

Review

Predictive Operational Safety Engineering, Part I: Foundations, Taxonomy, and Future Directions for Intelligent Industrial Process Safety

Feras Alrowaie 

Chemical Engineering Department, Jubail Industrial College, Royal Commission for Jubail & Yanbu, Jubail Industrial City 35718, Saudi Arabia; rowaie_fa@rcjy.edu.sa

Abstract

Industrial process safety systems are predominantly reactive: alarms activate after limits are crossed, faults are diagnosed after deviations develop, and HAZOP knowledge remains offline during operation. This paper proposes Predictive Operational Safety Engineering (POSE) as an emerging research paradigm in which operational safety is treated as a continuously forecastable state rather than a post-event classification, shifting the operational question from *what has gone wrong?* to *how much safe operating time remains, and which intervention is most urgent?* Four integrated predictive safety metrics anchor the framework: Remaining Safety Margin (RSM), quantifying the normalized distance between the predicted process trajectory and the nearest safety boundary; Remaining Safe Operating Time (RSOT), estimating when that boundary will be crossed under the current trajectory; the Operational Vulnerability Index (OVI), combining margin depletion rate, safeguard availability, and consequence severity into a single intervention-urgency signal; and Predictive Safety Confidence (PSC), the probability that a specific named operator intervention can be executed to completion before the predicted safety boundary is crossed, coupling prediction uncertainty with action execution time. The Predictive Operational Safety Twin (POST) is proposed as a three-layer reference architecture implementing POSE through predictive process intelligence, predictive safety intelligence, and human safety intelligence. The paper synthesizes six research streams, positions POSE against seven adjacent disciplines, states ten guiding principles, and formulates a research agenda. As a conceptual narrative review, the paper does not claim empirical validation of POSE. Instead, it establishes the foundational vocabulary, reference architecture, and research agenda required to advance predictive operational safety from an emerging concept toward benchmarked and industrially validated practice. This article constitutes the conceptual and evidence-synthesis phase of a staged research program; subsequent work must test the proposed constructs through benchmark simulation, uncertainty calibration, baseline comparison, operator studies, and industrial case studies.



Academic Editor: Sheng Du

Received: 1 July 2026

Revised: 25 July 2026

Accepted: 29 July 2026

Published: 30 July 2026

Copyright: © 2026 by the author.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: predictive operational safety; predictive risk; digital twin; process safety; alarm management; HAZOP; operator decision support; human-AI teaming; dynamic risk assessment; process monitoring; industrial artificial intelligence

1. Introduction

Industrial process plants are becoming increasingly complex due to higher production capacities, tighter operating constraints, advanced automation systems, plant-wide

integration, and extensive sensor networks. Modern facilities continuously generate large volumes of operational data from distributed control systems (DCS), advanced process control (APC), safety instrumented systems (SIS), process historians, and industrial Internet of Things (IIoT) platforms [1,2]. These developments have significantly improved plant visibility and enabled the widespread adoption of data-driven analytics, process monitoring, artificial intelligence, and digital twins.

Despite these technological advances, the dominant philosophy of industrial process safety remains largely reactive. Most existing safety technologies are designed to detect abnormalities after they have already developed [3,4]. Alarm management systems notify operators when predefined thresholds are exceeded, fault detection and diagnosis algorithms identify abnormal operating conditions after measurable deviations occur, and HAZOP studies provide structured hazard knowledge during plant design, modification, or periodic review rather than during real-time operation. Digital twins increasingly predict process behavior, but many implementations remain focused on monitoring, optimization, predictive maintenance, or asset performance rather than forecasting future operational safety.

Alarm floods illustrate the practical consequence of this reactive paradigm. International standards such as ISA-18.2 and EEMUA 191 have substantially improved alarm rationalization, prioritization, and lifecycle management [5,6]. However, alarm management improves alarm quality after alarms occur; it does not provide a framework for estimating how much safe intervention time remains before safe operation is lost.

A similar limitation appears in fault detection and diagnosis. Statistical process monitoring methods such as PCA, PLS, CUSUM, EWMA, and related multivariate techniques have made major contributions to abnormal event detection and process monitoring [7–10]. More recent machine learning and deep learning methods have improved diagnostic accuracy across benchmark processes such as the Tennessee Eastman Process [11]. Yet the main outputs of fault diagnosis remain fault indicators, fault classes, diagnostic delays, and classification accuracy. These outputs are important, but they do not directly answer the broader safety question: how will the future safety state of the plant evolve if no intervention occurs?

In parallel, digital twin technology has emerged as one of the most transformative developments in Industry 4.0. A digital twin can synchronize with real-time plant measurements, estimate hidden states, predict future process trajectories, and support decision making [12–15]. In process industries, digital twins have been increasingly applied to monitoring, optimization, predictive maintenance, and safety risk management [16–18]. Nevertheless, comparatively little attention has been devoted to using digital twins as proactive safety reasoning systems capable of continuously forecasting the future operational safety state of an industrial process. Recent work has also examined how digital twin and artificial intelligence technologies may support the evolution of HAZOP from a periodic, document-centered activity toward a dynamic, lifecycle-oriented safety-management layer [19].

This observation reveals an important conceptual gap: the six disciplines reviewed have advanced largely independently, and none has adopted the future operational safety state as the primary object of continuous estimation [20]. Figure 1 illustrates this gap and the conceptual position of POSE relative to the six contributing research streams. Section 9 maps this gap in detail across all six research streams and identifies the specific scientific problems that remain unresolved.

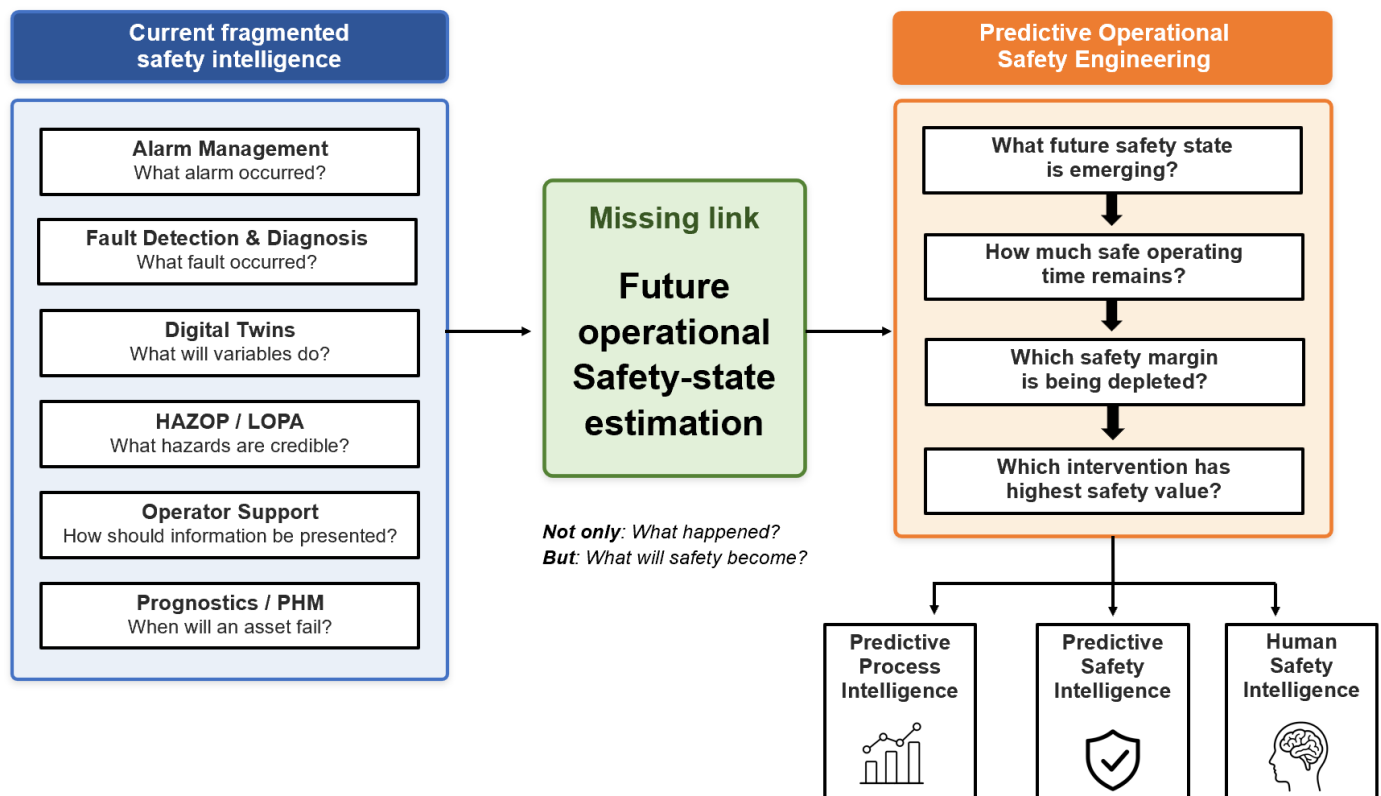


Figure 1. Conceptual motivation for Predictive Operational Safety Engineering (POSE). **(Left)** six established research streams (alarm management, fault detection and diagnosis, digital twins, HAZOP/LOPA, operator support, and prognostics/PHM) each address a partial reactive question but do not collectively produce a continuous estimate of the future operational safety state. **(Centre)** the missing link is future operational safety-state estimation, reframing the operational question from *what happened?* to *what will safety become?* **(Right)** POSE addresses four sequential predictive questions (what future safety state is emerging, how much safe operating time remains, which safety margin is being depleted, and which intervention has the highest safety value), implemented through Predictive Process Intelligence, Predictive Safety Intelligence, and Human Safety Intelligence.

This review argues that the next generation of industrial process safety should transition from reactive abnormal-event management toward predictive operational safety reasoning. Rather than merely identifying alarms or classifying faults after abnormalities become observable, future safety systems should continuously estimate the future safety condition of the operating process, forecast alarm evolution before alarm floods develop, quantify the remaining time available for safe operator intervention, and provide interpretable safety guidance that supports timely decision making.

To address this emerging need, this paper introduces Predictive Operational Safety Engineering (POSE) as a unifying research paradigm (the term paradigm is used here in the sense of a coherent research program: a shared vocabulary, a set of guiding principles, and an agenda of open problems that define a recognizable field of inquiry [21]; it does not claim a Kuhnian incommensurable break with prior work; existing safety disciplines are explicitly incorporated into POSE rather than displaced by it) for future industrial process safety. POSE is defined as the discipline concerned with continuously estimating, interpreting, and forecasting the future operational safety state of industrial processes through the integration of process prediction, alarm intelligence, safety knowledge, operational vulnerability assessment, and human-centered decision support. Unlike existing reviews that independently summarize developments in alarm management, digital twins, or fault diagnosis, this review synthesizes these research streams to establish a coherent scientific

foundation for predictive operational safety. POSE is not proposed as a replacement for existing safety systems, alarm management, FDD, digital twins, HAZOP, or operator decision support; rather, it provides an integrating predictive layer that connects these capabilities around future operational safety-state estimation.

Furthermore, a conceptual implementation referred to as the Predictive Operational Safety Twin (POST) is proposed as one possible realization of this paradigm. Rather than representing another digital twin architecture, POST illustrates how predictive process modeling, future alarm forecasting, operational vulnerability assessment, Remaining Safe Operating Time (RSOT), safety-aware alarm prioritization, and HAZOP-informed operator decision support can be integrated into a unified predictive safety framework.

The paper makes five conceptual contributions. First, it identifies a unifying gap in current process safety research: the absence of a framework that makes the future operational safety state the primary object of continuous estimation. Second, it introduces an integrated vocabulary for predictive safety reasoning, including Remaining Safety Margin (RSM), Remaining Safe Operating Time (RSOT), Operational Vulnerability Index (OVI), Predictive Safety Confidence (PSC), a probabilistic confidence measure associated with RSOT and RSM estimates rather than a competing safety index, Safety Horizon, and Operational Safety State. Third, it positions POSE relative to adjacent disciplines, including alarm management, fault detection and diagnosis, digital twins, HAZOP, dynamic risk assessment, prognostics, and model predictive control, distinguishing what each provides and what an integrated predictive safety paradigm is designed to deliver. Fourth, it proposes the Predictive Operational Safety Twin (POST) as a reference architecture illustrating how process prediction, safety-state estimation, vulnerability assessment, and human-centered safety reasoning may be integrated into a unified operational system. Fifth, it formulates a staged research agenda including benchmark systems, validation criteria, scientific challenges, and open questions required to mature POSE from a conceptual paradigm into an operational engineering discipline. This review is therefore the first, conceptual and evidence-synthesis phase of that program: it defines what must be tested, rather than claiming that the framework has already been validated.

The remainder of this review is organized as follows. Section 2 reviews the transition from reactive safety toward predictive operational safety. Sections 3–8 examine the current state of the art in alarm management, fault detection and diagnosis, digital twins, process safety knowledge, operator decision support, and prognostics. Section 9 synthesizes the major research gaps across these disciplines and positions the novelty of POSE relative to adjacent research areas. Sections 10–12 define POSE, formalize its core vocabulary, and present the Predictive Operational Safety Twin as a reference architecture. Section 13 states ten guiding principles. Section 14 presents the research agenda, validation criteria, and future challenges. Section 15 concludes the paper.

2. Evolution of Industrial Process Safety: From Reactive Protection to Predictive Operational Safety

Industrial process safety has progressed from mechanical protection and emergency shutdown toward distributed control, computerized alarms, multivariate monitoring, and, more recently, industrial AI and digital twins [1,2,5,6,13,14]. These developments have improved the detection and interpretation of abnormal conditions, but their dominant temporal logic remains reactive: alarms, diagnostics, and protective actions generally follow measurable deviations. Figure 2 contrasts this reactive logic with the anticipatory logic of predictive operational safety.

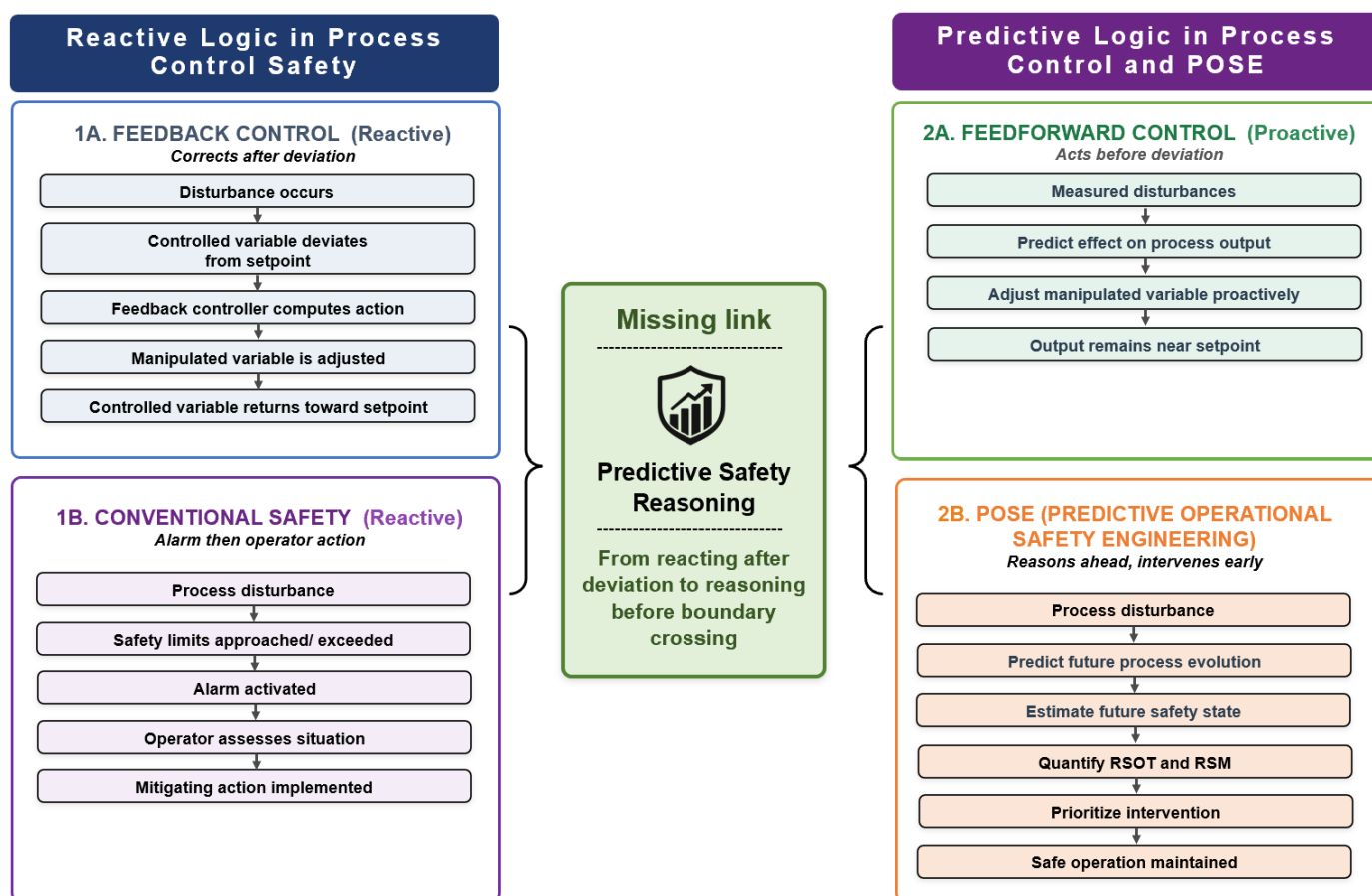


Figure 2. Conceptual analogy between classical process control and industrial process safety. Feed-back control (1A) and conventional safety (1B) respond after an observed deviation. Feedforward control (2A) and Predictive Operational Safety Engineering (2B) reason ahead of deviations by estimating RSM and RSOT before safety boundaries are reached. The analogy is conceptual rather than algorithmic: POSE is a safety-reasoning paradigm, not a control algorithm.

The resulting challenge is not a lack of individual safety technologies. Alarm systems identify deviations; FDD identifies abnormalities; HAZOP structures hazards and safeguards; and control and protection layers act once conditions demand intervention. Yet these functions are commonly deployed as separate capabilities and rarely report the future operational safety state, the safety margin being depleted, or the time available for intervention [3,4,20]. POSE is proposed as a predictive supervisory layer: it does not replace regulatory control, safety instrumented systems, alarms, or operator judgment, but organizes their information around a forecast of safety-state evolution; its supervisory role is conceptually distinct from predictive control [22,23].

3. Alarm Management and Alarm Flood Analytics

Alarm management has developed from threshold annunciation into a mature lifecycle discipline encompassing philosophy, rationalization, prioritization, shelving, performance assessment, and continuous improvement [5,6,24–26]. Research has extended this foundation through bad-actor analysis, flood detection, sequence mining, causal and propagation analysis, dynamic suppression, and AI-assisted operator support [27–35]. These methods reduce nuisance alarms and improve interpretation of developing disturbances.

Its remaining limitation is object-centered rather than methodological. Most methods begin with an active alarm, a historical alarm log, or an emerging flood. Even predictive alarm approaches commonly forecast messages or sequences rather than estimating how

the underlying process trajectory will deplete safety margins. Thus, alarm management supplies essential evidence about abnormality and operator workload, while POSE asks the additional safety-state question: what is likely to happen to operational safety before alarm overload or boundary crossing occurs?

4. Fault Detection and Diagnosis in Process Systems

FDD identifies abnormal process behavior and, where possible, its origin. Its major approaches include first-principles residual generation, multivariate statistical monitoring, and machine-learning classification [7–10,36–42]. Root-cause analysis, causal reasoning, and explainable diagnosis improve interpretability and operator use [34,43–46]. Recent reviews and physics-informed approaches further show the importance of combining data-driven detection with process knowledge when training data are limited or operating modes change [47–49].

FDD outputs a fault label, residual, probability, or diagnostic explanation. These are necessary inputs to predictive operational safety, but they do not by themselves specify how a disturbance will propagate, which safety boundary will become binding, or how long intervention remains feasible. POSE therefore treats diagnosis as an upstream intelligence layer whose evidence must be translated into a predicted, HAZOP-linked safety-state trajectory rather than as the endpoint of the analysis.

5. Digital Twins for Industrial Process Monitoring and Safety

Digital twins continuously synchronize models, measurements, and computational intelligence to estimate and predict process behavior across the asset lifecycle [12–14,50–52]. In process industries they support monitoring, optimization, fault detection, maintenance, training, and scenario analysis, enabled by IIoT infrastructure, hybrid modeling, and state estimation [16,53,54].

Safety-oriented applications include hazard visualization, emergency planning, consequence analysis, resilience, and links to risk assessment or HAZOP [17,18,55,56]. A safety-management review and a petrochemical monitoring architecture illustrate both growing industrial relevance and the need to substantiate safety benefit in specific applications [57,58]. Digital Risk Twins likewise update time-dependent risk quantities [59]. The limitation is not predictive capability but safety-state translation: a twin may forecast temperature or pressure without evaluating the forecast against formalized safety boundaries, safeguard status, and action time. POST is proposed to supply that translation through RSM, RSOT, OVI, and PSC.

6. Process Safety Knowledge: From Static HAZOP Studies to Dynamic Safety Reasoning

HAZOP and related process-safety methods provide structured engineering knowledge about deviations, causes, consequences, safeguards, and recommended actions [4,60–62]. Their strengths are interpretability, traceability, and the capture of expertise that cannot be inferred from data alone. LOPA and barrier-based approaches add a systematic view of safeguard sufficiency and consequence control [63,64].

Their usual limitation is temporal: Analyses are created during design, modification, or periodic review, rather than continuously interpreted from the evolving operating state. Digital HAZOP, knowledge graphs, and dynamic risk approaches begin to narrow this gap [65–69]. In POSE, this knowledge is not replaced; it becomes the machine-interpretable safety envelope against which predicted trajectories, barrier condition, and candidate interventions are evaluated. Figure 3 summarizes this transformation from static HAZOP knowledge to dynamic safety reasoning.

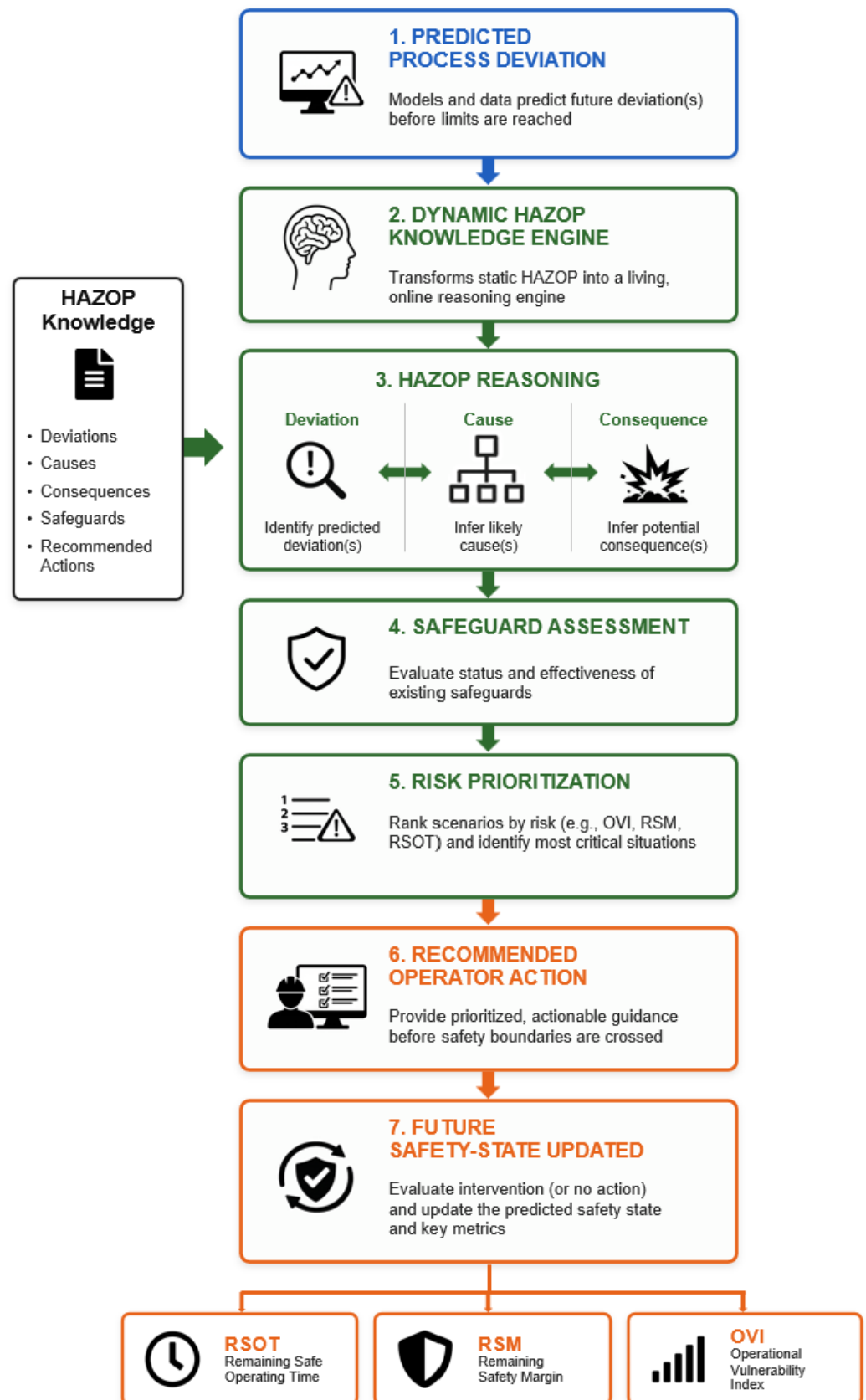


Figure 3. Conceptual transformation of static HAZOP knowledge into dynamic safety reasoning. A predictive safety workflow links process-state prediction, deviation recognition, causal and consequence reasoning, safeguard assessment, action ranking, and safety-state updating. The operator-facing outputs are RSOT, RSM, and OVI.

7. Operator Decision Support and Human Safety Intelligence

Process safety ultimately depends on human decisions during developing disturbances. Research on skills–rules–knowledge behavior, situation awareness, cognitive workload, ecological interface design, and explainable AI provides the basis for operator support [70–75]. These contributions improve alarms, displays, training, and the intelligibility of automated recommendations.

The open problem is not simply to show more information. Operators need a trustworthy explanation of what may happen next, why the trajectory is safety-relevant, what time remains, and which action should be considered first. Trust in AI-assisted operator decisions is shaped by advisory reliability, process dynamics, and error characteristics, so trust calibration and clear uncertainty communication are design requirements for POSE [76]. This motivates Human Safety Intelligence as a core, rather than optional, part of predictive safety design.

8. Prognostics, Remaining Useful Life, and the Safety Horizon Concept

Prognostics and Health Management estimates asset health and Remaining Useful Life (RUL) from model-based, data-driven, or hybrid degradation models [77–83]. Its transferable strengths include uncertainty-aware time-to-event estimation, condition monitoring, and degradation forecasting.

RSOT is deliberately not a relabeling of RUL. RUL concerns a component's degradation trajectory and its own failure threshold. RSOT concerns the first passage of a multi-variable process trajectory through a stated safety envelope, accounting for process dynamics, disturbances, safeguards, operating policy, and HAZOP-defined boundaries. A reactor approaching runaway or a vessel approaching overpressure may therefore have an urgent RSOT even when no individual asset has reached an end-of-life condition. Figure 4 contrasts the asset-focused RUL concept with the process-safety-focused RSOT concept.

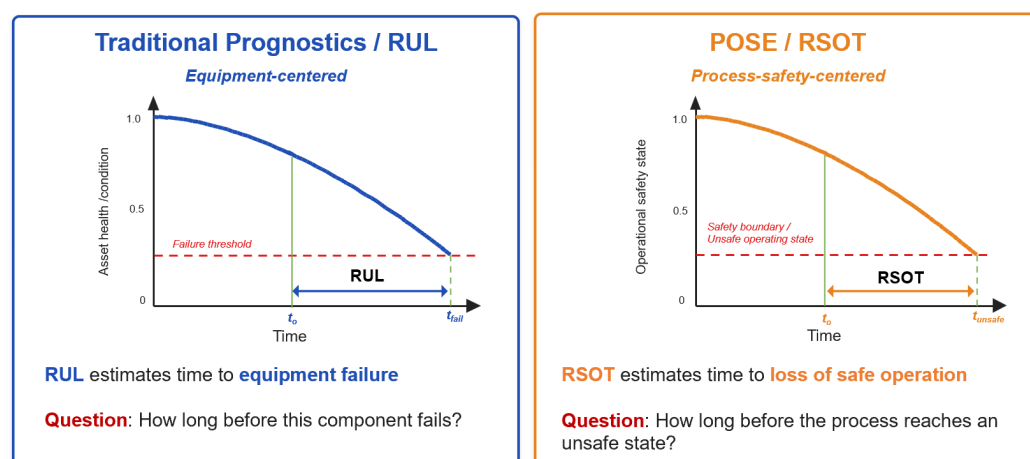


Figure 4. Conceptual contrast between traditional prognostics and POSE. Remaining Useful Life (RUL) estimates time until an individual asset crosses a degradation threshold. Remaining Safe Operating Time (RSOT) estimates time until the predicted multi-variable process state crosses a safety boundary. Safety Horizon and Predictive Safety Confidence are defined in Section 11.

The related Safety Horizon is the future interval over which a safety-state forecast remains sufficiently reliable for operational use; it shortens when uncertainty grows near critical boundaries. Together, these concepts motivate an integrated paradigm that combines trajectory prediction, safety knowledge, uncertainty treatment, and human-centered intervention support.

9. Toward Predictive Operational Safety Engineering: Synthesis of the Literature

9.1. Review Methodology and Scope

This paper follows a thematic narrative conceptual-review methodology rather than a pre-registered systematic or scoping-review protocol. It does not apply PRISMA screening, quantitative risk-of-bias appraisal, or meta-analytic pooling. Its purpose is to identify conceptual convergence, unresolved integration requirements, and research hypotheses across the literature relevant to predictive operational safety; it is not designed to estimate prevalence, effect sizes, or comparative method performance.

The final cited corpus contains 92 references, established by an exact count of the bibliography. It includes foundational contributions, peer-reviewed research articles, recognized standards, established reference texts, and recent methodological developments. Sources were retained when they made a direct contribution to at least one of the following functions: process-state prediction or monitoring; alarm intelligence; safety-boundary, barrier, or consequence reasoning; uncertainty-aware time-to-event estimation; or operator decision support. Foundational standards, benchmark descriptions, and conceptual works were retained where they were necessary to define the engineering or governance context. Sources with only tangential relevance to these functions were not used as substantive evidence in the synthesis.

For transparency, every cited source was manually assigned one primary thematic category according to its dominant contribution, while acknowledging that many sources have secondary relevance across categories. The resulting counts are descriptive of the cited corpus, not estimates of the size of the wider literature. Table 1 reports this source-level audit.

Table 1. Author-classified distribution of the 92 references cited in this thematic review. Each source is assigned one primary category by dominant contribution; categories are mutually exclusive for counting, although sources can be relevant to more than one topic.

Primary Thematic Category	Cited Sources, <i>n</i>	Typical Contribution
Alarm management	12	Alarm philosophy, rationalization, alarm floods, causal analysis, and early alarm prediction
Fault detection and diagnosis	14	Process monitoring, residual analysis, statistical methods, and fault diagnosis
Digital twins	16	Digital-twin concepts, process models, safety applications, and risk-twin architectures
HAZOP, LOPA, and process safety knowledge	13	Hazard identification, protection layers, safety knowledge, and HAZOP automation
Operator decision support and human factors	10	Situation awareness, operator cognition, trust, explanation, and human-centered design
Prognostics and PHM	7	Remaining useful life, degradation modeling, and prognostic uncertainty
Cross-cutting foundations, standards, and governance	20	Dynamic risk, control, cybersecurity, standards, benchmarks, general safety foundations, and research philosophy
Total	92	Mixed evidence base

Note: The final corpus contains 92 references after removal of one retracted source during proof correction.

The original thematic review was developed iteratively and did not retain a deduplicated historical screening ledger or a complete record of original database queries and raw

retrieval counts. To avoid retrospective fabrication, this paper does not report unrecoverable historical counts for unique records screened by title and abstract, full-text reports assessed, or exclusions. The present revision provides a source-level audit of the final cited corpus and a separate, documented revision-stage verification search.

Specifically, six revision-stage verification queries, one for each major POSE research stream, were executed manually in both Google Scholar and Scopus on 21 July 2026. In Scopus, the queries used the field-restricted TITLE-ABS-KEY() syntax and the publication-year filter 2000–2026. Older foundational works were retained in the cited corpus where necessary to establish concepts, standards, or benchmark context. Google Scholar was used as a complementary broad-discovery check; the raw-count sessions did not retain a documented date-filter setting. Table 2 reports the exact Boolean strings and raw retrieval counts. Results overlap across streams and databases, are not deduplicated, and are not historical screening counts; they are therefore not summed and were not used to derive the 92-reference final corpus.

Table 2. Revision-stage verification searches executed on 21 July 2026. Scopus searches used TITLE-ABS-KEY() and a 2000–2026 publication-year filter; Google Scholar counts are complementary broad-discovery results with no documented date filter. All counts are raw, query-specific, and overlapping rather than deduplicated screening totals.

Research Stream	Boolean Query (Inside TITLE-ABS-KEY() in Scopus)	Google Scholar	Scopus
Alarm management	("alarm management" OR "alarm flood" OR "alarm rationalization" OR "alarm prioritization" OR "alarm prediction" OR "root alarm") AND ("process industry" OR "chemical process" OR refinery OR petrochemical)	2300	199
Fault detection and diagnosis	("fault detection and diagnosis" OR "process monitoring" OR "anomaly detection") AND ("process systems" OR "chemical process" OR "process industry")	18,100	1312
Digital twins	("digital twin" OR "digital risk twin") AND ("process safety" OR "chemical industry" OR petrochemical OR refinery OR "process systems")	12,400	334
HAZOP, LOPA, and process safety knowledge	(HAZOP OR "hazard and operability" OR LOPA OR "layer of protection" OR "dynamic HAZOP") AND ("process safety" OR "safety knowledge" OR "dynamic risk" OR "safety reasoning")	12,200	752
Operator decision support and human factors	("operator decision support" OR "human factors" OR "situation awareness" OR "operator trust" OR "human-AI teaming") AND ("process safety" OR "process control" OR "process industry")	18,000	1144
Prognostics and PHM	(prognostics OR "remaining useful life" OR PHM OR "time to failure") AND ("process safety" OR "process systems" OR "chemical process" OR "industrial process")	16,100	207

A future systematic review should use a registered protocol, reproducible database queries, a defined publication window, explicit eligibility criteria, and dual-reviewer screening to test and extend the conceptual conclusions advanced here.

Synthesis is organized by thematic convergence across domains rather than by pooled quantitative estimates. Each domain is reviewed for its main achievements and its remaining gap relative to the unified predictive safety objective of POSE. The conclusions are interpretive and conceptual: they identify patterns in the cited corpus and use them to motivate a proposed research framework, not to establish empirical effect sizes. Readers

should interpret the findings as evidence-informed proposals and research hypotheses. The operational effectiveness of POSE and POST on real industrial data remains to be established through future empirical work.

9.2. Fragmentation of Current Research

The preceding sections reviewed mature research domains that collectively define the current state of industrial process safety. Alarm management has evolved into a comprehensive discipline for alarm lifecycle management and operator support. FDD has achieved substantial success in identifying abnormal operating conditions through physics-based, statistical, and data-driven approaches. Digital twins have transformed process monitoring by enabling continuous synchronization between physical assets and virtual models. HAZOP and LOPA provide structured repositories of safety knowledge. Operator decision support has improved human interaction with increasingly complex industrial systems. Prognostics has established rigorous methods for estimating remaining life, providing the conceptual foundations for RSOT. Figure 5 summarizes how these six streams converge toward Predictive Operational Safety Engineering.

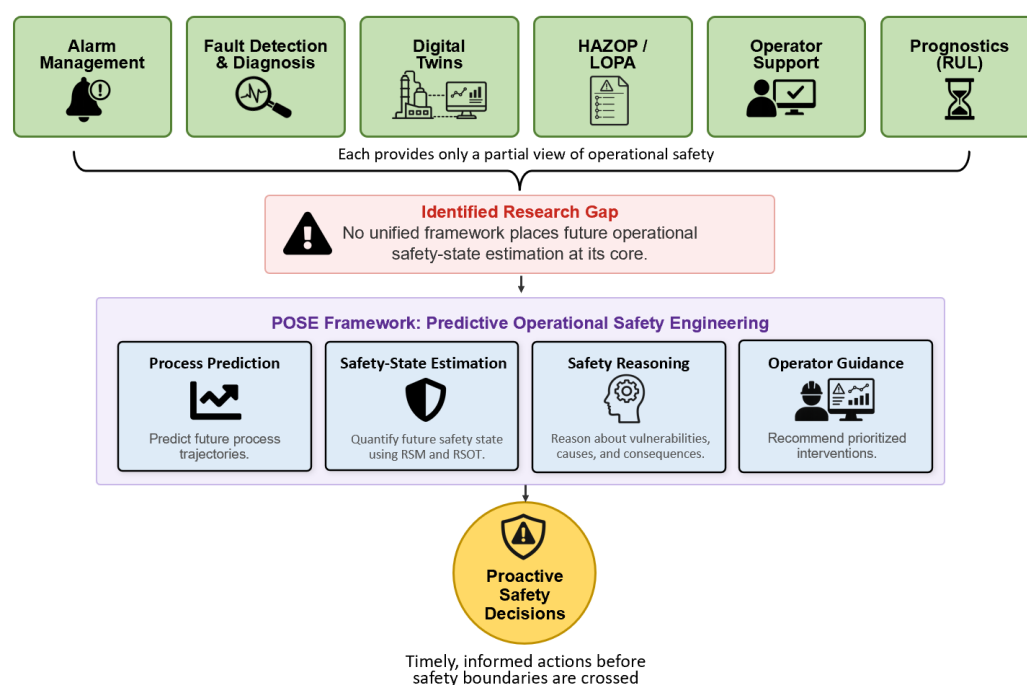


Figure 5. Synthesis of the six reviewed research streams and the critical research gap they collectively expose. Each stream (Alarm Management, Fault Detection and Diagnosis, Digital Twins, HAZOP/LOPA, Operator Support, Prognostics/RUL) provides a partial view of the operational safety problem. No unified framework presently makes the future operational safety state the central object of continuous estimation. POSE addresses this gap by integrating process prediction, safety-state estimation, safety reasoning, and operator guidance around the POSE metrics: RSOT, RSM, OVI, PSC, and Safety Horizon.

Individually, each discipline has achieved substantial scientific and industrial success. Collectively, however, they reveal a common observation: each discipline addresses only part of the operational safety problem.

The evidence reviewed identifies a recurring integration gap: no generally adopted framework in this purposive corpus combines these technologies into a coherent predictive safety architecture centered on the future operational safety state. This does not imply that adjacent fields lack predictive or safety-relevant capabilities; rather, their integration, common vocabulary, and operator-facing safety-state outputs remain uneven. The frag-

mentation is not a failure of individual disciplines but a structural consequence of their independent development. Table 3 summarizes the gap across each domain.

Table 3. Research streams and remaining gaps motivating Predictive Operational Safety Engineering.

Research Stream	Primary Object	Typical Output	Remaining Gap for POSE
Alarm management [5,6]	Alarm events	Alarm priority, sequences, floods	Does not estimate future safety-state trajectory
Fault diagnosis [36,40]	Fault class and occurrence	Fault label, residual, probability	Does not translate diagnosis into RSOT or RSM
Digital twins [13,14]	Process state	Predicted variables, virtual representation	Does not always translate prediction into safety state
HAZOP/LOPA [4,61]	Hazard scenarios/barriers	Deviations, causes, safeguards	Usually offline or periodically updated
Operator support [70,73]	Human decision quality	Displays, explanations, guidance	Lacks predictive safety-time reasoning
Prognostics & PHM [77]	Asset health	RUL, failure probability	Equipment-centered, not process-safety-centered

9.3. The Missing Scientific Link

The literature consistently demonstrates that industrial systems are becoming increasingly capable of answering questions regarding the current operating condition. Alarm management asks which alarms are active. Fault diagnosis asks which fault has occurred. Digital twins ask what the current or future process state may be. HAZOP asks what hazards exist under specific deviations. Operator decision support asks how information should be presented.

Although these questions are individually important, they do not directly answer the question that ultimately determines operational safety: *what will be the future operational safety state of the plant if the current trajectory continues?*

This question remains largely absent from the existing literature. It motivates a shift from reactive abnormal-event management toward predictive operational safety reasoning.

9.4. Positioning and Novelty of POSE

POSE does not replace alarm management, FDD, digital twins, HAZOP, dynamic risk assessment, operator support, or prognostics. Its novelty lies in redefining the primary object of operational safety intelligence as the future operational safety state and organizing existing capabilities around that object. Table 4 makes this concrete: for each established paradigm it shows what that discipline provides, where it stops, and what POSE adds.

The defining distinction is that POSE treats operational safety as a dynamic state variable that can be estimated and forecast. Existing methods may predict alarms, classify faults, simulate process variables, or update risk. POSE requires these outputs to be composed into a safety-state forecast: a structured statement about the future acceptability of operation, the margins being depleted, the time available for action, and the safety rationale for intervention.

Table 4. What POSE adds beyond existing safety paradigms. Each row identifies what an established discipline provides, where it stops, and the specific layer POSE adds.

Existing Paradigm	Primary Purpose	Typical Outputs	What It Cannot Provide	What POSE Adds
Alarm management [5,6]	Detect and prioritize alarm events	Alarm state, priority rank, flood pattern	Safety-state forecast before alarm activation	RSM and RSOT as pre-alarm leading indicators; OVI ranks urgency before alarms fire
Fault detection & diagnosis [36,40]	Identify fault occurrence and type	Fault label, fault probability, root-cause rank	How safety margins evolve if no action is taken	Converts fault evidence into a safety-state trajectory; PSC estimates whether an action can be completed in time
Digital twins [13,14,51]	Replicate and predict process state	Predicted T, P, flow, composition	HAZOP-linked safety-boundary assessment; operator intervention guidance	POST: reorients twin output from process fidelity to RSM, RSOT, OVI evaluated against the HAZOP safety envelope
HAZOP/LOPA [4,61,64]	Identify hazard scenarios and assess safeguard sufficiency	Deviation lists, safeguard evaluations, SIL requirements	Continuous online trajectory interpretation during operation	Activates HAZOP knowledge dynamically; barrier status propagated into OVI in real time
Dynamic risk assessment [20,65,84]	Update event probability from evidence	Updated event probability	Time-to-safety-boundary; ranked intervention guidance	RSOT: translates probability language into intervention-window language; not <i>how likely?</i> but <i>how long?</i> and <i>which action?</i>
Prognostics & PHM [77,78,80]	Estimate remaining useful life of assets	RUL distribution; degradation trajectory	Multi-variable, HAZOP-linked process safety horizon	RSOT: generalizes RUL from single-asset thresholds to multi-variable, HAZOP-defined safety boundaries
Operator decision support [70,73]	Improve human response to process information	Alarm displays, XAI explanations	Proactive safety-state forecasts; intervention-time ranking	PSC and OVI: time-stamped, safety-ranked intervention guidance before alarms activate (Layer III of POST)

This leads to three novelty claims that cannot be reduced to extensions of existing individual disciplines.

1. Paradigm novelty: POSE proposes the formalization of operational safety as a trajectory estimation problem, a framing not adopted as the primary objective by any of

the six disciplines reviewed. Existing disciplines produce alarms, fault labels, risk probabilities, or optimized control actions. POSE is designed to produce a forecast of the future safety state: a structured answer to how margins are depleting, when boundaries will be reached, and which action is most urgent, as the primary engineering output.

2. Metric novelty: RSM, RSOT, and OVI constitute an integrated class of safety signal qualitatively distinct from existing outputs. RSM is multi-variable and trajectory-based rather than single-sensor and threshold-based. RSOT is a time-to-boundary estimate derived from a predicted process trajectory and HAZOP-informed safety boundaries, not a linear extrapolation to an alarm limit. OVI aggregates margin, rate, safeguard availability, and consequence severity into a single decision-support metric. Together they are proposed as the primary operands of predictive safety reasoning, complementing rather than replacing existing alarm and diagnostic outputs during the transition to POSE-based practice.
3. Architecture novelty: POST is proposed as a reference architecture whose distinguishing design objective is to estimate the future operational safety state rather than the future process state. Unlike conventional digital twins (which forecast process variables), safety-constrained controllers (which enforce limits rather than inform humans), and dynamic risk assessment systems (which update risk probability scalars), POST is designed to produce a time-resolved, HAZOP-linked, uncertainty-aware safety-state trajectory communicated to an operator before any alarm is activated.

9.4.1. POSE as an Engineering Obligation, not a Terminology Restatement

A legitimate concern is whether RSM, RSOT, OVI, and PSC are genuinely new constructs or merely existing concepts (safety margins, time-to-alarm estimates, risk indices, confidence intervals) restated under new labels. The answer lies in the implementation obligations each metric imposes, which are qualitatively different from those of their apparent analogues.

RSM is required to be continuously estimated from a predicted multi-variable process trajectory, normalized against HAZOP-informed safety boundaries, and updated in real time as the process evolves. These requirements demand a digital twin or predictive process model, a formalized safety boundary representation, and an online computation architecture. An existing safety margin defined at design time, such as a SIL-derived safe operating limit, an MPC constraint bound, or a process capability index, carries none of these obligations. It is a fixed scalar assigned offline; RSM is a dynamic state variable that must be estimated continuously.

RSOT is not a renamed time-to-alarm. A trend-based time-to-alarm extrapolates a linear trajectory in a single measured variable to a preset threshold, requiring no process model, no safety knowledge, and no uncertainty quantification. RSOT is the first-passage time of a predicted multi-variable process trajectory through a HAZOP-informed safety envelope, represented as a conditional probability distribution with calibrated predictive bounds. Its implementation requires a digital twin, an uncertainty propagation method, and a structured safety boundary definition, which is a substantively different engineering task from setting an alarm setpoint.

OVI is not a renamed Risk Priority Number. RPN is assigned offline to failure modes and remains constant between design reviews. OVI is continuously updated as process margins evolve, safeguard availability changes, and prediction uncertainty grows or shrinks during an unfolding situation. Implementing OVI requires real-time integration of outputs from the process model, the safeguard monitoring system, and the uncertainty propagation layer, infrastructure that a static risk index does not require.

PSC has no direct predecessor in the process safety literature. Confidence intervals in prognostics quantify uncertainty about component remaining useful life. PSC quantifies the probability that a specific, named operator intervention can be executed to completion before the predicted safety boundary crossing, coupling prediction uncertainty with an action execution model, a required completion time, and a HAZOP-derived boundary definition. The question of whether sufficient time remains to complete a specific named action is not answered by any existing process safety metric.

These four metrics do not rename existing outputs. They collectively define a set of engineering obligations, specifically process prediction, safety knowledge formalization, uncertainty quantification, safeguard state monitoring, and action-time modeling, that no existing discipline presently fulfills as a unified design objective. POSE is the discipline that names, integrates, and takes responsibility for those obligations.

9.4.2. What POSE is Not

POSE is not dynamic risk assessment because it does not treat an updated event-probability or risk scalar as the complete operational output; it additionally requires margin depletion, first-passage time, and action feasibility to be made explicit. It is not safety-constrained or economic model predictive control because those methods compute or optimize control actions subject to constraints, whereas POSE provides a supervisory, HAZOP-linked safety-state interpretation that may inform operators and controllers without replacing their control objective. It is not a relabeling of remaining useful life or probabilistic risk assessment because RSOT concerns a multi-variable process boundary under a stated operating policy, while PSC concerns completion of a specific intervention. Finally, it is not advanced alarm management because it is intended to forecast safety-state degradation and intervention feasibility before alarm activation or alarm overload, while remaining complementary to alarm rationalization and prioritization.

Table 5 places six commonly cited analogues side by side to make these distinctions concrete.

Table 5. Why six commonly cited analogues are not POSE: each concept addresses a different primary question and imposes qualitatively different obligations from RSM, RSOT, OVI, and PSC.

Concept	Primary Question Answered	Typical Output	Why It Is Not POSE
Time-to-alarm	When will a single tag hit a preset alarm limit?	Extrapolated scalar time	Single-variable; no HAZOP linkage; no multi-variable trajectory; no uncertainty quantification
Remaining useful life (RUL)	When will an asset degrade to failure?	Asset-level life estimate	Equipment-centered; does not address safety-boundary proximity, safeguard status, or operator intervention windows
Risk Priority Number (RPN)	How severe, likely, and detectable is a failure mode?	Static ordinal score	Assigned offline at design time; not updated from the predicted process trajectory; no time-to-boundary output

Table 5. Cont.

Concept	Primary Question Answered	Typical Output	Why It Is Not POSE
Dynamic risk index	How has event likelihood changed given current evidence?	Updated probability or risk scalar	Probability-centered; does not produce margin depletion rates, RSOT estimates, or ranked operator-intervention guidance
Alarm priority	Which active alarm deserves the most immediate attention?	Priority class or rank	Reactive to alarm activation; does not forecast margin trajectories before alarms fire; no RSOT-equivalent output
MPC constraint margins	How close are process variables to control constraint bounds?	Constraint distance under active controller	Serves automatic control-enforcement; not designed for operator safety reasoning or HAZOP-linked intervention-time estimation

The ambition is not to claim that prediction, safety knowledge, or decision support are new individually. It is to establish the engineering discipline that integrates them around a single objective: making the future operational safety state visible, quantified, interpretable, and actionable before hazardous conditions develop.

Having established that the unresolved gap lies at the integration boundary of these disciplines rather than within any single one, the following sections define the POSE paradigm, formalize its core metrics, and present POST as a reference architecture for practical implementation.

9.4.3. Distinction from Dynamic Risk Assessment

Dynamic risk assessment (DRA) uses Bayesian networks, bow-tie models, or fault trees continuously updated from real-time operational evidence, including sensor readings, safeguard status, and near-miss reports, to estimate the conditional event probability $P(\text{event} \mid \text{evidence})$ [20,65–68,84]. Its primary output is a dimensionless probability scalar or a risk rank updated as evidence accumulates. POSE produces qualitatively different information: a safety-state trajectory represented as a vector $\{RSM(t+\tau), RSOT(t), OVI(t)\}$ evaluated from the predicted process path against HAZOP-defined safety boundaries. The distinction is both mathematical and operational. An operator confronting a DRA output of $P(\text{event}) = 0.08$ and a POSE output of $RSOT = 6$ min receives categorically different engineering guidance. The first is a statistical likelihood appropriate for regulatory reporting and design review; the second is an assessment of the intervention window available before corrective action is no longer feasible. DRA asks *how likely?* POSE asks *how long?* And *which action first?* These approaches are complementary: updated safeguard failure probabilities from DRA can inform the OVI computation within POSE. However, DRA alone does not produce margin depletion rates, RSOT estimates, or HAZOP-linked ranked intervention guidance; those constitute the layer POSE adds.

9.4.4. Distinction from Safety-Constrained Model Predictive Control

Advanced process control systems using model predictive control optimize operational objectives subject to safety limits treated as hard constraint bounds that must never be

violated. MPC acts on the process: it computes and implements control moves to stay within bounds. POSE diagnoses the safety state and informs humans: it estimates how margins are evolving and advises operators on interventions before limits are approached. MPC is most effective within its normal operating envelope under routine disturbances; POSE addresses the abnormal event regime where MPC may be inactive, overridden, or operating near the boundary of its validity. The two systems are complementary: MPC provides constraint-aware optimal control, while POSE provides the safety-state intelligence that operators need when the process evolves outside normal patterns.

9.4.5. Distinction from Digital Twins and Process Simulation

Conventional industrial digital twins are designed to achieve process fidelity: they replicate plant behavior by estimating and predicting state variables (temperature, pressure, composition, flow) and equipment conditions, and are validated by how accurately the simulated trajectory matches measured plant behavior [13,14,50,51]. POST adopts a fundamentally different design objective. It uses the predicted process trajectory as input to a safety-state estimator whose terminal output is not process variables but the future safety-state vector (RSM, RSOT, OVI, and PSC), derived by continuously evaluating the predicted path against the HAZOP-defined safety envelope. A digital twin that perfectly predicts reactor temperature provides no direct answer to whether that temperature trajectory is approaching a safety boundary, how rapidly the margin is depleting, or how much intervention time remains. POST is designed to answer precisely those questions. The two architectures share predictive modeling infrastructure but differ in their design objective and output: process fidelity (digital twin) versus safety-state trajectory (POST).

9.4.6. Distinction from Early Warning and Incipient Fault Detection

Early warning systems and incipient fault detection methods, including multivariate statistical process monitoring, abnormal situation management programs, and predictive alarm forecasting, aim to identify developing abnormalities before alarm limits are crossed [6,25,35]. These contributions are valuable and are incorporated within the POSE framework. However, early warning methods address the question *is something going wrong before the alarm fires?* POSE addresses the qualitatively different questions that follow: *how much safety margin remains across all binding boundaries?* (RSM) *How long before the nearest safety boundary is crossed under the current trajectory?* (RSOT) *Which intervention is most urgent given consequence severity, safeguard status, and prediction uncertainty?* (OVI) *And is there sufficient confidence that the chosen intervention can be completed before the boundary is reached?* (PSC). Early warning detection provides the signal that something is developing; POSE provides the time-resolved, HAZOP-informed safety reasoning and ranked operator guidance that must follow to make that signal actionable. In this sense, POSE begins where early warning ends. Table 6 summarizes this distinction.

9.5. An Engineering Analogy: From Feedback Control to Predictive Operational Safety

Note: The following analogy is intended to provide engineering intuition and should not be interpreted as a strict equivalence between control algorithms and safety management systems.

An intuitive way to understand the philosophy of Predictive Operational Safety Engineering is through its analogy with classical process control. Although process control and process safety address different engineering objectives, their underlying operational philosophies exhibit notable similarities.

Table 6. Comparison of early warning/incipient fault detection and POSE.

Aspect	Early Warning/Incipient Fault Detection	POSE
Primary question	Is an abnormality developing before the alarm fires?	How will safety margins evolve, and how much intervention time remains?
Main output to operator	Warning signal, fault precursor, or predictive alarm	RSM, RSOT, OVI, PSC, and ranked intervention guidance
Time logic	Before alarm activation or fault classification	Before safety boundary is approached
Safety knowledge	Detection thresholds or model-based signatures	HAZOP-informed safety boundaries, safeguard states, consequence severity
Operator value	Alerts operator to a developing situation	Quantifies urgency, ranks interventions, and estimates action feasibility

Conventional industrial safety systems operate primarily in a reactive manner. An alarm is generated only after a process variable exceeds a predefined limit, a fault diagnosis algorithm identifies an abnormal condition after measurable deviations become evident, and operator intervention follows the occurrence of these events. This sequence closely resembles the operation of a classical feedback controller, where corrective action is initiated only after the controlled variable deviates from its desired value.

In contrast, feedforward control anticipates the effect of measurable disturbances before they significantly influence the controlled variable. Rather than waiting for an error to develop, the feedforward controller predicts the future process response and applies corrective action proactively.

Predictive Operational Safety Engineering adopts an analogous philosophy at the safety-management level. Instead of waiting for alarms, fault classifications, or safety limit violations, POSE continuously predicts the future operational safety state by integrating process dynamics, alarm evolution, operational vulnerability, and engineering safety knowledge. The objective is to provide sufficient time for informed operator intervention before hazardous operating conditions fully develop. Figure 2 (panels 2A and 2B) illustrates this parallel.

It is important to emphasize that this analogy is conceptual rather than algorithmic. POSE is not a feedforward controller, nor is it intended to replace existing control or protection systems. Instead, it represents a higher-level predictive safety decision framework that complements conventional feedback-based regulatory control and safety instrumented systems. While regulatory control maintains process variables within operating limits, and SIS protects against defined failure scenarios, POSE monitors the predicted evolution of the operational safety state and supports earlier, safety-oriented operator decision making across the full range of developing abnormal situations.

This analogy highlights the fundamental paradigm shift proposed in this paper: from reacting to observed safety deviations toward anticipating future safety degradation and enabling proactive intervention. The following sections formalize this paradigm by proposing a taxonomy, an integrated set of operational safety metrics, and a reference architecture for practical implementation.

10. Definition and Taxonomy of Predictive Operational Safety Engineering

Predictive Operational Safety Engineering (POSE) is proposed as a research framework that treats operational safety as a continuously forecastable dynamic state and makes future safety-state estimation the primary objective of industrial process monitoring. In this context, the operational safety state refers to the evolving condition of the process relative

to safety boundaries, safeguard availability, credible consequences, operator response time, and prediction uncertainty. The central claim of POSE is that the future operational safety state, like temperature, pressure, or composition, can be estimated and forecast with useful confidence over a finite safety horizon. A safety system that waits for a limit to be crossed before acting has already forfeited part of the intervention time that prediction could have provided.

POSE is premised on the availability of digital twins, process models, safety knowledge, and human-centered reasoning sufficient to estimate RSM, RSOT, and OVI continuously, and to convert those estimates into timely operator guidance before safety boundaries are approached.

Table 7 contrasts current practice with the POSE paradigm across six operational dimensions.

Table 7. Contrast between current industrial safety practice and the POSE paradigm.

Dimension	Current Practice	POSE
What is monitored?	Process variables and alarm threshold crossings	Operational safety state as a dynamic, predicted trajectory
When does the system respond?	After alarm limits are exceeded	Before safety boundaries are approached
Primary output to operator	Alarm activation, fault label, risk probability	RSM, RSOT, OVI, and ranked intervention guidance
Time orientation	Present and near past	Predicted future under the current operating trajectory
Safety knowledge application	Offline in HAZOP documents	Connected online to trajectory predictions and deviation reasoning
Safety boundary definition	Single-variable alarm setpoints	Multi-variable, HAZOP-informed safety envelope
Treatment of uncertainty	Often implicit or separate from alarms and diagnosis	Explicitly linked to safety horizon and predictive safety confidence

The paradigm shift described in Table 7 is realized through three foundational pillars, each addressing a distinct layer of the predictive safety objective.

10.1. Predictive Process Intelligence

This pillar includes digital twins, hybrid models, process simulators, first-principles models, state estimation, and AI methods that forecast future process behavior. Its input is the evolving plant state, including process measurements, controller outputs, equipment condition, disturbance information, and operating context. Its output is not yet a safety judgment; it is a set of plausible future process trajectories with uncertainty [13,51].

10.2. Predictive Safety Intelligence

This pillar converts process forecasts into safety-state indicators. Key concepts include Operational Safety State, Operational Vulnerability, Remaining Safety Margin, Remaining Safe Operating Time, Safety Horizon, Alarm Flood Horizon, and safety-weighted alarm prioritization. Its role is to map future process trajectories onto safety boundaries, alarm limits, safeguard activation thresholds, and consequence regions.

10.3. Human Safety Intelligence

This pillar translates predictive safety information into interpretable, actionable decision support using HAZOP, process topology, safeguard knowledge, explainable AI,

and operator-centered interfaces. Its output should be suitable for operational action: what is likely to happen, why it matters, how urgent it is, which safeguard or barrier is relevant, and which intervention should be considered first [44,73]. Figure 6 shows how the three pillars interact within the POSE framework.

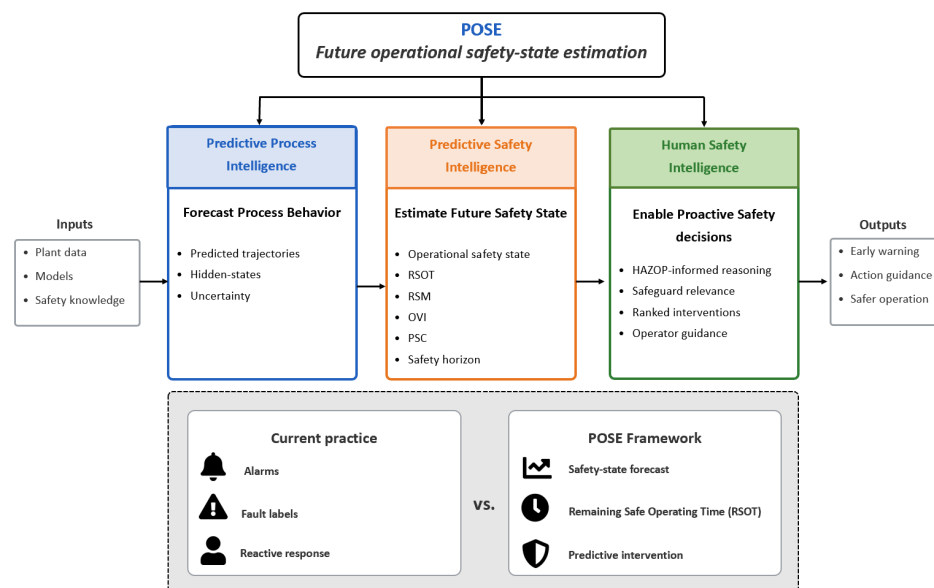


Figure 6. The three-pillar architecture of Predictive Operational Safety Engineering (POSE). Plant data, process models, and safety knowledge feed into Predictive Process Intelligence, which generates predicted trajectories, hidden-state estimates, and uncertainty bounds. Predictive Safety Intelligence translates these forecasts into safety-state indicators (RSM, RSOT, OVI, PSC, Safety Horizon). Human Safety Intelligence converts these indicators into HAZOP-informed, safeguard-aware, ranked operator guidance. The bottom panel contrasts current practice (alarms, fault labels, reactive response) with the POSE paradigm (safety-state forecast, time-to-unsafe-state, predictive intervention).

Table 8 provides a functional breakdown of each pillar, mapping representative inputs, core functions, and outputs.

Table 8. Functional taxonomy of POSE.

POSE Pillar	Representative Inputs	Core Functions	Representative Outputs
Predictive Process Intelligence	Process historian data, current measurements, control actions, equipment state, disturbances	State estimation, trajectory forecasting, uncertainty quantification	Predicted trajectories, estimated hidden states, uncertainty bounds
Predictive Safety Intelligence	Process trajectories, alarm limits, safe operating envelopes, safeguard thresholds, topology	Safety-state classification, margin estimation, time-to-boundary estimation, vulnerability scoring	RSM, RSOT, OVI, PSC, Safety Horizon, boundary-crossing forecasts
Human Safety Intelligence	Safety-state forecasts, HAZOP knowledge, procedures, safeguard status, operator context	Explanation, action ranking, consequence interpretation, interface design	Ranked interventions, HAZOP-linked explanations, safeguard status, operator guidance

The operational metrics that underpin Predictive Safety Intelligence are formalized in the following section.

11. Core Scientific Vocabulary

The equations in this section are intended as generic formalizations rather than universal fixed implementations. Practical applications may define boundary functions, normal-

ization constants, confidence thresholds, and vulnerability weights according to process-specific hazards, safeguards, operating modes, and safe operating envelopes.

11.1. Operational Safety State

The Operational Safety State is the dynamic condition of an operating process describing its ability to remain within acceptable safety boundaries under current and predicted future conditions. Let $x(t)$ denote the estimated process state, $u(t)$ the control or operator action, $d(t)$ disturbances, and $\mathcal{B}$ the set of safety boundaries. The operational safety state can be represented conceptually as

$$S(t) = \Phi(x(t), u(t), d(t), \mathcal{B}, \mathcal{K}_s), \quad (1)$$

where $\mathcal{K}_s$ represents safety knowledge such as HAZOP deviations, consequences, safeguards, and operating procedures. The function $\Phi(\cdot)$ may be implemented using rules, hybrid models, probabilistic inference, machine learning, or combinations of these methods. In this paper, operational safety state is used as an umbrella term: it may be represented either quantitatively, by margins and times-to-boundary (RSM, RSOT), or categorically, by regimes such as normal, vulnerable, critical, and unsafe, as formalized in the following subsection.

11.2. Safety State Space

The Safety State Space represents possible safety regimes of the process, such as normal, vulnerable, critical, and unsafe. A minimal state partition may be written as

$$\mathcal{S} = \{S_{\text{normal}}, S_{\text{vulnerable}}, S_{\text{critical}}, S_{\text{unsafe}}\}. \quad (2)$$

The boundaries between these regimes are not only alarm limits. They may include safe operating limits, equipment design limits, safeguard activation thresholds, environmental release thresholds, quality-related safety constraints, and human response-time constraints.

11.3. Remaining Safety Margin

Remaining Safety Margin is a generalized measure of the distance between the current or predicted process trajectory and the nearest defined safety boundary. Let $\mathcal{J}$ denote the index set of safety-relevant boundaries in $\mathcal{B}$. For a predicted trajectory $\hat{x}(t + \tau | t)$, a boundary function $g_j(\cdot)$, and prediction horizon τ , the margin to boundary $j \in \mathcal{J}$ can be represented as

$$M_j(t + \tau) = g_j(\hat{x}(t + \tau | t)), \quad (3)$$

where $M_j > 0$ denotes remaining safe margin and $M_j \leq 0$ denotes predicted boundary violation. The overall Remaining Safety Margin can be defined as the minimum normalized margin over relevant boundaries:

$$\text{RSM}(t + \tau) = \min_{j \in \mathcal{J}} \frac{M_j(t + \tau)}{M_j^{\text{ref}}}, \quad (4)$$

where j ranges over $\mathcal{J}$, and M_j^{ref} is the reference margin for boundary j , measured at the selected nominal reference condition. This normalization expresses remaining margin as a fraction of the margin available at that reference condition, so that $\text{RSM} = 1$ corresponds to the selected nominal reference condition and $\text{RSM} \leq 0$ indicates a predicted boundary violation. Because reference margins, operating modes, and boundary functions are plant-specific, RSM values are comparable only within a declared normalization convention; they

should not be treated as universal scores across equipment, plants, or operating modes without a documented mapping.

11.4. Remaining Safe Operating Time

Remaining Safe Operating Time is the time-based projection of Remaining Safety Margin. Let $X_t(\tau)$ denote the stochastic future process trajectory at lead time τ , conditional on the information $\mathcal{F}_t$ available at time t . Its random safety-margin trajectory is

$$\text{RSM}_t(\tau) = \min_{j \in \mathcal{J}} \frac{g_j(X_t(\tau))}{M_j^{\text{ref}}}. \quad (5)$$

The first-passage time to the governing safety boundary set, equivalently the first time at which this minimum normalized margin reaches zero, is

$$T_{\mathcal{B}}(t) = \inf\{\tau \geq 0 : \text{RSM}_t(\tau) \leq 0\}, \quad (6)$$

with $\inf \emptyset = \infty$. In a finite-horizon implementation, failure to observe a crossing means that $T_{\mathcal{B}}(t)$ is right-censored beyond the prediction horizon, not that the process is safe indefinitely. For a deterministic nominal trajectory $\hat{x}(t + \tau | t)$, the point estimate is

$$\widehat{\text{RSOT}}(t) = \inf\{\tau \geq 0 : \widehat{\text{RSM}}(t + \tau | t) \leq 0\}. \quad (7)$$

Uncertainty is represented by the conditional distribution of the first-passage time,

$$F_T(q | \mathcal{F}_t) = \Pr(T_{\mathcal{B}}(t) \leq q | \mathcal{F}_t), \quad Q_p(t) = \inf\{q \geq 0 : F_T(q | \mathcal{F}_t) \geq p\}, \quad (8)$$

where $Q_p(t)$ is the p -quantile of predicted boundary-crossing time for $p \in (0, 1)$. The distribution should distinguish, where feasible, measurement and state-estimation uncertainty, uncertain disturbances and operating conditions, parameter uncertainty, and structural model-form uncertainty. Monte Carlo or ensemble propagation represents the first three only to the extent that they are included in the sampling model; structural uncertainty additionally requires model ensembles, discrepancy terms, stress-test scenarios, or conservative margins. A conservative one-sided lower bound with predictive coverage $1 - \alpha$, $\alpha \in (0, 1)$, is therefore

$$\text{RSOT}_{1-\alpha}^{\text{L}}(t) = Q_{\alpha}(t), \quad \Pr(T_{\mathcal{B}}(t) \geq \text{RSOT}_{1-\alpha}^{\text{L}}(t) | \mathcal{F}_t) \geq 1 - \alpha, \quad (9)$$

subject to the usual quantile convention and equality for a continuous predictive distribution. Thus a 90 % lower predictive bound uses the 10th percentile $Q_{0.10}$, whereas the 90th percentile $Q_{0.90}$ answers the different question of when 90 % of predicted crossings have occurred. Reporting the median, a central predictive interval, and the conservative lower bound avoids the ambiguity of attaching a single “confidence level” to RSOT.

RSOT is conceptually distinct from the time-to-alarm concept sometimes computed in DCS trend analysis. Time-to-alarm typically extrapolates a linear trend in a single measured variable to a predefined alarm limit, and does so independently of prediction uncertainty, safeguard availability, and safety consequence severity. RSOT, as defined here, differs in four respects. First, it is multi-variable: it is derived from the minimum normalized margin across all boundaries indexed by $\mathcal{J}$, not from a single sensor trend. Second, it is trajectory-based: it depends on predicted future process evolution produced by a digital twin or process model, not on a linear extrapolation. Third, it is uncertainty-aware: it is represented by a conditional first-passage-time distribution and associated predictive bounds. Fourth, it is HAZOP-informed: the boundary set $\mathcal{B}$ reflects engineering

safety knowledge including safeguard activation thresholds, equipment design limits, consequence severity, and HAZOP deviation mappings, not only the alarm setpoint of the measured variable. These properties constitute a substantive methodological advance beyond univariate DCS trend extrapolation.

The central practical value of RSOT is the intervention window it reveals: the time between when a boundary crossing becomes predictable and when it actually occurs. Its value is conditional, not autonomous: a reported RSOT applies only under the explicitly stated prediction model, disturbance and control assumptions, boundary definition, and intervention policy. Figure 7 illustrates this window, showing how RSM, RSOT, and OVI relate to the process safety state trajectory and the safety boundary.

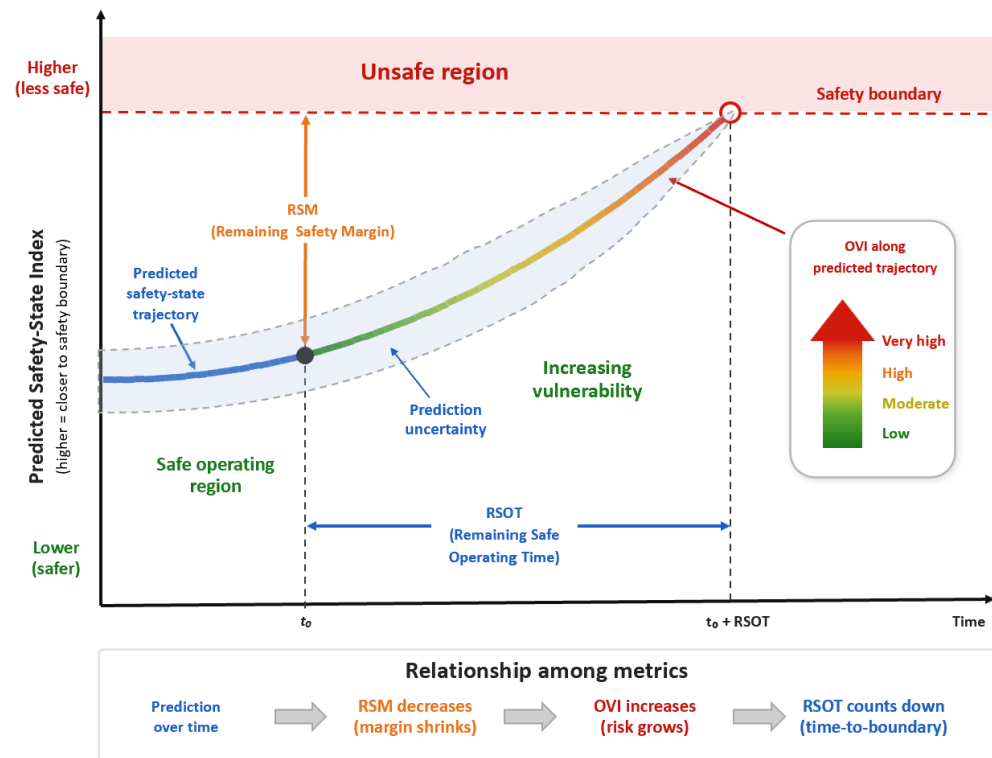


Figure 7. Relationship among Remaining Safety Margin (RSM), Remaining Safe Operating Time (RSOT), Operational Vulnerability Index (OVI), and the intervention window. The predicted safety-state trajectory (blue, with uncertainty band) rises toward the safety boundary (red dashed line). RSM is the vertical distance remaining at the current time t_0 ; RSOT is the horizontal interval from t_0 to the predicted boundary crossing at $t_0 + \text{RSOT}$; OVI increases as the trajectory accelerates toward the boundary. The shaded green region is the intervention window available to the operator before safe operation is lost.

11.5. Safety Horizon

Computational tractability of RSOT: The first-passage-time distribution $F_T(q \mid \mathcal{F}_t)$ is generally not available in closed form for nonlinear multi-variable processes. Three families of approximation are compatible with real-time industrial deployment. First, ensemble trajectory methods propagate a set of N model trajectories (e.g., $N = 100\text{--}1000$ particles or Monte Carlo draws) forward in time and estimate F_T empirically from the fraction of trajectories that cross $\mathcal{B}$ by each horizon q ; for moderate-dimensional systems, this may be computationally feasible at industrial process-control timescales [41,49]. Second, linearization-based bounds propagate first- and second-order uncertainty through the process model to obtain Gaussian or ellipsoidal approximations to the safety-margin distribution, from which conservative RSOT bounds can be derived analytically. Third, data-driven surrogate models, including recurrent neural

networks, Gaussian processes, and physics-informed neural networks, can be trained offline and evaluated online at millisecond timescales, making them suitable for high-frequency safety monitoring [49]. The appropriate approximation depends on the process dimensionality, model fidelity, and required update frequency; selecting and validating a tractable RSOT estimator for a given industrial application is itself a research challenge addressed in the research agenda (Section 14).

Safety Horizon is the future interval over which the safety-state forecast remains sufficiently reliable for operational use. It concerns forecast validity, not whether a predicted state belongs to the state space. Let

$$c_S(t, \tau) = \Pr(\|X_t(\tau) - \hat{x}(t + \tau | t)\|_W \leq \varepsilon_S | \mathcal{F}_t) \quad (10)$$

denote confidence that the future state lies within a safety-relevant tolerance ε_S of the predicted state, using a process-specific weighted norm $\|\cdot\|_W$. For prediction horizon H_p , the Safety Horizon is

$$H_s(t) = \sup\{h \in [0, H_p] : c_S(t, \tau) \geq \alpha \text{ for all } \tau \in [0, h]\}, \quad (11)$$

where α is the minimum acceptable forecast-confidence level and the weighting matrix W scales heterogeneous process variables into a dimensionally consistent norm. If the admissible set is empty, $H_s(t)$ is defined as zero. This definition prevents an isolated later increase in confidence from extending the horizon across an intervening interval of unreliable prediction. In a nonlinear implementation, $H_s(t)$ can be computed on a discrete prediction grid by forward-simulating an ensemble or Monte Carlo sample, estimating $c_S(t, \tau)$ at each lead time, and selecting the last consecutive grid point for which the confidence criterion is satisfied; reachability bounds or calibrated surrogate models may be substituted when they provide conservative coverage. Alternative implementations may replace the state-error event in Equation (10) with a calibrated safety-regime classification criterion, but they must evaluate predictive validity rather than the tautological event $S(t + \tau) \in \mathcal{S}$.

11.6. Predictive Safety Confidence

Predictive Safety Confidence (PSC) is the conditional probability that sufficient time remains to complete a specified intervention before the first safety-boundary crossing. For action a , the required time should be scenario- and policy-dependent: $T_{\text{req}}(a, t) = T_{\text{detect}} + T_{\text{diagnose}} + T_{\text{authorize}} + T_{\text{communicate}} + T_{\text{execute}} + T_{\text{process}} + T_{\text{buffer}}$, where the terms account respectively for detection, diagnosis, authorization, communication, field execution, process response, and a safety buffer. Under declared assumptions about controller behavior, disturbance evolution, safeguard condition, and operator response, then

$$\begin{aligned} \text{PSC}_a(t) &= \Pr(T_B(t) > T_{\text{req}}(a, t) | \mathcal{F}_t) \\ &= 1 - F_T(T_{\text{req}}(a, t) | \mathcal{F}_t). \end{aligned} \quad (12)$$

$\text{PSC}_a(t) = 1$ indicates complete predictive confidence, under the adopted model and uncertainty assumptions, that action a can be completed in time; $\text{PSC}_a(t) = 0$ indicates that timely completion is predicted to be impossible. Values between zero and one quantify the available intervention-time confidence. PSC is action-dependent: a rapid automated action and a slower operator procedure can have different PSC values under the same predicted process trajectory.

PSC is not intended to replace OVI or to represent consequence severity. OVI captures the overall urgency of a developing scenario, whereas PSC measures confidence that a particular intervention remains feasible in time. A low PSC may occur well before any

conventional alarm activates when the predicted intervention window is shorter than the required response time; conversely, alarm activation does not imply PSC is zero if effective action time remains. Facility-specific PSC decision thresholds must therefore be calibrated against prediction performance, intervention durations, and acceptable missed-warning and false-warning rates rather than assigned universally.

11.7. Operational Vulnerability

Operational Vulnerability describes the degree to which the process is approaching a condition where safety margins may be lost under predicted future evolution. It is not equivalent to risk probability alone; rather, it combines proximity to a safety boundary, rate of margin depletion, consequence severity, safeguard unavailability, and predictive uncertainty. The following bounded weighted model provides a computable reference implementation:

$$\begin{aligned} p_M(t) &= \text{clip}(1 - \text{RSM}(t), 0, 1), \\ p_R(t) &= \text{clip}\left(\frac{[-\dot{\text{RSM}}(t)]_+}{r_{\text{ref}}}, 0, 1\right), \\ \text{OVI}(t) &= w_M p_M(t) + w_R p_R(t) + w_C C(t) + w_A [1 - A_s(t)] + w_U U(t), \end{aligned} \quad (13)$$

where $[z]_+ = \max(z, 0)$, $\text{clip}(z, 0, 1) = \min(1, \max(0, z))$, and $r_{\text{ref}} > 0$ is a process-specific reference rate of normalized margin depletion. Because RSM is the minimum of boundary-specific margins, it may be nondifferentiable when the governing boundary changes; at such points, $\dot{\text{RSM}}$ denotes a conservative one-sided derivative or an equivalent backward finite-difference estimate. When a differentiable surrogate is required for optimization or learning, an implementation may replace the hard minimum with a smooth soft-min, for example $\text{RSM}_\beta(t) = -(1/\beta) \log \sum_{j \in \mathcal{J}} \exp[-\beta M_j(t) / M_j^{\text{ref}}]$, which approaches the hard minimum as $\beta \rightarrow \infty$. Active-set tracking and constrained-optimization formulations are alternative choices; each must still report the governing physical boundary for interpretability. The quantities $C(t)$, $A_s(t)$, and $U(t)$ are normalized to $[0, 1]$: $C = 1$ denotes maximum credible consequence severity, $A_s = 1$ denotes full availability of credited safeguards, and $U = 1$ denotes maximum admissible predictive uncertainty. The weights satisfy

$$w_k \geq 0 \quad \text{for } k \in \{M, R, C, A, U\}, \quad w_M + w_R + w_C + w_A + w_U = 1, \quad (14)$$

so that $\text{OVI} \in [0, 1]$, with larger values denoting greater intervention urgency. Weights and decision bands are not universal constants. A defensible elicitation procedure is to define candidate scenarios from HAZOP and LOPA records, obtain structured judgements from process-safety and operations experts (for example, through pairwise comparison or multi-criteria workshops), estimate or update priors using historical incident and near-miss records where available, and then test sensitivity of the urgency ranking to plausible weight changes. A calibration set containing representative upset trajectories should be held out to assess whether the selected bands rank scenarios consistently with expert safety judgement. This procedure is essential because inappropriate weights can invert the ranking of intervention urgency. For multiple credible scenarios or active boundaries, Equation (13) should be evaluated separately for each scenario, with the maximum reported as the plant-level OVI together with the identity of the governing scenario. This preserves interpretability and prevents aggregation from hiding the binding safety boundary.

OVI should be distinguished from the Risk Priority Number (RPN) used in Failure Mode and Effects Analysis (FMEA) [85]. RPN is a static, failure-mode-centered index (the product of severity, occurrence, and detectability) assigned offline during design review to rank corrective actions across failure modes. OVI is a continuously updated, trajectory-based signal: its inputs (RSM depletion rate, predicted margin, safeguard availability,

consequence severity, and predictive uncertainty) evolve at the timescale of the monitored process, not at the timescale of engineering review cycles. OVI therefore characterizes the time-varying urgency of intervention during a developing operational scenario, whereas RPN characterizes the relative design-time priority of failure modes for maintenance and engineering corrective action. OVI is accordingly not proposed as a universal risk score; it is a scenario-specific intervention-urgency signal whose weights, normalization conventions, and decision bands must be justified by plant-specific hazard analysis and calibrated against operational experience before use.

RSM, RSOT, OVI, and PSC together constitute the core quantitative vocabulary of POSE. Table 9 summarizes their definitions and intended interpretations. These metrics constitute the output specification for the POST reference architecture described in the following section.

Table 9. Core safety metrics of POSE and their operational interpretation.

Metric	Definition	Operational Question Answered
Remaining Safety Margin (RSM)	Normalized remaining distance to the nearest binding safety boundary	How far is the process from a safety boundary?
Remaining Safe Operating Time (RSOT)	Predicted time until RSM reaches zero under current trajectory	How much time remains before a boundary is crossed?
Operational Vulnerability Index (OVI)	Bounded weighted composite of proximity, depletion rate, consequence, safeguard unavailability, and uncertainty	How urgent is intervention for the governing scenario?
Predictive Safety Confidence (PSC)	Probability that first-passage time exceeds the time required for a specified intervention	What is the probability that action can be completed in time?

12. A Reference Architecture for Predictive Operational Safety Engineering: The Predictive Operational Safety Twin

12.1. Overview

The previous sections introduced POSE as an emerging research paradigm concerned with forecasting the future operational safety state of industrial processes. However, a scientific discipline requires practical implementations capable of translating conceptual principles into operational systems. To illustrate how POSE may be realized in practice, this section proposes the Predictive Operational Safety Twin (POST) as a reference architecture.

POST is presented as an example implementation rather than the only possible realization of POSE. Future researchers may develop alternative architectures based on different computational techniques, artificial intelligence models, industrial applications, or degrees of automation while remaining consistent with the fundamental principles of POSE. Accordingly, POST should be interpreted as a reference framework demonstrating how predictive process intelligence, predictive safety intelligence, and human safety intelligence can operate within a unified operational safety system.

12.2. Design Philosophy

Conventional industrial digital twins are primarily designed to estimate and predict process variables such as temperature, pressure, flow, composition, equipment condition, and production performance [13,51]. POST adopts a different objective. Rather than estimating only the future process state, POST continuously estimates the future operational safety state of the process.

Within POST, alarms, process variables, equipment conditions, safety boundaries, and safety knowledge are not treated as independent information sources. Instead, they collectively contribute to a continuously evolving representation of operational safety. The primary output of POST is therefore not a collection of predicted measurements but an engineering assessment of future operational safety.

POST should therefore be understood as a safety-reasoning architecture rather than a standalone automation system. Its purpose is not to replace DCS, APC, SIS, alarm systems, HAZOP, or operator authority, but to connect their outputs into a predictive safety-state estimate that can support earlier and better-informed intervention.

12.3. Reference Architecture

The proposed POST architecture consists of three interacting intelligence layers.

12.3.1. Layer I: Predictive Process Intelligence

The first layer represents the physical process using dynamic process models, hybrid digital twins, first-principles simulations, or artificial intelligence. Its objective is to continuously estimate future process trajectories under current operating conditions or candidate interventions. Representative outputs include predicted temperatures, pressures, flowrates, compositions, equipment states, inventories, and constraint distances.

12.3.2. Layer II: Predictive Safety Intelligence

The second layer transforms predicted process trajectories into quantitative measures describing future operational safety. Typical outputs include predicted alarm evolution, predicted alarm sequences, alarm flood horizon, Remaining Safe Operating Time (RSOT), Operational Vulnerability Index (OVI), Remaining Safety Margin (RSM), safety-aware alarm prioritization, and dynamic operational safety margins. Rather than simply reporting process deviations, this layer evaluates how future process evolution influences overall safety.

12.3.3. Layer III: Human Safety Intelligence

The third layer converts predictive safety information into engineering reasoning that supports human decision making. This layer integrates HAZOP knowledge, process topology, safeguard information, explainable artificial intelligence, operator guidance, and recommended intervention strategies [61,69]. Consequently, POST provides operators not only with predicted abnormalities but also with engineering interpretation and recommended operational responses.

The key architectural principle is that no single layer is sufficient: process prediction without safety interpretation is incomplete, safety metrics without human-readable explanation are unactionable, and decision support without an accurate underlying model is unreliable. A deployable implementation must also specify the interfaces and operating limits of each layer: data sources and timestamps, synchronization and update frequency, latency budgets, missing-data handling, model-fidelity and uncertainty status, degraded-mode behavior, cybersecurity controls, model versioning, and an auditable record of recommendations and operator responses. Model scale should match the safety decision: a low-order or data-driven model may be sufficient for short-horizon advisory screening, whereas safety-critical boundary forecasts require a validated hybrid or first-principles model with explicit uncertainty management. Update frequency must likewise be selected relative to the fastest relevant process and intervention time constant; there is no universal rate. Table 10 makes these implementation dependencies explicit. Figure 8 then shows how the three layers compose into a system whose output, operator guidance, emerges from the combination of all three, not from any one layer independently.

Table 10. Indicative POST implementation levels. These are design classes, not universal accuracy claims; PSC and RSOT require plant-specific calibration and uncertainty validation at every level.

Implementation Level	Model and Data Requirement	Appropriate Output	Update and Assurance Requirement
Advisory screening	Low-order first-principles, empirical, or data-driven model with basic state synchronization	Qualitative safety regime, trend direction, conservative screening of RSM or RSOT	Update relative to the relevant process time constant; do not present uncalibrated PSC as a quantitative action guarantee
Unit-level predictive advisory	Hybrid or validated data-assimilating model with boundary definitions, uncertainty propagation, and safeguard status	Scenario-specific RSM and RSOT intervals; action-specific PSC only with a documented action-time model	Update frequency and latency must be demonstrated adequate for the shortest credible intervention window
High-consequence research or pilot	Validated hybrid or first-principles dynamic model with model-discrepancy monitoring, version control, and degraded-mode handling	Calibrated uncertainty bounds and comparative action ranking for supervised decision support	Independent benchmark, sensitivity, and human-factors evaluation required; POST remains separate from SIS logic

Table 11 specifies the minimum deployment-assurance requirements for a future POST implementation.

Table 11. Minimum deployment-assurance specification for a future POST implementation. The table defines auditable design requirements, not capabilities demonstrated in this conceptual review.

Concern	Required Design and Evidence
Data interfaces and synchronization	Document source systems, units, timestamps, quality flags, reconciliation rules, and the maximum permitted clock skew for each safety-relevant tag.
Latency and update management	Define a latency budget from data acquisition through operator display; demonstrate that it is shorter than the minimum credible intervention window for the intended advisory use.
Missing or degraded data	Detect missing, stale, inconsistent, or out-of-range inputs; propagate their effect into uncertainty; and enter a labeled degraded or unavailable advisory mode rather than silently extrapolating.
Model mismatch and drift	Monitor residuals, prediction coverage, and out-of-distribution conditions; trigger conservative margins, recalibration, or withdrawal of quantitative RSOT/PSC outputs when validity is not supported.
Cybersecurity and access control	Segregate POST from SIS logic, use authenticated interfaces and least-privilege access, and record configuration changes according to the facility's cybersecurity governance.
Versioning, auditability, and failure management	Version models, boundary definitions, and knowledge bases; retain the input snapshot, uncertainty status, recommendation, and operator disposition; test fallback behavior and escalation procedures.

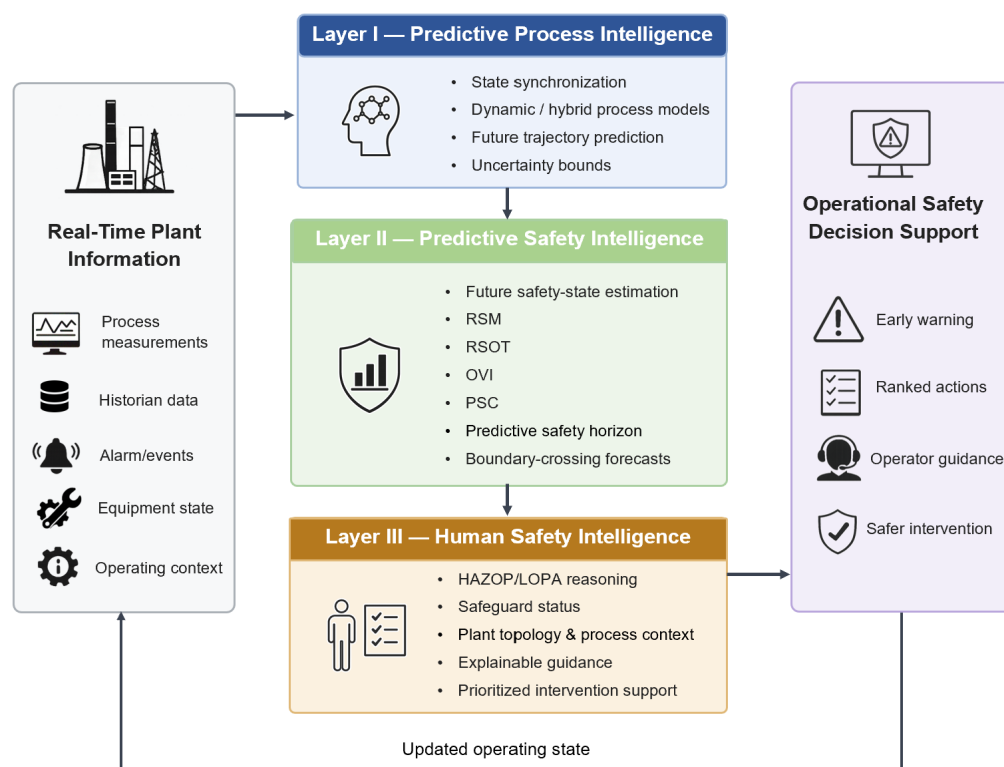


Figure 8. Three-layer reference architecture of the Predictive Operational Safety Twin (POST). Real-time plant information (process measurements, historian data, alarm events, equipment state, and operating context) enters Layer I (Predictive Process Intelligence), which performs state synchronization, hybrid process modeling, future trajectory prediction, and uncertainty quantification. Layer II (Predictive Safety Intelligence) translates these trajectories into safety-state metrics: RSM, RSOT, OVI, PSC, Alarm Flood Horizon, and boundary-crossing forecasts. Layer III (Human Safety Intelligence) applies HAZOP/LOPA reasoning, safeguard status assessment, and explainable guidance to produce ranked interventions and operator decision support. Updated plant response closes the feedback loop. POST integrates process prediction, predictive safety metrics, and human-centered safety reasoning into a unified operational safety architecture.

12.4. Operational Workflow

The operational workflow of POST consists of six sequential stages. First, time-stamped process measurements are acquired from industrial sensors, distributed control systems, process historians, laboratory systems, and advanced process control platforms, with missing, delayed, or inconsistent data explicitly flagged. Second, the digital twin synchronizes the state and estimates future process trajectories over a prediction horizon appropriate to the process dynamics. Third, predicted trajectories are translated into expected alarm evolution, safety-margin reduction, uncertainty, and operational vulnerability. Fourth, engineering knowledge, process topology, and HAZOP relationships interpret predicted abnormalities to identify probable causes, credible consequences, and safeguard involvement. Fifth, the framework ranks candidate actions using stated criteria: expected safety improvement, time to execute, probability of success, reversibility, side effects, production impact, operator workload, procedural compatibility, and interactions with concurrent actions. Sixth, following operator or controller intervention, plant measurements update the twin and its audit record so that the predictive safety assessment evolves with the physical process. Table 12 specifies the minimum information that a future action-ranking module should expose. It is a transparent multi-criteria decision specification, not a claim that this review has validated a universal ranking algorithm.

Table 12. Reference specification for POST action ranking. Hard safety and procedural constraints should be evaluated before any weighted preference score.

Decision Element	Evidence Required	Decision Role
Admissibility constraints	Procedure authorization, equipment availability, safeguard state, operator permissions, and incompatible concurrent actions	Exclude infeasible, unsafe, or procedurally prohibited actions before ranking.
Expected safety benefit	Predicted change in RSM, RSOT, boundary-crossing probability, and consequence exposure under the action scenario	Prefer actions that credibly restore or preserve the relevant safety margin.
Timeliness and success likelihood	Action completion-time distribution, PSC, field accessibility, and reliability of the required equipment or procedure	Penalize actions unlikely to complete or succeed within the intervention window.
Reversibility and side effects	Expected transients, secondary hazards, environmental effects, production impact, and recovery requirements	Avoid an apparent local safety benefit that creates a larger downstream safety or operability problem.
Human and organizational fit	Workload, training, communication burden, procedural clarity, and required coordination	Prefer actions that can be understood and executed reliably by the available operating team.
Conflict resolution and explanation	Scenario identity, active constraints, criterion values, uncertainty, and rationale for rejecting alternatives	Present the ranked recommendation as explainable advisory support, with operator authority retained.

12.5. Representative Safety Metrics

Predictive Operational Safety Engineering introduces safety-oriented performance indicators that differ from conventional process monitoring metrics because they describe future operational safety rather than current process conditions. Illustrative examples include RSOT, OVI, Alarm Flood Horizon, Predictive Alarm Sequence Confidence, Safety-aware Alarm Priority Index, Dynamic Safety Margin, Predicted Consequence Severity, and Expected Operator Intervention Window. The exact mathematical definitions of these quantities may vary by implementation and industrial application, but their common purpose is to make future safety degradation visible, interpretable, and actionable.

12.6. Relationship to Existing Technologies

POST integrates existing industrial technologies rather than replacing them: digital twins provide process prediction, alarm management provides alarm intelligence, fault diagnosis identifies abnormal conditions, HAZOP contributes structured safety knowledge, and operator support provides human-centered decision making. The key difference is that POST specifies how these capabilities can be combined into a continuously evolving estimate of the future operational safety state. Individual implementations in adjacent disciplines may offer some of these functions; POST is proposed as an integrative reference architecture rather than a claim that such capabilities are absent everywhere. Table 13 maps each POST reference function to its purpose and illustrative output, showing how the six functions span from state estimation to decision support.

Regulatory compliance positioning: POST is conceived as a decision-support layer that operates above, and is consistent with, existing functional safety standards. IEC 61511 (Functional Safety: Safety Instrumented Systems for the Process Industry Sector) governs

the design, installation, and operation of Safety Instrumented Systems (SIS) [64]; POST does not modify SIS logic or safety integrity levels. Instead, it provides early warning and operator guidance before SIS demand conditions are reached, thereby reducing the frequency of SIS demands, which is itself a recognized IEC 61511 objective. Similarly, ANSI/ISA-18.2 (Management of Alarm Systems for the Process Industries) [5] defines rationalization, prioritization, and performance metrics for alarm systems; POST is complementary to ISA-18.2 compliance in that RSM and RSOT estimates can inform alarm rationalization by quantifying the safety significance of individual alarm setpoints in terms of their proximity to safety boundaries. Future validation work should explicitly demonstrate that POSE implementations do not introduce unauthorized modifications to SIS logic, comply with Management of Change procedures, and satisfy the human factors requirements of ISA-18.2 for alarm presentation and operator workload.

Table 13. Reference functions of a Predictive Operational Safety Twin.

POST Layer	Purpose	Example Output
State synchronization	Estimate the current process and safeguard state from online data	Current operating state, active constraints, available safeguards
Trajectory prediction	Forecast future process evolution under current or candidate actions	Predicted pressure, temperature, flow, composition, or inventory trajectories
Safety-state translation	Convert predicted trajectories into safety regimes and margins	Normal, vulnerable, critical, or unsafe forecast; RSM; OVI
Time-to-boundary estimation	Estimate when relevant safety boundaries may be reached	RSOT, alarm flood horizon, trip horizon, confidence interval
Safety reasoning	Link predicted deviations to causes, consequences, and safeguards	HAZOP-linked explanation and consequence pathway
Decision support	Rank interventions according to urgency, feasibility, and safety value	Prioritized action guidance with rationale and uncertainty

12.7. Scientific Significance

The proposed architecture represents a conceptual transition in industrial safety. Traditional industrial systems estimate what is happening now. POST estimates what is likely to happen to operational safety in the near future if no intervention occurs. This distinction changes the objective of digital twins from process representation toward predictive safety reasoning. Consequently, POST should be viewed not simply as another digital twin architecture but as an operational realization of the broader POSE paradigm.

12.8. Illustrative Application Scenario

To illustrate the proposed POSE vocabulary without claiming empirical validation, the following hypothetical scenario considers a fixed-bed catalytic reactor in which temperature is monitored at multiple axial bed positions. A digital twin continuously estimates the peak bed temperature trajectory using an energy balance with online catalyst activity updating. At time t_0 , the twin predicts that the peak bed temperature will rise from 380 °C toward a selected operational boundary, illustrated here by a high-high alarm setpoint of 430 °C, over approximately 18 min if no intervention occurs. This setpoint is not asserted to be the physical hazard boundary.

Remaining Safety Margin: The reference margin M_j^{ref} is the gap between the nominal operating temperature of 350 °C and the alarm limit of 430 °C, giving $M_j^{\text{ref}} = 80$ °C. The currently predicted gap is 50 °C, so $\text{RSM}(t_0) = 50/80 = 0.625$. The process retains 62.5 % of its reference safety margin, but it is decreasing at an accelerating rate.

Remaining Safe Operating Time: The nominal trajectory gives $\widehat{\text{RSOT}}(t_0) = 18$ min. Assume that uncertainty propagation across the digital-twin ensemble produces first-passage-time quantiles $Q_{0.10} = 14$ min, $Q_{0.50} = 18$ min, and $Q_{0.90} = 23$ min. The 90 % conservative lower predictive bound is therefore $\text{RSOT}_{0.90}^L(t_0) = Q_{0.10} = 14$ min: conditional on the model and uncertainty assumptions, the boundary-crossing time is predicted to be at least 14 min away with 90 % probability. The interval $[14, 23]$ min is the central 80 % predictive interval; it should not be confused with the one-sided lower bound.

Operational Vulnerability Index: For illustration, let $\text{RSM} = -0.035 \text{ min}^{-1}$, $r_{\text{ref}} = 0.05 \text{ min}^{-1}$, normalized consequence severity $C = 0.90$, safeguard availability $A_s = 0.60$, and predictive uncertainty $U = 0.30$. With $(w_M, w_R, w_C, w_A, w_U) = (0.25, 0.25, 0.20, 0.20, 0.10)$, Equation (13) gives $p_M = 0.375$, $p_R = 0.70$, and $\text{OVI} = 0.25(0.375) + 0.25(0.70) + 0.20(0.90) + 0.20(0.40) + 0.10(0.30) \approx 0.56$. Under an illustrative advisory threshold of 0.50, POST identifies elevated intervention urgency. The threshold and weights would require process-specific calibration before operational use.

Predictive Safety Confidence: Suppose the selected cooling intervention requires $T_{\text{req}} = 10$ min, including diagnosis, authorization, execution, and an engineering buffer. If the first-passage ensemble gives $\Pr(T_B > 10 \text{ min} \mid \mathcal{F}_{t_0}) = 0.98$, then Equation (12) gives $\text{PSC}_a(t_0) = 0.98$. POST therefore reports high confidence that this action can still be completed, while the declining RSM and OVI of 0.56 indicate that delaying intervention is undesirable.

HAZOP-informed reasoning: The POST safety reasoning layer retrieves the HAZOP entry for HIGH TEMPERATURE in the reactor. The deviation is MORE TEMPERATURE; the probable cause is catalyst activity increase or cooling circuit fault; the credible consequence is thermal runaway; the critical safeguard is the cooling water control valve. POST presents the operator with the recommended action: verify and increase cooling water flow, confirm control valve position, and monitor the rate of change of peak bed temperature.

Contrast with conventional alarm response: A conventional alarm management system activates the high-high alarm when the measured temperature reaches 430°C . In this simplified example, that setpoint is used only as a selected operational boundary for demonstrating the metric calculations; it is not a substitute for a hazard boundary. An industrial implementation should distinguish advisory, alarm, trip, equipment-design, and hazardous-condition boundaries, and should formulate RSOT relative to the boundary appropriate to the decision being supported. POST forecasts the approach to the selected boundary before alarm activation and reports a 14-min conservative lower intervention window, an OVI of 0.56, and action-specific PSC of 0.98. This operational difference, specifically reasoning about intervention feasibility before versus after alarm activation, is the central practical value of predictive operational safety reasoning. The numerical values throughout this scenario (RSM, RSOT quantiles, OVI, and PSC) are illustrative only; they are not derived from real plant data, validated process models, or experimental measurements, and should not be interpreted as engineering thresholds or performance results. The following section distills the scientific arguments of this review into ten guiding principles for Predictive Operational Safety Engineering.

13. Ten Principles of Predictive Operational Safety Engineering

Table 14 presents the ten principles that synthesize the core scientific arguments of this review. Each is grounded in the evidence examined in Sections 3–8.

Table 14. Ten principles of Predictive Operational Safety Engineering and their meaning for POSE research and practice.

#	Principle	Meaning for POSE
1	Safety is a dynamic state, not merely an event	Forecast safety evolution continuously; systems that respond only to discrete events forfeit the prediction window available beforehand
2	Future safety-state prediction provides earlier intervention value than present alarm status alone	Predicting safety degradation before alarm limits are reached provides substantially greater intervention time [6]
3	Alarms are symptoms of an evolving safety state, not the safety state itself	Use alarms as evidence of safety-state change; alarms cannot describe how margins are evolving or when boundaries will be crossed [5,6]
4	Fault diagnosis is necessary but insufficient for safety reasoning	Convert diagnostic outputs into safety-state trajectories; detection accuracy alone does not quantify how rapidly safety margins are depleting [7,36]
5	For safety-critical applications, digital twins should predict safety meaning, not only process variables	Map process forecasts into RSM, RSOT, OVI, and PSC before they become actionable for operational safety [13,14]
6	Safety intelligence must integrate physics, data, knowledge, and human expertise	Combine first-principles models, data-driven methods, HAZOP/LOPA knowledge, safeguard states, and operator expertise; no single source is sufficient [3]
7	Prediction without interpretation does not improve safety	Apply HAZOP-informed consequence and safeguard reasoning to translate process forecasts into operator-relevant guidance [69,86] (Sections 6 and 7)
8	Operators remain central to industrial safety decisions	Design for situation awareness, trust, and action; predictive information must match human cognitive needs to improve rather than impair decision making [73] (Section 7)
9	Safety metrics must be predictive	RSM, RSOT, OVI, and PSC complement fault labels, alarm counts, and diagnostic outputs by translating them into predictive safety-state information that quantifies margin, intervention time, vulnerability, and confidence
10	Validation is mandatory	Benchmark, calibrate, explain, and test human impact; operators must understand and trust predictions for POSE to improve safety-critical decisions (Section 14)

The research agenda required to validate these principles and mature POSE into an operational engineering discipline is presented in the following section.

14. Research Agenda for Predictive Operational Safety Engineering

14.1. Grand Research Challenge

Industrial process safety has reached an important transition point. Alarm management, FDD, digital twins, process safety analysis, and operator decision support have each matured into successful research disciplines, yet they have largely progressed independently. POSE provides a conceptual framework for integrating these disciplines into a unified predictive safety paradigm, but its development is only beginning.

The central scientific challenge can be summarized by a single question: *how can industrial systems continuously estimate, interpret, and improve the future operational safety state of a process before hazardous operating conditions develop?*

Unlike conventional questions that focus on fault detection, alarm reduction, or process optimization independently, this challenge places future operational safety at the center of intelligent industrial decision making [87]. Addressing it requires contributions from process systems engineering, artificial intelligence, process safety, human factors, automation, and control engineering.

14.2. Future Research Directions

14.2.1. Physics-Informed Predictive Safety Twins

Future POST implementations should combine first-principles process models with machine learning and data-driven digital twins. Physics-informed predictive safety twins may improve prediction accuracy, robustness under previously unseen operating conditions, and engineering interpretability [13,49,51,88].

14.2.2. Uncertainty-Aware Operational Safety

Future predictive safety systems should quantify uncertainty associated with predicted process trajectories, operational vulnerability, and remaining intervention time. Rather than producing deterministic predictions alone, they should communicate prediction confidence in a form that supports risk-informed operator decisions [78,80]. A specific early milestone is an OVI weight-elicitation and calibration protocol combining structured safety-engineer judgement, historical operating evidence where available, sensitivity analysis, and a hold-out scenario set for checking whether rankings remain stable and safety-relevant.

14.2.3. Human-Centered Predictive Safety

Operators remain central to industrial safety. Future research should investigate how predictive safety information can be presented without increasing cognitive workload. Key topics include adaptive human-machine interfaces, explainable artificial intelligence, workload estimation, trust calibration, operator learning, and adaptive decision support [44,45,70,72,73,75].

14.2.4. Dynamic Safety Knowledge

Current HAZOP studies are largely static. Future investigations should develop methods capable of continuously updating and operationalizing engineering safety knowledge using operational data, digital twins, and expert reasoning [69]. Dynamic HAZOP represents an important research opportunity for combining engineering knowledge with predictive operational intelligence.

14.2.5. Multi-Unit and Plant-Wide Safety Twins

Most existing digital twins focus on individual process units. Future research should investigate plant-wide predictive safety frameworks capable of evaluating interactions among reactors, separation systems, utilities, storage facilities, environmental protection systems, and shared safeguards. Plant-wide operational safety remains largely unexplored.

14.2.6. Cyber-Physical Operational Safety

The increasing integration of industrial communication networks introduces new interactions between cybersecurity and operational safety [1,89]. Future predictive safety systems should evaluate process disturbances, cyber attacks, sensor failures, communication failures, and control system degradation within a unified operational safety framework.

14.2.7. Edge Computing and Real-Time Deployment

Deploying POSE in industrial environments requires safety reasoning to operate at the time scales of the process being monitored. For fast processes or emergency response sce-

narios, RSOT and OVI estimates must be updated within seconds. Future research should investigate lightweight model architectures, edge-deployable digital twins, and online learning methods capable of sustaining predictive safety reasoning on industrial hardware with limited computational resources, including DCS-integrated edge nodes operating close to process equipment.

14.2.8. Autonomous Safety Intelligence

Current safety systems mainly support operator decisions. Future systems may actively recommend optimized intervention strategies, evaluate multiple response alternatives, and cooperate with advanced process control or safety instrumented systems. Such systems would represent an important step toward autonomous operational safety management while maintaining appropriate human oversight.

14.3. Validation and Evaluation Agenda

For POSE to mature from a conceptual paradigm into an engineering discipline, it must be evaluated with metrics that go beyond fault classification accuracy or alarm count reduction [90]. This article is the first, conceptual and evidence-synthesis phase of a staged research program; the following validation agenda defines the subsequent phases rather than implying completed validation. A POSE system should be assessed on whether it predicts safety degradation early enough, whether its predicted safety state is correct, whether its time-to-boundary estimates are calibrated, and whether its recommendations improve operator response.

Six evaluation dimensions are proposed. The first five assess predictive and human-performance value; the sixth assesses whether a composite urgency signal is calibrated rather than merely numerically convenient.

1. Safety-state prediction accuracy: agreement between predicted and realized safety regimes over a defined horizon.
2. Intervention-time accuracy: error and calibration of RSOT estimates relative to actual or simulated boundary-crossing times.
3. Margin reliability: consistency between predicted RSM and observed approach to safety boundaries under disturbances and control actions.
4. Explanation fidelity: correctness of the link between predicted deviations, plausible causes, consequences, safeguards, and recommended interventions.
5. Human performance impact: improvement in operator prioritization, response time, workload, and abnormal-situation outcome compared with conventional alarm or FDD displays.

These criteria imply that benchmark development is a major research need. Existing process monitoring benchmarks often provide fault labels and process variables but do not provide explicit safety-state labels, dynamic HAZOP mappings, safeguard states, operator action windows, or ground-truth intervention-time metrics. Future benchmarks for POSE should therefore include process trajectories, alarm logs, safety boundaries, HAZOP-style consequence knowledge, safeguard availability, and scenario outcomes under alternative interventions.

14.4. Scientific Challenges and Open Questions

It is important to distinguish between the conceptual novelty of POSE and its empirical validation. POSE and POST are proposed frameworks. The metrics RSM, RSOT, OVI, and PSC have been formally defined here and their properties argued from synthesis of the literature, but their accuracy, calibration, and practical utility on real industrial systems have not yet been demonstrated. No field study, controlled experiment, or simulation benchmark

has been used to evaluate the operational effectiveness of any POSE implementation. The contributions of this paper are conceptual and should be interpreted as a research paradigm and a set of hypotheses awaiting experimental confirmation, not as validated engineering tools ready for industrial deployment.

This situation is consistent with how foundational engineering paradigms have historically been established. Model predictive control was proposed as a conceptual framework by Richalet et al. [91] and subsequently validated and refined over more than a decade before becoming the dominant advanced control methodology in the process industries. The digital twin concept was introduced by Grieves [12] as a formal conceptual definition without empirical demonstration; it is now a major industrial and research discipline. HAZOP was formalized as a structured methodology and disseminated widely for years before systematic comparative studies were conducted. These precedents do not reduce the obligation to validate POSE experimentally; they confirm that establishing a conceptual framework, a formal vocabulary, and a coherent set of research hypotheses is a recognized scientific contribution that precedes, and is necessary for, organized empirical investigation. This paper fulfills the first of those two obligations; the validation agenda in Section 14 defines the second.

Maturing POSE from a conceptual framework into a validated engineering discipline requires resolving several interconnected scientific and technical challenges.

14.4.1. Data Availability and Benchmark Limitations

Validating POSE methods requires process datasets that include not only measured variables and fault labels but also alarm logs, safety boundary definitions, HAZOP-derived consequence knowledge, safeguard availability records, and ground-truth operator action timelines. Such datasets are rarely available in the open literature. Building POSE-specific benchmark datasets that include safety-annotated trajectories, dynamic HAZOP mappings, and outcome labels is itself a major research contribution.

14.4.2. Model Uncertainty and Prediction Reliability

Predictive safety systems must produce trustworthy forecasts under conditions that may deviate substantially from training data. Uncertainty accumulates over prediction horizons and may render RSOT estimates unreliable precisely when they are most needed. Methods for uncertainty quantification, conformal prediction, and reliable extrapolation beyond the training distribution are therefore essential for safety-critical applications.

14.4.3. False Warning and Missed Warning Tradeoffs

A predictive safety system that generates excessive false warnings will be discarded or overridden by operators. One that misses genuine safety events provides no protection. Managing the tradeoff between false alarm rate and missed warning rate in safety-critical applications requires consequence-weighted cost functions and operating characteristic analysis specific to process safety contexts, going beyond standard classification metrics.

14.4.4. HAZOP Quality and Knowledge Currency

POSE depends on access to accurate, comprehensive, and current HAZOP and LOPA knowledge [61,69]. In practice, documentation quality varies substantially across facilities and may not reflect plant modifications made under Management of Change procedures. Methods for detecting when HAZOP knowledge has become stale and for updating safety knowledge representations from process topology changes represent important engineering challenges.

14.4.5. Computational Speed and Real-Time Deployment

Safety reasoning algorithms must produce RSOT, OVI, and operator guidance within the time scales relevant to the monitored process. Physics-based models, uncertainty propagation methods, and HAZOP reasoning engines must each be optimized for real-time deployment on industrial computing hardware, including edge devices with constrained resources and latency constraints imposed by DCS communication architectures.

14.4.6. Operator Trust and Explainability

Operators will not act on safety predictions they do not understand or trust. Methods for explaining RSOT estimates, OVI trends, and recommended actions in terms of familiar process engineering concepts, such as HAZOP deviations, safeguard status, and variable trends, are necessary for operational acceptance. Calibrating operator trust to avoid both over-reliance and under-reliance on AI-generated warnings requires dedicated human factors research.

14.4.7. Integration with DCS, APC, and SIS

Deploying POSE in industrial facilities requires integration with distributed control systems, advanced process control platforms, safety instrumented systems, process historians, and maintenance management systems. Technical challenges include data latency, communication reliability, cybersecurity, vendor interoperability, and the validation requirements imposed by safety lifecycle standards for software used in safety-related applications.

14.4.8. Regulatory Acceptance and Safety Lifecycle Governance

Safety-critical AI systems in process industries must satisfy regulatory expectations for reliability, auditability, and explainability. Existing process safety management frameworks and safety instrumented system standards were not designed with predictive AI in mind. Developing guidance for the validation, certification, and ongoing governance of POSE systems within existing regulatory frameworks represents a significant challenge at the intersection of safety engineering and regulatory policy.

14.5. Potential Industrial Applications

Although this review focuses on general process systems, POSE is broadly applicable. Representative application areas include chemical manufacturing, petroleum refining, sulfuric acid production, carbon capture, hydrogen production, ammonia synthesis, polymer manufacturing, pharmaceutical production, LNG processing, offshore production, power generation, water treatment, food processing, and mining and mineral processing. This diversity demonstrates that POSE should be regarded as a general engineering discipline rather than a technology developed for a single benchmark process. The ongoing energy transition, encompassing hydrogen production, battery systems, and process electrification, introduces new process safety challenges that further motivate the development of predictive safety frameworks [92].

Table 15 lists candidate benchmark systems for POSE validation, characterizing each by process type, key safety challenge, and relevance to POSE evaluation criteria.

Table 15. Candidate benchmark systems and industrial applications for POSE validation.

System	Process Type	Key Safety Challenge	POSE Relevance
Tennessee Eastman Process [11]	CSTR with recycle	Multiple fault modes, sensor failures	De facto FDD benchmark with rich alarm and measurement data
CO ₂ capture plant	Absorption and stripping	Solvent flooding, column pressure deviation	Dynamic safety margins under varying load conditions
Sulfuric acid converter	Catalytic SO ₂ oxidation	Catalyst deactivation, bed temperature runaway	RSOT relevant to progressive temperature excursion
Distillation column	Vapor–liquid separation	Hydraulic flooding, weeping, product deviation	Multi-variable safety boundary estimation
Ammonia synthesis loop	High-pressure catalytic reaction	Pressure safety, catalyst activity loss	RSOT estimation under high-pressure upsets
Hydrogen production unit	Steam reforming	Reformer tube integrity, H ₂ composition	Equipment–process coupled safety horizon

14.6. Grand Challenges

POSE introduces several grand challenges for future research:

1. Mathematical representation of operational safety: developing state-space, graph-based, probabilistic, and hybrid representations that can express safety boundaries, barriers, consequences, and operating context.
2. Reliable prediction of future safety states: forecasting not only process variables but safety regimes under uncertainty, nonlinear dynamics, and changing operating modes.
3. Uncertainty-aware RSOT estimation: estimating intervention time with calibrated uncertainty and communicating it in a form that supports action.
4. Dynamic HAZOP and online safety knowledge reasoning: converting static deviation-cause-consequence tables into machine-interpretable, online reasoning structures.
5. Plant-wide predictive operational safety: scaling from unit-level monitoring to inter-connected plant-wide propagation of disturbances and alarms.
6. Human-AI collaborative safety decision making: designing interfaces that preserve operator authority while improving anticipation, prioritization, and explanation.
7. Cyber-physical operational safety: accounting for cyber faults, sensor manipulation, controller compromise, and degraded data integrity in safety-state forecasts.
8. Transferable safety models across process industries: adapting POSE methods across reactors, separation systems, utilities, pipelines, and batch processes without excessive re-engineering.
9. Standard benchmarks and evaluation metrics: creating datasets that include safety boundaries, safeguard status, HAZOP knowledge, alarm logs, and outcome labels.
10. Regulatory and industrial acceptance: demonstrating reliability, explainability, auditability, and lifecycle governance sufficient for safety-critical deployment.

14.7. Limitations of This Review

Five limitations of this paper should be stated explicitly. First, the literature review is a purposive narrative review, not a systematic or scoping review; it is transparent about its thematic selection logic but is not claimed to be exhaustive or statistically reproducible. Second, the proposed metrics RSM, RSOT, OVI, and PSC are reference definitions and

mathematical frameworks, not validated performance indicators; their accuracy, calibration, cross-context comparability, and practical utility on real industrial systems have not yet been demonstrated. Third, POST is a reference architecture, not an implemented software system; no prototype, simulation, or pilot study has been conducted to evaluate its technical feasibility or operational effectiveness. Fourth, the illustrative reactor scenario in Section 12 is a hypothetical worked example whose numerical values are not derived from real plant data, validated models, or experimental measurements. Fifth, the paper does not establish POSE as a separate mature discipline; it proposes an integrative research framework and a testable agenda for its future development. Accordingly, this review should be interpreted as the first, conceptual and evidence-synthesis phase of a staged program whose subsequent phases require benchmark simulation, uncertainty calibration, baseline comparisons, operator-in-the-loop evaluation, and industrial case studies. These limitations define the boundary between what this paper claims and what remains to be established through future experimental, computational, systematic-review, and industrial research.

14.8. Ethical and Societal Considerations

The deployment of predictive safety systems in industrial environments raises ethical and societal considerations that should be addressed as POSE matures from a conceptual paradigm toward operational practice. Three issues are particularly important.

Automation bias and skill degradation: Operators who routinely receive POSE-generated guidance may become progressively dependent on algorithmic safety reasoning, reducing their capacity to diagnose abnormal situations independently when the system is unavailable, miscalibrated, or operating outside its training distribution [44,72]. This “irony of automation” is well-documented in aviation, nuclear, and process control contexts. POSE systems should therefore be designed to support and develop operator competence, not substitute for it: guidance should be accompanied by transparent HAZOP-linked reasoning that reinforces the operator’s own mental model, and periodic scenarios without POSE support should be incorporated into operator training programs.

Liability and accountability: When a POSE system recommends an intervention that is subsequently judged to have been incorrect, or fails to recommend an intervention before a safety boundary is crossed, questions of liability and accountability arise. The legal and regulatory frameworks governing software-assisted safety decision support in process industries are still evolving. Until these frameworks are established, POSE should be positioned as advisory rather than directive, and its outputs should be clearly distinguished from the recommendations of a certified Safety Instrumented System [64]. Governance documentation should record the basis for each POSE recommendation, the uncertainty bounds communicated to the operator, and the operator’s final decision.

Workforce and organizational implications: Introducing POSE into an operating plant requires not only technical integration but organizational change: operator training, role redefinition, alarm rationalization, and Management of Change procedures under IEC 61511 and ISA-18.2 [5]. Early stakeholder engagement with operators, safety engineers, and regulators is essential to ensure that POSE is perceived as a tool that enhances operator authority rather than a system that erodes human judgment. These societal dimensions of POSE deployment are as important to its long-term success as its technical performance.

14.9. POSE Research Roadmap

Figure 9 illustrates a phased research roadmap for the POSE discipline across three sequential phases.

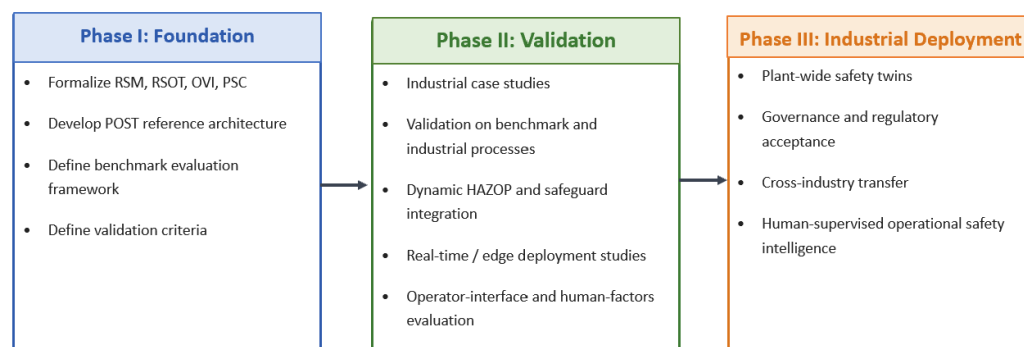


Figure 9. Research roadmap for Predictive Operational Safety Engineering (POSE) across three sequential phases. Phase I (Foundation) establishes metric definitions for RSM, RSOT, OVI, and PSC, develops the POST reference architecture, and benchmarks the framework on representative processes including the Tennessee Eastman Process. Phase II (Validation) conducts industrial case studies, integrates dynamic HAZOP and safeguard reasoning, investigates real-time and edge deployment, and evaluates operator-interface and human-factors dimensions. Phase III (Application) pursues plant-wide safety twins, regulatory and governance acceptance, cross-industry transfer, and semi-autonomous safety intelligence under human oversight. Five cross-cutting themes (uncertainty and confidence, explainability, data quality and governance, human-centered design, and lifecycle model management) span all phases.

14.10. Toward the Next Generation of Industrial Safety

The history of industrial process safety demonstrates a continuous evolution from mechanical protection toward intelligent operational support. The next stage of this evolution is expected to emphasize continuous prediction of future operational safety rather than reactive interpretation of abnormal events. Whether implemented through digital twins, hybrid artificial intelligence, first-principles modeling, or future computational technologies, the central objective remains unchanged: to provide operators with sufficient predictive safety intelligence to intervene before hazardous operating conditions fully develop.

The following section distills the contributions of this review into a set of conclusions and identifies the conditions under which POSE can be expected to mature into a validated engineering discipline.

15. Conclusions

Industrial process safety is entering a new stage. Alarm management, fault diagnosis, digital twins, HAZOP, prognostics, and operator decision support have each produced major advances. What they have not yet produced, individually or as an integrated ensemble, is a discipline whose primary objective is to continuously estimate and forecast the future operational safety state of an operating process. POSE is proposed to fill that role.

The central contribution of this paper is conceptual: operational safety can be treated as a forecastable dynamic state. Remaining Safety Margin, Remaining Safe Operating Time, Operational Vulnerability Index, and Predictive Safety Confidence are not merely renamings of existing metrics. They constitute an integrated class of predictive safety signals (multi-variable, trajectory-based, HAZOP-informed, uncertainty-aware, and intervention-oriented) that existing disciplines have not yet produced as a unified primary output. The Predictive Operational Safety Twin demonstrates how these signals can be integrated into a reference architecture whose output is interpretable operator guidance before safety boundaries are crossed, and in some cases before alarm activation.

The scientific tasks that follow from this first, conceptual and evidence-synthesis phase are concrete. The next phase should establish benchmark datasets and simulation environments capable of evaluating RSM depletion rates, RSOT accuracy, and OVI ranking

validity against alarm-based, dynamic-risk, and prognostic baselines. Subsequent phases should calibrate uncertainty and safety-boundary definitions, test whether RSOT and OVI improve operator intervention timing under realistic cognitive load, and finally evaluate the framework in industrial case studies. Until these phases are completed, POSE remains a proposed framework rather than validated engineering practice.

The foundational hypothesis of POSE is simple and directly testable: a safety system that continuously estimates remaining safe operating time and communicates it to an operator with ranked action guidance should enable earlier and more effective intervention than a system that notifies only after a limit has already been crossed. Demonstrating that hypothesis, on real industrial processes, under realistic operating conditions, and against safety outcome metrics that matter in practice, is the task to which this review calls the next generation of process safety researchers.

Funding: This research received no external funding.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: The author thanks the Chemical Engineering Department, Jubail Industrial College, for institutional support.

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Lee, J.; Bagheri, B.; Kao, H.A. A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems. *Manuf. Lett.* **2015**, *3*, 18–23. <https://doi.org/10.1016/j.mfglet.2014.12.001>.
2. Xu, L.D.; Xu, E.L.; Li, L. Industry 4.0: State of the art and future trends. *Int. J. Prod. Res.* **2018**, *56*, 2941–2962. <https://doi.org/10.1080/00207543.2018.1444806>.
3. Kletz, T.A. *What Went Wrong? Case Histories of Process Plant Disasters and How They Could Have Been Avoided*, 5th ed.; Gulf Professional Publishing: Oxford, UK, 2009.
4. Mannan, S., Ed. *Lees' Loss Prevention in the Process Industries: Hazard Identification, Assessment and Control*, 4th ed.; Butterworth-Heinemann: Oxford, UK, 2012.
5. ANSI/ISA-18.2; Management of Alarm Systems for the Process Industries. International Society of Automation (ISA): Research Triangle Park, NC, USA, 2016.
6. Engineering Equipment and Materials Users' Association. *EEMUA Publication 191: Alarm Systems—A Guide to Design, Management and Procurement*, 4th ed.; EEMUA: London, UK, 2024.
7. Qin, S.J. Survey on data-driven industrial process monitoring and diagnosis. *Annu. Rev. Control* **2012**, *36*, 220–234. <https://doi.org/10.1016/j.arcontrol.2012.09.004>.
8. Yin, S.; Ding, S.X.; Xie, X.; Luo, H. A review on basic data-driven approaches for industrial process monitoring. *IEEE Trans. Ind. Electron.* **2014**, *61*, 6418–6428. <https://doi.org/10.1109/TIE.2014.2301773>.
9. MacGregor, J.F.; Kourti, T. Statistical process control of multivariate processes. *Control Eng. Pract.* **1995**, *3*, 403–414. [https://doi.org/10.1016/0967-0661\(95\)00014-L](https://doi.org/10.1016/0967-0661(95)00014-L).
10. Chiang, L.H.; Russell, E.L.; Braatz, R.D. *Fault Detection and Diagnosis in Industrial Systems*; Springer: London, UK, 2001. <https://doi.org/10.1007/978-1-4471-0347-9>.
11. Downs, J.J.; Vogel, E.F. A plant-wide industrial process control problem. *Comput. Chem. Eng.* **1993**, *17*, 245–255. [https://doi.org/10.1016/0098-1354\(93\)80018-I](https://doi.org/10.1016/0098-1354(93)80018-I).
12. Grieves, M. *Digital Twin: Manufacturing Excellence through Virtual Factory Replication*; White Paper; Dassault Systèmes: Vélizy-Villacoublay, France, 2014.
13. Tao, F.; Zhang, H.; Liu, A.; Nee, A.Y.C. Digital twin in industry: State-of-the-art. *IEEE Trans. Ind. Inform.* **2019**, *15*, 2405–2415. <https://doi.org/10.1109/TII.2018.2873186>.
14. Fuller, A.; Fan, Z.; Day, C.; Barlow, C. Digital twin: Enabling technologies, challenges and open research. *IEEE Access* **2020**, *8*, 108952–108971. <https://doi.org/10.1109/ACCESS.2020.2998358>.
15. Jones, D.; Snider, C.; Nassehi, A.; Yon, J.; Hicks, B. Characterising the Digital Twin: A systematic literature review. *CIRP J. Manuf. Sci. Technol.* **2020**, *29*, 36–52. <https://doi.org/10.1016/j.cirpj.2020.02.002>.
16. Pal, P.K.; Hens, A.; Behera, N.; Lahiri, S.K. Digital twins: Transforming the chemical process industry—A review. *Can. J. Chem. Eng.* **2025**, *103*, 3611–3636. <https://doi.org/10.1002/cjce.25611>.

17. Virando, G.E.; Lee, B.; Kee, S.H.; Yee, J.J. Digital twin simulation and implementation in safety risk management process. *IEEE Access* **2024**, *12*, 190483–190494. <https://doi.org/10.1109/ACCESS.2024.3519704>.
18. Correa, O.C.; Mariano, J.L.S.; Amaro, E.P.; Santos, E.N. Process Safety Management in Oil and Gas Operating Units Through Digital Twin Platform. In Proceedings of the Offshore Technology Conference, Houston, TX, USA, 1–4 May 2023. <https://doi.org/10.4043/32894-MS>.
19. Alrowaie, F. Digital Twins and Artificial Intelligence for HAZOP Enhancement in Process Safety: A Critical Literature Review. *Preprints* **2026**, <https://doi.org/10.20944/preprints202606.1902.v1>.
20. Villa, V.; Paltrinieri, N.; Khan, F.; Cozzani, V. Towards dynamic risk analysis: A review of the risk assessment approach and its limitations in the chemical process industry. *Saf. Sci.* **2016**, *89*, 77–93. <https://doi.org/10.1016/j.ssci.2016.06.002>.
21. Kuhn, T.S. *The Structure of Scientific Revolutions*; University of Chicago Press: Chicago, IL, USA, 1962.
22. Mayne, D.Q.; Rawlings, J.B.; Rao, C.V.; Sokaert, P.O.M. Constrained model predictive control: Stability and optimality. *Automatica* **2000**, *36*, 789–814. [https://doi.org/10.1016/S0005-1098\(99\)00214-9](https://doi.org/10.1016/S0005-1098(99)00214-9).
23. Seborg, D.E.; Edgar, T.F.; Mellichamp, D.A.; Doyle, F.J. *Process Dynamics and Control*, 4th ed.; John Wiley & Sons: Hoboken, NJ, USA, 2016.
24. IEC 62682; Management of Alarm Systems for the Process Industries. International Electrotechnical Commission (IEC): Geneva, Switzerland, 2014.
25. Hollifield, B.R.; Habibi, E. *Alarm Management: A Comprehensive Guide*, 2nd ed.; ISA: Research Triangle Park, NC, USA, 2011.
26. Izadi, I.; Shah, S.L.; Shook, D.S.; Kondaveeti, S.R.; Chen, T. A framework for optimal design of alarm systems. *IFAC Proc. Vol.* **2009**, *42*, 651–656. <https://doi.org/10.3182/20090630-4-ES-2003.00108>.
27. Varga, T.; Szeifert, F.; Abonyi, J. Detection of safe operating regions: A novel dynamic process simulator based predictive alarm management approach. *Ind. Eng. Chem. Res.* **2010**, *49*, 3658–3669. <https://doi.org/10.1021/ie9005222>.
28. Kondaveeti, S.R.; Shah, S.L.; Izadi, I. Application of multivariate statistics for efficient alarm generation. *IFAC Proc. Vol.* **2009**, *42*, 657–662. <https://doi.org/10.3182/20090630-4-ES-2003.00109>.
29. Pariyani, A.; Seider, W.D.; Oktem, U.G.; Soroush, M. Dynamic risk analysis using alarm databases to improve process safety and product quality: Part I—Data compaction. *AIChE J.* **2012**, *58*, 812–825. <https://doi.org/10.1002/aic.12643>.
30. Rodrigo Marco, V. Alarm Flood Reduction Using Multiple Data Sources. Master's Thesis, Lund University, Lund, Sweden, 2014.
31. Manca, G.; Fay, A. Off-line analysis of dynamic causal dependencies in evolving industrial alarm floods. In Proceedings of the IEEE International Conference on Industrial Cyber-Physical Systems, Glasgow, UK, 24–26 May 2022; pp. 1–6. <https://doi.org/10.1109/ICPS51978.2022.9816853>.
32. Alinezhad, H.S.; Roohi, M.H.; Chen, T. A review of alarm root cause analysis in process industries: Common methods, recent research status and challenges. *Chem. Eng. Res. Des.* **2022**, *188*, 846–868. <https://doi.org/10.1016/j.cherd.2022.10.041>.
33. Hu, W.; Yang, G.; Wu, M. Root cause identification of industrial alarm floods using word embedding and few-shot learning. *IEEE Trans. Ind. Informatics* **2023**, *19*, 9841–9851. <https://doi.org/10.1109/TII.2023.3274223>.
34. Roohi, M.H.; Ramazi, P.; Chen, T. Towards accurate root-alarm identification: The causal Bayesian network approach. In Proceedings of the Conference on Control and Fault-Tolerant Systems, Virtual, 29 September–1 October 2021; pp. 7–12. <https://doi.org/10.1109/SYSTOL52990.2021.9595698>.
35. Jiang, W.; Hu, W.; Liu, Z.; Wang, F. An informer based alarm early prediction method over consecutive alarm monitoring periods. In Proceedings of the IEEE International Conference on Industrial Informatics, Beijing, China, 18–20 August 2024; pp. 1–6. <https://doi.org/10.1109/INDIN58382.2024.10774351>.
36. Venkatasubramanian, V.; Rengaswamy, R.; Yin, K.; Kavuri, S.N. A review of process fault detection and diagnosis: Part I: Quantitative model-based methods. *Comput. Chem. Eng.* **2003**, *27*, 293–311. [https://doi.org/10.1016/S0098-1354\(02\)00160-6](https://doi.org/10.1016/S0098-1354(02)00160-6).
37. Venkatasubramanian, V.; Rengaswamy, R.; Kavuri, S.N.; Yin, K. A review of process fault detection and diagnosis: Part III: Process history based methods. *Comput. Chem. Eng.* **2003**, *27*, 327–346. [https://doi.org/10.1016/S0098-1354\(02\)00162-X](https://doi.org/10.1016/S0098-1354(02)00162-X).
38. Venkatasubramanian, V.; Rengaswamy, R.; Kavuri, S.N. A review of process fault detection and diagnosis: Part II: Qualitative models and search strategies. *Comput. Chem. Eng.* **2003**, *27*, 313–326. [https://doi.org/10.1016/S0098-1354\(02\)00161-8](https://doi.org/10.1016/S0098-1354(02)00161-8).
39. Isermann, R. Model-based fault-detection and diagnosis—Status and applications. *Annu. Rev. Control* **2005**, *29*, 71–85. <https://doi.org/10.1016/j.arcontrol.2004.12.002>.
40. Ge, Z.; Song, Z.; Gao, F. Review of recent research on data-based process monitoring. *Ind. Eng. Chem. Res.* **2013**, *52*, 3543–3562. <https://doi.org/10.1021/ie302069q>.
41. Ge, Z.; Song, Z.; Ding, S.X.; Huang, B. Data mining and analytics in the process industry: The role of machine learning. *IEEE Access* **2017**, *5*, 20590–20616. <https://doi.org/10.1109/ACCESS.2017.2756872>.
42. Zhang, Z.; Zhao, J. A deep belief network based fault diagnosis model for complex chemical processes. *Comput. Chem. Eng.* **2017**, *107*, 395–407. <https://doi.org/10.1016/j.compchemeng.2017.02.041>.
43. Pearl, J. *Causality: Models, Reasoning, and Inference*; Cambridge University Press: Cambridge, UK, 2000.

44. Lee, J.D.; See, K.A. Trust in automation: Designing for appropriate reliance. *Hum. Factors* **2004**, *46*, 50–80. https://doi.org/10.1518/hfes.46.1.50_30392.
45. Arrieta, A.B.; Díaz-Rodríguez, N.; Del Ser, J.; Benetot, A.; Tabik, S.; Barbado, A.; Herrera, F. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Inf. Fusion* **2020**, *58*, 82–115. <https://doi.org/10.1016/j.inffus.2019.12.012>.
46. CaÃ§Ã£o, J.; Santos, J.; Antunes, M. Explainable AI for industrial fault diagnosis: A systematic review. *J. Ind. Inf. Integr.* **2025**, *47*, 100905. <https://doi.org/10.1016/j.jii.2025.100905>.
47. Zhao, S.; Yang, H.; Kareck, T.L.; Khan, F.; Wang, Q. Data-driven fault detection and diagnosis in industrial process systems: A systematic review and perspective. *Reliab. Eng. Syst. Saf.* **2025**, *270*, 112159. <https://doi.org/10.1016/j.res.2025.112159>.
48. Lai, C.; Baraldi, P.; Zio, E. Physics-Informed Deep Autoencoder for Fault Detection in New-Design Systems. *Mech. Syst. Signal Process.* **2024**, *215*, 111420.
49. Raissi, M.; Perdikaris, P.; Karniadakis, G.E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **2019**, *378*, 686–707. <https://doi.org/10.1016/j.jcp.2018.10.045>.
50. Grieves, M.; Vickers, J. Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems. In *Transdisciplinary Perspectives on Complex Systems*; Kahlen, F.J., Flumerfelt, S., Alves, A., Eds.; Springer: Cham, Switzerland, 2017; pp. 85–113. https://doi.org/10.1007/978-3-319-38756-7_4.
51. Rasheed, A.; San, O.; Kvamsdal, T. Digital Twin: Values, Challenges and Enablers from a Modeling Perspective. *IEEE Access* **2020**, *8*, 21980–22012. <https://doi.org/10.1109/ACCESS.2020.2970143>.
52. Liu, M.; Fang, S.; Dong, H.; Xu, C. Review of digital twin about concepts, technologies, and industrial applications. *J. Manuf. Syst.* **2021**, *58*, 346–361. <https://doi.org/10.1016/j.jmsy.2020.06.017>.
53. Kritzinger, W.; Karner, M.; Traar, G.; Henjes, J.; Sihn, W. Digital Twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnLine* **2018**, *51*, 1016–1022. <https://doi.org/10.1016/j.ifacol.2018.08.474>.
54. Bevilacqua, M.; Bottani, E.; Ciarapica, F.E.; Costantino, F.; Di Donato, L.; Ferraro, A.; Mazzuto, G.; MonteriÃ, A.; Nardini, G.; Ortenzi, M.; et al. Digital Twin Reference Model Development to Prevent Operators' Risk in Process Plants. *Sustainability* **2020**, *12*, 1088. <https://doi.org/10.3390/su12031088>.
55. Mane, S.; Dhote, R.R.; Sinha, A.; Thirumalaiswamy, R. Digital twin in the chemical industry: A review. *Digit. Twins Appl.* **2024**, *1*, 118–130. <https://doi.org/10.1049/dgt2.12019>.
56. Zio, E.; Miqueles, L. Digital twins in safety analysis, risk assessment and emergency management. *Reliab. Eng. Syst. Saf.* **2024**, *246*, 110040. <https://doi.org/10.1016/j.res.2024.110040>.
57. Agnusdei, G.; Elia, V.; Gnoni, M. Is Digital Twin Technology Supporting Safety Management? A Bibliometric and Systematic Review. *Appl. Sci.* **2021**, *11*, 2767. <https://doi.org/10.3390/app11062767>.
58. Hu, S.; Wang, S.; Su, N.; Li, X.; Zhang, Q. Digital Twin Based Reference Architecture for Petrochemical Monitoring and Fault Diagnosis. *Oil Gas Sci.-Technol.—Rev. D'Ifp Energies Nouv.* **2021**, *76*, 9. <https://doi.org/10.2516/ogst/2020095>.
59. Kabashkin, I. Mathematical Framework for Digital Risk Twins in Safety-Critical Systems. *Mathematics* **2025**, *13*, 3222. <https://doi.org/10.3390/math13193222>.
60. Seider, W.D.; Soroush, M.; Arbogast, J.E.; Oktem, U.G. Design for process safety—A perspective. In *Computer Aided Chemical Engineering*; Elsevier: Amsterdam, The Netherlands, 2014; Volume 33, pp. 457–462. <https://doi.org/10.1016/B978-0-444-63433-7.50117-6>.
61. Dunj3, J.; Fthenakis, V.; V3lchez, J.A.; Arnaldos, J. Hazard and operability (HAZOP) analysis: A literature review. *J. Hazard. Mater.* **2010**, *173*, 19–32. <https://doi.org/10.1016/j.jhazmat.2009.08.076>.
62. Crowl, D.A.; Louvar, J.F. *Chemical Process Safety: Fundamentals with Applications*, 3rd ed.; Prentice Hall: Upper Saddle River, NJ, USA, 2011.
63. Center for Chemical Process Safety. *Layer of Protection Analysis: Simplified Process Risk Assessment*; American Institute of Chemical Engineers: New York, NY, USA, 2001.
64. IEC 61511; Functional Safety—Safety Instrumented Systems for the Process Industry Sector. International Electrotechnical Commission (IEC): Geneva, Switzerland, 2016.
65. Khakzad, N.; Khan, F.; Amyotte, P. Safety analysis in process facilities: Comparison of fault tree and Bayesian network approaches. *Reliab. Eng. Syst. Saf.* **2011**, *96*, 925–932. <https://doi.org/10.1016/j.res.2011.03.012>.
66. Khakzad, N.; Khan, F.; Amyotte, P. Dynamic safety analysis of process systems by mapping bow-tie into Bayesian network. *Process Saf. Environ. Prot.* **2013**, *91*, 46–53. <https://doi.org/10.1016/j.psep.2012.01.005>.
67. Kalantarnia, M.; Khan, F.; Hawboldt, K. Dynamic risk assessment using failure assessment and Bayesian theory. *J. Loss Prev. Process Ind.* **2009**, *22*, 600–606. <https://doi.org/10.1016/j.jlp.2009.04.006>.
68. Paltrinieri, N.; Khan, F., Eds. *Dynamic Risk Analysis in the Chemical and Petroleum Industry*; Butterworth-Heinemann: Oxford, UK, 2016.

69. Elhosary, E.; Moselhi, O. Automation for HAZOP study: A state-of-the-art review and future research directions. *J. Inf. Technol. Constr.* **2024**, *29*, 665–697. <https://doi.org/10.36680/j.itcon.2024.033>.
70. Rasmussen, J. Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Trans. Syst. Man Cybern.* **1983**, SMC-13, 257–266. <https://doi.org/10.1109/TSMC.1983.6313160>.
71. Reason, J. *Human Error*; Cambridge University Press: Cambridge, UK, 1990.
72. Bainbridge, L. Ironies of automation. *Automatica* **1983**, *19*, 775–779. [https://doi.org/10.1016/0005-1098\(83\)90046-8](https://doi.org/10.1016/0005-1098(83)90046-8).
73. Endsley, M.R. Toward a theory of situation awareness in dynamic systems. *Hum. Factors* **1995**, *37*, 32–64. <https://doi.org/10.1518/001872095779049543>.
74. Vicente, K.J. *Cognitive Work Analysis: Toward Safe, Productive, and Healthy Computer-Based Work*; Lawrence Erlbaum Associates: Mahwah, NJ, USA, 1999.
75. Wickens, C.D. Multiple resources and performance prediction. *Theor. Issues Ergon. Sci.* **2002**, *3*, 159–177. <https://doi.org/10.1080/14639220210123806>.
76. Venger, A.L.; Dozortsev, V. Trust in Artificial Intelligence: Modeling the Decision Making of Human Operators in Highly Dangerous Situations. *Mathematics* **2023**, *11*, 4956.
77. Jardine, A.K.S.; Lin, D.; Banjevic, D. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mech. Syst. Signal Process.* **2006**, *20*, 1483–1510. <https://doi.org/10.1016/j.ymssp.2005.09.012>.
78. Si, X.S.; Wang, W.; Hu, C.H.; Zhou, D.H. Remaining useful life estimation—A review on the statistical data-driven approaches. *Eur. J. Oper. Res.* **2011**, *213*, 1–14. <https://doi.org/10.1016/j.ejor.2010.11.018>.
79. Sikorska, J.Z.; Hodkiewicz, M.; Ma, L. Prognostic modelling options for remaining useful life estimation by industry. *Mech. Syst. Signal Process.* **2011**, *25*, 1803–1836. <https://doi.org/10.1016/j.ymssp.2010.11.018>.
80. Lei, Y.; Li, N.; Guo, L.; Li, N.; Yan, T.; Lin, J. Machinery health prognostics: A systematic review from data acquisition to RUL prediction. *Mech. Syst. Signal Process.* **2018**, *104*, 799–834. <https://doi.org/10.1016/j.ymssp.2017.11.016>.
81. Peng, Y.; Dong, M.; Zuo, M.J. Current status of machine prognostics in condition-based maintenance: a review. *Int. J. Adv. Manuf. Technol.* **2010**, *50*, 297–313. <https://doi.org/10.1007/s00170-009-2482-0>.
82. Saxena, A.; Goebel, K.; Simon, D.; Eklund, N. Damage propagation modeling for aircraft engine run-to-failure simulation. In Proceedings of the 2008 International Conference on Prognostics and Health Management, Denver, CO, USA, 6–9 October 2008; pp. 1–9. <https://doi.org/10.1109/PHM.2008.4711414>.
83. Zhao, R.; Yan, R.; Chen, Z.; Mao, K.; Wang, P.; Gao, R.X. Deep learning and its applications to machine health monitoring. *Mech. Syst. Signal Process.* **2019**, *115*, 213–237. <https://doi.org/10.1016/j.ymssp.2018.05.050>.
84. Wang, Y.; Ji, Z.; Cao, Y.; Yang, S.H. Dynamic risk assessment for process operational safety based on reachability analysis. *Reliab. Eng. Syst. Saf.* **2025**, *253*, 110564. <https://doi.org/10.1016/j.ress.2024.110564>.
85. Marhavilas, P.K.; Koulouriotis, D.; Gemeni, V. Risk analysis and assessment methodologies in the work sites: On a review, classification and comparative study of the scientific literature of the period 2000–2009. *J. Loss Prev. Process Ind.* **2011**, *24*, 477–523. <https://doi.org/10.1016/j.jlp.2011.03.004>.
86. Robinson, C.; Brown, S.; Cordiner, J. Towards automated HAZOPs. In *Computer Aided Chemical Engineering*; Elsevier: Amsterdam, The Netherlands, 2021; Volume 50, pp. 491–496. <https://doi.org/10.1016/B978-0-323-88506-5.50080-2>.
87. Zio, E. The future of risk assessment. *Reliab. Eng. Syst. Saf.* **2018**, *177*, 176–190. <https://doi.org/10.1016/j.ress.2018.04.020>.
88. Wu, Y.; Sicard, B.; Gadsden, S.A. Physics-informed machine learning: A comprehensive review on applications in anomaly detection and condition monitoring. *Expert Syst. Appl.* **2024**, *255*, 124678. <https://doi.org/10.1016/j.eswa.2024.124678>.
89. Humayed, A.; Lin, J.; Li, F.; Luo, B. Cyber-physical systems security—A survey. *IEEE Internet Things J.* **2017**, *4*, 1802–1831. <https://doi.org/10.1109/JIOT.2017.2703172>.
90. Reiman, T.; Pietikäinen, E. Leading indicators of system safety—Monitoring and driving the organizational safety potential. *Saf. Sci.* **2012**, *50*, 1993–2000. <https://doi.org/10.1016/j.ssci.2011.07.015>.
91. Richalet, J.; Rault, A.; Testud, J.L.; Papon, J. Model predictive heuristic control: Applications to industrial processes. *Automatica* **1978**, *14*, 413–428. [https://doi.org/10.1016/0005-1098\(78\)90001-8](https://doi.org/10.1016/0005-1098(78)90001-8).
92. Pasman, H.; Sripaul, E.; Khan, F.; Fabiano, B. Energy transition technology comes with new process safety challenges and risks—What does it mean? *Process Saf. Prog.* **2024**, *43*, 226–230. <https://doi.org/10.1002/prs.12593>.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.