# Physics-Informed Machine Learning-Assisted State Estimation for Degrading System Considering Sensor Degradation Impacts

Trung-Thanh N. Thai[1], Phuc Do[2], Benoit Iung[3]

[1,3]*Université de Lorraine, Université de Lorraine, CNRS, CRAN, Nancy, France*
*thanh-thai@univ-lorraine.fr*
*benoit.iung@univ-lorraine.fr*

[2]*SyCoIA, IMT Mines Ales, Ales, France*
*phuc.do@mines-ales.fr*

## ABSTRACT

Degradation estimation is a fundamental component of prognostics, yet it is often compromised by the idealized assumption of perfect sensor fidelity. In harsh industrial environments, the concurrent degradation of both primary assets and monitoring sensors introduces severe observational ambiguity. Traditional state estimation methods, such as Kalman Filters (KF) or standard Sequential Monte Carlo (SMC), often fail to track states when sensor degradation causes the observation model to become non-stationary and complex. To address this, this paper proposes a hybrid degradation estimation framework that integrates Physics-Informed Machine Learning (PIML) into SMC inference. The joint evolution of asset and sensor degradation is modeled through stochastic Wiener processes, capturing both deterministic drift and diffusion. A learnable observation model is then implemented using a Multilayer Perceptron (MLP) to map the relationship between degradation states and measurements within a recursive Bayesian optimization framework. Numerical validation demonstrates that the proposed method achieves robust tracking accuracy under non-stationary conditions, performing competitively with model-based filtering approaches, presenting a promising approach in supporting prognostics frameworks considering sensor degradation.

**Keywords:** Prognostics and Health Management (PHM), Sensor Degradation, Measurement Degradation, Hybrid Approach, Online Learning, Monte Carlo Inference.

## 1. INTRODUCTION

Modern industrial frameworks increasingly rely on predictive maintenance strategies. At the core of these maintenance strategies lies the ability to accurately estimate current degradation states of systems and project its future trajectory to estimate the Remaining Useful Life (RUL). However, a pervasive challenge is that sensors deployed to monitor these systems are themselves undergo their own physical degradation by cumulative damage over long period of operating time (Cai, Liu, & Xie, 2017; Mo, Wang, Xie, & Xiong, 2017; Li & Ying, 2018). This often leads to the degradation of sensor measurement, indicated by additive, proportional, noise errors or in other forms (Li, Zhang, & Wang, 2013; Zhang, Si, Du, & Hu, 2018; Chen et al., 2018; Ma, Jiang, & Li, 2019; Yoo et al., 2020). Consequently, this makes it exceedingly difficult to distinguish between actual underlying system wear and spurious sensor malfunction (Jiang, Djurdjanovic, Ni, & Lee, 2006; Mehdizadeh, John, Wang, Ghorbani, & Rowe, 2012; Liu, Do, Iung, & Xie, 2019; Hachem, Vu, & Fouladirad, 2024), thereby, decrease the precision of RUL prediction and maintenance performance (Liu et al., 2019; Hachem, Vu, & Fouladirad, 2024).

Traditional PHM filtering approaches, such as standard Kalman Filters or Sequential Monte Carlo (SMC) methods, assume that the measurement function mapping the hidden states to the observations is explicitly formulated and invariant (Csuzdi, Bécsi, & Törő, 2026). These foundational assumptions often collapse when sensor degradation introduces time-varying or non-linear terms into the measurement model. To circumvent the limitations of rigid analytical models, the PHM domain has increasingly turned toward data-driven methodologies, with the integration with artificial neural networks and deep learning architectures (Cancelliere, Girard, Bourinet, & Broggi, 2023). However, purely data-driven architectures require large repositories of run-to-failure data for training that span all possible degradation modes, fault conditions, and operating regimes (Jung, 2022). In mission-critical industrial applications, exhaustive datasets are usually scarce, expensive to acquire, or unavailable due to the catastrophic consequences of

allowing a system to run to complete failure (Jung, 2022). Furthermore, unconstrained data-driven architectures can produce unphysical predictions when exposed to out-of-distribution data, violating the fundamental laws of thermodynamics or material science (Guo, You, Li, Tang, & Li, 2023; Farea, Yli-Harja, & Emmert-Streib, 2024).

The convergence of model-based physical principles with data-driven approaches has catalyzed the development of Physics-Informed Machine Learning (PIML) (Fink et al., 2025). PIML models embed structural knowledge, physical constraints into data-driven architectures and model-training paradigm (Farea et al., 2024). This synthesis reduces the dependency on vast historical datasets, enhances extrapolation capabilities outside the training domain, and guarantees that the network's outputs adhere to physical realities (Farea et al., 2024). However, static offline-trained data-driven models still struggle to adapt to the stochastic variations of individual assets in real-time operation. To track the stochastic evolution of complex non-linear and non-Gaussian degradation trajectories online, SMC methods, commonly known as Particle Filtering (PF), have proven highly effective (Wang & Gao, 2014). Particle filters are capable of handling non-linear transition densities without requiring local linearization. However, classic particle filters suffer from sample impoverishment and degeneracy, particularly when attempting to perform joint state and parameter estimation in high-dimensional spaces over extended temporal periods (Wang & Gao, 2014).

To bridge the gaps between rigid analytical measurement models, the vast data requirements of purely data-driven architectures, and the inherent limitation of offline models to adapt to real-time stochastic variations in individual assets, this paper introduces an estimation framework using PIML observation model with SMC inference engine, robustly compensates for sensor degradation impacts while accurately tracking the true system health. The remainder of this paper is organzed as follows: section 2 details the proposed hybrid degradation estimation framework, covering the stochastic modeling of system and sensor states, the measurement architecture, and the integration of physics-informed constraints within the SMC inference; section 3 presents numerical validations of the methodology; finally, section 4 provides conclusion and outlines potential future work.

## 2. METHODOLOGY: A HYBRID DEGRADATION ESTIMATION FRAMEWORK CONSIDERING SENSOR DEGRADATION IMPACTS AND IMPLEMENTATION

### 2.1. Stochastic Modelling of System and Sensor Degradation

Continuous degradation processes in mechanical, structural, and electrochemical systems are typically non-decreasing over time and subject to inherent stochasticity (Liu, Matias, Jäschke, & Vatn, 2022). The Wiener process, characterized by its continuous paths and independent, normally distributed increments, is widely utilized in reliability engineering and PHM to model such degradation trajectories (Tang, Jiang, & Mao, 2025). The mathematical properties of the Wiener process allow for the elegant incorporation of both deterministic degradation trends (drift) and stochastic environmental fluctuations (diffusion) (Kahle & Lehmann, 2009).

The degradation states at discrete time step $k$, denoted as $X_k$ and $Y_k$, can be modelled using discretized Wiener processes governed by the Euler-Maruyama approximation:

$$X_k = X_{k-1} + \mu_x \Delta t + \sigma_x \sqrt{\Delta t} W_{x,k} \quad (1)$$

$$Y_k = Y_{k-1} + \mu_y \Delta t + \sigma_y \sqrt{\Delta t} W_{y,k} \quad (2)$$

Here, $\mu_x$ and $\mu_y$ represent the deterministic drift coefficients, which map to the mean degradation rate of the respective system components. The parameters $\sigma_x$ and $\sigma_y$ represent the diffusion coefficients, which encapsulate the volatility or inherent stochastic noise in the degradation process. The term $\Delta t$ is the constant sampling interval, and $W$ denotes a standard normal random variable drawn from $\mathcal{N}(0,1)$.

### 2.2. Learnable Measurement Modelling

While underlying system and sensor degradation is effectively modelled by stochastic processes, mapping these latent states to sensor measurements may confront high complexity as highly nonlinear and non-stationary (Li, Si, Zhang, & Li, 2024). Consequently, traditional static or linear observation models, such as those used in standard filters, may be inadequate. To resolve this, this paper adopts a dynamic and learnable approach. The observation model can be formulated as $Z_k = h_\theta(X_k, Y_k) + \varepsilon_k$, where the parameterized function $h_\theta$ can continuously learn and adapt online as the system operates.

In this implementation, the measurement model used a multilayer perceptron (MLP) architecture, designed to map the observed sensor reading $Z$ with the latent states $(X, Y)$. The architecture of the proposed model is structured as a continuous feed-forward network. The input layer has a dimension of $2$, receiving the spatial coordinates of the latent degradation states, $X$ and $Y$. It then is followed by $3$ hidden layers comprising $64$ neurons each, a configuration that can provide sufficient parametric capacity to approximate complex mathematical degradation relationship while minimizing the risk of severe overfitting. For the activation function, Sigmoid Linear Unit (SiLU) was utilized which guarantees a smooth, non-vanishing, and infinitely differentiable gradient landscape, which is essential for the stable backpropagation of physical loss terms (Combette, Venaille, & Pustelnik, 2025). To mitigate

the risk of converging into local minima during early training, the weights of the layers are initialized using a Xavier normal distribution with a gain parameter of $0.5$ to maintain minimal pre-activation variance. Finally, the output layer has a dimension of $1$, yielding the predicted scalar sensor observation, $Z$. The architecture of the model can be visualized as in Figure 1.

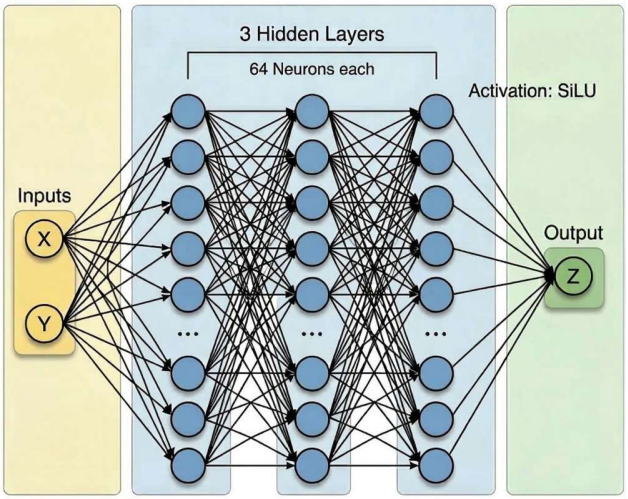


Figure 1. MLP measurement model architecture.

### 2.3. Physics Informed Machine Learning

While the data-driven architecture provides necessary capacity to capture complex relationships, it encounters under-training leading to lower performance when limited to moderate amounts of data. Integrating physics knowledge into data-driven parameter estimation ensures resulting solutions converge robustly while remaining physically meaningful. Varying industries can apply diverse range of physics domains. In this paper implementation, domain expertise is incorporated directly based on expert-defined prior distributions for the parameters: the deterministic drift coefficients $(\mu_x, \mu_y)$ and the stochastic diffusion coefficients $(\sigma_x, \sigma_y)$. Establish realistic bounding intervals for these parameters, domain experts leverage sources such as historical run-to-failure data, empirical wear models, manufacturer specifications, known environmental fluctuations, etc. By integrating prior physics, state estimation methods are guided toward a stable and efficient convergence, ensuring that the optimization process captures the underlying physical trends rather than overfitting to transient measurement anomalies.

The optimization process of the framework utilizes a composite loss function ($\mathcal{L}_{total}$) that regularizes both the observation network and the underlying physical parameters. To prevent catastrophic forgetting and ensure global trajectory consistency, the framework abandons the single-step update in favor of a full-window replay mechanism. The total loss combines data-driven and physics-informed objectives:

$$\mathcal{L}_{total} = \mathcal{L}_{NNL} + \mathcal{L}_{ancho} + \mathcal{L}_{mono} + \mathcal{L}_{grad} + \mathcal{L}_{curv} + \mathcal{L}_{drift} + \mathcal{L}_{diff} + \mathcal{L}_{prior} + \mathcal{L}_{bound} \quad (3)$$

The primary data-driven term is the Negative Log-Likelihood ($\mathcal{L}_{NLL}$), which computes the mean squared error between the network's predictions for the particle swarm $N_p$ and the scalar observation $Z_{obs}$ (Goodfellow, Bengio, & Courville, 2016), formulated as a global trajectory-fitting objective of the full history of scalar observations $Z_{obs\,1:k}$:

$$\mathcal{L}_{NLL} = \sum_{t=0}^{k} \frac{1}{2\sigma_{track}^2} \frac{1}{N_p} \sum_{i=1}^{N_p} \left(MLP(X_t^i, Y_t^i) - Z_{obs_t}\right)^2 \quad (4)$$

This is strictly anchored by $\mathcal{L}_{anchor}$ which forces the network to respect the known initial system conditions:

$$\mathcal{L}_{anchor} = (MLP(X_0, Y_0) - Z_{obs,0})^2 \quad (5)$$

To reflect the irreversible physical principle of damage accumulation, an active Rectified Linear Unit (ReLU)-based monotonicity constraint penalizes any negative gradients across both dimensions over a randomly sampled state space:

$$\mathcal{L}_{mono} = \mathbb{E}[ReLU(-\nabla_X MLP_\theta) + ReLU(-\nabla_Y MLP_\theta)] \quad (6)$$

The severe scale ambiguity inherent in mapping two latent variables to a scalar observation is addressed via a sensitivity band prior ($\mathcal{L}_{grad}$) *and a near-linearity prior* ($\mathcal{L}_{curv}$). $\mathcal{L}_{grad}$ restricts the instantaneous gradients within expert-defined bounds $[g_{lo}, g_{hi}]$ to prevent scale drift, while $\mathcal{L}_{curv}$) penalizes rapid topographic shifts, ensuring that the early-stage calibrated slope logically extrapolates to later degradation stages:

$$\mathcal{L}_{grad} = \mathbb{E}\left[\sum_{\chi \in \{X,Y\}} ReLU(g_{lo} - \nabla_\chi MLP_\theta)^2 + ReLU(\nabla_\chi MLP_\theta - g_{hi})^2\right] \quad (7)$$

$$\mathcal{L}_{curv} = Var(\nabla_X MLP_\theta) + Var(\nabla_Y MLP_\theta) \quad (8)$$

Furthermore, physics-consistency between the observed measurement statistics and the implied stochastic properties is rigorously enforced. The drift loss ($\mathcal{L}_{drift}$) aligns the predicted macroscopic trend with the empirically observed slope, and the diffusion loss ($\mathcal{L}_{diff}$) couples the predicted instantaneous variance with the actual sensor signal volatility:

$$\mathcal{L}_{drift} = (\nabla_X MLP_\theta \cdot \mu_x + \nabla_Y MLP_\theta \cdot \mu_y - Slope_{obs})^2 \quad (9)$$

$$\mathcal{L}_{diff} = \left(\frac{(\nabla_X MLP_\theta)^2 \sigma_x^2 + (\nabla_Y MLP_\theta)^2 \sigma_y^2 + \sigma_{noise}^2}{Var_{obs}} - 1\right)^2 \quad (10)$$

Finally, a Bayesian shrinkage penalty ($\mathcal{L}_{prior}$) and a hard boundary penalty ($\mathcal{L}_{bound}$) gently pull the unobservable components of the physical parameters toward their expert-defined prior means while strictly confining them within physical limits:

$$\mathcal{L}_{prior} = \sum_{\phi \in \Phi} \left( \frac{\phi - \bar{\phi}}{\Delta\phi} \right)^2, \Phi = \{\mu_x, \mu_y, \sigma_x, \sigma_y\} \quad (11)$$

$$\mathcal{L}_{bound} = \sum_{\phi \in \Phi} (ReLU(\phi_{lo} - \phi)^2 + ReLU(\phi - \phi_{hi})^2) \quad (12)$$

where $\bar{\phi}$ represents the expected prior mean, $\Delta\phi$ is the expert-defined prior range width, and $[\phi_{lo}, \phi_{hi}]$defines the physical bounding interval. $\mathcal{L}_{prior}$ gently acts as a soft Bayes estimate to minimize worst-case scale errors, while $\mathcal{L}_{bound}$ decisively prevents parameters from violating fundamental physical constraints during gradient descent.

### 2.4. Sequential Monte Carlo and Gradient Descent Parameter Updates

The dynamic estimation of the latent states $(X, Y)$ and the simultaneous optimization of model parameters are conducted sequentially through a SMC framework (Wang & Gao; 2014), integrated with a dual-stream Adam optimizer (Kingma & Ba, 2014). At each discrete time step $k$, the posterior distribution of the system states is approximated by a population of $N_p$ particles, denoted as $\left\{P_{x,k}^{(i)}, P_{y,k}^{(i)}\right\}_{i=1}^{N_p}$. The estimation cycle begins with multinomial resampling based on the likelihood weights $w_{k-1}^{(i)}$ to mitigate the issue of particle degeneracy. Immediately following resampling, thermodynamic jitter is injected into the spatial coordinates of the swarm, modelled as Gaussian noise $\mathcal{N}(0, \zeta^2)$. This artificial diffusion maintains structural diversity within the particle swarm, preventing dimensional collapse and ensuring that the subsequent gradient-based optimizer explores a sufficiently rich and continuous receptive field within the latent space.

Once the particle swarm is stabilized, the particles follow stochastic kinematic propagation using the neural network's currently estimated physical parameters $\widehat{\Phi} = (\hat{\mu}, \hat{\sigma})$. Each particle $i$ is propagated forward according to the discretized Euler-Maruyama approximation of the Wiener process:

$$P_{x,k}^{(i)} = P_{x,k-1}^{(i)} + \hat{\mu}_x \Delta t + \hat{\sigma}_x \sqrt{\Delta t} \epsilon_{x,k}^{(i)} \quad (13)$$

$$P_{y,k}^{(i)} = P_{y,k-1}^{(i)} + \hat{\mu}_y \Delta t + \hat{\sigma}_y \sqrt{\Delta t} \epsilon_{y,k}^{(i)} \quad (14)$$

where $\mathcal{N}(0,1)$. Concurrent with state propagation, the framework updates the MLP weights $(\theta)$ and the physical parameters $\widehat{\Phi}$ by minimizing the composite loss function $\mathcal{L}_{total}$. This dual-stream optimization utilizes differential learning rates: a base rate $\eta_{base}$ for the MLP to capture non-stationary nonlinearities, and a faster rate $\eta_{params}$ for the physical parameters to track macroscopic degradation trends. In addition, to prevent overfitting and ensure the smoothness of the learned mapping $h_\theta$, a weight decay regularization $\lambda$ is applied to the MLP parameters during the optimization process. This helps stabilize the Jacobian topography, which is critical for the physics-informed constraints.

Standard particle filters frequently suffer from weight collapse and sample impoverishment when the observation function $MLP_\theta$ is dynamically evolving. To robustly handle systematic neural network biases during the online learning phase, the likelihood width of our framework is not rigidly fixed. Instead, it is auto-tuned at each discrete time step via a vectorized bisection search. This tuning guarantees that the Effective Sample Size (ESS) strictly maintains a predefined target fraction of the total population, ensuring structural diversity.

In addition to the observation likelihood, the particle weighting incorporates a soft state marginal prior distribution, formulated as $\mathcal{N}(X_0 + \mu_x t, \lambda\sigma_x^2 t)$. This prior mathematically anchors the unobservable components to their expected physical drift trajectories. It suppresses the degrees of freedom in directions unconstrained by the scalar observation $Z$, preventing the latent states from arbitrarily drifting into non-physical dimensions.

## 3. NUMERICAL EXAMPLE AND RESULTS ANALYSIS

This section presents a numerical validation, conduct analysises on the accuracy and robustness of the proposed method. The following sections detail the simulation setup, the accuracy of the latent state estimation, and a comparative performance benchmark against established model-based filtering approaches.

### 3.1. Simulation Setup

The framework validation is illustrated with the specifications provided in Table 1. The simulation is conducted over $T$ steps with a sampling interval of $\Delta t$. In each trial, parameters of degradation processes $(X, Y)$ and observation model are initialized by uniform distributions. A sensor signal $Z$ is generated to illustrate, using a nonlinear test pattern $h_{test}(X, Y)$, as follows:

$$Z_k = \alpha X_k + \beta Y_k + \gamma X_k Y_k + \alpha_2 X_k^2 + \beta_2 Y_k^2 + v_k \quad (15)$$

- **Interdependency:** The term $\gamma X_k Y_k$ represents the mutual influence of the degradation processes to measurement quality. Decoupling $X$ and $Y$ from a single combined signal evaluates the MLP's ability to organize the latent space.
- **Variable Degradation Rates:** The quadratic terms $\alpha_2 X_k^2$and $\beta_2 Y_k^2$ simulate wear that accelerates during

the final stages of an asset's life. This tests how the monotonicity and variance conservation constraints perform when gradients change over time.

By varying the coefficients $(\alpha, \beta, \gamma, \dots)$ for every run, the framework is required to adapt the mapping $h_\theta$ without prior information regarding the functional form.

Table 1. Simulation, priors and algorithm parameters.

| Category | Parameter | Range/Value |
|---|---|---|
| Common | Timestep size $dt$ | 0.1 |
| | Timesteps $T$ | 3000 |
| Degradation Processes | Drift parameters $\mu_x, \mu_y$ | $\mathcal{U}(2,4)$, $\mathcal{U}(1,2.5)$ |
| | Diffusion parameters $\sigma_x, \sigma_y$ | $\mathcal{U}(2.5,5)$, $\mathcal{U}(1.5,3)$ |
| | Initial states $X_0, Y_0$ | $\mathcal{U}(0,50)$ |
| Test pattern $h_{test}$ | Linear Sensitivity $\alpha, \beta$ | $\mathcal{U}(0.5,1.5)$ |
| | Cross-coupling factor $\gamma$ | $\mathcal{U}(0.0003, 0.0015)$ |
| | Quadratic coefficients $\alpha_2, \beta_2$ | $\mathcal{U}(0.0005, 0.001)$ |
| | Measurement Noise $\sigma_{obs}$ | 0.1 |
| SMC & Optimizer | Number of particles $N_{particles}$ | 5000 |
| | MLP learning rate $\eta_{base}$ | $3 \times 10^{-3}$ |
| | Physical parameter LR $\eta_{params}$ | 0.03 |
| | Weight decay (L2) | $1 \times 10^{-5}$ |
| | Target ESS fraction | 0.5 |
| | Thermodynamic jitter std | 0.02 |
| | State prior variance inflation | 2.5 |
| Joint Loss Weights | Replay loss weight | 1.0 |
| | Initial condition anchor weight | 20.0 |
| | Monotonicity weight | 30.0 |
| | Drift consistency weight | 10.0 |
| | Diffusion consistency weight | 10.0 |
| | Sensitivity-band prior weight | 20.0 |
| | Sensitivity-band bounds | [0.5,2.2] |
| | Near-linearity curvature weight | 8.0 |
| | Bayesian prior weight | 1.0 |
| | Boundary penalty weight | 200.0 |

### 3.2. Accuracy and Consistency Testing

The Coefficient of Determination $(R^2)$ is used for evaluation, measuring how well the estimated states $(\hat{X}, \hat{Y})$ and predicted observations $(\hat{Z})$ match the true values $(X_{true}, Y_{true}, Z_{true})$. Figure 2 illustrates a tracking performance for a common representative trial. The hybrid method decouples well the interacting degradation modes, achieving $R^2$ scores $> 0.9$ for all variables $(X, Y, Z)$.

Consistency testing was also conducted over 50 independent initializations, with results summarized in Table 3 and Figure 3. The colored boxes in Figure 3 represent the interquartile range (IQR), containing the middle $50\%$ of the data, while the horizontal line within each box indicates the median performance. Although the majority of trials converge to high-accuracy solutions, minor occasional outliers (marked as open white dots) exhibit lower $R^2$ values, particularly for $Y$ estimation. These instances primarily occur when the MLP's stochastic weights fall into local minima or when measurement noise temporarily obscures the underlying physical monotonicity.

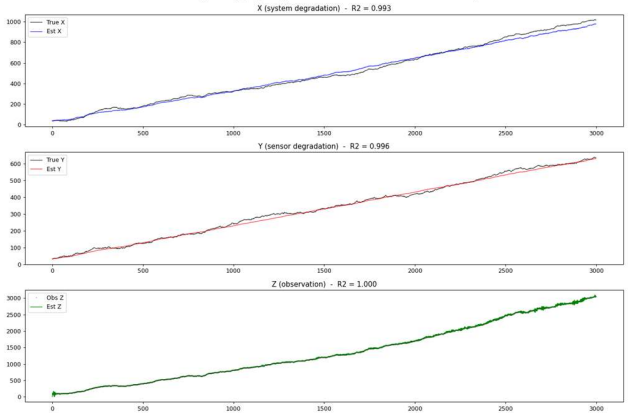


Figure 2. State estimation inference visualization.

Table 3. Statistical consistency $R^2$ analysis across 50 initializations.

| Variable | *Mean* | *Med* | *Min* | *Max* |
|---|---|---|---|---|
| Latent State *X* | 0.8757 | 0.9425 | −0.5852 | 0.9995 |
| Latent State *Y* | 0.7020 | 0.8943 | −3.3897 | 0.9970 |
| Estimated *Z* | 0.9993 | 0.9994 | 0.9988 | 0.9998 |

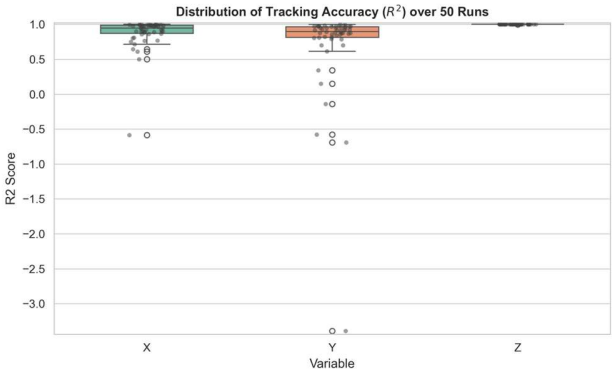


Figure 3. Tracking accuracy distribution of the proposed hybrid method over 50 initializations.

### 3.3. Comparative Analysis

To evaluate the effectiveness of the Proposed Hybrid Method, its tracking accuracy $(R^2)$ is benchmarked against established model-based approaches: the Model-based Particle Filter (PF), Particle Gibbs with Backward Sampling (PGBS), and Extended Kalman Filter (EKF). These three traditional baselines operate under the assumption that the underlying mathematical model is fully known, attempting to estimate the system states $(X, Y, Z)$ and unknown parameters simultaneously using their respective probabilistic filtering frameworks.

As summarized in Table 4, the proposed hybrid method demonstrates competitive performance, achieving results on par with the best-performing traditional baseline, the Model-based PF. For the state $X$, the hybrid approach yields a mean $R^2$ of $0.8757$, performing comparably to the Model-based PF $(0.8458)$. Notably, for the coupled latent state $Y$, the hybrid method actually outperforms all baselines $(0.7020$ vs. PF's $0.5028)$. Figure 4 illustrates that the tracking accuracy distribution for the Model-based PF, despite having full prior knowledge, exhibits a wide variance in estimating $X$ and $Y$. Meanwhile, Figure 3 in section 3.2 shows that the proposed hybrid method maintains a tighter interquartile range with fewer extreme outliers.

Table 4. Statistical consistency $R^2$ comparative analysis across 50 initializations.

| Method | $\overline{R}_X$ | $Med\ R_X$ | $Min\ R_X$ | $Max\ R_X$ |
|---|---|---|---|---|
| Proposed Hybrid | $\mathbf{0.8757}$ | $\mathbf{0.9425}$ | $-0.5852$ | $\mathbf{0.9995}$ |
| Model based PF | $\mathbf{0.8458}$ | $\mathbf{0.9111}$ | $\mathbf{0.2503}$ | $\mathbf{0.9942}$ |
| Model based PGBS | $0.6709$ | $\mathbf{0.8809}$ | $-0.5940$ | $\mathbf{0.9996}$ |
| Model based EKF | $0.6779$ | $0.8267$ | $-1.2276$ | $\mathbf{0.9879}$ |
| | | | | |
| Method | $\overline{R}_Y$ | $Med\ R_Y$ | $Min\ R_Y$ | $Max\ R_Y$ |
| Proposed Hybrid | $\mathbf{0.7020}$ | $\mathbf{0.8943}$ | $-3.3897$ | $\mathbf{0.9970}$ |
| Model based PF | $0.5028$ | $0.7177$ | $-2.6745$ | $\mathbf{0.9937}$ |
| Model based PGBS | $0.2068$ | $0.6901$ | $-3.8566$ | $\mathbf{0.9961}$ |
| Model based EKF | $-0.1685$ | $0.5933$ | $-7.6351$ | $\mathbf{0.9960}$ |

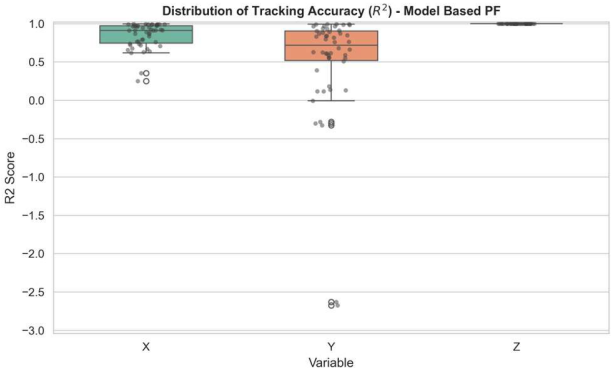


Figure 4. Tracking accuracy distribution of Model-based PF method over 50 initializations.

## 4. CONCLUSION

This paper presented a novel hybrid prognostic framework addressing the challenge of joint degradation state estimation in environments where both the primary asset and monitoring sensors undergo deterioration and compromise measurement integrity. By integrating a Multilayer Perceptron as a dynamic observation model within Sequential Monte Carlo (SMC) inference, the proposed methodology applies Physics-Informed Machine Learning (PIML) principles enforcing strict, domain-aware physical constraints such as monotonic damage progression and variance consistency, guides the neural network to interpret measurement function. Numerical simulations demonstrated that this hybrid approach is promising, achieving competitive tracking accuracy on par with established model-based baselines methods. Occasional susceptibility to local minima of the composite loss function remains a limitation. Future work can effectively mitigate this by integrating more specific industry-related physics priors, pretraining the learnable measurement model with available labelled data and exploring offline batch processing strategies to enhance convergence stability.

## REFERENCES

Cai, B., Liu, Y., & Xie, M. (2017). A dynamic-Bayesian-network-based fault diagnosis methodology considering transient and intermittent faults. *IEEE Transactions on Automation Science and Engineering*, 14(1), 276–285.

Cancelliere, F., Girard, S., Bourinet, J. M., & Broggi, M. (2023). A grey-box approach for the prognostic and health management of lithium-ion batteries. *Proceedings of the Annual Conference of the PHM Society* (pp. 1-8), New York, NY.

Chen, Z., Yang, C., Peng, T., Dan, H., Li, C., & Gui, W. (2018). A cumulative canonical correlation analysis-based sensor precision degradation detection method. *IEEE Transactions on Industrial Electronics*, 66(8), 6321-6330.

Combette, A., Venaille, A., & Pustelnik, N. (2025). A new initialisation to control gradients in sinusoidal neural network. *arXiv preprint arXiv:2512.06427*.

Csuzdi, D., Bécsi, T., & Törő, O. (2026). Physics-informed neural particle flow for the Bayesian update step. *arXiv preprint arXiv:2602.23089*.

Farea, A., Yli-Harja, O., & Emmert-Streib, F. (2024). Understanding physics-informed neural networks: Techniques, applications, trends, and challenges. *AI*, 5(3), 1534-1557.

Fink, O., Nejjar, I., Sharma, V., Niresi, K. F., Sun, H., Dong, H., ... & Zhao, M. (2025). From physics to machine learning and back: Part II-learning and observational bias in PHM. *arXiv preprint arXiv:2509.21207*.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. Cambridge, MA: MIT Press.

Guo, G., You, H., Li, C., Tang, Z., & Li, O. (2023). A physics-informed automatic neural network generation framework for emerging device modeling. *Micromachines*, 14(6), 1150.

Hachem, H., Vu, H. C., & Fouladirad, M. (2024). Different methods for RUL prediction considering sensor degradation. *Reliability Engineering & System Safety*, 243, 109897.

Jiang, L., Djurdjanovic, D., Ni, J., & Lee, J. (2006). Sensor degradation detection in linear systems. *Engineering Asset Management: Proceedings of the 1st World Congress on Engineering Asset Management (WCEAM)* (pp. 1252-1260), July 11–14, London, UK.

Jung, D. (2022). Automated design of grey-box recurrent neural networks for fault diagnosis using structural models and causal information. *Learning for Dynamics and Control Conference* (pp. 8-20), May, PMLR.

Kahle, W., & Lehmann, A. (2009). The Wiener process as a degradation model: Modeling and parameter estimation. In W. Kahle & A. Lehmann (Eds.), *Advances in degradation modeling: Applications to reliability, survival analysis, and finance* (pp. 127-146). Boston, MA: Birkhäuser Boston.

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Li, H., Si, X., Zhang, Z., & Li, T. (2024). A critical review on prognostics for stochastic degrading systems under big data. *Fundamental Research*.

Li, J., & Ying, Y. (2018). A method to improve the robustness of gas turbine gas-path fault diagnosis against sensor faults. *IEEE Transactions on Reliability*, 67(1), 3–12.

Li, Z., Zhang, Y., & Wang, C. (2013). A sensor-driven structural health prognosis procedure considering sensor performance degradation. *Structure and Infrastructure Engineering*, 9(8), 764-776.

Liu, B., Do, P., Iung, B., & Xie, M. (2019). Stochastic filtering approach for condition-based maintenance considering sensor degradation. *IEEE Transactions on Automation Science and Engineering*, 17(1), 177-190.

Liu, X., Matias, J., Jäschke, J., & Vatn, J. (2022). Gibbs sampler for noisy transformed gamma process: Inference and remaining useful life estimation. *Reliability Engineering & System Safety*, 217, 108084.

Ma, S. L., Jiang, S. F., & Li, J. (2019). Structural damage detection considering sensor performance degradation and measurement noise effect. *Measurement*, 131, 431-442.

Mehdizadeh, M., John, S., Wang, C. H., Ghorbani, K., & Rowe, W. S. (2012). Distinguishing the degradation of the interdigital piezoelectric fibre transducers from structural damage in multifunctional composites. *Smart Materials, Adaptive Structures and Intelligent Systems* (pp. 869-877), September 19-21, Stone Mountain, Georgia, USA.

Mo, H., Wang, W., Xie, M., & Xiong, J. (2017). Modeling and analysis of the reliability of digital networked control systems considering networked degradations. *IEEE Transactions on Automation Science and Engineering*, 14(3), 1491–1503.

Tang, J., Jiang, M., & Mao, Y. (2025). Reliability assessment for multivariate degradation system based on uncertainty and chatterjee correlation coefficient. *Systems*, 13(11), 953.

Wang, P., & Gao, R. X. (2014). Particle filtering-based system degradation prediction applied to jet engines. *Annual Conference of the PHM Society* (Vol. 6, No. 1), September.

Yoo, M., Kim, T., Yoon, J. T., Kim, Y., Kim, S., & Youn, B. D. (2020). A resilience measure formulation that considers sensor faults. *Reliability Engineering & System Safety*, 199, 106393. doi:10.1016/j.ress.2019.106393.

Zhang, J. X., Si, X. S., Du, D. B., & Hu, C. H. (2018). Specification analysis of the deteriorating sensor for required lifetime prognostic performance. *Microelectronics Reliability*, 85, 71-83.