International Journal of
Technology and
Systems
(IJTS)

AI-Enabled Process Safety Digital Twin for Predictive Risk Management in Process Industries



Journal of Technology and Systems
ISSN: 2788-6344 (Online)
Vol. 8, Issue No. 5, pp 1 – 22, 2026
www.carijournals.org

AI-Enabled Process Safety Digital Twin for Predictive Risk Management in Process Industries

Hafiz Abid Ali

MSC Safety Engineering, Mary Kay O'Connor Process Safety Center/TEES/Texas A&M University, USA

Accepted: 25th Oct, 2026, Received in Revised Form: 11th Nov, 2026, Published: 25th Nov, 2026

Abstract

Process safety management in process industries remains challenging due to complex operating conditions, equipment degradation, and limitations of conventional risk assessment methods. This study proposes a conceptual Artificial Intelligence (AI)-enabled process safety digital twin framework for predictive risk management. The architecture integrates real-time process monitoring, digital twin synchronization, machine learning-based anomaly detection, equipment reliability assessment, and dynamic risk prediction. An autoencoder is proposed for anomaly detection, while Long Short-Term Memory networks forecast process deviations and equipment degradation. A Dynamic Bayesian Network estimates evolving hazard probabilities using operational observations and predicted equipment conditions. The proposed evaluation protocol uses simulated industrial disturbances to assess detection accuracy, prediction lead time, false alarm rates, reliability estimation, and risk prediction performance. As the framework has not yet been implemented or experimentally validated, its potential benefits remain hypotheses requiring investigation. The study establishes a structured methodology for evaluating AI-assisted predictive safety decision support in industrial processing environments.

Keywords: *Artificial Intelligence; Digital Twin; Process Safety; Predictive Risk Management; Equipment Reliability.*

1. Introduction

Process safety management is a critical requirement in process industries, including petrochemical plants, oil and gas facilities, chemical manufacturing systems, and petroleum refineries, where complex operational conditions and hazardous materials create significant risks to personnel, equipment, and the environment. Industrial accidents may result from equipment deterioration, process deviations, control system failures, and inadequate identification of emerging hazards. Conventional process safety management relies on established methods such as Hazard and Operability Studies (HAZOP), Failure Mode and Effects Analysis (FMEA), and quantitative risk assessment (QRA). Although these approaches provide systematic methods for identifying hazards, they frequently depend on periodic assessments and predefined failure scenarios, limiting their responsiveness to changing operational conditions. Vairo et al. [7] emphasized the importance of dynamic risk assessment approaches that incorporate changing process conditions into industrial safety evaluation.

The increasing adoption of Industrial Internet of Things (IIoT) technologies, advanced sensing systems, and artificial intelligence (AI) has created opportunities to improve industrial process safety through continuous monitoring and predictive analysis. Machine learning and deep learning techniques can analyze complex operational datasets, identify abnormal patterns, and estimate potential equipment failures before critical operating limits are exceeded. For instance, Hartung et al. [16] investigated deep anomaly detection using the Tennessee Eastman process dataset, demonstrating the relevance of advanced analytical techniques for industrial fault identification. Similarly, Ma et al. [17] developed a data-driven digital twin approach for gas turbine anomaly detection while considering uncertainty in industrial operating conditions. These developments indicate the potential of predictive intelligence to support earlier hazard recognition and more informed operational decisions.

Digital twin technology provides an additional foundation for improving process safety by establishing a continuously updated virtual representation of physical industrial assets and operating processes. Unlike conventional simulation models, digital twins can incorporate operational measurements to reflect changing equipment conditions, process behaviour, and system performance. Kapteyn et al. [28] introduced a probabilistic graphical modelling foundation for predictive digital twins, highlighting their potential to integrate physical system information with uncertainty-aware predictions. Furthermore, Hassan et al. [23] examined the implementation of digital twins for maintenance in industrial processes, demonstrating their relevance to equipment condition monitoring and maintenance management. Integrating these capabilities with AI-based risk assessment may enable industrial operators to anticipate hazardous conditions instead of responding only after abnormalities have occurred.

Despite advances in AI-enabled monitoring, digital twin modelling, and dynamic risk assessment, challenges remain in translating uncertain process predictions into reliable, safety-relevant risk

estimates. Existing studies have investigated machine learning-based anomaly detection, equipment reliability modelling, and Bayesian risk assessment; however, the specific treatment of prediction uncertainty, temporal dependencies, and equipment degradation within an explicitly coupled digital twin risk-assessment architecture warrants further investigation. In particular, the extent to which uncertainty originating from sensor measurements, process-state estimation, and machine learning forecasts propagates into predicted hazard probabilities remains an important methodological concern.

To address this issue, the present study proposes an AI-enabled process safety digital twin framework that explicitly couples machine learning-based process prediction, equipment degradation assessment, and Dynamic Bayesian Network risk estimation. The principal proposed contribution is an uncertainty-propagation mechanism through which prediction uncertainty is incorporated into conditional hazard probabilities rather than treating AI-generated forecasts as deterministic inputs. This mechanism is intended to connect uncertain operational observations, predicted equipment conditions, and evolving process hazards within a synchronized digital twin architecture. The framework also includes anomaly detection, temporal state updates, and risk-informed supervisory decision support, with emphasis on traceable relationships between predicted deviations and estimated safety risks.

The proposed methodology will be evaluated through simulation-based experiments involving normal operation, process disturbances, sensor uncertainty, and equipment degradation scenarios. Evaluation will examine prediction accuracy, risk probability calibration, uncertainty sensitivity, false alarm rates, and warning lead time against conventional monitoring and deterministic risk-prediction approaches. Because the framework has not yet been implemented or experimentally validated, its anticipated advantages remain research hypotheses rather than demonstrated improvements. The study therefore contributes a conceptual and mathematical design for investigating whether explicit uncertainty propagation improves the reliability and interpretability of AI-enabled predictive risk assessment in process industries.

Despite advances in AI-enabled monitoring, digital twin modelling, and dynamic risk assessment, challenges remain in translating uncertain process predictions into reliable, safety-relevant risk estimates. Existing studies have investigated machine learning-based anomaly detection, equipment reliability modelling, and Bayesian risk assessment; however, the specific treatment of prediction uncertainty, temporal dependencies, and equipment degradation within an explicitly coupled digital twin risk-assessment architecture warrants further investigation. In particular, the extent to which uncertainty originating from sensor measurements, process-state estimation, and machine learning forecasts propagates into predicted hazard probabilities remains an important methodological concern.

To address this issue, the present study proposes an AI-enabled process safety digital twin framework that explicitly couples machine learning-based process prediction, equipment

degradation assessment, and Dynamic Bayesian Network risk estimation. The principal proposed contribution is an uncertainty-propagation mechanism through which prediction uncertainty is incorporated into conditional hazard probabilities rather than treating AI-generated forecasts as deterministic inputs. This mechanism is intended to connect uncertain operational observations, predicted equipment conditions, and evolving process hazards within a synchronized digital twin architecture. The framework also includes anomaly detection, temporal state updates, and risk-informed supervisory decision support, with emphasis on traceable relationships between predicted deviations and estimated safety risks.

The proposed methodology will be evaluated through simulation-based experiments involving normal operation, process disturbances, sensor uncertainty, and equipment degradation scenarios. Evaluation will examine prediction accuracy, risk probability calibration, uncertainty sensitivity, false alarm rates, and warning lead time against conventional monitoring and deterministic risk-prediction approaches. Because the framework has not yet been implemented or experimentally validated, its anticipated advantages remain research hypotheses rather than demonstrated improvements. The study therefore contributes a conceptual and mathematical design for investigating whether explicit uncertainty propagation improves the reliability and interpretability of AI-enabled predictive risk assessment in process industries.

2. Literature Review and Research Gap

Digital twin technology has emerged as an important approach for monitoring industrial operations, estimating equipment conditions, and supporting process safety decisions. By integrating physical system models with operational measurements, digital twins can update estimates of process behaviour as plant conditions change. Kapteyn et al. [28] developed a probabilistic graphical modelling foundation for predictive digital twins, demonstrating the importance of combining physical models, observational information, and uncertainty. More recently, Rebello and Nogueira [3] presented an integrated framework for chemical engineering digital twins incorporating online learning, model adaptation, and uncertainty assessment. Their work demonstrates that adaptive prediction and uncertainty management are already established capabilities in digital twin research. Consequently, the integration of artificial intelligence with process simulation cannot independently establish the novelty of a proposed safety framework.

Research on dynamic process safety assessment has also advanced beyond conventional static risk estimation. Vairo et al. [7] developed the DYN-RISK approach, integrating neural networks, Hidden Markov Models, and hierarchical Bayesian networks to support dynamic risk estimation in industrial facilities. Amin and Khan [31] proposed an adaptive Bayesian network methodology incorporating artificial neural networks for operating-mode identification, fault detection, event-tree modelling, and time-dependent consequence assessment. Their study established an important connection between process deviations and dynamic operational risk, indicating that combined machine learning and probabilistic risk modelling is not a new contribution by itself.

He et al. [32] developed a real-time probabilistic risk assessment method using Dynamic Bayesian Networks (DBNs) and online monitoring information for petrochemical applications. Their approach demonstrated how prior risk information and monitored evidence can be incorporated into continuously updated risk estimates. Similarly, Moradi et al. [33] integrated Bayesian deep learning and Bayesian networks for condition monitoring and operational risk assessment, including an application involving an industrial vapour recovery unit. These studies demonstrate that integrated monitoring and probabilistic risk updating have already been investigated using industrially relevant data.

Risk-informed fault detection has received further attention. Tarcsay et al. [34] combined Dynamic Principal Component Analysis (DPCA) with Failure Mode and Effects Analysis (FMEA)-based Bayesian networks to distinguish safety-critical faults from lower-risk abnormalities. Their methodology used risk information to support fault prioritization and reduce unnecessary alarms. This is particularly relevant because conventional anomaly detection models may identify abnormal behaviour without determining its operational consequences. Hartung et al. [16] investigated deep anomaly detection using Tennessee Eastman process data, supporting the applicability of data-driven fault monitoring. However, the identification of an anomaly alone does not provide a calibrated probability of hazardous-event occurrence.

Equipment reliability and process degradation represent another important research direction. Hassan et al. [23] investigated digital twin implementation for industrial maintenance, highlighting the practical relevance of updated equipment models and operational information. Predictive maintenance systems can estimate deterioration and support maintenance decisions, but their outputs require additional modelling to establish how equipment failure likelihood influences process-level hazard probabilities. The relationship between predicted operating states, time-dependent equipment failure rates, and safety barrier effectiveness therefore requires explicit mathematical treatment.

Braniff et al. [35] reviewed developments in real-time process safety and systems decision-making, demonstrating the growing convergence of dynamic risk analysis, predictive modelling, and digital technologies. Their review further reinforces the need to distinguish new research contributions from existing combinations of monitoring, machine learning, and risk assessment.

Table 1. Comparative Analysis of Existing Studies and Research Opportunities

Study	Main methodological contribution	Relevant limitation or unresolved question
Vairo et al. [7]	Dynamic risk assessment using neural networks, Markov models, and Bayesian networks	Establish whether prediction uncertainty is explicitly propagated into the reported hazard probabilities
Amin and Khan [31]	Adaptive Bayesian network integrating operating-mode recognition and time-dependent loss assessment	Explicit coupling with synchronized digital twin state estimation requires examination
He et al. [32]	DBN-based real-time probabilistic risk assessment using monitored evidence	Integration with forecast-dependent equipment hazard models requires further comparison
Moradi et al. [33]	Bayesian deep learning and Bayesian networks for condition and operational risk monitoring	Digital twin synchronization and component degradation coupling require further examination
Tarcsay et al. [34]	Risk-based fault detection combining DPCA, FMEA, and Bayesian networks	Predictive horizon and equipment reliability integration require further evaluation
Rebello and Nogueira [3]	Adaptive digital twin incorporating online learning and uncertainty assessment	Explicit process safety risk and consequence modelling are not its primary focus
Proposed study	Digital twin state estimation, anomaly detection, forecasting, reliability modelling, and DBN risk assessment	Integration and performance remain to be specified and experimentally validated

The comparative evidence indicates that the principal research gap cannot be described simply as a lack of integration between artificial intelligence and dynamic risk assessment. Existing studies already demonstrate substantial progress in these areas. A more precise research opportunity concerns the explicit and reproducible coupling of digital twin state estimates, anomaly indicators, forecast-dependent equipment hazard rates, and probabilistic safety risk assessment, including the propagation of predictive uncertainty between components.

Accordingly, the proposed research will investigate a coordinated methodology in which synchronized process-state estimates inform anomaly detection and forecasting, predicted degradation influences equipment failure probabilities, and these probabilities become evidence

within a Dynamic Bayesian Network. The intended contribution also includes defining risk prediction horizons, uncertainty treatment, and operator-facing risk indicators.

These features represent proposed methodological objectives rather than established innovations. Their novelty and practical value must be evaluated against relevant existing frameworks using common datasets, defined safety-event outcomes, comparative baselines, and statistical validation. This narrower research position provides a more defensible foundation for developing and assessing an AI-enabled process safety digital twin.

3. Proposed AI-Enabled Process Safety Digital Twin Framework

3.1. Overall System Architecture

This study proposes an uncertainty-aware, AI-enabled process safety digital twin framework that links industrial process monitoring, predictive modelling, equipment reliability estimation, and dynamic risk assessment. The architecture is intended to support early identification of hazardous conditions and evidence-based preventive decisions in petrochemical, chemical manufacturing, and related process facilities. Unlike a conventional monitoring pipeline, the proposed digital twin incorporates continuous state estimation, model discrepancy assessment, and feedback mechanisms for updating its predictions.

The architecture comprises six functional layers: (1) physical process monitoring, (2) data acquisition and preprocessing, (3) digital twin synchronization, (4) AI-based anomaly detection and forecasting, (5) equipment reliability and dynamic risk assessment, and (6) operator-supervised safety decision support. Data are transmitted from industrial equipment to the digital twin, where estimated process states and prediction residuals are generated. These outputs inform machine learning models, whose forecasts and uncertainty estimates are subsequently used for reliability and risk calculations. Kapteyn et al. [28] established a probabilistic foundation for integrating physical system information, observations, and uncertainty in predictive digital twins.

The proposed contribution is not the individual application of autoencoders, Long Short-Term Memory (LSTM) networks, or Dynamic Bayesian Networks (DBNs), which have been investigated previously [16], [31], [32]. Instead, the framework specifies a coordinated information pathway through which estimated process conditions and forecast uncertainty influence equipment failure probabilities and evolving process-level risk. Its effectiveness remains a hypothesis requiring experimental validation.

3.2. Industrial Data Acquisition and Preprocessing

The data acquisition layer collects time-stamped measurements from industrial sensors, programmable logic controllers (PLCs), supervisory control and data acquisition (SCADA) systems, and equipment condition-monitoring platforms. Core process measurements include temperature, pressure, flow rate, vessel level, and equipment operating states. Where available,

vibration measurements, inspection records, operating hours, and historical component failures provide additional reliability information.

Measurements undergo timestamp alignment, range validation, noise filtering, missing-value identification, and normalization. Data quality is recorded alongside individual measurements to distinguish genuine process abnormalities from sensor degradation or communication failures. Invalid observations are flagged rather than automatically interpreted as process faults.

The framework also separates process measurements from equipment reliability data. This distinction is essential because chemical process benchmarks, including the Tennessee Eastman Process, do not necessarily contain component failure times, inspection histories, or maintenance records. Consequently, additional datasets or explicitly defined simulated degradation mechanisms would be required to validate the reliability component independently.

3.3. Digital Twin Synchronization and State Estimation

The digital twin combines a process dynamics model with sequential state estimation to maintain an updated representation of plant operating conditions. For the proposed implementation, the process model will be expressed through nonlinear state-space equations, with an Unscented Kalman Filter (UKF) used to estimate unobserved process states from measured variables. The UKF is selected as an initial design because process relationships may be nonlinear and direct measurements of all relevant states are rarely available.

At each observation interval, the digital twin predicts the next process state using its current estimate, control inputs, and process parameters. Incoming measurements correct this prediction, while the difference between predicted and observed measurements forms an innovation residual. The estimated state, residual, and state-estimation covariance are then transferred to the predictive analytics layer.

Synchronization is evaluated using residual magnitude, state-estimation uncertainty, and consistency between simulated and measured process behaviour. Persistent discrepancies may indicate process abnormalities, sensor faults, or model–plant mismatch. Model parameters are reviewed and recalibrated only when sufficient evidence indicates model drift, with updates subject to validation before deployment.

This approach establishes a feedback relationship between the physical system, digital twin, and predictive models. Rebello and Nogueira [3] discussed online learning and uncertainty assessment in chemical engineering digital twins, providing relevant context for adaptive model maintenance.

3.4. AI-Based Anomaly Detection and Process Forecasting

The predictive analytics layer contains two complementary models. An autoencoder identifies abnormal operating patterns, while an LSTM forecasts future process states over a defined prediction horizon.

The autoencoder receives normalized measurements, digital twin residuals, and selected estimated process states. Its reconstruction error provides an anomaly score representing deviations from patterns observed during normal operation. The anomaly threshold will be selected using an independent validation dataset to achieve a predefined false-alarm constraint. Persistence checks distinguish isolated measurement disturbances from sustained abnormalities.

The LSTM uses historical sequences of estimated process states and operating inputs to forecast future process conditions. Its principal task is process-state forecasting, not direct prediction of equipment failure. Forecasted operating loads, temperature excursions, and other degradation-relevant variables are subsequently transferred to the reliability model.

To quantify predictive uncertainty, an ensemble of independently trained LSTM models is proposed. Variability among ensemble predictions represents an estimate of model uncertainty, while state-estimation covariance contributes information about measurement and process uncertainty. These uncertainty estimates are propagated into downstream reliability and risk calculations rather than discarded after forecasting.

Hartung et al. [16] investigated deep anomaly detection in chemical process monitoring, while Moradi et al. [33] examined the integration of Bayesian deep learning with probabilistic operational risk assessment. These studies provide relevant methodological foundations, although the proposed model interactions must be independently evaluated.

3.5. Equipment Reliability, Dynamic Risk Assessment, and Uncertainty Propagation

The reliability and risk assessment layer converts predicted operating conditions into equipment failure likelihoods and process-level risk estimates. The proposed reliability model employs a time-dependent proportional hazards formulation in which baseline equipment failure rates are modified by degradation indicators, operating loads, and forecast process conditions. Where sufficient failure-time data exist, model parameters will be estimated statistically. Otherwise, explicitly documented engineering assumptions or published reliability data will be required, with their uncertainty retained in subsequent calculations.

The Dynamic Bayesian Network represents dependencies among process deviations, equipment failure states, safety barrier availability, and hazardous events. Its initial structure will be derived from process hazard analysis, HAZOP findings, and bow-tie relationships. Conditional probability tables will be estimated from relevant operational records where available and supplemented by documented expert elicitation when empirical observations are insufficient.

At each monitoring interval, digital twin estimates, anomaly indicators, and equipment failure probabilities provide evidence for updating the DBN. The network subsequently estimates the probability of specified hazardous events within a defined prediction horizon. Risk is calculated by combining these probabilities with scenario-specific consequence estimates.

Importantly, uncertainty is propagated through the framework rather than represented by a single deterministic risk score. Samples from the estimated process-state and forecast distributions are transferred through the reliability and DBN models to obtain a distribution of possible risk outcomes. Reported outputs will include central risk estimates and uncertainty intervals. This enables operators to distinguish high predicted risk from poorly supported predictions.

Existing adaptive Bayesian risk assessment studies, including Amin and Khan [31] and He et al. [32], provide important comparators for evaluating whether the proposed coupling offers additional benefits. The intended contribution must therefore be demonstrated through comparative experiments rather than inferred from architectural integration alone.

3.6. Safety Decision Support, Industrial Assurance, and Model Governance

The final layer provides risk-informed recommendations while maintaining a clear separation between AI-based advisory functions and established industrial protection systems. The proposed digital twin is designated as an advisory decision-support system, not a safety instrumented function. It will not independently initiate shutdowns, modify safety instrumented system logic, or bypass established protective barriers.

This design recognizes the functional safety lifecycle requirements associated with IEC 61511 and IEC 61508. These standards govern safety-related systems and their engineering and operational management; the proposed AI layer must not be assumed to satisfy their requirements merely because it is separated from the safety instrumented system.

Operator-facing information will include the affected equipment, detected abnormality, estimated event probability, predicted risk horizon, associated uncertainty, relevant contributing variables, and recommended inspection or maintenance actions. Explanations will identify the principal measurements or model outputs responsible for elevated risk scores.

Alarm generation will follow a defined prioritization and suppression strategy to limit repeated or low-confidence alerts. This is consistent with the alarm lifecycle and rationalization principles established in ISA-18.2.

Alerts will be assessed using false-alarm frequency, alarm burden, detection lead time, and operator response requirements. The framework will distinguish informational notifications from alarms requiring timely operational action.

Model governance will include drift monitoring, version control, documented retraining criteria, validation before deployment, and management-of-change procedures. Cybersecurity measures will prioritize restricted data interfaces, authenticated access, network segmentation, and audit logging. Where possible, the analytics environment will receive process information through controlled, read-only interfaces.

If sensor reliability deteriorates, model uncertainty exceeds an approved limit, or digital twin synchronization becomes unreliable, predictive recommendations will be marked as unavailable or uncertain and escalated for operator review. Existing plant alarms and protective systems will continue to operate independently.

These safeguards are proposed engineering requirements, not demonstrated safety properties. Their adequacy must be established through technical verification, operational testing, and appropriate safety assurance before deployment.

Table 2. Proposed AI-Enabled Digital Twin Framework Components and Functional Roles

Component	Proposed Technology or Method	Functional Role
Physical process monitoring	Industrial sensors, PLC, SCADA	Collect operating measurements and control inputs
Data acquisition and preprocessing	Data validation, filtering, time synchronization	Prepare quality-controlled process data
Digital twin synchronization	Nonlinear process model, UKF	Estimate states, calculate residuals, and quantify uncertainty
Anomaly detection	Autoencoder	Detect deviations using measurements and twin residuals
Process forecasting	LSTM ensemble	Predict future states and quantify forecast uncertainty
Equipment reliability	Proportional hazards and survival models	Estimate failure probabilities
Dynamic risk assessment	Dynamic Bayesian Network	Update hazardous-event probabilities
Uncertainty propagation	Distribution sampling and probabilistic inference	Estimate risk uncertainty intervals
Safety decision support	Risk dashboard and alarm prioritization	Communicate risks and preventive recommendations
Model governance	Drift monitoring, change control, access control	Maintain model integrity and safety boundaries

Journal of Technology and Systems
ISSN: 2788-6344 (Online)
Vol. 8, Issue No. 5, pp 1 – 22, 2026
www.carijournals.org

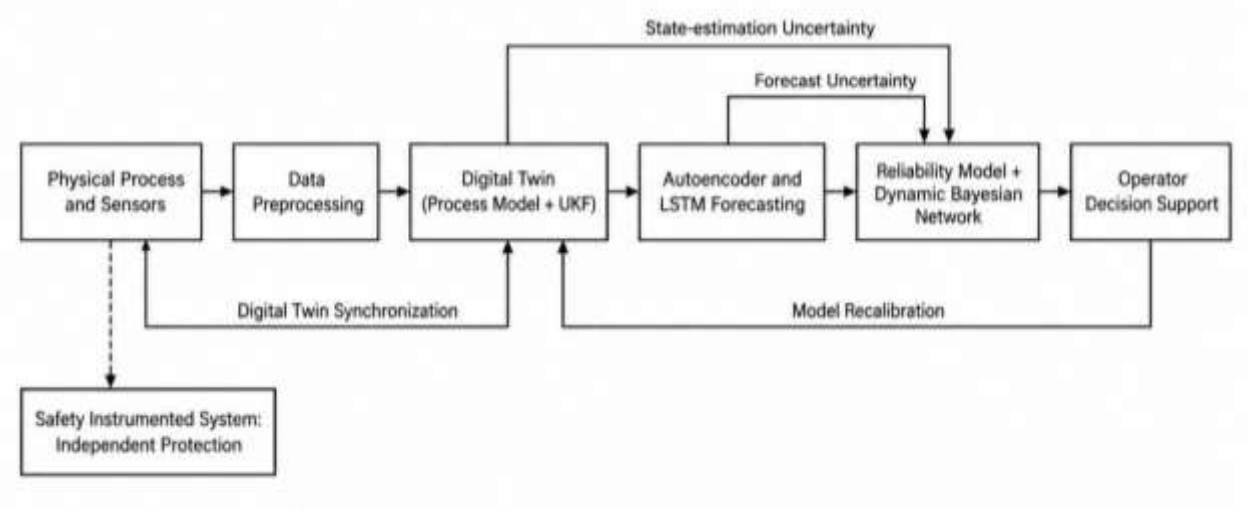

Figure 1. Proposed information flow, model coupling, uncertainty propagation, and separation between AI-enabled advisory functions and independent industrial protection systems.

4. Mathematical Modelling and Methodology

The proposed AI-enabled process safety digital twin integrates nonlinear state estimation, anomaly detection, process forecasting, equipment reliability modelling, and probabilistic risk assessment. The mathematical framework establishes an explicit relationship between uncertain process predictions, equipment-specific operating loads, and evolving hazard probabilities. At discrete sampling intervals, the physical process is represented by:

$$\begin{aligned} \mathbf{x}_{t+1} &= f(\mathbf{x}_t, \mathbf{u}_t, \boldsymbol{\theta}) + \mathbf{w}_t, \\ \mathbf{y}_t &= h(\mathbf{x}_t) + \mathbf{v}_t. \end{aligned} \qquad (1)$$

where $\mathbf{x}_t$ denotes process states, $\mathbf{u}_t$ represents control inputs, and $\boldsymbol{\theta}$ contains physical parameters.

Process disturbances $\mathbf{w}_t$ and measurement noise $\mathbf{v}_t$ have assumed covariance matrices $\mathbf{Q}$ and $\mathbf{R}$, respectively. These parameters require calibration or sensitivity analysis for the selected industrial process.

An unscented Kalman filter (UKF) synchronizes the digital twin with incoming observations. Sigma-point propagation provides predicted states, measurement estimates, and uncertainty covariance matrices. The measurement update is:

$$\begin{aligned} \mathbf{e}_t &= \mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1}, \\ \mathbf{K}_t &= \mathbf{P}_{xy,t}\mathbf{S}_t^{-1}, \\ \hat{\mathbf{x}}_{t|t} &= \hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t\mathbf{e}_t. \end{aligned} \qquad (2)$$

where $\mathbf{e}_t$ represents measurement innovation, $\mathbf{K}_t$ is the Kalman gain, and $\mathbf{S}_t$ denotes innovation covariance. The posterior covariance is $\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \mathbf{K}_t\mathbf{S}_t\mathbf{K}_t^T$. Innovation statistics are monitored to identify synchronization errors and potential model mismatch [28].

An autoencoder identifies abnormal process behaviour by reconstructing standardized sensor measurements and estimated states. Its reconstruction error is expressed as:

$$A_t = \frac{1}{n}\sum_{i=1}^{n}\left(z_{t,i} - \hat{z}_{t,i}\right)^2. \qquad (3)$$

where $z_{t,i}$ and $\hat{z}_{t,i}$ denote standardized input and reconstructed values. Anomaly thresholds are established using independent normal validation data, while persistent exceedances generate operational alerts. Anomaly scores are treated as diagnostic indicators rather than direct hazardous-event probabilities [16], [34].

A Long Short-Term Memory (LSTM) ensemble predicts future process states using historical operating conditions:

$$\hat{\mathbf{x}}_{t+k}^{(b)} = F_{\mathrm{LSTM}}^{(b)}(\hat{\mathbf{x}}_{t-m+1:t}, \mathbf{u}_{t-m+1:t}). \qquad (4)$$

where m denotes the historical observation window, k is the forecasting horizon, and b identifies an ensemble member. Forecast uncertainty combines ensemble disagreement, UKF state uncertainty, and process disturbances.

To establish a physically interpretable connection between process predictions and equipment reliability, a centrifugal pump supplying coolant to the reactor jacket is selected as an illustrative equipment component. Its predicted shaft-power demand is calculated as:

$$L_{\mathrm{pump},t+k} = \frac{\Delta p_{t+k} Q_{t+k}}{\eta_{\mathrm{pump}}}. \qquad (5)$$

where Δp_{t+k} represents predicted differential pressure across the pump (Pa), Q_{t+k} denotes volumetric coolant flow (m³/s), and η_{pump} represents pump efficiency. The resulting load is expressed in watts. Flow rate and pressure differential must be measured or obtained from an explicitly defined hydraulic model. This relationship provides the equipment-specific load input for reliability estimation, rather than assuming that reactor temperature alone determines pump degradation.

The pump failure hazard is modelled using a proportional-hazards formulation:

$$\lambda_i(t) = \lambda_{0,i}(t)\exp\left[\boldsymbol{\beta}_i^T \mathbf{L}_i(t) + \gamma_i d_i(t)\right]. \qquad (6)$$

where $\lambda_{0,i}(t)$ denotes baseline failure intensity, $\mathbf{L}_i(t)$ contains operating-load indicators, and $d_i(t)$ represents equipment degradation. Conditional failure probability over a future interval Δ is:

$$P_{f,i}(t,\Delta) = 1 - \exp\left[-\int_t^{t+\Delta} \lambda_i(\tau)\, d\tau\right]. \qquad (7)$$

Hazard parameters require equipment failure-time data or explicitly documented engineering assumptions; ordinary process fault labels cannot independently establish component failure probabilities.

A Dynamic Bayesian Network (DBN) integrates process deviations, equipment failure probabilities, safety barrier conditions, and hazardous-event states. Its posterior event probability is:

$$P(H_{t+\Delta}|E_{1:t}) = \sum_{\mathbf{s}} P(H_{t+\Delta}, \mathbf{s}|E_{1:t}; \Theta). \qquad (8)$$

where $E_{1:t}$ represents accumulated evidence, $\mathbf{s}$ denotes unobserved network states, and Θ contains conditional probability parameters. Network dependencies are defined using documented hazard analysis rather than inferred solely from anomaly detection outputs [31]–[33].

Finally, uncertainty is propagated through Monte Carlo sampling of process-state estimates, forecast distributions, equipment reliability parameters, and DBN probabilities. The expected risk is calculated as:

$$\mathbb{E}[\mathcal{R}_{t+\Delta}] \approx \frac{1}{M}\sum_{r=1}^{M}\left[\sum_{j=1}^{J} P_r\left(H_j|E_{1:t}\right)C_j\right]. \qquad (9)$$

where M denotes simulation samples and C_j represents the consequence of hazardous-event category j, with mutually exclusive categories used to avoid double counting. Percentile intervals describe model-dependent uncertainty. The proposed methodology will be evaluated using prediction error, probability calibration, Brier score, false alarm rate, warning lead time, and computational latency. Its effectiveness remains subject to experimental validation.

5. Experimental Setup and Performance Evaluation

The experimental methodology is designed to evaluate the proposed AI-enabled process safety digital twin under normal operating conditions, process disturbances, sensor degradation, and evolving equipment faults. The evaluation incorporates reproducible benchmark experiments, comparative baseline models, uncertainty quantification, and separate assessments of anomaly detection, equipment reliability estimation, and hazardous-event risk prediction. Particular emphasis is placed on distinguishing process fault detection from equipment failure prediction and safety risk assessment, as these tasks require different datasets, performance metrics, and validation procedures. The experimental design combines Tennessee Eastman Process benchmark data with supplementary reliability information and explicitly defined simulated hazardous-event scenarios to assess the proposed framework's predictive capabilities.

The Tennessee Eastman Process (TEP) will serve as the primary benchmark for process monitoring and fault detection. Originally introduced by Downs and Vogel [36], TEP represents a simulated chemical production process involving interconnected unit operations, feedback control, measured process variables, and predefined disturbances. Its established fault scenarios permit controlled assessment of abnormal-condition detection and process forecasting. However, the standard TEP benchmark does not provide complete equipment degradation histories, component failure-time observations, consequence estimates, or safety barrier availability records. Therefore, it cannot independently validate all elements of the proposed framework. These limitations will be addressed through a supplementary reliability evaluation and explicitly defined simulated safety-event scenarios.

The experimental implementation will employ a nonlinear digital twin model synchronized with process observations using an Unscented Kalman Filter (UKF). The twin will generate estimated process states, measurement residuals, and state-estimation covariance. These outputs will support autoencoder-based anomaly detection and LSTM-based process forecasting. Predicted operating conditions will subsequently inform proportional hazards reliability models, while the resulting

component failure probabilities and process deviations will provide evidence for Dynamic Bayesian Network (DBN) risk estimation. This procedure follows the integrated modelling architecture established in Sections 3 and 4.

To prevent information leakage, the selected TEP dataset version, fault definitions, sampling interval, and simulation settings will be documented. Training, validation, and testing observations will be separated at the simulation-run level rather than through random division of individual time-series samples. Model architecture, hyperparameters, random seeds, and preprocessing settings will also be reported. The digital twin will be tested under deliberately introduced parameter mismatch, sensor noise, missing measurements, and operating-regime changes to assess synchronization robustness rather than assuming perfect agreement between the plant and its virtual representation.

For equipment reliability evaluation, a separate dataset containing documented equipment degradation histories, component condition measurements, and observed failure or censoring times will be required. The Tennessee Eastman Process (TEP) benchmark will be used exclusively for process-state prediction and anomaly detection, as its predefined fault labels do not represent verified equipment failures or industrial accident outcomes. Equipment reliability models will therefore be calibrated and evaluated independently using suitable reliability records. If such records are unavailable, degradation and failure-time scenarios may be generated using explicitly documented engineering assumptions; however, their results will be classified as simulation-based sensitivity analyses rather than empirical validation of equipment reliability.

Performance comparisons will include conventional threshold monitoring, Principal Component Analysis (PCA), Dynamic Principal Component Analysis (DPCA), standalone autoencoders, and relevant risk-aware Bayesian approaches [31], [32], [34]. Ablation experiments will progressively evaluate the anomaly detector alone, integration with digital twin residuals, addition of LSTM forecasting, and inclusion of reliability and DBN risk estimation. These comparisons will help determine whether each component contributes measurable improvements beyond established methods.

Detection performance will be assessed using precision, recall, F1-score, false alarms per operating hour, and detection delay. Early-warning lead time will be defined as the interval between the first valid warning and a specified event onset or constraint violation. Forecasting accuracy will be measured using mean absolute error and root mean square error. Risk prediction will be assessed using the Brier score, calibration plots, and uncertainty interval coverage when event labels are available. Computational latency will be evaluated separately against the selected process sampling interval.

Experiments will use repeated simulation runs and independently initialized model training trials. Performance estimates will be reported with confidence intervals, while statistical comparisons will account for variation between test runs. Sensitivity analyses will examine the influence of

model mismatch, sensor faults, degradation assumptions, DBN conditional probabilities, and risk thresholds. These procedures establish a testable evaluation protocol, although actual performance conclusions will require implementation and experimental results.

Table 3. Proposed Experimental Scenarios and Performance Evaluation Metrics

Experimental Scenario	Evaluation Objective	Performance Metrics
Normal operation	Measure false-alarm burden under fault-free conditions	False alarms per operating hour; specificity
Process fault detection	Identify abnormalities across TEP fault runs	Precision, recall, F1-score, detection delay
Early warning	Measure warning time before defined event onset	Lead time; recall at fixed false-alarm rate
Digital twin synchronization	Evaluate state estimation under model–plant mismatch	RMSE; innovation consistency
Process forecasting	Assess predictions across defined horizons	MAE, RMSE; prediction interval coverage
Equipment degradation	Evaluate reliability using supplementary failure data	Brier score; survival calibration
Dynamic risk prediction	Assess probabilities of defined safety-event proxies	Brier score; reliability diagrams
Robustness testing	Assess sensor noise, missing data, and model drift	Performance degradation; uncertainty coverage
Computational performance	Assess compatibility with process sampling interval	Inference latency; memory use
Ablation and baselines	Measure contribution of each model component	Paired metric differences; confidence intervals

6. Expected Outcomes and Discussion

The proposed AI-enabled process safety digital twin framework is expected to provide an integrated approach to industrial risk monitoring by combining real-time operational measurements, machine learning predictions, equipment reliability estimates, and dynamic safety assessment. Its anticipated contribution is the ability to identify developing process abnormalities and translate their predicted consequences into information that supports preventive operational decisions. Unlike conventional threshold-based monitoring, which primarily responds when measured variables exceed predefined operating limits, the proposed framework is designed to recognize evolving deviations and assess their potential implications before hazardous conditions become critical.

One expected outcome is improved early identification of abnormal operating conditions through the combined application of autoencoder-based anomaly detection and LSTM forecasting. The autoencoder identifies deviations from learned operational patterns, while the LSTM estimates future process behaviour using historical observations. Combining these approaches may provide

additional warning time when process abnormalities develop gradually. Hartung et al. [16] demonstrated the relevance of deep anomaly detection for chemical process monitoring, while Ma et al. [17] investigated uncertainty-aware anomaly detection using a data-driven industrial digital twin. However, the expected benefits must be assessed against false alarms, changing operating regimes, and incomplete fault representations.

A second anticipated contribution concerns dynamic risk prediction. The integration of a Dynamic Bayesian Network allows predicted process deviations and equipment conditions to influence hazardous-event probability estimates. This approach may provide a more responsive representation of industrial risks than assessments based exclusively on fixed historical probabilities. Vairo et al. [7] discussed dynamic risk assessment as an approach for evaluating evolving industrial hazards. Nevertheless, accurate risk prediction will depend on the quality of conditional probability estimates, model assumptions, and available information about safety barriers and accident consequences.

Equipment reliability assessment represents another important expected outcome. Incorporating degradation indicators and time-dependent failure probabilities may support earlier maintenance planning and improved identification of vulnerable components. Hassan et al. [23] examined industrial digital twins for maintenance applications, emphasizing the practical importance of model updating and operational data integration. The proposed framework extends this relationship by connecting equipment reliability information with process-level risk assessment rather than treating maintenance prediction as an isolated function.

Despite these potential benefits, implementation presents several technical challenges. Industrial sensor measurements may contain missing values, noise, communication delays, and systematic errors that affect digital twin synchronization. Machine learning models may also produce unreliable predictions when operating conditions differ substantially from training data. Furthermore, equipment failure information is often limited, creating uncertainty in reliability estimation. Computational requirements, cybersecurity vulnerabilities, and the interpretability of predicted risk levels must also be considered.

The proposed framework therefore emphasizes uncertainty estimation, model validation, transparent risk communication, and operator supervision. AI-generated recommendations are intended to support existing safety management procedures rather than replace established protective systems.

Overall, the study is expected to establish a practical research foundation for integrating predictive analytics and digital twin technology into process safety management. Its actual effectiveness will be determined through comparative experimentation, sensitivity analysis, and subsequent validation using representative industrial operating data.

7. Conclusion

This study presented a conceptual AI-enabled process safety digital twin framework for predictive risk management in process industries. The proposed architecture integrates real-time process monitoring, uncertainty-aware state estimation, machine learning-based anomaly detection, equipment reliability modelling, and Dynamic Bayesian Network risk assessment. Its principal methodological contribution is the explicit coupling of uncertain process forecasts with equipment-specific operating loads, conditional failure probabilities, and time-dependent hazard estimates. The mathematical formulation establishes a structured approach for propagating uncertainty across these components, while the proposed evaluation protocol distinguishes process anomaly detection from equipment reliability and hazardous-event prediction. As the framework has not been implemented or experimentally validated, no conclusions can be drawn regarding predictive superiority, risk reduction, or improved industrial safety. Future research will require reproducible simulations, independent reliability and hazardous-event data, uncertainty calibration, and comparative evaluation before practical effectiveness can be established.

REFERENCES

[1] Zio, E., & Miqueles, L. (2024). Digital twins in safety analysis, risk assessment and emergency management. Reliability Engineering & System Safety, 246, 110040. https://doi.org/10.1016/j.ress.2024.110040

[2] Alrowaie, F. (2026). Digital twins and artificial intelligence for HAZOP enhancement in process safety: A critical literature review. Processes, 14, 2334. https://doi.org/10.3390/pr14142334

[3] Rebello, C. M., & Nogueira, I. B. R. (2025). Digital twins in chemical engineering: An integrated framework for identification, implementation, online learning, and uncertainty assessment. Computers & Chemical Engineering, 200, 109178. https://doi.org/10.1016/j.compchemeng.2025.109178

[4] Digital twins: Transforming the chemical process industry—A review. (2025). The Canadian Journal of Chemical Engineering, 103, 3611–3636. https://doi.org/10.1002/cjce.25611

[5] Digital twin in the chemical industry: A review. (2024). Digital Twins and Applications, 1(2), 118–130. https://doi.org/10.1049/dgt2.12019

[6] How digital twin technology may improve safety management: A multi-industry perspective. (2025). Safety Science, 189, 106837. https://doi.org/10.1016/j.ssci.2025.106837

[7] Vairo, T., Bragatto, P., Milazzo, M. F., Pettinato, M., & Fabiano, B. (2022). DYN-RISK – Design and development of a dynamic risk assessment tool. Chemical Engineering Transactions, 90, 325–330. https://doi.org/10.3303/CET2290055

[8] Dynamic risk assessment of chemical process systems using the System-Theoretic Accident Model and Process approach (STAMP) in combination with cascading failure propagation model (CFPM). (2024). Safety Science, 171, 106375. https://doi.org/10.1016/j.ssci.2023.106375

[9] Lee, S., Haskins, C., & Paltrinieri, N. (2022). Digital twin concept for risk analysis of oil storage tanks in operations: A systems engineering approach. Chemical Engineering Transactions, 90, 157–162. https://doi.org/10.3303/CET2290027

[10] Kabashkin, I. (2025). Mathematical framework for digital risk twins in safety-critical systems. Mathematics, 13(19), 3222. https://doi.org/10.3390/math13193222

[11] Artificial intelligence in process safety: A review of opportunities, challenges, and future directions for the chemical process industries. (2026). Journal of Loss Prevention in the Process Industries, 100, 105917. https://doi.org/10.1016/j.jlp.2026.105917

[12] Process safety in the era of AI: Assurance, deployment, and measurable risk reduction. (2026). Journal of Loss Prevention in the Process Industries, 101, 105956. https://doi.org/10.1016/j.jlp.2026.105956

[13] Artificial intelligence-enabled process safety: Advances, challenges, and future directions for high-risk industrial systems. (2026). Journal of Loss Prevention in the Process Industries, 102, 106026. https://doi.org/10.1016/j.jlp.2026.106026

[14] Digital twin-enabled safety management: A bibliometric and knowledge mapping analysis with core challenge insights. (2026). Journal of Loss Prevention in the Process Industries, 103, 106105. https://doi.org/10.1016/j.jlp.2026.106105

[15] Manzi, A., Iaiani, M., Tugnoli, A., Antonioni, G., & Cozzani, V. (2026). A digital twin-based approach supporting the calculation of cyber-attack success probability in process plants. Chemical Engineering Transactions, 126, 313–318. https://doi.org/10.3303/CET26126053

[16] Hartung, F., et al. (2023). Deep anomaly detection on Tennessee Eastman process data. Chemie Ingenieur Technik, 95, 1077–1082. https://doi.org/10.1002/cite.202200238

[17] Ma, Y., Zhu, X., Lu, J., Yang, P., & Sun, J. (2023). Construction of data-driven performance digital twin for a real-world gas turbine anomaly detection considering uncertainty. Sensors, 23(15), 6660. https://doi.org/10.3390/s23156660

[18] Latsou, C., Farsi, M., & Erkoyuncu, J. A. (2023). Digital twin-enabled automated anomaly detection and bottleneck identification in complex manufacturing systems using a multi-agent approach. Journal of Manufacturing Systems, 67, 242–264. https://doi.org/10.1016/j.jmsy.2023.02.008

[19] A collaborative network of digital twins for anomaly detection applications of complex systems: Snitch Digital Twin concept. (2023). Computers in Industry, 144, 103767. https://doi.org/10.1016/j.compind.2022.103767

[20] Wu, Y., Yang, G., Sun, J., Cui, L., & Wang, M. (2024). Digital twin modeling and leak diagnosis of temperature and stress fields in LNG storage tanks. Measurement, 228, 114374. https://doi.org/10.1016/j.measurement.2024.114374

[21] Mousavi, M., Li, B., Zhang, Z., Heidary, M. S., Barati, K., & Shen, X. (2026). A digital twin-oriented Dynamic Bayesian Network framework for reliable environmental safety monitoring in underground spaces. Process Safety and Environmental Protection, 213, 109027. https://doi.org/10.1016/j.psep.2026.109027

[22] Hassan, M., & Björsell, N. (2026). System and system-of-systems digital twins for predictive maintenance and root cause analysis. Journal of Intelligent Manufacturing. https://doi.org/10.1007/s10845-025-02789-w

[23] Hassan, M., Svadling, M., & Björsell, N. (2024). Experience from implementing digital twins for maintenance in industrial processes. Journal of Intelligent Manufacturing, 35, 875–884. https://doi.org/10.1007/s10845-023-02078-4

[24] Predictive maintenance using digital twins: A systematic literature review. (2022). Information and Software Technology, 151, 107008. https://doi.org/10.1016/j.infsof.2022.107008

[25] Harries, T., Hartnoll, M., Hafezianrazavi, M., Meek, H., & Nassehi, A. (2023). Digital twins for predictive maintenance. Procedia CIRP, 118, 306–311. https://doi.org/10.1016/j.procir.2023.06.053

[26] Integrating production and maintenance planning in process industries using Digital Twin: A literature review. (2024). IFAC-PapersOnLine, 58(19), 151–156. https://doi.org/10.1016/j.ifacol.2024.09.124

[27] van Dinter, R., Catal, C., Arends, O., et al. (2026). Reference architecture for IoT-based predictive maintenance systems using digital twins. Results in Engineering, 30, 111115. https://doi.org/10.1016/j.rineng.2026.111115

[28] Kapteyn, M. G., Pretorius, J. V. R., & Willcox, K. E. (2021). A probabilistic graphical model foundation for enabling predictive digital twins at scale. Nature Computational Science, 1, 337–347. https://doi.org/10.1038/s43588-021-00069-0

[29] Soesanto, J. F., Maciszewski, B., Mirmontazeri, L., Romero, S., Michonski, M., Milne, A., & Huang, B. (2024). Digital twin and control of an industrial-scale bitumen extraction process. Computers & Chemical Engineering, 185, 108671. https://doi.org/10.1016/j.compchemeng.2024.108671

[30] Leveraging Digital Twin for operational resilience in the oil and gas industry. (2025). Array, 27, 100443. https://doi.org/10.1016/j.array.2025.100443

[31] Amin, M. T., & Khan, F. (2022). Dynamic process safety assessment using adaptive Bayesian network with loss function. Industrial & Engineering Chemistry Research, 61(45), 16799–16814. https://doi.org/10.1021/acs.iecr.2c03080

[32] He, R., Zhu, J., Chen, G., & Tian, Z. (2022). A real-time probabilistic risk assessment method for the petrochemical industry based on data monitoring. Reliability Engineering & System Safety, 226, 108700. https://doi.org/10.1016/j.ress.2022.108700

[33] Moradi, R., Cofre-Martel, S., Lopez Droguett, E., Modarres, M., & Groth, K. M. (2022). Integration of deep learning and Bayesian networks for condition and operation risk monitoring of complex engineering systems. Reliability Engineering & System Safety, 222, 108433. https://doi.org/10.1016/j.ress.2022.108433

[34] Tarcsay, B. L., Bárkányi, Á., Németh, S., Chován, T., Lovas, L., & Egedy, A. (2024). Risk-based fault detection using Bayesian networks based on failure mode and effect analysis. Sensors, 24(11), 3511. https://doi.org/10.3390/s24113511

[35] Braniff, A., Akundi, S. S., Liu, Y., Dantas, B., Niknezhad, S. S., Khan, F., Pistikopoulos, E. N., & Tian, Y. (2025). Real-time process safety and systems decision-making toward safe and smart chemical manufacturing. Digital Chemical Engineering, 15, 100227. https://doi.org/10.1016/j.dche.2025.100227

[36] Downs, J. J., & Vogel, E. F. (1993). A plant-wide industrial process control problem. Computers & Chemical Engineering, 17(3), 245–255. https://doi.org/10.1016/0098-1354(93)80018-I

2026 by the Authors. This Article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/)