisy. A generalization of this work to less regular functions would be worthwhile. Finally, it would be interesting to include in the model a way to handle discontinuities in order to reduce the dimension and have a more global representation of the deformation.

References

- [1] Allasonniere, S., Bigot, J., Glaunes, J., Maire, F., F., R.: Statistical models for deformable templates in image and shape analysis. *Annales Mathematiques Blaise Pascal* **20**(1), 1–35 (2013)
- [2] Barron, A.R., Cohen, A., Dahmen, W., DeVore, R.A.: Approximation and learning by greedy algorithms. *Ann. Statist.* **36**(1), 64–94 (2008). DOI 10.1214/009053607000000631. URL <https://doi.org/10.1214/009053607000000631>
- [3] Bigot, J., Gadat, S.: A deconvolution approach to estimation of a common shape in a shifted curves model. *Ann. Statist.* **38**(4), 2422–2464 (2010). DOI 10.1214/10-AOS800. URL <https://doi.org/10.1214/10-AOS800>
- [4] Bigot, J., Gamboa, F., Vimond, M.: Estimation of translation, rotation, and scaling between noisy images using the fouriermellin transform. *SIAM Journal on Imaging Sciences* **2**, 614–645. ISSN 2936-4954
- [5] Dimeglio, C., Loubes, J.M., Maza, E.: Manifold embedding for curve registration (2011). URL <https://hal.archives-ouvertes.fr/hal-00593180>. Working paper or preprint
- [6] Fournier, E., Grihon, S., Klein, T.: A case study : Influence of Dimension Reduction on regression trees-based Algorithms -Predicting Aeronautics Loads of a Derivative Aircraft (2018). URL <https://hal.archives-ouvertes.fr/hal-01700314>. Working paper or preprint
- [7] Gamboa, F., Loubes, J.M., Maza, E.: Semi-parametric estimation of shifts. *Electron. J. Statist.* **1**, 616–640 (2007). DOI 10.1214/07-EJS026. URL <https://doi.org/10.1214/07-EJS026>
- [8] Grenander, U.: *General pattern theory*. Oxford Science Publications (1993)
- [9] Mallat, S.G., Zhang, Z.: Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing* **41**(12), 3397–3415 (1993). DOI 10.1109/78.258082
- [10] McGuire, M.: An image registration technique for recovering rotation, scale and translation parameters. Tech. rep., NEC Tech Report (1998). URL <http://casual-effects.com/research/McGuire1998ParameterRecovery/index.html>
- [11] Nocedal, J., Wright, S.J.: *Numerical Optimization*, second edn. Springer, New York, NY, USA (2006)
- [12] Ramsay, J.O., Li, X.: Curve registration. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **60**(2), 351–363. DOI 10.1111/1467-9868.00129

- [13] Ritter, M., Dillinger, J.: Nonlinear numerical flight dynamics for the prediction of maneuver loads. IFASD 2011 pp. 1–10 (2011)
- [14] Sancetta, A.: Greedy algorithms for prediction. *Bernoulli* **22**(2), 1227–1277 (2016). DOI 10.3150/14-BEJ691. URL <https://doi.org/10.3150/14-BEJ691>
- [15] Van der Vaart, A.W.: Asymptotic statistics. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press (1998)