Detection of anomalous vibrations in an aircraft gas turbine engine

Eustache Besançon, Andrea Bondesan, Aurélien Citrain, Marwa Dridi

DETECTION OF ANOMALOUS VIBRATIONS IN AN AIRCRAFT GAS TURBINE ENGINE

EUSTACHE BESANÇON[1], ANDREA BONDESAN[2],AURÉLIEN CITRAIN,[3] AND MARWA DRIDI[4]

Abstract. We consider the compressor disk of a gas turbine engine composed of 13 blades showing anomalous vibrations around the normal vibratory modes. We use a statistical approach to study the correlation between the anomalous vibrations and the geometrical properties of the blades. We eventually detect the defective blade and provide a mathematical and numerical justification of this choice.

1. Introduction

An aircraft gas turbine engine is the component of the propulsion system for an aircraft that generates mechanical power from chemical energy. It is formed by an upstream rotating compressor coupled to a downstream turbine, and a combustion chamber in between. The compressor part of the propulsion system is made up of stages that consist of rotating bladed disks and stationary stators, or vanes. As the air moves through the compressor, accelerated by an initial centrifugal force, its temperature and pressure increase, until it reaches the combustion chamber where it is used to ignite the fuel sprayed inside, generating a high-temperature flow. This high-temperature high-pressure gas enters a turbine where it expands down to the exhaust pressure, producing a shaft work output in the process that is used in part to propulse the aircraft and in part to drive the compressor, in order to restart the process.

The main problem that we want to solve is the following. A new aircraft turbine engine is conceived, built and finally tested in order to study its efficiency. Inside the engine, a compressor disk composed of 13 blades shows a malfunction in each test that is made: while the engine is activated and the compressor disk rotates, by means of a tip-timing analysis, an anomalous vibration is detected for one of the blades (see Figures 3 and 4). Since the engine is a closed environment, we are not able to identify the defective blade in an easy way. In fact, if we number the blades in ascending order from 1 to 13 during each test, we can only deduce that one blade (and always the same) produces the anomalous vibration, but we cannot establish its correct position in the numbering (since we could have a circular shift of the blades from one test to the other). Nevertheless, fundamental geometrical properties (94 distinct measures of thickness for each blade, as in Figure 1, and main geometrical data on leading and trailing edges, as in Figure 2) are available for each of the blades and the idea is to find a way to link these data to the vibrations in all their possible circular shifts, in order to understand which one is the most probable (i.e. which one is the defective blade). We will solve this problem making use of well known statistical techniques.

The purpose of multivariate statistical analysis is to simultaneously analyse multiple measurements on a population of individuals or objects of interest. It is commonly assumed that there exists a direct relationship between each object under investigation and the set of observations represented by the values of the variables being measured, i.e. that it is possible to tell precisely for each measurement, to which individual it belongs. If this assumption is broken, for instance

[1]Telecom ParisTech, École doctorale de mathématiques Hadamard (MH, ED 574), Université Paris-Saclay (eustache.besancon@telecom-paristech.fr).

[2]MAP5 UMR CNRS 8145 – Sorbonne Paris Cité – Université Paris Descartes 75270 Paris Cedex 6, France (andrea.bondesan@parisdescartes.fr).

[3]INSA Rouen-Normandie Université, LMI EA 3226, 76000, Rouen & Team project Magique.3D, INRIA.UPPA.CNRS, Pau, France (aurelien.citrain@insa-rouen.fr).

[4]Laboratoie de Mathématiques - Analyse, Probabilités, Modélisation - Orléans (MAPMO), UMR CNRS 7349 – Fédération Denis-Poisson (marwa.dridi@univ-orleans.fr).

because of technical constraints on the way data are being measured, statistical literature gives little attention to how statistical methods can still be applied. This article focuses on an instance where this link is broken but where the type of relationship between the objects under investigation and the variables being measured is known: there exists a circular rotation between the lines of the array of variables and the set of objects. We devise a procedure which enables to identify the right rotation and to ultimately relate the observations to the objects under investigation.

Our work is organized as follows. In Section 2 we perform a principal component analysis and we settle a suitable statistical test on our data to identify the defected blade. In Section 3 we justify mathematically and numerically our former analysis, giving a complete answer to the problem.

2. Preliminary analysis: PCA and statistical tests

We begin introducing the following useful notation. We will refer to the numbering of the vibrations in Figure 4 as S_0, the 'S' standing for (circular) "shift"; any possible circular shift of $1 \leq n \leq 12$ positions of the vibrations will be denoted S_n, meaning that the vibration 1 in S_0 becomes vibration $13-(n-1)$ in S_n (we stop at $n=12$ since we easily check that $S_{13}=S_0$).

The problem then becomes to understand which of the circular shifts S_n, for $0 \leq n \leq 12$, corresponds to the geometrical data known for the blades (ordered as in S_0). To do this, we will perform a statistical analysis studying their correlation.

Since we are dealing with a lot of data, we first need to understand which ones contain most of the statistical information. Namely, we carry out a Principal Component Analysis (PCA) to identify possible useless data that can be ignored. In fact, Figure 5 shows that almost 80% of the cumulative variance is explained by the first principal component calculated by the PCA test, meaning that F_1 accounts for the most variability in the data. In particular, if we analyse this first eigenvector we observe that the most important weigth of informations comes from the thicknesses: from Figure 6 we immediately see that for F_1 the thicknesses are concentrated in the far right region of the circle where the weight is close to 1, while the other geometrical variables are diffused inside the circle with a negligible weight. As a consequence, we decided to focus our attention only on the geometrical data coming from the 94 points of thickness on the blades.

The next step in order to identify the defective blade is then to choose an adapted non-parametric statistic to measure the correlation between these thicknesses and the vibrations in all the possible circular shifts S_n.

To do so, we take the correlation coefficient to be the most likely to fit our data. Given the complexity of the interactions between all the blades in the disk, one of the first observations we have made is that the relation between vibrations and thicknesses is certainly nonlinear and moreover the set of vibrations does not follow a Gaussian distribution, which means that correlation coefficients like Pearson's r are not adapted to our situation and we need to take a look at a rank correlation coefficient (RCC).

Among all possible RCCs the largest used is Spearman's ρ, but it fails to give a good measure of correlation when dealing with a lot of (nearly) equivalent data, which is the case for most of the variables in our data set. This is why our choice falls on Kendall's τ RCC, firstly introduced by Maurice G. Kendall in 1938 in [4]. In fact, Kendall's τ has been classically used to test the significance of cross-correlation between two variables X and Y when their distributions significantly deviate from the normal law.

Following [3], we know that Kendall's τ is a statistic defined as the difference between the probabilities of concordance and discordance between two observed variables X and Y, namely

$$\tau = \mathbb{P}(y_i < y_j | x_i < x_j) - \mathbb{P}(y_i > y_j | x_i < x_j).$$

When dealing with finite sample of observations we only need an estimate for this statistic and one possible choice is the following.

Definition 2.1. *For a sample of n observations of X and Y, $\{(x_1, y_1), \ldots, (x_n, y_n)\}$, an estimate t for τ can be defined as*

$$t = \frac{2T}{n(n-1)}, \tag{1}$$

where the statistic T is given by

$$T = \sum_{i<j} sign(x_j - x_i) sign(y_j - y_i).$$

In other words, the estimate t measures the quantity

$$t = \frac{n_c - n_d}{n_0},$$

where, for two pairs (x_i, x_j) and (y_i, y_j), n_c denotes the number of concordant pairs (i.e. $x_i < x_j,\ y_i < y_j$ or $x_i > x_j,\ y_i > y_j$), n_d denotes the number of discordant pairs (i.e. $x_i < x_j,\ y_i > y_j$ or $x_i > x_j,\ y_i < y_j$) and $n_0 = \binom{n}{2}$ denotes the number of all possible pair combinations.

Remark 2.2. *The range of Kendall's coefficient is $t \in [-1, 1]$, which means that*

- *if the agreement between the two rankings is perfect (the rankings are the same), the coefficient has value 1,*
- *if the disagreement between the two rankings is perfect (one ranking is the reverse of the other), the coefficient has value -1,*
- *if X and Y are independent, then the coefficient is expected to be approximately zero.*

We are now able to set a statistical test using Kendall's coefficient. Under the classical null hypothesis of independence between the two random variables X and Y, the test reads

- H_0 : The variables X and Y are independent (no correlation between the samples).
- H_1 : The variables X and Y are dependent (correlation between the samples).

From the computation of all Kendall's coefficients between the eigenvector F_1 (coming from the PCA test) and the vibrations (coming from all the possible circular shifts of the blades in the disk) we observe that two particular circular shifts show strong correlation between their corresponding vibrations and the vector F_1. We summarize in Table 1 the most relevant computations.

	Kendall's t	p-value
S_1 (Blade 10)	0.452	0.04
S_6 (Blade 5)	0.452	0.04

Table 1: Correlation between F_1 and the vibrations in the circular shifts S_1 and S_6.

In particular, we obtain that the p-values corresponding to the circular shifts S_1 and S_6 are lower than 0.1, which allows to ensure that in these two cases hypothesis H_0 can be rejected in favour of hypothesis H_1, that is to say, we have correlation between F_1 and the vibrations.

The last step is to infer which of the two circular shifts S_1 and S_6 is the correct one. Going back to the PCA test, we recall that the most important contributions to the principal component F_1 came from the thicknesses of the blades. This is why we decided to perform another Kendall's test, this time trying to understand the correlation between the vibrations in the two circular shifts S_1 and S_6 and each of the 94 sets of thicknesses. Doing so, the aim is to localize a strong correlation between vibrations and thicknesses in a precise area of the defective blade, allowing not only to conclude which of the two remaining circular shifts is the correct one (i.e. which of the 13 blades is the defective one), but also to understand the problem causing the anomalous vibration.

Studying the variation of Kendall's coefficient over the defective blade in each circular shift S_n, $0 \leq n \leq 12$, we easily check that the maximum of t (with a value of $t = 0.71$) is reached for S_1 on the points of thickness 24 and in the close surrounding area, leading us to deduce that the anomalous vibration is produced by blade 10 (see Figure 11).

A nice visual description of these variations is given in Figures 7, 8, 9 and 10. The defective blade is represented in every possible circular shift, where each of its 94 points of thickness is coloured depending on the value of Kendall's coefficient assumed in that particular point: points coloured in yellow (respectively green/blue) show a strong agreement (respectively independence/strong disagreement) between thicknesses in those points and vibrations in the corresponding circular shift. In other words, the points coloured in yellow exhibit the maximum values of Kendall's coefficient and we immediately see that the area with the highest values of t is localized on the defective blade for the circular shift S_1, this phenomenon strongly appearing only in this case.

The statistical analysis presented in this section allows to infer that the blade producing the anomalous vibration is actually blade 10 (circular shift S_1), the problem coming from the defective area around the point 24 of thickness.

We conclude our study in the next section, providing a mathematical and numerical justification of our study.

3. Mathematical and numerical validation

We are interested in justifying mathematically the fact that the anomalous vibration comes from blade 10 (corresponding to the circular shift S_1) in the area surrounding the point 24 of thickness. To do so, the correct approach is certainly the probabilistic one. As a consequence, we start choosing a suitable distribution for Kendall's t, since it is the only statistical tool we used in the previous section to study the problem.

Following [3], under the null hypothesis that the two series X and Y are independent, and assuming that observations in each time series are independent, the mean and the variance of T in (1) are given by

$$\mathbb{E}(T) = 0, \qquad \mathrm{Var}(T) = \frac{n(n-1)(2n+5)}{18}$$

and accordingly, the mean and the variance of t are

$$\mathbb{E}(t) = 0, \qquad \mathrm{Var}(t) = \frac{2(2n+5)}{9n(n-1)}. \tag{2}$$

The authors in [3] show that the distribution of T (and accordingly that of t) tends to normality as the number of observations becomes large enough. In particular, for $n > 10$, the normal curve gives a satisfactory approximation of the distributions of T and t.

The problem becomes now to establish which is the correct distribution for the maximum value of Kendall's coefficients between vibrations and thicknesses. Fortunately, the theory of extreme value distributions helps us in giving a complete answer to the problem.

Let us consider a set of N i.i.d. variables $\{X_1, \ldots, X_N\}$ with common law F and their maximum value $M_N = \max\{X_1, \ldots, X_N\}$. We introduce the following definitions.

Definition 3.1. *We say that a distibution G is of exteme value-type if it has one of the following forms:*

1) $G(x) = e^{-e^{-x}}, \quad x \in \mathbb{R}, \qquad$ *(Gumbel)*

2) $G(x) = \begin{cases} 0 & x \leq 0 \\ e^{-x^{-\alpha}} & x > 0, \end{cases} \qquad$ *(Fréchet)*

3) $G(x) = \begin{cases} e^{-(-x)^{\alpha}} & x \leq 0 \\ 1 & x > 0, \end{cases} \qquad$ *(Weibull)*

for some $\alpha > 0$.

Definition 3.2. *Let F be the common distribution of the i.i.d. variables $\{X_1, \ldots, X_N\}$, with maximum value M_N. We say that F belongs to the Max-Domain of Attraction (MDA) of an extreme value distribution G if there exist two constants $a_N > 0$ and $b_N \in \mathbb{R}$ such that*

$$\lim_{N \to +\infty} \mathbb{P}\left(\frac{M_N - b_N}{a_N} \leq x \right) = G(x) \quad \forall\, x \in \mathbb{R}.$$

In particular, for a set $\{X_1, \ldots, X_N\}$ of i.i.d. variables distributed with common standard normal law $F \sim \mathcal{N}(0,1)$, the following result holds (see [5], Theorem 1.5.3):

Theorem 3.3. *If $\{X_i\}_{i=1}^N$ is an i.i.d. standard normal sequence of random variables, then the asymptotic distribution of $M_N = \max\limits_{1 \leq i \leq N} X_i$ is of Gumbel-type. specifically*

$$\lim_{N \to +\infty} \mathbb{P}\left(\frac{M_N - b_N}{a_N} \leq x \right) = e^{-e^{-x}} \quad \forall\, x \in \mathbb{R},$$

with

$$a_N = (2 \log N)^{-1/2}$$

and

$$b_N = (2 \log N)^{1/2} - \frac{1}{2}(2 \log N)^{-1/2}(\log \log N + \log 4\pi).$$

We are interested in applying to our problem the theory that we have presented. To do so, let us consider the set $\{t_i\}_{i=1}^{94}$ of all Kendall's coefficients in a single correlation test on the data (i.e. for one possible permutation of the vibrations), where t_i corresponds to Kendall's coefficient between the vibrations and the thicknesses at the points i of each blade in the disk (captured during the tip-timing analysis).

Reminding that each t_i follows a normal distribution $t_i \sim \mathcal{N}(0, \sigma^2)$ with $\sigma^2 = \dfrac{62}{1404}$ (using (2) with $n = 13$ observations), we deduce from Theorem 3.3 that the random variable

$$M = \max\left\{ \frac{t_1}{\sigma}, \ldots, \frac{t_{94}}{\sigma} \right\} \sim G(x) = e^{-e^{-\frac{x-b}{a}}} \quad \forall x \in \mathbb{R}$$

follows a Gumbel distribution of parameters

$$a = (2\log 94)^{-1/2}$$

and

$$b = (2\log 94)^{1/2} - \frac{1}{2}(2\log 94)^{-1/2}(\log\log 94 + \log 4\pi).$$

Now, we have seen in the previous section that, for the statistical test evaluating the correlation between vibrations and thicknesses in all the possible circular shifts of the blades in the disk, Kendall's coefficients are maximized in the area surrounding the point of thickness 24 on blade 10 (circular shift S_1). In particular, the maximum is reached exactly for the point 24 and we have

$$t_{24} = 0.71 \quad \text{and} \quad \frac{t_{24}}{\sigma} \approx 3.38 \ (\sigma \approx 0.21).$$

We conclude from our analysis that

$$\mathbb{P}(M \leq 3.38) = e^{-e^{-\frac{3.38-b}{a}}} \approx 0.96,$$

which means that the point 24 on blade 10 (circular shift S_1) corresponds to the point where the correlation between vibrations and thicknesses is maximized, with probability almost 1. This justifies our conclusion at the end of the previous section, allowing us to localize the precise area of the disk where the defects give rise to the anomalous vibration.

We conclude providing numerical evidence to our mathematical study with a simple test. We simulate 10^4 random permutations of the vibrations (thus, not only the circular shifts), evaluating for each of these permutations the maximum value of Kendall's coefficient between thicknesses and vibrations in the particular random sorting. We observe the result in Figure 12: the blue dots represent the maximum values of t calculated in each of the 10^4 random permutations of the vibrations, while the red line represents the maximum value atteined by t in the circular shift S_1 for the point 24 of thickness on blade 10. We immediately check that in 96% of the cases the specific permutation S_1 provides the maximal correlation between vibrations and thicknesses.

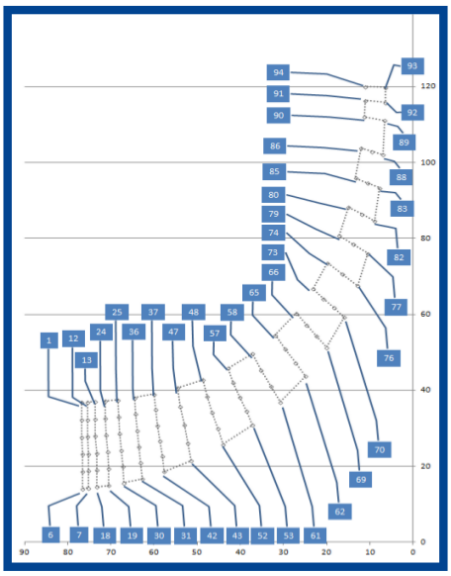

Figure 1: The 94 points on each blade where the thickness is measured.

Données géométriques

	Défaut de forme - Rayon de arccordement						Ray	
N° Pale	BA in	MI in	BF intra	BA ex	MI Ex	BF ex	BA in	MI in
1	0,077	0,064	0,047	0,054	0,035	0,053	4,43	2,03
2	0,068	0,073	0,059	0,056	0,039	0,053	4,44	2,02
3	0,074	0,072	0,055	0,059	0,038	0,047	4,44	2,03
4	0,07	0,076	0,058	0,056	0,037	0,062	4,46	2,02
5	0,086	0,06	0,041	0,042	0,043	0,065	4,42	2,03
6	0,066	0,066	0,046	0,048	0,047	0,063	4,47	2,02
7	0,079	0,059	0,05	0,049	0,048	0,05	4,46	2,04
8	0,076	0,075	0,056	0,053	0,042	0,051	4,43	2,01
9	0,08	0,066	0,058	0,046	0,05	0,053	4,42	2,02
10	0,086	0,062	0,054	0,049	0,047	0,07	4,42	2,02
11	0,087	0,066	0,047	0,054	0,04	0,046	4,42	2,03
12	0,081	0,068	0,045	0,051	0,039	0,05	4,42	2,02
13	0,086	0,061	0,049	0,047	0,048	0,073	4,41	2,02

Figure 2: Table containing some of the main geometrical data known on the blades.

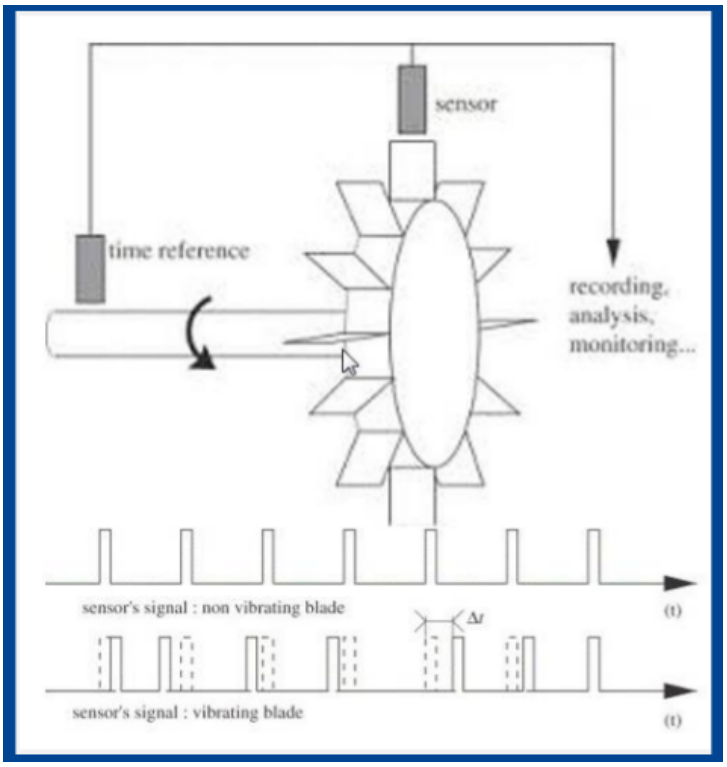

Figure 3: Tip-timing analysis: test for the monitoring of blades' vibrations.

Essai vibratoire		
N° Pale	Niveau vibratoire 1FPP (mm)	Vitesse de rotation - Régime (RPM)
1	0,54	16532,3
2	0,6	16538,4
3	0,61	16543,6
4	0,53	16560,2
5	0,43	16550,8
6	0,65	16605
7	0,38	16662,8
8	0,43	16615,3
9	0,52	16635
10	0,46	16639,5
11	0,99	16635,5
12	0,39	16633
13	0,41	16520,5

Figure 4: Vibrations and angular velocities of the blades during a single test: anomalous vibration detected on the blade 11 (for one of the 13 possible circular shifts of the disk).

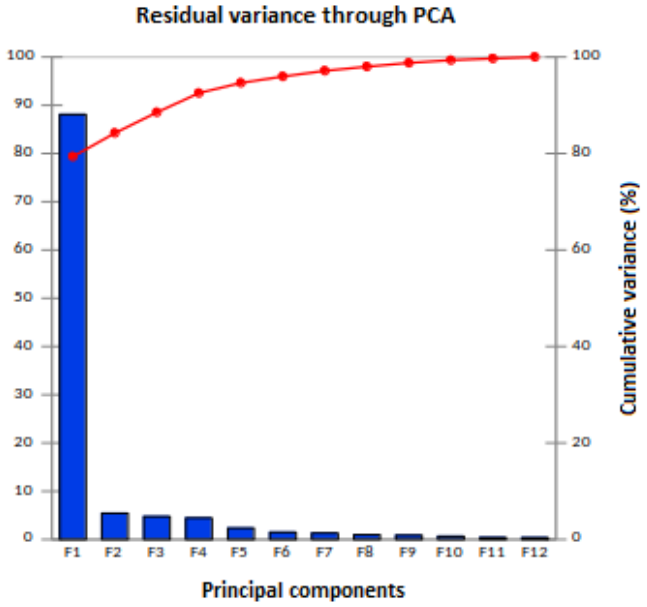

Figure 5: The principal components F_1 and F_2 explain 84.25% of the cumulative variance.

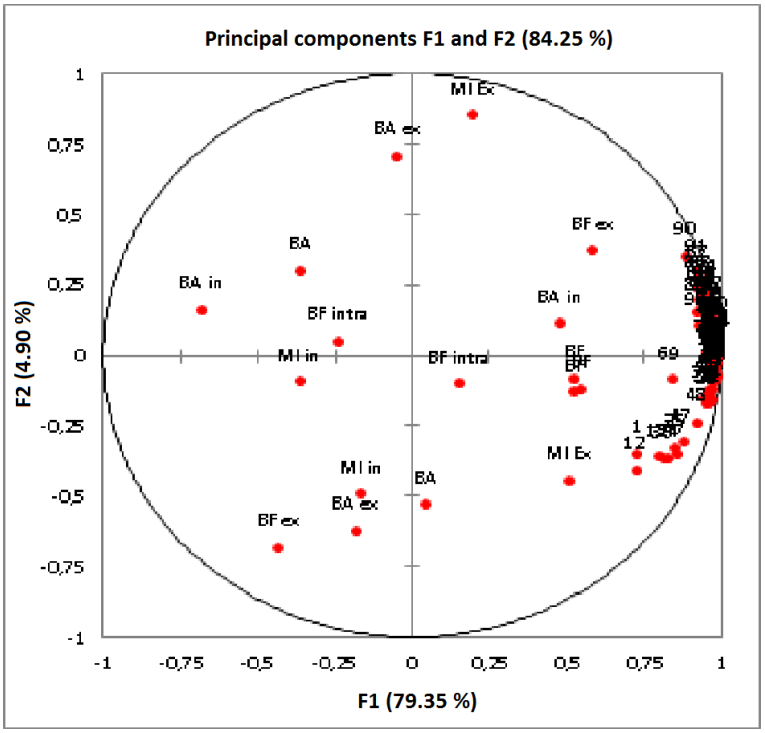

Figure 6: Weight of the geometrical variables in the first two principal components.

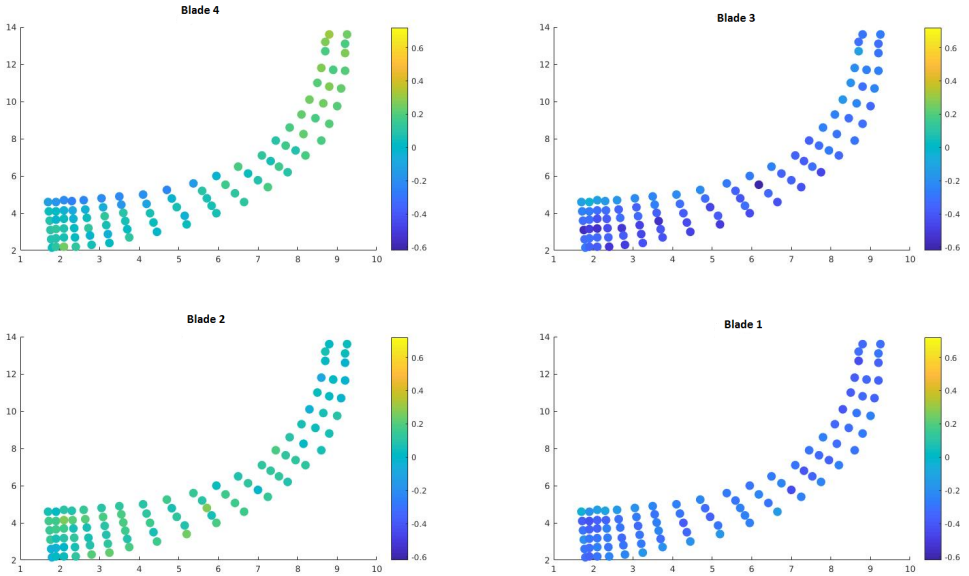

Figure 7: Variation of Kendall's t over the defective blade in the circular shifts S_7, S_8, S_9, S_{10}.

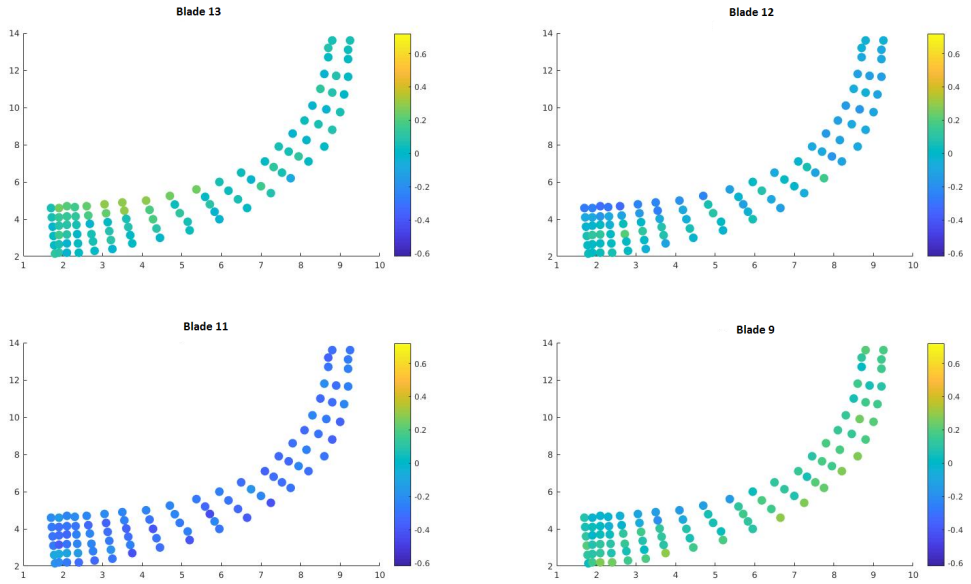

Figure 8: Variation of Kendall's t over the defective blade in the circular shifts S_{11}, S_{12}, S_0, S_2.

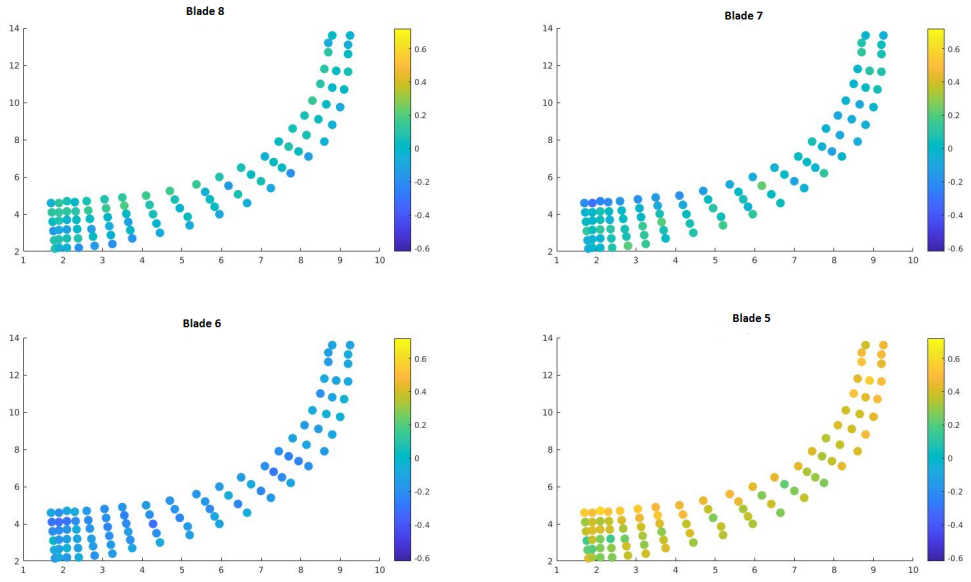

Figure 9: Variation of Kendall's t over the defective blade in the circular shifts S_3, S_4, S_5, S_6.

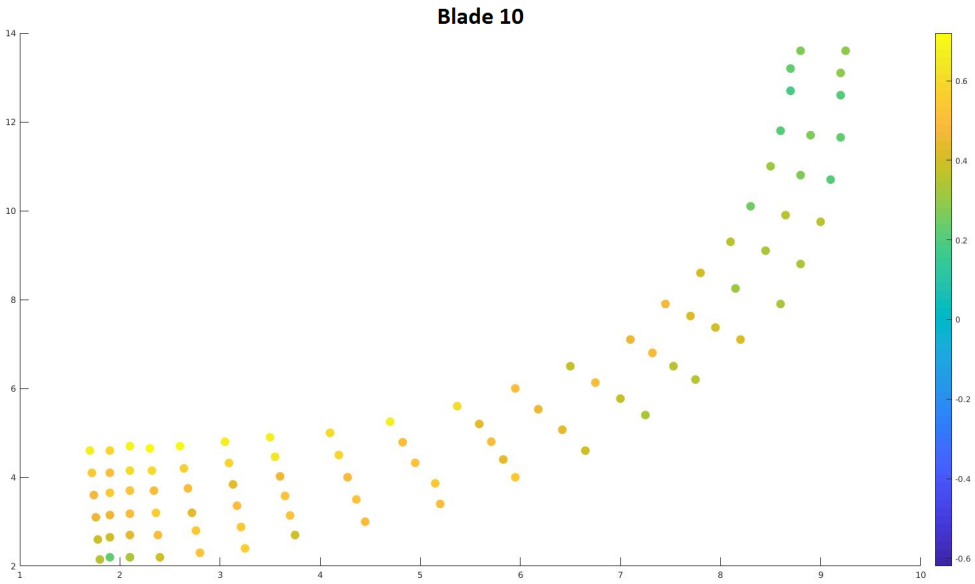

Figure 10: Variation of Kendall's t over the defective blade in the circular shift S_1.

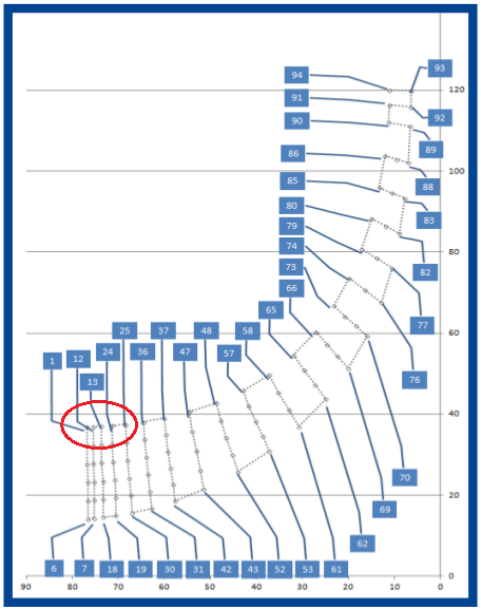

Figure 11: Precise area where the defect is localized on blade 10.

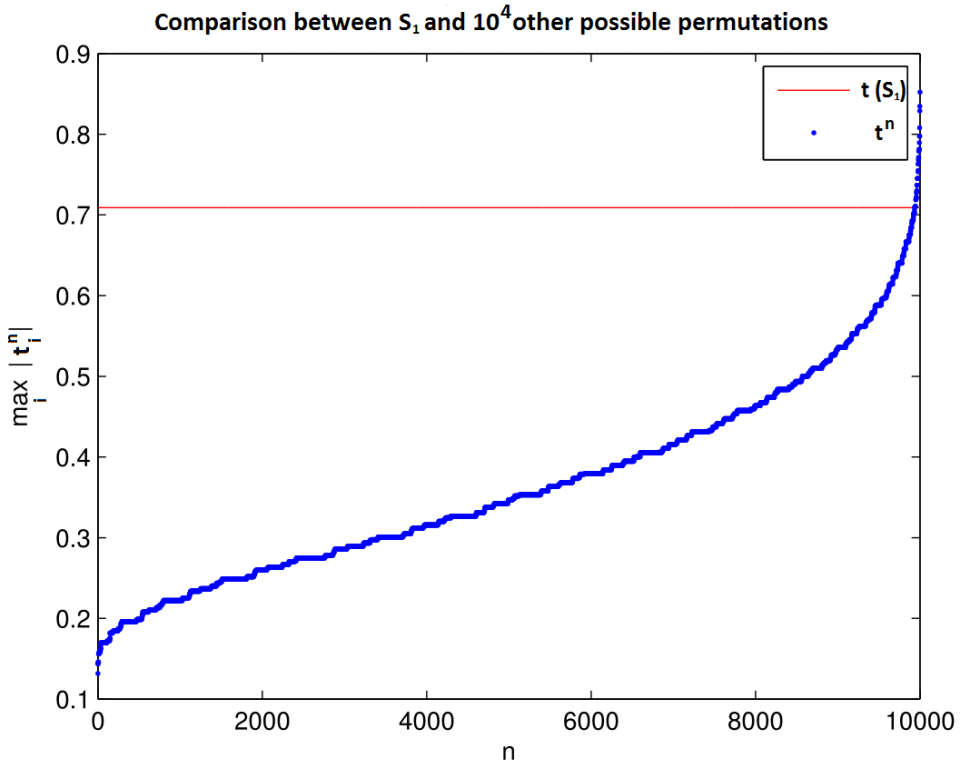

Figure 12: Variation of the maximum of Kendall's t between vibrations and thicknesses, for 10^4 random permutations of the vibrations.

Aknowledgements. The authors want to acknowledge AMIES for funding important scientific collaborations between universities and industries, SAFRAN Aircraft Engines for providing the interesting subject that lead to this study, Jérôme Lacaille for the fruitful discussions and Laboratoire MAP5 of Université Paris Descartes for its hospitality during the project and for its amazing organization of the SEME week.

References

[1] J. D. Gibbons and S. Chakraborti. *Nonparametric statistical inference*, volume 168 of *Statistics: Textbooks and Monographs*. Marcel Dekker, Inc., New York, fourth edition, 2003.

[2] K. H. Hamed. The distribution of Kendall's tau for testing the significance of cross-correlation in persistent data. *Hydrological Sciences Journal*, 56(5):841–853, 2011.

[3] M. Kendall and J. D. Gibbons. *Rank correlation methods*. A Charles Griffin Title. Edward Arnold, London, fifth edition, 1990.

[4] M. G. Kendall. A new measure of rank correlation. *Biometrika*, 30(1/2):81–93, 1938.

[5] M. R. Leadbetter, G. Lindgren, and H. Rootzén. *Extremes and related properties of random sequences and processes*. Springer Series in Statistics. Springer-Verlag, New York-Berlin, 1983.

[6] T. Rabenoro. *Statistical tools based on indicators aggregation for diagnostic and prognostic of aircraft engines.* Theses, Université Paris 1 Panthéon Sorbonne, September 2015.

[7] A. C. Sall. *Comportement dynamique d'un redresseur de turbomachine aéronautique : effets du désaccordage.* Theses, Ecole Centrale de Lyon, December 2011.