Fault Detection in Non Gaussian Problems Using Statistical Analysis and Variable Selection

João P. P. Gomes[1], Bruno P. Leão[1], Roberto K. H. Galvão[2] and Takashi Yoneyama[2]

[1] *EMBRAER, São Jose dos Campos, São Paulo, 12227-901, Brazil*
joao.pordeus@embraer.com.br
bruno.leao@embraer.com.br

[2] *ITA – Instituto Tecnológico de Aeronáutica, São José dos Campos, São Paulo, 12228-900, Brazil*
kawakami@ita.br
takashi@ita.br

ABSTRACT

This work concers the problem of fault detection using data-driven methods without the assumption of gaussianity. The main idea is extend the Runger's U^2 statistical distance measures to the case where the monitored variables are not gaussian. The proposed extension is based on Gaussian Mixture Models and Parzen windows classifiers to estimate the required conditional probability distributions. The proposed methodology was applied to an APU dynamic model and showed better results when compared to classical fault detection techniques using Multivariate Statistical Process control with Hotelling's T^2 metrics.*

1. INTRODUCTION

Data-driven methods comprise a powerful set of tools for performing failure prognosis and diagnosis. Such group of methods includes clustering and classification techniques, where the data is divided into groups on the basis of some specific distance measure (Duda et al., 2001). Statistical measures are a usual choice for such methods. The origins of clustering and classification methods based on statistical measures may be linked to the works of Mahalanobis (1936) and Hotelling (1933), which are related, respectively, to the Mahalanobis distance (MD) (De Maesschalck et al., 2000) and the T^2 statistic (Kourti and MacGregor, 1995). Such statistical distance measures are the basis of Multivariate Statistical Process Control (MSPC), which consists of a group of multivariate analysis techniques that can be used in health monitoring and diagnosis in industrial environment, such as chemical plants (Kourti and MacGregor, 1995) and mining enterprises (Yacher and Orchard, 2003). Statistical measures are also employed in other fields of knowledge such as image processing and pattern recognition (Webb, 2002). In these fields, similar concepts are used for the definition of the Gaussian. Both in the Gaussian classifier and in the MSPC techniques, an usual assumption is to considerthe underlying joint distributions of the monitored variables as Gaussian (or at least can be approximated as). Literature presents various examples of the use of such types of methods for Prognostics and Health Management (PHM): Kumar et al. (2008) present the use of MD for monitoring electronic systems; Mimnagh et al. (2000) present the use of Hotelling's T^2 statistic for the diagnostics of a helicopter drive system; Leão et al. (2009) show the application of MSPC for monitoring the health of electro-mechanical systems.

The abovementioned methods may provide poor performance when the gaussianity assumption is not verified. Since many practical problems do not satisfy such gaussianity assumption, extensions of these methods have been proposed to address non-Gaussian problems (Webb, 2002). One solution of this type is the use of Gaussian Mixture Models (GMM) to approximate the joint probability density of the variables of interest. Another possibility is the use a composition of Gaussian kernels for approximating the joint density in a non-parametric way. Such method is commonly referred as Parzen windows.

This work presents extensions to GMM or Parzen windows classifiers, which can provide better results for PHM solutions. Such extensions are inspired by the U^2 statistical distance (Runger, 1996), which was introduced by Runger in the context of MSPC (on the assumption of Gaussian joint distributions). Runger's

U^2 statistic is based on the division of the set of monitored variables (**x**) into two subsets: the first one (**y**) includes variables which are affected by the failure under consideration; the second one (**z**) encompasses variables which are not affected by the failure but are correlated to the variables of the first subset. Examples of the latter include operational and environmental conditions. The proposed methods are extensions of Runger's work to non-Gaussian problems, which are based on the calculation of the conditional likelihood $p(\mathbf{y}|\mathbf{z})$ using the densities estimated through GMM or Parzen windows.

In order to illustrate the use of the proposed methods, a simulation model of an aircraft Auxiliary Power Unit (APU) is employed. Different failure modes are simulated using such model and the proposed methods are used for failure diagnosis. Their performance is compared to that of the aforementioned traditional methods described in literature.

The remaining sections are organized as follows: section 2 describes the theoretical background associated to MSPC, Gaussian, GMM and Parzen windows classifiers; section 3 presents the novel methods proposed in this work; section 4 presents the simulated tests and results and section 5 is the conclusion.

2. THEORETICAL BACKGROUND

2.1 Statistical Distances, MSPC and the Gaussian Classifier

The application of statistical theory for fault detection relies on the assumption that the characteristics of the data variations are relatively unchanged unless a fault occurs in the system. This is a reasonable assumption under the definition of a fault as an abnormal process condition. It implies that the statistical properties of the data are repeatable for the same operating conditions, although the actual values may not be predictable (Chiang et al, 2001). The repeatability of the statistical properties allows the use of statistical measures, based on statistical distances, for the detection of abnormal behaviors on a process.

Eq. (1) presents the well known Mahalanobis Distance (MD):

$$M(\mathbf{x}) = (\mathbf{x} - \mu)\sum^{-1}(\mathbf{x} - \mu)^T \qquad (1)$$

where **x** is the feature vector associated to an observation and μ and Σ are respectively the mean values and the covariance matrix of a given dataset. These statistical properties can be estimated as

$$\mu = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i \qquad (2)$$

and

$$\Sigma = \frac{1}{n-1}\sum_{i=1}^{n}\left[(\mathbf{x_i} - \mu)(\mathbf{x_i} - \mu)^T\right] \qquad (3)$$

where $\{\mathbf{x}_i, i = 1, 2, ..., n\}$ is a given set of observations.

The MD is used to define one of the most popular MSPC methods called Hotelling's T^2 statistic. In Hotelling's T^2 statistic, a statistical model is built using Eq. (2) and Eq. (3) given a dataset **X** containing n instances of a feature vector **x**. Each instance is composed by k monitored variables. All feature vectors in **X** are obtained for a healthy system (without faults).

After this training stage, the MD is calculated for each new instance $\mathbf{x_{new}}$ and the result is compared with a threshold in order to detect anomalies. Hotelling's T^2 statistic is defined as:

$$T^2 = (\mathbf{x_{new}} - \mu)\sum^{-1}(\mathbf{x_{new}} - \mu)^T \qquad (4)$$

where μ and Σ are estimated using Eq. (2) and Eq. (3) for the dataset **X** and $\mathbf{x_{new}}$ is a new instance of the feature vector **x** that needs to be classified as healthy or faulty.

The same principles are involved in a popular classification method, mainly employed in the pattern recognition literature, the Quadratic Gaussian Classifier (QGC, or simply Gaussian Classifier). For fault detection, the QGC can be formulated to solve the problem of one class classification, that is, to classify the operation of a system as healthy or not.

Using Bayes' theorem one could obtain the probability of a system being healthy ($H=1$) given a feature vector **x** according to Eq. (5).

$$P(H=1|\mathbf{x}) = \frac{p(\mathbf{x}|H=1)P(H=1)}{p(\mathbf{x})} \qquad (5)$$

Since the unconditional probability density $p(\mathbf{x})$ is not related to the health state of the system, it is not useful to decide if the system operation is healthy or faulty ($H=0$). Therefore, it can be ignored in the statistical measure.

Considering the prior probabilities of the system being healthy or faulty ($P(H=1)$ or $P(H=0)$) are not affected by the **x**, these can be also be ignored, resulting on:

$$P(H=1|\mathbf{x}) \propto p(\mathbf{x}|H=1) \qquad (6)$$

Assuming $p(\mathbf{x}|H=1)$ to be a Gaussian distribution one can use the following statistical measure to detect anomalies based on $\mathbf{x_{new}}$.

$$p(H=1|\mathbf{x}_{new}) \propto \frac{1}{(2\pi)^k|\Sigma|^{1/2}}\exp\left[-\frac{1}{2}(\mathbf{x}_{new} - \mu)\Sigma^{-1}(\mathbf{x}_{new} - \mu)^T\right] \qquad (7)$$

It is possible to notice in Eq. 7 the presence of a term identical to that of expression Eq. 1, which is the

MD. It is the only term that depends on $\mathbf{x}_{new}$, so that the result is similar to that presented by Hotelling's T^2 statistic.

An MSPC method proposed by (Runger, 1996) provides relevant improvements in Hotelling's T^2 statistic and QGC. In many real applications it is possible that only a subset of the monitored variables are affected by a failure. The main idea in this method is to restrict the analysis to these fault-sensitive variables but without excluding the influences of other non fault-sensitive variables in system behavior.

Consider an observation $\mathbf{x}_i$ of the feature vector:

$$\mathbf{x}_i = [\mathbf{x}_{i1} \quad \cdots \quad \mathbf{x}_{ij} \quad \cdots \quad \mathbf{x}_{ik}] \tag{8}$$

Assuming that a fault only affects variables $\mathbf{x}_{i1}$ up to $\mathbf{x}_{ij}$, one can divide $\mathbf{x}_i$ into two sets:

$$\mathbf{y}_i^{\mathbf{T}} = [\mathbf{x}_{i1} \quad \ldots \quad \mathbf{x}_{ij}] \tag{9}$$

and

$$\mathbf{z}_i^{\mathbf{T}} = [\mathbf{x}_{i(j+1)} \quad \ldots \quad \mathbf{x}_{ik}] \tag{10}$$

where $\mathbf{y}_i$ contains the features that are sensitive to an incipient failure and $\mathbf{z}_i$ contains those that are not sensitive to the failures.

The idea of Runger's U^2 statistic is to calculate Hotelling's T^2 statistic and subtract the influence of $\mathbf{z}_i$ in the final calculated distance while keeping $\mathbf{z}_i$ influence in $\mathbf{y}_i$ behavior.

Runger's U^2 statistic can be defined as:

$$U^2 = T^2 - (\mathbf{z}_{\mathbf{new}} - \mu_z)^T \sum\nolimits_z^{-1} (\mathbf{z}_{\mathbf{new}} - \mu_z) \tag{11}$$

It can be noticed that the MD is employed to compare $\mathbf{z}_{\mathbf{new}}$ with a statistical model built using a subset of **X** comprising only the variables not affected by faults. The result is subtracted from Hotelling's T^2 statistic.

2.2 GMM and Parzen Windows Classifiers

All methods presented in section 2.1 have the assumption that the healthy data follows a Gaussian distribution. However, that assumption may be invalid in some real applications.

In order to overcome this problem, many authors have proposed methodologies mostly based on the usage of nonparametric estimators for the distribution of the healthy data (Webb, 2002), (Duda et al., 2001). With that estimation, it is possible to approximate $p(\mathbf{x}_{new}|H=1)$. In these cases, the statistical measure can be defined by Eq. (6) with no need for assumptions on the particular distribution for the data.

Two of the most commonly used nonparametric estimation methods are the Parzen windows and the GMM.

Parzen windows is a non parametric estimator based on the idea of approximating the distribution to be estimated by a superposition of kernel functions centered on each of the $\mathbf{x}_i$ samples in **X**. Based on that, and using the formulation presented in Eq. (6), it is possible to estimate $p(\mathbf{x}_{new}|H=1)$ according to Eq. (12) .

$$p(\mathbf{x}_{new} \mid H=1) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h}K\left(\frac{\mathbf{x}_{new} - x_i}{h}\right) \tag{12}$$

where h is the smoothing parameter and $K(.)$ is the kernel function, chosen *a priori*.

One important drawback when applying Parzen windows is the curse of dimensionality occurring when dealing with high dimension data. In these cases, a limited number of data vectors can result in a sparse dataset witch could difficult the task of distribution estimation. One way to overcome this problem is the application of so called semi-parametric estimators such as GMM.

The GMM approach models the distribution to be estimated as a composition of a set of weighted Gaussian distributions. The general expression for $p(\mathbf{x}_{new}|H=1)$ can be written as:

$$p(\mathbf{x}_{new} \mid H=1) = \sum_{l=1}^{m}\pi_l G(\mathbf{x}_{new}, \theta_l) \tag{13}$$

where π_l are the weights of each of the l Gaussian models whose parameters μ_l and Σ_l are expressed in θ_l.

The values for parameters π_l and θ_l can be obtained according to the Expectation-Maximization algorithm as presented originally in (Dempster et al, 1977).

Implementation details and information about other nonparametric estimators based classifiers can be found in many references such as (Webb, 2002) and (Duda et al., 2001).

3. PROPOSED METHOD

Although some authors proposed methodologies to monitor systems that provide non-Gaussian data, no previous work exploited the differentiation of variables which are affected by failure from those that are not. This feature was the main contribution in the development of Runger's U^2 when compared to Hotelling's T^2 in the Gaussian case.

Using the same definitions presented in Eq. (8), Eq. (9) and Eq. (10), one can rewrite Eq. (5) as:

$$P(H=1 \mid \mathbf{y},\mathbf{z}) = \frac{p(\mathbf{y},\mathbf{z} \mid H=1)P(H=1)}{p(\mathbf{y},\mathbf{z})} \tag{14}$$

The joint distribution of **y** and **z** can be rewritten in terms of the conditional probability. This substitution leads to.

$$P(H=1 \mid \mathbf{y},\mathbf{z}) = \frac{p(\mathbf{y} \mid \mathbf{z}, H=1)p(\mathbf{z} \mid H=1)P(H=1)}{p(\mathbf{y},\mathbf{z})} \tag{15}$$

The hypothesis assumed for Runger U^2 is that the fault only affects the subset of the feature vector defined by **y**. This hypothesis can be reformulated by saying that the distribution of **z** does not change whenever the system is healthy or faulty. In this case we have:

$$p(\mathbf{z} \mid H=1) = p(\mathbf{z}) \tag{16}$$

With Eq. (15) and Eq. (16) we have:

$$P(H=1 \mid \mathbf{y},\mathbf{z}) = \frac{p(\mathbf{y} \mid \mathbf{z}, H=1)\, p(\mathbf{z}) P(H=1)}{p(\mathbf{y},\mathbf{z})} \tag{17}$$

The joint probability $p(\mathbf{y},\mathbf{z})$ can be expressed by:

$$p(\mathbf{y},\mathbf{z}) = p(\mathbf{y} \mid \mathbf{z})\, p(\mathbf{z}) \tag{18}$$

Resulting in:

$$P(H=1 \mid \mathbf{y},\mathbf{z}) = \frac{p(\mathbf{y} \mid \mathbf{z}, H=1)\, p(\mathbf{z}) P(H=1)}{p(\mathbf{y} \mid \mathbf{z})\, p(\mathbf{z})} \tag{19}$$

That leads to:

$$P(H=1 \mid \mathbf{y},\mathbf{z}) = \frac{p(\mathbf{y} \mid \mathbf{z}, H=1) P(H=1)}{p(\mathbf{y} \mid \mathbf{z})} \tag{20}$$

Using the same simplification procedures described for the Gaussian case, it yields:

$$P(H=1 \mid \mathbf{y},\mathbf{z}) \propto p(\mathbf{y} \mid \mathbf{z}, H=1) \tag{21}$$

The conditional probability of **y** given **z** in a healthy system can be obtained by:

$$P(H=1 \mid \mathbf{y},\mathbf{z}) \propto \frac{p(\mathbf{y},\mathbf{z} \mid H=1)}{p(\mathbf{z})} \tag{22}$$

For the classification of $\mathbf{x}_{new}$ as healthy or faulty we have:

$$P(H=1 \mid \mathbf{x}_{new}) \propto \frac{p(\mathbf{y}_{new},\mathbf{z}_{new} \mid H=1)}{p(\mathbf{z}_{new})} \tag{23}$$

where $p(\mathbf{y}_{new},\mathbf{z}_{new}|H=1l)$ and $p(\mathbf{z}_{\mathbf{new}})$ can be estimated using any nonparametric estimation method as the ones presented in section 2.2.

Analyzing the result obtained in Eq. (23) it may be noticed that the basis of the proposed method is to estimate the conditional probability of **y** given **z** instead of the joint probability of **y** and **z** as presented in Eq. (6). The new method is expected to provide greater sensitivity and therefore better performance for fault diagnosis and health monitoring applications.

4. SAMPLE APPLICATION

To demonstrate how the proposed method could be applied in a real system and to compare the results against some classical methods, a sample application will be presented. The application consists of the detection of faults in an Auxiliary Power Unit (APU).

The Auxiliary Power Unit is a gas turbine device on a vehicle with the purpose of providing power to other systems when main engines are turned off. This power can be either pneumatic, obtained through the bleeding of compressed air, or electrical, obtained by coupling a generator to the APU shaft. They are commonly found on medium and large aircraft, as well as some large land vehicles. Its primary purpose is usually to provide bleed air to start the main engines. It is also used to run accessories such as air conditioning units and hydraulic pumps. A simplified APU representation is illustrated in Figure 1.

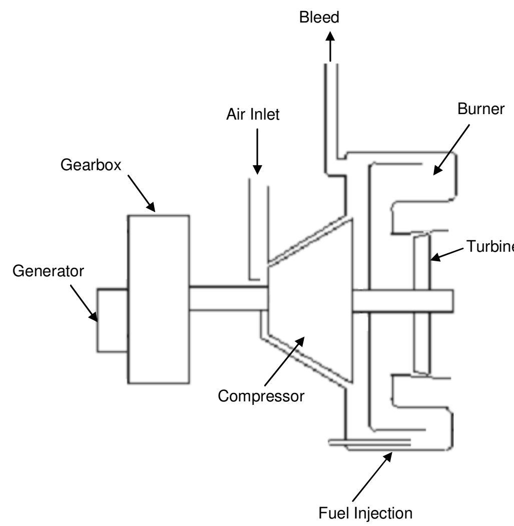

Figure 1: A simplified representation of an APU

In order to provide data for the APU fault detection, a mathematical model was developed using Matlab/ Simulink. Figure 1 shows a schematic view of the mathematical model developed. The main modeled blocks are represented in Figure 2.

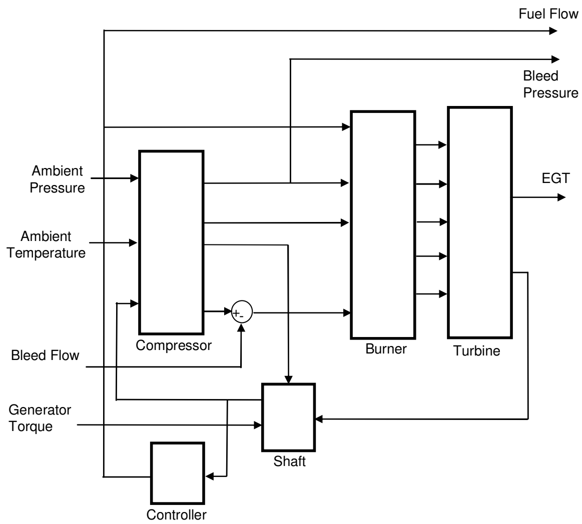

Figure 2: Simulation blocks and their relations in the depeloped APU model.

In the model, the compressor, burner and turbine model blocks were designed according to thermodynamic principles and information provided by nonlinear maps that describe the behavior of a real APU. The controller block comprises the control of shaft speed using fuel flow. The shaft block receives torque values from mechanical loads coupled to the APU shaft and calculates shafts rotation.

The main variables present in the model and used as measurements in the sample applications are the Exhaust Gas Temperature (EGT), Bleed Pressure (BP) and Fuel Flow (FF). The influences of ambient pressure (P_{amb}) and ambient temperature (T_{amb}) were also modeled.

Six different failure modes were seeded into the model, one at a time. These failure modes are:

- Bleed pressure sensor bias
- Fuel flow sensor bias
- Shaft speed sensor bias
- Exhaust gas temperature sensor bias
- Loss of compressor efficiency
- Loss of turbine efficiency

For the present study, four different fault detection methodologies were applied. Hotelling's T^2, Runger U^2, the GMM classifier (GMMC) and the GMM classifier with selection of variables of interest (GMMC-SV). The GMMC-SV classifier is the proposed method described in section 3 using GMM to estimate $p(\mathbf{y}_{new},\mathbf{z}_{new}|H=1)$ and $p(\mathbf{z}_{new})$.

The feature vectors comprised the steady state values for EGT, BP, FF, T_{amb} and P_{amb} during a simulation of APU startup. Considering Eq. (23), values of EGT, BP and FF were selected to form **y** and T_{amb} and P_{amb} composed vector **z**. All signals were corrupted with gaussian noise.

To characterize the behavior of the APU without faults, 1,000 simulations of APU startups were performed for different condition of pressure and ambient temperature. Both ambient conditions were simulated as following Gaussian distributions.

In Hotelling's T^2 and Runger's U^2 the generated dataset was used to estimate the mean vectors and covariance matrices presented in Eq. (4) and Eq. (11) respectively. In GMMC and GMMC-SV the distributions were estimated using a composition of five Gaussian distributions. The weights and parameters of each Gaussian were estimated using the EM algorithm.

For the generation of the test dataset, 12,000 simulations were performed, being 6,000 simulations of a healthy system and 6,000 simulations of the system with fault in different levels of severity (1,000 simulations for each failure mode).

To verify the performance of each method the Receiver Operating Characteristic (ROC) curve was used. The ROC curve was generated by varying the fault detection threshold and collecting false alarm rate and correct detection rate for each of the methods. In a ROC curve it is possible to classify the performance of the methods by evaluating the area under the curve. Better methods yield greater areas under the curve.

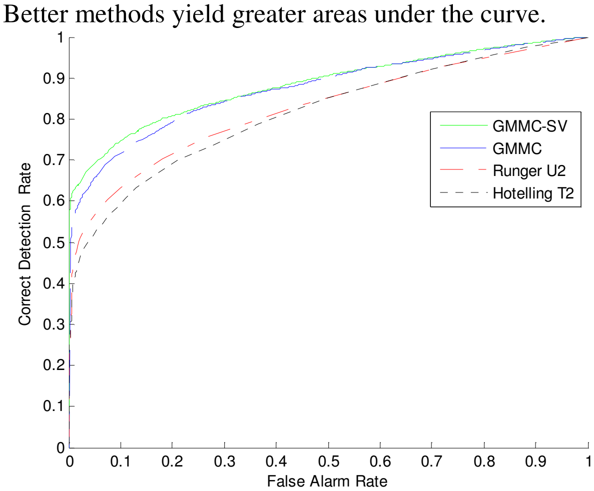

Figure 3: ROC curves for the implemented fault detection methods

The area under the ROC curve for each method is presented in Table 1.

Table 1: Area under ROC curve

Method	Hotelling's T^2	Runger U^2	GMMC	GMMC-SV
Area	0.838	0.860	0.875	0.884

It is possible to notice a significant difference in the performance when comparing methods that assume gaussianity in the data and methods that do not rely on that assumption. Methods that can deal with non-

Gaussian distributions achieve better scores. In the present case, such a finding can be ascribed to the nonlinearity of the simulation model employed to generate the data.

Analyzing Gaussian and non Gaussian methods separately, one can notice a superior performance of the methods where the subset of the monitored variables are selected. This result was already mentioned in (Runger, 1996) for Gaussian data and was extended for non Gaussian data in this work. The proposed GMMC-SV presented the better performance overall.

5. CONCLUSION

This work presented a novel data-driven methodology for fault detection. The concept of anomaly detection in a subset of the monitored variables proposed by (Runger, 1996) was extended to the case where the monitored variables do not followed a Gaussian distribution.

The method was tested using an APU dynamic model and showed betters results when compared to classical fault detection methods.

REFERENCES

Chiang L. H., Russel E. L. and Braatz R. D. (2001) *Fault Detection and Diagnosis in Industrial Systems*. 1st ed. Springer-Verlag London.

Duda, R. O., Hart, P. E., and Stork, D. G. (2001). *Pattern Classification*. 2nd ed. New York: Wiley.

De Maesschalck , R., Jouan-Rimbaud, D., and Massart, D.L. (2000). The Mahalanobis Distance, *Chemometrics and Intelligent Laboratory Systems*, 50, 1–18, 2000.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of Royal Statistical Society B*, 39, 1–38.

Hotelling, H. (1933). Analysis of a Complex of Statistical Variables into Principal Components, *Journal of Educational Psychology*, 24, 498–520.

Kourti, T. and MacGregor, J. F. (1995). Process Analysis, Monitoring and Diagnosis, Using Multivariate Projection Methods", *Chemometrics and Intelligent Laboratory Systems* 28, 3–21.

Kumar, S., Sotiris, V., and Pecht, M. (2008), Mahalanobis Distance and Projection Pursuit Analysis for Health Assessment of Electronic Systems, *in Proceedings IEEE Aerospace Conference*, Big Sky, MO.

Leão, B. P., Gomes, J. P. P., Galvão, R. K. H., and Yoneyama, T. (2009). Aircraft Flap and Slat Systems Health Monitoring Using Statistical Process Control Techniques, *in Proceedings of IEEE Aerospace Conference*, Big Sky, MO.

Mahalanobis, P. C. (1936). On the Generalized Distance in Statistics, *Proceedings of the National Institute of Science of India*, 12, 49–55.

Mimnagh, M. L., Hardman, W., and Sheaffer, J. (2000), Helicopter Drive System Diagnostics Through Multivariate Statistical Process Control, *in Proceedings IEEE Aerospace Conference*, Big Sky, MO.

Runger, G. C. (1996). Projections and the U2 Multivariate Control Chart", *Journal of Quality Technology*, 28, 313–319.

Webb, A (2002), *Statistical Pattern Recognition*. 2nd ed. West Sussex: John Wiley and Sons Ltd.

Yacher, L., and Orchard, M. (2003), Statistical Multivariate Analysis and Dynamics Monitoring for Plant Supervision Improvement, *in Proceedings Copper International Conference*.

João Paulo Pordeus Gomes holds a bachelor's degree on Electrical Engineering (2004) from Universidade Federal do Ceará (UFC), Brazil, and Master Degree on Aeronautical Engineering (2006) from Instituto Tecnológico de Aeronáutica (ITA), Brazil. He is currently pursuing his Ph.D. from ITA. He is with Empresa Brasileira de Aeronáutica S.A (EMBRAER), São José dos Campos, SP, Brazil, since 2006. He works as a Development Engineer of a R&T group at EMBRAER focused on PHM technology applications in aeronautical systems

Bruno P. Leão holds a bachelor's degree on Control and Automation Engineering (2004) from Universidade Federal de Minas Gerais (UFMG), Brazil, and a master's degree on Aeronautical Engineering (2007) from Instituto Tecnológico de Aeronáutica (ITA), Brazil. He is currently pursuing his Ph.D. on the theme of failure prognosis from ITA. He is with EMBRAER S.A. in Brazil since 2005. He has worked as a Systems Engineer on the areas of Flight Controls and Automatic Flight Controls. Since 2007 he is with the PHM research group at EMBRAER developing diagnosis, prognosis and health monitoring solutions for aircraft systems. He has published over 10 PHM related papers in peer-reviewed conferences.

Roberto Kawakami Harrop Galvão holds a bachelor's degree on Electronic Engineering (Summa cum Laude, 1995) from Instituto Tecnológico de Aeronáutica (ITA), Brazil. He also obtained the master's (1997) and doctorate (1999) degrees in Systems and Control from the same institution. Since 1998 he has been with the Electronic Engineering Department of ITA as a full-time academic. Dr. Galvão is a Senior Member of the IEEE and an Associate

Member of the Brazilian Academy of Sciences. He has published more than 100 papers in peer-reviewed journals and conferences. His main areas of interest are fault diagnosis and prognosis, wavelet theory and applications, and model predictive control.

Takashi Yoneyama is a Professor of Control Theory with the Electronic Engineering Department of ITA. He received the bachelor’s degree in electronic engineering from Instituto Tecnológico de Aeronáutica (ITA), Brazil, the M.D. degree in medicine from Universidade de Taubaté, Brazil, and the Ph.D. degree in electrical engineering from the University of London, U.K. (1983). He has more than 250 published papers, has written four books, and has supervised more than 50 theses. His research is concerned mainly with stochastic optimal control theory. Prof. Yoneyama served as the President of the Brazilian Automatics Society in the period 2004-2006.