Turbomolecular high-vacuum pump bearings diagnostics using temperature and vibration measurements

Turbomolecular high-vacuum pump bearings diagnostics using temperature and vibration measurements

*Note: Sub-titles are not captured in Xplore and should not be used

Alessandro Paolo Daga
Politecnico di Torino DIMEAS
Torino, Italy
alessandro.daga@polito.it

Luigi Garibaldi
Politecnico di Torino DIMEAS
Torino, Italy

Luca Bonmassar
Agilent Technologies Italia Spa
Vacuum Products Division
Leinì, Italy

***Abstract*—Turbomolecular pumps are devices designed to obtain and maintain high vacuum. These are based on the principle that momentum can be given to gas molecules by collision with a moving solid surface. Hence, a multiple-stages fan rotor must spin at very high speeds to obtain the desired effect, leading to critical issues related to friction. When rolling elements bearings are employed to support the rotor shaft, these become the weakest link in the chain, so that diagnostic systems could be desirable to monitor the health state of such components. In this work temperature and vibration sensors are tested for this purpose using an endurance test rig running a turbomolecular pump to failure.**

Keywords— Turbomolecular pumps bearings, temperature monitoring, vibration monitoring

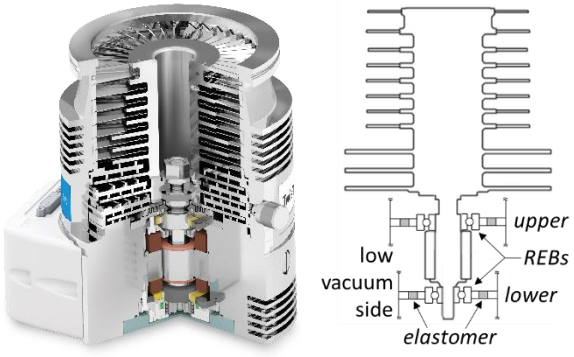
I. INTRODUCTION

Turbomolecular pumps (TMPs) are key apparatus for obtaining and maintaining high and ultrahigh vacuum (i.e., 10^{-3} to 10^{-11} mbar), which is fundamental in many industrial and scientific processes, from electronics and semiconductor manufacturing [1] to laboratory equipment as particle colliders and electron microscopes [2]. Classical TMPs are axial multistage compressors and achieve high vacuum transferring momentum (i.e., mass by velocity $m\bar{v}$) to the molecules through collision with the solid surface of a turbine rotor. Stator blades canalize then the molecules, leading them to the next stage, so that the operation is repeated as many times as the number of stages [3]. In order to ensure effective and efficient operation, the rotor is required to rotate up to 90.000 rpm, so that its supports must be accurately designed. In general, three architectures can be found: a) fully active magnetic suspension, b) one passive magnetic bearing on the high vacuum side of the rotor and a mechanical rolling bearing on the low vacuum side, c) two mechanical Rolling Element Bearings (REBs) on the low vacuum side [4]. If active magnetic suspensions are able to introduce damping in the system by proper control, passive and hybrid solutions are comparable to the coupling obtained with the more traditional REBs, so that a damping element becomes fundamental to lower the otherwise intolerable vibration levels induced by even small, unavoidable static unbalance, combined to such high rotational speeds. The state of the art involves then the addition of an annular elastomeric element (i.e., the square ring) press-fitted between the bearing and its non-rotating support [5].

If, on one hand, vibration is unavoidable, and generally associated with negative, undesired effects (i.e., noise, wear, fatigue, etc.), on the other hand, it can be used to take diagnostic information out of a rotating machine. Thanks to cost effective sensors (i.e., accelerometers) and minimum impact on machines operation, Vibration Monitoring (VM) is becoming the state of the art for implementing damage detection in most of maintenance regimes of machines involving rotating shafts [6-10].

With these premises, the present work aims at assessing the ability of VM to infer the state of health of REBs in Agilent TwisTorr 304 FS turbomolecular pumps, featuring architecture c) and Agilent Floating Suspension, as displayed in fig.1. Exploiting vibration acquisitions from the test rig described in section II, the diagnostic ability of VM will be evaluated in terms of pattern recognition [7] (semi-supervised or supervised techniques, as described in section III) and compared to a reference methodology based on temperature sensors integrated in the pump REBs supports.

As demonstrated by the results in section IV, using the diagnostic information from acceleration acquisitions, it is possible to anticipate REB failure of a couple of weeks, outmatching the diagnostic information obtained by the temperature acquisitions.



(A) (B)

Fig. 1. A) Cutaway image of an Agilent turbomolecular pump, B) Scheme of its cantilever rotor supported by two REBs on the low vacuum side and the elastomeric dampers characterizing the Agilent Floating Suspention [4,5].

II. DESCRIPTION OF THE TEST RIG AND OF THE ACQUISITIONS

In order to assess the diagnostic ability of temperature and vibration for TMPs, an accelerated-life, endurance experimental test rig was set up. In particular, several

sensorized TMPs were connected in a vacuum circuit (fig. 2) and run to failure, while temperature and vibration level were continuously acquired.

The TMPs featured one radial accelerometer mounted on each bearing support (i.e., the upper and lower bearing, accordingly to fig.1) and an additional radial accelerometer mounted on the body. The temperatures of the supports were also acquired (i.e., upper and lower thermistor), as well as the body temperature and the temperature of the high vacuum side of the rotor (i.e., rotor temperature, acquired with an infrared temperature sensor).

Three TMPs were selected for this analysis: T4, whose upper bearing failed on 2017/05/15, was chosen to be compared to T1 and T6, which were used as a healthy reference in the considered period of time from 2017/01/01 (which is taken as the initial time of the test, even if some machine was ran at reduced load starting from 2016/09/16) to 2017/05/30.

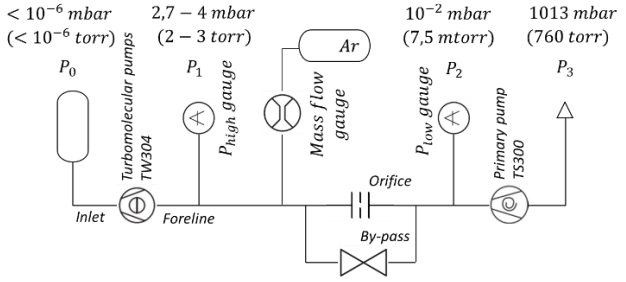


Fig. 2. Test rig vacuum circuit.

Temperatures were measured once every couple of minutes, while 10s of accelerations at 25,6 ksps were acquired every hour. Anyway, a more balanced dataset was obtained by extracting six features to summarize each 10s vibration acquisition, while the temperature measurements were down-sampled so as to match the one-per-hour sampling frequency. The selected features are well known level indicators such as ***RMS***, ***Peak value*** and ***crest factor***, as well as time-series statistical moments such as: ***Variance***, ***Skewness***, and ***Excess kurtosis*** (refer to appendix A for the mathematical formulation). The average temperature (i.e., the mean of the four acquired temperatures: body, rotor, lower bearing and upper bearing) for the three selected TDMs are shown in fig.4, while the RMS vibration trends (given as z-scores: centered over their mean value and normalized by their standard deviation) are reported in fig.5.

In order to accelerate the bearings' life, the test rig was initially set so as to keep the bearings working at their upper limit temperature (i.e., T_{set}), while the rotational speed was kept stationary at its nominal value of 1000Hz. This implied an almost stationary load due to a P_{high} always around $2 - 3\ torr$. Anyway, as several TMPs were run in parallel, the real temperatures of the different pumps could differ, as shown in fig.4, where the temperature trends normalized over T_{set} are given. Notice that the downward temperature spikes correspond to run-up transients, as the TMPs were periodically stopped. These transients were removed from the data (i.e., the rotational speed is always the nominal one), but still remain visible in the temperatures due to the thermal capacity of the pumps.

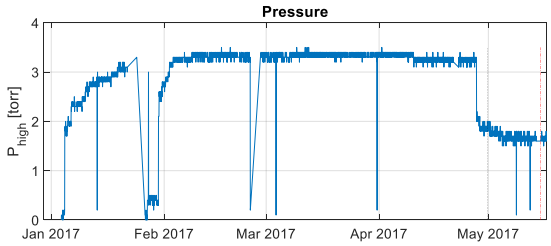


Fig. 3. P_{high} pressure. The vertical red dash-dotted line corresponds to T4 upper bearing failure.

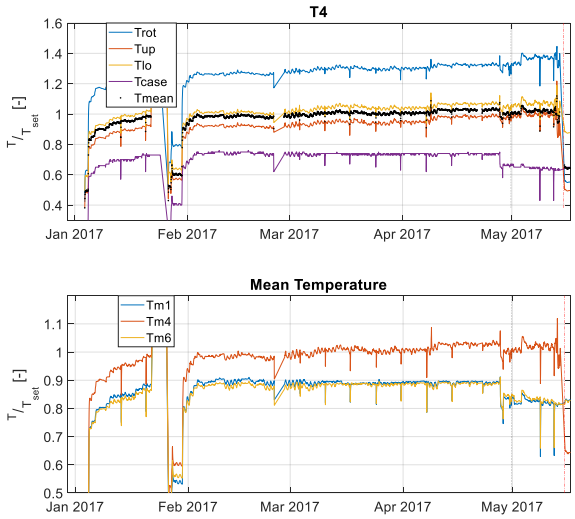


Fig. 4. The different temperatures for T4 and the average temperature trends for all the three TMPs. The vertical red dash-dotted line corresponds to T4 failure.

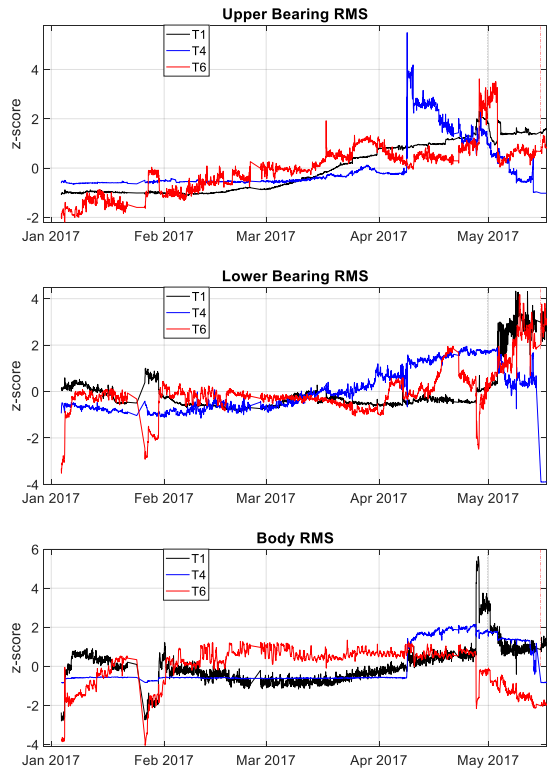


Fig. 5. RMS z-scores trends for the 3 TMPs under analysis. The vertical red dash-dotted line corresponds to T4 failure.

Even from such quite raw data, a couple of interesting considerations can be deduced:

- At the end of April (27/4), the load is reduced (in fig.3 P_{high} visibly drops); this does not influence much the temperatures, and just marginally affects the acceleration trends, so that the hypothesis of stationarity does not fail.
- At the beginning of April (8/4) the mean temperature $Tm4$ shows a local peak at 1,09 °C/°C (fig.4). If the mean temperature was stable around 1,01 °C/°C before this event, it settles to 1,03 °C/°C later. Even if this could seem a minor event, the vibration signals (fig.5) clearly highlight that a huge variation in the machine condition happened, having an impact also on the body accelerometer.
- On 13/5 a second and more important temperature peak at 1,12 °C/°C anticipates the bearing failure of a couple of days, but no other particular indication can be obtained from temperature.

As the scope of the health monitoring is that of picturing the damage evolution of the TMPs, vibration signals seemed to be more suitable for obtaining a more resolved monotonic trend from the healthy state to failure. With this target in mind, some data processing strategies are proposed in the next section.

III. Data Processing and Results

For the purpose of diagnostics, the temperature data was processed so as to produce the average temperature trend visible in fig.4. computed as the mean value of the body temperature, the two bearings temperatures and the rotor temperature. In particular, the rotor temperature, which was acquired with an infrared temperature sensor, was cleaned of the many unphysical outliers (probably due to disconnections of the sensor) by means of an Hampel median filter. Furthermore, the data was later down-sampled to the frequency of 1 datum per hour.

Considering the accelerometric signals on the contrary, given the very high rotational speeds of the shaft, a black box approach based on Pattern Recognition was preferred to the more traditional envelope demodulation. In other words, assuming that the operating and environmental conditions are sufficiently stationary not to affect the acquisitions (i.e., not to behave as confounders), any difference detected with respect to a normal reference condition can be reliably attributed to the presence of damage (fig. 6). If just data from the healthy state are used to train the difference-detecting algorithm (i.e., a semi-supervised technique), we will speak about Novelty Detection, while we will speak about classification if both data from the healthy and the damaged state are fed to the training together with the membership information (i.e., the label "healthy" or "damaged").

Hence, each 10s accelerometric acquisition will be substituted by the six selected features described in section II and appendix A. In this M-dimensional space (i.e., $M = 18$ if three accelerometers and 6 features are considered) some transform can be later used to compress the multivariate information to a univariate variable which will take the name of Novelty Index (NI). Ideally, the final NI, sampled at the frequency of 1 datum per hour, should be a monotonic increasing function depicting the wear of the bearing, degrading over time until failure. Hence, the proposed transforms which could enable to obtain such a result are briefly described hereinafter.

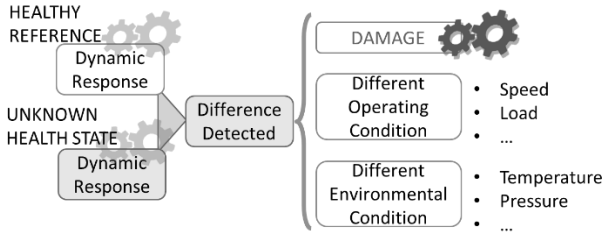


Fig. 6. Novelty Detection scheme.

A. Non-linear Combination of features: Mahalanobis Distance NI

Derived from probability theory, a data point can be considered abnormal (i.e., novel) if it is far enough from the centre of the reference distribution. For example, in the 1D case, points further than 3 times the standard deviation from the mean can be considered outliers according to the 3σ rule. If this consideration is extended to a multivariate case, the problem corresponds to finding the 3σ -equivalent, limit hyper-ellipsoid, beyond which a data-point can be considered novel. This problem can be turned into an equivalent 1-D problem if the multivariate space is whitened (i.e., the hyper-ellipsoid is transformed into a hypersphere via rotation and rescaling), as in this case, the only relevant information becomes the distance from the centre (i.e., the radius in polar coordinates). Hence, the limit hyper-ellipsoid simplifies to a limit radius or a simple threshold. According to such considerations, the Mahalanobis Distance (MD) is the perfect candidate, as it exactly corresponds to the points-to-centre distance in such a whitened space [7]. In fact, if X is a matrix collecting in its columns the extracted features over time (centred by removing their mean values μ), and S is the covariance matrix of X, it is easy to write:

$$NI \equiv MD = \sqrt{X' S^{-1} X} \qquad (1)$$

Notice that MD is dimensionless as the data is whitened (i.e., it is standardized in the multivariate space).

Finally, what is left is the selection of the reference data on which to estimate the covariance matrix S and the mean values μ to centre the data, and the selection of a proper threshold [7,11,12].

B. Linear Combination of features: Principal Component Analysis and Linear Discriminant Analysis NIs

Even if the MD-NIs are not too complex to compute, the MD remains a non-linear transform, so that it could be interesting to find whether the use of a linear combination of features is able to lead to satisfactory results [13,14]. One possibility is to conduct a Principal Component Analysis (PCA) [15], which is strictly related to MD as it is meant to find the rotation matrix used in the first step of the whitening (i.e., the eigenvector matrix V), and the rescaling coefficients involved in the whitening (i.e., the eigenvalues λ), solving the eigenproblem:

$$SV = V\lambda \qquad (2)$$

The first eigenvector is then summarizing most of the variance of the dataset, which, excluding the presence of

confounders, is related to the health condition. Hence, the corresponding eigenvector is a good candidate for transforming the multivariate data into a 1-D PC1-NI. Notice that this is possible because we have available acquisitions from TMP – T4 from the healthy state until failure, so that PC1 eigenvector points towards the damage-evolution direction. Furthermore, in order to have a dimensionless measure, PCA should be applied to standardized data (N.B., standardized univariately, i.e., each column-data is divided by its own standard deviation after having removed its mean value). This characterization should be done on TMP – T4 and extended to the other healthy pumps T1 and T6 for validation.

Notice that this approach is still semi-supervised as the previously proposed MD-NI analysis, but in this case the whole T4 data (i.e., healthy to damaged) is used to find the linear combination weights (i.e., for training). Anyway, it could be interesting to test also a linear supervised approach. In particular, Fisher's Linear Discriminant Analysis [16,10,8] can be easily implemented for finding the direction in the multivariate space along which the separation among the damaged and the healthy T4 data is maximum. If the data before failure are labelled as damaged, then, it is easy to find the optimal LDA weights [16,10]:

$$S_w = \sum_{h=1}^{n_1} (x_h - \mu_1)(x_h - \mu_1)' + \sum_{k=1}^{n_2} (x_k - \mu_2)(x_k - \mu_2)'$$

$$d(w) = \frac{w'(\mu_2 - \mu_1)(\mu_2 - \mu_1)'w}{w' S_w w}$$

$$\arg\max_w \; d(w) : w \propto S_w^{-1}(\mu_2 - \mu_1) \qquad (3)$$

Hence, as stated by (3), the maximization of the measure of separation *d(w)* results in a direction *w* that can be computed as the inverse within-class covariance matrix S_w by the distance of the two classes (i.e., damaged vs healthy) centroids $\mu_2 - \mu_1$. The corresponding weights can then be used for transforming the multivariate data into a 1-D LDA-NI.

IV. Vibration Monitoring Results

According to what described in section III, the dataset was first analysed by computing MD-NIs referred to a training from 2017/01/01 to 2017/02/28. In particular, the analysis was first conducted using all the three available accelerometers (i.e., Upper & Lower bearing + Body), and later reduced neglecting the body information. As it can be clearly noticed in fig.7, the two results are very similar, apart from some peak (e.g., T1 peak around 2017/05/01). This suggests that the body information is not particularly relevant for diagnostic purposes, and is just introducing noise from the outside, so that it will be neglected in the further analyses.

The second step was to run a PCA on the standardized data of each pump. At first, PCA was run independently on the 3 TMPs, to explore the dataset. The result, limited to the first two principal components (explaining roughly the 50% of the data variability) is shown in fig.8. As it is easy to notice, a clear evolution is present along the first principal component of T4. Since the only source of variability left uncontrolled is related to the wear of the TMP, the variance pictured along PC1 must correspond to the health-state evolution of T1. According to this consideration, the weights describing the PC1 related eigenvector could be used to produce a PC1-NI, if the data is standardized according to the mean value and the standard deviation obtained from T4 pump.

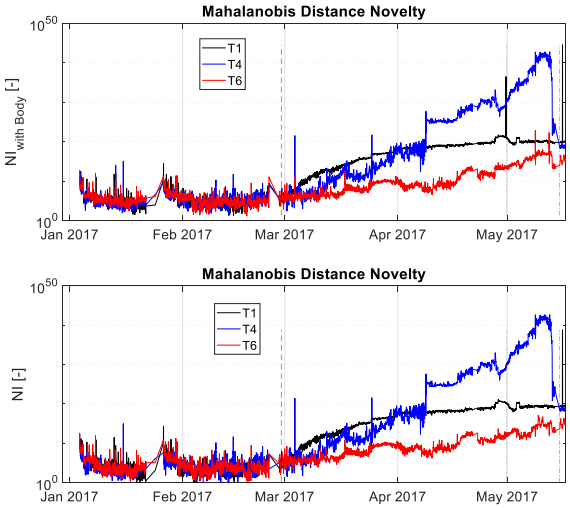


Fig. 7. MD-NIs considering all the three accelerometers compared to the result neglectinf the body accelerometer. The training involves data until February 2017 (dashed line).

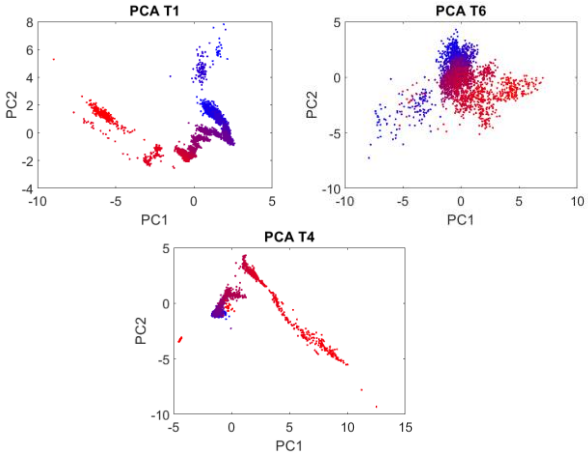


Fig. 8. PCA: 2D plot (PC2 vs PC1) of the whole standardized dataset of each TMP. Notice that the first two components explain 55% of the variance for T1, 60% of the variance for T4 and 40% of the variance for T6. The color, running from blue to red, denotes the evolution over time.

If PC1-NIs are computed for T1 and T6 using the training information obtained from T4, then, the plot in fig.9 can be produced. Comparing fig.9 to fig.7, two important considerations can be made:

- T1 and T6 (healthy data) are flatter for PC1-NIs rather than considering MD-NIs. This enables a better early detection of the critical situation for T4, which stands out more clearly. In particular, from PC1-NIs, the critical health condition can be noticed at the very beginning of April, while MD-NIs enable a clear detection only after April 8 (N.B., corresponding to the first peak in the mean temperature signal Tm4 in fig.4).
- It must be remembered that PC1_NIs were computed exploiting the entire T4 data (i.e., until failure) in the training, while each turbine was trained only on its own healthy data (i.e., until the end of February) for the MD-NIs computation. The improvement obtained with PC1-NIs is then justified.

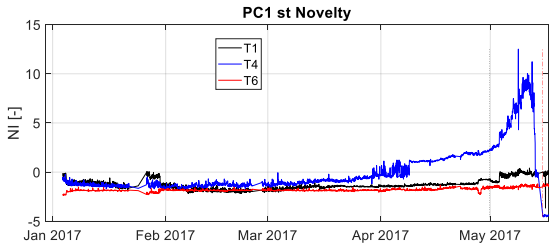


Fig. 9. PC1_NI trained using T4, and validated using T1 and T6.

To conclude LDA-NIs were tested. Exploiting the knowledge produced from the previous analyses, the healthy and the damaged data were labelled taking the beginning of April as a discriminant. LDA was then trained on T4 standardized data and the result (i.e., the computation of the linear combination weights w) was validated on T1 and T6 standardized data, using the training information obtained from T4.

The result is reported in fig.10. Comparing it to fig.9 and fig.7 it can be seen that the so obtained result is slightly (even if just marginally) improved, as a consequence of the addition of further external information to the training (i.e., the T4 labels of healthy or damaged).

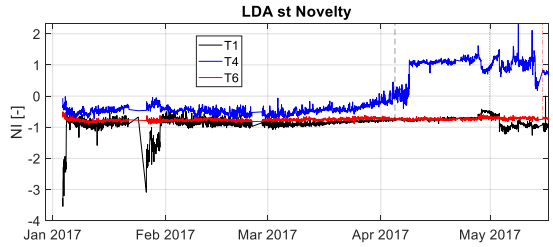


Fig. 10. LDA_NI trained using T4 data and labels created using the healthy to damaged threshold represented as a dashed line (2017/04/4), validated using T1 and T6.

V. Conclusions

To summarize, Agilent TMPs were monitored in terms of temperature and vibration. The mean temperature of the turbine was able to highlight two critical moments of T4 life (i.e., two anomalous peaks), nevertheless it was not good for tracking the damage evolution over time, nor it could be used to find unique limit values for all the pumps, which were proved to work at very different mean temperatures.

On the contrary, vibration monitoring data seemed to better picture the evolution of the TMPs' health state, leading to a more resolved monotonic trend from the healthy reference state to the final failure of T4. In particular, three Pattern Recognition strategies were adopted to reduce the multivariate vibration-features dataset to a 1-D novelty information:

- Semi-supervised MD-NIs with a training on the healthy data alone (i.e., from 2017/01/01 to 2017/02/28).
- Semi-supervised PC1-NIs with a training on T4 data from healthy-state to failure.
- Supervised LDA-NIs with training on T4 data from healthy-state to failure and membership labels.

As the amount of information used in the three approaches increases from the first to the last, this is reflected by the moment when the T4 anomaly can be detected, which occurs in advance. Nevertheless, the PC1-NIs describe a better health-state evolution, so that this methodology results the most promising.

VI. Appendix A

The extracted vibration features were:

- The peak value of the signal:

$$p\big(x(t)\big) = \frac{1}{2}[max(x(t)) - min(x(t))] \quad (4)$$

- The Root Mean Square (RMS) of the signal:

$$RMS(x(t)) = \sqrt{E[x^2(t)]} \quad (5)$$

- The crest factor, defined as:

$$C.F.(x(t)) = \frac{p(x(t))}{RMS(x(t))} \quad (6)$$

- The standard deviation of the signal:

$$std\,\big(x(t)\big) = E[(x(t) - E[x(t)])^2]^{0,5} \quad (7)$$

- The skewness of the signal:

$$skew(x(t)) = E\left[\frac{(x(t)-E[x(t)])^3}{std\,(x(t))^3}\right] \quad (8)$$

- The excess kurtosis of the signal:

$$kurt(x(t)) = E\left[\frac{(x(t)-E[x(t)])^4}{std\,(x(t))^4}\right] - 3 \quad (9)$$

Acknowledgment

We would like to thank Agilent Technologies for ideating and running the tests, sharing the acquisitions and for their fundamental support to this work.

References

[1] Sun K., Zhang W., Han F., Zhao F., Zhang Z. J., Han J., "A new modeling method to reveal pumping mechanism of turbomolecular pump", Journal of Applied Fluid Mechanics, Vol. 14, No. 1, pp. 165-173, 2021. DOI: 10.47176/jafm.14.01.31340

[2] Degasperi F. T., Ricotta R. M., "Pressure field and its gradient in electron microscopes", Vacuum, Volume 188, 2021. DOI: 10.1016/j.vacuum.2021.110162.

[3] Munawar I., Zia A. W., Batani D., Rashid H., Lodhi M. A. K., "Design modification in rotor blade of turbo molecular pump". Nuclear Instruments and Methods in Physics Research Section A Accelerators Spectrometers Detectors and Associated Equipment. 678. 88-90. DOI: 10.1016/j.nima.2012.02.030.

[4] Quartarone C., Beqari J., Detoni J.G., Cometti F., Emelli E. "Balancing of Turbomolecular Pumps: Modal Balancing Approach and Experimental Results". In: Cavalca K., Weber H. (eds) Proceedings of the 10th International Conference on Rotor Dynamics – IFToMM. IFToMM 2018. Mechanisms and Machine Science, vol 63. Springer, Cham. 2019. DOI: 10.1007/978-3-319-99272-3_9

[5] Quartarone C. et al. "Modeling and Experimental Characterization of Resilient Supports for Turbomolecular Pumps". In: Pennacchi P. (eds) Proceedings of the 9th IFToMM International Conference on Rotor Dynamics. Mechanisms and Machine Science, vol 21. Springer, Cham. 2015. DOI: 10.1007/978-3-319-06590-8_178

[6] Antoni J., et al., "Feedback on the Surveillance 8 challenge: Vibration-based diagnosis of a Safran aircraft engine", Mechanical Systems and Signal Processing, pagine 112-144, Vol.97, ISSN:0888-3270, 2017. DOI:10.1016/j.ymssp.2017.01.037

[7] 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, Volume 120, 2018. DOI: 10.1016/j.ymssp.2018.10.010.

[8] Daga A. P., Garibaldi L., "Machine vibration monitoring for diagnostics through hypothesis testing", Information, Vol.10, 2019. DOI:10.3390/info10060204

[9] Castellani F., Garibaldi L., Daga A.P., Astolfi D., Natili F., "Diagnosis of Faulty Wind Turbine Bearings Using Tower Vibration Measurements", Energies, Vol.13, 2020. DOI:10.3390/en13061474

[10] Daga A.P., Fasana A., Garibaldi L., Marchesiello S., "Big Data management: A Vibration Monitoring point of view", IEEE, Metrology for Industry 4.0 and IoT 2020, pagine 548-553, 2020. DOI:10.1109/MetroInd4.0IoT48571.2020.9138196

[11] K. Worden, D. W. Allen, H. Sohn, C. R. Farrar, "Damage detection in mechanical structures using extreme value statistics", Proceedings of SPIE - The International Society for Optical Engineering, 4693, 289–299 (2002). DOI: 10.1117/12.475226

[12] K. Worden, G. Manson, N. R. J. Fieller, "Damage detection using outlier analysis", Journal of Sound and Vibration (2000). DOI: 10.1006/jsvi.1999.2514

[13] A. Deraemaeker, K. Worden, "A comparison of linear approaches to filter out environmental effects in structural health monitoring", Mechanical Systems and Signal Processing, Volume 105, 2018. DOI: 10.1016/j.ymssp.2017.11.045.

[14] A.M. Yan, G. Kerschen, P. De Boe, J.C. Golinval, "Structural damage diagnosis under varying environmental conditions – Part I: a linear analysis", Mech. Syst. Signal Process. 19 (2005) 847–864, DOI: 10.1016/j.ymssp.2004.12.002.

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

[16] C. Bishop, "Pattern Recognition and Machine Learning", Springer-Verlag New York, 2006. ISBN: 978-0-387-31073-2