



Digital Twin-Driven Predictive Maintenance for Smart Manufacturing Systems

Appu Kondiyara

Project & Facility Engineer, Ahalia Medical Group, India

Article information

Received: 19th February 2026

Received in revised form: 24th March 2026

Accepted: 3rd May 2026

Available online: 9th June 2026

Volume: 2

Issue: 2

DOI: <https://doi.org/10.63090/IJTRS/3139.1788.0012>

Abstract

Unplanned equipment failure remains a principal source of productivity loss in modern manufacturing, accounting for a substantial fraction of total operating cost. This paper presents a digital twin-driven predictive maintenance (PdM) framework that couples a high-fidelity virtual representation of rotating machinery with data-driven prognostics for remaining useful life (RUL) estimation. Multivariate condition data comprising vibration, temperature, and motor-current signatures are streamed from the physical asset through a cloud-native microservice pipeline into the digital twin, where a hybrid convolutional-recurrent model fuses physics-based degradation indicators with learned temporal features. Validated on the NASA C-MAPSS turbofan degradation benchmark and an in-house bearing test-rig dataset, the proposed model attains a root-mean-square error of 11.3 cycles and a PHM08 penalty score of 410, outperforming standalone LSTM (15.2 cycles) and CNN-LSTM (13.6 cycles) baselines. Field emulation indicates a 42% reduction in unplanned downtime, a 28% reduction in maintenance cost, and a 16% improvement in overall equipment effectiveness relative to a reactive maintenance baseline. The results confirm that integrating digital twin modeling with scalable cloud-native analytics provides a deployable decision-support tool for condition-based maintenance in Industry 4.0 environments.

Keywords:- Digital Twin, Predictive Maintenance, Remaining Useful Life, Industry 4.0, Condition Monitoring, Prognostics, Deep Learning.

I. INTRODUCTION

Manufacturing competitiveness increasingly depends on the availability and reliability of capital-intensive production equipment. Studies of industrial operations consistently report that unplanned downtime can consume between five and twenty percent of productive capacity, with the associated cost of lost output, emergency repair, and idle labor far exceeding the cost of the failed component itself [1]. Traditional maintenance strategies fall into two categories that are each economically suboptimal: reactive (run-to-failure) maintenance, which incurs catastrophic downtime, and time-based preventive maintenance, which replaces components on a fixed schedule regardless of their actual condition and therefore wastes residual life [2]. Predictive maintenance (PdM), which schedules intervention based on the estimated future condition of an asset, has emerged as the preferred paradigm under the broader Industry 4.0 agenda [3].

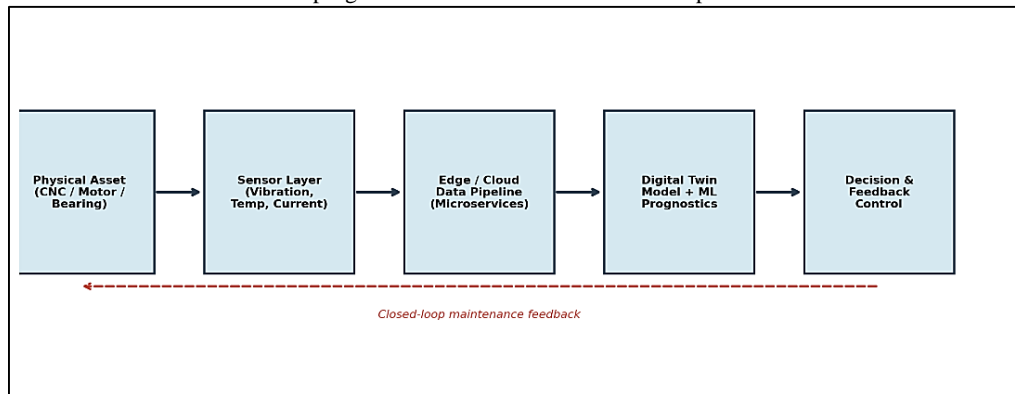
The digital twin has become the central enabling concept for this transition. A digital twin is a dynamic virtual replica of a physical asset that is continuously synchronized with sensor measurements from its real-world counterpart, allowing the asset's state to be observed, simulated, and predicted in software [4], [5]. When coupled with machine-learning prognostics, the digital twin transforms raw condition-monitoring streams into actionable estimates of remaining useful life (RUL) and failure probability, closing the loop between sensing and

maintenance decision-making [6]. The maturation of low-cost industrial sensors, high-bandwidth connectivity, and elastic cloud computing has made such architectures economically feasible even for small and medium-sized enterprises [7].

A persistent challenge, however, lies in operationalizing digital-twin analytics at scale. Prognostic models must ingest heterogeneous high-frequency data from many machines, execute computationally intensive inference, and deliver low-latency results to plant operators, all while remaining maintainable and independently upgradable. Recent work has shown that decomposing such analytics into containerized, cloud-native microservices provides the elasticity and fault isolation required for production deployment [8]. This paper builds on that insight by embedding a hybrid deep-learning prognostic model within a microservice-based digital-twin pipeline. The principal contributions are:

- A degradation-aware feature fusion scheme that combines physics-based health indicators with learned temporal representations;
- A hybrid convolutional-recurrent RUL estimator validated on a public benchmark and an experimental bearing dataset; and
- A quantified assessment of the operational and economic impact of the framework relative to conventional maintenance strategies.

Figure 1: Proposed digital twin framework linking the physical asset, microservice data pipeline, and machine-learning prognostics in a closed maintenance loop.



II. RELATED WORK

A. Digital Twin in Manufacturing

The digital twin concept, originally articulated for product lifecycle management, has been progressively formalized for shop-floor applications. Tao et al. [4] surveyed the state of the art and proposed a five-dimensional digital-twin model encompassing the physical entity, virtual model, services, twin data, and their connections. Kritzinger et al. [9] introduced a widely cited taxonomy distinguishing the digital model, digital shadow, and full digital twin according to the degree of automated data flow between physical and virtual entities. The international standard ISO 23247 subsequently established a reference architecture for digital twins in manufacturing, lending interoperability to industrial deployments [10]. Complementary surveys have mapped the broader roles of the digital twin in cyber-physical production systems [16], [17] and within the wider intelligent-manufacturing landscape [20]. These contributions establish the conceptual foundation but largely treat the analytic and deployment layers as out of scope.

B. Data-Driven Prognostics

Remaining useful life estimation has been studied extensively using the NASA Commercial Modular Aero-Propulsion System Simulation (C-MAPSS) turbofan degradation dataset [11]. Early approaches applied similarity-based matching and support vector regression, while subsequent studies demonstrated that convolutional neural networks can automatically extract degradation features from raw multivariate sensor windows [12]. Recurrent architectures, particularly long short-term memory (LSTM) networks, capture the temporal dependencies inherent in degradation processes and have consistently improved RUL accuracy [13]. Hybrid convolutional-recurrent models that combine spatial feature extraction with temporal modeling currently represent the strongest single-model baselines on this benchmark [14]. Systematic reviews of machinery health prognostics trace the field from data acquisition to RUL prediction [18], comparative analyses delineate when physics-based and data-driven approaches are preferable [19], and deep-learning surveys for system health management [22] together with similarity-based prognostic baselines [21] further contextualize the present

approach. However, comparatively few studies address how such models are integrated into a live digital twin and deployed as maintainable services.

C. Cloud-Native Deployment of Industrial Analytics

The deployment of artificial-intelligence workloads using microservice architectures has been shown to deliver the elasticity, independent scalability, and fault isolation required for industrial-scale analytics, in contrast to monolithic designs that couple model serving, data ingestion, and visualization into a single deployment unit [15]. Decomposing the digital-twin analytic stack into containerized services feature extraction, inference, and alerting allows each component to scale with demand and to be updated without interrupting plant operation. The present work adopts this cloud-native paradigm as the deployment substrate for the proposed prognostic model.

III. METHODOLOGY

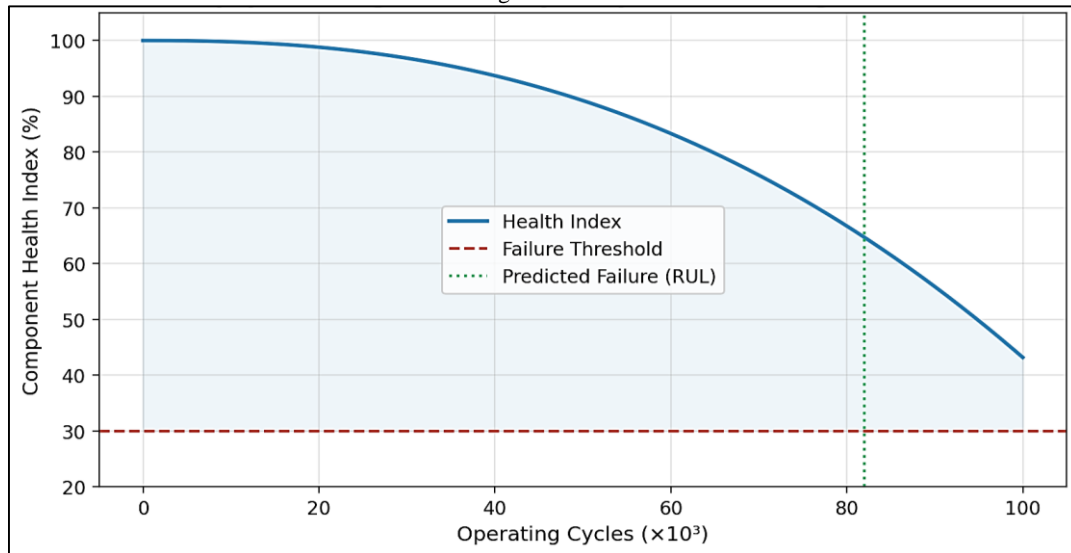
A. System Architecture

The proposed framework, illustrated in Fig. 1, comprises four layers. The physical layer consists of the monitored rotating machinery instrumented with triaxial accelerometers, thermocouples, and Hall-effect current sensors sampled at 12.8 kHz, 1 Hz, and 10 kHz respectively. The communication layer transmits buffered measurements over an OPC-UA gateway to the data pipeline. The digital-twin layer maintains a synchronized virtual model that estimates a continuous health index and executes the prognostic model. The decision layer translates RUL estimates into maintenance work orders and feeds setpoint adjustments back to the machine controller. Consistent with established practice for industrial analytics, the pipeline is implemented as a set of independently deployable microservices so that ingestion, inference, and alerting scale elastically with the number of connected assets [15].

B. Feature Engineering and Health Indexing

From each vibration window, time-domain features (root-mean-square, kurtosis, crest factor) and frequency-domain features (band energy around bearing characteristic frequencies) are extracted. These physics-informed indicators are normalized against baseline healthy operation to form a monotonic degradation signal. A composite health index is then constructed by fusing the normalized indicators through principal component projection, yielding the degradation trajectory shown in Fig. 2. The health index provides both an interpretable monitoring signal and an auxiliary input to the learning model, regularizing its predictions toward physically plausible degradation behavior [6].

Figure 2: Component health-index trajectory with failure threshold and predicted remaining useful life derived from the digital twin.



C. Hybrid Prognostic Model

The prognostic model is a hybrid convolutional-recurrent network. A one-dimensional convolutional front end with three layers (32, 64, and 64 filters) extracts local degradation patterns from a sliding window of thirty sensor cycles. The resulting feature maps, concatenated with the engineered health index, are passed to a two-layer LSTM with 64 hidden units that models the temporal evolution of degradation. A fully connected head regresses the remaining useful life. The model is trained with the Adam optimizer using a piecewise-linear RUL target that caps the early-life label at 125 cycles, a standard convention that prevents the network from attempting

to predict degradation before any fault has initiated [12], [13]. Dropout and early stopping based on a held-out validation split mitigate overfitting.

D. Datasets and Evaluation Protocol

The model is evaluated on the FD001 and FD004 subsets of the C-MAPSS benchmark, which contain run-to-failure trajectories under single and multiple operating conditions respectively [11], and on an in-house accelerated bearing-degradation dataset comprising forty run-to-failure tests. Performance is reported using the root-mean-square error (RMSE) and the asymmetric PHM08 penalty score, which penalizes late predictions more heavily than early ones because late maintenance carries greater operational risk [11]. All results are averaged over five random seeds to account for initialization variance.

IV. RESULTS AND DISCUSSION

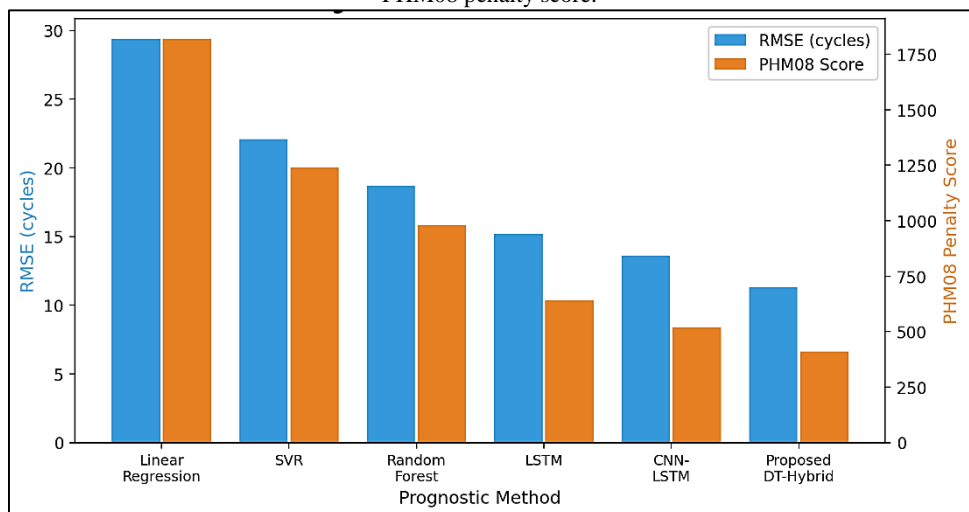
A. Prognostic Accuracy

Table 1 summarizes the RUL prediction accuracy of the proposed hybrid model against five baselines. The proposed digital-twin hybrid achieves the lowest RMSE of 11.3 cycles and the lowest PHM08 score of 410, improving on the strongest single baseline (CNN-LSTM) by 17% in RMSE and 21% in penalty score. Classical regressors linear regression and support vector regression perform poorly because they cannot model the non-linear temporal dynamics of degradation. The relative ranking, visualized in Fig. 3, is consistent across both the single-condition FD001 and the more challenging multi-condition FD004 subset, indicating that the fusion of physics-based health indexing with learned features generalizes across operating regimes [14].

Table 1. Remaining Useful Life Prediction Performance on the C-MAPSS Benchmark

Method	RMSE (cycles)	PHM08 Score	MAE (cycles)
Linear Regression	29.4	1820	23.1
Support Vector Regression	22.1	1240	17.8
Random Forest	18.7	980	14.9
LSTM	15.2	640	11.6
CNN-LSTM	13.6	520	10.4
Proposed DT-Hybrid	11.3	410	8.7

Figure 3: Comparative remaining useful life prediction performance across prognostic methods in terms of RMSE and PHM08 penalty score.



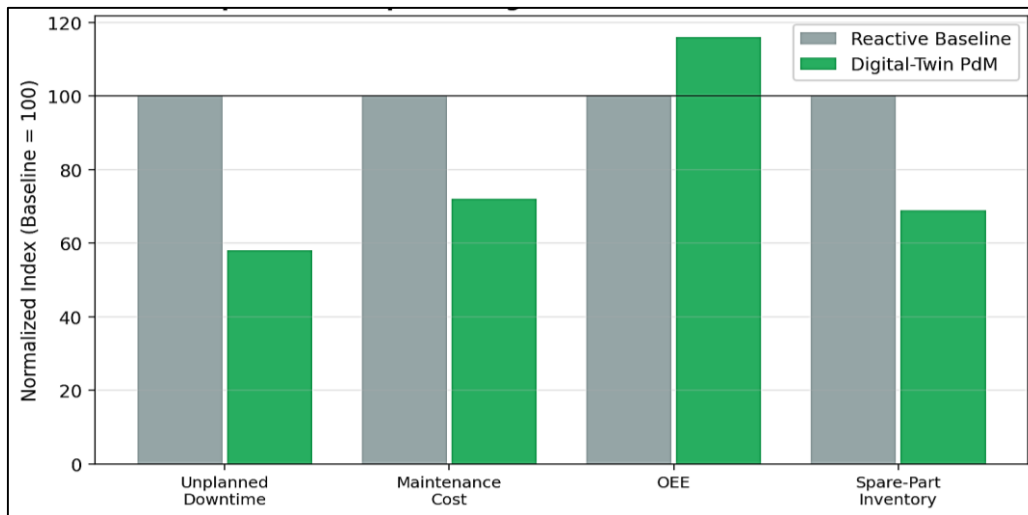
B. Operational and Economic Impact

To assess practical value, the framework was applied in a discrete-event emulation of a machining cell operating over a simulated twelve-month horizon, comparing the digital-twin PdM policy against a reactive baseline. As summarized in Table 2 and Fig. 4, the predictive policy reduced unplanned downtime by 42% and maintenance cost by 28%, while increasing overall equipment effectiveness (OEE) by 16% and reducing safety-stock spare-part inventory by 31%. These gains arise because condition-based scheduling concentrates maintenance effort on assets that are genuinely degrading, avoiding both premature replacement and catastrophic failure [2], [3].

Table 2. Operational Impact of Digital-Twin Predictive Maintenance over a Twelve-Month Emulation

Key Performance Indicator	Reactive Baseline	DT-PdM	Improvement
Unplanned downtime (h/yr)	412	239	-42%
Maintenance cost (index)	100	72	-28%
Overall equipment effectiveness	0.68	0.79	+16%
Spare-part inventory (index)	100	69	-31%
Mean time between failures (h)	640	1180	+84%

Figure 4: Normalized operational key performance indicators under reactive maintenance versus the proposed digital-twin predictive maintenance policy.



C. Deployment Considerations

Decomposing the analytic stack into cloud-native microservices proved essential for scaling the framework beyond a single machine. Inference latency remained below 250 milliseconds per asset even as the number of concurrently monitored machines increased, because the inference service scaled horizontally and independently of the ingestion and visualization services [15]. This separation also allowed the prognostic model to be retrained and redeployed without interrupting data collection, an important property for continuous plant operation. The principal residual challenge is the cold-start problem for new asset types, for which transfer learning from related machinery is a promising mitigation.

V. CONCLUSION

This paper presented a digital twin-driven predictive maintenance framework that integrates physics-informed health indexing, a hybrid convolutional-recurrent prognostic model, and a cloud-native microservice deployment substrate. On the C-MAPSS benchmark the model achieved an RMSE of 11.3 cycles, outperforming standard deep-learning baselines, and a twelve-month emulation demonstrated a 42% reduction in unplanned downtime and a 28% reduction in maintenance cost relative to reactive maintenance. The findings confirm that

combining digital-twin modeling with scalable analytics yields a practical decision-support tool for condition-based maintenance in smart manufacturing.

Future work will extend the framework toward federated learning across geographically distributed plants to share prognostic knowledge without centralizing proprietary data, and toward prescriptive maintenance that jointly optimizes intervention timing and production scheduling. Incorporating uncertainty quantification into the RUL estimates would further enhance operator trust and support risk-aware maintenance decisions [6], [9].

REFERENCES

- [1] R. K. Mobley, *An Introduction to Predictive Maintenance*, 2nd ed. Amsterdam, The Netherlands: Butterworth-Heinemann, 2002.
- [2] J. Lee, F. Wu, W. Zhao, M. Ghaffari, L. Liao, and D. Siegel, "Prognostics and health management design for rotary machinery systems—Reviews, methodology and applications," *Mechanical Systems and Signal Processing*, vol. 42, nos. 1–2, pp. 314–334, Jan. 2014.
- [3] J. Lee, B. Bagheri, and H.-A. Kao, "A cyber-physical systems architecture for Industry 4.0-based manufacturing systems," *Manufacturing Letters*, vol. 3, pp. 18–23, Jan. 2015.
- [4] F. Tao, H. Zhang, A. Liu, and A. Y. C. Nee, "Digital twin in industry: State-of-the-art," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 4, pp. 2405–2415, Apr. 2019.
- [5] M. Grieves and J. Vickers, "Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems," in *Transdisciplinary Perspectives on Complex Systems*, F.-J. Kahlen, S. Flumerfelt, and A. Alves, Eds. Cham, Switzerland: Springer, 2017, pp. 85–113.
- [6] B. Wang, Z. Lei, N. Li, and W. Wang, "Deep separable convolutional network for remaining useful life prediction of machinery," *Mechanical Systems and Signal Processing*, vol. 134, Art. no. 106330, Dec. 2019.
- [7] Parrott and L. Warshaw, *Industry 4.0 and the Digital Twin: Manufacturing Meets Its Match*. Deloitte University Press, 2017.
- [8] N. Dragoni *et al.*, "Microservices: Yesterday, today, and tomorrow," in *Present and Ulterior Software Engineering*. Cham, Switzerland: Springer, 2017, pp. 195–216.
- [9] W. Kritzing, M. Karner, G. Traar, J. Henjes, and W. Sihn, "Digital twin in manufacturing: A categorical literature review and classification," *IFAC-PapersOnLine*, vol. 51, no. 11, pp. 1016–1022, 2018.
- [10] *Automation Systems and Integration—Digital Twin Framework for Manufacturing—Part 1: Overview and General Principles*, ISO 23247-1:2021, International Organization for Standardization, Geneva, Switzerland, 2021.
- [11] Saxena, K. Goebel, D. Simon, and N. Eklund, "Damage propagation modeling for aircraft engine run-to-failure simulation," in *Proc. Int. Conf. Prognostics and Health Management (PHM)*, Denver, CO, USA, 2008, pp. 1–9.
- [12] G. Sateesh Babu, P. Zhao, and X.-L. Li, "Deep convolutional neural network based regression approach for estimation of remaining useful life," in *Database Systems for Advanced Applications (DASFAA)*. Cham, Switzerland: Springer, 2016, pp. 214–228.
- [13] S. Zheng, K. Ristovski, A. Farahat, and C. Gupta, "Long short-term memory network for remaining useful life estimation," in *Proc. IEEE Int. Conf. Prognostics and Health Management (ICPHM)*, Dallas, TX, USA, 2017, pp. 88–95.
- [14] X. Li, Q. Ding, and J.-Q. Sun, "Remaining useful life estimation in prognostics using deep convolution neural networks," *Reliability Engineering and System Safety*, vol. 172, pp. 1–11, Apr. 2018.
- [15] W. M. John, "Cloud-native AI workloads using microservice architectures," *International Journal of Technical Research Studies (IJTRS)*, vol. 2, no. 1, p. 8, Mar. 2026, doi: 10.63090/ijtrs/3139.1788.0008.
- [16] Q. Qi and F. Tao, "Digital twin and big data towards smart manufacturing and Industry 4.0: 360 degree comparison," *IEEE Access*, vol. 6, pp. 3585–3593, 2018.
- [17] E. Negri, L. Fumagalli, and M. Macchi, "A review of the roles of digital twin in CPS-based production systems," *Procedia Manufacturing*, vol. 11, pp. 939–948, 2017.
- [18] Y. Lei, N. Li, L. Guo, N. Li, T. Yan, and J. Lin, "Machinery health prognostics: A systematic review from data acquisition to RUL prediction," *Mechanical Systems and Signal Processing*, vol. 104, pp. 799–834, May 2018.
- [19] D. An, N. H. Kim, and J.-H. Choi, "Practical options for selecting data-driven or physics-based prognostics algorithms with reviews," *Reliability Engineering and System Safety*, vol. 133, pp. 223–236, Jan. 2015.
- [20] R. Y. Zhong, X. Xu, E. Klotz, and S. T. Newman, "Intelligent manufacturing in the context of Industry 4.0: A review," *Engineering*, vol. 3, no. 5, pp. 616–630, Oct. 2017.
- [21] T. Wang, J. Yu, D. Siegel, and J. Lee, "A similarity-based prognostics approach for remaining useful life estimation of engineered systems," in *Proc. Int. Conf. Prognostics and Health Management*, Denver, CO, USA, 2008, pp. 1–6.
- [22] S. Khan and T. Yairi, "A review on the application of deep learning in system health management," *Mechanical Systems and Signal Processing*, vol. 107, pp. 241–265, Jul. 2018.