Available online at www.sciencedirect.com

ScienceDirect

IFAC Papers Online CONFERENCE PAPER ARCHIVE

IFAC PapersOnLine 50-1 (2017) 13544–13549

# A Data-Driven Approach For Actuator Servo Loop Failure Detection

**S. Urbano** *, **E. Chaumette** **, **P. Goupil** *, **J-Y. Tourneret** ***

* *Airbus, 316 route de Bayonne,Toulouse,France,31060*
** *ISAE/TéSA, 10 Avenue Edouard-Belin,Toulouse,France,31055*
*** *ENSEEIHT/TéSA, 2 rue Charles Camichel,Toulouse,France,31071*

**Abstract:** This paper studies a data-driven approach to detect faults in flight control systems of civil aircraft. A particular class of failures, referred to as Oscillatory Failure Cases (OFC), impacting the actuator servo loop has motivated the authors to consider a data-driven approach based on distance and correlation measures (see reference [Goupil et al.(2016). A data-driven approach to detect faults in the Airbus flight control system. IFAC-PapersOnLine, 49(17), 52-57] of this paper) leading to promising results compared to the state-of-the-art methods based on analytical redundancy. The present paper extends the formulation and the results of the considered OFC detection approach investigating Support Vector Machine (SVM) techniques to define a more accurate detector based on distance and correlation measures.



*Keywords:* Flight Control, Fault Detection and Diagnosis, Oscillatory Failure Case, Distance, Correlation, Classification, Support Vector Machines.

## 1. INTRODUCTION

Economic and environmental efficiency has nowadays an increasing role in the design of a modern civil aircraft. Reducing the structural airframe weight for a given passenger capacity without compromising the structural strength of the aircraft is a main research axis. The unwanted oscillation of a control surface represents a significant design case in the development of a fault-tolerant system. Indeed, it may not only induce local structural load augmentation, but can also match the frequency of a flight mechanical mode (deteriorating the flight handling qualities) or a weakly damped structural mode (leading to a divergent behaviour in the worst case). For this reason, the impact of a spurious control surface oscillation to the airframe structural design is not negligible. The ability to detect an oscillation beyond a given amplitude in a given time interval, whatever its frequency, means directly to reduce the associated loads and thus lighter structural reinforcement. The Oscillatory Failure Case (OFC) is the name given to a class of failures in the actuator servo loop that cause undesired oscillations of the control surface. OFC can result from the faulty behaviour of an electronic component or a mechanical failure inside the actuator control loop. Model-based techniques are in-service on Airbus aircraft and have proved their efficiency and maturity (Goupil, 2010). However, different models have to be designed and tested for different actuators (Hydraulic, Electro-Hydrostatic (EHA) and Electro-Backup-Hydrostatic (EBHA)), different control surfaces and different aircraft. For this reason a different approach with a higher level of genericity has been investigated in (Goupil et al., 2016). In particular, a data-driven technique comparing the input and output of the aircraft control surface servo loop has been considered. The current Airbus servo loop principle is shown in Fig. 1. It consists of a dual channel scheme where the so-called COM (command) channel is dedicated mainly to the flight control law computation and to the control surface servo-loop. The so-called MON channel (monitoring) is primarily dedicated to the monitoring of all Electronic Flight Control System (EFCS) components. The OFC effects can be simulated through the injection of a periodic signal at two specific points of the control loop: the ANalogical Output(ANO) and the ANalogical Input(ANI) (see Fig. 1). The OFC can appear at the output of the actuator servo loop (the control surface position) as an additive or substitutive signal, also called liquid and solid OFC (Goupil, 2010). The actuator acts as a low-pass filter. Thus, the OFC phenomenon is limited to signals whose frequencies are usually contained in the interval 0-10 Hz.

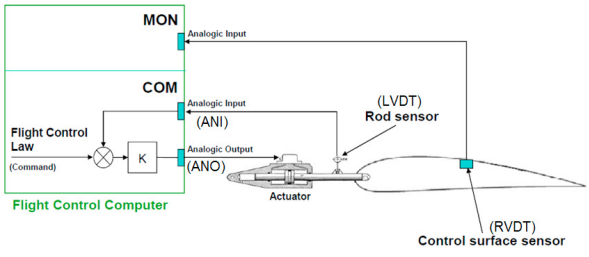


Fig. 1. Airbus control surface servo loop.

The main aim of the present study is to improve the detector proposed in (Goupil et al., 2016) using Support Vector Machines (SVMs). A simple OFC detector has been proposed based on distance and correlation measures. This detector has the structure of a similarity index that combines the information of different parameters. The structure of this similarity index corresponds to a particular parameter combination. Indeed, the detector accuracy can


10.1016/j.ifacol.2017.08.2353

be improved by identifying the optimal detector structure (characterized by an optimal parameter combination) and the SVMs are well adapted to this purpose.
The paper is organized as follows: Section 2 presents the latest state-of-the-art OFC detection methods based on analytical redundancy and the recent data-driven method based on distance and correlation measures. Section 3 details the proposed data-driven OFC detection method and introduces the way in which it can be improved. Section 4 focuses on SVMs and its application to OFC detection. Finally, sections 5 and 6, are devoted to simulation results and conclusions.

## 2. STATE-OF-THE-ART

The latest OFC detection strategy used on board Airbus aircraft (for example the modern A380 and A350) is based on analytical redundancy. The principle of this detection consists of comparing the actual functioning of the control surface with an ideal one expected in the absence of failure. The comparison can be summarized in two steps: residual generation and residual evaluation. The residual is built by comparing the real position of the control surface delivered by a sensor with an estimated position produced by the actuator model (where the input is the command produced by the flight control law). The estimated position $\hat{p}$ is the result of a discrete integration of the estimated velocity $\dot{\hat{p}}$ expressed by the following equation

$$\dot{\hat{p}}(t) = V_0(t)\sqrt{\frac{\Delta P - \frac{F_{\text{aero}}+F_{\text{damp}}}{S}}{\Delta P_{\text{ref}}}} \tag{1}$$

where $\Delta P$ is the hydraulic pressure delivered in input of the actuator, $F_{\text{aero}}$ represents the aerodynamic forces, $F_{\text{damp}} = k_d\dot{p}^2$ is the servo control load of the adjacent actuator in damping mode [1], $S$ is the surface area of the piston, $\Delta P_{\text{ref}}$ is the differential pressure corresponding to the maximum rod speed and $V_0(t)$ is the maximal speed of one actuator without any load. The parameter vector $\boldsymbol{\theta} = (\Delta P, F_{\text{aero}}, k_d)^T$ is assumed to be constant (fixed to its most probable value) to reduce the computational complexity of the OFC detection algorithm or can be estimated by means of an extended Kalman filter to reduce modelling errors (Simon, 2011) (Zolghadri et al., 2014). The OFC detection is classically achieved by counting successive and alternate crossings of the residual around a given threshold (for a liquid OFC) or around the opposite of the estimated position (for a solid OFC). Analytical redundancy approaches have shown to provide a good trade-off between robustness, performance and computational cost in many test campaigns (Goupil, 2011)(Goupil, 2010) and are currently used on in-service aircraft. However, the model-based approaches demand a long test and calibration period and have to be adapted each time the system changes. In this context, it is also important to consider methods that provide acceptable results without making any hypothesis on the system. Data-driven approaches are interesting since they generally offer more genericity. The idea of a data-driven approach is to exploit the correlation between the order sent to the actuator and the position of the aerodynamic surface, which is generally large in the fault-free case. In fact, in the fault-free case, the position of the aerodynamic surface should show only small attenuation (or amplification) and delay compared to the order because of actuator dynamics. In other words, the technical problem to be addressed is the detection of an additive or substitutive periodic signal affecting the aerodynamic surface position using a specific comparison with the order. From a general point of view, the techniques that we will present in this study can be applied to detect a fault in a generic system whose input is fault-free and whose output is similar (in both time and frequency domains) to the input in the absence of a fault. A very simple way to define a fault detection technique is then to directly compare the system input and output.
The OFC detection constraints [2] have motivated the authors to consider a simple approach based on distance and correlation measures (Goupil et al., 2016) (Urbano, 2017). In particular, a novel similarity index was proposed as test statistic for OFC detection. This index is evaluating the similarity between the command and the position of the actuator servo loop by means of a fault probability. The greater its value, the higher the probability of fault. More precisely, this index is simply a statistic that quantifies the dissimilarity between command and position combining the effect of three main parameters of two signals acquired in a given observation window:

i) the correlation coefficient $\rho$
ii) the Euclidean distance $d_E$
iii) an amplification factor $a$ based on the ratio of the estimated curvilinear distances of the two signals

The choice of these parameters and the choice of a specific combination of them is analysed in the next section.

## 3. SIMILARITY MEASURE

The main purpose of defining a similarity measure between two signals is to compare these two signals by means of a single number which quantifies their resemblance. In the signal processing community, one of the most common similarity measures is the correlation coefficient. The correlation coefficient embodies the classical geometrical concept of similarity, in the sense that it provides a score between -1 and 1 to the statistical relationships between two signals (in particular 1 means full agreement, 0 full disagreement, -1 phase opposition). On the other hand, more classical distances can be used to quantify the proximity between the values of two time series. An interesting collection of the historically proposed metrics can be found in (Deza and Deza, 2009).
The properties of the correlation coefficient and of more sophisticated distances can be interesting in many practical applications. For the special case of OFC detection, observing command and position signals in the faulty and fault-free cases, we can see that these signals are very similar in shape and have close amplitudes in the fault-free case. Conversely, in the presence of a fault, there is a reduced correlation, an augmented distance and an increased signal complexity [3]. Based on this observation,

[1] In case of a dual active/passive scheme.

[2] For example, the required detection time, on board sampling time and computational capacity limitations make difficult to achieve a good frequency resolution.

[3] In the sense of having more peaks, valleys and features or in other terms having a bigger curvilinear distance (Batista et al., 2014).

it make sense to consider the parameters $\rho$, $d_E$ and $a$ as test statistics for OFC detection.
The parameter $\rho$ is the maximal Pearson correlation coefficient between the command $x$ and the control surface position $y$

$$\rho_\tau = \rho\left(x, y; \widehat{\tau}\right)$$

where the delay $\tau$ between command and position is estimated as [4]

$$\widehat{\tau} = \arg\max_{\tau} \left\{\rho\left(x, y; \tau\right)\right\}$$

The information about $\tau$ can be used to realign the measured position signal to the command signal and compute a more robust version of the Euclidean distance

$$d_{E_\tau} = \sqrt{\sum_{i=1}^{N} (x_i - y_{i-\tau})^2}$$

and of the amplification factor $a_\tau$, where $a_\tau$ is build as the ratio between the maximum and the minimum of the estimated curvilinear distance of $x$ and $y$ as

$$a_\tau = \frac{\max(e_x, e_{y_\tau})}{\min(e_x, e_{y_\tau})} \geq 1$$

where $e_x$ is an estimation of the curvilinear distance of $x$ in the observation window. For example, we have considered

$$e_x = \sum_{i=1}^{N-1} \sqrt{(x_i - x_{i+1})^2 + T_s^2}$$

where $T_s$ is the sampling time. We can observe that $a_\tau$ is always greater than 1 and it increases if the curvilinear distance of one of the two signals differs substantially from the other.
A very simple OFC detector might be based on one of these parameters or on a particular combination of them. Building a detector monitoring one or multiple attributes in an instance in order to characterize the fault free neighbour is quite common in the literature (Chandola et al., 2009) (Pimentel et al., 2014).
It is important to observe that the parameters $\rho$, $d_E$ and $a$ have some specific limitations

i) if the command is constant, the position is also constant. In this case the correlation coefficient can have a small value since we are considering two uncorrelated measurement noises. Moreover, the correlation coefficient is not always a reliable discriminating factor, in the sense that there is a wide range of signal pairs with the same correlation coefficient.
ii) the difference between the values of the measured position and the command can reach a few degrees for fast manoeuvres or large deflections of the control surface. As a consequence, the Euclidean distance can locally increase in the fault free case, which is not desired.
iii) the ratio of the curvilinear distance in the faulty case depends (as the Euclidean distance) on the OFC amplitude and frequency. More precisely, the lower the OFC amplitude and frequency, the lower the increase in the curvilinear distance of the measured position. Thus, this ratio could not differ substantially from 1 even in the faulty case. It is possible to argue the same conclusion applies to the Euclidean distance, even if this distance is generally more sensitive to the failure (due to the fact that it is proportional to the energy of the faulty signal).

As a conclusion, considering the parameters $\rho$, $d_E$ and $a$ separately can lead to poor performances in terms of minimal detectable OFC amplitude. The intuitive idea behind the proposed statistic is that the information supplied by these indexes is somehow complementary and thus a combination of them can enable the definition of a more reliable statistic. In a previous study, we investigated a particular combination of $\rho$, $d_E$ and $a$ for OFC detection. The chosen dissimilarity measure results from an incremental approach inspired by ideas in the literature (Batista et al., 2011)(Batista et al., 2014)(Chouakria and Nagabhushan, 2007)(Frambourg, 2013).
In (Goupil et al., 2016) and (Urbano, 2017) two particular combinations of $\rho_\tau$, $a_\tau$ and $d_{E_\tau}$ were proposed, i.e.,

$$D_1^*\left(x, y\right) = \left(1 - \rho_\tau\right) a_\tau d_{E_\tau}$$

and [5]

$$D_2^*\left(x, y\right) = \left[\ln\left(2 - \rho_\tau\right) + \ln\left(a_\tau\right)\right] d_{E_\tau}$$

In a first test campaign conducted on simulator [6], we have verified that the similarity measures $D_1^*$ and $D_2^*$ are compliant with the OFC detection constraints [7] and we have shown that they represent a promising alternative to the model-based approaches. For further details see (Goupil et al., 2016) or (Urbano, 2017). However, the proposed statistic only correspond to a particular combination of the chosen attributes $\rho$, $d_E$, $a$. This paper goes a step further by conducting a more general analysis of the OFC phenomenon in the feature space represented by the chosen parameters, in order to evaluate the best statistic that we can consider for OFC detection.

### *3.1 A different decision surface*

As already mentioned, the measure $D^*$ defines a specific decision surface in the feature space $(d_{E_\tau}, \rho_\tau, a_\tau) \in ([0, \infty], [-1, 1], [1, \infty])$. For the considered attributes, we expect the fault free case to be in the neighbour of the point $(d_{E_\tau}, \rho_\tau, a_\tau) = (0, 1, 1)$, or equivalently near the origin of the transformed space

$$\mathbf{\Sigma} = (d_{E_\tau}, 1 - \rho_\tau, a_\tau - 1)$$

The statistics $D_1^*$ and $D_2^*$ consider specific decision surfaces (detectors) in the space spanned by the vectors $\mathbf{\Sigma}$, that are described respectively by the Cartesian equations $xyz = \gamma^3$ and $[\log(x) + \log(y)]z = \mu$ (where $\gamma$ and $\mu$ are detection thresholds). A priori, the number of possible decision surfaces that we can consider is infinite. However,

[4] We can consider a discrete variation of $\tau$ ($\tau = \frac{1}{Fs}, \frac{2}{Fs}, \frac{3}{Fs}, \ldots, \frac{N_{max}}{Fs}$) to simplify the computation.

[5] Considering a logarithmic scale for the first two parameters

[6] ADDSAFE simulator delivered by Airbus in the corresponding European project (Goupil and Marcos, 2014).

[7] In terms of detection time, performance and computational time

knowing that the ideal fault free case is in the origin of the space $\mathbf{\Sigma}$ and that in the faulty case all the three factors have to be at the same time higher than a certain threshold, the detectors $D_1^*$ and $D_2^*$ are two possible choices. The next section proposes a strategy for optimizing the decision surface for OFC detection using the theory of SVMs (applied to simulated data). The purpose is to better characterize the OFC phenomenon in the $\mathbf{\Sigma}$ space and to identify the most suitable decision function for OFC detection.

## 4. SUPPORT VECTOR MACHINES

Providing a sufficiently representative database on the OFC phenomenon, a classification technique such as the two-class SVMs is a suitable way to identify the maximum margin hyperplane that divides faulty and fault-free classes. The main principle of SVMs is separating datasets into two classes according to a decision boundary (binary classifier). This boundary is the hyperplane that has the maximum distance between the so-called support vectors in each class. If we call $\mathbf{w}$ the normal vector to the separating hyperplane and $b$ the offset of the hyperplane from the origin along the normal vector, any hyperplane can be written as the set of points $\mathbf{x}$ satisfying:

$$\mathbf{w}^T\mathbf{x} + b = \mathbf{0}$$

If the training data are linearly separable, we can select the optimal hyperplane that separate the data by maximizing the minimum distance between this hyperplane and the training samples (the samples associated with this minimum distance are the support vectors). The decision function associated with the hyperplane is:

$$f(\mathbf{x}) = \text{sgn}(\mathbf{w}^T\mathbf{x} + b)$$

where $\text{sgn}(x)$ is the sign function of $x$. The method was originally thought only for linearly separable data. However, it was then extended to non-linearly separable data using the feature space technique, also called "kernel trick". The idea of this technique is to map the initial data in a higher dimensional space in which the transformed data are linearly separable. This higher dimensional space is generally identified using kernel methods. The kernels enable one to operate in high-dimensional feature spaces without ever computing the coordinates of the data in these spaces, but rather by simply computing the inner products between two transformed vectors $\mathbf{x}_i$ and $\mathbf{x}_j$ using the following function

$$K(\mathbf{x}_i, \mathbf{x}_j) = \Phi(\mathbf{x}_i)^T\Phi(\mathbf{x}_j) \quad \Phi : \mathbb{R}^n \to \mathbb{R}^m$$

where $m > n$. Computing $K(\mathbf{x}_i, \mathbf{x}_j)$ directly generally requires much smaller computational cost than computing the inner product between $\Phi(\mathbf{x}_i)$ and $\Phi(\mathbf{x}_j)$. Determining the appropriate kernel function to use is important since it defines the feature space where the training data is classified. Three of the most commonly used kernels are recalled below

i) Linear: $K(\mathbf{x}_1, \mathbf{x}_2) = \mathbf{x}_1^T\mathbf{x}_2$

ii) Gaussian: $K(\mathbf{x}_1, \mathbf{x}_2) = e^{-\frac{||\mathbf{x}_1 - \mathbf{x}_2||^2}{2\sigma^2}}$

iii) Polynomial: $K(\mathbf{x}_1, \mathbf{x}_2) = (\mathbf{x}_1^T\mathbf{x}_2)^p$

Fig. 2 explains the principle of kernel methods. It can be shown that the decision function using the feature space technique can be computed as

$$\begin{aligned} f(\mathbf{x}) &= \text{sgn}(\mathbf{w}^T\Phi(\mathbf{x}) + b) \\ &= \text{sgn}\left(\left[\sum_{n=1}^{N} \alpha_n y_n K(\mathbf{s}_n, \mathbf{x}_n)\right] + b\right) \end{aligned}$$

where the $\boldsymbol{s} = (\mathbf{s}_1, \ldots, \mathbf{s}_n)$ are the support vectors, $(y_1, \ldots, y_n)$ the related classes with $y_i = \pm 1$ and $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$ the Lagrangian multipliers related to the dual optimization problem (Boyd and Vandenberghe, 2004). In fact, the computation of the decision function requires the resolution of a particular optimization problem in order to identify the support vectors $\boldsymbol{s}$, the bias $b$ and the weight vector $\boldsymbol{\alpha}$. The shape of the computed hyperplane (that we would like to use to classify new data) also depends on the considered training set and the tuning parameters of the optimization algorithm. In order to guarantee the convergence and to have an additional degree of freedom in the optimization process, a soft margin parameter $C$ is usually considered. The soft-margin SVM classifier is a generalization of the hard-margin SVM classifier defined to guarantee that the SVM can converge to a decision function even if the data are not linearly separable. The parameter $C$ determines the trade-off between increasing the margin size and ensuring that the support vector lies on the correct side of the margin. It can be shown that considering a soft margin parameter $C$ with the kernel trick is equivalent to defining a linear constraint in the optimization process of the Lagrangian multipliers ($0 \leq \alpha_i \leq C$). Since the optimization problem is a quadratic function of $\boldsymbol{\alpha}$ subject to linear constraints, it is efficiently solvable by well known quadratic programming algorithms (Platt et al., 1998). For further details on the SVM mathematical background and the related optimization problem see (Cristianini and Shawe-Taylor, 2000) and Appx. B. For the purpose of this study we have decided to train the decision function on simulated data from a specific desktop simulator that enables us to generate the command and position signals in the fault and fault-free case. The next section presents some simulation results associated with this study.

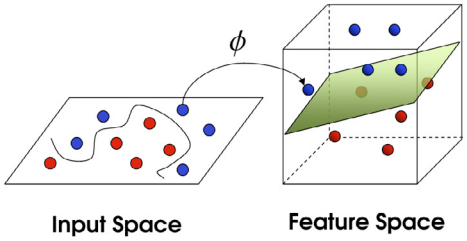


Fig. 2. The kernel trick

## 5. RESULTS

The ADDSAFE simulator (Goupil and Marcos, 2014) can simulate the command and position of the actuator servo loop of an Airbus aircraft (under different side-stick inputs). Moreover, we can simulate the OFC consequences adding a sinusoidal signal in the loop at different points. For this paper, two scenarios (corresponding to the training and test sets) of a few minutes of flight have been generated to tune and test the SVM classifier for different OFC frequencies[8] and amplitudes. The two classes

[8] [0.5,1,2,3,5,7,10] Hz

have been represented with the same proportion in both datasets [9]. The attributes $(d_{E_\tau}, \rho_\tau, a_\tau)$ associated with the two scenarios were computed using a sliding window and the related training sets and test sets were stored to be processed with SVMs (see also Fig. A.1). The SVM method was then performed in two steps: training and performance evaluation [10].

It is worth mentioning that this case study is limited to a sliding window of 1sec (at 100Hz) based on the half period of the smallest OFC frequency considered i.e., 0.5Hz. For a higher sensitivity to the failure (that is frequency dependent), different window sizes should be considered as in (Goupil et al., 2016) and different SVM classifier should be trained.

### 5.1 Training

The training of the SVM classifier was accomplished with the following steps

i) A soft-margin SVM with a Gaussian kernel was considered [11].
ii) A validation step to evaluate the best parameters $C$ and $\sigma$ was considered to avoid data over-fitting (Hsu et al., 2003). A 10-fold cross-validation method was finally used on the training set for the SVM tuning. The score considered in the grid search for $C$ and $\sigma$ was the accuracy [12]. A first guess for the parameter $\sigma$ was obtained using the Jaakkola heuristic (Jaakkola et al., 1999). The optimization process was then conducted for a wide range of values $(C, \sigma)$ leading to optimal performance in terms of accuracy. In this range, we chose $C = 1000$ and $\sigma = \sigma_{\text{Jaakkola}}$.
iii) Considering the optimal parameters $C$ and $\sigma$, the decision function (weight vector $\boldsymbol{\alpha}$) was computed using all the training set.

In Fig. 3 we can see the shape of the decision function computed using the training set (in blue).

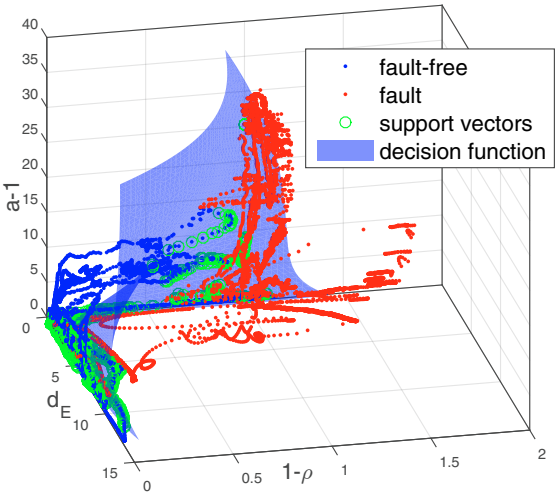


Fig. 3. SVM hyperplane in the space Σ (computed using the training set and the cross-validated values of C and $\sigma$)

It is interesting to observe that as expected the fault-free behaviour of the system is localized close to the origin and the axes of the space $\boldsymbol{\Sigma}$. On the other hand, the points describing the faulty behaviour, as a function of the OFC frequency and amplitude are more or less far from the origin. In particular, it can be shown that $a_\tau$ is more sensitive to the OFC frequency and $d_{E_\tau}$ to the OFC amplitude. In order to be able to compare the tuned decision function to the one defined by $D_1^*$, the training set is used also to evaluate the best threshold (in terms of accuracy) for $D_1^*$.

### 5.2 Performance evaluation

The performance of the computed decision function was evaluated on the test set. In Fig. 4, we can see the computed hyperplane obtained with the test set (compared also to the $D_1^*$ surface tuned on the training set). Table 1 displays the results in terms of accuracy, precision [13] and recall [14] for the two presented methods.

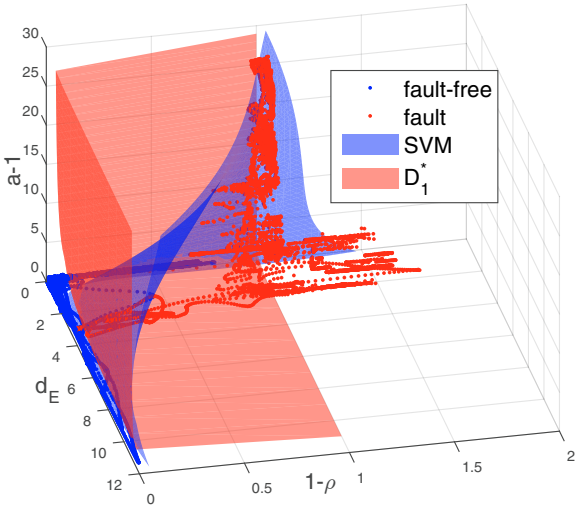


Fig. 4. SVM hyperplane obtained with the test set (compared to the surface defined by $D_1^*$)

Table 1. Accuracy of the proposed method on the test set (compared to the similarity index $D_1^*$)

| Decision function | Accuracy | Precision | Recall |
|---|---|---|---|
| SVMs | 0.96 | 0.95 | 0.97 |
| $D_1^*$ | 0.92 | 0.87 | 0.99 |
| $D_2^*$ | 0.88 | 0.85 | 0.94 |

We can observe that, for the considered test set, the SVM classifier has better performance in terms of accuracy than $D_1^*$. However, one can argue that the values corresponding to the SVM classifier and $D_1^*$ do not differ too much. In this case, it is important to mention that the $D_1^*$ maximal accuracy drops to 0.85 on the training set, while the SVM performances is similar between the training and test phases. It is interesting to notice here that an on-board algorithm for OFC detection has higher requirements in terms of false alarm rate (compared to the obtained results). However, considering filtering and consolidation steps in the detection algorithm, it is possible to obtain results that

[9] Half of the data is fault-free and the other half is faulty
[10] For the computation of the maximum margin hyperplane in the space $\boldsymbol{\Sigma}$ the MATLAB STATISTIC TOOLBOX was used.
[11] A Gaussian Kernel is often considered as a reasonable first choice (Hsu et al., 2003).
[12] Accuracy $= \frac{\text{Number of True Positives+Number of True Negatives}}{\text{Total Number of Samples}}$
[13] Precision $= \frac{\text{Number of True Positives}}{\text{Number of True Positives+Number of False Positives}}$
[14] Recall $= \frac{\text{Number of True Positives}}{\text{Number of True Positives+Number of False Negatives}}$

are compliant with the specifications [15]. As a consequence of the impossibility to define a fully representative test set for OFC detection, it is difficult to generalize the previous results. However, they still represent a good reference for further studies.

## 6. CONCLUSION

The aim of the present study was to show how we can identify a test statistic (decision function) for OFC detection using the theory of Support Vector Machines (SVMs) based on distance and correlation measures. The proposed method achieved higher detection performance in terms of accuracy than a previously proposed statistic $D_1^*$ (that proved to be a promising alternative to the state-of-the-art methods (Goupil et al., 2016)). The proposed approach has a computational cost that is similar to $D_1^*$ (because the on-board application only require the separating hyperplane obtained after the training phase of the SVM classifier). However, it relies on simulated data and thus on a system model. Further studies can be carried out with the dual objective of reducing the minimal OFC detectable amplitude and avoiding the use of a system model. In this case, a One-Class Support Vector Machine (OC-SVM) technique might be used directly on flight data in order to define a suitable test statistic.

## ACKNOWLEDGEMENTS

The authors would like to thank Prof. Corinne Mailhes for her support and advice.

## REFERENCES

Batista, G.E., Keogh, E.J., Tataw, O.M., and de Souza, V.M. (2014). CID: an efficient complexity-invariant distance for time series. *Data Mining and Knowledge Discovery*, 28(3), 634–669.

Batista, G.E., Wang, X., and Keogh, E.J. (2011). A Complexity-Invariant Distance Measure for Time Series. In *SDM*, volume 11, 699–710. SIAM.

Boyd, S. and Vandenberghe, L. (2004). *Convex optimization*. Cambridge university press.

Chandola, V., Banerjee, A., and Kumar, V. (2009). Anomaly detection: A survey. *ACM computing surveys (CSUR)*, 41(3), 15.

Chouakria, A.D. and Nagabhushan, P.N. (2007). Adaptive dissimilarity index for measuring time series proximity. *Advances in Data Analysis and Classification*, 1(1), 5–21.

Cristianini, N. and Shawe-Taylor, J. (2000). *An introduction to support vector machines and other kernel-based learning methods*. Cambridge university press.

Deza, M.M. and Deza, E. (2009). Encyclopedia of distances. In *Encyclopedia of Distances*, 1–583. Springer.

Frambourg, C. (2013). *Apprentissage d'appariements pour la discrimination de séries temporelles*. Ph.D. thesis, Université de Grenoble.

Goupil, P. (2010). Oscillatory failure case detection in the A380 electrical flight control system by analytical redundancy. *Control Engineering Practice*, 18(9), 1110–1119.

Goupil, P. (2011). Airbus state of the art and practices on FDI and FTC in flight control system. *Control Engineering Practice*, 19(6), 524–539.

Goupil, P. and Marcos, A. (2014). The European ADDSAFE project: Industrial and academic efforts towards advanced fault diagnosis. *Control Engineering Practice*, 31, 109–125.

Goupil, P., Urbano, S., and Tourneret, J.Y. (2016). A Data-Driven Approach to Detect Faults in the Airbus Flight Control System. *20th IFAC Symposium on Automatic Control in Aerospace, IFAC-PapersOnLine*, 49(17), 52–57.

Hsu, C.W., Chang, C.C., Lin, C.J., et al. (2003). *A practical guide to support vector classification*. Technical report, Department of Computer Science, National Taiwan University.

Jaakkola, T.S., Diekhans, M., and Haussler, D. (1999). Using the Fisher kernel method to detect remote protein homologies. In *ISMB*, volume 99, 149–158.

Pimentel, M.A., Clifton, D.A., Clifton, L., and Tarassenko, L. (2014). A review of novelty detection. *Signal Processing*, 99, 215–249.

Platt, J. et al. (1998). Sequential minimal optimization: A fast algorithm for training support vector machines.

Simon, P. (2011). *Détection robuste et précoce des pannes oscillatoires dans les systèmes de commandes de vol*. Ph.D. thesis, Bordeaux 1.

Urbano, S. (2017). Early and Robust Detection of Oscillatory Failure Cases (OFC) in the Flight Control System: A Data Driven Technique. In *55th AIAA Aerospace Sciences Meeting*, 2017.

Zolghadri, A., Henry, D., Cieslak, J., Efimov, D., and Goupil, P. (2014). *Fault Diagnosis and Fault-tolerant Control and Guidance for Aerospace Vehicles*. Springer.

## Appendix A. DATASET GENERATION

The training set (and test set) generation is accomplished in three steps, as shown in Fig. A.1.

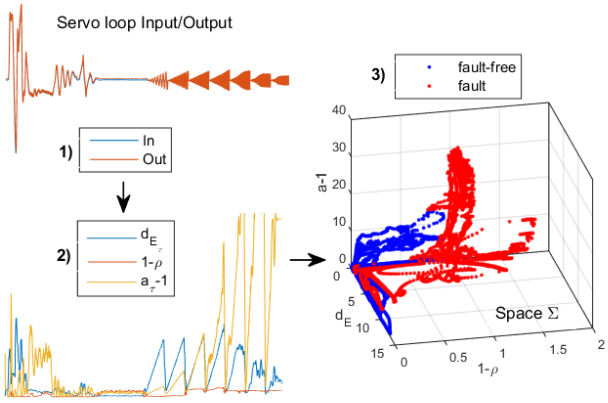


Fig. A.1. Training set generation process.

## Appendix B. DUAL PROBLEM

A soft-margin SVM classifier can be trained by solving a quadratic programming problem that can be expressed in the dual form as (Cristianini and Shawe-Taylor, 2000)

$$\max_{\alpha} \left[ \sum_{n=1}^{N} \alpha_n - \frac{1}{2} \sum_{n=1}^{N} \sum_{m=1}^{N} y_n y_m K(\mathbf{x}_n, \mathbf{x}_m) \alpha_n \alpha_m \right]$$

$$\text{subject to} \quad 0 \leq \alpha_i \leq C, \quad \sum_{n=1}^{N} y_n \alpha_n = 0$$

The sequential minimal optimization (SMO) algorithm (Platt et al., 1998) can be used to solve this problem.

[15] It is the case for the results presented in (Goupil et al., 2016)).