Development of a hybrid PHM system for flight control actuators fusing data analytics with physical knowledge

Development of a hybrid PHM system for flight control actuators fusing data analytics with physical knowledge

Sylvain Autin[1], Corentin Boitard[1], Nicolas Morizet[2], Andrea De Martin[3], Antonio Carlo Bertolino[3], Roberto Guida[3] and Giovanni Jacazio[3]

[1] *Safran Electronics and Defense, 13, Avenue de l'Eguillette, - BP7186 Saint-Ouen l'Aumône, France*
sylvain.autin@actuation.safrangroup.com
corentin.boitard@actuation.safrangroup.com

[2] *Forvis Mazars, 45 rue Kléber 92300 Levallois-Perret, France*
nicolas.morizet@forvismazars.com

[3]*Department of Mechanical and Aerospace Engineering, Politecnico di Torino, Torino, 10129, Italy*
andrea.demartin@polito.it
antonio.bertolino@polito.it
roberto.guida@polito.it
giovanni.jacazio@formerfaculty.polito.it

ABSTRACT

It is well known that the development of PHM systems can be approached with a large variety of techniques. In general, the different prognostics techniques are classified as "data driven", or "physics model based", depending on whether the assessment of a system health status and its evolution with time is performed with an intelligent statistical analysis of present and historical data (data driven approach), or the system actual operating status derived from different sensors is compared with the expected system status in the same operating conditions (model based approach). When a large database of historical data is available, data driven techniques can be accurate and do not require a knowledge of the underlying physics as for the case of model-based techniques. On the other hand, a data driven approach detects only anticipated faults, while a physics-based model approach can also detect unanticipated faults, that never occurred in the past. Flight control actuators of aircraft in revenue service are a typical application in which a large historical database can be available from maintenance, repair and overhaul departments, still the prediction of their health status may fall short of the necessary accuracy without a model description of their physics. For this reason, Safran Actuation, a leading manufacturer of flight control actuators, is conducting an extensive R&D work, together with Politecnico di Torino and Forvis Mazars, aimed at developing an effective PHM system with specific reference to the spoiler actuators of the Airbus A320. This use case is of a particular interest due to the very large number of this type of aircraft in service, to their expected continued use in the years to come and to the number of actuators per aircraft. With reference to this use case, this paper shows how an intelligent fusion of data analytics with physical knowledge can be a multiplying factor in improving accuracy and reliability of the PHM system, with the objective of developing a technological demonstrator for its validation in the lab prior to the implementation of a prototype.

Keywords: PHM, flight control, particle filtering, Echo-State Networks, Physics-Informed Machine Learning

1. INTRODUCTION

Flight control systems, along with their associated servo-actuators, are safety-critical components in aircraft; as such, the occurrence of failure events associated with them is considered a major factor in operational disruptions for both civil and military aviation. The establishment of a robust Prognostics and Health Management (PHM) framework for flight control actuators, could bring substantial technological advancements, benefiting both individual aircraft and entire fleets. The spoiler actuators addressed in this study are Electro-Hydraulic Servo-Actuators (EHSAs), that continue to be by far the preferred technology for primary flight control actuation systems. The successful implementation of this technology could thus improve the aircraft availability, reducing the significant costs related to flight delays, cancellations, re-routing, or in-flight return to base (Sprong et al., 2020). Literature on prognostics for Electro-Hydraulic Servo Actuators (EHSA) is rather limited and mostly focused on a few single faults scenarios. A recent systematic review, (Baldo, De Martin, Jacazio, et al., 2025), found a limited pool of around 30 contributions focusing specifically on flight control actuators. Between these, it is worth mentioning the work by Byington et al. (2004), in which authors presented the possibility of developing a PHM system for the F/A-18 stabilizer focusing on a fully data-driven approach. Guo & Gan (2017), Guo & Sui (2019) proposed the application of Hellinger's distance and verified their results on accelerated fatigue tests on a few structural component of the actuator. Autin et al. (2021) provided the results of an early feasibility study for a PHM system for the primary flight-controls of a wide-body commercial aircraft, while Shahkar & Khorasani (2022) proposed a multidimensional Bayesian methodology specifically tailored to servovalves. More recently, Baldo, De Martin, Terner, et al., (2025) proposed a framework to estimate the RUL of primary flight control EHSAs based only on operational data. The PHM framework herein proposed combines physics-based and data-driven methodologies with engineering knowledge to apply advanced signal processing techniques, in-depth feature extraction processes, diagnostic and prognostic algorithms sporting advanced uncertainty management tools. The end purpose is a computationally efficient PHM suited to be implemented in a dedicated technological demonstrator, posing the basis for future developments for onboard aircraft use. The paper is organized as follows: at first the project organization is introduced and justified (Section 2). The selected case study, the A320 spoiler actuator, is then introduced and analyzed with specific attention given to the characteristics pertaining to the development of a PHM application (Section 3). As the preliminary work was based upon simulation activities, an overview of the simulation tools and process is then provided, with specific care reserved to the measures taken to cover the most prominent uncertainty sources (Section 4). The results of the simulation campaign are then used to inform the feature selection process, conducted combining the results of physics-based and data driven considerations (Section 5). The fault diagnosis process is then outlined (Section 6), with the prognostic routine described in the following section (Section 7). Finally, a few considerations pertaining the deployment of the PHM code to the dedicated technological demonstrator are presented together with a summary of the results.

2. PROJECT ORGANIZATION

The project is organized as depicted in Figure 1 and is roughly divided into three parallel swim lanes. The first pertains both preliminary and simulation activities, the second is dedicated to the design of the PHM system and the third is related to the experimental activities. The program started with the analysis of the case study, with specific attention to the actuator type, architecture, configuration and sensor suite. Then, a Failure Mode Effects and Criticality Analysis (FMECA) analysis tailored for PHM was prepared to identify the most prominent failure modes. As datasets pertaining specific failure modes in controlled test environment were not available, a high-fidelity simulation model was developed leveraging the availability of previously obtained experimental dataset for healthy conditions. Two simulation campaigns were then prepared. The first, defined as the "on-ground" data set, was obtained for a predefined command sequence designed to be run on a test-rig to optimize troubleshooting procedures in presence of a certain number of degradations of fixed severity level. The second, the "in-flight" simulation dataset, was instead obtained considering the possible failure mode evolution as a function of time and of the actuator usage using a set of command and load patterns representative of the actuator behavior during flight. The simulated datasets were then employed to inform the PHM system design phase, which is being pursued with a combination of purely data-driven considerations and physics-based engineering judgment. The main purpose of this approach is to leverage the benefits of the latest machine learning technologies to supplement the results obtainable through the physical knowledge of the system, and conversely to find physical meaning to the results of the purely data driven algorithm. Once trained on the "on-ground sequence", the performance of the diagnostic and of the prognostic algorithms will be verified versus the data provided by the "inflight" simulation database. In parallel with these theoretical and numerical activities, a significant effort is dedicated to the definition of a dedicated test-rig with the main purpose of performing experimental runs to verify the results of both the high-fidelity model and of the PHM routines. The test-rig, which will be equipped with a real-time digital-twin of the actuator under test, will finally act as a technological demonstrator embedding PHM functionalities.

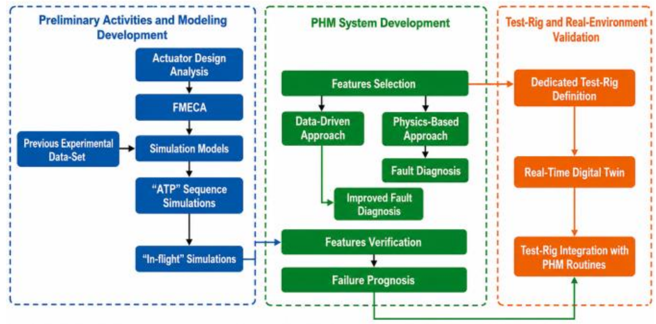

Figure 1. Project organization.

3. CASE STUDY ANALYSIS

The case study of the proposed analysis is the electro-hydraulic spoiler actuator of the Airbus A-320. Its architecture, depicted in Figure 2, is based upon a hydraulic actuator with unbalanced areas, favoring the return stroke, commanded by a two-stage jet-pipe servovalve. A blocking valve is employed to perform safety functions, ensuring the

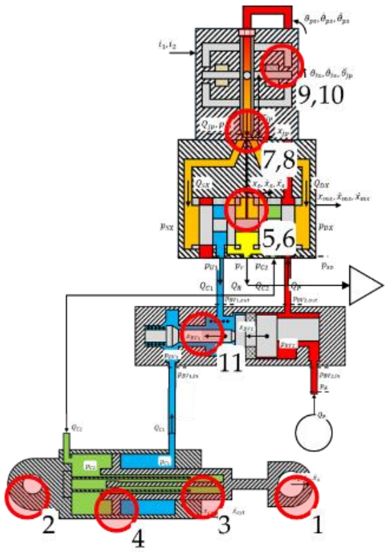

Figure 2. Case-study architecture.

return of the actuator to a fully retracted position under an external load following pressure loss in the supply line or in response to a manual command from the maintainers. The system is position controlled with a proportional control law leveraging the measurement of the piston travel provided by an LVDT (Linear Variable Differential Transducer) sensor integral with the actuator. The actuator does not employ additional control loops, nor additional sensors. This leaves only a handful of signals usable for PHM purposes, namely the actuator position, the actuator command signal and the servovalve currents. A dedicated FMECA analysis was conducted to organize the possible failure modes as a function of their severity, observability, frequency of occurrence and repairability. Only failure modes associated with progressive degradations were considered. The following faults were then identified as the most prominent:

1. Wear in the surface-side rod-end.
2. Wear in the airframe-side rod-end.
3. LVDT short/winding unbalance.
4. Wear of the sealing elements in the EHSA
5. Crack of the servovalve feedback spring. The loss of the feedback spring causes instability conditions in the servovalve, which leads to a critical high-frequency limit cycle oscillation of the actuator.
6. Wear of the feedback spring attachment points.
7. Blockage of the jet-pipe channel
8. Deformation of the jet-pipe channel. A mechanical deformation of the jet-pipe causes a drift in the servovalve behavior.
9. Demagnetization of the torque motor.
10. Short in the servovalve windings due to degradation of the insulating material, mechanical strain (vibration), thermally driven ageing or ingress of foreign particles.
11. Misalignment in the blocking valve.

4. SIMULATION ACTIVITIES

This section presents the high-fidelity dynamic modeling and simulation-based assessment framework developed for the next-generation EHSA of the A320 spoiler actuator. The first section describes the multi-domain actuator model, integrating electrical control, hydraulic network, servovalve dynamics, mechanical actuation, and aerodynamic load interaction, derived from the one proposed in (De Martin et al., 2018). The second subsection outlines the structured simulation campaign designed to evaluate actuator performance over its operational life under both nominal and degraded conditions. A combination of on-ground qualification tests and statistically representative in-flight mission profiles is employed to generate a comprehensive dataset, supporting performance analysis and the development of physics-based health monitoring indicators.

4.1. Dynamic Model of the Electro-Hydraulic Servoactuator

The dynamic model developed for the EHSA depicted in Fig. 1 is structured according to a high-fidelity modular architecture designed to consistently reproduce the interaction among the electrical, hydraulic, mechanical, and aerodynamic domains. The overall model is organized into six main interconnected subsystems (Fig. 3) within the simulation environment: the Command Generator, the Fluid Physical Properties, the Control System, the Hydraulics System, Aerodynamic Surface and Aerodynamic Load Model. The Command Generator block provides the position reference signal generated by the Flight Control Computer (FCC). It represents the functional interface between the flight control system and the actuator.

The Fluid Physical Properties subsystem computes the thermophysical properties of the hydraulic fluid as a function of operating conditions, namely supply temperature and pressure.

The Control System processes the commanded position together with position feedback signals.

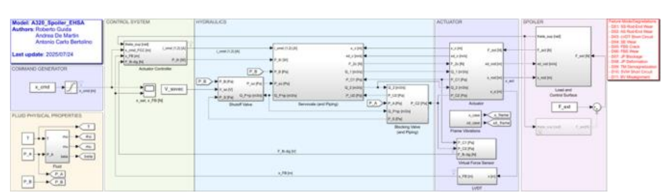

Figure 3. High-fidelity model in MATLAB/Simulink Environment.

It generates the current driving the servovalve and the control signal for the shutoff valve.

The control system incorporates:

- Gain scheduling as a function of the spoiler angular position to ensure consistent performance across the operating envelope,
- An open-loop current mode for ground testing,
- A bias term to compensate for intrinsic servovalve behavior.

The Hydraulics System (Fig. 2) includes the hydraulic piping and three main valves: shutoff valve, servovalve (Li, 2016), and mode (or blocking) valve. The cylinder body dynamics is also included and coupled with the aircraft structural vibrations, capturing the interaction between actuator and airframe.

The control surface model receives rod position and velocity and computes the mechanical interaction through an equivalent spring-damper coupling. The external aerodynamic load acting on the spoiler is also applied. The model provides angular acceleration, velocity, and position, ensuring bidirectional coupling between actuator dynamics and aerodynamic loading.

Figure 4 compares the experimental actuator position with the corresponding model prediction. Overall, the model reproduces the main dynamic features of the measured response with high fidelity, including the ramp transitions, steady-state plateaus, and the sine wave segment. The peak amplitudes and phase of the oscillations are closely matched, indicating that the dominant system dynamics are well captured. The position error remains bounded within approximately ±2.5 %, with the largest deviations occurring during rapid transients and at switching points, where unmodeled nonlinearities effects are likely to be more pronounced. In steady-state regions, the error is near zero with low noise, demonstrating accurate representation.

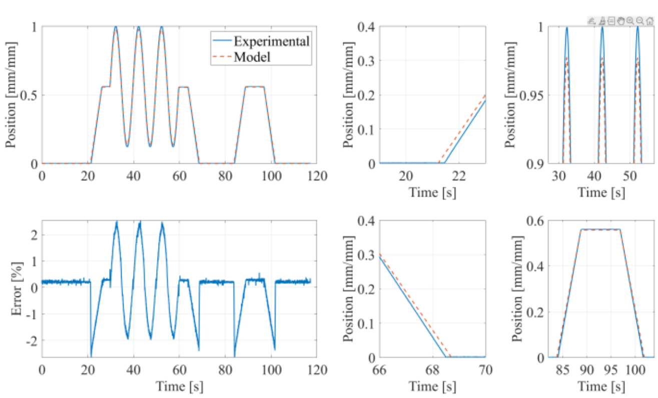

Figure 4. Comparison between experimental and model feedback position of the actuator

4.2. Simulation Campaign and Performance Assessment Methodology

The simulation campaign is designed to comprehensively assess the actuator performance under both nominal and degraded conditions, with particular emphasis on in-flight behavior. The overarching objective is twofold: firstly, to characterize the system response in healthy and faulty states; secondly, to support the development and validation of physics-based health indicators for condition monitoring and prognostics.

4.2.1. Objectives and uncertainty sources

The first objective is a systematic system behavior analysis, aimed at understanding how the actuator responds across its operational envelope, including the presence of degradation mechanisms and fault scenarios. The analysis focuses on closed-loop tracking performance, hydraulic response, force generation capability, and dynamic stability. The second objective concerns the implementation of physics-based health indexes, derived from measurable quantities and physically interpretable internal states. These indicators are evaluated on simulated in-flight data to enable continuous health monitoring, predictive maintenance strategies, and improved mission-level decision-making.

To achieve this, significant effort was dedicated to the identification and characterization of the most prominent uncertainty sources, divided between those associated with intrinsic actuator variability and those due to external operating conditions. Despite the nominal uniformity of considered spoiler actuators, deviations arise from manufacturing tolerances, variability in third-party components, and assembly processes. This variability manifest in performance dispersion within the actuators population. On the other hand, environmental factors, including hydraulic supply and return pressures, fluid temperature and gust occurrence, are treated as uncertain inputs, with temperature modeled as a stochastic variable following a normal distribution around a representative operating condition. Additional parametric uncertainties are introduced through pseudo-random variations in leakage characteristics, servovalve internal features, and friction. These parameters are sampled within admissible ranges that preserve compliance with performance requirements bands as defined by the actuator Acceptance Test Procedures (ATPs). Moreover, the values of such parameters have been tuned to build a population of simulated actuators coherent with the expected distribution of performance indexes. These very same simulated actuators have then been used for the entire simulation campaign.

4.2.2. Simulation Campaign

The simulation campaign is structured to replicate the full operational life cycle of the actuator. Each simulated life cycle consists of two main cycles, each comprising:

- An on-ground test phase;
- An in-flight operational phase.

This structure ensures repeatability and allows progressive degradation to accumulate over successive cycles. During each on-ground phase, a predefined sequence of command inputs is applied. The test profile (Fig. 5) includes:

- Positive and negative position ramps, with and without external load;
- Positive and negative step commands, with and without load;
- A frequency sweep test;
- An open-loop current test with positive, negative, and zero current inputs.

The on-ground test sequence is designed to isolate specific dynamic characteristics of the actuator and control loop. The Very Low-Rate Ramp Position Command (in both directions) evaluates steady-state tracking accuracy and servovalve current demand under quasi-static conditions. The Low-Amplitude Square Wave at Low Frequency allows evaluation of transient performance metrics such as rise time, overshoot, and steady-state error. The Sine Wave Command with Increasing Frequency characterizes the frequency response of the closed-loop system. The Zero Current Command to the Servovalve evaluates actuator free response and residual motion. The Rated Current Command, applied in open-loop control conditions, is applied to assess maximum speed capability and hydraulic performance limits. The subsequent in-flight phase consists of a sequence of spoiler deflections and hinge moment loads statistically sampled from experimental flight data. Each macro-cycle includes 100 simulated flights, ensuring sufficient variability in operating conditions and load histories. The in-flight simulation phases are constructed starting from real operational datasets. Equivalent time histories of position setpoints and hinge moments are generated to reproduce realistic mission profiles within a computationally efficient framework. To reduce simulation time while preserving dynamic fidelity, the mission profile is condensed without altering the statistical and dynamic characteristics of the original data. Command and load sequences are systematically reorganized to maintain physically consistent actuator behavior representative of real flight conditions. A distinctive aspect of the methodology is that degradation factors are not imposed a priori as fixed scenarios. Instead, degradation evolves dynamically according to:

- The selected physics-based degradation models,
- The actual simulated usage history of the actuator (load levels, frequency of actuation, cumulative stress).

This usage-driven degradation approach ensures that wear mechanisms, performance drift, and efficiency losses are intrinsically linked to operational demand, thereby enhancing realism and predictive capability.

The entire simulation campaign is executed via an automated MATLAB script that systematically runs all predefined test combinations across the life-cycle scenarios. Upon completion of each simulation run in the Simulink environment, relevant variables, including commanded position, commanded current, actuator states, pressures, and health indicators, are stored in a structured database.

Within the time histories (Fig. 6), the initial time window corresponds to the on-ground test sequence. The open-loop segment is clearly identifiable by the transition from position command to direct current command input. Subsequent time intervals represent the simulated in-flight phases, where

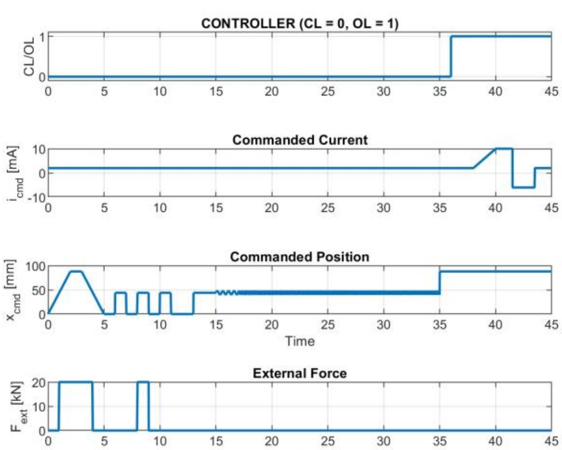

Figure 5. On-Ground Test Sequence.

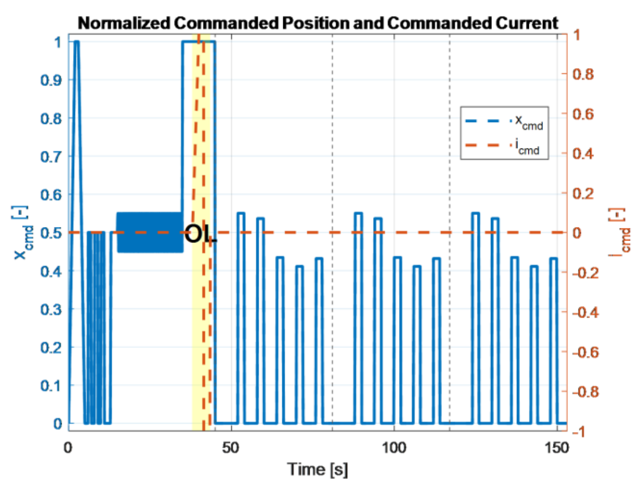

Figure 6 Command Position and current for the equivalent flight simulation.

realistic mission profiles and load distributions are applied. Overall, the simulation campaign provides a controlled yet highly representative virtual test environment. It enables systematic exploration of actuator behavior across its operational life, supporting performance assessment, degradation modeling, and the validation of physics-based health monitoring methodologies under realistic flight conditions.

5. FEATURE SELECTION PROCESS

The simulation campaign results were investigated through two different and complementary approaches. First, a physics-based approach leveraging expert engineering knowledge of the actuator design and behavior was followed. Then, a purely data-driven analysis was conducted and a structured search for the best features combination to achieve high-quality diagnostics was pursued. This section discusses both approaches and then presents the outcome of their combination.

5.1. Physics-based approach

The physics-based approach provided the basis for the feature-selection process and was conducted following the scheme depicted in Figures 7 and 8. Over 70 feature candidates were at first defined in both time and time-frequency domain looking both at the PHM-dedicated command sequence and the in-flight simulation database. Such feature candidates were developed actively searching for noticeable (and often expected) variations in the available signals behavior. For the "on-ground" test sequence, results were then at first pruned following cross-correlation analysis and then optimized ranking their performance according to metrics of correlation with the fault growth and signal-to-noise ratios, leaving 37 features as feasible candidates. This approach, while successful for dedicated test sequences, is not however directly applicable to the in-flight database, as features behavior can be significantly affected by the actuator operations at any particular time instant. As such, ten "context vectors" were defined. Each "context vector" is associated with certain signals, or combination thereof, useful to classify at any moment the operational conditions seen by the actuators (i.e., whether the actuator is extending or retracting, whether any sudden acceleration occur, whether load is present and, if so, which are its characteristics). Each feature candidate was then evaluated for each considered operational condition through correlation with the fault size

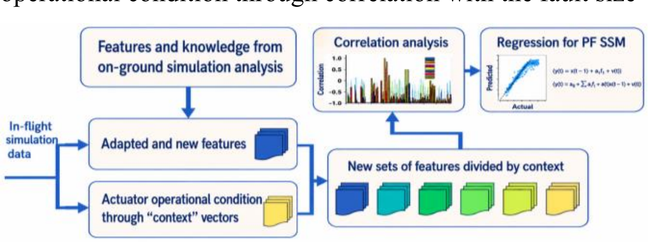

Figure 7 Physics-based feature selection approach

to better understand which combination of context vectors and features worked best for any given failure mode. A regression analysis was then pursued on the best-performing feature for each failure mode, aiming at defining observation models necessary for the definition of physics-based routines for prognosis introduced in Section 7.

5.2. Data driven approach

The 37 original expert features (mainly coming from the time domain) are enriched with an additional set of wavelet features. More specifically, a multilevel wavelet packet decomposition (WPD) (Percival & Walden, 2000) is applied on time series derived from only two signals available in flight (the actuator feedback position and the servovalve currents); this approach was adopted to better observe the occurrence of failure modes associated with sudden low amplitude changes in the signals behavior (such as high frequency impacts) distinguishing their effect from the measurement noise.

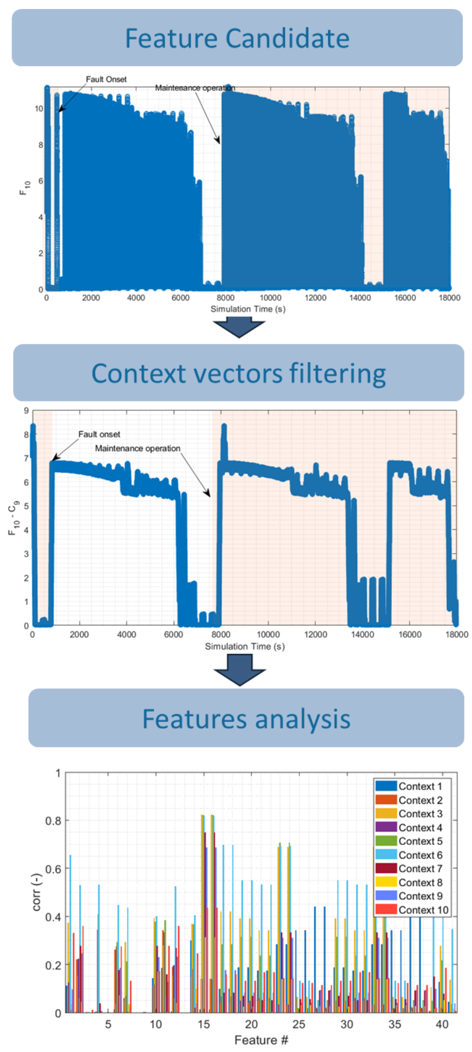

Figure 8 Features engineering for in-flight data.

The relative energy percentage of the terminal nodes of the wavelet packet tree is computed, leading to 24 wavelet packet energy (WPE) features. Wavelet packets involving frequency analysis above 500 Hz are removed, since such frequencies will not be observable given the current sensor suite.

5.3. Hybrid approach

Physics-based features (PBF) alone are not sufficient to fully address prognostic requirements. In practice, significant interdependencies exist among the extracted indicators, and WPE features by themselves do not provide adequate predictive accuracy for reliable fault prognosis. A promising strategy is therefore to combine physics-based indicators with WPE features to construct a hybrid feature set capable of capturing complementary information. However, this expanded feature space introduces redundancy and increased dimensionality, making feature reduction necessary. Consequently, an appropriate feature selection process must be applied to retain the most informative variables while preserving a high level of prediction accuracy. For fault classification, such operation was performed in conjunction with the algorithm definition and is described in Section 6.

6. FAULT DIAGNOSIS

The objective of this study is to accurately classify the eleven fault modes (stepped degradation simulations) described in Section 3, *Case Study Analysis*. Initial classification experiments were conducted using a Linear Support Vector Machines (SVM), considering fault progression higher than 25%. Linear SVMs were first chosen as a simple tool to first assess the suitability of the selected features for diagnostics; as the proposed features were identified on simulation data only, the choice of a simple classifier finds justification in the attempt to avoid overfitting on non-validated data. Analysis of the resulting confusion matrix (Fig. 9) revealed several systematic misclassifications, particularly between fault pairs which exhibit similar effects on the actuator performance faults (1,2), (4,5), and failures whose effect are negligible on the macroscopic behavior of the monitored system (failure mode 11).

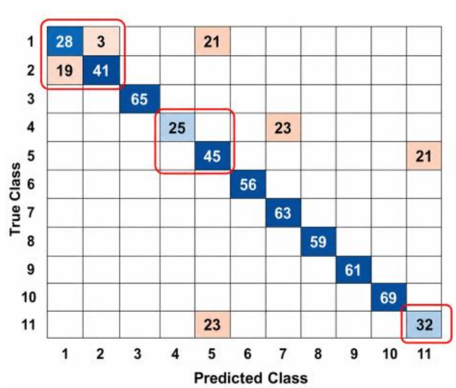

Figure 9 Confusion matrix for LSVM.

Table 1. Prediction accuracy across different feature sets (second experiment).

Feature Set	#Features	Accuracy (%)				
		Train	Test	Class 1	Aggregated Classes	Class 11
PBF	37	91.05	74.55	76.00	*N/A*	57.00
PBF +WPE	61	**98.68**	**92.00**	**89.00**	*N/A*	**89.00**
PBF +WPE (Agg.)	61	**98.91**	**93.45**	**96.00**	**93.00**	**94.00**

Following this study, it was then decided to assess the performance of the features-set through more-advanced classifiers, with the intention of building knowledge towards the application of such techniques on the technological demonstrator. A second experiment was then performed using an augmented dataset combining expert and WPE features, resulting in a total of 61 indicators. For this experiment, the selected classifier was eXtreme Gradient Boosting (XGBoost) (Chen & Guestrin, 2016). Given the large number of hyperparameters associated with the XGBoost algorithm, an efficient hyperparameter optimization strategy was required. The Tree-structured Parzen Estimator (TPE) (Bergstra et al., 2011) was therefore adopted as the Bayesian optimization method to efficiently explore the high-dimensional hyperparameter search space. Using an 80/20 train/test split, results obtained on the test dataset indicate that expert features alone already account for ~74% classification accuracy across the eleven failure modes. The inclusion of wavelet-based features, combined with a finely tuned XGBoost classifier, leads to an additional improvement of ~18% in overall accuracy. Further performance gains can be achieved by aggregating several classes related to the servovalve behavior, which provides an additional 2% improvement and results in a final classification accuracy of ~94%. Comparison results are gathered in Table 1 and corresponding confusion matrices are depicted in Figures 10 and 11.

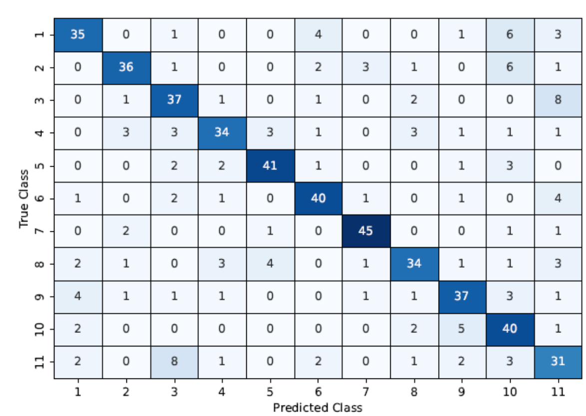

Figure 10. Confusion matrix (PBF only).

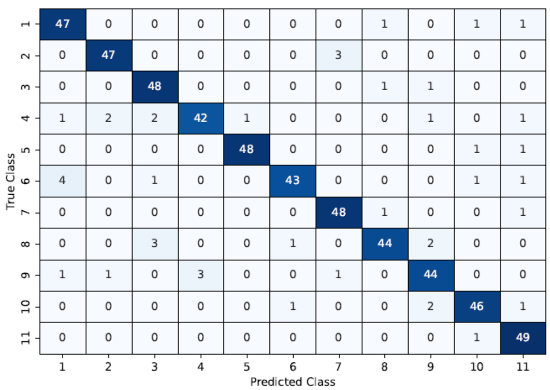

Figure 11. Confusion matrix (PBF+WPE).

In a third experiment, two additional signals available from the test rig (namely actuator chamber pressures 1 and 2) were incorporated into the analysis. WPD was applied to these signals, increasing the total number of extracted features to 73. This expanded feature set introduced the need for feature selection to reduce dimensionality while maintaining high predictive performance. To address this challenge, SHAP-based feature importance was employed. SHAP (SHapley Additive exPlanations) (Lundberg et al., 2017), derived from cooperative game theory through the concept of Shapley values, quantifies the average contribution of each feature to the model's predictions. It provides a consistent and model-agnostic measure of feature influence while accounting for interactions among predictors. Using a 75/15/15 train/validation/test split, results obtained on the test set indicate that expert features alone already account for ~70% classification accuracy across the eleven failure modes. By incorporating chamber pressure signals, wavelet-based features, and employing a fine-tuned XGBoost classifier, the overall accuracy improves by ~20%. By aggregating classes corresponding to servovalve behavior, an additional 2% accuracy is achieved, resulting in a final accuracy of ~92%. Moreover, employing only the top 15 features determined by SHAP-based feature importance markedly improves generalization performance compared to the 'All features' baseline.

Table 2. Prediction accuracy across different feature sets (third experiment).

Feature Set	Configuration	#Features	Accuracy (%)		
			Train	Validation	Test
PBF	All features	37	**91.04**	69.96	70.00
	Top15 SHAP	15 (40%)	90.76	**71.77**	**72.54**
PBF +WPE	All features	73	**98.93**	87.77	87.58
	Top15 SHAP	15 (20%)	98.89	**88.47**	**87.97**
PBF +WPE (Agg.)	All features	73	**99.06**	90.83	90.71
	Top15 SHAP	15 (20%)	98.57	**92.18**	**91.26**

Considering the PBF feature set, only 40% of the original feature set is required to achieve performance comparable to the full set. Regarding the PBF+WPE feature set, selecting the most pertinent 15 features (merely 20% of the original set) still enhances generalization. Similarly, for the last feature set, despite the reduction in class count through the consolidation of five failure modes, the use of a selective subset (again only 20% of the original features) continues to yield performance improvements.

Comparison results are shown in Table 2 and corresponding confusion matrices are depicted in Figures 12, 13, and 14.

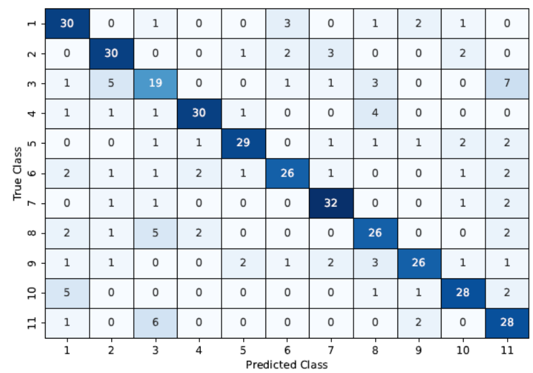

Figure 12. Confusion matrix (PBF, Top15 SHAP).

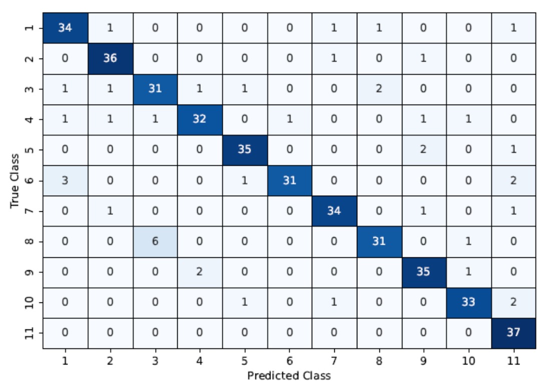

Figure 13. Confusion matrix (PBF + WPE, Top15 SHAP).

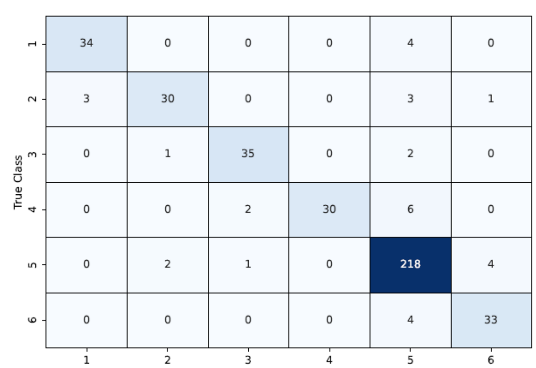

Figure 14. Confusion matrix (PBF + WPE, aggregated classes, Top15 SHAP).

These results demonstrate that focusing on a compact subset of highly informative predictors not only reduces model complexity and mitigates overfitting but also improves robustness and enhances suitability for real-time applications.

7. FAILURE PROGNOSIS

Having assessed the feasibility of the fault diagnosis process relying only on the sensors currently employed on the spoiler actuator, the next natural step is to study the possible application of prognostics techniques to the selected features set. It was decided to use only the data provided by the in-flight simulations, as on-ground tests do not provide a number of samples compatible with a robust reconstruction of the failure growth in time. Similar to the fault diagnosis activity it was decided to start first with a traditional, well-established technique – in this case a well-established declination of the particle filtering approach – and then explore alternative options trying to solve some of the main weaknesses of traditional particle filtering while retaining its capabilities to robustly describe the prediction uncertainty. Traditional Particle Filtering relies on a State-Space Model $f(x)$ (SSM) of the degradation process and a non-linear observation model $h(x_t)$ linking the computed features values y with the estimated fault size x (Arulampalam et al., 2009) through non-Gaussian process and measurement noises $\omega(t)$ and $\nu(t)$.

$$\begin{cases} x_t = f(x_{t-1}) + \omega(t) \\ y_t = h(x_t) + \nu(t) \end{cases} \tag{1}$$

Prognosis is achieved by performing two sequential steps, prediction and filtering. Prediction uses both the knowledge of the previous state estimate and the process model to generate the a priori state probability density function (pdf) estimate for the next time instant. Equations of the SSM are defined a-priori and tuned on-line through an embedded Recursive Least Square (RLS) routine, while filtering is achieved in this study through a Sequential Importance Resampling (SIR) scheme (Orchard & Vachtsevanos, 2009). RUL is computed through Acuña's definition of probability of failure, which allows to rigorously characterize the RUL pdf and the associated Risk of Failure (Acuña & Orchard, 2017, 2018). As stated, traditional particle filtering approaches rely on strong prior assumptions, notably the availability of a well-defined state-space model (SSM) and an explicit fault-size signal. However, such requirements are rarely satisfied in real operational environments, where degradation dynamics are often complex, partially observed, or poorly characterized. This is especially relevant for high performance actuation systems, for which small variations in the mechanical tolerances can have a significant impact on the end effect of the fault progression on the features evolution in time.

To tackle this issue, a possible solution is to integrate Echo State Networks (ESNs) to solve the regression problem and exploit the Particle Filtering framework for long term Prognosis. ESNs (Lukoševičius, 2012) constitute a particular class of Reservoir Computing (RC) framework, derived from the Recurrent Neural Networks (RNNs) theory. ESNs provide a computationally efficient yet expressive framework for Remaining Useful Life (RUL) prediction. Their architecture relies on a fixed, randomly connected recurrent reservoir combined with a linear readout layer. In contrast to many conventional dynamical modeling approaches, only the output weights are trained (typically through linear regression) while the internal reservoir dynamics remain unchanged (Figure 15). This property leads to extremely fast training and low computational cost, while still enabling the modeling of complex temporal dependencies. The motivation for combining ESNs with particle filtering stems from the observation that ESN dynamics can naturally be formulated as an SSM (Singh & Balasubramanian, 2025). In particular, the standard leaky-integrator ESN recursion can be rewritten as a nonlinear discrete-time SSM incorporating both process and measurement noises. Such expression can be written as Equation (1), giving the following meaning to each term:

- $f(x_{t-1}) = (1-\alpha)x_{t-1} + \alpha \tanh(W_X x_{t-1} + W_{in} u_t)$
- $h(x_t) = W_{out} x_t$
- α: leaking rate $\in (0, 1]$
- u_t: input vector at time instant t
- W_{in} , W_X , W_{out}: input weight matrix, reservoir weight matrix, output weight matrix, respectively (Figure 15)

Once trained, the ESN state recursion is interpreted as the nonlinear transition model of a state-space formulation and embedded into a particle filtering scheme to sequentially infer the posterior distribution of the latent reservoir state from observed degradation measurements. The particle ensemble thus provides a real-time estimate of the system health state together with its associated uncertainty. RUL prediction is performed by propagating this filtered state distribution forward through the ESN dynamics until a predefined failure threshold on the predicted health indicator is reached.

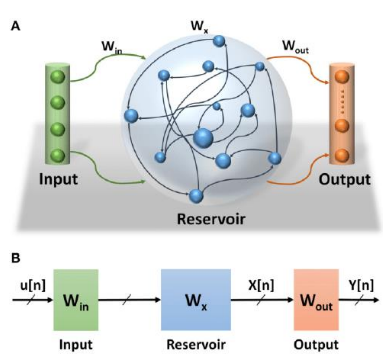

Figure 15. **(A)** An ESN consists of three layers: input layer, reservoir layer, and output layer (the only *trainable* one). **(B)** The ESN abstract structure.

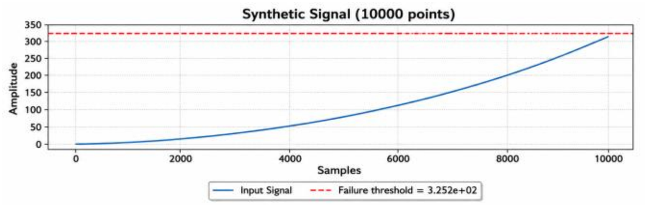

Figure 16. Synthetic input signal (10 000 samples).

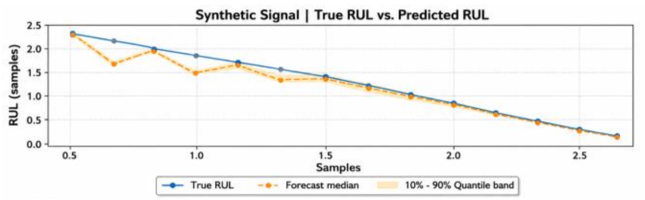

Figure 17. True vs. Predicted RUL (synthetic signal).

The resulting distribution of threshold-crossing times directly yields probabilistic RUL estimates, enabling both point predictions and confidence bounds while preserving the computational efficiency and long-horizon forecasting capability of the ESN. As a first experimental validation of the ESN with particle filtering approach, a synthetic signal is built as a smooth increasing degradation trend with a low-amplitude periodic component and additive stochastic noise (Figure 16). Comparison between true and predicted RUL, alongside the 10%-90% quantile uncertainty band, is depicted in Figure 17. The method quickly converges towards the ground truth and maintain coherence with the simulated End of Life (EoL) until failure. The algorithm was then stressed with simulated fault-to-failure processes using the high-fidelity model. In the most difficult conditions (noisy features, highly non-linear fault growth), the PF/ESN algorithm tend to converge more slowly (Figure 18) to the ground truth; coherently with this, the algorithm uncertainty grows as the PF/ESN framework estimates higher values of measurement and process noises.

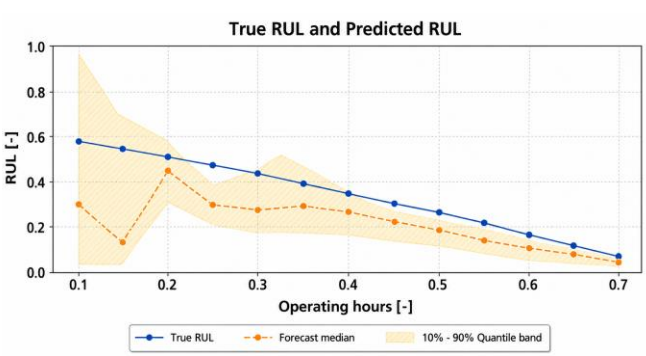

Figure 18. True vs. Predicted RUL (worst result – simulated failure mode on the control valve).

8. TECHNOLOGICAL DEMONSTRATOR

To verify the proposed PHM framework and solve possible implementation challenges a technological demonstrator, in the form of a dedicated test rig, has been prepared. The rig is composed of an electro/hydraulic actuator which exerts a controlled load on the spoiler actuator under test, a complete sensor suite to monitor the experimental activities and a dedicated control rack, depicted in Figure 19, which embeds the PHM routines proposed in this paper. A specific attention was given on how to transform development oriented code to an embedded code with respects to real time constraints. It led to dedicated optimization of the diagnostic code, supported by integration and validation activities to ensure performances stability in presence of possible variations of the rig sampling time.

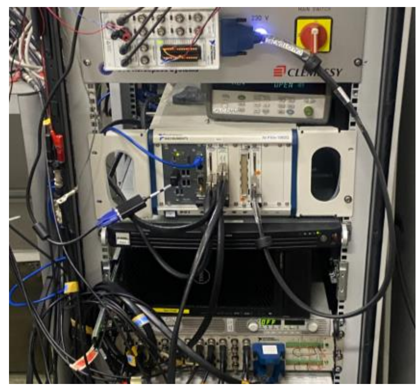

Figure 19. The Control Rack with PHM functionalities.

9. CONCLUSIONS

The activities pursued so far have been effective in demonstrating the feasibility of a PHM system for a legacy flight control actuator operating with an extremely low number of signals relevant to failure prognosis. The main outcome of the project is the clear understanding that the application of a hybrid PHM approach, fusing engineering judgement and physics of failure with advanced data-driven techniques is currently the most promising solution for designing and implementing PHM for systems operating with a minimal number of sensors. This consideration is directly applicable both to the type of algorithms employed for fault diagnosis/failure prognosis and for the definition of the features themselves. With regards to features performances, it emerged that neither a purely data-driven nor a purely physics-based approach were capable of providing optimal results; the performance of physics-based features can be improved by filtering them depending on the contextual operations being performed by the monitored equipment, but efficient diagnostics for failure modes with overlapping (or almost negligible) effects on the main dynamics was not possible without the introduction of data-driven indicators. On the contrary, data-driven only features are heavily dependent on the training set, and if used alone risk providing inconclusive results in the case of unexpected variations in the signals characteristics (variable or irregular sampling frequency). Similar considerations can be transferred to the fault diagnosis and prognostic functions, both of which provided the best outcome when combining expert-features with advanced data-driven techniques. At the time of writing all the presented activities are being subjected formal experimental verification, which is being pursued through a dedicated test-rig. As the proposed routines have been already converted for real-time usage and embedded on the rig. The experimental campaign will focus on the verification of the simulated features behavior, both in term of correlation with the fault size and in terms of uncertainty characterization and on evaluating the robustness and diagnostic/prognostic performances of the proposed PHM framework.

REFERENCES

Acuña, D. E., & Orchard, M. E. (2017). Particle-filtering-based failure prognosis via sigma-points: Application to Lithium-Ion battery State-of-Charge monitoring. *Mechanical Systems and Signal Processing*. https://doi.org/10.1016/j.ymssp.2016.08.029

Acuña, D. E., & Orchard, M. E. (2018). A theoretically rigorous approach to failure prognosis. *Proceedings of the 10th Annual Conference of the Prognostics and Health Management Society 2018 (PHM18), Philadelphia, PA, September 24-27*.

Arulampalam, M. S., Maskell, S., Gordon, N., & Clapp, T. (2009). A Tutorial on Particle Filters for Online Nonlinear/NonGaussian Bayesian Tracking. In *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*. IEEE. https://doi.org/10.1109/9780470544198.ch73

Autin, S., De Martin, A., Jacazio, G., Socheleau, J., & Vachtsevanos, G. J. (2021). Results of a feasibility study of a Prognostic System for Electro-Hydraulic Flight Control Actuators. *International Journal of Prognostics and Health Management*, *12*(3), 1–18. https://doi.org/DOI https://doi.org/10.36001/ijphm.2021.v12i3.2935

Baldo, L., De Martin, A., Jacazio, G., & Sorli, M. (2025). A Systematic Literature Review on PHM Strategies for (Hydraulic) Primary Flight Control Actuation Systems. In *Actuators* (Vol. 14, Number 8). Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/act14080382

Baldo, L., De Martin, A., Terner, M., Jacazio, G., & Sorli, M. (2025). Implementing PHM for legacy flight control actuators through operational aircraft data: Approach and lessons learned. *Results in Engineering*, *28*. https://doi.org/10.1016/j.rineng.2025.107214

Bergstra, J., Bardenet, R., Bengio, Y., & Kégl, B. (2011). Algorithms for Hyper-Parameter Optimization. *Adv. Neural Inf. Process. Syst.*, *24*.

Byington, C. S., Watson, M., & Edwards, D. (2004). Data-driven neural network methodology to remaining life predictions for aircraft actuator components. *IEEE Aerospace Conference Proceedings*, *6*, 3581–3589. https://doi.org/10.1109/AERO.2004.1368175

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. https://doi.org/10.1145/2939672.2939785

De Martin, A., Dellacasa, A., Jacazio, G., & Sorli, M. (2018). High-Fidelity Model of Electro-Hydraulic Actuators for Primary Flight Control Systems. *Proceedings of the 2018 Bath/ASME Symposium on Fluid Power and Motion Control FPMC2018 September 12-14, 2018, University of Bath, United Kingdom*, (4), V001T01A058. https://doi.org/10.1115/fpmc2018-8917

Guo, R., & Gan, Q. (2017). Prognostics for a Leaking Hydraulic Actuator Based on the F-Distribution Particle Filter. *IEEE Access*, *5*, 22409–22420. https://doi.org/10.1109/ACCESS.2017.2759119

Guo, R., & Sui, J. (2019). Remaining Useful Life Prognostics for the Electro-Hydraulic Servo Actuator Using Hellinger Distance-Based Particle Filter. *IEEE Transactions on Instrumentation and Measurement*. https://doi.org/10.1109/tim.2019.2910919

Li, Y. (2016). Mathematical modelling and characteristics of the pilot valve applied to a jet-pipe/deflector-jet servovalve. *Sensors and Actuators, A: Physical*, *245*, 150–159. https://doi.org/10.1016/j.sna.2016.04.048

Lukoševičius, M. (2012). A Practical Guide to Applying Echo State Networks. *Neural Networks*.

Lundberg, S. M., Allen, P. G., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. *Proceedings of the 31st International Conference on Neural Information Processing Systems*. https://github.com/slundberg/shap

Orchard, M. E., & Vachtsevanos, G. J. (2009). A particle-filtering approach for on-line fault diagnosis and failure prognosis. *Transactions of the Institute of Measurement and Control*. https://doi.org/10.1177/0142331208092026

Percival, D. B., & Walden, A. T. (2000). *Wavelet Methods for Time SeriesAnalysis*. Cambridge University Press. https://doi.org/10.1017/CBO9780511841040

Shahkar, S., & Khorasani, K. (2022). A Multidimensional Bayesian Methodology for Diagnosis, Prognosis, and Health Monitoring of Electrohydraulic Servo Valves. *IEEE Transactions on Control Systems Technology*, *30*(3), 931–943. https://doi.org/10.1109/TCST.2021.3079198

Singh, P., & Balasubramanian, R. (2025). Echo State Networks as State-Space Models: A Systems Perspective. In *arXiv preprint arXiv:2509.04422*.

Sprong, J. P., Jiang, X., & Polinder, H. (2020). Deployment of Prognostics to Optimize Aircraft Maintenance – A Literature Review. *JOURNAL OF INTERNATIONAL BUSINESS RESEARCH AND MARKETING*, *5*(4), 26–37. https://doi.org/10.18775/jibrm.1849-8558.2015.54.3004