Confounding factors analysis and compensation for high-speed bearing diagnostics

Alessandro Paolo Daga[1], Luigi Garibaldi[1], Alessandro Fasana[1], Stefano Marchesiello[1]

[1]Politecnico di Torino, Italy
Department of Mechanical and Aerospace Engineering (DIMEAS),
Dynamic & Identification Research Group (DIRG),

Abstract

In recent years, machine diagnostics through vibration monitoring is gaining a rising interest. Indeed, in the literature many advanced techniques are available to disclose the fault establishment as well as damage type, location and severity. Unfortunately, in general, these high-level algorithms are not robust to operational and environmental variables, restricting the field of applicability of machine diagnostics. Most of industrial machines in fact, work with variable loads, at variable speeds and in uncontrolled environments, so that the finally measured signals are often non-stationary. The very common time-series features based on statistical moments (such as root mean square, skewness, kurtosis, peak value and crest factor) undergo variations related to changes in the machine operational parameters (e.g. speed, load, …) or in the environmental parameters (e.g. temperature, humidity, …), which can be seen as non-measured, and then latent, confounding factors with respect to the health information of interest.
In order to face such issue, statistical techniques like (in a first exploratory stage) the Principal Component Analysis, or the Factor Analysis, are available. The pursuit of features insensitive to these factors, can be also tackled exploiting the cointegration property of non-stationary signals.
In this paper, the most common methods for reducing the influence of latent factors are considered, and applied to investigate the data collected over the rig available at the DIRG laboratory, specifically conceived to test high speed aeronautical bearings monitoring vibrations by means of 2 tri-axial accelerometers while controlling the rotational speed (0 – 30000 RPM), the radial load (0 to 1800 N) and recording the lubricant oil temperature. The compensation scheme is based on two procedures which are established in univariate analyses, but not so well documented in multivariate cases: the removal of deterministic trends by subtraction of a regression, and the removal of stochastic trends in difference stationary series by subtraction of the one-step ahead prediction from an autoregressive model. The extension of these methods to the multivariate case is here analysed to find an effective way of enhancing damage patterns when the masking effect due to the non-stationarities induced by latent factors is strong.

Keywords: trend stationary, difference stationary, regression, autoregressive prediction, residuals, orthogonal regression, PCA whitening, Mahalanobis distance, cointegration, vector autoregression, novelty detection, damage detection, vibration monitoring, bearings.

1 Introduction

Vibration Monitoring (VM) is a particular kind of condition monitoring which exploits vibration as a condition indicator. Indeed, an online non-destructive testing (NDT) based on vibration can be set up to monitor the health condition of the machine while in operation. This turns out to be fundamental in Condition-Based Maintenance (CBM) regimes, in which the maintenance is not programmed but preventive, and must then rely on diagnostics and prognostics. The advantage of VM against other techniques such as Oil debris analysis, Performance analysis, thermography, Acoustic analysis or Acoustic Emissions AE, etc. is related to the high reactivity to sudden changes in a machine, and to the flexibility of the accelerometers (i.e. the vibration sensors), which are not only cost effective and reliable, but also small and light so that almost any machine can be easily instrumented.

A vibration-CBM is basically a Data-to-Decision (D2D) process [1] but the present work will focus mainly on the signal and the pattern processing parts, namely that of selecting and extracting damage-sensitive features and that of building and validating a statistical model based on the data whose scope is the detection of a damaged condition (***data-based modelling***). Damage detection is commonly considered the first fundamental step of diagnostics and consists in producing an indication of the presence of a damage, possibly at a given confidence, so that an alarm can be eventually triggered in case of danger. This can be performed by looking for the symptoms which indicates the presence of a faulty condition (i.e., recognizing patterns in the data). Such damage distinguishing characteristics can be effectively highlighted only by the extraction of proper features, namely quantities which show:

- Damage Consistency (i.e. they vary with damage)
- Damage Sensitivity and Noise-Rejection ability (i.e. they are sensitive also to small, incipient damage),
- Low sensitivity to unmonitored confounding factors.

A perfect feature is then able to reject any influence other than damage, producing stationary sets of indices for which the departure from the normal condition can only be ascribed to a malfunctioning. In this case the detection of novelty corresponds to the detection of damage and is then a relatively easy task.
In reality, researchers will always deal with features affected by operational (e.g. speed, load, …) and environmental (e.g. temperature, humidity, …) variations, which can be seen as latent (i.e., non-measured), confounding factors that can compromise the correct damage detection.
The scope of this paper is to highlight some technique for compensating such confounders with a particular focus on damaged bearings data. In order to cope with the need of a fast and automated real-time damage detection, an analysis is here proposed, based on the common time-series features (i.e., root mean square, skewness, kurtosis, peak value and crest factor (peak/RMS)) which are known to be sensitive to bearing damage but also to the operational conditions of the machine under analysis. The Novelty Detection then, must be preceded by some algorithm compensating for the confounders. In particular, the **regression** and the **cointegration** will be discussed in section 2. The experimental data used in this work refers to the high-speed aeronautical bearings test rig available at the Department of Mechanical and Aerospace Engineering of Politecnico di Torino, shortly introduced in next section.

1.1 The test rig and the dataset

The dataset considered in this analysis comes from a test rig built by the Dynamic & Identification Research Group (DIRG), part of the Department of Mechanical and Aerospace Engineering of Politecnico di Torino, to test high-speed aeronautical bearings. The rig is fully described in [2], but the main information is summarized hereinafter. The rig is made by a single direct-drive rotating shaft supported by two identical high-speed aeronautical roller bearings (B1 and B3 in Figure 1). B3 is known to be healthy while B1 is damaged by purpose with indentations of different size in different parts of the bearing (Rolling Element and Inner Ring) as described in Table 2. The third central bearing B2 is mounted on a sledge meant to load the shaft with an increasing force of 0, 1000, 1400 and 1800 N, while the speed is reducing from 470 to 0 Hz (run-down acquisitions). Table 1 summarizes the operational conditions. Two tri-axial accelerometers located respectively on the B1 bearing support (accelerometer A1, as reported in Figure 1) and on the loading sledge (accelerometer A2). The acquisitions last for a duration of about $T = 50\,s$ at a sampling frequency fs = 102400 Hz. In order to perform a significant analysis, the five selected features *root mean square*, *skewness*, *kurtosis*, *peak value* and *crest factor* are extracted on one hundred independent chunks (about 0,5 s each) for each of the 6 channels of the 4 original acquisitions in all the 7 health conditions (from 0A, healthy, to 6A).
Finally, 100 observations in a 30-dimensional space (6 channels, 5 features) per each health conditions are obtained. A part of the dataset is visually summarized in Figure 2.

Table 1. The operational conditions: the different loads while the speed is decreasing from 470 to 0 Hz (run-down acquisitions).

Label	1	2	3	4
F [kN]	0	1	1,4	1,8

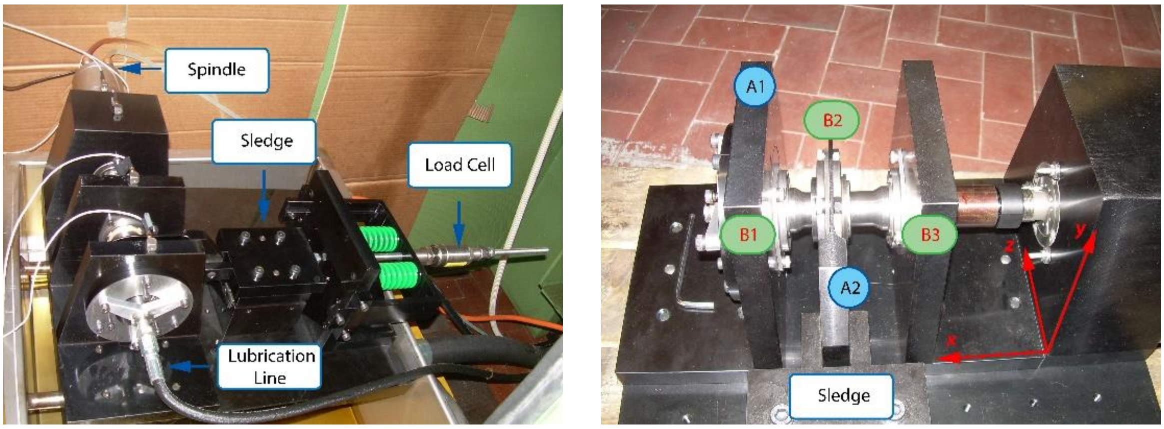


Figure 1. The experimental setup, the triaxial accelerometers location (A1 and A2) and orientation.

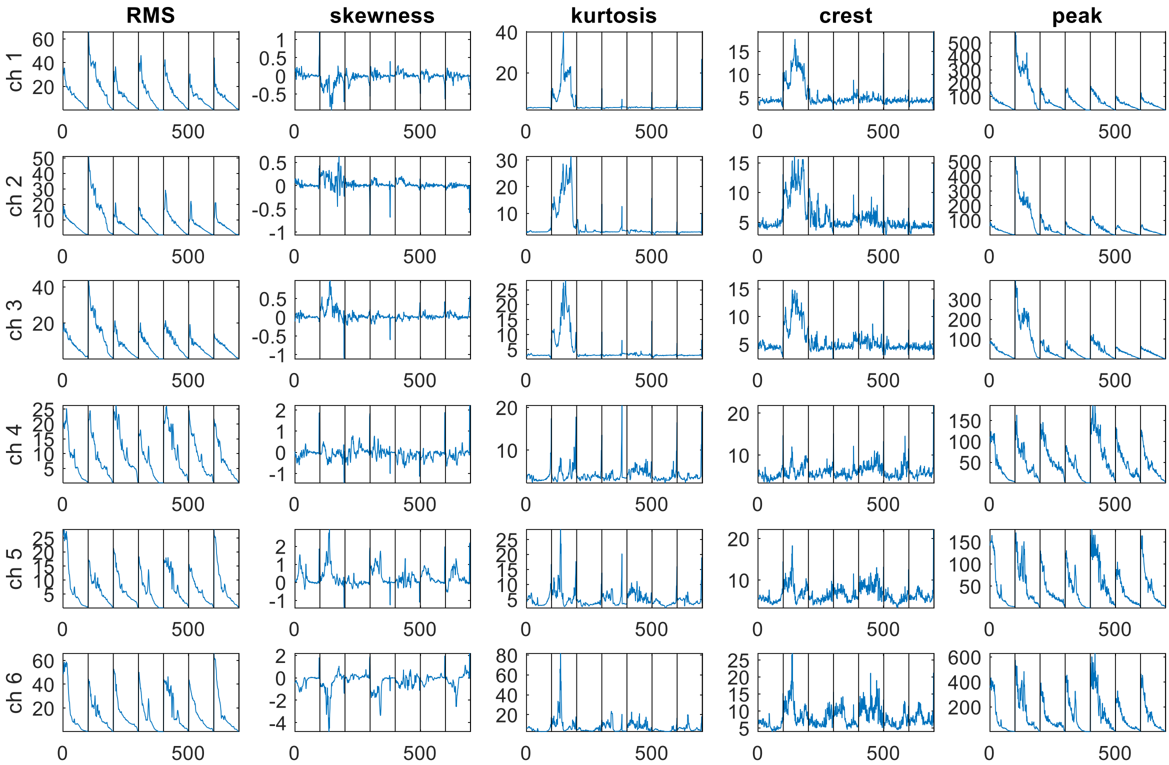


Figure 2. The considered dataset after features extraction for load condition 1 (0 N) while the speed is decreasing until a stop starting from 470 Hz. The black dotted lines divide the different damage conditions (0A to 6A). For each, 100 observations are plotted sequentially.

Table 2. Bearing B1 codification according to damage type (Inner Ring or Rolling Element) and size. The damage is obtained through a Rockwell tool producing a conical indentation of maximum diameter reported as characteristic size.

Code	0A	1A	2A	3A	4A	5A	6A
Damage type	none	Inner Ring	Inner Ring	Inner Ring	Rolling Element	Rolling Element	Rolling Element
Damage size [μm]	-	450	250	150	450	250	150

Figure 3. Stationary stochastic process and the biconditional relationship of novelty and damage.

2 Methodology

This work is devoted to the application of Novelty Detection on time-series features extracted from the raw acceleration data of a test rig. The idea of "novelty" is commonly related to that of "outlier", a discordant measure inconsistent with the others and then believed to be generated by an alternate mechanism.
The judgment on discordancy will depend on a measure of distance from a reference distribution (e.g., healthy), usually called Novelty Index (NI), on which a threshold can be defined [3].
This very simple idea can be exploited for Damage Detection when the healthy vibration signal can be modelled as a stationary stochastic process, meaning that the joint probability distribution function is invariant under time translation, so that damage is left as the only possible cause of discordancy (Figure 3).
Unfortunately, hidden latent (non-measured) factors will always affect the measurements. When their effect is important, non-stationarities will arise, leading to misinterpretations of the novelty (and then damage), so that they are often referred to as confounders (Figure 4).

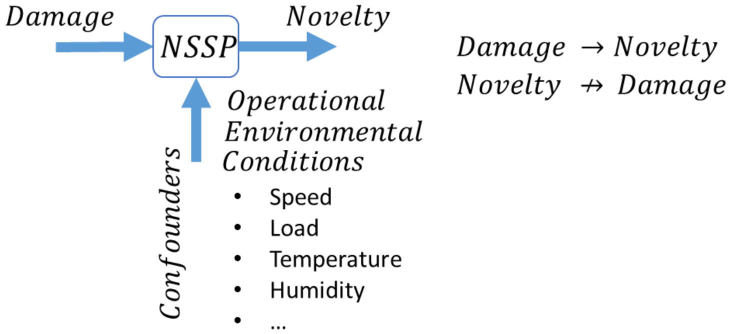


Figure 4. Non-stationary stochastic process and the effect of confounders

In these cases, a wise feature selection able to reduce the effect of the confounders may be important but is usually not enough for solving this issue. Algorithms for compensating such effects become then essential.
In this work, the temperature is controlled, and the load is kept constant so that the only confounder is the variable speed, which strongly affect time features as the peak level and the RMS of the acceleration signal (see Figure 2). The measurement involves an uncontrolled braking of the machine from full speed (470 Hz) to a stop. The features from the first channel of the first accelerometer with a null load are reported in Figure 5.

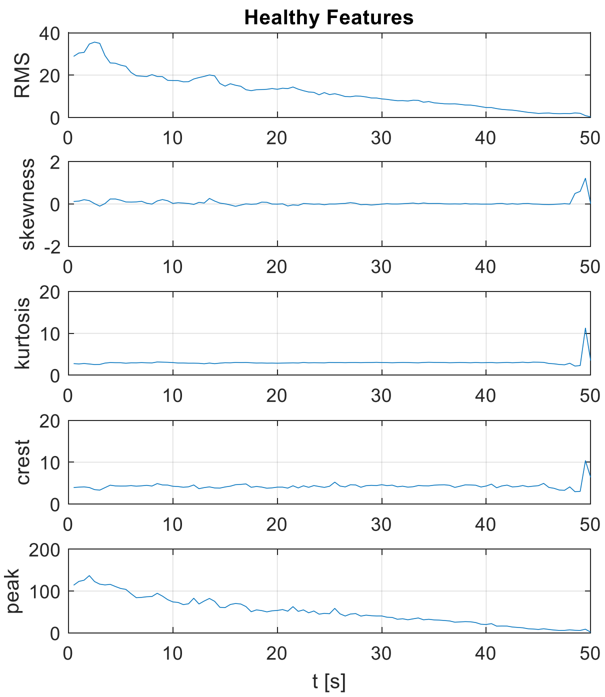


Figure 5. The five selected features from the first channel of the first accelerometer – 0N.

Stationarity is noticeably violated as a trend is clearly visible in Figure 5. In the literature [4] two simple models for such violations can be found.

The first involves a deterministic trend, so that the resulting signal takes the name of **trend stationary**. A polynomial fitting can be used in this case to find and subtract the trend, leading to a stationary residual which is said to be "white" as the resulting frequency spectrum turns out to be flat (i.e. the residual is a white noise) or "decorrelated" as its autocorrelation is null for any lag different from 0.

$$y_t = \beta t + \varepsilon_t \qquad\qquad \varepsilon_t \sim N(0, \sigma^2)$$

A second model on the contrary involves a stochastic trend. In the simplest case, this means that the increment in the signal from time to time (innovation) is defined as a stochastic process ε such that

$$y_t - y_{t-1} = \varepsilon_t$$

In this case the signal y is the result of the integration of the considered stochastic process ε and is then called integrated of order 1 or $I(1)$. This process which corresponds to a random walk is **difference stationary** as its first difference is stationary. Again, it is possible to get a stationary signal which can be considered white.

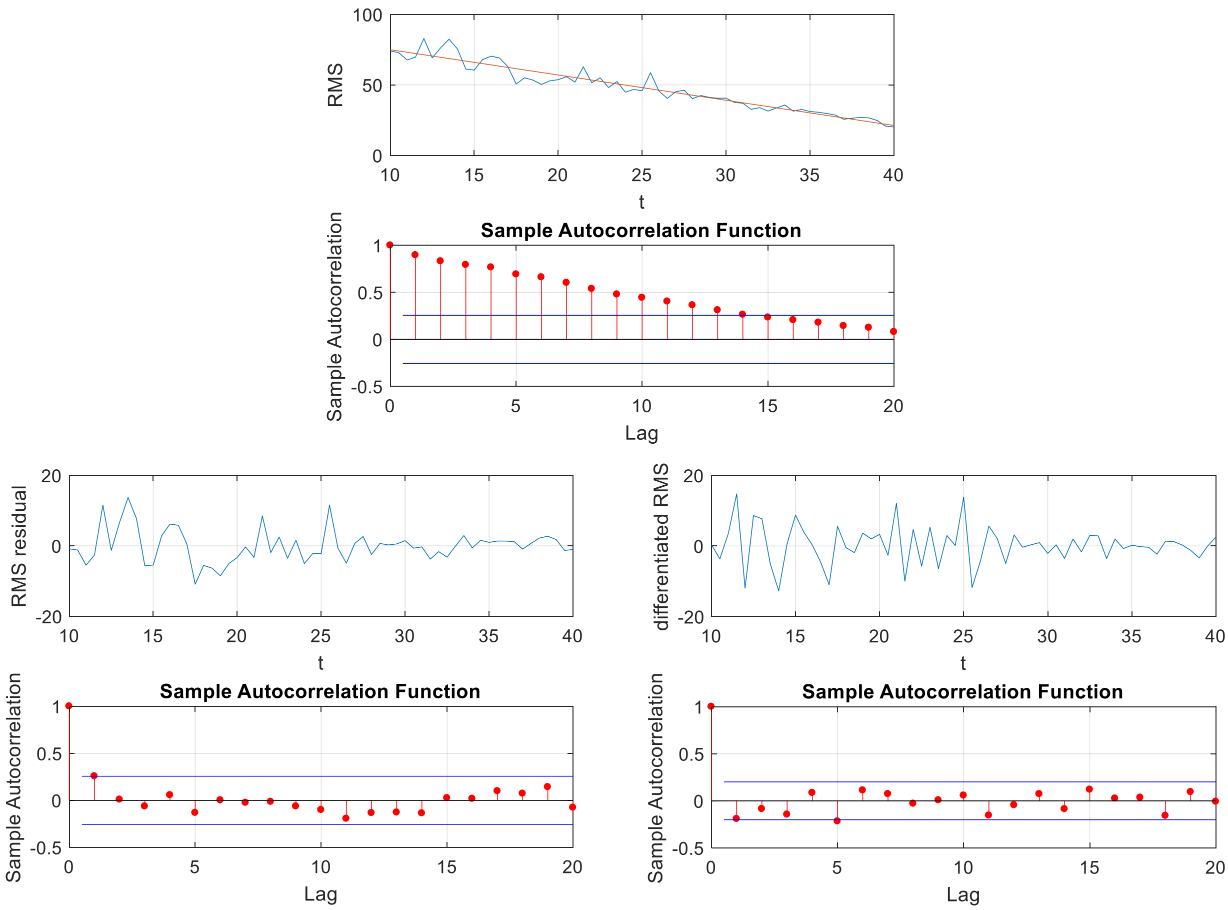


Figure 6. The raw RMS from 10 to 40s and its autocorrelation function (ACF). Below, on the left the residual from removing the linear regression and its ACF, on the right, the differencing (equivalent to the residual after an AR(1) fit) and its ACF.

To generalize, the random walk can be considered as a particular case of an autoregressive AR(1) model with a unitary coefficient. That explains why it is very common in the literature to whiten the data by fitting an AR(1) to the series and focus on the residual, as done in [5] to highlight the damaged bearing signature.

These concepts can be extended to multivariate spaces. When the features are affected by the same confounder in fact, they turn out to be strongly correlated (in simple terms, they vary in sympathy).

Under the first assumption (trend stationarity) then, a multivariate regression can be used. In this case, considering that both the variables are affected by measurement errors and that it is not easy to find a dependent and an independent quantity, the orthogonal regression [6] based on PCA is proposed.

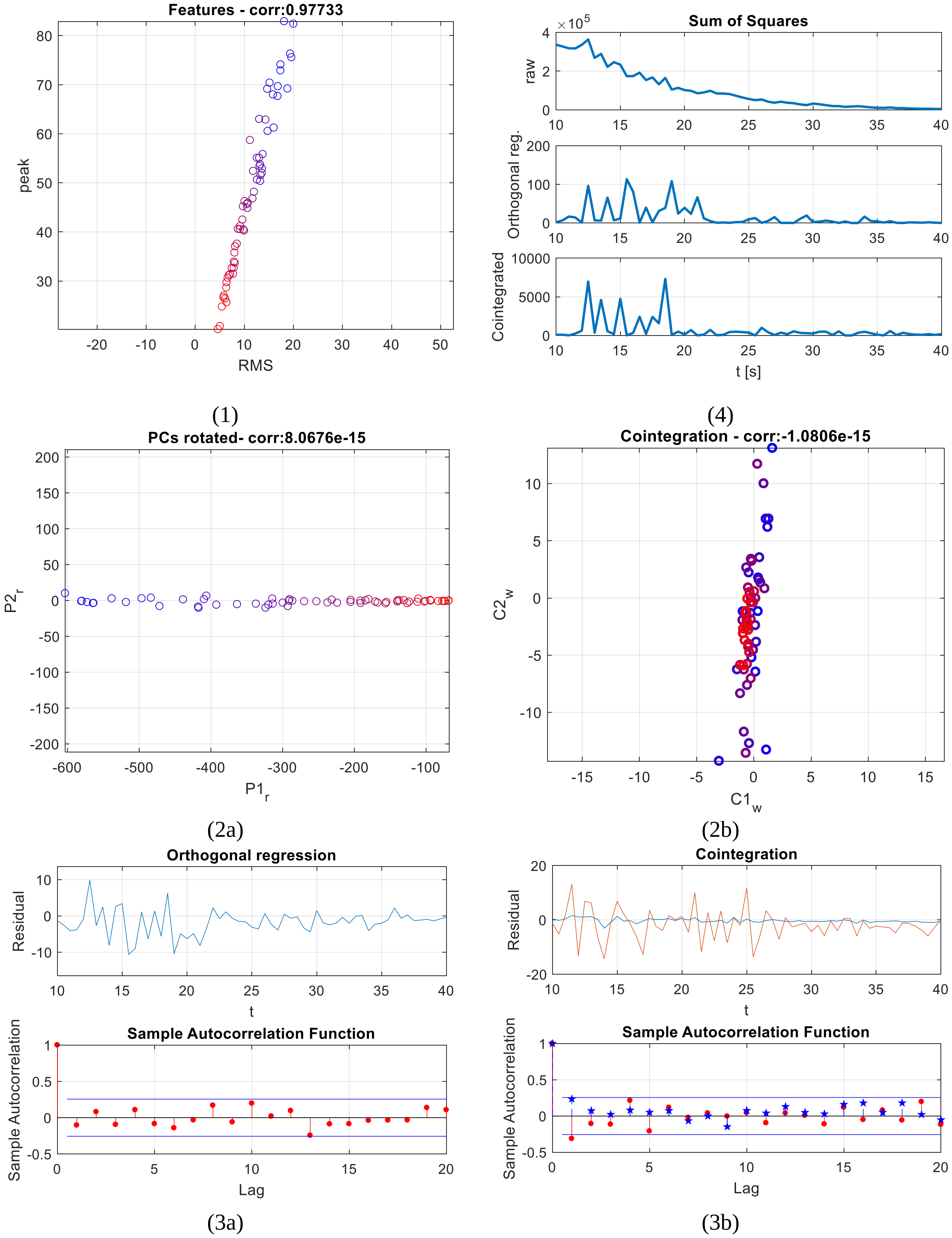


Figure 7. (1) The bivariate scatterplot with time evolution (from blue – 10s to red – 40s)
(2a) PCA rotation (3a) Rotated PC2 corresponding to the residual and its ACF
(2b) coitegration scatterplot (3b) stationary residual from cointegration and its ACF
(4) sum of squares of 3a and 3b components compared to the raw Euclidean squared distance

Orthogonal regression is fundamentally a reconstruction of the dataset in a subspace obtained by neglecting the first principal component. This methodology can be merged to the PCA whitening, to directly obtain a white, unitary covariance residual.
PCA orthogonal regression and whitening are mathematically tackled in section 2.1.
On the contrary, difference stationary multivariate series can be whitened by fitting a vector-auto-regressive VAR(1) model and computing the multivariate residual. In this case the dimensionality is not affected. Nevertheless, the considered variables must be I(1), so that a statistical hypothesis test such as the Augmented Dickey-Fuller is needed [4]. This procedure commonly takes the name of cointegration [7,8] and is described in section 2.2.

The results of a simple bivariate analysis on the RMS and the Peak value of the first channel are reported in Figure 7 to show the ability of the two methods on a real acquisition. In case of real measurements in fact, as confirmed by this simple analysis, it may be difficult to confidently identify the underlying model as both may work in a quite proper way.
Nevertheless, keeping focused on the final scope of detecting novelty (and damage), a relevant consideration can be made about novelty indices (NI). Novelty detection in fact, is commonly based on the Mahalanobis distance [3,9] which is known to be equivalent to a Euclidean distance on a features space rotated to match the principal components and normalized to obtain unitary variance PCs [2].
Hence, the squared Mahalanobis distance equals the sum of the squared whitened principal components. Obviously, it involves also the first PC which pictures the confounding factor, so that it is not stationary, as shown in Figure 8. A good idea then is to use as novelty index the sum of the squared principal components rejecting the first or first few.
The same idea of summarizing the residuals with a single novelty index computed as the sum of squares of the cointegrated series can be extended to cointegration (Figure 7.4 and 8).

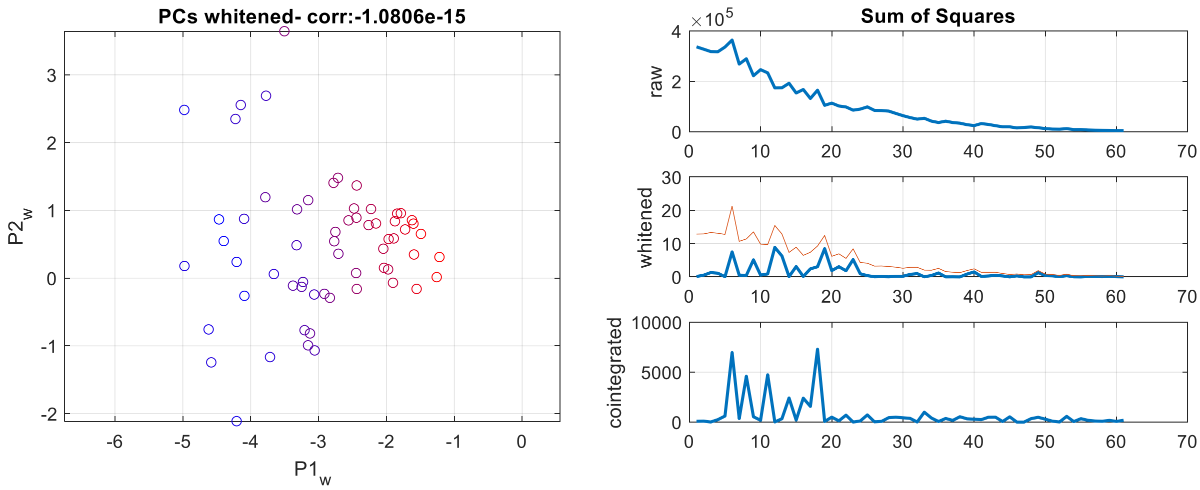


Figure 8. Standardized principal components (whitening) and sum of squares of the two PCs (in red) and of the second alone (blue, mid graph) compared to the squared Eulerian distance (raw) and to the sum of squares of the cointegrated residuals.

2.1 PCA orthogonal regression and whitening

Orthogonal regression is an extension of traditional regression for datasets in which the independent variable is not assumed to be perfectly known but admits errors. Indeed, in statistical literature this is known as "errors-in-variables" model, or also, "total least squares". A simple but effective way to perform it is based on Principal component analysis PCA [10,11].
Mathematically, given a d-dimensional centred dataset of n observations $X \in R^{d \times n}$, an unbiased estimator for the covariance can be used to obtain:

$$S = \frac{1}{n-1} XX'$$

PCA corresponds to the solution of the eigenproblem:

$$S\,V = V\Lambda$$

where V is the orthogonal matrix whose columns are the d eigenvectors v_j while Λ is the diagonal matrix of the d eigenvalues λ_j (usually sorted in descending magnitude) of the matrix S.
The matrix V can be used then to decorrelate the dataset X, that is, to rotate the reference frame to the one identified by the eigenvectors (i.e. the principal components, PCs) of matrix S:

$$Z = V'X$$

If the eigenvectors in V are normalized to have unit length ($v_j' v_j = 1$), the transform is a pure rotation, and it can be proved that $\sigma_j^2 = var(z_j) = \lambda_j$. Namely, the diagonal matrix Λ is the covariance of Z.
Focusing on linear orthogonal regression, the direction given by the first eigenvalue corresponds to the regression line, so that the residuals X_L can be simply found as a projection on the subspace generated by the $L = d - 1$ components other than the first:

$$z_{\mathrm{j}} = v_{\mathrm{j}}'X = v_{\mathrm{j}1}x_1 + v_{\mathrm{j}2}x_2 + \cdots + v_{\mathrm{j}d}x_d = \sum_{k=1}^{d} v_{\mathrm{j}k}x_k$$

$$Z_L = V_L'X$$

$$X_L = V_L Z_L = V_L V_L' X$$

The orthogonal regression is visualized in Figure 9.

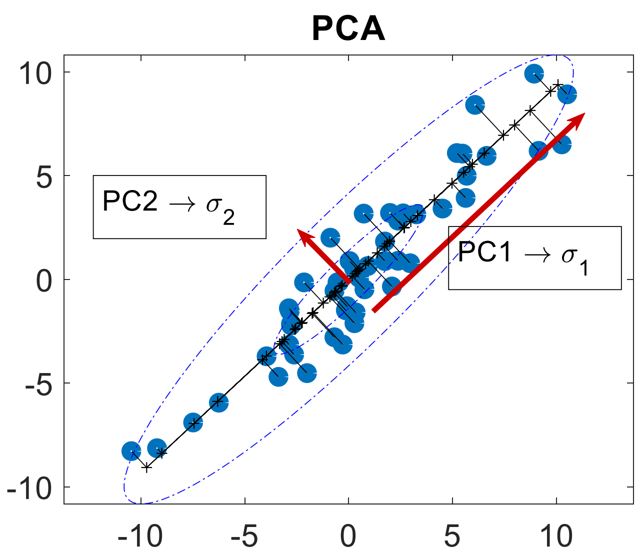


Figure 9. Visualization of the PCA orthogonal regression – the residuals corresponding to PC2 are highlighted.

Different normalizations for the eigenvectors are obviously possible. Another quite common one consists in normalizing for $v_j' v_j = \lambda_j$. In this case $var(z_j) = 1$ so that the covariance matrix of Z is the identity matrix I. In this case, on top of the rotation, a rescaling on the principal component occurs. V is then commonly called a “whitening matrix” W or also sphering matrix as it transforms the data covariance ellipsoid to a spheroid [6].

$$Z_W = W'X = \Lambda^{-1/2}V'X = \Lambda^{-1/2}Z$$

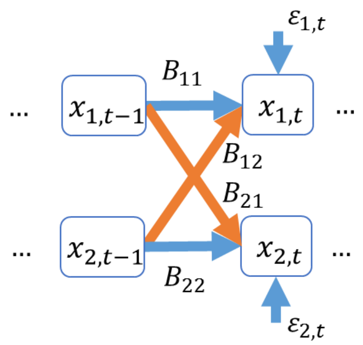
Finally, the squared Mahalanobis distance can be then written as

$$SMD = X'S^{-1}X = Z'V'S^{-1}VZ = Z'\Lambda^{-1}Z = \sum_j \frac{z_j^2}{\lambda_j} = Z_W'Z_W = \sum_j z_{Wj}^2$$

This proves that the squared Mahalanobis distance corresponds to the sum of squares of the whitened features. Hence, removing the first whitened component(s) from the sum corresponds to merging orthogonal regression and PCA-whitening: the so found distance is therefore robust to confounders. This makes it a good candidate to substitute the Mahalanobis distance as NI in the presence of non-stationary operational or environmental conditions.

2.2 Cointegration

Cointegration is a property of multiple nonstationary time series which can be linearly combined through a cointegrating matrix B to produce stationary series as the residual from a one-step-ahead prediction from a Vector Auto Regressive (VAR) model:



$$\hat{X}(t) = \mathrm{B}\, X(t-1)$$

$$Z(t) = X(t) - \hat{X}(t)$$

Figure 10. The VAR(1) model and residual.

The resulting $Z(t)$ can be proved to be stationary if the considered series in X share the same order of integration equal to 1. This can be verified by an Augmented Dickey Fuller univariate test [4,7] run on all the series.
In simple terms, ADF is based on the estimation of an AR(1) model

$$y_t = \rho\, y_{t-1} + \varepsilon_t$$

Obviously,
Obviously, different cases can be found:

- If $\rho = 0$ the signal is a pure white noise,
- If $|\rho| < 1$ the signal is stationary, or I(0),
- If $\rho = 1$ the signal is a pure random walk, or I(1),

Therefore, ADF tests the null hypothesis $H_0\colon \hat{\rho} = 1$ against the alternative $H_a\colon \hat{\rho} \neq 1$.
Hence, a confidence interval can be built around the estimated

$$\hat{\rho} = \frac{\sum_t y_t y_{t-1}}{\sum_t y_{t-1}^2} = \rho + \frac{\sum_t \varepsilon_t y_{t-1}}{\sum_t y_{t-1}^2}$$

If $\hat{\rho} - 1$ falls within the interval centred around 0, then the signal is proved to be I(1) at the selected confidence.

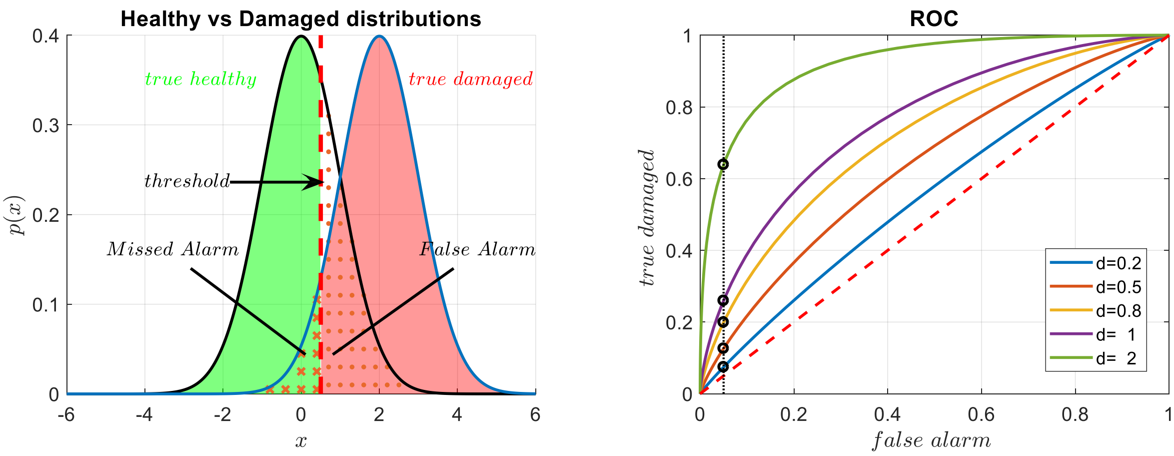
Once the I(1) series are highlighted, the VAR model can be fitted, and the residual vector $Z(t)$ computed.
By analogy to Mahalanobis distance and PCA orthogonal regression & whitening, the sum of squares of the Z residuals can be used as a NI, proved that these residuals are stationary and uncorrelated.
In practical cases, PCA orthogonal regression & whitening performs decorrelation, but not ensures necessarily the stationarity of the series. On the contrary cointegration enforces stationarity but does not guarantee uncorrelatedness. That is why it may be a good idea to merge the to procedures into a single approach.
By exploiting cointegration on pre-whitened series in fact, both uncorrelatedness and stationarity can be obtained.

2.3 Novelty detection and classification: performance assessment

Diagnostics, and in particular damage detection, can be considered a classification problem. The simplest binary classification can be tackled via novelty detection: when novelty is found (i.e. when the NI exceeds a threshold) the measure is assigned to the "damaged" class and an alarm is triggered, otherwise the acquisition is believed to be "healthy".
This implies the possibility of two kind of errors:

- type I error, which corresponds to triggering a False Alarms (FA or False Positive)
- type II error, which is a missed indication of damage although present (Missed Alarm, MA or 1-True Positive).

These error rates are usually collected in tables such as Table 8, which are very common when binary classification is considered. If the classification involves more than two groups, larger tables can be found with the name of confusion matrices.
On the contrary, in the field of Operational Research (OR), a discipline that deals with the application of analytical methods for making better decisions, the Receiver Operating Characteristic (ROC) is usually preferred for assessing the diagnostic ability of a binary classifier while its discrimination threshold is varied. Figure 7(b) summarizes the true damaged rate as a function of the false alarm rate for some relevant effect sizes $d = {}^{\mu_1-\mu_2}\!/_{\sigma}$ (the variance-normalized distance of the healthy and the damaged distributions) while the threshold takes all the possible values. The threshold corresponding to the very common 5% false alarm rate is highlighted. In general, anyway, the farthest is the ROC curve from the 1st – 3rd quadrant bisector, the better the classification, which obviously improves as the effect size is increasing.



(a) **(b)**

Figure 11. Receiver Operating Characteristic (ROC) as a function of the threshold (Gaussian distributions). **(a)** graphical summary of the table of type I and type II errors in yellow (Table 3). **(b)** ROC for binary classification with different effect sizes d and the position of the 5% false alarm rate. For $d = 0{,}2$ the performance is very poor as the ROC is near to the 1st-3rd quadrant bisector (nearly a random classifier).

Table 3. Type I and II errors in damage detection.

		True Health Condition:	
		Healthy	Damaged
Classification:	Healthy	No Alarm true healthy	Missed Alarm type II error
	Damaged	False Alarm type I error	Alarm true damaged

3 Results and Discussion

The methodology introduced in section 2 is tested on the experimental acquisition regarding high speed aeronautical bearings described in section 1.1. The dataset collecting 5 simple time features per each of the 6 channels (i.e., 30 features) is analysed separately for the 4 different load conditions. The confounding effect of the reducing speed (from 470 Hz to 0 Hz) will be compensated during the healthy training with 5 different approaches:

- Plain Euclidean distance (raw)
- Cointegration of the standardized features,
- Mahalanobis novelty,
- PCA orthogonal detection and whitening,
- Cointegration of the PCA-pre-whitened features.

The standardization which precedes the cointegration is necessary as the considered features have different order of magnitude, and this could lead to wrong estimates of the cointegrating matrix (poorly conditioned problem).
The novelty indices for the healthy reference and for the damaged conditions are reported in Figure 12 for condition 4 (load 1800N, decreasing speed). This graph highlights relevant considerations previously introduced in section 2.

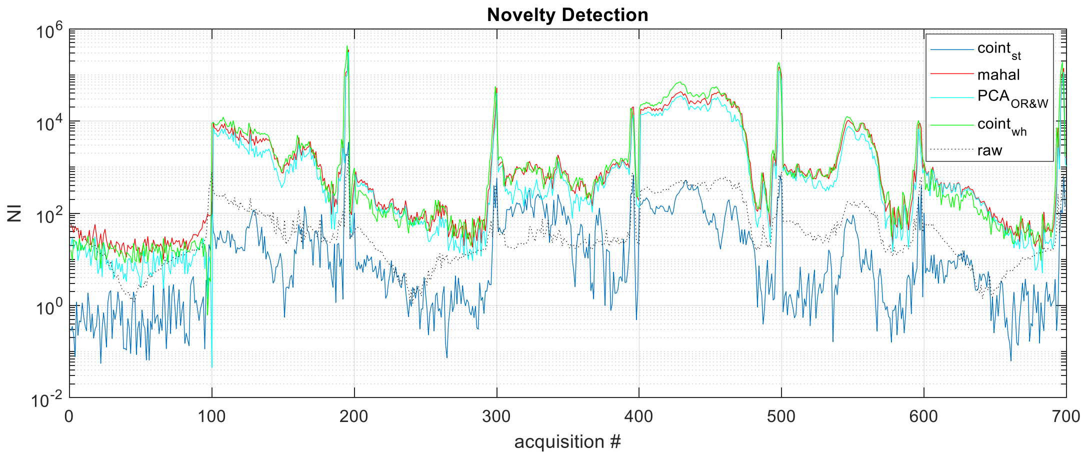


Figure 12. The Novelty Detection with the different NIs. 0-100 samples are the healthy reference, 100-200 corresponds to 1A damage, and so on until 600-700 coming from 6A damage.

In particular, cointegration based healthy NIs are stationary, while the Mahalanobis NIs are not. A trend is clearly visible, as this algorithm is targeted on decorrelation. An improvement in stationarity is obtained by neglecting the first 20 whitened principal components and focusing on the subspace individuated by the last 10. An additional note should be added to explain that the behaviour of the NIs at the end of all the run-down is ascribable to the fact that the record is not stopped exactly when the machine stops, so that the last points are practically acquiring just noise as the machine is already at a stop.

Despite the NIs curve already gives a qualitative impression of good detectability of almost all the different damages (from 1A to 7A), a quantitative comparison of the performances of the different methods is necessary. At this scope, Figure 13 reports the ROC curves for all the 4 different load conditions.

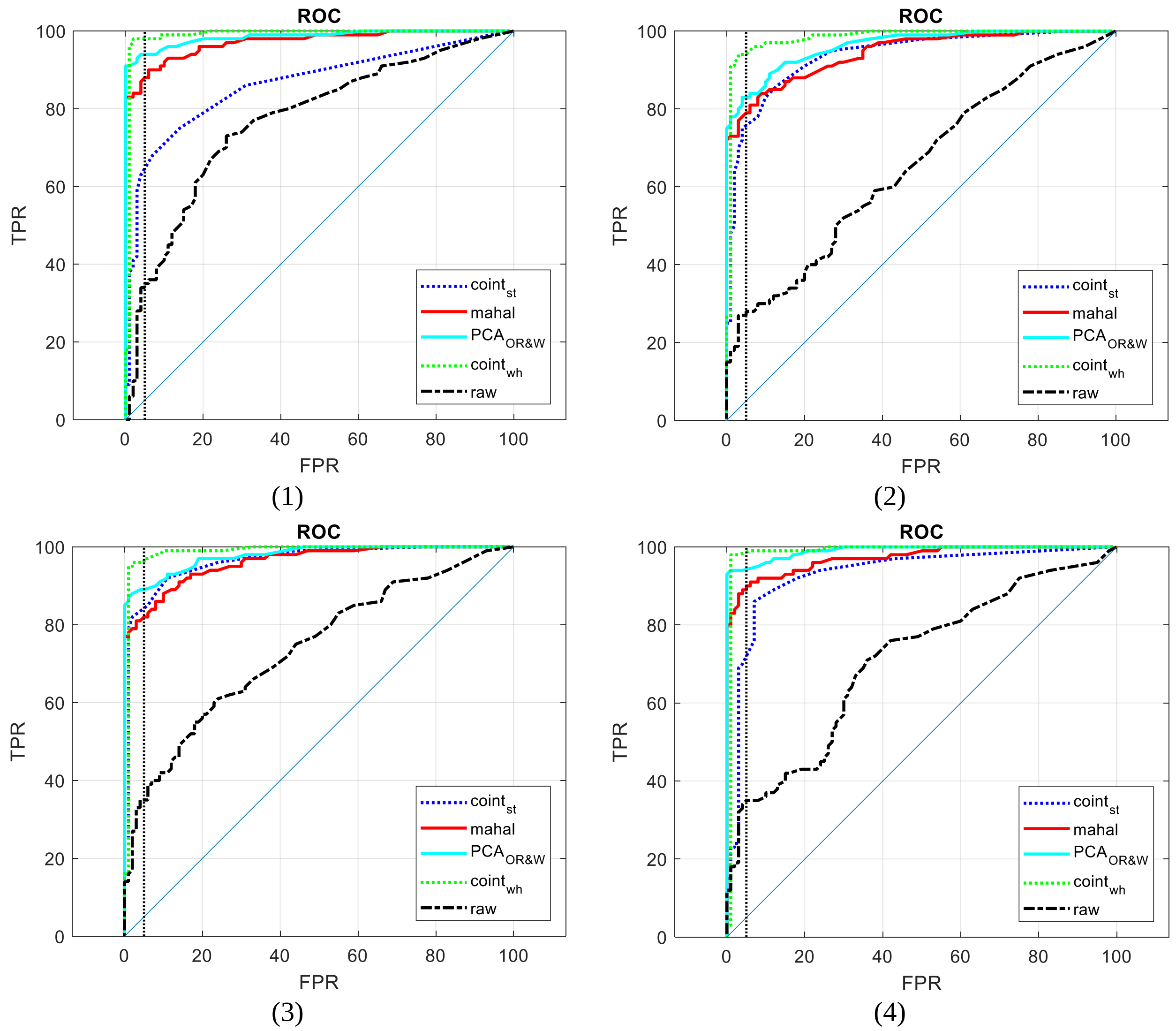


Figure 13. The ROC for Novelty Detection in the 4 loading conditions, respectively 1800, 1400, 1000, 0N

All the 4 graphs lead to a similar result: in this particular application, cointegration of the standardized features can enhance the damage detection in case of non-stationary rotational speed, but its performances are always comparable or inferior to Mahalanobis novelty detection. Removing the first 20 principal whitened components from the NIs computation (PCA-Ortogonal Regression & whitening), is able to further improve the results. The best performances anyway are given by cointegration of the PCA-pre-whitened features, as, accepting a 5% of false alarms, the missed alarms are always lower than 5% (100-95 % in the graphs).
The only issue with cointegration in this application is that decreasing the acceptable false alarm rate to 1% a knee of the curve is reached, so that the missed alarm rate increases dramatically. This phenomenon is not evidenced my Mahalanobis or PCA-Orthogonal Regression & whitening NIs.

4 Conclusions

This work focused on the compensation of confounders through two procedures which are established in univariate analyses, but not so well documented in multivariate cases:

- removal of linear deterministic trends in series by subtraction of a linear regression,
- removal of stochastic trends of difference stationary series by subtraction of the one-step ahead prediction from an autoregressive model.

In order to extend these approaches to the multivariate case, PCA orthogonal regression was used in the first case, while a vector autoregressive VAR(1) model was estimated in the second case.

In order to obtain consistent results from the VAR estimation, the analysed features were first standardized.

This is unnecessary for PCA, as a standardization can be more wisely performed on the principal space at a paltry expense. Indeed, just by normalizing the eigenvectors of the data covariance matrix so as to have the modulus equal to the corresponding eigenvalue, a sphering transform can be obtained. This is called PCA whitening and is “hidden” inside the Mahalanobis distance.

On the contrary, focusing only on the 1-D reduced dimensionality space individuated by the first principal component, orthogonal regression was performed. Therefore, a residual was found by removing the first principal component and focusing on the reduced dimensionality space.

Doing this on the whitened principal space, a Novelty index analogous to the Mahalanobis distance was found by summing the squares of the residuals (PCA-Orthogonal Regression & whitening).

Finally, cointegration was performed also on the PCA-pre-whitened dataset, to get the best of the two detrending strategies.

All the introduced methods enhanced the damage detection with respect to the raw Euclidean-distance-novelty-detection. The best damage detection in terms of reducing the missed alarms at a fixed maximum false alarm rate of 5% is without doubts the PCA-pre-whitened cointegration, which ensures missed alarm rates lower than 5% in all the loading conditions. Nevertheless, if the acceptable false alarm rate is decreased to 1%, the PCA-Orthogonal Regression & whitening proved to outperform all the other methods.

In general, the here proposed methodology, gives a quick, human independent and simple but effective way of performing damage detection also in case of non-stationary operational conditions. The nice compensation of the confounders in fact allows to enhance the damage, which can be easily highlighted by novelty detection.

References

[1] Daga A. P., Fasana A., Marchesiello S., Garibaldi L., *The Politecnico di Torino rolling bearing test rig: Description and analysis of open access data*, Mechanical Systems and Signal Processing, 2019. DOI: 10.1016/j.ymssp.2018.10.010.

[2] Farrar C.R., Doebling S.W., *Damage Detection and Evaluation II*. In: Silva J.M.M., Maia N.M.M. (eds) Modal Analysis and Testing. NATO Science Series (Series E: Applied Sciences). Springer, Dordrecht, 1999. ISBN 978-0-7923-5894-7.

[3] Worden K., Manson G., Fieller N. R. J., *Damage detection using outlier analysis*, Journal of Sound and Vibration (2000).

[4] Stadnitski T., *Deterministic or Stochastic Trend: Decision on the Basis of the Augmented Dickey-Fuller Test*. Methodology European Journal of Research Methods for the Behavioral and Social Sciences. 6. 83-92. 2010. DOI: 10.1027/1614-2241/a000009.

[5] Sawalhi N., Randall R.B., *Vibration response of spalled rolling element bearings: Observations, simulations and signal processing techniques to track the spall size*, Mechanical Systems and Signal Processing, 2011, DOI: 10.1016/j.ymssp.2010.09.009.

[6] Kessy A., Lewin A., Strimmer K., *Optimal Whitening and Decorrelation*, The American Statistician. 2015.DOI: 10.1080/00031305.2016.1277159.

[7] Cross E. J., Worden K., Chen Q., *Cointegration: a novel approach for the removal of environmental trends in structural health monitoring data,* Proc. R. Soc. A, 2011. DOI: 10.1098/rspa.2011.0023

[8] K. Worden, T. Baldacchino, J. Rowson and E. J. Cross, *Some Recent Developments in SHM Based on Nonstationary Time Series Analysis,* Proceedings of the IEEE, vol. 104, no. 8, pp. 1589-1603, Aug. 2016. DOI: 10.1109/JPROC.2016.2573596

[9] A. Deraemaeker, K. Worden, *A comparison of linear approaches to filter out environmental effects in structural health monitoring*, Mechanical Systems and Signal Processing, 2018, DOI: 10.1016/j.ymssp.2017.11.045.

[10] Jolliffe I.T., *Principal Component Analysis*, Springer, 2002. DOI: 10.2307/1270093

[11] Van Huffel S., Vandewalle J., *The Total Least Squares Problem: Computational Aspects and Analysis*, SIAM, 1991, ISBN: 978-0-898712-75-9