

Industrial Deployment of a Decision-Oriented Digital Twin for Predictive Maintenance of a Reverse Osmosis System: An End-To-End Case Study in a Beverage Filling Plant



Soufiane Embarki^{1*}, Yousra El Kihel², Bachir El Kihel¹

¹Laboratory of Industrial Engineering and Seismic Engineering, Mohammed First University, Oujda 60000, Morocco

²Laboratoire d'Innovation Numérique pour les Entreprises et les Apprentissages au service de la Compétitivité des Territoires, LINEACT-CESI, Bordeaux 33000, France

Corresponding Author Email: s.embarki@ump.ac.ma

Copyright: ©2026 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/jesa.590321>

ABSTRACT

Received: 18 January 2026

Revised: 20 March 2026

Accepted: 27 March 2026

Available online: 31 March 2026

Keywords:

digital twin, predictive maintenance, vibration analysis, reverse osmosis, interpretable alarm logic, gated recurrent unit, industry 4.0

Digital twin (DT) is increasingly proposed as an enabler of predictive maintenance, yet few studies document their full industrial deployment in continuous-process settings. This paper reports the end-to-end implementation and empirical evaluation of a decision-oriented DT for vibration-based predictive maintenance of an industrial reverse osmosis (RO) system in a beverage filling plant. The deployed architecture combines IO-Link (Input/Output Link) vibration sensing, historian-based time-series storage, asset modelling, supervision dashboards, alarm rules aligned with ISO 20816-3:2022, alarm-triggered spectral diagnosis, and CMMS-connected maintenance execution. Multi-horizon forecasting models were benchmarked using five-minute V-RMS vibration data, and a direct gated recurrent unit (GRU) model was selected because it provided the best compromise between predictive accuracy, horizon stability, and low-resource deployment constraints. The operational evaluation compared a pre-go-live period (January–May 2025) with a post-go-live period (June–December 2025) using auditable maintenance logs. After go-live, the share of failure-free operating days increased from 72.8% to 84.5%, while unplanned downtime decreased and system availability improved. These results should be interpreted as improvements associated with deployment rather than as strict causal effects, because the study used a quasi-experimental before-and-after design without a contemporaneous control group. The contribution of the paper lies in documenting how engineering-grounded alarm logic, low-latency forecasting, and maintenance workflow integration can be combined into a decision-oriented DT for a critical water-treatment asset. The findings provide implementation evidence relevant to process industries seeking operationally credible pathways from monitoring pilots to maintainable DT.

1. INTRODUCTION

Industry 4.0 has accelerated the convergence of instrumentation, connectivity, analytics, and cyber-physical systems in industrial operations. One of the most promising applications of this convergence is predictive maintenance, where continuously acquired field data are converted into actionable signals that can reduce unscheduled downtime and improve asset availability [1, 2]. In practice, however, many industrial initiatives still stall between pilot monitoring and sustained operational use. This gap is particularly visible in continuous-process environments, where maintenance interventions must be integrated into tightly constrained production systems and where the cost of a false alarm may be almost as important as the cost of a missed failure.

The digital twin (DT) has emerged as a useful conceptual and operational bridge between physical assets and predictive models. In its mature form, a DT is not merely a digital replica: it combines real-time synchronization, contextual data management, simulation or prediction, and decision support

across the asset life cycle [3, 4]. Yet the deployment of a decision-oriented DT remains technically demanding. Reliable sensing, traceable data pipelines, low-latency analytics, interpretable outputs, and workflow-compatible interfaces must all function together if a maintenance team is to trust and use the system in daily operations [5, 6].

The present study addresses these deployment challenges through an industrial case study on a reverse osmosis (RO) system used for water treatment in a beverage filling plant. This asset contains critical rotating components whose degradation directly affects production continuity and water supply quality. The paper focuses on vibration-based condition monitoring because vibration remains one of the most established and information-rich techniques for diagnosing faults in rotating machinery [7, 8]. Instead of proposing a new theoretical framework, the paper documents a complete implementation pathway, from field instrumentation to maintenance execution, and evaluates the observed operational changes after go-live.

This work is guided by three research questions. RQ1 asks

how a DT for vibration-based maintenance can be architected so that it remains technically coherent and operationally usable in a continuous-process environment. RQ2 asks which forecasting approach offers the best practical trade-off between predictive quality, false-alarm control, latency, and memory footprint under edge-oriented deployment constraints. RQ3 asks whether the deployment is associated with measurable improvements in auditable operational indicators once the system enters routine use.

The article is structured as follows. Section 2 positions the work within the literature on DT and vibration-based predictive maintenance and identifies the specific gap addressed by the case. Section 3 presents the industrial context, system architecture, instrumentation, analytics, deployment logic, and evaluation protocol. Section 4 reports the predictive and operational results. Section 5 discusses the engineering implications, adoption conditions, and validity limits of the findings. Section 6 concludes.

2. LITERATURE BACKGROUND AND RESEARCH POSITIONING

2.1 Digital twins in maintenance systems

The DT concept has broadened considerably since its early

formulations in lifecycle engineering and aerospace systems. Across definitions, three characteristics are recurrent: a structured digital representation of a physical asset, a live data link with the asset, and a capacity to support prediction or decision-making over time, which helpfully distinguish among a digital model, a digital shadow, and a full DT according to the direction and richness of data exchange [3-5]. This distinction is not merely semantic. In maintenance applications, systems described as DT often remain digital shadows because they provide monitoring without closing the loop toward intervention planning or asset management.

Recent reviews confirm the rapid growth of industrial DT research but also show that implementation maturity remains uneven. Fuller et al. [6] highlighted the central roles of interoperability, data quality, and model validation. More recent work emphasizes the need for modular engineering methods, explicit governance, and cloud-to-edge deployment patterns that can survive real operational constraints [9-11]. Maintenance-focused DT studies report promising results, but many remain limited to simulated environments, laboratory demonstrators, or architectures that do not document integration with maintenance execution systems [12, 13].

Table 1 synthesizes the main technological pillars that enable maintenance-oriented DT and clarifies their role in maintenance applications.

Table 1. Representative technological enablers of maintenance-oriented digital twin (DT)

Technological Pillar	Definition	Role in Maintenance-Oriented Digital Twins
Artificial intelligence	Computational methods capable of learning patterns and supporting decision-making from operational data.	Forecasting, anomaly detection, fault classification, and model adaptation
Internet of Things	Sensor and actuator infrastructure interconnected through standardized communication protocols.	Field acquisition, state synchronization, and equipment-level connectivity
Cyber-physical systems	Tight coupling of computational, networking, and physical processes through feedback loops.	Execution environment for monitoring, control, and synchronization between physical and digital entities
Big data	Large, heterogeneous operational datasets requiring dedicated storage and analytics.	Historical learning, trend analysis, and continuous model improvement
Internet of Services	Service-oriented software environment for integrating distributed functionalities.	Interoperability, orchestration of analytics services, and third-party integration
Virtual/augmented reality and 3D simulation	Digital visualization technologies used to represent asset state and context.	Training, remote support, intuitive visualization, and scenario exploration

2.2 Vibration-based predictive maintenance in process industries

Vibration-based maintenance remains a cornerstone of condition monitoring for rotating machinery because it captures information related to imbalance, misalignment, bearing damage, looseness, and hydraulic anomalies. Conventional practice combines global indicators such as RMS vibration velocity with expert interpretation of frequency spectra. This approach is well established, but it does not scale easily when many assets must be monitored continuously and when maintenance teams need early warnings rather than post hoc diagnosis [7, 8]. Data-driven predictive models can extend the anticipation horizon, yet they also introduce new requirements: labelled histories, stable operating contexts, drift management, and explainable alarm criteria [14].

Continuous-process industries introduce additional constraints that are sometimes underrepresented in generic DT discussions. Hygiene requirements, regulatory traceability, frequent operational changes, and the need for offline continuity all influence sensor selection, data architecture, and alarm policy. In such environments, the relevant question is

not only whether a model is accurate, but whether the overall DT remains interpretable, maintainable, and actionable under plant conditions [15].

Table 2. Main paradigms of vibration-based predictive maintenance and their limitations

Approach	Principle	Main Limitations
Normative thresholds	Comparison of global vibration levels with standardized thresholds	Limited specificity to operating conditions; often detects problems relatively late
Expert spectral analysis	Interpretation of FFT spectra by experienced analysts	Low scalability, dependence on specialist expertise, and higher cost
Predictive models	Use of learned patterns for anomaly detection, forecasting, or classification	Dependence on representative histories, drift management, and explainability challenges

Table 2 compares the principal vibration-based maintenance paradigms and highlights the limitations that motivated the hybrid strategy adopted in this study.

To translate these general observations into the case-study context, Table 3 summarizes the continuous-process constraints that directly shaped the deployed architecture.

Table 3. Continuous-process constraints that shaped the deployed architecture

Constraint	Manifestation in the Plant	Design Implication for the Digital Twin
Hygiene and cleaning	Daily high-pressure washing and exposure to aggressive cleaning routines	Need for protected sensors, robust mounting, and stable connectivity
Regulatory traceability	HACCP and quality-traceability requirements	Need for auditable data storage, configuration evidence, and event retention
Product and load variability	Recipe changes and operating variations	Need for careful normalization and cautious interpretation of vibration drift
Continuous criticality	Downtime immediately affects production and quality risk	Need to control false alarms while preserving early detection capability

2.3 Research gap and contribution

The gap addressed in this paper is therefore practical and methodological. The study does not claim a novel digital-twin theory or a new forecasting family. Instead, it contributes an industrial implementation study with four linked elements: an end-to-end architecture connecting field sensing, historian-based storage, asset modelling, supervision, forecasting, spectral diagnosis, and CMMS-linked action; a transparent benchmark of alternative forecasting models under deployment-oriented criteria; an alarming strategy that keeps maintenance decisions expressed through ISO 20816-3:2022 condition zones and engineering units [16]; and an auditable evaluation of operational indicators before and after go-live [17]. In this paper, interpretability is therefore claimed at the system level rather than at the internal-model level: the gated recurrent unit (GRU) is used as a forecasting component, while the maintenance-facing logic remains grounded in measurable vibration levels, explicit persistence rules, and fault labels derived from spectral analysis. Together, these elements position the work as a decision-oriented industrial case study with direct relevance for adoption in continuous-process environments [18].

3. MATERIALS AND METHODS

3.1 Study design, site, and monitored asset

This research uses an embedded industrial case-study design combined with a before-and-after evaluation. The DT went live on 1 June 2025, which separates a pre-deployment period (January–May 2025) from a post-deployment period (June–December 2025). The chosen design is appropriate for a real plant setting in which randomization and experimental controls were not feasible. It enables operational measurement while preserving traceability to field conditions. The observed differences are therefore interpreted as changes associated with deployment rather than as definitive causal effects.

The study site is a RO water-treatment system located in a soft-drink bottling plant. The RO unit produces permeate water used as a principal process input. The operating regime is quasi-continuous and includes critical rotating assets whose degradation can affect both production continuity and process quality. The high-pressure pump and its drive train were prioritized because they combine high criticality, high mechanical loading, and a direct influence on the availability of treated water. Deployment subsequently extended to other rotating elements within the RO line.

Table 4 lists the rotating equipment included in the monitoring scope together with the technical characteristics and operational criticality that justified their prioritization.

Table 4. Critical rotating equipment is monitored in the reverse osmosis (RO) system

Equipment	Technical Specifications	Operational Criticality
High-pressure pump	multi-stage centrifugal pump; 55 kW; 1480 rpm nominal speed; 20 bar discharge pressure	critical asset; failure causes immediate production interruption
Main motor	three-phase asynchronous motor; 55 kW; 1500 rpm synchronous speed; ip55 protection	critical because it drives the high-pressure pump
Booster pump	centrifugal pump; 15 kW; 1450 rpm nominal speed	important asset for feed-pressure continuity
Booster-pump motor	asynchronous motor; 15 kW; 1500 rpm synchronous speed	important supporting asset

3.2 Digital twin architecture

The implemented DT combines a physical layer and a virtual layer that interact continuously through a structured data pipeline. The physical layer comprises the monitored RO assets, the installed vibration and temperature sensors, acquisition gateways, and the plant operating context. The virtual layer comprises the historian server, the asset repository, calculation and analytics services, supervision interfaces, notification services, and maintenance system connectors [19]. Figure 1 summarizes the architecture adopted in the case study.

A central design choice was to treat the physical asset as the sole source of truth while using the virtual layer to contextualize, process, and operationalize the incoming data. This separation helped avoid a common implementation problem in which dashboards and models proliferate without a clear chain of traceability back to the field signals [20]. The historian stores time-series data, the asset repository structures equipment context and derived indicators, and the supervision layer provides role-based views for operations and maintenance. The analytics layer computes health indicators, forecasts short-term trajectories, evaluates threshold-crossing risk, and triggers diagnosis-support functions. Maintenance integration then closes the loop by linking confirmed alerts to work-order generation and execution feedback.

Supplementary Figures S1 and S3 provide additional visual documentation of the deployed supervision environment and the operator-facing monitoring, maintenance, and prediction views used during implementation.

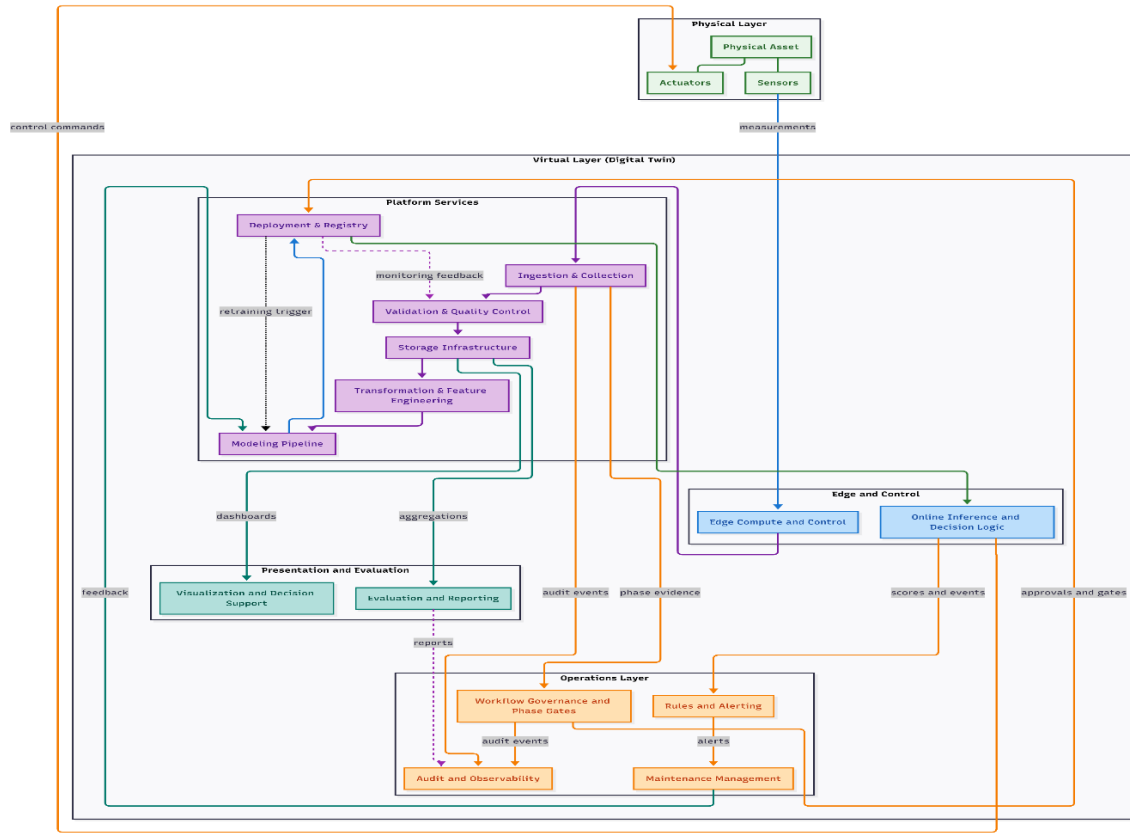


Figure 1. End-to-end architecture of the deployed digital twin (DT) for the reverse osmosis (RO) system

3.3 Instrumentation and data pipeline

Vibration instrumentation was designed to balance diagnostic value, robustness, and integration simplicity. Sensors with analogue and IO-Link connectivity were selected for online condition monitoring. Depending on location, these sensors provide vibration velocity, supplementary condition indicators, and temperature information over a frequency range suitable for rotating-equipment surveillance. Measurement points were chosen to capture the most informative mechanical paths while remaining compatible with access, hygiene, and protection requirements in the plant. Table 5 details the vibration instrumentation selected for the case study and the connectivity options used for online condition monitoring.

Data acquisition follows a sequential chain from field sensing to decision support. Sensor values are transmitted through IO-Link gateways to an acquisition connector, historized as time series, contextualized in the asset repository, exposed to supervision services, and used by the notification and maintenance connectors. This design ensures that the same

trusted data backbone supports monitoring, analytics, and action. Table 6 summarizes the end-to-end flow and its operational purpose.

Before model training and evaluation, the vibration series were subjected to a preprocessing workflow designed to preserve auditability and avoid temporal leakage. Short communication gaps were treated conservatively, obvious transients and signal-quality anomalies were screened during commissioning, and non-representative operating periods were excluded when they were clearly attributable to maintenance or wash-down activities rather than to machine condition. For model development, scaling parameters were estimated on the training partition only and then carried unchanged to validation and test data. This combination of signal-quality control, operating-context filtering, and training-only normalization was adopted to make the benchmark fair across models while keeping the deployed pipeline consistent with plant practice.

Supplementary Figure S2 provides additional visual documentation of the field instrumentation and QR-enabled documentation support used during deployment.

Table 5. Vibration instrumentation used in the case study

Sensor Type	Main Function	Connectivity	Frequency Range	Outputs
Velocity sensor with analogue and IO-Link output	online condition monitoring of rotating equipment.	4–20 mA and IO-Link	0.5–10,000 Hz	vibration velocity (mm/s) and condition indicators
Velocity sensor with software integration	continuous vibration measurement and indicator transmission.	IO-Link	0.5–10,000 Hz	vibration velocity and machine-health indicators
Condition-monitoring velocity sensor	vibration-speed measurement under equipment-specific configuration.	IO-Link	Equipment-dependent	vibration-speed indicators and configured health variables

Table 6. End-to-end data flow from sensing to maintenance action

Step	Source	Transport	Target	Operational Usage	Step
Field acquisition	vibration sensors	IO-Link gateway	acquisition connector	collection of vibration and velocity measurements	field acquisition
Historization	acquisition connector	connector to historian server	time-series historian	storage, timestamp integrity, and traceability	historization
Asset modelling	historized signals	platform services	asset repository	context, attributes, calculations, and indicators	asset modelling
Supervision	asset repository	interface services	web supervision interface	dashboards, trends, and spectrum access	supervision
Alerts	analytics engine	notification services	notification layer	email/SMS alerts and escalation logic	alerts

3.4 Analytics and alarm logic

The analytics layer is organized into complementary modules rather than a single black-box model. A first module computes a health index that combines vibration, peak behaviour, and temperature into a compact condition score for prioritization. A second module maps measured V-RMS values to severity zones using ISO 20816-3:2022, which provides an interpretable bridge between signal levels and maintenance decisions. A third module forecasts future V-RMS trajectories over multiple horizons. A fourth module performs fault-oriented diagnosis after alarm confirmation through the extraction of spectral signatures and a classifier or rule base operating on characteristic frequencies [21].

For forecasting, the study benchmarked several candidate models on five-minute vibration-velocity series. The candidate set included direct GRU, direct long short-term memory (LSTM), causal one-dimensional convolutional neural network (CNN), histogram-based gradient boosting, simple recurrent neural network (RNN), random forest, linear regression, support vector regression (SVR), and a persistence baseline. To keep model comparison transparent, tuning effort was harmonized across model families using the same training/validation/test chronology and an expanding-window validation logic on the training partition. The deployed GRU used a direct multi-horizon formulation with recurrent layers followed by a dense output layer, and hyperparameters were

selected on validation performance rather than on the test set. Model selection was based not only on global fit but also on horizon-specific performance, false-negative and false-positive behaviour around the 240-minute horizon, inference latency, and memory footprint, because a deployed maintenance model must be both accurate and operationally efficient [22].

The alarm policy combines measurement, forecast, persistence, and hysteresis. Immediate inspection is requested when measured V-RMS exceeds the abnormal zone threshold. A pre-alarm is triggered when the predicted trajectory crosses the faulty-to-abnormal boundary within the forecasting horizon or when failure probability rises persistently [17]. Alarm confirmation requires repeated exceedance over two to three successive acquisitions in order to filter one-off transients associated with operating changes, while return to normal status requires three acquisitions below a reset threshold. The chosen persistence window was established during commissioning as an engineering compromise: single exceedances generated nuisance alarms, whereas longer confirmation windows delayed actionable intervention. Confirmed alarms trigger short-duration high-frequency acquisitions for spectral diagnosis, thereby preserving interpretability in the final maintenance recommendation [23].

Table 7 summarizes the analytical modules embedded in the digital twin (DT) and the operational outputs expected from each module.

Table 7. Analytical modules and expected outputs of the deployed digital twin (DT)

Module	Algorithm or Logic	Main Output	Operational Use
Health index	Composite formula combining vibration, peaks, and temperature	Condition score (0–100)	Prioritization, drift tracking, and comparison across assets
Condition mapping	ISO 20816-3:2022 thresholds applied to measured or forecast V-RMS	Normal/acceptable/faulty/abnormal zone	Interpretable alarm handling and maintenance communication
Trend forecasting	Direct GRU multi-horizon predictor selected after benchmarking	Short-term trajectory forecast	Threshold anticipation and intervention planning
Fault-oriented diagnosis	Rules and classifier on characteristic frequencies after triggered high-frequency acquisition	Probable fault class	Root-cause support and maintenance prioritization
Alarm orchestration	Persistence rule, hysteresis, and confirmation logic	Pre-alarm, confirmed alarm, return-to-normal status	False-alarm filtering and workflow triggering

Note: GRU = gated recurrent unit

3.5 Deployment batches and evaluation protocol

Deployment was organized in successive batches to preserve auditability. The first batch established historian infrastructure, data points, and connectivity tests. The second batch instrumented the assets and validated measurement quality. The third batch configured supervision views, thresholds, and notifications. The fourth batch connected alerts to the CMMS or ERP maintenance workflow. The fifth batch

added advanced analytics and periodic performance review. Each batch produced technical deliverables and concrete evidence, such as test reports, configuration captures, generated work orders, or usage logs.

System evaluation used both predictive and operational indicators. Predictive indicators included RMSE, MAE, coefficient of determination, latency, memory demand, and false-negative/false-positive rates. Operational indicators were extracted from maintenance logs and included monthly

unplanned downtime, total downtime, availability, and the proportion of failure-free days. A day was classified as failure-free when no unplanned downtime event was recorded. The primary comparative analysis contrasted the pre- and post-go-live periods. A two-proportion z-test and relative risk were used for the failure-free-day metric, and the corresponding absolute difference in proportions is also reported to make the practical effect size explicit. Mann-Whitney tests and

interrupted time-series summaries were used descriptively for monthly downtime and availability trends. Because the study is observational and based on a single line, all inferential statements are made cautiously and are interpreted as evidence of association rather than strict causation.

Table 8 summarizes the successive deployment batches, the deliverables produced at each stage, and the auditable evidence retained during implementation.

Table 8. Deployment batches and associated evidence

Deployment Batch	Contents	Main Deliverables	Auditable Evidence
Historian platform infrastructure	servers, interfaces, and asset model	architecture, asset database, historian points	installation logs, configuration captures, connectivity tests
Industrial sensor instrumentation	sensors, gateways, and tags	sensor list, tag mapping, measurement points	sensor test reports, signal stability checks, loss-rate verification
Supervision	role-based web views	dashboards, thresholds, notifications	view captures, usage logs, alert history
CMMS integration	connector, API, workflow	alert-to-work-order flow	generated work orders, response-time traces, feedback records
Advanced analytics	health index, detection, forecasting	rules, parameters, performance monitoring	monthly reports, alarm reviews, model adjustments

4. RESULTS

4.1 Forecasting benchmark and model selection

Table 9 reports the forecasting benchmark. The direct GRU model ranked first on the test set with RMSE = 0.096 mm/s, MAE = 0.074 mm/s, overall R² = 0.874, and R² = 0.775 at the 240-minute horizon. It also remained compatible with low-

resource deployment, requiring 1.20 ms inference time and 78 KB of memory in the benchmarked configuration. The causal 1D-CNN delivered competitive latency but lower horizon accuracy, whereas the direct LSTM incurred higher memory and latency costs without outperforming the GRU. Tree-based and linear baselines were easier to serve but showed materially weaker fit and poorer decision robustness.

Table 9. Forecasting benchmark on the test set

Rank	Model	Family	RMSE (mm/s)	MAE (mm/s)	R ² (Overall)	R ² (H240)	Latency (ms)	Memory (KB)	FNR (H240)	FPR (H240)	FNR (Max)	FPR (Max)
1	Direct GRU	Gated RNN	0.096	0.074	0.874	0.775	1.20	78	0.08	0.06	0.05	0.09
2	Direct causal 1D-CNN	Temporal CNN	0.129	0.076	0.775	0.608	0.85	92	0.10	0.06	0.06	0.10
3	Direct LSTM	Gated RNN	0.208	0.148	0.410	0.320	1.55	110	0.09	0.07	0.06	0.10
4	HistGB multi-output	Tree boosting	0.230	0.169	0.277	0.211	0.65	420	0.12	0.08	0.08	0.12
5	Direct SimpleRNN	RNN	0.248	0.200	0.152	0.098	0.95	55	0.14	0.07	0.10	0.11
6	ExtraTrees	Tree ensemble	0.346	0.283	-0.639	-0.752	5.20	8200	0.12	0.09	0.09	0.13
7	RandomForest	Tree ensemble	0.617	0.467	-4.173	-4.563	6.80	9600	0.13	0.09	0.10	0.14
8	GradBoost multi-output	Tree boosting	0.665	0.526	-5.044	-5.903	1.10	1500	0.16	0.10	0.12	0.15
9	AdaBoost multi-output	Tree boosting	0.673	0.531	-5.192	-5.895	1.00	1300	0.18	0.10	0.13	0.16
10	DecisionTree multi-output	Tree	0.675	0.538	-5.228	-5.458	0.55	980	0.21	0.12	0.16	0.18
11	SVR-RBF multi-output	Kernel	0.722	0.590	-6.106	-6.804	4.10	650	0.24	0.14	0.18	0.20
12	KNN multi-output	Instance-based	0.880	0.716	-9.643	-	3.20	2200	0.28	0.15	0.22	0.21

Note: GRU = gated recurrent unit; RNN = recurrent neural network; CNN = convolutional neural network; LSTM = long short-term memory; KNN = k-nearest neighbors; SVR = support vector regression

These results supported the operational choice of the direct GRU as the forecasting core of the DT. The model provided the strongest overall compromise between predictive sharpness, horizon stability, and deployability. Table 10

further shows that performance degrades gradually with horizon length rather than collapsing abruptly. This behaviour is advantageous for maintenance planning because it preserves directional information over several hours while keeping

short-horizon estimates reliable.

Table 10. Prediction performance by horizon for the selected direct gated recurrent unit (GRU) model

Prediction Type / Horizon	RMSE	R ²
Overall	0.096	0.874
t + 1	0.060	0.952
t + 60	0.096	0.874
t + 120	0.113	0.826
t + 180	0.118	0.810

These results supported the operational choice of the direct GRU as the forecasting core of the DT. The model provided the strongest overall compromise between predictive sharpness, horizon stability, and deployability. Table 10 further shows that performance degrades gradually with horizon length rather than collapsing abruptly. This behaviour

is advantageous for maintenance planning because it preserves directional information over several hours while keeping short-horizon estimates reliable.

4.2 Online monitoring, alarm handling, and diagnosis support

Figure 2 presents the internal decision logic that links data processing, alarm handling, and maintenance action. In practice, the DT did not rely on prediction alone. Alarm handling remained anchored in interpretable severity zones derived from ISO 20816-3:2022, while forecasting extended the anticipation horizon and spectral diagnosis supported fault discrimination after alarm confirmation. This layered design reduced the risk that maintenance recommendations would be perceived as opaque or disconnected from established engineering practice.

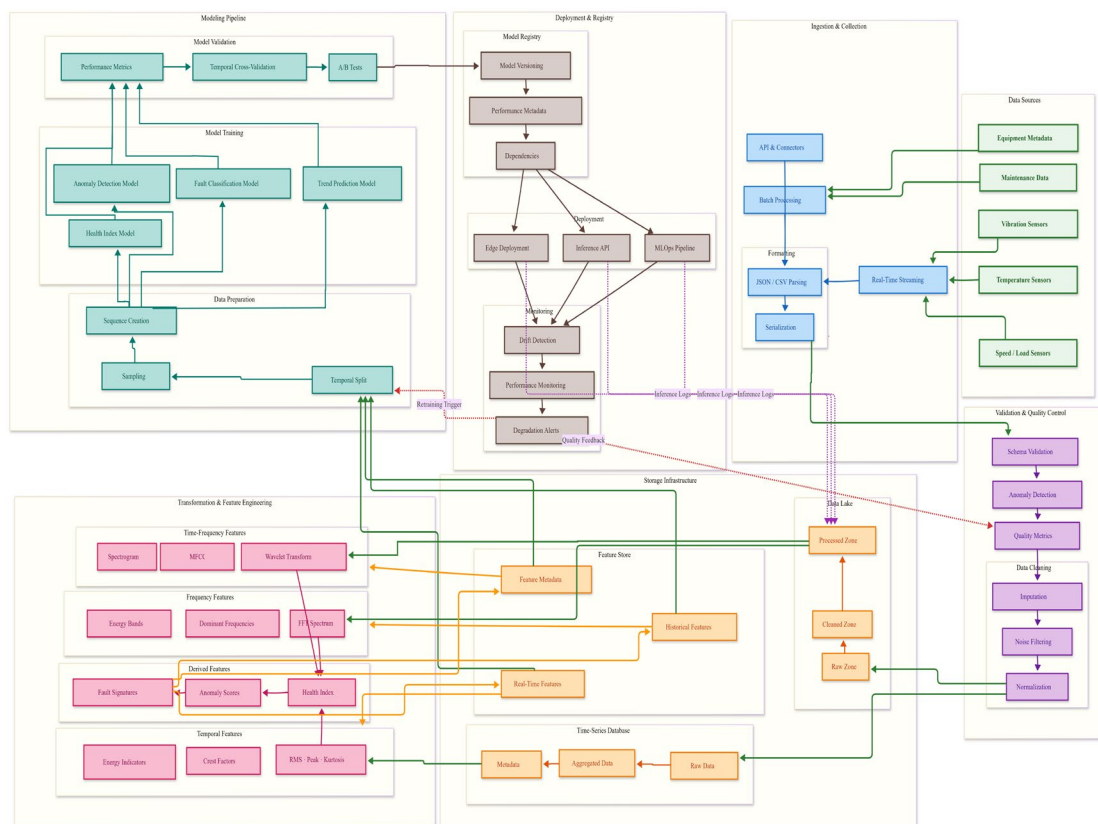


Figure 2. Internal data-processing and decision logic of the deployed digital twin (DT)

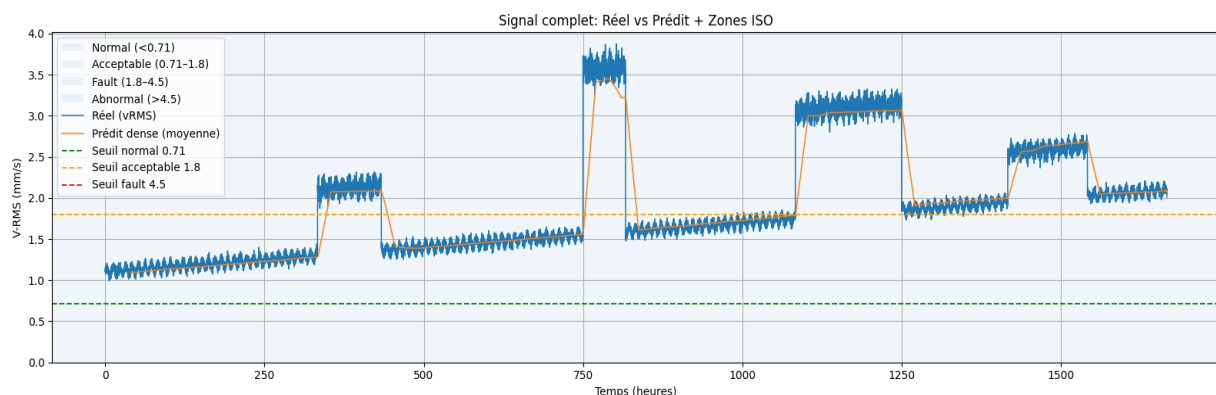


Figure 3. Representative comparison between measured and predicted V-RMS trajectories for the selected model

Figure 3 illustrates a representative measured-versus-predicted vibration trajectory for the selected model. The forecast follows the main trend and preserves the major excursions relevant to threshold anticipation [24]. The purpose of the forecast is not to reproduce every short-lived fluctuation, but to provide actionable foresight on whether the vibration trajectory is approaching a condition zone that warrants intervention. This qualitative reading is supported by the quantitative benchmark of Table 10: The selected GRU achieved an overall RMSE of 0.096 mm/s and an overall R^2 of 0.874, while remaining at $R^2 = 0.826$ at 120 minutes and $R^2 = 0.810$ at 180 minutes. Figure 3 should therefore be read as a representative operational example of a pattern that is already documented quantitatively across horizons, rather than as a standalone proof point. Once a confirmed alarm is issued, the system triggers short-duration high-frequency acquisitions to inspect characteristic frequencies linked to imbalance,

misalignment, bearing defects, cavitation, or electrical anomalies.

A representative maintenance episode further illustrates the decision chain implemented in the plant. A progressive increase in vibration on the high-pressure pump first appeared as a sustained upward drift in the health indicators, then as repeated pre-alarms, and finally as a confirmed alarm once the persistence rule was satisfied. The confirmed alarm triggered focused spectral acquisition and maintenance inspection, which pointed to a bearing-related fault and led to a targeted intervention rather than reactive continuation until stoppage. After the intervention, vibration returned to the normal or acceptable zone. Although the full maintenance record cannot be reproduced here for confidentiality reasons, this case shows how the deployed workflow translated forecast information into diagnosis support and maintenance action.

Table 11. Fault-specific performance after alarm-triggered spectral diagnosis

Fault Type	Frequency	Prediction Accuracy (%)	Early Detection Rate (%)	False Positive Rate (%)
Bearing defects	30	94.2	89.7	3.1
Misalignment	20	91.8	85.3	4.2
Mechanical imbalance	20	93.5	87.9	2.8
Pump cavitation	15	89.6	82.4	5.7
Electrical defects	10	87.3	78.1	6.9
Fixation problems	5	92.1	86.2	3.9
Weighted average	100	92.1	86.4	4.1

Fault-specific performance is summarized in Table 11. The best results were obtained for bearing defects, mechanical imbalance, and misalignment, with prediction accuracies above 91%. Hydraulic and electrical faults were more difficult to classify, which is consistent with their lower event counts and the fact that some signatures can overlap with process-induced variability. Even so, the diagnosis layer remained sufficiently discriminative to support maintenance prioritization and root-cause investigation.

4.3 Operational impact after go-live

The operational evaluation indicates a favourable change

after go-live. Over the 245 recorded operating days in 2025, the proportion of failure-free days increased from 72.8% in the pre-deployment period (75 of 103 days) to 84.5% in the post-deployment period (120 of 142 days). The corresponding two-proportion comparison yielded $p = 0.025$ and an estimated relative risk of 1.16 [1.01, 1.33], which is consistent with improved day-level reliability after deployment. The absolute improvement was 11.7 percentage points, which helps interpret the practical magnitude of the change alongside the relative effect. Table 12 provides the aggregate before-and-after comparison of failure-free days and days affected by unplanned downtime around the digital-twin go-live date.

Table 12. Failure-free days before and after digital-twin go-live

Period	Recorded Days	Failure-Free Days	Days with Unplanned Downtime	% Failure-Free
Pre (Jan–May 2025)	103	75	28	72.8
Post (Jun–Dec 2025)	142	120	22	84.5
Overall (Jan–Dec 2025)	245	195	50	79.6

Table 13. Monthly stability profile during 2025

Month	Period	Recorded Days	Failure-Free Days	Days with Unplanned Downtime	% Failure-Free
2025-01	Pre	21	17	4	81.0
2025-02	Pre	20	15	5	75.0
2025-03	Pre	20	14	6	70.0
2025-04	Pre	21	16	5	76.2
2025-05	Pre	21	13	8	61.9
2025-06	Post	19	16	3	84.2
2025-07	Post	22	18	4	81.8
2025-08	Post	18	15	3	83.3
2025-09	Post	21	19	2	90.5
2025-10	Post	22	19	3	86.4
2025-11	Post	18	14	4	77.8
2025-12	Post	22	19	3	86.4

Table 14. Summary of main operational outcomes after go-live

Metric	Observed Change after Go-Live	Statistical Summary	Measurement Basis
Unplanned downtime	strong decrease	Mann–Whitney $p = 0.00535$; interrupted time-series level change -44.4 h/month and trend -5.96 h/month	maintenance logs
System availability	increase	Mann–Whitney $p = 0.00568$; interrupted time-series level change $+21.3$ percentage points and trend $+3.23$ percentage points/month	availability = $100 \times (\text{production time} / \text{planned time})$
Reliability (failure-free days)	72.8% \rightarrow 84.5%	two-proportion z-test $p = 0.025$; relative risk = 1.16 [1.01, 1.33]	daily logs; day classified as failure-free when unplanned downtime = 0 h

Table 13 shows that the improvement was not confined to a single month. Although the post-go-live period still contains variability, most months after June 2025 display a higher share of failure-free days than the most unstable pre-deployment months. Figure 4 complements this view by showing the monthly evolution of unplanned downtime, failure-free-day percentage, and availability around the go-live date. The visual pattern is coherent with a shift toward fewer severe stoppages and more stable operation after deployment.

Table 14 consolidates the main operational indicators. Maintenance logs indicate a strong reduction in unplanned downtime, an increase in availability, and an improvement in the reliability metric defined by failure-free days. These findings should be interpreted as convergent operational evidence rather than as proof that the DT alone caused the entire improvement. In an industrial setting, maintenance planning discipline, operator learning, and parallel technical adjustments may contribute to the observed pattern.

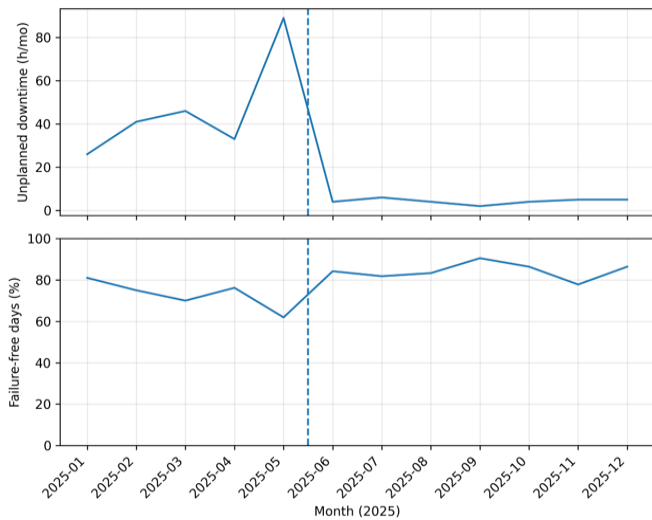


Figure 4. Monthly evolution of unplanned downtime, failure-free days, and availability around the go-live date

5. DISCUSSION

The first engineering contribution of the study lies in the combination of interpretability and predictive reach. Many maintenance systems force an artificial choice between standards-based monitoring and data-driven forecasting. In the present case, interpretability was preserved by anchoring alarm logic in ISO 20816-3:2022 severity zones, while predictive reach was added through the direct GRU model and downstream spectral diagnosis. This hybrid architecture appears particularly suitable for plants that need decision

support without abandoning established maintenance heuristics.

The second contribution concerns operationalization. The DT described here is not limited to data display. Its practical value depends on the maintenance loop that links alerts to work-order generation, execution feedback, and subsequent adjustment. This loop moves the implementation beyond a pure digital shadow toward a more operational DT. The result is not full autonomy, nor should it be in a safety- and quality-critical environment. Instead, the DT functions as a structured decision-support system that improves the timeliness and traceability of human intervention.

The model-selection results also carry a broader methodological message. For industrial predictive maintenance, the best model is not necessarily the most complex model in abstract benchmarking terms. The chosen model must satisfy a multi-criteria requirement set that includes false-negative control, latency, memory demand, and maintainability. In this case, the direct GRU outperformed the alternatives not because recurrent architectures are universally superior, but because its accuracy-deployment balance was the most credible for the operating context.

The study also illustrates the importance of contextual constraints in process industries. Sensor ruggedness, cleaning routines, regulatory traceability, production variability, and the need for continuity during network or platform disturbances all shaped the final architecture. These constraints are often treated as implementation details, yet they determine whether a DT survives beyond a demonstration phase. The case, therefore, supports the view that industrial DT research should pay greater attention to deployment governance and lifecycle maintainability rather than to modelling performance alone.

Generalisability should be interpreted at the architectural level rather than as a statistical claim about all plants. The combination of online vibration sensing, historian-based traceability, explicit alarm zones, and workflow integration is transferable to other continuous-process settings with critical rotating assets. By contrast, threshold calibration, persistence settings, and the final forecasting configuration remain site-dependent and should be commissioned locally rather than copied unchanged.

The operational change is also economically meaningful even when confidential plant-cost data cannot be disclosed. Using the interrupted time-series estimate reported in Table 14, a reduction of about 44.4 hours of unplanned downtime per month corresponds to an avoided-loss term of approximately 44.4 cost savings per month per month, where C denotes the site-specific cost of one lost production hour, including stoppage, restart, quality, and labour effects. Expressing the result in this parametric form avoids over-claiming a monetary figure while still making clear that the observed downtime

reduction has direct managerial relevance.

Several limitations must be acknowledged. First, the analysis is based on a single industrial site and one principal asset family, which limits external validity. Second, the before-and-after design does not control for confounders such as seasonal demand, changing production schedules, operator learning, or parallel maintenance improvements. Third, the event base used for fault-specific diagnosis remains moderate in size for some fault classes, which may affect the stability of class-wise estimates. Fourth, the current implementation uses vibration as the primary predictive signal; integrating process variables such as flow, pressure, and temperature should improve normalization under changing loads. These limitations are the reason the paper frames the results as operational evidence associated with deployment rather than as a definitive causal estimate.

Future work should therefore focus on multi-site validation, concept-drift monitoring, controlled retraining strategies, and a fuller economic evaluation of avoided downtime and maintenance effort. Extending the DT to additional rotating assets and integrating process context more tightly would also strengthen generalizability. Despite these limitations, the case provides rare implementation evidence on how an industrial DT can be assembled into a coherent maintenance workflow and how its impact can be evaluated using auditable plant records.

6. CONCLUSIONS

This paper documented the industrial deployment of a decision-oriented DT for vibration-based predictive maintenance of an RO system in a beverage filling plant. The reported contribution is practical rather than purely conceptual: it combines field instrumentation, historian-based data management, asset modelling, supervision, explicit alarm logic, multi-horizon forecasting, alarm-triggered spectral diagnosis, and CMMS-linked maintenance action within a single operational architecture.

The forecasting benchmark showed that a direct GRU provided the best compromise between predictive performance and low-resource deployment constraints. After go-live, the plant recorded a higher share of failure-free days and more favourable operational indicators, consistent with a transition toward earlier and more proactive maintenance. Because the evaluation is observational, these changes should be interpreted cautiously; nevertheless, they form a coherent body of evidence that the DT supported a more stable operation.

For practitioners, the main lesson is that predictive maintenance succeeds when data acquisition, model selection, interpretability, and maintenance workflow integration are designed together. For researchers, the case highlights the value of documenting deployment detail and auditable operational outcomes, especially in process industries where implementation barriers remain high.

ACKNOWLEDGMENT

This research was supported by the National Centre for Scientific and Technical Research, Morocco through the Project: ALKHARIZMI No. 2020/24.

REFERENCES

- [1] Giret, A., Garcia, E., Botti, V. (2016). An engineering framework for service-oriented intelligent manufacturing systems. *Computers in Industry*, 81: 116-127. <https://doi.org/10.1016/j.compind.2016.02.002>
- [2] Gunduz, M.A., Paksoy, T., Çopur, E.H. (2024). The Internet of Things and cyber-physical systems: Aviation industry applications. In *Smart and Sustainable Operations Management in the Aviation Industry*, pp. 69-76.
- [3] Grieves, M., Vickers, J. (2016). Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. In *Transdisciplinary Perspectives on Complex Systems: New Findings and Approaches*, pp. 85-113. https://doi.org/10.1007/978-3-319-38756-7_4
- [4] Glaessgen, E., Stargel, D. (2012). The digital twin paradigm for future NASA and US air force vehicles. In *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 20th AIAA/ASME/AHS Adaptive Structures Conference 14th AIAA*, Honolulu, Hawaii, p. 1818. <https://doi.org/10.2514/6.2012-1818>
- [5] Kritzinger, W., Karner, M., Traar, G., Henjes, J., Sihn, W. (2018). Digital twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnline*, 51(11): 1016-1022. <https://doi.org/10.1016/j.ifacol.2018.08.474>
- [6] Fuller, A., Fan, Z., Day, C., Barlow, C. (2020). Digital twin: Enabling technologies, challenges and open research. *IEEE Access*, 8: 108952-108971. <https://doi.org/10.1109/ACCESS.2020.2998358>
- [7] Randall, R.B. (2021). *Vibration-Based Condition Monitoring: Industrial, Automotive and Aerospace Applications*. John Wiley & Sons.
- [8] Shah, R., Mittal, V., Lotwin, M. (2025). Recent advances in vibration analysis for predictive maintenance of modern automotive powertrains. *Vibration*, 8(4): 68. <https://doi.org/10.3390/vibration8040068>
- [9] Bellavista, P., Di Modica, G. (2023). The IoTwins methodology and platform to implement and operate digital twins-based 14.0 applications in the cloud continuum. In *2023 26th Euromicro Conference on Digital System Design (DSD)*, Golem, Albania, pp. 176-183. <https://doi.org/10.1109/DSD60849.2023.00034>
- [10] Kerrouchi, S., Aghezzaf, E.H., Cottyn, J. (2024). Production digital twin: A systematic literature review of challenges. *International Journal of Computer Integrated Manufacturing*, 37(10-11): 1168-1193. <https://doi.org/10.1080/0951192X.2024.2314792>
- [11] Latsou, C., Ariansyah, D., Salome, L., Erkoyuncu, J.A., Sibson, J., Dunville, J. (2024). A unified framework for digital twin development in manufacturing. *Advanced Engineering Informatics*, 62: 102567. <https://doi.org/10.1016/j.aei.2024.102567>
- [12] Aivaliotis, P., Georgoulas, K., Chryssolouris, G. (2019). The use of digital twin for predictive maintenance in manufacturing. *International Journal of Computer Integrated Manufacturing*, 32(11): 1067-1080. <https://doi.org/10.1080/0951192X.2019.1686173>
- [13] Wynn, M., Irizar, J. (2023). Digital twin applications in manufacturing industry: A case study from a german multi-national. *Future Internet*, 15(9): 282. <https://doi.org/10.3390/fi15090282>

- [14] Embarki, S., El Kihel, Y. (2024). Integrating digital twins into manufacturing execution systems (MES) case study on wastewater filtration pilot. In International Conference on Electronic Engineering and Renewable Energy Systems, Singapore, Springer Nature Singapore, pp. 207-216. https://doi.org/10.1007/978-981-97-9975-6_21
- [15] Aminzadeh, A., Dimitrova, M., Meiabadi, M.S., Sattarpanah Karganroudi, S., Taheri, H., Ibrahim, H., Wen, Y. (2023). Non-contact inspection methods for wind turbine blade maintenance: Techno-economic review of techniques for integration with Industry 4.0. *Journal of Nondestructive Evaluation*, 42(2): 54. <https://doi.org/10.1007/s10921-023-00967-5>
- [16] International Organization for Standardization. (2022). ISO 20816-3:2022 - Mechanical vibration — Measurement and evaluation of machine vibration — Part 3: industrial machinery with a power rating above 15 kW and operating speeds between 120 R/min and 30 000 R/min, Geneva, Switzerland. <https://nexgen-cableway.com/wp-content/uploads/2025/11/ISO-20816-3-2022.pdf>.
- [17] Abdalah, R.W., Abdulateef, O.F., Hamad, A.H. (2025). A predictive maintenance system based on industrial Internet of Things for multimachine multiclass using deep neural network. *Journal Européen des Systèmes Automatisés*, 58(2): 373-381. <https://doi.org/10.18280/jesa.580218>
- [18] Hassan, M., Svadling, M., Björzell, N. (2024). Experience from implementing digital twins for maintenance in industrial processes. *Journal of Intelligent Manufacturing*, 35(2): 875-884. <https://doi.org/10.1007/s10845-023-02078-4>
- [19] El Kihel, Y., Embarki, S., El Kihel, B. (2025). Enhancing Industrial process optimization through digital twins. *IFAC-PapersOnLine*, 59(10): 1283-1288. <https://doi.org/10.1016/j.ifacol.2025.09.216>
- [20] Pulcini, V., Modoni, G. (2024). Machine learning-based digital twin of a conveyor belt for predictive maintenance. *The International Journal of Advanced Manufacturing Technology*, 133(11): 6095-6110. <https://doi.org/10.1007/s00170-024-14097-3>
- [21] Bensaoucha, S., Gharib, G.M., Soudi, M.A., Teta, A., et al. (2025). Robust bearing fault detection and classification using deep neural networks: A comprehensive study on the CWRU dataset. *Journal Européen des Systèmes Automatisés*, 58(11): 2435-2443. <https://doi.org/10.18280/jesa.581120>
- [22] Cho, K., Van Merriënboer, B., Gulçehre, Ç., Bahdanau, D., Bougares, F., Schwenk, H., Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, pp. 1724-1734. <https://doi.org/10.3115/v1/D14-1179>
- [23] Soufiane, E., El Mahdi, B. (2023). Contribution to maintenance 4.0 by monitoring and prognosis of industrial installations by digital twin: Case study on wastewater filtration pilot. In 2023 IEEE 12th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), Dortmund, Germany, pp. 113-118. <https://doi.org/10.1109/IDAACS58523.2023.10348937>
- [24] Touil, A., Babaa, F., Kratz, F., Bennis, O. (2024). Bearing fault diagnosis in induction machines based on electromagnetic torque spectral frequencies analysis. *Journal Européen des Systèmes Automatisés*, 57(1): 255-261. <https://doi.org/10.18280/jesa.570124>

APPENDIX

The following supplementary figures provide additional visual documentation of the deployed supervision environment, field instrumentation, and operator-facing views of the implemented digital twin.

Figure S1. illustrates the supervision layer of the implemented digital twin, where process variables and condition-monitoring indicators are visualized in an integrated operational interface.

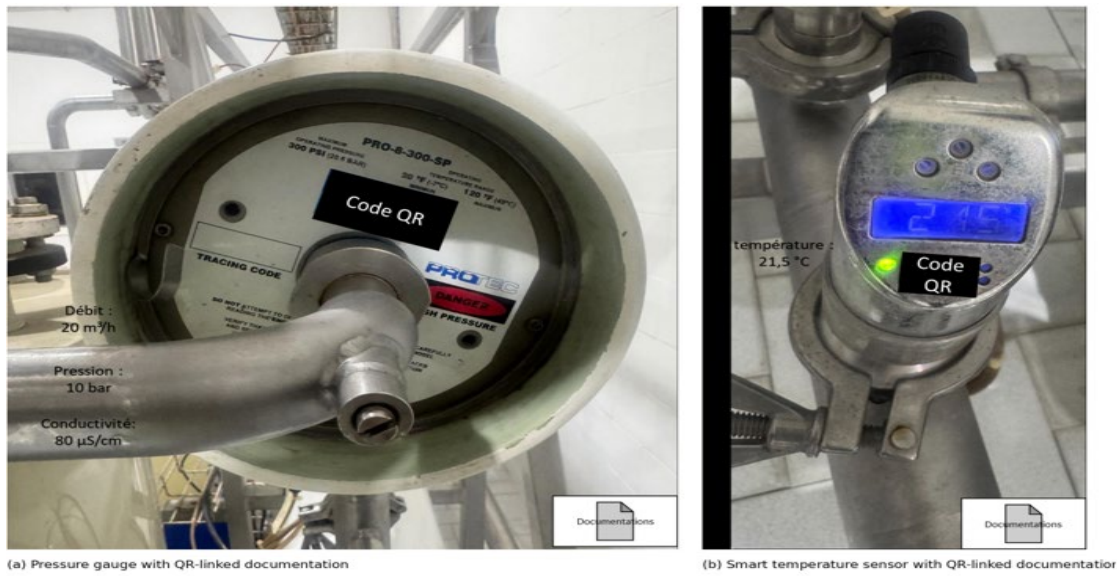


Figure S1. Example of the deployed supervision interface used for reverse osmosis monitoring and decision support

Figure S2 illustrates the field devices and documentation-access mechanism used during implementation. The QR-

enabled traceability approach supported maintenance operations, equipment identification, and practical use of the

deployed monitoring architecture in the plant environment.



Field instrumentation and documentation access used in the deployed reverse-osmosis monitoring system.

Figure S2. Example of field instrumentation and QR-enabled documentation support in the deployed reverse osmosis monitoring system

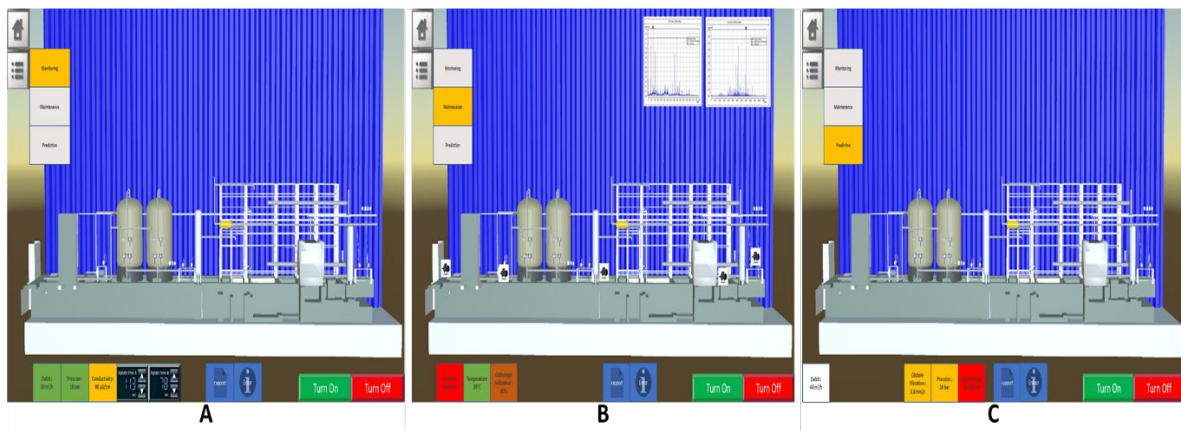


Figure S3. Operator-facing views of the deployed digital twin ((A) Monitoring view used for real-time supervision of the reverse osmosis system. (B) Maintenance view with alarm-centered information and spectral-analysis support for fault diagnosis. (C) Prediction view used for trend anticipation and what-if scenario exploration)