A Proposal for Application of Physics-Informed Digital Twin and Particle Filtering for the Detection and Prognosis in Harmonic Drives

Roberto Guida[1], Andrea De Martin[1], Antonio Carlo Bertolino[1], Giovanni Jacazio[1], and Massimo Sorli[1]

[1]Politecnico di Torino, Torino, 10139, Italy

roberto.guida@polito.it
andrea.demartin@polito.it
antonio.bertolino@polito.it
giovanni.jacazio@polito.it
massimo.sorli@polito.it

ABSTRACT

Electro-mechanical actuators (EMAs) in aerospace and robotics increasingly rely on harmonic drives, whose compliant architecture introduces nonlinear dynamics and specific degradation mechanisms. Conventional Prognostics and Health Management (PHM) approaches remain limited: data-driven methods require extensive fault datasets and lack interpretability, while physics-based models are often too computationally demanding for embedded real-time use. This work proposes a physics-informed digital-twin-based framework for fault detection and prognosis in harmonic drives. The digital twin is implemented through a Physics-Informed Neural Network (PINN), so that governing mechanical relations are embedded directly into the training process. Wear evolution is described through a physics-based degradation model, while Remaining Useful Life (RUL) is estimated via particle filtering to provide probabilistic prognostic predictions. The study focuses on the digital twin and prognostic modules within a broader PHM architecture. Preliminary results show accurate reconstruction of nonlinear dynamics, physically coherent degradation tracking, and consistent probabilistic RUL prediction.

1. INTRODUCTION

Electro-Mechanical Actuators (EMAs) are increasingly adopted in safety-critical applications, like aerospace and robotics (Clochiatti, 2024), due to their high efficiency, reduced maintenance requirements, and compatibility with electrified architectures (Behbahani, 2010). In aerospace systems, EMAs are employed ad example in secondary flight control surfaces and landing gear actuation (Mazzoleni, 2021). Within EMAs, harmonic drives (Fig. 1) play a pivotal role. Their unique operating principle, based on the elastic deformation of the flexspline induced by the elliptical wave generator, enables high reduction ratios in a single stage, near-zero backlash, and exceptional torque density within compact envelopes (Tuttle, 2002).

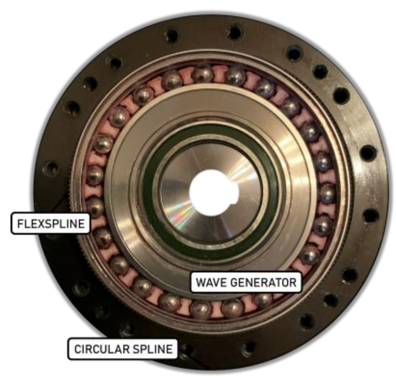

Figure 1. Harmonic Drive.

These characteristics make harmonic drives particularly attractive in applications where precision, stiffness, and weight constraints coexist. However, the same compliant architecture that enables high performance also introduces complex nonlinear dynamics and specific degradation mechanisms. Harmonic drives are susceptible to a variety of fault-to-failure mechanisms (Guida et al., 2024).

Typical failure modes alter meshing stiffness, increase frictional losses, and modify load distribution along the engagement arc (Yang et al., 2021). In aerospace, but not only, such degradations directly impact safety, positional accuracy, and system availability. Undetected faults may lead to catastrophic mechanical failure or unsafe operating conditions. To mitigate these risks, Prognostics and Health Management (PHM) strategies are increasingly adopted.

Existing approaches exhibit significant limitations, purely data-driven methods require large, representative fault datasets, which are rarely available for highly reliable systems like harmonic drives (Tsui et al., 2015), and often lack physical interpretability, limiting robustness beyond the training domain. Conversely, physics-based models provide detailed mechanical insight but are typically too computationally demanding for embedded real-time PHM applications (Yucesan et al., 2021). These challenges motivate strategies that combine the interpretability of physics-based modeling (Biondani et al., 2024) with the adaptability of data-driven learning.

Accordingly, this work proposes a PHM architecture for harmonic drives in EMAs comprises a layered PHM structure separating real-time monitoring, ground-based analytics, and prognostic estimation; a harmonic drive digital twin (Liu et al., 2024) implemented via a Physics-Informed Neural Network (PINN) embedding governing mechanical relations; and a prognostic module for degradation tracking and Remaining Useful Life (RUL) estimation (Ahmadzadeh et al., 2014). By integrating physics-based modeling with constrained machine learning (Karniadakis et al., 2021), the methodology delivers physically interpretable health indicators, robust state estimation under limited data, and forward-looking degradation prediction, and is applicable beyond harmonic drives to broader actuation systems.

1.1. Harmonic Drive Degradation Mechanisms

The FMECA-based analysis by Guida et al. (2024) identifies harmonic drive degradation as a multifactorial process arising from intrinsic factors, such as material fatigue, cyclic plasticity, surface fatigue, and wear, and extrinsic factors including poor lubrication, thermal stresses, assembly errors, and overloads. These causes initiate progressive damage that degrades stiffness, meshing quality, and load distribution, ultimately reducing accuracy and structural integrity. The flexspline is particularly vulnerable, experiencing plastic strain accumulation, instability phenomena (e.g., ratcheting and buckling), and fatigue-driven crack initiation and propagation, while inadequate lubrication and contamination accelerate tribological damage and system-wide wear. Repeated Hertzian contact stresses further drive micro-pitting and material removal at the tooth interface, increasing backlash and kinematic errors, with additional component damage compounding performance decline. Overall, understanding these coupled degradation mechanisms enables the development of physically grounded prognostic models and digital twin frameworks with interpretable health indicators.

1.2. PHM Approaches for Harmonic Drive Systems

PHM of motion transmission systems has been extensively investigated across aerospace, robotics, and industrial automation. In aerospace EMAs, where gear reducers and screw mechanisms are integral to flight control chains, transmission degradation directly affects stiffness, efficiency, and safety margins, motivating the development of advanced diagnostic and prognostic methodologies.

Early PHM approaches for gear transmissions were predominantly model-based. Tuttle et al. (1992) proposed one of the first nonlinear dynamic models of harmonic drives, capturing compliance and hysteresis effects in strain wave gears. Later works on gear crack propagation and vibration response analysis (Ma et al., 2015; Meng et al., 2020) demonstrated how stiffness reduction due to tooth damage can be identified through dynamic response modeling. For aerospace EMAs, De Martin et al. (2017) and Nesci et al. (2020) showed the effectiveness of physics-based models for fault detection and prognosis in flight control actuators. Finite element-based approaches (Chen et al., 2017), provide high-fidelity representations of torsional stiffness and structural compliance in harmonic drives. While accurate, these methods are computationally expensive and unsuitable for real-time PHM deployment. Similarly, wear and backlash modeling in strain wave gears (Raviola et al., 2022) highlighted the progressive stiffness loss induced by degradation but remained primarily simulation-oriented.

Parallel to model-based strategies, data-driven methods have been widely adopted. In robotic transmissions, Jaber and Bicker (2016) employed wavelet transforms combined with neural networks for gear fault detection. However, in aerospace systems, purely data-driven approaches face significant limitations. As highlighted by Yucesan et al. (2021), prognostic models require representative degradation datasets, which are rarely available for high-reliability components such as harmonic drives and flight control transmissions.

Overall, the literature indicates that while model-based approaches ensure interpretability and extrapolation capability, and data-driven methods offer adaptability and pattern recognition power, neither paradigm alone fully satisfies the requirements of aerospace-grade motion transmission systems. To bridge this gap, hybrid PHM approaches combining physics-based modeling and data-driven adaptation have emerged. Ayankoso et al. (2024) proposed a hybrid digital twin scheme for collaborative robots, integrating model-based simulation with data-driven monitoring. Similarly, recent works on harmonic drive modeling (Guida et al., 2025) emphasized the value of computationally efficient nonlinear dynamic models capable of simulating faults for PHM purposes.

2. PROPOSED PHM ARCHITECTURE

The proposed PHM architecture is structured according to a layered and functionally decoupled paradigm, consistent with established condition-based maintenance and aerospace

PHM frameworks. In particular, the overall organization is conceptually aligned with the principles of the OSA-CBM architecture (Swearingen et al., 2007).

2.1. System Overview

The proposed solution separates data acquisition, signal processing, health assessment, and decision support functions into distinct modules. This division improves robustness, scalability, and certification alignment in safety-critical systems (Ranasinghe, 2022). Furthermore, consistent with digital twin layered architectures, a physics-consistent model core is embedded within the framework and interacts bidirectionally with both monitoring and prognostic modules. The resulting PHM system is organized into four main functional modules (Fig. 2): Real-Time On-Board Module; Non-Real-Time On-Ground Module; Partially-Real-Time Prognostic Module; and Physics-Informed Digital Twin Core.

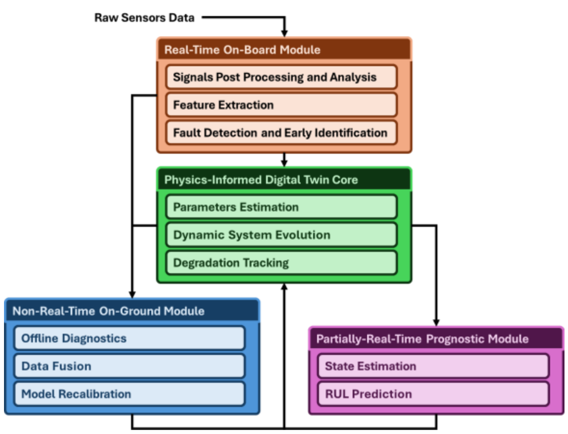

Figure 2. Proposed PHM architecture.

The overall PHM architecture is presented for completeness, as it defines the reference framework within which the proposed methodology is intended to operate. However, the present paper focuses primarily on two key building blocks that can be investigated at this stage independently of application-specific implementation constraints: the physics-informed digital twin and the prognostic module. Accordingly, the real-time and non-real-time layers are discussed here mainly to clarify the functional role of the proposed framework, while the methodological developments and results reported in the following sections are concentrated on the digital-twin formulation and on degradation tracking/RUL estimation.

2.2. Real-Time Module and Non-Real-Time Module

The Real-Time On-Board Module constitutes the first operational layer of the PHM architecture and is responsible for in-flight monitoring and immediate fault detection under strict computational and timing constraints typical of embedded EMA controllers. The module acquires high-frequency measurements, including motor currents, angular position and velocity, and temperature, together with reference signals such as commanded position and current. Signal conditioning is performed and from the processed signals, a set of physically interpretable features is extracted. Fault detection is based on residual evaluation, defined as the deviation between measured signals and the expected behavior provided by the digital twin. These residuals are compared against predefined thresholds to generate fault flags for immediate awareness. In addition, the module outputs time-stamped feature vectors and residual histories, which are transmitted to higher-level modules for prognostic processing and long-term analysis. The Real-Time Module therefore acts as the high-frequency diagnostic interface of the PHM architecture, providing reliable and computationally efficient detection capabilities while ensuring continuous synchronization with the digital twin.

The Non-Real-Time Module represents the high-level analytical layer of the PHM architecture and operates on ground-based infrastructure, where computational constraints are significantly relaxed. It processes mission-aggregated datasets, including feature histories, residual statistics, and selected raw data segments collected across multiple operating cycles and load conditions. In this environment, advanced diagnostic analyses and multi-mission data fusion are performed to identify long-term degradation trends and refine fault classification. The module also integrates on-ground test data and current digital twin parameters, enabling cross-validation between measured behavior and model-based predictions. A key function of this module is the recalibration of the physics-informed digital twin. Based on accumulated operational data, it updates model parameters. These updates are then propagated to the digital twin core and, when appropriate, to the onboard module in the form of adaptive thresholds or normalization factors. Unlike the real-time layer, this module is not constrained by deterministic execution requirements and is therefore designed to enhance model fidelity, diagnostic depth, and long-term prognostic consistency. It serves as the main learning and validation environment of the PHM framework, supporting scalability and alignment with certification-oriented aerospace PHM architectures (Ranasinghe, 2022).

2.3. Partially-Real-Time Module

The Partially-Real-Time Module provides the intermediate layer between in-flight fault detection and ground-based analytics, focusing on health-state estimation and RUL prediction. It operates at a lower update rate than the Real-Time Module, while avoiding the latency associated with ground processing. The module receives data from the onboard layer, as well as updated model parameters and degradation laws from the digital twin core. When available, corrections from the Non-Real-Time Module are also incorporated, ensuring consistency between measurement-driven information and physics-based modeling. Health-state estimation is performed by modeling degradation as a hidden state variable. In this work, a particle filtering approach is adopted. The output of the module consists of health indicators and probabilistic RUL estimates, which can be used for maintenance decision support.

2.4. Digital Twin Core

The Physics-Informed Digital Twin Core constitutes the model-based nucleus of the PHM architecture, providing a physics-consistent and dynamically updated representation of the harmonic drive within the electro-mechanical actuator. Unlike other modules structured by execution rate or computational constraints, the digital twin is defined by its methodological function: it serves as the unified reference model that synchronizes monitoring, diagnostics, and prognostics. It integrates processed measurements and feature data from the Real-Time Module, aggregated datasets and recalibrated parameters from the Non-Real-Time Module, and health state estimates from the Partially-Real-Time Module to maintain an adaptive dynamic model of the transmission. In operation, the twin supplies expected system responses to support residual generation and adaptive thresholds in real-time monitoring, while providing state predictions and sensitivity information for health estimation and RUL forecasting in the prognostic layer. Through periodic updates from the ground module, it incorporates parameter corrections. Functioning as a continuously synchronized virtual counterpart rather than an offline simulation tool, the digital twin centralizes physical consistency within a single adaptive core, preventing fragmentation between data-driven estimators and analytical models and enhancing robustness.

3. PHYSICS-INFORMED DIGITAL TWIN

The DT proposed in this work is built upon a Physics-Informed Neural Network (PINN) framework, enabling the integration of first-principles mechanical modeling with data-driven learning. Unlike conventional neural networks, which approximate input-output mappings solely from data, PINNs embed the governing physical equations of the system directly into the training process, ensuring that the learned solution satisfies both measurement consistency and model-based dynamics. As a result, the network does not act as a black-box regressor but as a physics-constrained estimator.

PINNs have been successfully applied in several engineering domains, including structural health monitoring (Al-Adly et al., 2024), fluid dynamics (Cai et al. 2021), battery state estimation (Tian et al, 2025), and rotating machinery degradation modeling. Their ability to infer latent states and unknown parameters under partial observability makes them particularly suitable for systems where direct measurement of degradation variables is not feasible. In recent years, PINN-based formulations have been increasingly explored within digital twin architectures, where they provide a mechanism to continuously synchronize virtual models with physical assets while preserving interpretability and stability (Wang et al. 2025).

The physical model adopted in this work is based on a computationally efficient multibody formulation capable of capturing the nonlinear dynamic behavior of harmonic drive reducers while remaining suitable for integration within PHM and digital twin frameworks (Guida, 2025). Unlike purely kinematic or high-fidelity finite element approaches, the model represents each main component of the harmonic drive (the wave generator, flexspline, and circular spline) as distinct bodies interacting through equivalent compliant elements. This multibody structure allows the model to reproduce both macroscopic transmission behavior and internal nonlinear interactions, including compliance, friction, hysteresis, and meshing dynamics.

3.1. PINN Structure

The Physics-Informed Long Short-Term Memory (LSTM) architecture was selected as the backbone of the PINN due to its capability to capture temporal dependencies and internal state evolution in nonlinear dynamical systems. In the context of harmonic drive dynamics, where the system response is strongly influenced by history-dependent effects such as hysteresis, compliance, and load redistribution, recurrent architectures provide a natural framework for modeling time-correlated behaviors. While alternative architectures could also be considered, the LSTM offers a good trade-off between modeling capability, training stability, and integration with physics-based constraints, making it suitable for this preliminary study. The Physics-Informed LSTM used as the digital-twin estimator was trained with a purpose-built MATLAB pipeline designed to balance physical consistency, temporal expressiveness and computational robustness. Time-series outputs from multiple simulation runs (that could be substituted by experimental or historical data) were collected, optionally anti-aliased and downsampled, and split at experiment level into training and validation sets (75%/25%). Training was carried out by minimizing a composite loss function combining a data-consistency term and a physics-consistency term:

$$\mathcal{L}_{tot} = \mathcal{L}_d + \lambda_c \, \mathcal{L}_c \tag{1}$$

where $\mathcal{L}_{\mathrm{d}}$ is the mean squared error between the predicted and measured reducer output torque, and $\mathcal{L}_{\mathrm{c}}$ is the mean squared residual of the governing mechanical relations embedded in the PINN. In particular,

$$\mathcal{L}_d = \frac{1}{NBT} \sum_{b=1}^{B} \sum_{t=1}^{T} \left(\hat{Q}_{red}^{(b,t)} - Q_{red}^{(b,t)} \right)^2 \tag{2}$$

while $\mathcal{L}_c$ is obtained from the residuals (r_1, r_2, r_3, r_4) of the multibody equations used to describe the harmonic-drive dynamics:

$$\begin{aligned} r_1 &= Q_{res} - F_b r_b - J_{WG}\,\ddot{\theta}_m \\ r_2 &= F_b - F_{m,x} - m_{FS,x}\,\ddot{x}_{FS} \\ r_3 &= -F_{FS} - F_{m,y} - m_{FS,y}\,\ddot{y}_{FS} \\ r_4 &= F_m r_m - Q_{red} - J_{CS}\ddot{\theta}_f \end{aligned} \tag{3}$$

$$\mathcal{L}_c = \frac{1}{4\,NBT}\sum_{j=1}^{4}\sum_{b=1}^{B}\sum_{t=1}^{T} r_j^2 \tag{4}$$

The physics weight λ_c was treated as a learnable parameter jointly optimized with the network parameters during training. This choice allows the model to automatically balance data fitting and physics enforcement during training. Since the data and physics terms involve quantities with different units and magnitudes, each residual component was individually scaled by an adaptive algorithm. Adaptive residual scaling is useful because the different physics residuals often exhibit very different numerical magnitudes, units, and sensitivities, especially when some terms depend on latent variables or higher-order derivatives. In such cases, the largest residual may dominate the physics loss and prevent the network from learning a balanced solution, even when the model structure is correct. Automatically rescaling each residual according to its observed magnitude helps keep their contributions comparable during training, improves numerical conditioning, and enhances optimization stability. Although this strategy is not strictly rigorous from a physical modeling perspective, it is often an effective practical regularization tool for achieving convergence in PINN training. In addition, a ramping strategy was adopted for the physics contribution, so that its effect was introduced progressively during the early stages of training rather than being enforced at full strength from the beginning. This gradual activation helps the network first capture the dominant data-driven trends and then increasingly satisfy the physical constraints, further reducing optimization instabilities and improving convergence robustness. Moreover, the initial value of the physics weight was treated as a tunable hyperparameter and optimized together with the remaining architectural and training settings. As a result, the final trade-off between the data-driven and physics-informed contributions was not imposed a priori, but learned from the training process itself.

A preliminary sensitivity analysis was implicitly performed through Bayesian optimization over the main hyperparameters of the PINN architecture and training process. This exploration was aimed at identifying configurations providing a suitable compromise between prediction accuracy, convergence stability, and physics consistency. Although a full parametric study is outside the scope of this preliminary work, the adopted optimization procedure provided evidence of the robustness of the selected final configuration. The final selected hyperparameter configuration is summarized in Table 1.

Table 1. Hyperparameters of the tuned LSTM PINN.

Hyperparameter	**Value**
Number of LSTM layers	3
Hidden units	512
Bidirectionality	no
Dropout	0.41873
Learning rate	4.51×10^{-3}
Physics weight	8×10-2

3.2. Degradation Tracking and RUL Estimation

Meshing teeth wear evolution was modeled through a physics-based approach grounded on Archard's law (Archard, 1953), enabling a direct relationship between sliding contact conditions, normal meshing force, and material removal over time. Within this context, specific health indicators were defined to capture the macroscopic effects of microscopic material loss. For on-ground monitoring, two features demonstrated strong correlation with wear severity: kinematic error indicator quantifying the deviation between expected and actual output position, and current-speed gain metric reflecting variations in transmission efficiency. These features were normalized and analyzed across multiple degradation levels, showing clear monotonic trends and strong separability between early-stage and advanced wear conditions.

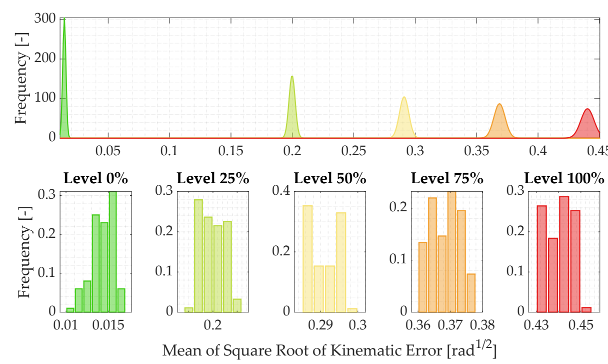

Figure 3. Statistical evolution of the selected kinematic-error-based health feature at increasing wear levels.

For in-flight monitoring, a dedicated kinematic-error-based feature was selected as the main health indicator, since it provides a direct and physically interpretable measure of the progressive mismatch between the expected and actual transmission behavior. As shown in Fig. 3 as wear increases, the selected feature exhibits a monotonic trend, making it suitable for health-state tracking and prognostic purposes. In addition to its average growth with degradation, the feature distribution progressively shifts and broadens, indicating not only an increase in its mean value but also a degradation-related increase in variability. This statistical behavior is consistent with the progressive alteration of meshing conditions and load transfer mechanisms within the reducer. From a prognostic standpoint, these properties are particularly relevant: monotonicity supports robust state evolution modeling, while the distributional changes provide information on uncertainty growth as the system approaches end-of-life. For this reason, the selected feature was adopted as the observation variable for the particle-filter-based prognostic framework. Once degradation is detected and tracked, RUL estimation is performed using a particle filtering framework. Particle filtering (PF) was selected in this preliminary study because it provides a flexible Bayesian framework for estimating hidden degradation states in the presence of nonlinear dynamics and non-Gaussian uncertainty. In addition, it naturally yields probabilistic RUL estimates rather than single deterministic predictions, which is advantageous in safety-critical applications. At this stage, the PF is adopted as a suitable baseline prognostic solution, while more extensive comparative analyses with alternative filtering or state-estimation approaches are left for future developments.

The resulting RUL is not a single deterministic value but a probability distribution, from which risk-based metrics can be extracted. By combining statistically robust health indicators, and particle-filter-based state estimation, the degradation tracking framework will enable continuous synchronization between measurable actuator signals and internal wear evolution.

4. RESULTS

As aforementioned the dataset used in this study was generated through a high-fidelity nonlinear simulation model of an electro-mechanical actuator including the harmonic drive transmission. The model captures coupled electrical, mechanical, and thermal dynamics, together with physics-based degradation mechanisms, ensuring consistency with realistic aerospace operating conditions.

To ensure variability and generalization capability, the simulation campaign included a population-based approach, where key model parameters were varied within physically meaningful ranges to account for uncertainties and system dispersion. The resulting dataset consists of time-series signals including motor currents, angular positions, velocities, and derived quantities such as output torque and kinematic error, which are consistent with signals typically available in real electro-mechanical actuator systems. Prior to training, the data were processed through a dedicated preprocessing pipeline. Signals were filtered and, when necessary, downsampled to ensure numerical stability and reduce computational burden. The dataset was then organized at experiment level and split into training and validation subsets, preserving the independence of different simulation runs. This simulation-driven dataset provides a physically consistent and controllable environment for the development and preliminary validation of the proposed methodology, while maintaining strong adherence to realistic actuator behavior and degradation mechanisms.

The first set of results evaluates the dynamic reconstruction capability of the proposed LSTM-PINN DT architecture. In the training set, the predicted signal closely follows the true torque profile across the full range of operating conditions, including high-amplitude transients and rapid oscillations (Fig. 4). Peak values exceeding ±2500 units (Nm) are accurately captured, with no evident phase lag or systematic amplitude bias.

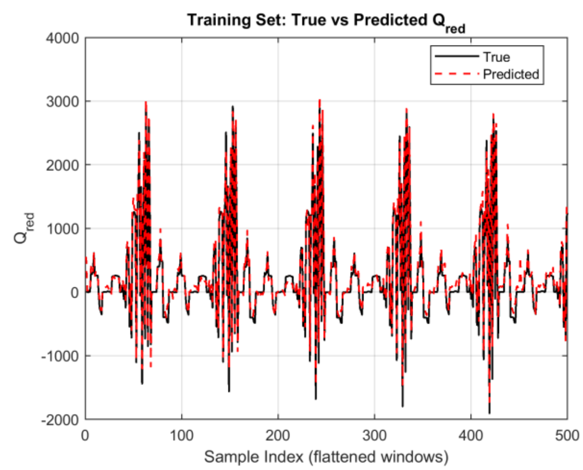

Figure 4. Comparison between the measured reducer output torque and the network prediction for the training set.

The model reproduces both low-level oscillatory behavior and large load spikes, indicating that the recurrent architecture successfully encodes the temporal dependencies of the actuator dynamics.

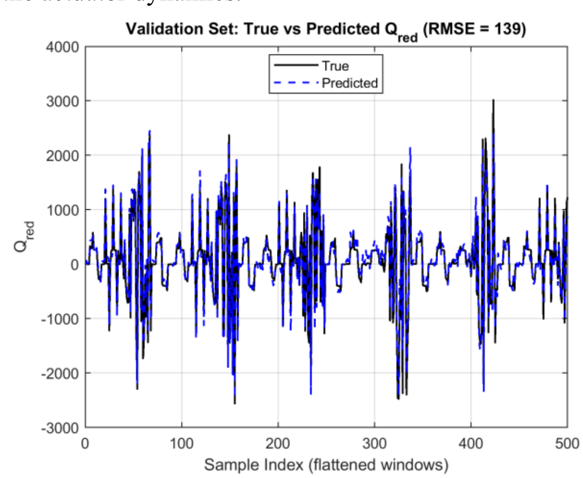

Figure 5. Comparison between the measured reducer output torque and the network prediction for the validation set.

The validation results confirm strong generalization performance. The predicted torque remains well aligned with the measured signal, even in dynamic regimes characterized by rapid sign changes and high-frequency content (Fig. 5).

The scatter plot (Fig. 6) of predicted versus true torque values further demonstrates this behavior: the samples are tightly clustered around the ideal 45° line, with limited dispersion across the full dynamic range (approximately -3000 to +3000). This confirms that the digital twin is not merely interpolating the training patterns but retains predictive consistency on unseen data.

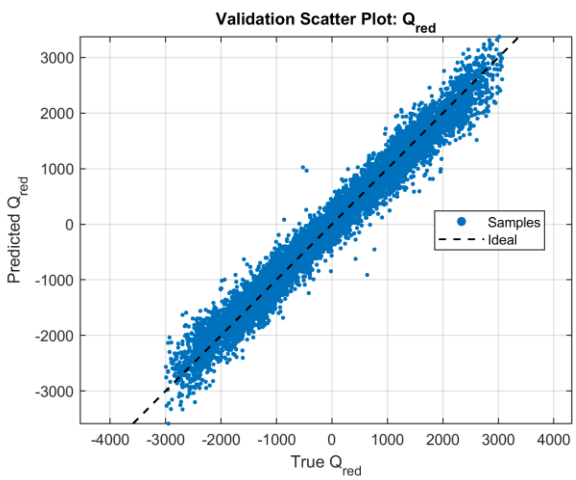

Figure 6. Scatter plot of predicted versus true torque.

Overall, these results demonstrate that the hybrid LSTM–PINN structure is capable of reconstructing nonlinear transmission dynamics while preserving stability and temporal coherence. In addition to the prediction accuracy results, the convergence behavior of the PINN training process was analyzed by monitoring the evolution of the loss terms over the epochs.

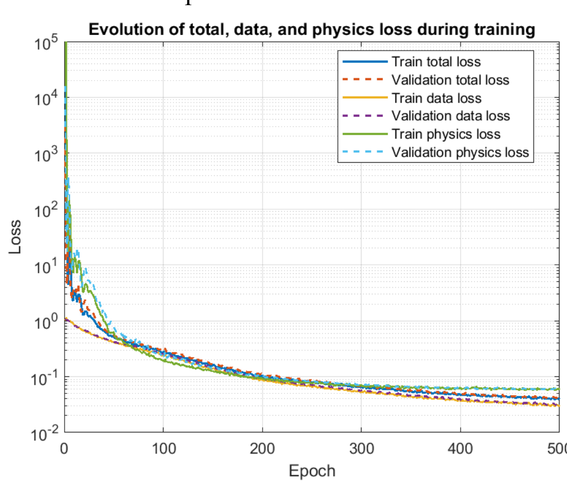

Figure 7. Evolution of the total, data and physics training and validation losses.

Fig. 7 reports the training and validation total loss, the results indicate stable convergence of the optimization process, with progressive reduction of both the data mismatch and the physics residuals. This behavior confirms that the proposed training strategy successfully enforces physical consistency without compromising the quality of the signal reconstruction.

The evolution of the learned physics weight was also monitored during training. Its progressive adaptation confirms that the relative importance of the physics-informed contribution is automatically adjusted during optimization, thus avoiding an arbitrary fixed balance between the purely data-driven and physics-based components.

4.1. Health Tracking Performance

The second set of results focuses on degradation tracking capability under progressive meshing teeth wear. The selected in-flight health feature $\left(\mathrm{HF}_{\mathrm{MTW,IF}}\right)$ is derived from the kinematic error between input and output shafts.

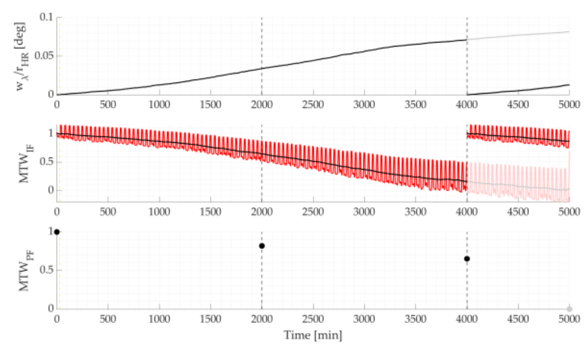

Figure 8. Fault-to-failure evolution of the wear level and health feature.

The fault-to-failure evolution plot (Fig. 8) further illustrates monotonic behavior of both the physical wear variable and the corresponding health feature. As the wear depth increases, the in-flight feature follows a consistent and progressive upward trend, with limited noise-induced variability. The integrated tracking of wear and feature evolution confirms that the proposed architecture maintains coherence between internal degradation mechanisms and externally measurable quantities. The health feature therefore provides a consistent mapping from latent wear dynamics to observable signals, enabling robust state estimation prior to failure.

4.2. RUL Prediction

The final set of results evaluates the Remaining Useful Life (RUL) estimation obtained through the particle filter (PF) framework applied to the tracked health indicator.

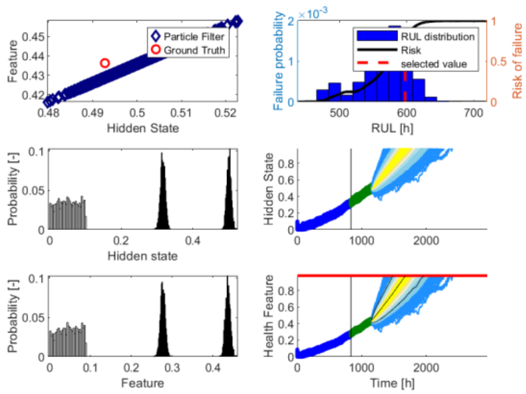

Figure 9. Intermediate degradation stages PF algorithm application.

At intermediate degradation stages, the particle cloud accurately tracks the hidden wear state, with the estimated state closely aligned with ground truth (Fig. 9). The probability distributions of both the hidden state and the observed feature show concentrated peaks, indicating low estimation uncertainty. The corresponding RUL distribution is well-defined, with moderate variance reflecting limited long-horizon uncertainty.

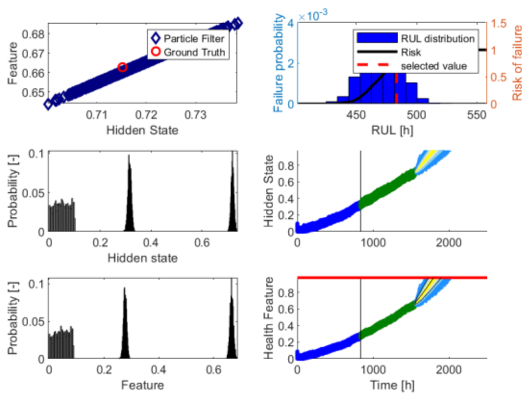

Figure 10. Near-to-end degradation stages PF algorithm application.

As degradation progresses toward later stages (Fig. 10), the PF maintains coherent state tracking. The RUL distribution shifts toward lower values, and the risk-of-failure curve steepens as the predicted failure threshold approaches. The temporal evolution plots show particle trajectories converging toward the failure threshold, providing an intuitive visualization of degradation dynamics.

From a quantitative standpoint, prognostic performance in PHM applications is commonly evaluated using metrics such as prediction error, uncertainty dispersion, and α–λ performance indicators, which jointly assess the accuracy and timeliness of RUL predictions.

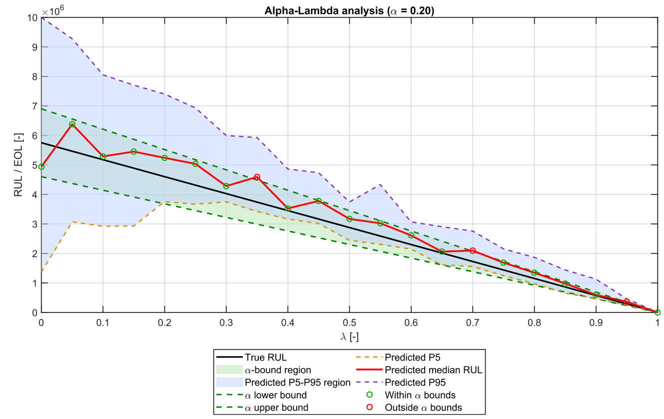

Figure 11. PF α–λ analysis.

In order to provide a preliminary quantitative assessment of the prognostic routine, an α–λ analysis was performed with $\alpha = 0.2$ (Fig. 11). The α–λ analysis indicates that the prognostic model captures the overall degradation trend with reasonable fidelity, as the predicted median RUL closely follows the true RUL across most of the operational horizon. However, the uncertainty bounds (P5–P95) are notably wider at lower λ values, reflecting increased epistemic uncertainty during the early stages of the lifecycle, and progressively narrow as failure approaches. The α-bounded region shows that the majority of predictions fall within acceptable error margins, although several deviations occur in the initial phase, where the model occasionally under- or over-estimates the RUL. This behavior suggests a degree of heteroscedasticity in the prediction error, with reduced reliability in early-life predictions and improved calibration near end-of-life. Overall, the model demonstrates satisfactory probabilistic calibration and convergence, but highlights the need for enhanced early-stage prognostic sensitivity and uncertainty quantification.

5. CONCLUSION

This work presented a hybrid PHM architecture for harmonic drives operating within electro-mechanical actuators, combining physics-based modeling and constrained machine learning within a unified digital twin framework.

By structuring the system into real-time, partially-real-time, and ground-based layers, the architecture satisfies embedded computational constraints while preserving long-term model fidelity and prognostic capability.

The physics-informed LSTM-based digital twin demonstrated accurate reconstruction of nonlinear transmission dynamics and strong generalization performance. The proposed health indicators showed monotonic and physically interpretable correlation with wear progression, enabling reliable degradation tracking.

The particle-filter-based prognostic module provided probabilistic RUL estimates with coherent uncertainty evolution from intermediate to near-failure stages.

Overall, the results confirm that embedding mechanical consistency within adaptive learning architectures enhances robustness, interpretability, and extrapolation capability under limited fault data availability.

The present study should be considered as a preliminary methodological contribution, primarily aimed at establishing the feasibility of a physics-informed digital-twin-based prognostic framework for harmonic drives. Future work will extend the quantitative validation of the prognostic layer through standard PHM metrics, comparative analyses of state-estimation strategies, and a broader assessment of sensitivity to model and training hyperparameters.

NOMENCLATURE

F_b	Deformable bearing force
F_m	Meshing force
F_{FS}	Flexspline restoring force
J_{CS}	Circular Spline inertia
J_{WG}	Wave Generator inertia
$\mathcal{L}_d$	Data loss
$\mathcal{L}_c$	Physics loss
m_{FS}	Flexspline mass
Q_{red}	HD output torque
Q_{res}	HD input torque
r_b	Bearing force radius
r_m	Meshing force radius
x_{FS}	Flexspline equivalent radial deformation
y_{FS}	Flexspline equivalent torsion
λ_c	Physics loss weight
θ_f	Flap angular position
θ_m	Motor angular position

REFERENCES

Ahmadzadeh, F., & Lundberg, J. (2014). Remaining useful life estimation. International Journal of System Assurance Engineering and Management, 5(4), 461–474.

Al-Adly, A. I., & Kripakaran, P. (2024). Physics-informed neural networks for structural health monitoring: A case study for Kirchhoff–Love plates. Data-Centric Engineering, 5, e6.

Archard, J. (1953). Contact and rubbing of flat surfaces. Journal of Applied Physics, 24(8), 981–988.

Ayankoso, S., Kaigom, E., Louadah, H., Faham, H., Gu, F., & Ball, A. (2024). A hybrid digital twin scheme for the condition monitoring of industrial collaborative robots. Procedia Computer Science, 232, 1099–1108.

Behbahani, A. R., & Semega, K. J. (2010). Control strategy for electro-mechanical actuators versus hydraulic actuation systems for aerospace applications. SAE Technical Paper. Power Systems Conference.

Biondani, F., Dall'Ora, N., Tosoni, F., Fraccaroli, E., Migliore, D. F., Acerra, F., & Fummi, F. (2024). Fault injection for synthetic data generation in aircraft: A simulation-based approach. Proceedings of the 2024 IEEE 22nd International Conference on Industrial Informatics (INDIN) (pp. 1–8). IEEE.

Cai, S., Mao, Z., Wang, Z., Yin, M., & Karniadakis, G. E. (2021). Physics-informed neural networks (PINNs) for fluid mechanics: A review. Acta Mechanica Sinica, 37(12), 1727–1738.

Chen, Y. C., Cheng, Y. H., Tseng, J. T., & Hsieh, K. J. (2017). Study of a harmonic drive with involute profile flexspline by two-dimensional finite element analysis. Engineering Computations, 34(7), 2107–2130.

Clochiatti, E., Scalera, L., Boscariol, P., & Gasparetto, A. (2024). Electro-mechanical modeling and identification of the UR5 e-series robot. Robotica, 42(7), 2430–2452.

Guida, R., Bertolino, A. C., De Martin, A., & Sorli, M. (2024). Comprehensive analysis of major fault-to-failure mechanisms in harmonic drives. Machines, 12(11), 776.

Guida, R., Bertolino, A. C., De Martin, A., & Sorli, M. (2025). A new computationally efficient model of the non-linear dynamics in harmonic drive reducers. Mechanism and Machine Theory, 209, 105992.

Jaber, A. A., & Bicker, R. (2016). Fault diagnosis of industrial robot gears based on discrete wavelet transform and artificial neural network. Insight: Non-Destructive Testing and Condition Monitoring, 58(4), 179–186.

Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L. (2021). Physics-informed machine learning. Nature Reviews Physics, 3(6), 422–440.

Liu, Y., Liu, Y., McCrory, J., & Guo, X. (2024). High-fidelity digital twin modelling for predictive maintenance state-of-the-art.

Ma, H., Pang, X., Zeng, J., Wang, Q., & Wen, B. (2015). Effects of gear crack propagation paths on vibration responses of the perforated gear system. Mechanical Systems and Signal Processing, 62, 113–128.

Mazzoleni, M., Di Rito, G., & Previdi, F. (2021). Electro-mechanical actuators for the more electric aircraft (Vol. 2021). Cham, Switzerland: Springer.

Meng, Z., Shi, G., & Wang, F. (2020). Vibration response and fault characteristics analysis of gear based on time-varying mesh stiffness. Mechanism and Machine Theory, 148, 103786.

Nesci, A., De Martin, A., Jacazio, G., & Sorli, M. (2020). Detection and prognosis of propagating faults in flight control actuators for helicopters. Aerospace, 7(3), 20.

Ranasinghe, K., Sabatini, R., Gardi, A., Bijjahalli, S., Kapoor, R., Fahey, T., & Thangavel, K. (2022). Advances in integrated system health management for mission-essential and safety-critical aerospace applications. Progress in Aerospace Sciences, 128, 100758.

Raviola, A., De Martin, A., & Sorli, M. (2022). A preliminary experimental study on the effects of wear on the torsional stiffness of strain wave gears. Actuators, 11(11), 305.

Swearingen, K., Majkowski, W., Bruggeman, B., Gilbertson, D., Dunsdon, J., & Sykes, B. (2007). An open system architecture for condition based maintenance overview. Proceedings of the 2007 IEEE Aerospace Conference (pp. 1–8). IEEE.

Tian, A., He, L., Ding, T., Dong, K., Wang, Y., & Jiang, J. (2025). A generic physics-informed neural network framework for lithium-ion batteries state of health estimation. Energy, 332, 137215.

Tsui, K. L., Chen, N., Zhou, Q., Hai, Y., & Wang, W. (2015). Prognostics and health management: A review on data driven approaches. Mathematical Problems in Engineering, 2015(1), 793161.

Tuttle, T. D. (1992). Understanding and modeling the behavior of a harmonic drive gear transmission. Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Tuttle, T. D., & Seering, W. P. (2002). A nonlinear model of a harmonic drive gear transmission. IEEE Transactions on Robotics and Automation, 12(3), 368–374.

Wang, Z., Jahanshahi, M. R., Lund, A., Shahriar, A., & Montoya, A. (2025). Physics-informed machine learning for hybrid digital twin–enhanced damage detection and localization. Journal of Engineering Mechanics, 151(12), 04025080.

Yang, L. T., Shao, Y. M., Jiang, W. W., Zhang, L. K., Wang, L. M., & Xu, J. (2021). Effects of tooth surface crack propagation on meshing stiffness and vibration characteristic of spur gear system. Applied Sciences, 11(4), 1968.

Yucesan, Y. A., Dourado, A., & Viana, F. A. (2021). A survey of modeling for prognosis and health management of industrial equipment. Advanced Engineering Informatics, 50, 101404.

BIOGRAPHIES

Roberto Guida is a Research Assistant at Politecnico di Torino, where he earned his PhD in Mechanical Engineering in 2025. He is a member of the Mechatronics and Servosystems Research Group within the Department of Mechanical and Aerospace Engineering. His research focuses on actuation systems, aerospace control, motion transmission mechanisms, and industrial and collaborative robotics, with a particular emphasis on dynamic modeling and the application of diagnostics, prognostics, and health management methodologies.

Antonio Carlo Bertolino got his Mechanical Engineering degree from Politecnico di Torino in 2016, followed by a Ph.D. in Mechanical Engineering in 2020. Currently serving as an Assistant Professor in Applied Mechanics at Politecnico di Torino, his research focuses on high-fidelity PHM-oriented dynamic modeling of mechanical components and servosystems, electro-hydraulic, electromechanical and hybrid actuators for flight control systems in both fixed and rotary-wing aircraft, with a particular emphasis on ball screw dynamics. Antonio has collaborated on research and consulting projects with various industrial and aerospace companies.

Andrea De Martin servers as an assistant professor in the Mechanical and Aerospace Engineering department at Politecnico di Torino. He obtained his master's degree in mechanical engineering from the same institution in 2013 and has since been a member of the Mechatronics and Servosystems research group. He holds a PhD degree in Mechanical Engineering. His main research interests are in the area of actuation systems and aerospace equipment, with a particular focus on the application of Prognostics and Health Management techniques to flight-control systems.

Giovanni Jacazio is former full professor of applied mechanics and of mechanical control systems. His main research activity is in the area of aerospace control and actuation systems and of prognostics and health management. He is a member of the SAE A-6 Committee on Aerospace Actuation Control and Fluid Power Systems, and a member of the International Society of Prognostics and Health Management.

Massimo Sorli born in Sanremo (Italy) in 1956, has been Full Professor of Applied Mechanics in the Politecnico di Torino, Italy, since 2000. He graduated in Mechanical Engineering at the Politecnico di Torino in 1981 and received his PhD in Applied Mechanics in 1987 at the same university. Since October 2007 until December 2011, he has been Director of the Department of Mechanics, from January 2012 until October 2015 he has been Director of the Department of Mechanical and Aerospace Engineering (DIMEAS), both at Politecnico di Torino. His scientific activity is mainly focused in the area of mechatronics, mechanical and fluid (pneumatic, hydraulic and electromechanical) servo-systems for automotive and aeronautical applications, innovative devices in flight control systems, prognostics of servo-actuators, Rendez-Vous & Docking Systems for spacecraft. He has been the author or co-author of more than 230 papers.