

Fault Identification and Isolation in Dynamic Systems Using Multiple Models

Olívia M. A. Coelho¹, Wlamir O. L. Vianna², and Takashi Yoneyama³

^{1,2} *EMBRAER S.A., São José dos Campos, São Paulo, 12227-000, Brazil*
olivia.coelho@embraer.com.br
wlamir.vianna@embraer.com.br

³ *Instituto Tecnológico de Aeronáutica, São José dos Campos, São Paulo, 12228-900, Brazil*
takashi@ita.br

ABSTRACT

The demand for more reliability, safety and performance in industrial systems is rapidly increasing every day. The early detection of faults can avoid catastrophic events and the identification of the fault nature and severity can lead to the most appropriated and efficient maintenance task. Thus, an enhanced system diagnosis feature has the potential to increase safety and reduce the operational costs. In this context, fault detection and isolation techniques are used as the basis for building powerful decision making tools. This work's objective is to identify and isolate multiple faults in dynamic systems through signal processing. An approach based on a multiple-models architecture is considered whereas the plant output signals are compared with simulation data from a set of models representing the failure modes being analysed. The Autonomous Multiple Models (AMM) technique is chosen for further residue estimation and fault isolation. A case study using computational models representing an electro-mechanical system is carried out in order to validate the proposed method and evaluate its performance and limitations such as failure modes not mapped through the models and its capability to handle concurrent faults.

1. INTRODUCTION

The term fault refers to any deviation from a system normal operating state. A fault may lead to the system non-optimal operation and, eventually, to a failure state, where it can no longer achieve its required function. For the past decades, great effort has been put into developing novel techniques for systems Fault Detection and Isolation (FDI) (Vachtsevanos, Lewis, Hess, & Wu, 2006). FDI consists in identifying that the monitored system is in a faulty state (detection) and clas-

sifying which type of fault has occurred (isolation). The FDI techniques are generally classified as (Thirumarimurugan, Bagyalakshmi, & Paarkavi, 2016):

- **Model-Based Approaches:** methods that use mathematical or knowledge based models to detect faults;
- **Data-Driven Approaches:** methods that use the measurement data to extract information of the system health without having prior knowledge of the system model.

This work focuses on model-based FDI approaches. In the model-based approaches, features are calculated based on the measured signals and compared with the system expected behaviour. The expected behaviour is obtained by evaluating the features calculated from a mathematical model response. If the features calculated for the real plant at some point have a considerable deviation from the ones calculated for the model, this indicates that the system is in a faulty state. These features may be, for instance, residuals, parameter estimates or state estimates (Isermann, 2005). Figure 1 illustrates the residual calculation on a simple model-based FDI architecture. The system input is defined by $u(t)$, the output by $y(t)$ and the measured output by $y_v(t)$. Additionally, two disturbances are considered: the process noise $w(t)$ and the measurement noise $n(t)$.

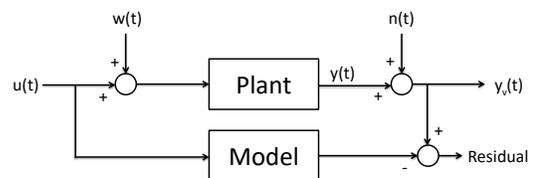


Figure 1. Model-Based Residual Generation

In an ideal process, if the residual is different from zero, the plant is in a faulty state. However, there are many uncertain-

Olívia M. A. Coelho et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 United States License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ties in the residual generation process. Besides the process and measurement noises, illustrated on Figure 1, there are also the model uncertainties. There will always be a trade-off between the model fidelity and the computational power required to process it.

Alongside with the advance in model-based FDI theory, the rapid growth of computational power over the past decades made feasible the use of these techniques in on-line fault diagnosis systems with industrial application. Today, there are studies for its application in several process in industries such as aeronautical (Vianna, Gomes, Galvão, Yoneyama, & Matsuura, 2011) (Xue & Guo, 2010), automobilistic (Ompusunggu, Papy, & Vandenas, 2016) and nuclear power (Coble, Ramuhalli, Bond, Hines, & Ipadhyaya, 2015).

Condition Based Maintenance (CBM) is the term used for the maintenance strategy where prognostics and health monitoring methods are used to determine maintenance actions based on the components current health conditions (Vachtsevanos et al., 2006). This concept comes as an alternative for the traditional maintenance approaches such as reactive (repair/replace when broken) and reliability centred (establishment of safe minimum levels of maintenance) (Goebel et al., 2012).

In summary, the residual based FDI process comprises two steps: residual generation and residual evaluation. The robust residual generation is the use of a robust filter or estimator that generates residuals which are unresponsive to uncertainties, without losing its sensitivity to the system faults. Some of the most popular methods used in this approach are the observer-based (Ding, 2008) (Hammouri, Kinnaert, & El Yaagoubi, 1999) and Kalman filter based (Ompusunggu et al., 2016) (Xue & Guo, 2010). Robust residual evaluation is the use of robust statistical analysis for evaluating signals or system parameters changes which correspond to faults. Some of these methods are based on generalized likelihood ratio test (Willsky & Jones, 1976) or sequential probability ratio test (Malladi & Speyer, 1999). Residual generation based on Kalman filters and residual evaluation based on multiple models algorithms are the main focus of this work.

2. KALMAN FILTER BASED RESIDUAL GENERATION

The Kalman filter is an effective state estimation tool initially proposed by (Kalman, 1960). The algorithm implements a predictor-corrector type estimator. For linear dynamic systems which are corrupted by white noise, the linear Kalman Filter solution is optimal in the sense that it minimizes the error covariance of the system states estimate (Vachtsevanos et al., 2006). The Kalman filter output prediction error, or innovation sequence, can be used as residual in FDI systems. If the filter is correctly predicting the system behaviour, the innovation sequence will be white noise.

Therefore, by analysing the innovation sequence properties, it is possible to detect faulty states.

Considering the discrete time linear system described by Eq. (1), the innovation sequence is determined by Eq. (2).

$$\begin{cases} x_{k+1} = Ax_k + B(u_k + w_k) \\ z_k = Hx_k + v_k \end{cases} \quad (1)$$

$$\text{InnovationSequence} = \tilde{z}_k^{(i)} = z_k - H\hat{x}_k^- \quad (2)$$

Where $\tilde{z}_k^{(i)}$ represents the innovation sequence for model (i) at instant k, while z_k is the system measured output and \hat{x}_k^- is the Kalman Filter *a priori* state estimate at instant k.

3. MULTIPLE MODELS FDI ARCHITECTURES

Up until this point, the discussion focused on how it is possible to detect that the system is in a faulty state by evaluating the residuals obtained using a plant model in its healthy state and filters generated based on this model. The next step would be to isolate the fault. In a multiple models FDI architecture, the fault isolation capability is achieved by assembling a models bank comprising models for the system in its healthy and faulty states. The residuals are then calculated simultaneously for all models. If the residual resulting from the comparison of the plant output with a faulty system model is close to zero, then it is likely that the system is in the faulty state represented by this particular model.

A similar approach can be used where the residuals are generated from a bank of filters, each one built from a different model from the models bank (see Figure 2). In this case, the filter built from the model that best represents the system behaviour will generate the most accurate output estimate and, therefore, have the prediction error that most resembles white Gaussian noise.

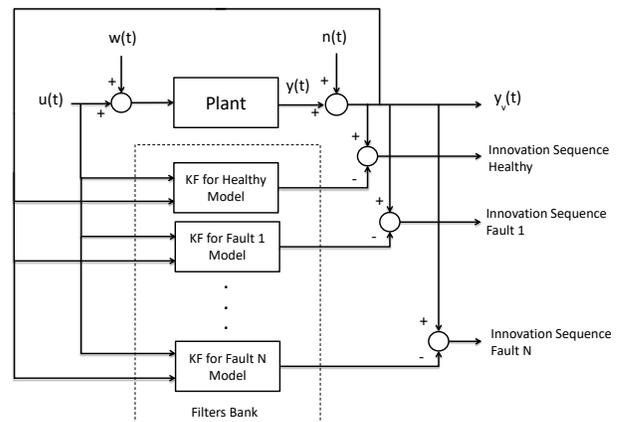


Figure 2. Schematic Description of Model-Based Fault Detection with Multiple Models and Filters

With the configuration proposed on Figure 2, the output produced as residual for the FDI analysis will be a discrete time series for each filter prediction error. In order to process the information from these time series and translate it into meaningful information for the system prognostics, a set of algorithms denominated Multiple Models Algorithms (MMA) can be applied.

3.1. Multiple Models Algorithms

The MMA are estimate-decision based algorithms commonly applied to target tracking problems (Li & Jilkov, 2005) (Pitre, 2004) and also suitable to system FDI (Maybeck, 1999) (Efe & Atherton, 1997). In general, the implementation of MMA comprises three main steps (Li & Jilkov, 2005):

- To build a models bank and assume these models as possible candidates of the system state;
- To run a set of filters in parallel with the process, each based on a unique model in the bank;
- To combine the filters output and generate the overall estimate for the system state.

The result of the process is the likelihood that the system current state matches each one of the models in the models set. The algorithm will be capable of correctly identifying the system state if the model set contains the true system mode, therefore, the performance of the MMA depends deeply on the models bank assembled.

The MMA can be classified in three generations (Bar-shalom & Blair, 2000): First Generation - Autonomous; Second Generation - Cooperative and Third Generation - Variable Structure. The later generations come as more sophisticated algorithms, with increased performance but also increased complexity (Pitre, 2004). They inherit characteristics from the earlier generations while introducing new ones. For this work the Autonomous Multiple Model (AMM), a first generation MMA, was the residual evaluation technique chosen.

Since the AMM is classified as a first generation MMA, it is characterized by independence between filters. The filters operate without any information from the other models in the models set. One Kalman filter is run for each model independently and the model-based estimates are fused to generate an overall estimate. The following assumptions are made for first generation MMA (Li & Jilkov, 2005): the system is time invariant and the true state of the system corresponds to a model in the models set.

The output of the AMM algorithm will be the probability that the system current state corresponds to each model in the models bank. The probabilities are updated in each time step using information up to the current time step. They are calculated based on each model's likelihood, denoted by $L_k^{(i)}$ (likelihood for model (i) at instant k) and given by Eq. (3).

$$L_k^{(i)} = p[z_k | m^{(i)}, z_{k-1}] = \mathcal{N}(\tilde{z}_k^{(i)}; 0; S_k^{(i)}) \quad (3)$$

Eq. (3) gives the likelihood function of model $m^{(i)}$ (or model likelihood), where $m^{(i)} \in \mathbb{M}$ (models bank set). The function $\mathcal{N}(x; \mu; \sigma^2)$ is the Gaussian (normal) probability density function of x with mean μ and standard deviation σ . For the definition of $L_k^{(i)}$ on Eq. (3), the variable $\tilde{z}_k^{(i)}$ represents the residual or prediction error (as defined in Eq. (2)), the mean μ is considered to be zero and the standard deviation is given by the measurement prediction error covariance $S_k^{(i)}$ of model (i) at instant k .

The measurement of how probable a given model (i) is in effect at time k is called mode probability ($\mu_k^{(i)}$) and is given by Eq. (4). The model probability is calculated as a weighted average using the models likelihood as the weights assigned to each filter and normalized to guarantee that the sum of all probabilities will be 1 ($\sum_{i \in \mathbb{M}} \mu_k^{(i)} = 1$).

$$\mu_k^{(i)} = \frac{\mu_{k-1}^{(i)} L_k^{(i)}}{\sum_{j \in \mathbb{M}} \mu_{k-1}^{(j)} L_k^{(j)}} \quad (4)$$

The main advantage of the AMM when compared with other non-MMA approaches is the fact that it fuses outputs of the different elemental filters to yield the overall estimate, however, the other filters results are only combined in the final estimate fusion step. The AMM, as a first generation MMA, does not implement filter interaction. Consequently, its performance may be inferior when compared with later generation MMA, especially when handling systems with frequent mode jumps (Gomes, 2008). As expected, the decision to use first, second or third generation MMA is a trade-off between performance and complexity.

4. AN STUDY OF THE APPLICATION OF THE AMM ALGORITHM FOR FDI

This work aims to evaluate the application of multiple model algorithms in the fault isolation and identification of dynamic systems with single and multiple failure modes. To illustrate and assess the application of the proposed method, a case study of an electro mechanic system is used.

4.1. Case Study

The system considered as case study for this work is a laboratory bench consisting of an electro-mechanic rotational system. The system comprises an electric motor, a pulley and belt set, the motor and load disks, bearings and a flexible shaft. This laboratory bench is used on the case study presented in (Jha, Dauphin-Tanguy, & Ould-Bouamama, 2016). Figure 3 (adapted from (Jha et al., 2016)) brings a schematic

with the elements from the laboratory bench considered for the system modelling.

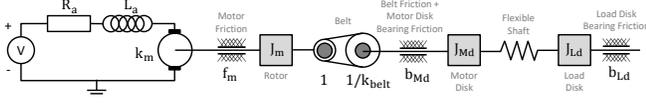


Figure 3. Case study system diagram

The technique chosen to derive the mathematical model was the bond graph representation, which allows the system to be easily converted into a state-space representation. The system bond graph is shown in Figure 4.

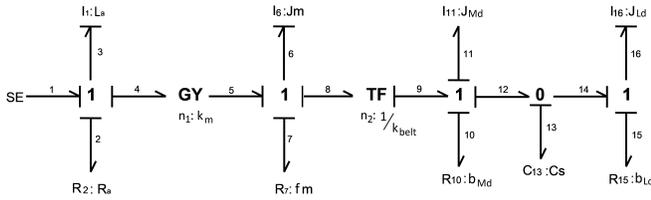


Figure 4. System bond graph representation

The DC motor current and the load bearing angular speed are assumed as the system outputs. The system input is defined as the effort source, i.e. the voltage applied in the DC motor. A system state-space representation was derived from Figure 4 bond graph, the state vector defined is $x = [p_3 \ p_6 \ q_{13} \ p_{16}]^T$ and the state-space matrices are shown in Eq. (5).

$$A = \begin{bmatrix} -\frac{R_a}{L_a} & -\frac{n_1}{J_m} & 0 & 0 \\ \frac{n_1}{L_a} & -\frac{1}{J_m}(f_m + n_2^2(J_{Md} + b_{Md})) & -\frac{n_2}{C_s} & 0 \\ 0 & \frac{n_2}{J_m} & 0 & -\frac{1}{J_{Ld}} \\ 0 & 0 & \frac{1}{C_s} & -\frac{1}{J_{Ld}} \end{bmatrix} \quad (5)$$

$$B = [1 \ 0 \ 0 \ 0]^T \quad H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad D = 0$$

The state space equations were implemented on MATLAB, from MathWorks $\text{\textcircled{R}}$, and the following values were assigned to the variables:

$$\begin{aligned} R_a &= 1.23\Omega & J_{Md} &= 9.00 \times 10^{-2} \text{Kg.m}^2/\text{rad} \\ L_a &= 1.34 \times 10^{-3} \text{H} & b_{Md} &= 4.22 \times 10^{-1} \text{N.m.s}/\text{rad} \\ J_m &= 6.76 \times 10^{-3} \text{Kg.m}^2/\text{rad} & C_s &= 1.79 \times 10^{-1} \text{N.m}/\text{rad} \\ n_1 &= 2.57 \text{A}/\text{N.m} & J_{Ld} &= 6.70 \times 10^{-3} \text{Kg.m}^2/\text{rad} \\ n_2 &= 2.70 & b_{Ld} &= 5.10 \times 10^{-1} \text{N.m.s}/\text{rad} \\ f_m &= 2.00 \times 10^{-1} \text{N.m.s}/\text{rad} \end{aligned}$$

Initially, two failure modes are considered, they are referred as motor and bearing failures. Two resistances were added to the bond-graph representation shown on Figure 4. The first one corresponding to an electrical motor failure, which is represented by adding an electrical resistance (R_{deg1}) to the first

1-junction and the second one corresponding to an increased friction in the pulley bearing, represented as another dissipative resistance (R_{deg2}) in the third 1-junction. These two failure modes were selected because they provide enough separation since they affect different model parameters and different entries in the matrix A.

For the case study, R_{deg1} is considered as a degradation of +65% in R_a and R_{deg2} is considered as a degradation of +150% in b_{Md} . The severity of the degradation was selected based on the observation of the system response in several simulations. The values were selected to produce notable changes in the system outputs when compared to the healthy system simulation.

A third failure mode, referred as shaft failure, was also considered later on this study. A capacitance was added to the 0-junction where C_s is, representing a change in the flexible shaft elasticity. The degradation C_{deg} is considered to cause a reduction of 50% in C_s . The models and failure modes presented herein are the basis of the study case and will be explored more in the following subsections.

4.2. FDI Using the AMM Algorithm

This work focuses on residual based FDI process with residual generation based on Kalman filters and residual evaluation based on MMA. Therefore, one Kalman filter was designed for each model in the models bank and the output prediction error, or innovation sequence, is used as residual in the FDI architecture. Initially, the models bank considered comprises: the healthy system, the system with a motor failure, the system with a bearing failure and the system with simultaneous motor and bearing failures. Figure 5 brings the multiple-models architecture considered for the case study.

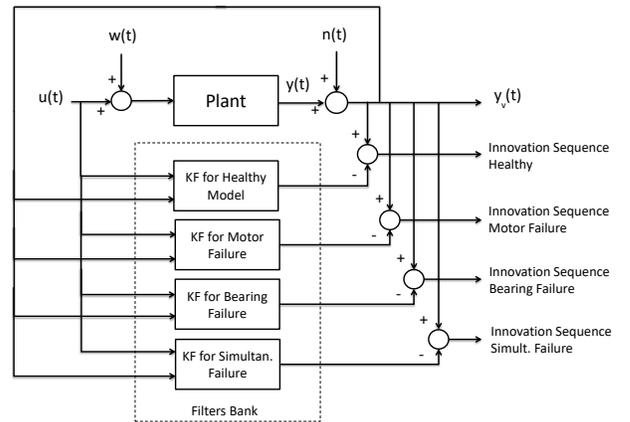


Figure 5. Innovation Sequence Calculation for Case Study

The AMM algorithm detailed in Section 3.1 was used for assessing the innovation sequence properties for each filter estimate and for generating a single overall estimate of the system current state. One Kalman filter runs for each model in

the models set and the model-based estimates are used for calculating the models likelihood at each instant (Eq. (3)). The model probability at instant k corresponds to the probability that the system current state matches that particular model and is calculated based on the model likelihood as shown in Eq. (4). As more output samples are collected, the algorithm estimate becomes more accurate.

Considering the multiple-models architecture presented on Figure 5 and the AMM algorithm, several simulations were carried out. For the first results presented herein, the system was considered in its healthy state. Figure 6 brings the outputs estimated with each one of the filters in the filters bank and Figure 7 brings the models estimated probability over time. By inspecting the filter's predictions it is clear that the one from the healthy model filter fits better the measured output, therefore, the algorithm quickly converges to indicate 100% probability of the system being healthy.

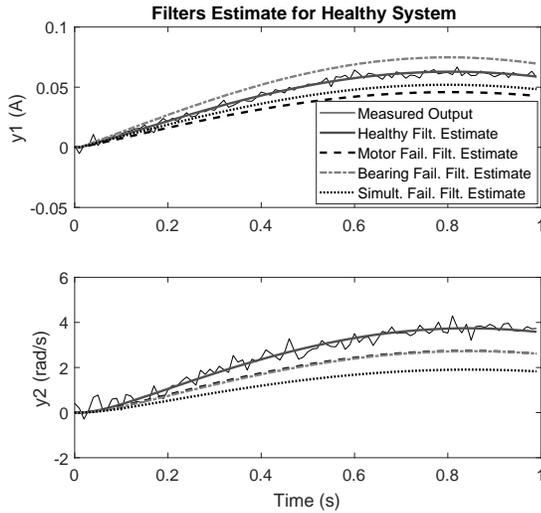


Figure 6. Filters Estimate on Healthy System Simulation

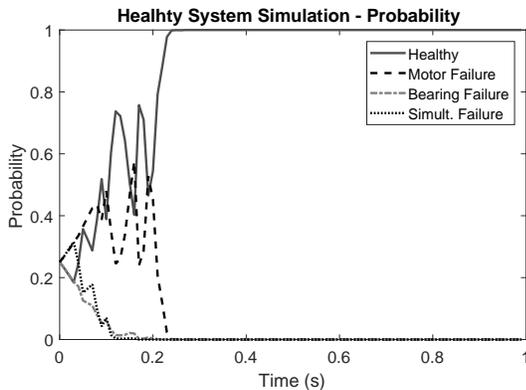


Figure 7. Healthy System Simulation

The simulation was repeated with the system in each one of

the faulty states (motor failure, bearing failure and simultaneous motor and bearing failures). The algorithm was capable of correctly isolating all failure modes in approximately 0.2 to 0.4 seconds. As an example, Figure 8 brings the result for the simulation with the system with a motor failure. It is important to highlight that for this first round of simulations, the current state of the system matched exactly one of the models considered in the models bank, thus, representing an ideal scenario.

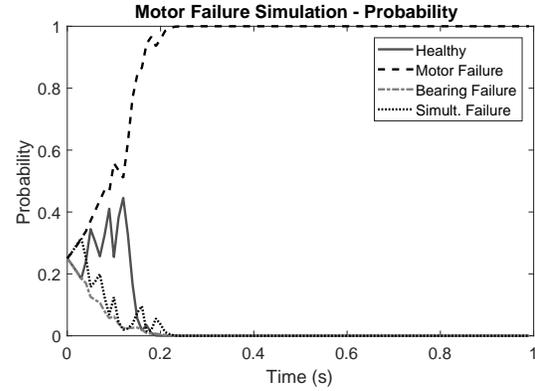


Figure 8. System with Motor Failure Simulation

4.3. Algorithm Performance vs. Failure Severity

In the first simulations, an ideal scenario was considered where the plant model used in the simulation matched exactly one of the models in the models bank. However, in a real system, the motor and bearing failures would most likely happen gradually, leading to a progressive increase of the resistive parameters. In order to assess how the algorithm would perform with different failure severities other than the standard value used in the failure modelling and identify what is the severity threshold for the failure to be detected, the following study was carried out. First, a metric for assessing the algorithm performance was defined. The metric chosen was the time to correctly detect and isolate the failure. This time was defined as the instant where the correct model probability reached 90% and stayed above 90% until the end of the simulation.

A range of degradation severities was selected for the simulations. One simulation was run for each test case and the time to isolate the failure was computed for the cases where the algorithm was able to successfully isolate the correct failure. For the motor failure degradation parameter, values of 0 to +100% of R_a , with a step of 1%, were considered and for the bearing failure degradation parameter, values of 0 to +200% of b_{Md} , with a step of 2.5%. The results are presented on Figures 9 and 10.

In Figures 9 and 10, the x-axis represents the percentage of degradation in R_a (Figure 9) and b_{Md} (Figure 10). The y-axis

represents the time that the algorithm took to correctly isolate the failure. The grey area represents the range in which no failure was detected by the algorithm within the simulation time. A bold line marks the 0% degradation simulation, which corresponds to the healthy system while the dashed line marks the +65% and +150% degradation case, which corresponds to the standard degradation used for the failure model in the models bank. The bars are plotted for each test case in which the failure was correctly isolated and they bring information of how much time it took for the algorithm to detect the failure.

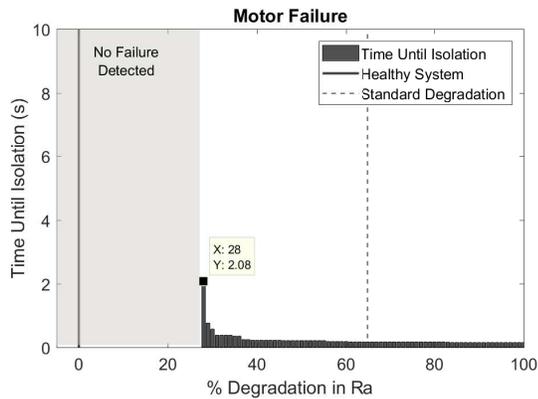


Figure 9. Time Until Isolation versus Degradation Severity for Motor Failure

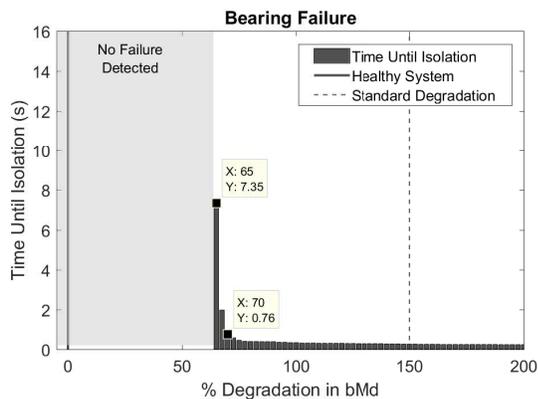


Figure 10. Time Until Isolation versus Degradation Severity for Bearing Failure

As it can be seen, the motor failure was correctly isolated for degradations starting from +28% and the bearing failure for degradations starting from +65%. The complete simulation of all the test cases was run multiple times and the results obtained were slightly different in each run. However, the threshold for the failure detection was always close to the results presented in Figures 9 and 10.

For this second study the algorithm performance was assessed in a non-ideal scenario, where the failure severity does not

match exactly the standard value used in the models bank. Whether or not the algorithm was capable of isolating the failure and also the time until the correct isolation happened, were used as metrics. For both the motor and bearing failures, it was possible to identify a range in which the degradation severity is not enough for the algorithm to detect that the system is in a faulty state. However, this can be a desired characteristic for a FDI system and some reasons are listed below:

- The true system does not match exactly the model. There are many uncertainties regarding the modelling process and if the algorithm is too sensitive this may lead to false positives;
- Small degradations that do not affect the system overall behaviour usually do not require a maintenance action. The intention of FDI in real systems is to anticipate and support maintenance actions, however, if any mild degradations are flagged as system failures, unnecessary maintenance stops and procedures may be started;
- Disturbances can trigger false alarms in high sensitivity fault detection algorithms. A degradation in its earliest stage can be so small that it becomes hard to distinguish it from disturbances like measurement noise.

It is important to notice that the desired sensitivity of the fault diagnosis technique will be highly dependent on the application, specially on the hazard of the failure. For FDI systems that assist safety related features, high sensitivity is crucial, therefore, more sophisticated methods are usually applied.

As one of the assumptions of the AMM algorithm is that the true state of the system corresponds to a model in the models set, when this assumption is not met the algorithm behaviour is unknown. Besides the variation on the degradation severities, another scenario in which the AMM assumption discussed herein will not be met is when there is a failure in the system that is not modelled in the models bank.

4.4. Algorithm Response to a Failure not Modelled

The results presented up until this point considered only scenarios in which the system parameter degradations were restricted to the failure cases modelled in the models bank. This represents degradation in only two parameters, while the system model studied has 11 parameters. At the same time that any of those parameters are prone to degradations, there must be a trade-off between algorithm complexity and FDI capability. Additionally, considering the modelling limitations, there will always be modelling uncertainties. As a consequence, in model-based FDI, there will be the possibility of the system to be in a failure state that is not considered in the models bank.

To study the response of the AMM to a failure mode not

present in the models bank, the flexible shaft failure described in Section 4.1 was considered. The same architecture illustrated in Figure 5 was used for the innovation sequence calculation and a simulation was carried out using the plant model with the flexible shaft failure. The results are presented on Figures 11 and 12. The flexible shaft failure was isolated as a simultaneous motor and bearing failure after approximately 0.5 seconds of simulation. The result indicates that the Kalman filter built with the simultaneous failure model was the one that produced the closest estimate for the system outputs in the flexible shaft failure simulation. This can be visualized in Figure 11. The top plot shows that for the DC motor current, the closest estimate was produced by the healthy model filter. While the bottom plot shows that for the load bearing speed, the simultaneous failure filter produced the best estimate.

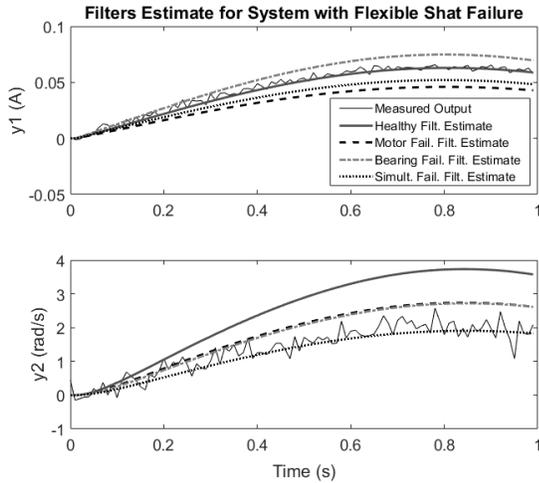


Figure 11. Filters Estimate on Simulation of the System with a Shaft Failure

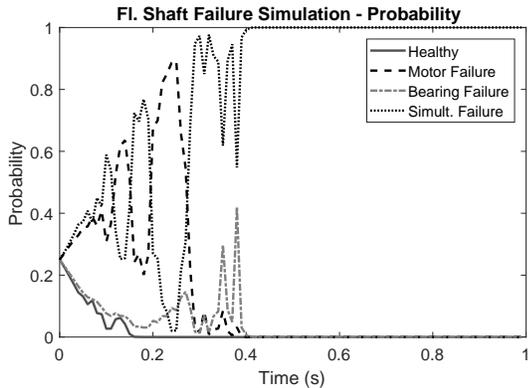


Figure 12. Shaft Failure System Simulation - Failure not included in Models Bank

For the AMM algorithm to be able to correctly isolate this new failure mode, a new model must be added to the models

bank. A new simulation was run with an extended models bank, including the shaft failure model. As expected, in this scenario the failure was successfully isolated (see Figure 13).

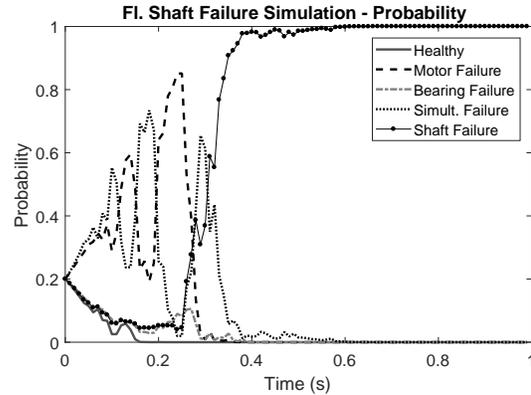


Figure 13. Shaft Failure System Simulation - Failure Included in Models Bank

However, for this simulation, only one extra model was added to the models bank, if the intention is to cover all single and multiple failure scenarios, four models would have to be added (flexible shaft failure, simultaneous flexible shaft and motor failures, simultaneous flexible shaft and bearing failures and simultaneous flexible shaft, motor and bearing failures). To handle multiple failures with the proposed FDI architecture and cover all possible combinations, the number of models needed in the models bank would increase significantly for each new single failure mode added.

4.5. Alternative Multiple-Model Architectures

The multiple-model based FDI architecture presented herein is based on the AMM algorithm and, therefore, is tied to its limitations and assumptions. One of those being that the true system state is assumed to be modelled in the models bank. This means that if the system current state does not match any of the models, the AMM result will point to the Kalman filter that produces the most accurate estimate, regardless of it being an adequate estimate. As a consequence, if one's intent is to cover all expected failures, the number of models needed can be significantly large. This becomes specially challenging when multiple failures are being handled, since for each new failure mode added to the FDI scheme, its combination with the other failure modes will also result in new models in the models bank. Besides the monitoring of simultaneous failures, the monitoring of different magnitudes of failure severity can also lead to a substantial increase in the number of models in the bank.

Having a large number of models in the bank will allow the user to cover more possible system states. However, this will come with considerable computational cost and may jeopardize the FDI system performance. To address this problem,

two alternative multiple-model architectures are proposed in (Maybeck & Hentz, 1987) and (Maybeck & Stevens, 1991): the moving bank and hierarchical structure approaches. Both aiming to reduce the number of models being computed in each time step.

In the moving bank approach, only a subset of models is considered in each time instant. This subset is composed by the models that are more similar to the current system state. Once the system state changes, the subset of models also changes, resulting in a moving bank (Maybeck & Hentz, 1987). On the hierarchical structure, the models are separated into different levels and only one level is computed in each time step. If a level 0 failure state is confirmed, the level 1 bank immediately below that state is brought on-line. This level 1 bank will comprise models which already take into consideration that the level 0 failure state was confirmed (Maybeck & Stevens, 1991). In this work, both moving bank and hierarchical structure approaches were studied as alternatives for handling multiple failures.

A brief proof of concept was built considering one of these alternative architectures, the hierarchical approach. The depth of the structure depends on how many simultaneous failures needs to be covered (e.g.: two levels for double failures). In the level 0, there are $n+1$ Kalman filters running in parallel, one for each of the n single failure modes plus one for the healthy system model. Figure 14 brings a schematic of the concept considering the case study and the three failure modes presented. Only one branch of the level two is shown for simplification purposes, however, analogue branches can be built for all double-failure nodes in level one.

The healthy system is included in each level to allow the FDI system to move back to the level 0 if the failure is no longer detected. The following bank-switching decision rule is proposed in (Maybeck, 1999): the model probability must exceed 0.95 or be above 0.9 for 10 sample periods. The example presented illustrates the application of the hierarchical structure for AMM based FDI algorithms.

The complexity of building the hierarchical tree for defining the models bank pays off with the reduced computational cost since a smaller set of filters needs to be run in each time step. The depth of the three will depend on how many multiple failures the user needs to detect at the same time. Too many levels may be impractical and unnecessary since in real applications it is unlikely that the system will continue to operate regardless of how many failures have occurred. Ultimately, this will be another design decision highly dependent on the system and application.

5. CONCLUSION

The growing demand for performance with safety and reliability in the modern industry gives strength to the advent

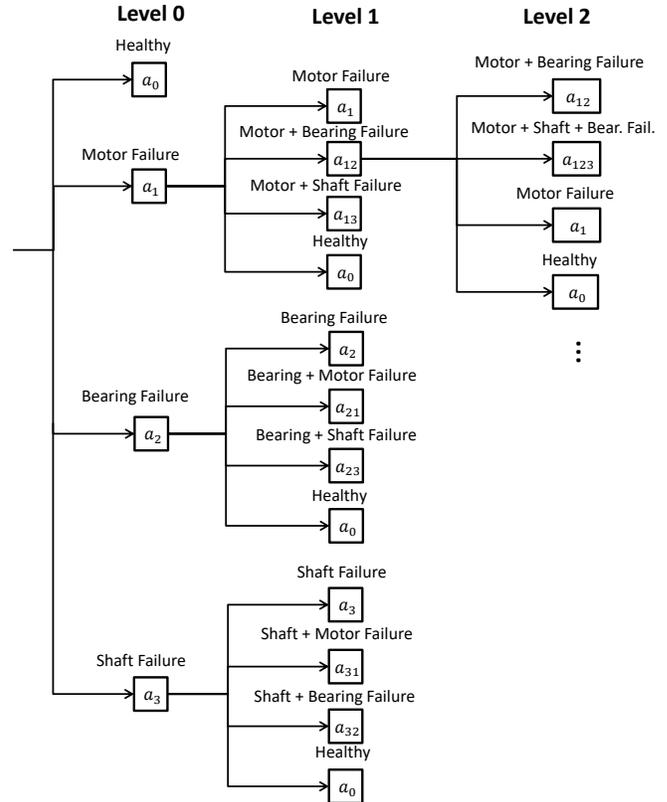


Figure 14. Hierarchical Structure Concept Applied to Case Study Considering up to Three Simultaneous Failures

of fault detection and isolation systems. The early detection of faults and failures can prevent catastrophic events and the ability to not only detect, but also isolate the failure, can allow the system reconfiguration to a fault tolerant state and reduce maintenance tasks time. This work focused on the application of a model-based FDI algorithm based on multiple models. In summary, the simulations performed using the case study show positive results for the application of the AMM algorithm in FDI and the technique applied can be generalized to other systems. The algorithm limitations and challenges bring special attention to the importance of the models bank selection. It is impractical to supervise all possible failure modes using a model-based FDI technique because it would lead to a large models bank, a complex FDI architecture and costly computational power. Additionally, the failure modes considered must have enough separation to allow the algorithm to distinguish between them.

The main motivation of this work is the application of the studied model-based FDI in the industry. The interest of the industry in health monitoring techniques can be narrowed down to two main factors: safety and maintenance related advantages. For safety enhancement, the AMM based FDI algorithm can be used not only for failure isolation but also

to allow the construction of reconfigurable control systems that are fault tolerant. As for the maintenance advantages, the operator can use the technique to allow the optimization of maintenance tasks and to reduce the time of maintenance routines, allowing the introduction of new maintenance and logistic concepts and reducing the operational cost. In this context, FDI techniques are used as the basis for building powerful decision making tools.

NOMENCLATURE

A	State Space Matrix
B	State Space Input Matrix
D	State Space Feed-Through Matrix
e	Effort Variable
f	Flow Variable
GY	Gyrator
H	State Space Output Matrix
K_k	Kalman Gain at instant k
$L_k^{(i)}$	Likelihood of model (i) at instant k
\mathbb{M}	Model Set
$m^{(i)}$	i -th model in the model set
$\mu_k^{(i)}$	Mode Probability
$\mathcal{N}(x; \mu; \sigma)$	Normal distribution function of x with mean μ and standard deviation σ
$\hat{n}(t)$	Measurement Noise Vector
TF	Transformer
$S_k^{(i)}$	Measurement Prediction Error Covariance of Model (i) at Instant k
$u(t)$	Input Vector
$v(t)$	Measurement Noise Vector
$w(t)$	Process Noise Vector
x_k	Discrete Time State Vector
\hat{x}_k	Kalman Filter State Estimate at Instant k
\hat{x}_k^-	Kalman Filter State Estimate Prediction at Instant k
$\hat{x}_{k k}$	Overall State Estimate at Instant k
$x(t)$	State Vector
$\dot{x}(t)$	State Vector Derivate
$y(t)$	Output Vector
y_k	Discrete Time Output Vector
$y_v(t)$	Measured Output Vector
z_k	Discrete Time Measured Output Vector
$\tilde{z}_k^{(i)}$	Innovation Sequence for Model (i) at Instant k

REFERENCES

- Bar-shalom, Y., & Blair, W. (2000). *Multitarget-multisensor tracking*. Artech Print.
- Coble, J., Ramuhalli, P., Bond, L. J., Hines, J., & Ipadhyaya, B. (2015). A review of prognostics and health management applications in nuclear power plants. *International Journal of prognostics and health management*, 6(16), 1–22.
- Ding, S. X. (Ed.). (2008). *Model-based fault diagnosis techniques, design schemes, algorithms, and tools*. Springer.
- Efe, M., & Atherton, D. (1997). The IMM approach to the fault detection problem. *IFAC System Identification*, 30(11), 603–608.
- Goebel, K., Saxena, A., Daigle, M., Celaya, J., Roychoudhury, I., & Clements, S. (2012). Introduction to prognostics. In *1st european conference of the prognostics and health management society*. Retrieved from <http://ftp.phm.society.org/sites/phm.society.org/files/Tutorialprognostics.pdf>
- Gomes, J. B. B. (2008). *An overview on target tracking using multiple model methods* (Unpublished master's thesis). Universidade Técnica de Lisboa.
- Hammouri, H., Kinnaert, M., & El Yaagoubi, E. H. (1999). Observer-based approach to fault detection and isolation for nonlinear systems. *IEEE Transactions on Automatic Control*, 44, 1879–1884.
- Isermann, R. (2005). Model-based fault-detection and diagnosis—status and applications. *Annual Reviews in Control*, 29, 71–85.
- Jha, M. S., Dauphin-Tanguy, G., & Ould-Bouamama, B. (2016). Particle filter based hybrid prognostics for health monitoring of uncertain systems in bond graph framework. *Mechanical Systems and Signal Processing*, 75, 301–329.
- Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, 82, 35–45.
- Li, X. R., & Jilkov, V. P. (2005). Survey of maneuvering target tracking. part v. multiple-model methods. *IEEE Transactions on Aerospace and Electronic Systems*, 41, 1255–1321.
- Malladi, D. P., & Speyer, J. L. (1999). A generalized shiryayev sequential probability ratio test for change detection and isolation. *IEEE Transactions on Automatic Control*, 44, 1522–1534.
- Maybeck, P. S. (1999). Multiple model adaptive algorithms for detecting and compensating sensor and actuator/surface failures in aircraft flight control systems. *International Journal of Robust and Nonlinear Control*, 9, 1051–1070.
- Maybeck, P. S., & Hentz, K. P. (1987). Investigation of moving-bank multiple model adaptive algorithms. *Journal of Guidance, Control, and Dynamics*, 10(1), 90–96.
- Maybeck, P. S., & Stevens, R. D. (1991). Reconfigurable flight control via multiple model adaptive control methods. *IEEE Transactions on Aerospace and Electronic systems*, 27, 470–480.
- Ompusunggu, A. P., Papy, J.-M., & Vandenplas, S.

(2016). Kalman-filtering-based prognostics for automatic transmission clutches. *IEEE/ASME Transactions on Mechatronics*, 21, 419–430.

Pitre, R. (2004). *A comparison of multiple-model target tracking algorithms* (Unpublished master's thesis). University of New Orleans.

Thirumarimurugan, M., Bagyalakshmi, N., & Paarkavi, P. (2016). Comparison of fault detection and isolation methods: A review. In *Proceedings of 10th IEEE international conference on intelligent systems and control*.

Vachtsevanos, G. J., Lewis, F., Hess, A., & Wu, B. (Eds.). (2006). *Intelligent fault diagnosis and prognosis for engineering systems*. Springer.

Vianna, W. O., Gomes, J. P., Galvão, R. K., Yoneyama, T., & Matsuura, J. P. (2011). Health monitoring of an auxiliary power unit using a classification tree. In *Proceedings of the annual conference of the prognostics and health management society*.

Willsky, A., & Jones, H. (1976). A generalized likelihood ratio approach to the detection and estimation of jumps in linear systems. *IEEE Transactions on Automatic Control*, 21, 108–112.

Xue, W., & Guo, Y. q. (2010). Application of kalman filters for the fault diagnoses of aircraft engine. In *Kalman filter*. IntechOpen.

BIOGRAPHIES

Olívia Maria Alves Coelho holds a bachelor's degree in Control and Automation Engineering (2015) from Univer-

sidade Federal de Minas Gerais (UFMG), Brazil, and Master's Degree in Aeronautical Engineering (2018) from Instituto Tecnológico de Aeronáutica (ITA), Brazil. She is with Empresa Brasileira de Aeronáutica S.A (EMBRAER), São José dos Campos, SP, Brazil, since 2016. She works as a Product Development Engineer with focus on fuel and inerting systems control.

Wlamir Olivares Loesch Vianna holds a bachelor's degree in Mechanical Engineering (2005) from Universidade de São Paulo (USP), Brazil, Master's Degree in Aeronautical Engineering (2007) from Instituto Tecnológico de Aeronáutica (ITA), Brazil, and the Ph.D degree in Electronic Engineering and Computer Science (2017) from Instituto Tecnológico de Aeronáutica (ITA), Brazil. He is with Empresa Brasileira de Aeronáutica S.A (EMBRAER), São José dos Campos, SP, Brazil, since 2007. He works as a Development Engineer of a R&T group at EMBRAER focused on PHM technology applications in aeronautical systems.

Takashi Yoneyama is a Professor of Control Theory with the Electronic Engineering Department of the Instituto Tecnológico de Aeronáutica (ITA). He received his bachelor's degree (1975) and master's degree (1979) in Electronic Engineering from Instituto Tecnológico de Aeronáutica (ITA), Brazil, the M.D. degree in Medicine (1993) from Universidade de Taubaté (UNITAU), Brazil, and the Ph.D. degree in Electrical Engineering (1983) from the Imperial College London, U.K. (1983). He has more than 400 published papers, has written four books, and has supervised more than 100 Ph.D. and masters thesis. His research is concerned mainly with stochastic optimal control theory. He served as the President of the Brazilian Automatics Society in the period of 2004-2006.