International Journal of
Computing and
Engineering
(IJCE)

A Physics-Informed Multi-Agent Reinforcement Learning Framework for Autonomous Process Safety Barrier Management in Petrochemical Facilities



International Journal of Computing and Engineering
ISSN 2958-7425 (online)
Vol. 8, Issue No. 4, pp. 94 - 120, 2026
www.carijournals.org

A Physics-Informed Multi-Agent Reinforcement Learning Framework for Autonomous Process Safety Barrier Management in Petrochemical Facilities

[1*]Hafiz Abid Ali, [2]Tehreem Ahmad
[1]MSC Safety Engineering, Mary Kay O'Connor Process Safety Center/TEES/Texas A&M University, USA
[2]Computer Science, Lahore University of Management Sciences, Lahore Pakistan

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

Abstract

Process safety management in petrochemical facilities remains challenging due to complex operating conditions, equipment degradation, and limitations in conventional monitoring systems. This study proposes a Physics-Informed Multi-Agent Reinforcement Learning (PI-MARL) framework for supervisory process safety barrier management. The architecture integrates physics-informed neural networks, cooperative learning agents, equipment reliability assessment, and safety-constrained decision support. A continuous stirred-tank reactor is selected as the representative process for a proposed simulation-based evaluation involving operating disturbances, equipment degradation, and impaired protective barriers. The framework incorporates physical governing equations, safety-aware reward functions, and supervisory action screening while maintaining the independence of established safety instrumented systems. The evaluation protocol specifies comparisons with conventional control, model predictive control, and safe reinforcement learning methods using safety violations, barrier availability, prediction errors, and computational performance. The study establishes a conceptual and mathematical foundation for evaluating coordinated, risk-informed supervisory decisions. Experimental validation is required to determine its effectiveness and practical applicability.

Keywords: *Physics-Informed Machine Learning; Multi-Agent Reinforcement Learning; Process Safety Barrier Management; Equipment Reliability; Supervisory Control.*

1. Introduction

Petrochemical facilities operate under potentially hazardous conditions involving elevated temperatures, high pressures, flammable substances, toxic chemicals, and interconnected processing equipment. These conditions require systematic process hazard identification, equipment integrity management, and reliable protective measures to prevent or minimize industrial accidents (Occupational Safety and Health Administration [OSHA], 1992). Process safety barriers, including pressure relief devices, emergency shutdown systems, safety instrumented systems, alarms, and containment mechanisms, play essential roles in preventing hazardous events and limiting their consequences. The OSHA Process Safety Management standard, 29 CFR 1910.119, establishes requirements for process hazard analysis, mechanical integrity, operating procedures, inspection, and preventive maintenance. However, periodic hazard assessments and equipment inspections may not independently capture rapidly evolving operational conditions or interactions between equipment degradation and process disturbances. These limitations motivate research into predictive monitoring and safety-constrained decision-support approaches for industrial process safety management.

Artificial intelligence (AI) and machine learning have emerged as promising technologies for improving industrial monitoring, fault detection, and operational decision-making. Reinforcement learning (RL), in particular, enables intelligent agents to develop decision-making policies through interactions with dynamic environments. Nian et al. (2020) highlighted the potential of reinforcement learning for industrial process control, particularly in applications requiring adaptive optimization under changing operating conditions. Nevertheless, conventional reinforcement learning approaches may require extensive training data and can generate unsafe actions when physical operating constraints are insufficiently represented. These limitations become especially significant in petrochemical environments, where inappropriate control decisions may result in equipment damage, process instability, or serious safety incidents.

Physics-informed machine learning offers a potential solution by incorporating established physical principles into data-driven modelling. Rather than relying exclusively on historical observations, physics-informed models integrate governing equations, operational boundaries, and process dynamics into their learning procedures. Raissi et al. (2019) demonstrated the effectiveness of physics-informed neural networks for modelling nonlinear physical systems, while Karniadakis et al. (2021) established their broader relevance to scientific and engineering applications. Incorporating such principles into industrial AI systems may improve model consistency, reduce dependence on extensive datasets, and support more reliable predictions under varying operating conditions.

Multi-agent reinforcement learning (MARL) provides an additional opportunity to address the distributed nature of petrochemical safety management. Instead of assigning all monitoring and decision-making responsibilities to a single intelligent controller, MARL enables specialized

agents to coordinate activities across interconnected systems. Gronauer and Diepold (2022) examined the foundations of multi-agent deep reinforcement learning, while Sass et al. (2022) demonstrated its application to chemical reactor operation. Furthermore, Szatmári et al. (2024) explored resilience-based reinforcement learning for chemical process safety, highlighting opportunities for intelligent decision support in safety-critical industrial environments.

Although previous studies have investigated physics-informed learning, cooperative reinforcement learning, and dynamic safety barrier assessment, questions remain regarding their combined effectiveness in safety-constrained supervisory decision-making. Yuan et al. (2023) demonstrated dynamic risk-informed safety barrier management, while Tayal et al. (2025) explored decentralized physics-informed multi-agent control. However, the comparative benefits of integrating physics-informed state prediction, condition-dependent barrier reliability assessment, and cooperative supervisory decisions under plant-model mismatch and interacting equipment failures remain insufficiently established. The specific research gap therefore concerns whether this integrated architecture provides measurable advantages over established safe reinforcement learning, conventional estimation, and model predictive control approaches.

To investigate this question, the present study proposes a Physics-Informed Multi-Agent Reinforcement Learning (PI-MARL) framework for supervisory process safety barrier management in petrochemical systems. The framework combines physics-informed process prediction, condition-dependent equipment and barrier assessment, cooperative decision-making, and independent supervisory safety validation. Three research hypotheses guide the proposed evaluation: physics-informed prediction improves state estimation under plant-model mismatch; multi-agent coordination improves responses to interacting equipment and barrier conditions; and supervisory safety filtering reduces unsafe control actions. A simulation-based evaluation protocol using a continuous stirred-tank reactor is developed to test these hypotheses against conventional and safety-constrained control approaches. The framework remains a proposed methodology whose effectiveness requires experimental validation.

2. Literature Review

The integration of artificial intelligence (AI), reinforcement learning (RL), and physics-informed modelling into petrochemical process safety has attracted increasing research attention. However, the use of learning-based methods in safety-critical environments requires more than accurate process prediction or improved control performance. It requires explicit consideration of operating constraints, uncertainty, equipment degradation, protective barrier reliability, and the consequences of inappropriate control actions. Existing research provides important foundations in safe reinforcement learning, multi-agent coordination, physics-informed learning, and dynamic safety barrier management. This review critically examines these approaches, their limitations, and their relevance to the proposed Physics-Informed Multi-Agent Reinforcement Learning (PI-MARL) framework.

2.1. Safe Reinforcement Learning in Chemical Process Control

Reinforcement learning has been investigated as an alternative to conventional optimization and control methods for nonlinear industrial processes. Nian et al. (2020) reviewed its applications in industrial process control, identifying opportunities for adaptive decision-making and optimization under changing operating conditions. Nevertheless, reinforcement learning policies trained primarily to maximize cumulative rewards may generate unsafe actions when process constraints are represented inadequately. This limitation is particularly important in petrochemical facilities, where temperature excursions, pressure increases, and equipment malfunctions may lead to hazardous events.

Recent research has consequently focused on safe reinforcement learning techniques that incorporate operating constraints into policy development and execution. These approaches include constrained Markov decision processes, control barrier functions, model predictive control (MPC), action shielding, and constraint tightening. Devarakonda et al. (2025) reviewed reinforcement learning developments in chemical process control and identified constraint tightening, control barrier functions, and control-invariant sets as important approaches to safety-aware learning.

Kim and Oh (2024) proposed a model-based safe reinforcement learning method for nonlinear systems under uncertainty using constraint tightening. Their approach incorporated Gaussian process models to estimate constraint errors and demonstrated chance-constraint handling using a four-tank simulation. This study shows that safety-aware reinforcement learning can explicitly account for uncertainty instead of relying exclusively on reward penalties. However, its principal focus was constrained process control rather than coordinated equipment degradation and protective barrier management.

More recently, Gerold and Lucia (2026) developed an adaptive robust model predictive shielding approach that evaluates reinforcement learning actions and substitutes a backup control policy when necessary. Their continuous stirred-tank reactor simulations demonstrated the relevance of robust shielding for safe learning-based process control. Similarly, Durkin et al. (2026) investigated safe deployment of offline reinforcement learning through input-convex action correction in an exothermic polymerization reactor. These studies establish stronger methodological benchmarks for the proposed PI-MARL framework than conventional threshold-based monitoring alone.

The implications for PI-MARL are substantial. Its proposed supervisory mechanism must be assessed against established safe reinforcement learning approaches, particularly regarding constraint satisfaction, uncertainty handling, intervention frequency, and computational requirements. The combination of multiple learning agents cannot itself establish superior operational safety.

2.2. Physics-Informed Machine Learning and Predictive Process Modelling

Physics-informed machine learning incorporates physical relationships into model development to improve the consistency of predictions with established governing equations. Raissi et al. (2019) introduced physics-informed neural networks (PINNs) for solving forward and inverse problems involving differential equations. Karniadakis et al. (2021) further discussed the integration of physical knowledge with machine learning across scientific and engineering applications.

In petrochemical processing, physical relationships include mass conservation, energy balances, reaction kinetics, heat transfer, and equipment operating constraints. These relationships can be represented within neural network training through residual-based loss functions. Incorporating such information may improve process state estimation when observations are limited or measurement uncertainty is significant.

However, the necessity of a PINN depends on the modelling problem. When a system is described adequately by known differential equations, conventional numerical integration, state observers, or extended Kalman filters may provide simpler alternatives. The potential value of a PINN becomes more relevant when process parameters are uncertain, measurements are incomplete, or physical models differ from actual plant behaviour.

Accordingly, the proposed PI-MARL framework should not assume that incorporating physics-informed neural networks automatically improves prediction accuracy or control safety. Their contribution must be evaluated through comparison with conventional model-based estimation methods, particularly under parameter uncertainty, sensor degradation, and plant-model mismatch.

Physics-informed learning and safety constraint enforcement must also remain conceptually distinct. Minimizing violations of physical governing equations can improve model consistency, but it does not guarantee that subsequent reinforcement learning actions will satisfy industrial safety limits. An independently justified action-screening mechanism remains necessary.

2.3. Multi-Agent Reinforcement Learning and Coordinated Safety Decisions

Multi-agent reinforcement learning extends conventional reinforcement learning by distributing decision-making responsibilities among multiple interacting agents. Gronauer and Diepold (2022) provided a comprehensive examination of multi-agent deep reinforcement learning, including cooperation, communication, coordination, and challenges associated with non-stationary learning environments.

Within chemical processing, Sass et al. (2022) investigated multi-agent reinforcement learning for optimizing semi-batch reactor operation. Their work demonstrated the applicability of coordinated learning to chemical process operation. However, the objectives of reactor optimization differ from those of protective barrier management, where equipment integrity, failure probabilities, and safety requirements must influence operational decisions.

Szatmári et al. (2024a) investigated the integration of resilience engineering and reinforcement learning in chemical process safety, demonstrating how Deep Q-learning can support intervention

decisions in exothermic reactor operations. Subsequently, Szatmári et al. (2024b) extended this research by incorporating explainable reinforcement learning techniques, including decision trees and Shapley values, to improve the interpretability of mitigation system activation decisions. Together, these studies establish a foundation for resilience-oriented intelligent safety interventions, although coordinated equipment reliability assessment and multi-agent barrier management remain outside their primary scope.

Recent developments further demonstrate that safety-aware multi-agent systems are an active research area. González-Potes et al. (2026) developed a safety-aware multi-agent deep reinforcement learning approach for fault-tolerant control in industrial cleaning systems. Their work combined decentralized decision-making, safety mechanisms, and adaptation to sensor limitations. Although the application differs from petrochemical barrier management, it establishes an important competing approach involving industrial safety and coordinated learning.

Tayal et al. (2025) also proposed MAD-PINN, a decentralized physics-informed framework for safe multi-agent control. Its reported application concerned multi-agent navigation rather than petrochemical facilities, and the available version is a preprint. Nevertheless, the study demonstrates that combining physics-informed learning with decentralized safety-aware control is not, by itself, an unexplored research direction.

These developments indicate that the proposed PI-MARL framework requires a more precise justification for its agent structure. Process monitoring and equipment reliability estimation may be performed by dedicated predictive models rather than reinforcement learning agents. Therefore, the value of four specialized agents should be established through comparisons with simpler single-agent and estimator-assisted architectures.

2.4. Dynamic Safety Barrier Management and Equipment Reliability

Safety barriers are critical protective measures intended to prevent hazardous events or reduce their consequences. In petrochemical facilities, these include pressure relief devices, emergency shutdown systems, safety instrumented functions, containment systems, and associated operational procedures. Their effectiveness depends on component reliability, degradation, diagnostic coverage, maintenance, proof testing, and potential common-cause failures.

Traditional reliability assessments commonly employ historical failure information and predefined probability estimates. Although useful for engineering design and regulatory compliance, fixed estimates may not fully represent changing equipment conditions.

Yuan et al. (2023) developed a dynamic-risk-informed safety barrier management methodology combining condition monitoring, periodic proof-test information, accident precursor data, and cost-effectiveness analysis. Their Simulink-based framework demonstrated how changing barrier failure probabilities could inform risk assessment and barrier optimization. Importantly, the study already integrated several functions that the proposed PI-MARL framework intends to address,

including degradation-sensitive barrier modelling and dynamic risk updating. Therefore, the research gap cannot reasonably be defined as the absence of integrated barrier assessment.

The distinction between component reliability and safety barrier availability also requires attention. For safety instrumented functions operating in low-demand mode, probability of failure on demand (PFD) is an important measure. Barrier configurations may incorporate redundant elements and voting arrangements, making a simple series reliability formulation insufficient for all applications.

IEC 61511 provides the functional safety framework for safety instrumented systems in process industries. It addresses requirements associated with specification, design, operation, maintenance, and the safety lifecycle. Learning-based supervisory systems must therefore be evaluated within established industrial safety responsibilities rather than treated as replacements for independently engineered protective functions.

The specific opportunity for PI-MARL concerns linking dynamically updated barrier conditions with coordinated, safety-constrained operational decisions. Whether this integration produces measurable improvements beyond existing dynamic risk assessment and conventional control methods remains an empirical question.

2.5. Research Synthesis and Identified Gap

The literature demonstrates that the principal technological components proposed for PI-MARL already exist in related research areas. Safe reinforcement learning provides methods for constrained decision-making, physics-informed learning supports process modelling, multi-agent methods enable coordinated policies, and dynamic barrier management incorporates equipment degradation into risk assessment.

The remaining research question is whether combining these capabilities produces measurable benefits when equipment conditions, protective barriers, and operating decisions interact.

For the present study, the proposed contribution is therefore defined more narrowly as the development and evaluation of a supervisory architecture integrating physics-informed process prediction, condition-dependent safety barrier assessment, and coordinated learning-based response selection under explicitly specified operating constraints.

This contribution must be evaluated against alternative architectures that combine conventional process models, predictive maintenance estimators, single-agent safe reinforcement learning, and independent supervisory safety filters.

Table 1. Comparative Analysis of Relevant Research Approaches and Implications for PI-MARL

Study	Primary approach	Safety or reliability contribution	Relevance and remaining question for PI-MARL
Raissi et al. (2019)	Physics-informed neural networks	Physical consistency in predictive modelling	Whether PINNs outperform conventional process estimation under realistic uncertainty
Nian et al. (2020)	Review of industrial reinforcement learning	Adaptive industrial process control	Need for explicit constraint enforcement
Sass et al. (2022)	Multi-agent reinforcement learning	Coordinated reactor operation	Whether multiple agents improve barrier-related decisions
Yuan et al. (2023)	Dynamic barrier modelling and risk assessment	Updated barrier failure probabilities and risk-informed optimization	Whether learning-based coordination adds benefits beyond established barrier optimization
Kim and Oh (2024)	Model-based safe RL with constraint tightening	Uncertainty-aware constraint handling	Strong comparison for proposed safety supervision
Szatmári et al. (2024)	Resilience-based explainable RL	Hazard mitigation and interpretable decisions	Need to evaluate distributed barrier and equipment interactions
Tayal et al. (2025)	Decentralized physics-informed multi-agent control	Safety-aware coordination	Related method in a different application; preprint evidence
Gerold and Lucia (2026)	Robust model predictive shielding	Safety screening and backup control	Strong benchmark for the proposed supervisory filter
Durkin et al. (2026)	Safe offline RL with action correction	Safer deployment of learned reactor control	Alternative to PI-MARL action correction
González-Potes et al. (2026)	Safety-aware multi-agent fault-tolerant control	Distributed control under sensor limitations	Need to demonstrate petrochemical-specific advantages
Proposed PI-MARL	Physics-informed prediction, barrier reliability modelling, and supervised cooperative learning	Intended coordination of predictive risk and safety-related responses	Requires implementation, ablation experiments, and quantitative validation

The comparison presented in Table 1 provides a targeted methodological synthesis of existing research rather than a systematic literature review. It highlights the contributions and limitations of relevant approaches while identifying opportunities for integrating physics-informed prediction, cooperative decision-making, and dynamic safety barrier assessment.

Based on the identified research gaps, four testable hypotheses are proposed. H1: Physics-informed prediction reduces process-state estimation errors under plant-model mismatch compared with conventional estimation methods. H2: Cooperative multi-agent decision-making reduces cumulative process risk and safety constraint violations under interacting equipment degradation and process disturbances compared with an equivalent single-agent architecture. H3: Independent supervisory safety validation reduces unsafe action execution and operating-limit violations compared with reinforcement learning without supervisory filtering. H4: Integrating

condition-dependent equipment reliability and protective barrier assessment improves barrier availability and reduces risk-weighted barrier unavailability compared with conventional condition-based barrier management. These hypotheses will be evaluated through comparative simulations using predefined performance metrics, numerical acceptance thresholds, and statistical significance criteria.

3. Proposed Framework and System Architecture

The proposed Physics-Informed Multi-Agent Reinforcement Learning (PI-MARL) framework is designed as a supervisory decision-support architecture for process safety management in petrochemical systems. Its objective is not to replace safety instrumented systems or emergency shutdown logic, but to support earlier detection of abnormal conditions, improved interpretation of equipment degradation, and more coordinated preventive decision-making under explicitly defined safety constraints. The framework combines physics-informed state prediction, condition-based barrier assessment, equipment reliability estimation, cooperative reinforcement learning, and an independent supervisory safety filter. This integration is intended to address the interaction between process deviations, protective barrier condition, and equipment health, which are often treated separately in conventional monitoring and control strategies (Raissi et al., 2019; Karniadakis et al., 2021; Gronauer & Diepold, 2022).

The architecture contains five connected functional layers: (1) petrochemical process and sensing, (2) data acquisition and state estimation, (3) physics-informed predictive modelling, (4) cooperative multi-agent decision-making, and (5) supervisory safety validation and response execution. The first layer represents the physical process environment, including the reactor unit, associated cooling system, feed system, pumps, valves, pressure relief devices, sensors, and logic-based protective functions. For the initial case study, a continuous stirred-tank reactor (CSTR) is used as the representative unit because it captures nonlinear reaction behaviour, thermal instability, and disturbance-response interactions relevant to petrochemical safety. The monitored variables include reactant concentration, reactor temperature, coolant temperature, reactor pressure, flow conditions, alarm status, and selected degradation indicators.

The second layer performs data acquisition and state estimation. Sensor measurements, alarm signals, maintenance information, and protective barrier status are collected and preprocessed before they are used by the predictive and decision-making modules. Preprocessing includes timestamp alignment, normalization, missing-value handling, and uncertainty tagging. State estimation is used to reconstruct unmeasured or noisy states and to provide a consistent observation vector for downstream modules. This layer is necessary because reinforcement learning and reliability assessment depend on timely and coherent plant-state information, especially when disturbances or sensor degradation occur.

The third layer is the physics-informed predictive module. This component uses first-principles process relations and a physics-informed neural network to estimate future process behaviour and

detect deviations from expected dynamics. The motivation for this component is not that first-principles models are unavailable, but that in realistic settings process parameters, disturbance magnitudes, and measurement quality may be uncertain. The physics-informed model therefore acts as a predictive estimator under imperfect information rather than as a replacement for classical simulation. Its role is to improve robustness under plant-model mismatch and noisy observations. The module receives process measurements and control inputs, then outputs predicted short-horizon process states and residuals that indicate inconsistency between measured behaviour and expected physical dynamics. This predictive layer provides structured information to the decision-making agents but does not itself authorize control actions.

The fourth layer incorporates a cooperative decision-making architecture supported by three specialized estimation modules and a reinforcement learning-based Coordination and Response Agent. The Process Monitoring Module evaluates deviations between observed and predicted process states, generating anomaly indicators associated with abnormal temperature, concentration, and pressure conditions. The Safety Barrier Assessment Module estimates the operating condition and availability of protective functions, including pressure relief devices, safety instrumented functions, and cooling-related protective systems. The Equipment Reliability Module estimates degradation states, component failure likelihood, and remaining useful life using equipment condition indicators. These modules employ physics-informed estimation, probabilistic reliability modelling, and diagnostic algorithms rather than reinforcement learning because their primary responsibilities involve state estimation rather than action selection. Their outputs are combined by the Coordination and Response Agent, which applies reinforcement learning to select permissible supervisory responses. This arrangement distinguishes predictive assessment from sequential decision-making and provides a more technically justified architecture. However, the revised configuration should be described as a physics-informed, reinforcement learning-based supervisory framework, rather than a four-agent MARL system, unless additional learning agents with explicit action spaces are implemented.

This multi-agent structure is implemented using centralized training with decentralized execution. During training, a shared critic evaluates the joint effect of agent decisions on safety, stability, and equipment stress, while each agent has a specialized observation space. The Process Monitoring Agent observes process states and prediction residuals. The Barrier Assessment Agent observes barrier condition variables, trip status, and protection demand indicators. The Equipment Reliability Agent observes degradation variables, operating stress, and maintenance-related signals. The Coordination and Response Agent observe summarized outputs from all other agents together with current risk indicators. The action spaces are discrete or hybrid supervisory actions rather than direct manipulation of safety-critical shutdown logic. A suitable learning algorithm for this structure is Multi-Agent Proximal Policy Optimization (MAPPO), since it supports cooperative learning with centralized training and decentralized policies. This is more appropriate than leaving the learning method unspecified.

The reward structure is shared but operationally decomposed. It penalizes predicted safety violations, repeated excursions toward hazardous states, excessive equipment stress, unnecessary interventions, and loss of barrier availability, while rewarding stable operation and timely preventive action. This design is intended to align learning with process safety objectives. However, because reward shaping alone cannot guarantee safety, the framework includes a separate supervisory layer.

The fifth layer is the supervisory safety validation module. This component screens every candidate action proposed by the Coordination and Response Agent before implementation. The safe-action set is defined using operating constraints on temperature, pressure, coolant conditions, and allowable supervisory interventions. If a proposed action would violate a constraint or increase the risk of unsafe evolution beyond an accepted threshold, it is rejected and replaced with the nearest admissible supervisory action or with operator escalation. This design ensures that learning-based decisions remain subordinate to established process safety rules. Importantly, safety instrumented systems, emergency shutdown logic, and pressure relief functions remain external protective layers and are never overridden by PI-MARL. Therefore, the architecture should be interpreted as a bounded supervisory decision-support framework rather than fully autonomous barrier management.

A continuous feedback loop links all layers. New sensor information updates the state estimator, which informs the physics-based predictor, which then updates the agent observations and supervisory evaluation. Through this loop, the framework can be evaluated under normal operation, cooling disturbances, pressure excursions, sensor faults, equipment degradation, and reduced barrier availability. The architecture therefore provides a structured basis for testing three main hypotheses: that physics-informed prediction improves state estimation under mismatch, that cooperative agent coordination improves supervisory response quality under interacting hazards, and that safety filtering reduces unsafe decisions relative to unconstrained reinforcement learning.

International Journal of Computing and Engineering
ISSN 2958-7425 (online)
Vol. 8, Issue No. 4, pp. 94 - 120, 2026
www.carijournals.org

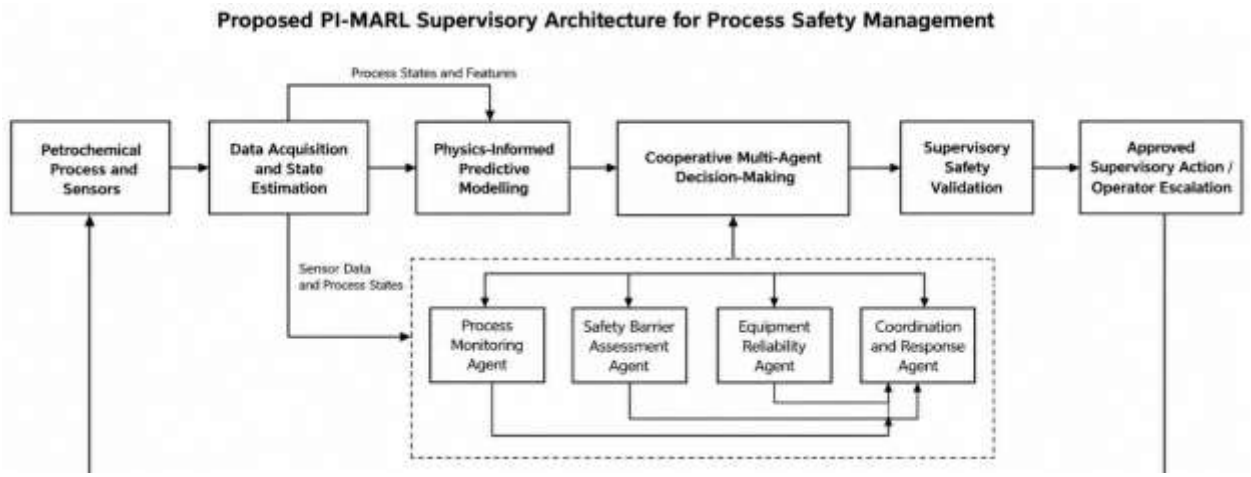

Figure 1: Proposed PI-MARL supervisory architecture for process safety management, showing information flow from petrochemical process sensing to state estimation, physics-informed prediction, cooperative multi-agent decision-making, supervisory safety validation, and bounded response execution.

4. Mathematical Modelling and Methodology

The proposed Physics-Informed Multi-Agent Reinforcement Learning (PI-MARL) framework is formulated as a **constrained, partially observable multi-agent decision problem** for a simulated continuous stirred-tank reactor (CSTR) and its associated cooling and protective functions. Its purpose is to test whether physically informed state prediction, condition-aware barrier assessment, and coordinated supervisory decisions offer advantages over simpler safe-control architectures. The equations below define a proposed simulation model, not a plant-calibrated implementation or a demonstrated safety guarantee. Physics-informed modelling follows the general principles of Raissi et al. (2019) and Karniadakis et al. (2021), while cooperative learning is informed by Gronauer and Diepold (2022).

A representative exothermic, first-order liquid-phase reaction is considered, with a jacket cooling loop and a simplified gas-containing headspace. The proposed state and input vectors are

$$\mathbf{x}(t) = \left[C_A(t),\ T(t),\ T_c(t),\ P(t),\ h(t)\right]^{\mathrm{T}}, \qquad \mathbf{u}(t) = \left[F(t),\ F_c(t),\ u_v(t)\right]^{\mathrm{T}}. \qquad (1)$$

Here, C_A is reactant concentration (mol m $^{-3}$), T reactor temperature (K), T_c jacket temperature (K), P headspace pressure (Pa), and $h \in [0,1]$ a dimensionless equipment-health indicator, with $h = 1$ representing nominal condition. F and F_c are reactor-feed and coolant-flow rates (m ^{3} s

$^{-1}$), and u_v is a bounded valve command. A fixed reactor liquid volume V and well-mixed conditions are assumed. The material balance is

$$\frac{dC_A}{dt} = \frac{F}{V}(C_{A,f} - C_A) - k(T)C_A, \qquad k(T) = k_0 \exp\left(-\frac{E_a}{R_g T}\right). \qquad (2)$$

In Eq. (2), $C_{A,f}$ is feed concentration, k_0 the Arrhenius pre-exponential factor (s $^{-1}$), E_a the activation energy (J mol $^{-1}$), and R_g the universal gas constant. Assuming constant density ρ and heat capacity C_p, the reactor energy balance is

$$\frac{dT}{dt} = \frac{F}{V}(T_f - T) + \frac{(-\Delta H)k(T)C_A}{\rho C_p} - \frac{UA(h)}{\rho C_p V}(T - T_c). \qquad (3)$$

Here, T_f is feed temperature; $-\Delta H > 0$ is exothermic heat release (J mol $^{-1}$); and $UA(h)$ is the effective heat-transfer conductance (W K $^{-1}$). To couple degradation to heat-removal capacity, an explicitly assumed relationship is $UA(h) = UA_0[\eta + (1-\eta)h]$, with a floor factor $0 < \eta \leq 1$. This relationship is a test assumption requiring calibration, not an established universal degradation law. Coolant dynamics are represented by

$$\frac{dT_c}{dt} = \frac{F_c}{V_c}(T_{c,in} - T_c) + \frac{UA(h)}{\rho_c C_{p,c} V_c}(T - T_c), \qquad (4)$$

where V_c, $T_{c,in}$, ρ_c, and $C_{p,c}$ are jacket volume, coolant inlet temperature, coolant density, and coolant heat capacity. The pressure state is retained only if the simulated reactor includes a physically specified headspace and vapor/gas-generation mechanism. Under an illustrative ideal-gas, fixed-volume headspace approximation,

$$\frac{dP}{dt} = \frac{R_g T_g}{V_g}\left[\dot{n}_{g,gen}(\mathbf{x}) - \dot{n}_{g,out}(P, u_v)\right] + \frac{P}{T_g}\frac{dT_g}{dt}. \qquad (5)$$

Here, V_g and T_g are headspace volume and gas temperature, and the molar generation and discharge functions must be specified and validated for the selected chemistry and vent design. For a non-gas-generating benchmark, pressure should instead be omitted from the state rather than

International Journal of Computing and Engineering
ISSN 2958-7425 (online)
Vol. 8, Issue No. 4, pp. 94 - 120, 2026
www.carijournals.org

assigned unsupported dynamics. Equipment health is described by a bounded phenomenological degradation law,

$$\frac{dh}{dt} = -\alpha_0 - \alpha_T[\max(0, T - T_{\text{stress}})]^2 - \alpha_P[\max(0, P - P_{\text{stress}})]^2, \qquad 0 \le h \le 1. \qquad (6)$$

The nonnegative α parameters set degradation rates and require independent calibration or sensitivity analysis. When pressure is omitted, its stress term is also omitted. Repair or maintenance, if simulated, is a discrete reset event and is not implied by Eq. (6). Equations (2)–(6) thus specify the five proposed states, subject to the explicitly stated conditional pressure assumption. Initial states, disturbance distributions, control limits, solver tolerances, and physical parameters must be documented in the experimental configuration; numerical values are not claimed here.

The measurement vector available to estimation modules is $\mathbf{y}_t = H\mathbf{x}_t + \mathbf{v}_t$, with documented sampling interval Δt, observation matrix H, and sensor noise $\mathbf{v}_t$. Sensor bias, dropout, and delay are separate test disturbances. A physics-informed neural network $\hat{\mathbf{x}}_\vartheta(t)$ is proposed to estimate selected states from available observations and input histories. Training minimizes

$$\mathcal{L}_{\text{PINN}} = \lambda_y \mathcal{L}_{\text{obs}} + \lambda_f \mathcal{L}_{\text{phys}} + \lambda_0 \mathcal{L}_{\text{initial}}, \qquad \mathcal{L}_{\text{phys}} = \frac{1}{N_f}\sum_{q=1}^{N_f} \left\| \frac{d\hat{\mathbf{x}}_\vartheta}{dt}(t_q) - \mathbf{f}(\hat{\mathbf{x}}_\vartheta(t_q), \mathbf{u}(t_q)) \right\|_W^2 . \qquad (7)$$

Only physically defined state derivatives enter the residual, and the weighting matrix W and loss coefficients $\lambda_y, \lambda_f, \lambda_0$ must be disclosed and tuned using validation data. Collocation points t_q cover the sampled operating envelope, including transient disturbances; adaptive sampling and stiff-solver comparisons are planned for runaway-prone trajectories. The predicted states are not treated as certified bounds. To establish whether the PINN is necessary, it must be compared with direct numerical integration and a conventional state observer or extended Kalman filter under deliberately introduced plant–model mismatch.

The cooperative control problem is represented as a constrained decentralized partially observable Markov decision process (Dec-POMDP),

International Journal of Computing and Engineering
ISSN 2958-7425 (online)
Vol. 8, Issue No. 4, pp. 94 - 120, 2026
www.carijournals.org

$$\mathcal{M} = \langle \mathcal{I}, \mathcal{S}, \{\mathcal{O}_i\}, \{\mathcal{A}_i\}, \mathrm{P}, \{\mathrm{O}_i\}, r, \{c_j, d_j\}, \gamma \rangle, \quad \mathbb{E}_\pi \left[\sum_{t=0}^{\infty} \gamma^t \, c_j(s_t, \mathbf{a}_t) \right] \leq d_j. \qquad (8)$$

Here, $\mathcal{I}$ is the agent set; $\mathcal{S}$ the latent process and barrier state; $\mathcal{O}_i$ and $\mathcal{A}_i$ each agent's observation and action spaces; P the transition kernel; O_i its observation function; r a shared reward; c_j constraint costs; d_j cost budgets; and $\gamma \in (0,1)$ the discount factor. Physical limit exceedance, excessive barrier demand, and unapproved interventions are separately defined constraint costs. Each agent observes only measured or estimated quantities within its assigned scope. The monitoring and reliability modules primarily produce diagnostic or prognostic outputs; whether they need **learned actions** rather than standard estimators is an open architectural question to be tested by ablation. The coordination/response agent selects supervisory recommendations or bounded commands, while the barrier-assessment module supplies estimates of proof-test status and unavailability. Centralized training with decentralized execution can be implemented using multi-agent proximal policy optimization (MAPPO), with the algorithm, network sizes, communication protocol, and hyperparameters fixed and reported before evaluation.

For an interpretable training objective, the proposed shared reward is

$$r_t = -w_R \tilde{\mathcal{R}}_t - w_V V_t - w_D D_t - w_C C_t - w_B \left(1 - A_{b,t}\right) - w_I I_t, \qquad (9)$$

where $\tilde{\mathcal{R}}_t$ is normalized estimated process risk, V_t the magnitude and duration of operating-limit violations, D_t incremental health loss, C_t normalized intervention cost, $A_{b,t}$ estimated barrier availability, and I_t unnecessary protective intervention count. All terms must be calculated from declared models and normalized to comparable ranges; weights $w_\cdot \geq 0$ require sensitivity analysis. Reward penalties do **not** establish hard safety guarantees.

Equipment reliability is described conditionally through a calibrated, nonnegative hazard-rate model,

$$R_i(t|\mathcal{H}_t) = \exp \left[- \int_0^t \lambda_i \left(\tau|\mathcal{H}_\tau\right) d\tau \right], \qquad (10)$$

International Journal of Computing and Engineering
ISSN 2958-7425 (online)
Vol. 8, Issue No. 4, pp. 94 - 120, 2026
www.carijournals.org

where $\mathcal{H}_t$ denotes observed stress and maintenance history. This operational reliability measure is **not** interchangeable with the probability of failure on demand (PFD) of a dormant safety function. For an illustrative low-demand, single-channel function with approximately constant dangerous undetected failure rate λ_{DU} and perfect periodic proof testing every T_{PT}, the average PFD is approximated by

$$\mathrm{PFD}_{\mathrm{avg},1oo1} \approx \frac{\lambda_{DU} T_{PT}}{2}, \qquad A_b \approx 1 - \mathrm{PFD}_{\mathrm{avg}}. \qquad (11)$$

For an illustrative 1-out-of-2 arrangement with small independent channel PFD q and a simplified common-cause fraction β, one screening approximation is $\mathrm{PFD}_{1oo2} \approx \beta q + (1-\beta)^2 q^2$. This expression is not a substitute for full safety integrity level (SIL) verification: proof-test coverage, diagnostic coverage, repair times, common-cause mechanisms, voting logic, demand mode, and maintenance bypasses must be modelled in accordance with the applicable safety lifecycle (IEC 61508/61511). The manuscript's former series-product expression applies only to genuinely series-dependent functions and is not used as a generic barrier equation.

Finally, a supervisory filter checks proposed actions against a constructively defined admissible set. For an identified uncertainty set $\mathcal{W}$, engineering limit set $\mathcal{X}_{\mathrm{lim}}$, actuator limits $\mathcal{U}$, and independently validated one-step transition model F_{safe},

$$\mathcal{U}_{\mathrm{safe}}(\hat{s}_t) = \{u \in \mathcal{U}:\ F_{\mathrm{safe}}(\hat{s}_t, u, w) \in \mathcal{X}_{\mathrm{lim}}\ \forall w \in \mathcal{W}\}, \quad u_t^* = \arg\min_{u \in \mathcal{U}_{\mathrm{safe}}(\hat{s}_t)} \| u - u_t^{\mathrm{RL}} \|^2. \qquad (12)$$

This one-step screen does not prove recursive feasibility or closed-loop invariance. The plant's independent safety instrumented systems and emergency shutdown logic remain outside RL authority. If the admissible set is empty, state confidence is inadequate, or computation exceeds its allowed time budget, the system must revert to an engineered fallback and operator escalation. Any future safety case must verify the fallback, supervisor independence, model uncertainty bounds, and multi-step reachability for the actual application.

The resulting experimental methodology will compare PI-MARL against matched PID-plus-trip, MPC, shielded single-agent RL, a single agent with the same PINN and supervisor, and physics-uninformed MARL. Ablations will remove physics-informed prediction, agent decomposition, and safety screening one at a time. Tests must include held-out disturbances, combined faults, sensor

degradation, varying proof-test assumptions, plant–model mismatch, multiple random seeds, and confidence intervals. A single CSTR can validate the preliminary method but cannot alone establish coordinated safety-barrier management across an entire petrochemical facility. No numerical improvement or verified safety guarantee is claimed before these simulations are performed.

5. Experimental Setup and Performance Evaluation

The experimental evaluation of the proposed Physics-Informed Multi-Agent Reinforcement Learning (PI-MARL) framework is designed to investigate the effectiveness of physics-informed prediction, cooperative decision-making, and supervisory safety validation in petrochemical process safety management. The evaluation employs a simulation-based protocol to assess process stability, equipment reliability, protective barrier performance, and safety constraint compliance under normal and abnormal operating conditions. Comparative experiments are specified using conventional process controllers, model predictive control, and safety-constrained reinforcement learning approaches. The methodology incorporates disturbance scenarios, plant-model uncertainty, baseline comparisons, ablation studies, and statistical performance evaluation to determine the contributions of individual framework components. As computational experiments have not yet been conducted, the specified procedures constitute a reproducible evaluation design rather than experimentally validated findings.

A continuous stirred-tank reactor (CSTR) is selected as the representative process for evaluating the proposed framework under nonlinear reaction dynamics, thermal disturbances, and equipment degradation. The proposed simulation adopts a jacket-cooled exothermic reactor with a gas-containing headspace and a pressure relief device. Accordingly, the process state vector comprises reactant concentration, reactor temperature, coolant temperature, headspace pressure, and equipment health. Headspace pressure is represented through gas accumulation, temperature-dependent gas behaviour, and discharge through an actuated relief mechanism, as formulated in Section 4. The simulation additionally incorporates cooling disturbances, feed variability, sensor degradation, actuator delays, equipment deterioration, and protective barrier impairment. Plant-model mismatch is introduced through controlled variations in reaction kinetics, heat-transfer efficiency, and degradation parameters to evaluate the robustness of the physics-informed predictive component.

The experimental scenarios are organized into four groups: normal operation, isolated disturbances, combined disturbances, and degraded-protection scenarios. Normal operation assesses control stability and monitoring consistency under expected conditions. Isolated disturbances include cooling failure, abnormal feed conditions, partial sensor malfunction, pump degradation, and transient pressure excursions. Combined disturbances include simultaneous cooling degradation and sensor noise, or equipment degradation occurring under abnormal thermal loading. Degraded-protection scenarios examine cases in which a barrier is unavailable or partially

degraded while the reactor is subject to a process upset. These scenarios are necessary to evaluate whether the proposed framework improves supervisory decisions when process deviations, barrier condition, and equipment reliability interact.

To provide meaningful comparison, PI-MARL should be evaluated against stronger reference approaches than those originally proposed. The revised baseline set therefore includes: (1) PID control with alarm and trip logic, representing conventional practice; (2) model predictive control with constraint handling; (3) single-agent safe reinforcement learning with supervisory action filtering; (4) multi-agent reinforcement learning without physics-informed prediction; and (5) PI-MARL without the safety supervisory layer as an ablation case. These baselines enable assessment of whether performance gains arise from multi-agent coordination, physics-informed state prediction, or independent safety filtering. In addition, a single-agent version using the same predictive model and safety filter should be included to test whether four-agent decomposition is genuinely beneficial.

Training and testing should follow a centralized-training, decentralized-execution strategy for PI-MARL. Each learning-based controller is trained over repeated simulation episodes covering randomized disturbance conditions, initialization states, and degradation profiles. A separate held-out test set should be used for final evaluation and must include disturbance magnitudes and combinations not seen during training. This is necessary to assess generalization beyond memorized scenarios. Rare-event stress testing should also be included, particularly for severe temperature excursions and pressure escalation.

The evaluation metrics should be defined explicitly. Safety performance should be assessed using violation frequency, violation duration, and cumulative violation severity for temperature, pressure, and coolant constraints. Barrier management performance should be measured using predicted barrier availability, protective demand frequency, and missed or unnecessary interventions. Equipment reliability performance should include degradation estimation error, remaining useful life prediction error where applicable, and the frequency of failure-preventive decisions. Process prediction should be evaluated using root mean square error and mean absolute error for key states. Decision-making performance should include cumulative reward, response time, and intervention efficiency. Computational performance should be reported as mean decision latency per control interval. Since safety is the main claim, primary emphasis should be placed on constraint violations and successful mitigation under disturbance.

To support statistical validity, all evaluated approaches should be tested across multiple random seeds and repeated simulation runs. Results should be summarized using means, standard deviations, and 95% confidence intervals. Statistical comparisons should be conducted using appropriate non-parametric or parametric tests depending on distributional assumptions, with effect sizes reported alongside p-values. Multiple-comparison correction should be applied when several controllers are compared simultaneously. Sensitivity analysis should also be performed on

reward weights, mismatch magnitude, and degradation parameters to determine whether the conclusions remain stable under different assumptions.

Ablation studies are essential to justify the architecture. At minimum, separate ablations should remove the physics-informed predictor, the multi-agent structure, and the safety supervisory layer. This design directly tests the three hypotheses established in Section 2. If PI-MARL does not consistently outperform simpler alternatives under matched conditions, then its added complexity may not be justified.

The experimental design therefore provides a stronger basis for evaluating the framework than the original protocol. Nevertheless, until the simulations are actually implemented and the results are reported, this section should be presented as an evaluation plan rather than evidence of performance improvement.

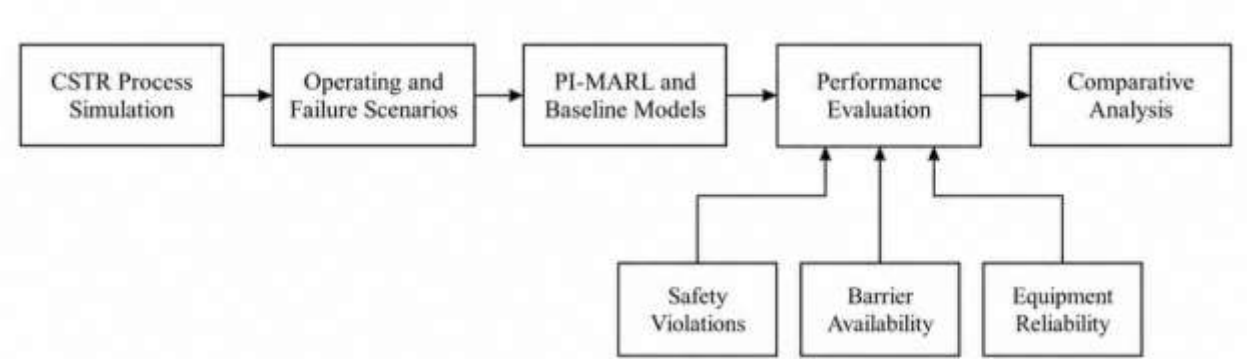

Figure 2: Proposed Experimental Evaluation Workflow for the PI-MARL Framework

6. Expected Outcomes and Evaluation Criteria.

The proposed Physics-Informed Multi-Agent Reinforcement Learning (PI-MARL) framework is intended to integrate process-state prediction, equipment degradation assessment, safety barrier reliability, and coordinated supervisory decision-making. Building on the experimental protocol described in Section 5, this section establishes testable hypotheses, quantitative evaluation criteria, and predefined decision rules for assessing the framework against established control and reinforcement learning approaches. Since computational experiments have not yet been conducted, the criteria represent proposed validation requirements rather than demonstrated performance outcomes.

The first hypothesis (H1) states that physics-informed prediction will reduce process-state estimation errors under plant-model mismatch compared with conventional state estimation and purely data-driven prediction methods. This hypothesis will be evaluated using root mean square error (RMSE) and mean absolute error (MAE) for reactor temperature, reactant concentration, coolant temperature, and pressure. Testing will include nominal operating conditions, parameter uncertainty, measurement noise, and previously unseen disturbances. H1 will be supported only if

PI-MARL's predictive component produces statistically distinguishable improvements under the predefined evaluation conditions.

The second hypothesis (H2) states that cooperative multi-agent decision-making will improve supervisory responses to interacting equipment degradation and safety barrier impairment compared with an equivalent single-agent architecture. To isolate the contribution of agent coordination, both approaches will use identical process models, safety filters, observation information, and evaluation scenarios wherever technically feasible. Performance will be assessed through cumulative risk, intervention effectiveness, violation duration, and equipment degradation indicators. The comparison will also examine whether improved decision quality justifies the additional computational complexity of coordinating multiple agents.

The third hypothesis (H3) states that independent supervisory safety validation will reduce unsafe action execution compared with an equivalent reinforcement learning controller without action filtering. Safety performance will be assessed using constraint violation frequency, cumulative violation severity, rejected unsafe actions, and fallback activation frequency. The supervisory mechanism will also be tested under sensor faults, model uncertainty, and combined process disturbances. Importantly, reduced observed violations in simulation will not constitute proof of safety under all operating conditions.

Equipment reliability and protective barrier performance will be assessed separately using component failure probabilities, estimated barrier unavailability, and preventive intervention frequency. Where sufficient degradation information exists, remaining useful life estimation error will also be evaluated. Barrier performance calculations will account for the assumed voting architecture, proof-test intervals, diagnostic coverage, and common-cause failure mechanisms. These assumptions must be documented to prevent misleading reliability comparisons.

Statistical evaluation will employ repeated independent simulation trials, predefined test scenarios, and consistent performance metrics. Results will be presented using mean estimates, 95% confidence intervals, effect sizes, and appropriate statistical tests. Acceptance thresholds and minimum practically meaningful improvements must be established before model evaluation. Sensitivity analyses will investigate the effects of reward weighting, prediction uncertainty, degradation parameters, and disturbance severity.

The framework will be considered experimentally promising only if it demonstrates measurable safety or reliability advantages over appropriate baselines without unacceptable computational overhead or increased safety violations in critical scenarios. Otherwise, the results should identify which architectural components provide limited benefits or require redesign. These criteria establish a falsifiable evaluation basis for determining the practical value of PI-MARL in supervisory petrochemical process safety management.

Table 2. Proposed Hypotheses, Performance Metrics, and Validation Criteria

Hypothesis	Primary Evaluation Metric	Comparison Baseline	Proposed Numerical Acceptance Criterion
H1: Process-state prediction	RMSE and MAE	Conventional estimator and data-driven model	At least 10% lower RMSE on held-out mismatch scenarios
H2: Multi-agent coordination	Cumulative process risk and constraint violations	Equivalent single-agent safe RL	At least 10% lower cumulative risk, with no increase in severe violations
H3: Safety supervision	Unsafe actions executed and safety violations	Equivalent RL without supervisory filtering	At least 90% reduction in executed unsafe actions
H4: Barrier availability and reliability	Risk-weighted barrier unavailability and equipment failure probability	Conventional condition-based barrier management	At least 10% lower risk-weighted unavailability, with no increase in missed protective demands
Computational feasibility	Decision latency	Predefined supervisory decision deadline	99th-percentile latency below the specified decision deadline

7. Discussion

The proposed Physics-Informed Multi-Agent Reinforcement Learning (PI-MARL) framework addresses the challenge of integrating process prediction, equipment reliability assessment, and coordinated supervisory decisions in petrochemical safety management. Previous studies have established the applicability of reinforcement learning to industrial process control (Nian et al., 2020) and physics-informed neural networks to modelling nonlinear physical systems (Raissi et al., 2019; Karniadakis et al., 2021). However, the potential contribution of PI-MARL lies in determining whether combining these techniques with dynamic barrier assessment provides measurable advantages over established safety-constrained control approaches.

The proposed multi-agent architecture enables specialized functions to exchange information concerning process abnormalities, equipment degradation, and protective barrier conditions. Although Sass et al. (2022) demonstrated multi-agent learning for chemical reactor operation, its additional value for safety barrier management requires experimental verification. Comparisons with single-agent reinforcement learning, model predictive control, and conventional estimation methods are necessary to determine whether the proposed architecture justifies its increased computational complexity.

Safety assurance represents another consideration. Physics-informed predictions do not independently guarantee safe actions, particularly under sensor faults, parameter uncertainty, and plant-model mismatch. Accordingly, the supervisory safety filter must enforce explicitly defined

operating constraints and provide validated fallback responses. Its implementation should also respect IEC 61511 functional safety principles, including the independence of safety instrumented systems, proof-testing requirements, and established management-of-change procedures. Learning-based supervisory decisions must not replace certified protective functions.

The proposed framework has several limitations. First, the CSTR model represents a simplified industrial environment and cannot fully demonstrate coordination across interconnected petrochemical units. Second, barrier reliability estimates depend on assumptions concerning component degradation, voting configurations, proof-test intervals, and common-cause failures. Third, multi-agent learning may introduce policy instability, communication delays, and increased computational requirements. Furthermore, the framework has not undergone simulation testing or industrial validation, preventing conclusions regarding safety improvement or operational superiority.

Future research should implement reproducible CSTR simulations, evaluate stronger baseline controllers, and conduct ablation studies under unseen disturbances and plant-model mismatch. Extending validation to interconnected process units and incorporating representative industrial reliability data would further strengthen practical relevance. Ultimately, PI-MARL remains a proposed supervisory architecture whose safety performance, reliability benefits, and operational feasibility require rigorous experimental demonstration before industrial deployment can be considered.

8. Conclusion

This study proposed a Physics-Informed Multi-Agent Reinforcement Learning (PI-MARL) framework for autonomous process safety barrier management and equipment reliability enhancement in petrochemical facilities. The framework integrates physics-informed neural networks, cooperative reinforcement learning agents, dynamic safety barrier assessment, and equipment degradation modelling to support intelligent, risk-informed decision-making. By incorporating physical governing equations and operational constraints, the proposed approach aims to improve process state prediction, hazard identification, and preventive intervention while maintaining established industrial safety requirements. The research presented a unified system architecture, mathematical formulation, and simulation-based evaluation methodology using a continuous stirred-tank reactor as a representative petrochemical process. The proposed evaluation considers safety violations, barrier availability, equipment reliability, prediction accuracy, and decision-making performance. However, as computational experiments have not been conducted, the effectiveness of the framework remains subject to experimental validation.

Future research should focus on implementing the proposed architecture, conducting comparative simulations, validating equipment degradation models, and assessing robustness under abnormal operating conditions. Ultimately, the proposed PI-MARL framework provides a foundation for

developing adaptive, reliable, and safety-constrained intelligent management systems capable of supporting proactive risk mitigation and industrial operational resilience.

References

[1]. Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, 378, 686–707.

DOI: https://doi.org/10.1016/j.jcp.2018.10.045

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

DOI: https://doi.org/10.1038/s42254-021-00314-5

[3]. Meng, C., Griesemer, S., Cao, D., Seo, S., & Liu, Y. (2025). When physics meets machine learning: A survey of physics-informed machine learning. Machine Learning for Computational Science and Engineering, 1, Article 20.

DOI: https://doi.org/10.1007/s44379-025-00016-0

[4]. Michaloglou, A., Papadimitriou, I., Gialampoukidis, I., Vrochidis, S., & Kompatsiaris, I. (2026). Physics-informed neural networks in materials modeling and design: A review. Archives of Computational Methods in Engineering, 33, 5223–5260.

DOI: https://doi.org/10.1007/s11831-025-10448-9

[5]. Wang, E., et al. (2026). A physics-constrained Bayesian neural network for machinery remaining useful life prediction and uncertainty quantification. Reliability Engineering & System Safety, 266, Article 111778.

DOI: https://doi.org/10.1016/j.ress.2025.111778

[6]. Gronauer, S., & Diepold, K. (2022). Multi-agent deep reinforcement learning: A survey. Artificial Intelligence Review, 55, 895–943.

DOI: https://doi.org/10.1007/s10462-021-09996-w

[7]. Sass, Á., Kummer, A., & Abonyi, J. (2022). Multi-agent reinforcement learning-based exploration of optimal operation strategies of semi-batch reactors. Computers & Chemical Engineering, 162, Article 107819.

DOI: https://doi.org/10.1016/j.compchemeng.2022.107819

[8]. Li, C., Chang, Q., & Fan, H.-T. (2024). Multi-agent reinforcement learning for integrated manufacturing system-process control. Journal of Manufacturing Systems, 76, 585–598.

DOI: https://doi.org/10.1016/j.jmsy.2024.08.021

[9]. Nian, R., Liu, J., & Huang, B. (2020). A review on reinforcement learning: Introduction and applications in industrial process control. Computers & Chemical Engineering, 139, Article 106886.

DOI: https://doi.org/10.1016/j.compchemeng.2020.106886

[10]. Gao, Q., & Schweidtmann, A. M. (2024). Deep reinforcement learning for process design: Review and perspective. Current Opinion in Chemical Engineering, 44, Article 101012.

DOI: https://doi.org/10.1016/j.coche.2024.101012

[11]. Szatmári, K., Horváth, G., Németh, S., Bai, W., & Kummer, A. (2024). Resilience-based explainable reinforcement learning in chemical process safety. Computers & Chemical Engineering, 191, Article 108849.

DOI: https://doi.org/10.1016/j.compchemeng.2024.108849

[12]. Bras, E. H., Louw, T. M., & Bradshaw, S. M. (2024). Safe, visualizable reinforcement learning for process control with a warm-started actor network based on PI-control. Journal of Process Control, 144, Article 103340.

DOI: https://doi.org/10.1016/j.jprocont.2024.103340

[13]. Bloor, M., Mowbray, M., del Rio Chanona, E. A., & Tsay, C. (2026). A survey and tutorial of reinforcement learning methods in process systems engineering. Computers & Chemical Engineering, 206, Article 109515.

DOI: https://doi.org/10.1016/j.compchemeng.2025.109515

[14]. Dynamic-risk-informed safety barrier management: An application to cost-effective barrier optimization based on data from multiple sources. (2023). Journal of Loss Prevention in the Process Industries, 83, Article 105034.

DOI: https://doi.org/10.1016/j.jlp.2023.105034

[15]. Dynamic and integrated safety and security barrier management: A new framework to manage major event risks in chemical plants. (2025). Journal of Loss Prevention in the Process Industries, 96, Article 105632.

DOI: https://doi.org/10.1016/j.jlp.2025.105632

[16]. A physics-of-failure (PoF) model-based dynamic Bayesian network for considering the aging of safety barriers in the risk assessment of industrial facilities. (2024). Journal of Loss Prevention in the Process Industries, 91, Article 105402.

DOI: https://doi.org/10.1016/j.jlp.2024.105402

[17]. Tamascelli, N., Dal Pozzo, A., Scarponi, G. E., Paltrinieri, N., & Cozzani, V. (2024). Assessment of safety barrier performance in environmentally critical facilities: Bridging

conventional risk assessment techniques with data-driven modelling. Process Safety and Environmental Protection, 181, 294–311.

DOI: https://doi.org/10.1016/j.psep.2023.11.021

[18]. Bayesian network based dynamic operational risk assessment. (2016). Journal of Loss Prevention in the Process Industries, 41, 399–410.

DOI: https://doi.org/10.1016/j.jlp.2015.11.024

[19]. Developing a framework for dynamic risk assessment using Bayesian networks and reliability data. (2017). Journal of Loss Prevention in the Process Industries, 50, 142–153.

DOI: https://doi.org/10.1016/j.jlp.2017.09.011

[20]. Kim, J., Shah, A. U. A., & Kang, H. G. (2020). Dynamic risk assessment with Bayesian network and clustering analysis. Reliability Engineering & System Safety, 201, Article 106959.

DOI: https://doi.org/10.1016/j.ress.2020.106959

[21]. Lin, Z., Qin, S., He, S., Li, D., Zhu, H., & Shi, S. (2026). Dynamic risk assessment and emergency response effectiveness analysis for hazardous chemicals explosion based on dynamic Bayesian network and cloud model. Reliability Engineering & System Safety, Article 113633. Advance online publication.

DOI: https://doi.org/10.1016/j.ress.2026.113633

[22]. Li, X., Williams, J., Swanson, C., & Berg, T. (2025). A machine learning approach to predictive maintenance: Remaining useful life and motor fault analysis. Computers & Industrial Engineering, 206, Article 111222.

DOI: https://doi.org/10.1016/j.cie.2025.111222

[23]. Remaining useful lifetime prediction for predictive maintenance in manufacturing. (2023). Computers & Industrial Engineering, 184, Article 109566.

DOI: https://doi.org/10.1016/j.cie.2023.109566

[24]. Gao, Y. (2024). Machine learning models for equipment preventive maintenance and predictive analytics. Procedia Computer Science, 247, 1080–1088.

DOI: https://doi.org/10.1016/j.procs.2024.10.130

[25]. Chen, Z., Ran, Q., Chen, N., Zhang, Y., Tang, J., & Long, J. (2026). Spatio-temporal MLP model incorporating physical-informational neural networks for remaining useful life prediction. Mechanical Systems and Signal Processing, 244, Article 113815.

DOI: https://doi.org/10.1016/j.ymssp.2025.113815

[26]. A novel local enhanced channel self-attention based on transformer for industrial remaining useful life prediction. (2025). Engineering Applications of Artificial Intelligence, 141, Article 109815.

DOI: https://doi.org/10.1016/j.engappai.2024.109815

[27]. Reynoso-Donzelli, S., & Ricardez-Sandoval, L. A. (2025). An integrated reinforcement learning framework for simultaneous generation, design, and control of chemical process flowsheets. Computers & Chemical Engineering, 194, Article 108988.

DOI: https://doi.org/10.1016/j.compchemeng.2024.108988

[28]. Development of AI-based process controller of sour water treatment unit using deep reinforcement learning. (2024). Journal of the Taiwan Institute of Chemical Engineers, 157, Article 105407.

DOI: https://doi.org/10.1016/j.jtice.2024.105407

[29]. How to support decision making with reinforcement learning in hierarchical chemical process control? (2025). Chemical Engineering Journal Advances, 22, Article 100753.

DOI: https://doi.org/10.1016/j.ceja.2025.100753

[30]. Devarakonda, V. S., Sun, W., Tang, X., & Tian, Y. (2025). Recent advances in reinforcement learning for chemical process control. Processes, 13(6), Article 1791.

DOI: https://doi.org/10.3390/pr13061791

[31]. Kim, Y., & Oh, T. H. (2024). Model-based safe reinforcement learning for nonlinear systems under uncertainty with constraints tightening approach. Computers & Chemical Engineering, 183, 108601. https://doi.org/10.1016/j.compchemeng.2024.108601

[32]. Szatmári, K., Németh, S., & Kummer, A. (2024). Integration of resilience engineering and reinforcement learning in chemical process safety. Process Safety and Environmental Protection, 181, 343–353. https://doi.org/10.1016/j.psep.2023.11.027

[33]. Gerold, H., & Lucia, S. (2026). Safe reinforcement learning via adaptive robust model predictive shielding. Computers & Chemical Engineering, 206, 109521. https://doi.org/10.1016/j.compchemeng.2025.109521

[34]. Durkin, A., Stolte, J., Jones, M., Pitchumani, R., Li, B., Michler, C., & Mercangöz, M. (2026). Safe deployment of offline reinforcement learning via input convex action correction. Computers & Chemical Engineering, 206, 109535. https://doi.org/10.1016/j.compchemeng.2025.109535

[35]. González-Potes, A., Félix-Cuadras, R. A., Mena, L. J., Félix, V. G., Martínez-Peláez, R., Ostos, R., Velarde-Alvarado, P., & Ochoa-Brust, A. (2026). Safety-aware multi-agent deep reinforcement learning for adaptive fault-tolerant control in sensor-lean industrial systems:

Validation in beverage CIP. Technologies, 14(1), 44. https://doi.org/10.3390/technologies14010044

[36]. Tayal, M., Singh, A., Kolathaya, S., & Bansal, S. (2025). MAD-PINN: A decentralized physics-informed machine learning framework for safe and optimal multi-agent control https://arxiv.org/abs/2509.23960

[37]. Yuan, S., Reniers, G., & Yang, M. (2023). Dynamic-risk-informed safety barrier management: An application to cost-effective barrier optimization based on data from multiple sources. Journal of Loss Prevention in the Process Industries, 83, 105034. https://doi.org/10.1016/j.jlp.2023.105034

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/)