

Article

Digital Twin Framework for Nuclear Reactor Monitoring and Anomaly Detection

Veska Gancheva 

Department of Computer Systems and Technologies, Technical University of Sofia, 1000 Sofia, Bulgaria;
vgan@tu-sofia.bg

Abstract

Artificial intelligence and digital twin technologies are increasingly being explored for reactor process analysis, anomaly detection, predictive maintenance, and operator decision support in nuclear engineering. However, many existing systems remain limited by static model behavior, reduced robustness to changing operating conditions, and insufficient integration of complementary monitoring functionalities within unified monitoring workflows. In this paper, a digital twin architecture is proposed for reactor condition assessment under multivariate operating conditions. The architecture supports integration of temporal analysis and reconstruction-based anomaly detection within a layered reactor-monitoring workflow, while defining adaptive recalibration support, uncertainty-estimation and explainability modules, and physics-informed integration components intended for future extension of the analytical monitoring architecture. Experimental validation is performed using the publicly available PWR Abnormality Dataset, containing multivariate measurements from a pressurized water reactor environment. The implemented and experimentally validated component corresponds to the data-driven monitoring module of the proposed architecture. A limited proof-of-concept adaptive threshold recalibration experiment under simulated operating drift is additionally performed to illustrate one adaptive monitoring mechanism. The comparative evaluation includes temporal data partitioning, sliding-window sequence generation, controlled synthetic perturbations, anomaly score normalization, temporal smoothing, threshold optimization, and evaluation using Precision, Recall, F1-score, and ROC-AUC metrics. Several anomaly detection approaches are comparatively evaluated, including temporal residual analysis, robust multivariate monitoring, PCA-based reconstruction, anomaly score fusion, and classification using Gradient Boosting. The results show that PCA-based reconstruction achieves the highest F1-score, while score-level classification achieves the highest ROC-AUC value. The performed experiments validate the monitoring layer of the proposed architecture and demonstrate the feasibility of integrating multiple anomaly-detection approaches within a unified reactor-monitoring workflow. The study contributes a digital twin architecture for reactor monitoring together with comparative validation of its analytical monitoring layer under controlled multivariate reactor-monitoring conditions.



Academic Editors: Fabio Moro,
Yang Liu and Xingang Zhao

Received: 10 May 2026

Revised: 27 June 2026

Accepted: 10 July 2026

Published: 12 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 \(CC BY\)](#) license.

Keywords: digital twin; nuclear reactor monitoring; reactor process analysis; anomaly detection; multivariate time-series analysis; condition assessment; reconstruction-based monitoring; AI-assisted monitoring; PWR abnormality dataset

1. Introduction

Artificial intelligence and digital twin technologies are increasingly investigated as promising tools for reactor process monitoring, condition assessment, and operational support in nuclear engineering [1,2]. Recent studies have explored digital twin concepts and condition assessment frameworks for nuclear power applications, highlighting the growing importance of integrated monitoring architectures for safety-critical environments [2,3]. Reactor systems generate large volumes of multivariate measurement data, creating a need for monitoring approaches capable of identifying abnormal process behavior under changing operating conditions.

Conventional monitoring and diagnostic systems in nuclear power plants typically rely on fixed thresholds, predefined rules, or separate models for predicting specific parameters. Such approaches can provide reliable operation under steady-state conditions, but often show limited robustness to changing operating conditions, sensor drift, incomplete measurements, or complex transient conditions. In multivariate reactor data, abnormal conditions frequently manifest themselves through correlated variations across multiple process parameters, making reliable analysis through isolated evaluation of individual signals difficult.

In recent years, various applications of machine learning in nuclear engineering have been investigated, including anomaly detection, fault diagnosis, predictive maintenance, and temporal analysis of reactor data [4–7]. Statistical models, neural networks, reconstruction-based anomaly detection methods, and hybrid approaches combining data-driven and physics-based models have been explored [8–13]. Despite the reported results, much of the existing research remains focused on individual analytical tasks and does not consider the integration of different approaches into a unified framework for reactor monitoring, condition assessment, and anomaly detection [1,7]. Existing reactor-monitoring studies are frequently focused on isolated analytical tasks such as anomaly detection, forecasting, or fault classification. Less attention has been given to architectural integration of complementary monitoring approaches within unified monitoring workflows capable of supporting multivariate analysis, reconstruction-based monitoring, and operator-oriented analytical support.

At the same time, the practical deployment of AI-based methods in nuclear power applications requires a conservative approach consistent with the high demands for safety, traceability, and reliability [14,15]. Therefore, analytical and predictive systems should be considered primarily as operator-support tools rather than autonomous reactor control mechanisms. Capabilities related to reactor condition assessment, process variable trend analysis, and early identification of slowly developing abnormal conditions are of particular importance.

Existing reactor-monitoring and digital twin approaches still face limitations related to static model behavior, limited adaptive recalibration, insufficient uncertainty and explainability integration, and weak comparative validation of complementary analytical methods under common multivariate conditions.

In this work, a digital twin framework for nuclear reactor monitoring is proposed for multivariate reactor process analysis and anomaly detection. The framework defines a modular architecture that combines temporal analysis, reconstruction-based monitoring, multivariate anomaly detection, and anomaly score fusion within a common reactor-monitoring workflow. It also defines modules for adaptive recalibration, uncertainty estimation, explainability, and physics-informed integration.

The experimental part of the study focuses on the analytical monitoring layer of the proposed architecture. A comparative evaluation is performed using a publicly available pressurized water reactor (PWR) dataset containing multivariate reactor measurements.

Several complementary analytical approaches are comparatively evaluated, including temporal residual analysis, robust multivariate monitoring, PCA-based reconstruction, anomaly score fusion, and score-level Gradient Boosting classification.

The main contributions of the present study can be summarized as follows:

1. A digital twin framework for reactor monitoring and condition assessment is proposed, defining a modular architecture with components intended to support multivariate analytical monitoring, adaptive recalibration, uncertainty estimation, explainability analysis, physics-informed monitoring integration, and operator-feedback mechanisms within a unified monitoring workflow.
2. An analytical monitoring layer for reactor-monitoring anomaly detection is implemented and comparatively evaluated, including temporal residual analysis, robust multivariate monitoring, PCA-based reconstruction, anomaly-score fusion, and score-level classification.
3. An experimental evaluation workflow based on the PWR Abnormality Dataset is implemented using temporal data partitioning, sliding-window sequence generation, controlled synthetic perturbations, anomaly-score normalization, temporal smoothing, and threshold calibration.
4. Comparative evaluation of complementary analytical approaches for multivariate reactor-monitoring conditions is performed using Precision, Recall, F1-score, and ROC-AUC metrics. The experimental study is further extended through a proof-of-concept adaptive recalibration experiment under simulated operating drift.
5. Technical and implementation considerations for modular reactor-monitoring systems are discussed, including conceptual pathways for adaptive recalibration, uncertainty estimation, explainability integration, computational scalability, and operator oversight in safety-critical environments.

The remainder of the paper is organized as follows. Section 2 provides an overview of existing approaches for digital twins, anomaly detection, and AI applications in nuclear engineering. Section 3 presents the proposed digital twin framework and describes its architectural components. Section 4 discusses technical considerations and implementation aspects. Section 5 presents experimental validation and comparative results. Section 6 discusses limitations of the study and possible directions for future work.

2. Related Work

This section reviews research directions that are directly relevant to the proposed reactor-monitoring architecture. In response to the need for a more systematic positioning, the literature is organized into four groups: reactor process monitoring and diagnostic systems, machine learning methods for anomaly detection and fault diagnosis, digital twins and physics-informed modelling, and uncertainty-aware and explainable AI for operator support. The review is used to identify which capabilities have been demonstrated experimentally in prior work and which capabilities are usually treated only as architectural or post-processing extensions.

2.1. Reactor Process Monitoring, Diagnostics, and Predictive Maintenance

Conventional monitoring and diagnostic systems in nuclear power plants are primarily based on instrumentation signals, alarm thresholds, diagnostic rules, surveillance tests, and condition-monitoring procedures. Such systems are essential for operational safety and remain the reference basis for reactor monitoring. International guidance on online monitoring and condition-based maintenance emphasizes data quality, sensor reliability, validation of diagnostic outputs, and conservative interpretation of alarms before maintenance or operational decisions are made [16–19].

The main limitation of conventional threshold-based monitoring is not the lack of reliability under nominal operating conditions, but its reduced ability to identify slowly developing multivariate deviations before individual variables exceed predefined limits. In reactor systems, degradation, sensor drift, abnormal heat-transfer behavior, flow deviations, or transient operating states may appear as correlated changes across several measured variables. These dependencies are difficult to capture when signals are evaluated independently. This motivates the use of multivariate monitoring methods that can analyze relationships among temperature, pressure, flow, power-related variables, and other process variables.

Predictive maintenance studies in nuclear engineering have therefore investigated data-driven condition assessment, signal processing, machine learning, and hybrid modelling to support early warning, fault diagnosis, and maintenance planning [18–22]. These methods are intended to complement, rather than replace, established engineering monitoring and safety functions.

2.2. Machine Learning Methods for Anomaly Detection and Fault Diagnosis

Machine learning applications in nuclear engineering can be grouped according to their primary monitoring objective. First, temporal models such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks are used to learn sequential dependencies in reactor signals and to support forecasting or early identification of abnormal trends [4,5]. Second, signal-processing and statistical-monitoring approaches are used to identify disturbances, changes in operating regimes, and localized fault signatures [6,7]. Third, reconstruction-based methods, including PCA-based reconstruction and related latent-space models, compare measured behavior with reconstructed nominal behavior and use reconstruction errors as anomaly indicators. Fourth, supervised classifiers and ensemble methods are used when labelled fault or anomaly classes are available [23–26].

These categories differ in their assumptions and validation requirements. Unsupervised or weakly supervised methods are appropriate when only nominal data are reliably available, but their thresholds must be calibrated carefully and their false-alarm behavior must be reported. Supervised classifiers can provide strong discrimination when labels are available, but their performance depends on the representativeness of the labelled data. Therefore, score-level Gradient Boosting trained using evaluation labels corresponding to controlled perturbation intervals should be interpreted separately from unsupervised PCA reconstruction, residual scoring, and robust multivariate monitoring. In this manuscript, score-level Gradient Boosting is retained as a score-level classification baseline rather than as a directly equivalent unsupervised anomaly detector.

Although many machine learning methods report promising results on task-specific datasets, several limitations remain. Models trained offline may degrade under changing operating conditions, sensor drift, long-term ageing, or previously unseen transients. In addition, comparisons across studies are difficult when datasets, anomaly definitions, variable sets, sampling intervals, and evaluation metrics differ.

2.3. Digital Twins, Adaptive Updating, and Physics-Informed Modelling

Digital twins provide a conceptual and computational basis for linking measurements, simulation models, data-driven algorithms, and operator feedback within a unified monitoring environment. In nuclear engineering, digital twin concepts have been discussed as solutions for condition monitoring, in-service inspection, predictive maintenance, safety analysis, and operational support [1–3,27]. Recent reviews also emphasize that digital twins for nuclear systems require validated data pipelines, uncertainty handling, traceability, and integration with domain knowledge rather than isolated predictive models.

Recent review studies have further emphasized the growing integration of digital twins with artificial intelligence and machine learning for condition monitoring, anomaly detection, fault classification, virtual sensing, and predictive maintenance in nuclear power systems. These studies also highlight the importance of validation, qualification, and regulatory acceptance for AI-enabled digital twin technologies in safety-critical nuclear applications [27].

Hybrid and physics-informed approaches are particularly relevant because they can preserve physical consistency while reducing the computational cost of detailed simulations. Surrogate models, neural operators, and physics-informed neural networks have been used to approximate complex physical processes, support near-real-time inference, and incorporate governing equations or constraints into the modelling process [8–13]. These approaches are valuable for reactor monitoring because they can reduce the risk of purely data-driven extrapolation outside the training domain. However, their practical deployment still requires careful validation against reactor-specific operating regimes and uncertainty sources. Similar adaptive digital twin concepts have also been investigated in gas turbine and aero-engine monitoring, where model updating and performance-fault warning are integrated within operational digital twins [28]. Recent digital twin studies have also investigated real-time operational monitoring and online state correction in nuclear-energy-related systems, demonstrating the potential of adaptive monitoring workflows that combine simulation models with continuously acquired operational data [29].

Cross-domain digital twin studies in power and propulsion systems provide useful methodological analogies for adaptive updating and early warning. For example, the gas turbine digital twin model uses measured data to improve performance-model accuracy and applies self-learned warning thresholds for performance fault warning [28]. Although gas turbines and nuclear reactors differ substantially in physics, safety requirements, and regulatory contexts, this type of work illustrates how model updating and threshold learning can be implemented in an engineering digital twin workflow.

2.4. Uncertainty Estimation, Explainability, and Operator Interpretation

In safety-critical environments, anomaly scores or predicted labels are insufficient without information about confidence, traceability, and interpretation. Uncertainty estimation methods, including Monte Carlo dropout and ensemble-based techniques, can provide confidence estimates or predictive variance, together with model outputs [14,30]. Such information is particularly important when measurements are incomplete, operating states are underrepresented in training data, or the system experiences distribution shifts.

Explainable AI methods, including SHapley Additive exPlanations (SHAP) and Local Interpretable Model-Agnostic Explanations (LIME), can help identify variables and temporal features that contribute most strongly to a warning or classification result [31,32]. In nuclear applications, explainability is connected not only to operator trust but also to auditability, review of abnormal-event hypotheses, and regulatory acceptability [15,33,34]. However, many studies apply explainability only after model training and prediction. This limits real-time usefulness when explanations, uncertainty estimates, and anomaly scores should be available together for operator review.

Uncertainty estimation and explainability are therefore regarded as important components of advanced reactor-monitoring workflows. Their integration can improve confidence assessment, transparency of analytical decisions, and operator interpretation of monitoring results, particularly under changing operating conditions and incomplete information.

2.5. Comparison Criteria and Positioning of the Present Study

Table 1 presents representative research directions related to reactor process analysis, anomaly detection, and digital twin applications in nuclear engineering. The reviewed studies demonstrate the increasing use of machine learning methods, hybrid models, physics-informed approaches, and uncertainty-estimation techniques for reactor monitoring, operational-state prediction, anomaly detection, and diagnostic support.

Table 1. Comparison of representative AI-based and digital twin approaches for reactor process analysis.

Approach/Research Direction	Main Contribution	Adaptation	Uncertainty	Explainability	Main Limitation/Relevance to This Study
Conventional online monitoring and condition-based maintenance [16–19]	Operational monitoring, diagnostics, and maintenance support based on plant signals and engineering rules	Limited/plant-procedure dependent	Limited or external	Moderate through engineering interpretation	Robust reference basis, but limited detection of correlated multivariate deviations before thresholds are exceeded
Temporal ML models for reactor signals [4,5]	Learning time-dependent patterns and forecasting abnormal trends	Limited after offline training	Usually not integrated	Limited	Useful for sequence modelling, but may degrade under drift or unseen transients
Signal-processing and statistical diagnostics [6,7]	Detection of disturbances and abnormal signal behavior	Limited	Usually not integrated	Moderate	Often focused on specific signals, disturbances, or fault types
Reconstruction-based monitoring, including PCA-type methods [23,24]	Detection of deviations from nominal multivariate behavior through reconstruction error	Limited	External threshold-based	Moderate	Relevant to the implemented analytical layer; threshold and false-alarm behavior require careful calibration
Supervised/ensemble classifiers [25,26]	Fault or anomaly classification when labelled data are available	Limited unless retrained	Partial in some ensembles	Moderate	Should be distinguished from unsupervised methods; label representativeness is critical
SHAP/LIME-based explainable anomaly analysis [31–34]	Post hoc or integrated interpretation of model outputs	None or limited	Not primary objective	High	Improves traceability, but often added after prediction rather than embedded in real-time workflow
Hybrid digital twins and surrogate models [1–3,8–12,25,26]	Integration of measurements, simulations, and fast surrogate inference	Partial	Limited/partial	Partial	Often validated for simulation acceleration or specific use cases rather than closed-loop monitoring
Physics-informed neural networks [13]	Embedding differential equations or physical constraints in learning	Partial	Limited/partial	Moderate	Promising for physical consistency, but scaling and reactor-specific validation remain challenging
Cross-domain adaptive digital twins [28,29]	Model updating and self-learned warning thresholds in engineering digital twin workflows	Partial to high in the specific application	Limited/partial	Moderate	Useful methodological analogy; physics and safety constraints differ from nuclear reactors
Present study	Integrated reactor-monitoring architecture with implemented comparative analytical anomaly-detection layer	Architectural support; limited proof-of-concept validation	Architectural support; not experimentally validated	Architectural support; not experimentally validated	Analytical layer and a limited adaptive recalibration proof-of-concept are validated; broader framework capabilities remain subject to future implementation and validation

To improve consistency of interpretation, the qualitative ratings used in Table 1 are defined as follows. “None” indicates that a capability is not discussed or supported in the cited work. “Limited” indicates that the capability is mentioned or indirectly sup-

ported but not integrated into the primary workflow. “Partial” indicates implementation within a specific model, dataset, or offline scenario without comprehensive operational validation. “High” indicates that the capability is a central implemented component of the approach. “Architectural support” indicates that the capability is defined within the proposed framework architecture. “Validated in analytical layer” indicates that the capability is implemented and experimentally evaluated within the analytical monitoring layer.

Table 1 provides a structured comparison of representative research directions related to reactor monitoring, anomaly detection, and digital twin applications in nuclear engineering. The reviewed studies demonstrate growing interest in adaptive monitoring, uncertainty-aware analysis, explainable AI, and physics-informed modelling, while also highlighting the challenges associated with their integration into unified reactor-monitoring frameworks. Based on the reviewed literature, the proposed framework combines multivariate analytical monitoring with adaptive recalibration, uncertainty estimation, explainability support, and physics-informed integration at the architectural level.

3. Digital Twin Framework for Nuclear Reactor Monitoring

3.1. System Design

The proposed framework defines a modular architecture for nuclear reactor monitoring organized around several complementary functional layers, including multivariate analytical monitoring, adaptive recalibration support, uncertainty-aware analysis, explainability support, and operator-oriented monitoring integration. The architecture is designed to integrate these capabilities within a unified monitoring environment that supports reactor condition assessment, anomaly detection, and operational decision support.

The core principle of the framework is the integration of data acquisition, reactor process analysis, and operational feedback. This architectural organization is intended to support integration of data acquisition, analytical monitoring, and operational feedback within a unified reactor-monitoring workflow.

3.2. Data Acquisition and Preprocessing

The first module of the framework is intended for heterogeneous data acquisition and preprocessing from various data sources, including real-time sensor measurements, Supervisory Control and Data Acquisition (SCADA) system outputs, maintenance logs, nondestructive testing data, and simulation-generated datasets.

Due to the multimodal nature of nuclear reactor data, preprocessing plays an important role in ensuring stable and reliable data analysis. The preprocessing stage includes noise reduction, normalization, feature extraction, missing-data handling, temporal synchronization, and physics-informed preprocessing, transforming the data into features suitable for subsequent multivariate analysis.

The preprocessing procedures are designed to reduce the effects of sensor noise, operational variability, and incomplete measurements, which are characteristic of the real-world environments of reactor process analysis and condition assessment.

3.3. Hybrid Modeling Layer

At the core of the framework is a hybrid modeling layer intended to support integration between data-driven analytical approaches and physics-informed extensions for reactor process analysis. The layer is designed to support several complementary strategies for reactor process analysis and operational state assessment, including

- Temporal models for analysis of reactor time-series sequences, anomaly detection, and operational deviation analysis;

- Reconstruction-based and statistical methods for condition assessment and deviation analysis relative to nominal behavior;
- Physics-informed extensions for aligning model outputs with reactor physics and simulation constraints.

This organization enables condition assessment, temporal analysis, and incorporation of physical constraints within a unified architecture for reactor process analysis and operator support. Within the proposed framework, the physics-informed integration module is intended to operate by comparing analytical monitoring outputs with reactor-physics constraints and simulation-based knowledge in order to identify potential inconsistencies before subsequent adaptive recalibration and decision-support processes. This staged interaction defines the intended role of the conceptual module within the overall architecture and provides a technical roadmap for its progressive integration with the implemented analytical monitoring layer.

3.4. Adaptive Updating and Recalibration

One of the main limitations of conventional digital twin systems is their predominantly static behavior after deployment. To address this limitation, the proposed framework includes mechanisms for model updating and recalibration using newly acquired operational data. Rather than relying on periodic full retraining, the architecture supports adaptive recalibration based on newly acquired reactor measurements and validated analytical results. This design enables adaptation to

- Reactor aging effects;
- Sensor drift;
- Changing operating conditions;
- Previously unseen transient states.

This mechanism is intended to support the long-term robustness of the framework and to reduce performance degradation under changing operating conditions.

3.5. Uncertainty Estimation and Explainability

The proposed architecture includes dedicated modules for uncertainty estimation and explainability analysis within the monitoring workflow. The architecture allows integration of probabilistic or ensemble-based uncertainty-estimation mechanisms capable of providing confidence estimates together with model outputs. This enables operators to assess not only the predicted reactor condition but also the reliability of the prediction itself. In parallel, explainability mechanisms are considered through interpretable AI techniques that identify the variables and subsystem interactions contributing most significantly to model outputs and anomaly detection results.

Unlike conventional post-processing interpretation approaches, the architecture considers integration of uncertainty estimation and explainability analysis directly within the analytical workflow. Such integration may improve transparency and support operator interpretation and decision-making in safety-critical operating environments.

3.6. Decision Support and Operational Integration

The final layer of the framework is intended to transform analytical outputs into actionable information for reactor condition assessment and operator support, including

- Anomaly detection and warning generation;
- Predictive maintenance recommendations;
- Reactor state monitoring and operational deviation analysis;
- Operator-support functionalities.

The framework is organized as a modular analytical architecture intended to support integration of validated analytical results and operator feedback into monitoring and recalibration workflows. This feedback mechanism supports the continuous refinement of anomaly detection thresholds, updating procedures, and reactor condition assessment models. This approach is intended primarily as an operator-support and reactor process analysis framework rather than a fully autonomous reactor control mechanism. Therefore, human oversight remains an essential element for safe decision-making and regulatory accountability.

3.7. Architectural Perspective and System Integration

Figure 1 presents the high-level architecture of the proposed digital twin framework for nuclear reactor monitoring and illustrates the main architectural layers, including multi-source data acquisition, preprocessing and feature engineering, reactor process monitoring, and decision-support and operational integration, as well as conceptual adaptive extensions related to uncertainty estimation, recalibration, and physics-informed integration. The presented architecture is intended as a general reference framework for digital twins in nuclear engineering rather than a reactor-specific implementation. The framework provides a basis for integrating analytical monitoring, adaptive recalibration, uncertainty-aware analysis, explainability support, and future physics-informed extensions within a unified reactor-monitoring architecture.

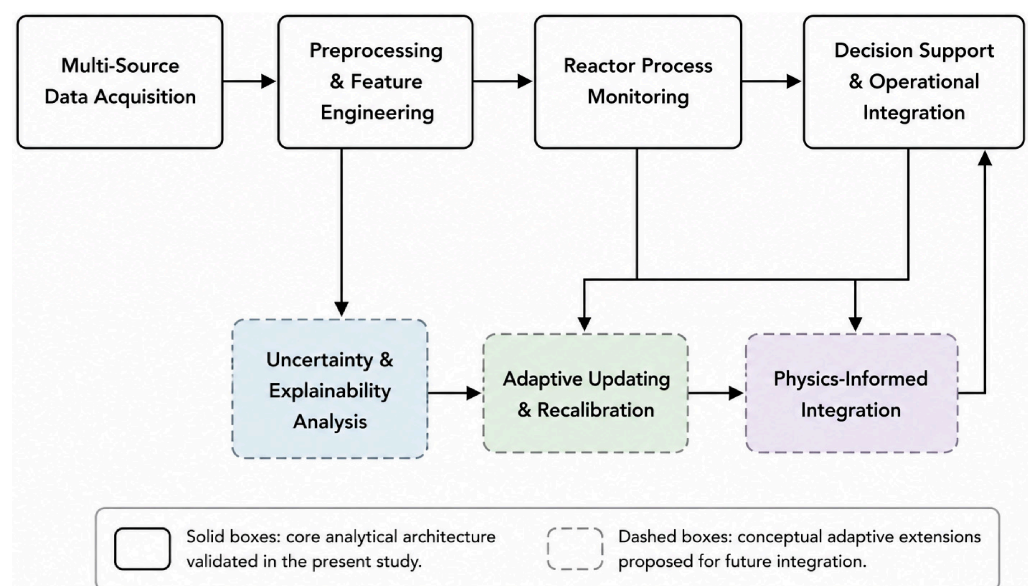


Figure 1. Conceptual digital twin architecture for nuclear reactor analysis and monitoring.

4. Architectural Integration and Technical Considerations

4.1. Comparative Architectural Positioning

Existing approaches for reactor process analysis and monitoring can generally be divided into conventional machine learning models, static digital twins, and hybrid physics-informed analytical approaches. Although these approaches can support anomaly detection, simulation-based analysis, and predictive modeling, limitations remain with respect to long-term condition assessment, changing operating conditions, sensor drift, and previously unseen transient states.

Conventional machine learning models are typically designed for specific tasks such as anomaly detection, fault classification, or signal prediction. They can provide good accuracy under stable operating conditions but often require periodic offline retraining and show limited robustness under distribution shifts in operational data.

Static digital twins provide better physical consistency through the use of simulation models and engineering representations of the reactor. They are suitable for analysis, operator training, and offline studies but usually operate as periodically updated simulation systems with limited adaptability.

Hybrid physics-informed approaches combine simulation-based relationships with machine learning methods and can reduce computational costs in the modeling of thermal-hydraulic and neutron-physics processes. However, many of these approaches remain primarily focused on simulation acceleration or offline anomaly detection rather than long-term condition assessment and reactor process analysis.

The proposed framework defines mechanisms for adaptive recalibration, uncertainty estimation, explainability, operator feedback, and integration of physical constraints. A comparison of the corresponding functional characteristics is presented in Table 2.

Table 2. Comparison of functional characteristics of reactor process analysis and monitoring approaches.

Functional Characteristic	Conventional ML Models	Static Digital Twins	Hybrid Physics-Informed Models	Proposed Framework
Primary objective	Task-specific prediction	Simulation-based representation	Physics-constrained prediction	Adaptive analysis and condition assessment
Multivariate time-series analysis	Partially supported	Limited	Partially supported	Validated in the analytical layer
Reconstruction-based analysis	Limited use	Limited use	Partially supported	Included in the analytical layer
Adaptive updating	Limited	Usually absent	Partially supported	Architecturally supported
Uncertainty estimation	Post-processing or external	Limited	Partially supported	Architecturally integrated
Explainability	Limited	Indirect	Partially supported	Architecturally integrated
Operator feedback integration	Rarely used	Limited	Limited	Included in the architecture
Real-time operation	Task-dependent	Limited	Limited	Requires additional validation

4.2. Expected Architectural Benefits for Nuclear Reactor Monitoring

The proposed digital twin framework for reactor monitoring can support long-term reactor process analysis by combining continuously acquired measurements, predictive models, and adaptive recalibration mechanisms. One potential benefit is earlier identification of deviations associated with changing operating conditions, sensor drift, or gradual equipment degradation.

Predictive condition assessment may support identification of developing operational trends before conventional alarm thresholds are exceeded. Reconstruction-based analytical methods provide additional capability for detecting complex multivariate deviations that may be difficult to identify through isolated evaluation of individual process variables.

Uncertainty estimation can provide operators with additional information regarding the reliability of generated warnings and analytical outputs, particularly in the presence of incomplete measurements, previously unseen operating states, or limited data availability. Explainability mechanisms can further support interpretation of analytical results by identifying variables and temporal relationships that contribute most strongly to detected deviations or anomaly indications.

More broadly, multivariate analytical approaches can support monitoring of reactor systems, auxiliary systems, and critical equipment through analysis of relationships among multiple process variables. Such analysis may assist in identifying gradual changes

in process behavior that are not evident from individual measurements alone and can complement existing monitoring and condition-assessment practices.

4.3. Regulatory and Engineering Considerations

Deployment of digital twins and machine learning-based analytical systems in nuclear engineering requires rigorous verification, validation, traceability, and documentation. These requirements are particularly important for adaptive systems, in which model characteristics may change after updating with newly acquired operational data. Comprehensive evaluation of the proposed framework should include

- Simulation-based validation;
- Comparative evaluation using public and synthetic datasets;
- Testing under sensor drift and changing operating conditions;
- Assessment of prediction reliability;
- Verification of interpretability and traceability of analytical outputs.

The proposed architecture is intended for operator support rather than autonomous reactor control. Therefore, human oversight remains necessary for warning interpretation, risk assessment, and final operational decision-making.

In practical deployment, such an analytical monitoring system derived from the proposed framework would operate as an additional analytical module integrated with existing monitoring and operator-support systems, without direct participation in reactor control functions. The analytical models could process continuously acquired monitoring data from existing instrumentation and SCADA systems and provide early warnings, deviation assessments, and information regarding changes in process relationships.

In such a configuration, final interpretation of analytical results and operational decisions would remain under the responsibility of operators and engineering personnel. This is consistent with the conservative approach typically adopted for deployment of artificial-intelligence-based technologies in safety-critical environments such as nuclear energy systems.

Although the proposed framework defines five complementary evaluation dimensions for comprehensive digital twin assessment, the present study experimentally validates primarily the analytical monitoring layer, under controlled multivariate perturbation scenarios, using the publicly available PWR Abnormality Dataset. Validation involving physics-based reactor simulations, extended operational regimes (including startup, shutdown, and transient conditions), and long-term operational deployment is considered the next stage of the proposed research and therefore remains outside the scope of the present work.

4.4. Computational Implementation Considerations

The implemented analytical monitoring workflow requires consideration of computational complexity and latency during multivariate reactor-monitoring analysis. In practical monitoring environments, continuous execution of computationally intensive simulation models for every newly acquired measurement may not be feasible.

For this reason, the proposed architecture should separate monitoring tasks into

- Fast, continuously acquired data analysis;
- Local anomaly detection;
- Periodic or event-driven recalibration;
- Computationally intensive offline simulations and retraining.

The implemented analytical workflow relies on computationally lightweight statistical and machine-learning methods executed using standard scientific-computing libraries. The

reported experiments were performed in a controlled offline evaluation setting and did not require specialized computational infrastructure.

Additional validation remains necessary with respect to computational latency, robustness under incomplete measurements, and long-term operational stability before practical deployment in safety-critical reactor-monitoring environments.

4.5. Broader Evaluation and Integration Considerations

Evaluation of digital twin frameworks should extend beyond predictive accuracy alone. In addition to anomaly-detection performance, comprehensive assessment requires consideration of adaptability, uncertainty estimation, explainability, computational efficiency, and deployment-related robustness. Representative prediction-quality metrics include root mean square error (RMSE), mean absolute error (MAE), and mean squared error (MSE), while anomaly-detection performance is commonly assessed using metrics such as Precision, Recall, F1-score, and ROC-AUC. Table 3 summarizes representative evaluation categories and corresponding metrics relevant to adaptive digital twin-based reactor-monitoring systems. Future evaluation should include scenarios involving sensor degradation, incomplete measurements, transient operating conditions, changing operating regimes, and previously unseen abnormal states.

Table 3. Representative evaluation metrics for adaptive digital twin-based reactor monitoring.

Category	Example Metrics	Significance
Prediction	RMSE, MAE, MSE	Prediction quality
Anomaly Detection	Precision, Recall, F1-score, ROC-AUC	Reliability of anomaly detection
Adaptation	Adaptation rate under drift, recalibration time	Robustness under changing data distributions
Uncertainty	Calibration error, predictive entropy	Reliability of confidence estimates
Explainability	Stability of feature attribution, operator interpretability	Transparency
Computational Efficiency	Inference latency, memory usage	Real-time applicability
Robustness	Performance under sensor drift, noise, or missing data	Reliability under real operating conditions

5. Experimental Scope and Validation Strategy

The experimental part of the present study focuses on implementation and validation of the data-driven analytical monitoring module within the proposed reactor-monitoring architecture. The evaluated module includes temporal residual analysis, robust multivariate monitoring, PCA-based reconstruction, anomaly-score fusion, and score-level classification under controlled multivariate monitoring conditions.

Accordingly, the experimental study presented in this section focuses on validation of the implemented analytical monitoring layer, while the remaining evaluation dimensions defined in Section 4.3 are discussed as future stages of framework validation.

The adaptive recalibration, uncertainty estimation, explainability support, feedback mechanisms, and physics-informed integration components are retained primarily as defined modules of the architecture. Within the present study, the main experimental validation focuses on the analytical monitoring module. In addition, a limited proof-of-concept adaptive threshold recalibration experiment is included to illustrate one possible mechanism under simulated operating drift.

The selected analytical approaches were chosen in order to represent different principles for reactor-process monitoring and anomaly detection under multivariate operating conditions. Temporal residual analysis evaluates short-term deviations in sequential reactor behavior, robust multivariate monitoring evaluates statistical deviations from normal

operating conditions, and PCA-based reconstruction evaluates changes in correlated multivariate process structure through reconstruction residual analysis.

PCA-based reconstruction was selected because of its interpretability, suitability for correlated multivariate reactor-process variables, and relatively low computational requirements for comparative monitoring evaluation. More complex nonlinear reconstruction approaches, such as deep autoencoder architectures, are not included in the present study because they would require additional architecture optimization and validation under broader operating conditions.

Score-level Gradient Boosting was selected as a supervised classification stage operating on anomaly-related scores generated by the evaluated analytical approaches. Isolation Forest was additionally included as a conventional unsupervised anomaly-detection baseline for comparative evaluation.

5.1. Objectives of the Experimental Validation

The objective of the experimental validation is to evaluate various anomaly detection and reactor process analysis approaches that can be integrated into the analytical layer of the proposed digital twin architecture. The evaluation is conducted using multivariate time-series data from a pressurized water reactor (PWR).

The experiments conducted include:

- Statistical anomaly detection;
- Temporal sequence analysis;
- Reconstruction-based deviation analysis;
- Fusion of results from multiple analytical approaches;
- Discrimination between normal and abnormal operating conditions through multivariate monitoring.

Comparative evaluation is performed using Precision, Recall, F1-score, Receiver Operating Characteristic Area Under the Curve (ROC-AUC), and Precision–Recall Area Under the Curve (PR-AUC) metrics.

5.2. Dataset Description

The experimental validation is performed using a multivariate reactor-monitoring dataset derived from the publicly available PWR Abnormality Dataset [35], which provides baseline reactor sensor measurements acquired under controlled operating conditions. In the present study, the dataset serves as a source of baseline monitoring signals and is subsequently processed through feature extraction, temporal sequence generation, and controlled synthetic perturbation procedures to create a reproducible environment for comparative evaluation of the analytical monitoring approaches. Evaluation labels are assigned to the perturbed temporal intervals for comparative anomaly-detection analysis. Although the publicly available dataset includes anomaly labels, the present study employs controlled synthetic perturbations together with corresponding evaluation labels in order to establish a consistent and reproducible benchmark for comparative assessment of all implemented analytical monitoring approaches under identical disturbance conditions. Publicly available reactor datasets containing fully annotated abnormal operational transients remain limited because of safety, confidentiality, and regulatory restrictions associated with nuclear operational data. The available variables include parameters associated with:

- Thermal behavior;
- Pressure conditions;
- Flow-related reactor process behavior;
- Operational reactor dynamics.

The utilization of all available numerical variables preserves temporal and multivariate dependencies among correlated reactor signals during analytical monitoring evaluation. Unlike simplified univariate benchmark scenarios, the utilized dataset supports comparative evaluation of reactor-monitoring and anomaly-detection approaches under controlled multivariate operating conditions.

The publicly available dataset documentation does not explicitly report the physical acquisition sampling interval of the reactor-monitoring measurements. Consequently, the present study focuses primarily on preservation of temporal ordering and multivariate temporal dependencies during comparative evaluation. The experimental analysis is therefore performed using sliding windows consisting of 20 consecutive temporally ordered samples.

The use of temporally correlated reactor measurements enables assessment of analytical sensitivity to coupled process variations, while the controlled perturbation strategy described in Section 5.4 provides a reproducible basis for comparative evaluation of the implemented monitoring approaches.

After preprocessing, the dataset contains approximately 12,266 reactor-monitoring samples and 16 numerical reactor-related variables. The reported dataset dimensions correspond to the processed analytical dataset obtained after preprocessing, numerical feature extraction, and temporal sequence generation. All available numerical variables are included in the experiments in order to preserve dependencies among correlated reactor signals and subsystem behaviors. The preprocessing workflow includes removal of incomplete records; extraction of numerical process variables; temporal partitioning into training, validation, and testing subsets; feature normalization; and sliding-window sequence generation for temporal analysis. Temporal partitioning is performed prior to window generation in order to avoid information leakage between the evaluation subsets.

After temporal splitting and sliding-window sequence generation the dataset contains approximately

- 8567 training sequences;
- 1821 validation sequences;
- 1821 test sequences.

A temporal window length of 20 time steps is used. Controlled synthetic perturbations are introduced into approximately 12% of the validation and test sequences in order to support controlled and reproducible comparative evaluation. The generated perturbations include drift, bias, noise, localized deviations, and combined multivariate anomalies affecting one or more process variables. These perturbations are intended primarily as analytical test scenarios and do not fully reproduce all physical coupling relationships characteristic of real reactor fault transients.

The utilized publicly available PWR dataset primarily represents multivariate monitoring measurements obtained under limited operating conditions and does not include fully validated startup/shutdown procedures, long-term sensor drift behavior, or broad sets of physically validated transient operating scenarios. Consequently, the present study focuses primarily on comparative evaluation under controlled perturbation conditions rather than comprehensive operational reactor validation across multiple power regimes.

The main parameters of the dataset and experimental configuration used for comparative evaluation of the monitoring layer under controlled multivariate reactor-monitoring conditions are summarized in Table 4. Representative categories of reactor-related variables, together with their analytical interpretation, are summarized in Table 5, while the numerical variables extracted from the experimental dataset are additionally presented in Table 6.

Table 4. Dataset and experimental configuration.

Parameter	Value
Total reactor-monitoring samples	12,266
Number of reactor-related variables	16
Training split	70%
Validation split	15%
Test split	15%
Sliding-window length	20
Training sequences	8567
Validation sequences	1821
Test sequences	1821
Synthetic anomaly ratio	12%

Table 5. Representative categories of reactor-related variables used in the experiments.

Variable Category	Representative Interpretation	Representative Analytical Role
Thermal variables	Reactor thermal and temperature-related behavior	Detection of thermal deviations and temperature-related anomalies
Pressure-related variables	Reactor pressure-related behavior	Monitoring of pressure stability and transient deviations
Flow-related variables	Flow-related reactor process and operational behavior	Detection of flow-related process deviations
Operational variables	Reactor operational and subsystem dynamic behavior	Multivariate operational-state assessment
Derived analytical variables	Reconstruction residuals and anomaly-related indicators	Comparative anomaly-score generation and supervisory fusion

Table 6. Numerical variables included in the experimental dataset.

Variable	Representative Interpretation
Temperature	Reactor thermal behavior
Pressure	Reactor pressure-related behavior
Flow1	Flow-related reactor process measurement
Flow2	Flow-related reactor process measurement
VRR12–VRR33	Reactor response-related monitoring variables
VRS01–VRS31	Reactor subsystem monitoring variables
VRI01–VRI03	Reactor instrumentation-related variables

5.3. Preprocessing and Experimental Procedure

Before training and evaluation, the dataset is processed using a sequential procedure designed to ensure reproducibility and numerical stability.

The main processing stages include

- Removal of incomplete or invalid measurements;
- Extraction of numerical reactor-related variables;
- Normalization using StandardScaler;
- Temporal splitting into training, validation, and test subsets;
- Generation of time-series sequences using a sliding window;
- Controlled synthetic perturbation generation;

- Normalization of anomaly scores;
- Threshold optimization on the validation subset;
- Temporal smoothing of anomaly scores;
- Fusion of anomaly scores from multiple analytical approaches.

The dataset is divided into 70% training, 15% validation, and 15% test subsets. Temporal splitting rather than random partitioning is used in order to avoid information leakage between past and future reactor operating states. Normalization statistics were computed exclusively using the training subset and subsequently applied to the validation and test subsets in order to avoid information leakage during comparative evaluation. For experiments based on temporal sequence analysis and reconstruction-based analysis, sequence lengths of 20 time steps are used.

The overall experimental workflow is presented in Figure 2 and includes preprocessing of multivariate reactor data, generation of temporal sequences, controlled synthetic perturbation generation, application of analytical monitoring models, anomaly score normalization, threshold optimization, and comparative evaluation.

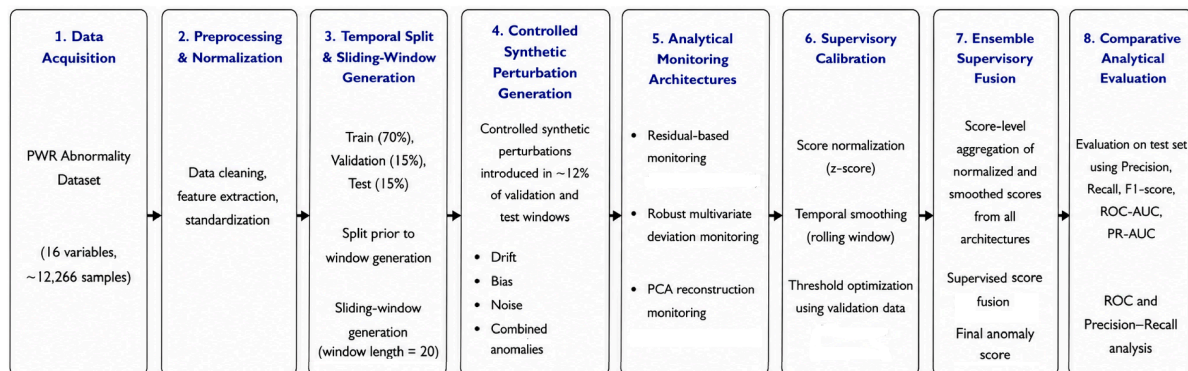


Figure 2. Experimental workflow for validation of the analytical monitoring layer.

All experiments were implemented in Python 3.11, using standard scientific computing and machine learning libraries. For PCA-based reconstruction, 95% of the cumulative explained variance is retained. Temporal smoothing is implemented using moving-average filtering applied to consecutive anomaly scores. Threshold optimization is performed on the validation subset through F1-score maximization. Synthetic anomalies are introduced after generation of temporal sequences in order to preserve the original temporal structure of reactor measurements. The generated perturbations include short-duration spikes, gradual drift, signal-level shifts, noise disturbances, and combined multivariate process deviations applied across consecutive temporal windows. Each perturbation type is applied independently to selected temporal windows affecting one or more process variables while preserving the original temporal ordering of the reactor-monitoring sequences.

The experimental procedure is implemented using a fixed processing pipeline and fixed random seeds in order to improve reproducibility and consistency of the comparative evaluation. To additionally evaluate robustness with respect to perturbation generation and evaluation variability, the comparative experiments were repeated using multiple random seeds controlling perturbation initialization and experimental evaluation procedures. The reported comparative metrics correspond to the mean values obtained across five repeated experimental runs.

Figure 3 presents representative reactor-monitoring variables, using their original dataset signal scales rather than normalized values in order to improve interpretability of temperature-, pressure-, and flow-related temporal patterns.

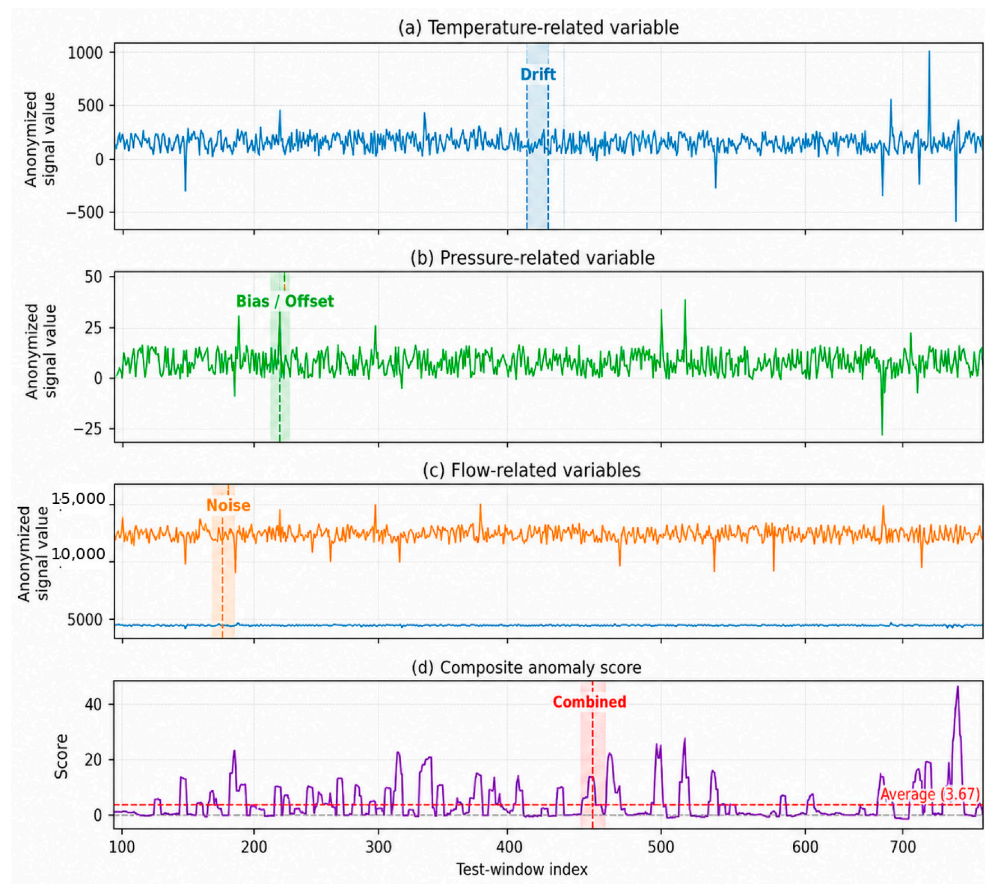


Figure 3. Representative reactor-monitoring variables presented using their original anonymized dataset signal scales together with representative controlled perturbation examples, and the corresponding composite anomaly score obtained during comparative evaluation.

The presented trajectories correspond to selected segments of the processed experimental dataset derived from the publicly available PWR Abnormality Dataset. The displayed trajectories illustrate representative temporal variability and anomaly-score responses observed under the controlled perturbation scenarios used for comparative evaluation. Subfigure (a) presents a representative temperature-related variable, while subfigures (b) and (c) illustrate representative pressure-related and flow-related variables, respectively. Subfigure (d) presents the composite anomaly score obtained through score fusion and supervisory threshold calibration during evaluation of the monitoring layer.

The presented variables correspond to anonymized reactor-monitoring signals extracted from the processed experimental dataset and should not be interpreted as absolute physical process measurements. Consequently, the displayed values represent original anonymized dataset signal scales rather than engineering units. The relatively large variations observed for the temperature-related variable reflect the characteristics of the anonymized dataset, together with the preprocessing and feature-extraction workflow, rather than absolute reactor coolant temperature measurements. The highlighted perturbation intervals illustrate representative analytical test scenarios generated using the implemented synthetic perturbation procedure. Perturbation magnitudes are randomly selected within the range of approximately 5–8 normalized standard deviations and applied over temporal windows of 6–16 consecutive samples while affecting one or more process variables and preserving the original temporal ordering of the evaluation sequences.

5.4. Synthetic Anomaly Generation

Due to the limited availability of publicly accessible datasets providing fully annotated abnormal-event labels suitable for the comparative evaluation protocol adopted in this study, controlled synthetic disturbances are introduced into the validation and test subsets. Although the utilized PWR Abnormality Dataset includes anomaly labels, controlled synthetic perturbations are employed to establish a consistent and reproducible comparative evaluation of all implemented analytical monitoring approaches under identical disturbance conditions. These perturbations provide a controlled experimental environment for systematic comparative evaluation of the implemented analytical monitoring approaches.

Several types of perturbations are considered:

- Short-duration spikes;
- Signal-level shifts;
- Gradual drift;
- Noise disturbances;
- Combined multivariate process deviations.

The anomalies are implemented as temporal segments spanning multiple consecutive windows rather than isolated single-point disturbances. Perturbation magnitudes are generated using randomized scaling factors within predefined ranges while preserving realistic relative variability of the affected variables. This approach enables evaluation under varying anomaly intensities while maintaining consistency with the underlying characteristics of the reactor-monitoring signals. This makes the experimental evaluation more representative of gradually developing deviations in reactor process data. The generated perturbations do not represent physically validated accident scenarios. They are used solely as controlled deviations for comparative evaluation of different approaches for reactor process analysis and anomaly detection.

5.5. Analytical Architectures for Anomaly Detection

Several complementary approaches for anomaly detection are included in the comparative evaluation. The implemented approaches differ in terms of their analytical principles and supervision requirements. The first four approaches operate in an unsupervised or reconstruction-based manner, whereas the score-level Gradient Boosting classifier represents a lightweight, supervised classification stage operating on normalized anomaly-related scores generated by the preceding analytical approaches.

5.5.1. Temporal Residual Baseline

A baseline approach based on short-term prediction residuals is implemented for anomaly detection.

The residual is defined as

$$e_t = |y_t - \hat{y}_t|$$

where

y_t is the observed reactor-state vector (normalized units);

$\hat{y}_t$ is the predicted reactor-state estimate (normalized units).

The resulting residual values are subsequently transformed into anomaly scores.

5.5.2. Robust Multivariate Monitoring

This approach evaluates deviations relative to the nominal multivariate distribution of reactor process variables. The resulting anomaly scores are calibrated and normalized using the validation subset prior to threshold selection.

5.5.3. PCA-Based Reconstruction

A reconstruction-based approach using Principal Component Analysis (PCA) is implemented. The anomaly score is estimated through the reconstruction error:

$$r_t = \|X_t - \hat{X}_t\|_2$$

where

X_t is the original multivariate reactor-state sequence (normalized units);

$\hat{X}_t$ is the reconstructed multivariate sequence (normalized units).

5.5.4. Unsupervised Anomaly Score Fusion

To exploit the complementary characteristics of the individual analytical approaches, normalized anomaly scores are temporally smoothed and combined into a unified anomaly score. The fused anomaly score is computed as the arithmetic mean of the normalized anomaly scores generated by the individual analytical approaches after temporal smoothing and validation-based normalization.

5.5.5. Score-Level Gradient Boosting

In addition, a lightweight gradient boosting classifier implemented using the GradientBoostingClassifier model from scikit-learn, with 80 estimators and a maximum tree depth of 2, is trained on the normalized outputs from the individual analytical approaches. The objective is to evaluate whether combining different anomaly-related scores improves discrimination between normal and abnormal operating conditions. The classifier is trained on the training subset using the corresponding perturbation labels and subsequently evaluated on the disjoint validation and test subsets, in accordance with the adopted temporal data partitioning.

5.5.6. Isolation Forest Baseline

As an additional conventional unsupervised anomaly-detection baseline, an Isolation Forest model is included in the comparative evaluation. The model is trained using flattened multivariate temporal sequences derived from the training subset and evaluated using the same preprocessing, normalization, temporal smoothing, and validation-threshold calibration procedures applied to the other analytical approaches.

5.6. Threshold Optimization and Calibration

Anomaly detection thresholds are determined through optimization on the validation subset rather than by using fixed percentile-based thresholds.

The procedure includes

- Calibration of anomaly scores on validation data;
- Temporal smoothing;
- Anomaly score normalization;
- Threshold selection through F1-score maximization.

Temporal smoothing reduces isolated false alarms and improves stability during sequential anomaly detection. This reflects practical monitoring requirements, in light of which an appropriate balance should be maintained between sensitivity to process deviations and robustness against false alarms.

5.7. Comparative Results

The evaluated approaches utilize different anomaly-detection and classification strategies, including reconstruction-based, residual-based, statistical, score-fusion, and supervised score-level classification procedures. The reported results correspond to mean values

obtained across five repeated experimental runs using different perturbation initialization seeds in order to improve robustness assessment of the comparative analytical evaluation. Table 7 therefore presents both the comparative analytical performance and the corresponding detection strategy associated with each evaluated monitoring approach.

Table 7. Comparative results for multivariate reactor-process anomaly detection (mean values across five repeated experimental runs).

Model	Detection Strategy	Precision	Recall	F1-Score	ROC-AUC	PR-AUC	Balanced Accuracy	False Alarm Rate
Temporal Residual Baseline	Residual-based anomaly scoring with validation-threshold calibration	0.222	0.330	0.266	0.593	0.215	0.586	0.157
Robust Multivariate Monitoring	Statistical anomaly scoring with validation-threshold calibration	0.218	0.752	0.338	0.722	0.331	0.693	0.367
PCA-Based Reconstruction	Reconstruction-based anomaly scoring with validation-threshold calibration	0.220	0.913	0.354	0.729	0.345	0.736	0.440
Unsupervised Anomaly Score Fusion	Unsupervised score fusion with validation-threshold calibration	0.221	0.835	0.349	0.731	0.350	0.717	0.401
Score-Level Gradient Boosting	Supervised score-level classification using synthetic labels	0.224	0.564	0.320	0.740	0.339	0.649	0.266
Isolation Forest		0.179	0.495	0.263	0.632	0.162	0.594	0.307

PCA-based reconstruction achieved the highest F1-score, indicating strong sensitivity to multivariate process deviations in reactor data. Score-level Gradient Boosting achieved the highest ROC-AUC, demonstrating the best overall capability for discriminating between normal and abnormal operating conditions. Unlike the preceding anomaly-scoring approaches based primarily on analytical deviation monitoring and validation-threshold calibration, the score-level Gradient Boosting model operates as a supervised classification stage using normalized anomaly-related scores generated by the monitoring pipeline. Isolation Forest was additionally included as a conventional unsupervised anomaly-detection baseline for comparative evaluation. Consequently, the reported performance of the score-level Gradient Boosting model should not be interpreted as directly equivalent to the unsupervised analytical approaches, but rather as a complementary supervised score-level classification stage within the proposed monitoring workflow.

The high Recall values obtained by PCA-based reconstruction and unsupervised anomaly score fusion indicate good sensitivity to abnormal conditions. At the same time, the reported False Alarm Rate and PR-AUC values further illustrate the trade-off between anomaly sensitivity and false-alarm reduction under imbalanced anomaly-detection conditions. The moderate Precision values reflect the conservative threshold-calibration strategy adopted in the experiments, which prioritizes sensitivity to abnormal conditions and early deviation detection at the expense of increased false-alarm rates. The obtained Precision values therefore correspond to a relatively high false-alarm fraction and should be interpreted within the context of comparative analytical evaluation rather than as an operationally optimized monitoring configuration. From an operational perspective, elevated false-alarm rates may increase operator workload, contribute to alarm fatigue, and reduce confidence in automated monitoring recommendations. Consequently, the conservative threshold configuration adopted in the present study should be interpreted as appropriate for comparative

analytical assessment rather than direct deployment in safety-critical reactor-monitoring environments. The repeated experiments indicate relatively stable comparative behavior of the evaluated analytical approaches under varying perturbation realizations.

5.8. ROC and Precision–Recall Analysis

Figure 4 presents the comparative ROC curves for the implemented analytical approaches. The ROC analysis indicates that PCA-based reconstruction, anomaly-score fusion, and score-level gradient boosting provide the strongest discrimination capability between normal and abnormal operating conditions, with the highest ROC-AUC achieved by the score-level Gradient Boosting classifier.

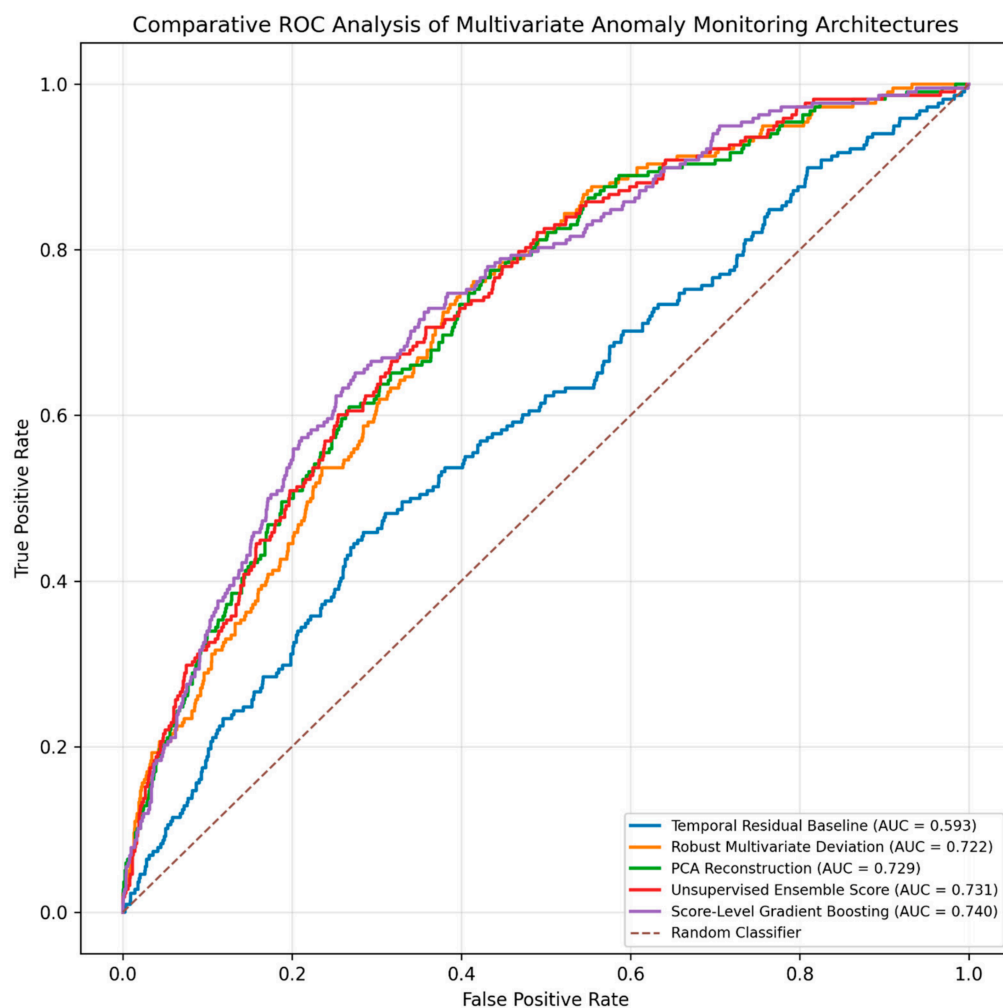


Figure 4. Comparative ROC analysis of the implemented anomaly detection approaches.

Figure 5 presents the Precision–Recall curves for the same models. The analysis further confirms the trade-off between anomaly sensitivity and false-alarm reduction in imbalanced anomaly detection tasks. The repeated experimental evaluation indicates generally consistent comparative Precision–Recall behavior across varying perturbation realizations.

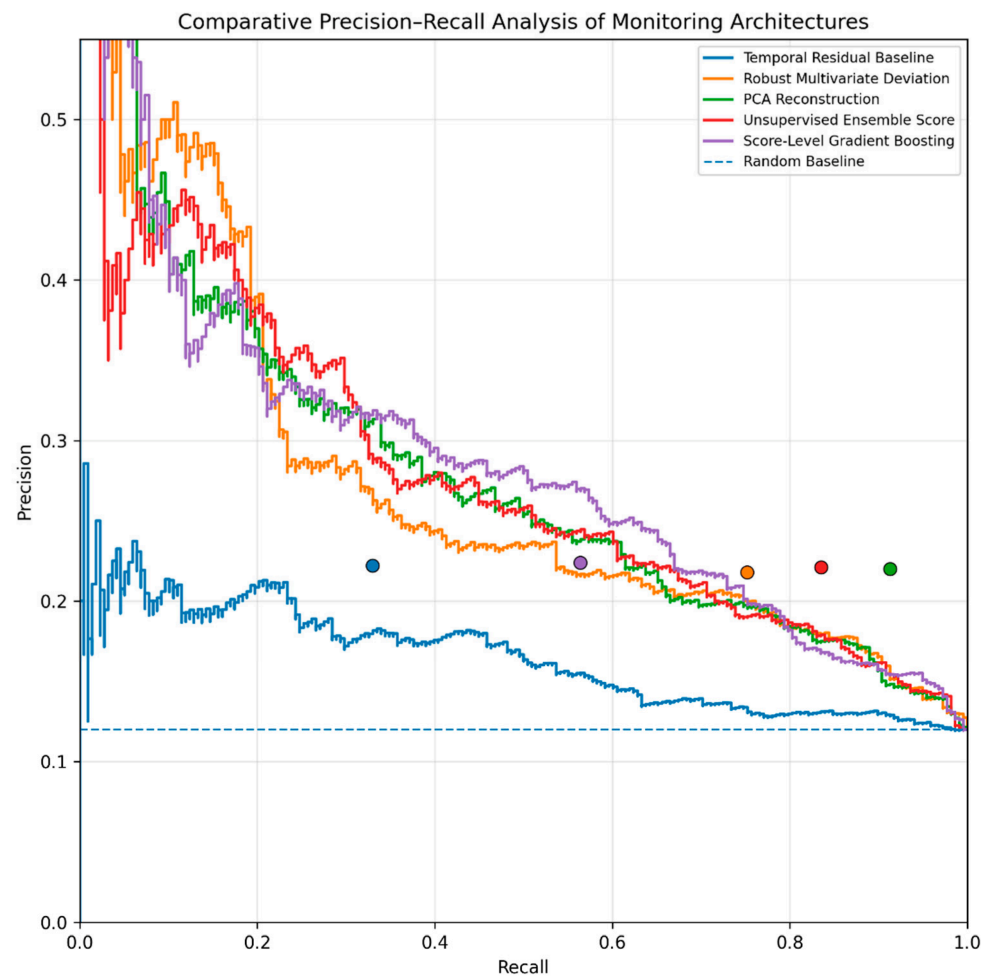


Figure 5. Comparative Precision–Recall analysis for the implemented models.

5.9. Adaptive Recalibration Under Simulated Operating Drift: A Proof-of-Concept Experiment

Although the experimental validation presented in the previous sections focuses primarily on the analytical monitoring layer, an additional proof-of-concept experiment was conducted to investigate one adaptive monitoring mechanism relevant to future digital twin deployment. The objective of this experiment was not to implement a complete digital twin, but rather to evaluate whether a simple recalibration procedure could improve monitoring robustness under changing operating conditions.

5.9.1. Experimental Setup

A gradual distribution shift was introduced into the monitoring data stream in order to emulate a change in nominal operating conditions. Unlike the synthetic anomaly scenarios used in the comparative benchmark evaluation, the simulated drift represented a persistent change in the statistical properties of selected reactor-monitoring variables without corresponding abnormal reactor behavior. Such effects may arise in practical monitoring environments because of sensor aging, calibration offsets, long-term operational evolution, or gradual changes in operating regimes.

The experiment was based on the PCA reconstruction-error monitoring approach, which achieved the highest F1-score among the evaluated analytical methods. Two monitoring configurations were considered:

- Static monitoring, where the anomaly-detection threshold remained fixed throughout the monitoring period;

- Adaptive monitoring, where the anomaly-detection threshold was recalibrated after detection of a statistically significant operating-point shift using a drift-verified nominal adaptation window.

The recalibration procedure was designed to emulate a future adaptive monitoring workflow in which newly acquired operational data indicates that an observed distribution shift corresponds to a modified nominal operating condition rather than an abnormal reactor state.

5.9.2. Performance Evaluation

Table 8 summarizes the results obtained for the static and adaptive monitoring configurations under the simulated operating-drift scenario. The False Alarm Rate of 1.000 observed for the static-threshold configuration represents a degenerate fixed-threshold limit, in which all post-drift sequences exceed the fixed decision threshold.

Table 8. Performance comparison between static and adaptive monitoring under simulated operating drift.

Monitoring Strategy	Precision	Recall	F1-Score	False Alarm Rate
Static Threshold	0.119	1.000	0.213	1.000
Adaptive Recalibration	0.189	0.902	0.312	0.525

Figure 6 illustrates the evolution of the anomaly score and adaptive threshold during the simulated operating-drift scenario. Following drift detection, threshold recalibration adapts the monitoring system to the new nominal operating regime, reducing the number of false-positive detections while maintaining anomaly sensitivity.

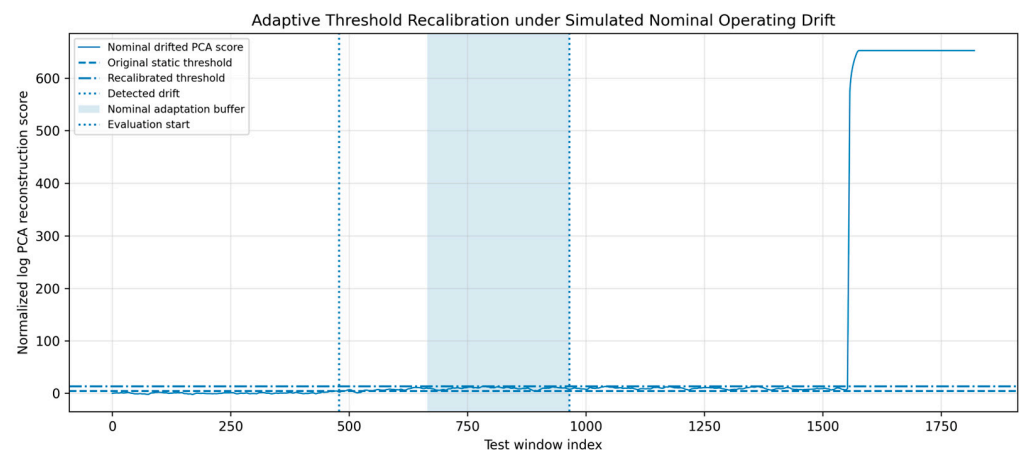


Figure 6. Adaptive threshold recalibration under simulated operating drift.

The comparison of the obtained monitoring performance of the static and adaptive configurations is shown in Figure 7. The static monitoring configuration exhibited extremely high sensitivity and successfully identified nearly all injected abnormal conditions. However, after the operating-point drift occurred, the fixed threshold generated a large number of false-positive detections, resulting in a False Alarm Rate of 1.000.

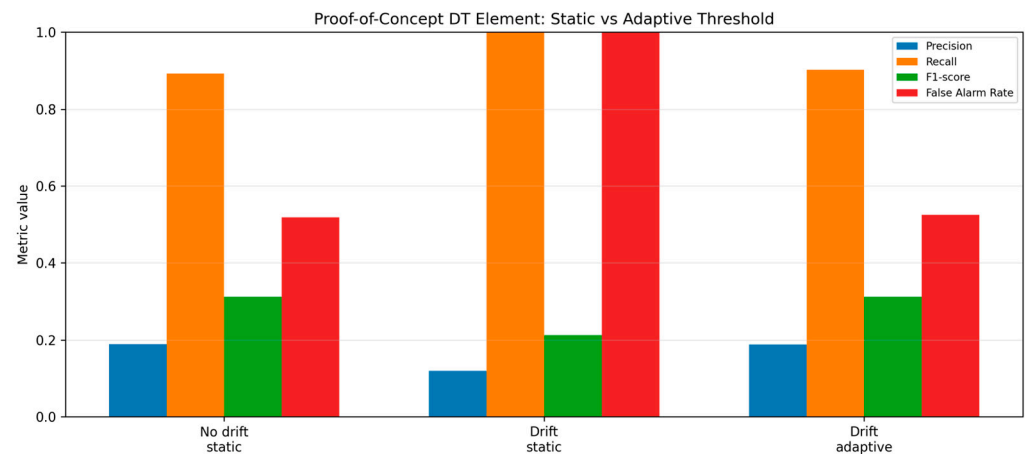


Figure 7. Comparison between static and adaptive monitoring performance.

The adaptive recalibration strategy reduced the False Alarm Rate from 1.000 to 0.525 and improved the F1-score from 0.213 to 0.312. Although a moderate decrease in Recall was observed, the resulting monitoring behavior exhibited a more balanced trade-off between sensitivity and false-positive generation.

5.9.3. Analysis

The obtained results indicate that even a relatively simple adaptive threshold-update mechanism can improve monitoring robustness when the statistical properties of reactor-monitoring data evolve over time. The reduction in false alarms and improvement in F1-score suggest that adaptive recalibration may improve monitoring robustness under changing operating conditions, including sensor aging effects, calibration shifts, and gradual operating-regime variations. The experiment demonstrates one possible adaptive recalibration mechanism that could support future digital twin applications for reactor monitoring. However, the presented results should not be interpreted as validation of the complete adaptive recalibration module defined in the proposed architecture. The implemented procedure is limited to threshold adaptation under simulated operating drift. Online state estimation, bidirectional synchronization, uncertainty-aware adaptation, explainability-assisted recalibration, and physics-informed model updating were not implemented and remain subjects for future research.

5.10. Interpretation of Comparative Analytical Results

The comparative evaluation demonstrates that complementary analytical and anomaly-detection approaches can be integrated within a common analytical monitoring layer for multivariate reactor process analysis. The main observations are as follows:

- Reconstruction-based analysis achieves the best performance according to F1-score;
- Fusion of anomaly scores from multiple analytical approaches provides stable behavior under sensitivity-oriented detection settings;
- Score-level Gradient Boosting achieves the highest ROC-AUC;
- Repeated experimental evaluation demonstrates relatively stable comparative behavior across varying perturbation realizations;
- Combined use of different anomaly-related scores improves discrimination between normal and abnormal operating conditions;
- Adaptive threshold recalibration under simulated operating drift reduces false alarms while preserving high anomaly-detection sensitivity, demonstrating one adaptive capability relevant to future digital twin deployment.

The obtained results support the integration of complementary analytical approaches within the monitoring layer of the proposed digital twin framework and provide initial experimental evidence for adaptive monitoring behavior under changing operating conditions.

6. Discussion

6.1. Interpretation of Experimental Results

The experimental results demonstrate the feasibility of comparatively evaluating complementary analytical approaches for multivariate reactor-monitoring conditions within a unified analytical workflow. The performed experiments highlight the advantages and limitations of different anomaly-detection strategies under controlled synthetic perturbation scenarios and illustrate the trade-off between anomaly sensitivity and false-alarm behavior in safety-oriented monitoring environments.

PCA Reconstruction achieved the highest F1-score, indicating comparatively stronger sensitivity to complex multivariate process deviations in reactor data. Score-level Gradient Boosting achieved the highest ROC-AUC and demonstrated the strongest overall discrimination capability for distinguishing between normal and abnormal sequences. The anomaly-score-fusion approach also showed stable performance, particularly in scenarios requiring high anomaly sensitivity.

The results support the concept that the analytical layer of a reactor-monitoring digital twin architecture should not rely on a single model or analytical method. Different approaches capture different manifestations of process deviations in reactor operation and can complement each other in process analysis and reactor condition assessment.

The observed differences between the evaluated analytical approaches are associated primarily with the different monitoring principles employed by the models. Reconstruction-based approaches, such as PCA-based monitoring, are particularly sensitive to correlated multivariate deviations because reconstruction residuals increase when relationships among reactor-process variables change simultaneously. In contrast, temporal residual analysis is more sensitive to localized sequential deviations, while score-fusion approaches combine complementary anomaly indicators and therefore improve sensitivity under heterogeneous perturbation conditions. The obtained results additionally indicate that increased anomaly sensitivity is frequently accompanied by increased false-alarm generation under conservative threshold configurations and imbalanced anomaly detection conditions.

The relatively low Precision values observed across several approaches are likely associated with the conservative threshold calibration strategy adopted during the experiments, in which anomaly sensitivity and Recall were prioritized in order to minimize the probability of missed abnormal conditions. This trade-off is common in safety-critical monitoring environments, in which false negatives may be considered more critical than false positives. Consequently, the reported performance should be interpreted as comparative validation of the analytical monitoring layer rather than as evidence of deployment-ready operational performance.

6.2. Importance of Multivariate and Temporal Analysis

Reactor process variables are related both temporally and physically. Temperature, pressure, coolant flow, and other operational parameters evolve jointly and should not be treated as independent signals. In PWR systems, thermal, hydraulic, and pressure-related parameters are physically coupled through the primary circuit and reactor control systems. As a result, changes in coolant flow rate, temperature distribution, or pressure conditions are typically reflected simultaneously across multiple correlated measurements rather than appearing as isolated deviations in individual signals.

For this reason, reconstruction-based approaches are particularly suitable for reactor process analysis because they evaluate the consistency of the overall system state instead of relying solely on local thresholds or temporal residual deviations. This is especially important for slowly developing process deviations, in which the relationships among parameters may change before significant deviations appear in individual measurement channels.

The dataset used in this study preserves the multivariate structure of reactor measurements and enables more realistic reactor process analysis. The obtained results indicate that approaches such as PCA-based reconstruction and anomaly score fusion are more effective for detecting complex multivariate process deviations, relative to methods based solely on residual analysis.

Temporal data splitting and the use of sliding time windows additionally reduce the risk of unrealistically optimistic results that may arise from random mixing of past and future reactor states. In this way, the experimental procedure remains closer to realistic conditions for process analysis and reactor condition assessment.

During real transient operating conditions in PWR systems, process deviations frequently appear not as abrupt isolated changes but as gradual modifications in the relationships among multiple process variables. For example, changes in circulation regime or transient thermal loading may produce correlated variations in temperature, pressure, and coolant flow rate.

In such cases, reconstruction-based approaches may be more sensitive to disruptions in normal parameter relationships, relative to methods relying solely on local deviations in individual signals.

A similar situation may occur during gradual degradation of instrumentation channels or sensors, where absolute values remain within acceptable limits while temporal consistency among correlated reactor parameters progressively deteriorates. This explains the strong performance of multivariate reconstruction-based and ensemble-based approaches observed in the conducted experiments.

In real operating conditions, multivariate analytical and machine learning-based approaches can support monitoring of key reactor-process parameters and early detection of abnormal conditions. Such methods are particularly useful for analyzing relationships among coolant flow, temperature, and pressure in the primary circuit during both steady-state and transient operating conditions.

Under normal operating conditions, these parameters follow expected thermohydraulic relationships determined by reactor design and coolant properties. If these relationships gradually change over time—for example, if the correlation between temperature, pressure, and coolant flow becomes inconsistent—this may indicate developing process deviations. In practice, such effects may be associated with degradation of reactor coolant pump performance or gradual reduction of heat-transfer efficiency in steam generators. In many cases, individual parameters may remain within acceptable operating ranges while the multivariate relationships among them progressively change.

Techniques such as PCA-based analysis, PLS modeling, and autoencoder-based approaches are commonly used to characterize multivariate operational behavior and detect deviations from nominal process conditions.

In some cases, neural-network-based models, including autoencoders, are also applied, particularly when the relationships among variables are nonlinear or highly coupled. These approaches are typically used to characterize nominal operating behavior from historical data. Deviations from this reference behavior can subsequently be identified through statistical indicators or reconstruction errors, even when no individual parameter exceeds predefined limits.

A similar principle can be applied to reactor control and instrumentation systems. Simultaneous deviations across multiple correlated signals may indicate abnormal behavior or degradation mechanisms. Analysis of temporal dependencies and delays between related signals can help distinguish normal transient effects from phenomena such as sensor drift, actuator degradation, or control-system-related abnormal operating conditions.

Sensor drift represents a particularly important challenge in long-term reactor monitoring. In practice, drift corresponds to gradual and systematic deviation of sensor measurements over time, even when the underlying physical parameter remains stable. Such effects may influence both sensor offset and sensitivity, leading to persistent measurement inaccuracies.

In nuclear power plant operation, such effects are particularly relevant for temperature, pressure, flow, and neutron flux measurements. Even when the deviation remains within acceptable limits for a certain period, it can still distort the relationships between physically connected parameters. As a result, the overall behavior of the system may appear consistent at first glance, while underlying inconsistencies gradually develop.

The underlying causes are usually linked to ageing and environmental conditions, including thermal cycling, radiation exposure, mechanical stress, and gradual degradation of electronic components or sensing elements, as discussed in the literature on instrumentation-ageing in nuclear power plants [27]. In some cases, contamination, fouling, or changes in material properties may also contribute to the effect.

From a monitoring perspective, sensor drift is difficult to identify using simple threshold-based methods because individual signals may remain within acceptable limits for extended periods. Instead, drift often manifests itself as gradual degradation of correlations among physically related variables or as increasing model residuals over time. For this reason, multivariate analytical approaches are well suited for detecting such effects, since they rely on the consistency of relationships among multiple process variables rather than on isolated measurements alone.

In this context, detecting sensor drift is not only important for measurement accuracy, but also for maintaining the reliability of higher-level monitoring, diagnostic, and control functions that depend on these signals.

Similar analytical approaches can also support monitoring of auxiliary systems and critical equipment. The combination of vibration measurements with thermal and hydraulic variables can provide improved insight into equipment condition and degradation trends. This becomes increasingly important during long-term operation and ageing, in which changes are often gradual and difficult to detect through individual measurements alone. Even when all monitored variables remain within acceptable operating limits, the overall multivariate behavior of the system may still progressively change. In this context, condition-based monitoring and predictive maintenance approaches, including remaining useful life estimation, may provide practical operational benefits.

6.3. Conservative Threshold Selection and the Precision–Recall Trade-Off

The high Recall values achieved by several models indicate good sensitivity to anomalies and process deviations in reactor operation. This is particularly important in nuclear applications, in which missed abnormal operating conditions may lead to consequences that are more severe than the presence of a moderate number of false alarms.

The lower Precision values indicate that parts of the normal sequences are classified as abnormal. This reflects the conservative experimental configuration, in which priority is given to early detection of potential deviations and anomalies.

In real operating environments, threshold selection should be adapted according to the required sensitivity, the acceptable number of false alarms, and operator workload. Such a

trade-off is particularly important in decision-support systems and predictive condition assessment applications.

6.4. Connection to the Proposed Adaptive Framework

The results presented in Section 5 validate the analytical layer of the proposed framework and demonstrate that statistical, predictive, reconstruction-based, and score fusion approaches can be integrated within a unified workflow for reactor data analysis.

Several architectural components remain subjects for future implementation, including adaptive online updating, uncertainty estimation, integrated explainability, and incorporation of physical constraints and reactor-process models.

The adaptive recalibration experiment provides an initial demonstration of one adaptive capability relevant to future digital twin deployment. Although the implemented mechanism is limited to threshold adaptation and does not include online state estimation, bidirectional synchronization, uncertainty-aware adaptation, or physics-informed updating, the observed reduction in false alarms under operating drift suggests that adaptive monitoring can improve long-term robustness. These results support the feasibility of extending the proposed architecture toward more advanced adaptive functionality.

In practical deployment, adaptive recalibration should be activated only after the observed drift has been verified to correspond to a benign operating-point change or sensor-related variation rather than an incipient process fault. Such verification would require higher-level monitoring or diagnostic mechanisms, and remains outside the scope of the present proof-of-concept implementation.

Future extensions of the architecture may incorporate uncertainty-estimation mechanisms, explainability support, online model adaptation, and physics-informed integration. Additional software-architecture developments may include modular analytical services, streaming-based monitoring pipelines, and operator-support interfaces intended to improve monitoring reliability, interpretability, scalability, and adaptation under changing operating conditions.

Nevertheless, the experiments conducted indicate that the analytical layer can serve as a practical foundation for future adaptive digital twin systems intended for process analysis, anomaly detection, and operator decision support in nuclear engineering.

At the same time, it should be kept in mind that in nuclear applications these methods are generally used as support tools, not as part of safety-critical functions. Their use requires proper validation and testing and a clear understanding of how the results are interpreted. This is important in order to remain consistent with established safety principles and regulatory expectations, including those defined in the IAEA safety standards [28].

In this context, data-driven monitoring approaches may strengthen diagnostic and monitoring capabilities within the broader concept of defense-in-depth without replacing deterministic safety functions. Such systems should also remain aligned with instrumentation and control requirements related to reliability, independence, and traceability of information used for operational decision-making.

Similar concepts are discussed in IAEA publications related to online monitoring, condition-based maintenance, and diagnostic systems for nuclear power plants [29,30]. These documents describe established approaches for detecting equipment degradation through signal analysis and for integrating monitoring systems into plant operation and maintenance practices. They explicitly address the use of data-driven methods for early fault detection, diagnosis of abnormal conditions, and support of maintenance decision-making. At the same time, they clearly define boundary conditions for their application, including requirements for data quality, model validation, and consistent interpretation of results. In this sense, data-driven approaches are not treated as standalone solutions,

but as structured extensions of conventional engineering methods within a controlled and well-defined framework.

6.5. Limitations and Future Work

The present study has several limitations. First, the original dataset does not contain fully annotated abnormal-event labels for the evaluated monitoring scenarios. Therefore, controlled perturbation scenarios are introduced for comparative evaluation, and the corresponding evaluation labels are assigned to the perturbed temporal intervals. Publicly available fully annotated datasets for abnormal events in nuclear reactor environments remain limited, which constrains direct validation under realistic fault conditions. In addition, the use of controlled synthetic perturbations, while useful for comparative evaluation, cannot fully reproduce the complexity and physical coupling mechanisms of real reactor fault conditions.

Second, the experiments primarily validate the analytical monitoring layer of the reactor-monitoring architecture rather than a complete system ready for practical deployment. The experiments implemented in this study focus on comparative evaluation of complementary anomaly-detection approaches under multivariate reactor-monitoring conditions and should therefore be interpreted within the scope of the implemented analytical monitoring subsystem.

Third, several important architectural components remain outside the scope of the current implementation, including

- Online adaptive retraining;
- Prediction reliability assessment;
- Explainability and interpretation of analytical outputs;
- Integration of physical constraints and simulation-based dependencies.

These components remain defined modules of the architecture, and were not experimentally validated within the scope of the present study. Uncertainty estimation, explainability mechanisms, and integration of reactor-process models are currently considered primarily at the architectural level and are not experimentally implemented in the present work.

Fourth, the comparative evaluation environment does not yet include

- Physically validated reactor transient simulations;
- Long-term operational datasets;
- Prolonged condition assessment under real operating conditions;
- Real-time adaptive model updating.

Direct quantitative comparison with previous studies also remains limited because different publications utilize different datasets, anomaly-generation procedures, preprocessing workflows, and evaluation metrics.

The present experimental evaluation focuses primarily on comparative evaluation under controlled multivariate perturbation scenarios. Although additional operationally relevant metrics, including PR-AUC and False Alarm Rate, were included in the revised comparative evaluation, further investigation of anomaly-detection delay, sensitivity with respect to perturbation magnitude, and long-term operational robustness remains necessary for practical deployment validation.

Such evaluation metrics are particularly important in safety-critical reactor-monitoring environments, because excessive false alarms, delayed anomaly detection, or sensitivity variations under different transient conditions may significantly influence operator interpretation and deployment feasibility. Future work should therefore include more comprehensive operational validation using physically validated transient simulations,

reactor-specific operational constraints, sensor-drift scenarios, long-term operational sequences, evaluation under incomplete or noisy measurements, and extended temporal performance analysis.

The present study does not include experimental validation of real-time deployment performance, hardware resource consumption, or online adaptive recalibration under operational nuclear power plant conditions. The proof-of-concept adaptive recalibration experiment is limited to simulated operating drift and threshold adaptation; it does not represent online adaptive model updating under operational plant conditions. Several evaluated analytical approaches, including PCA-based reconstruction and statistical monitoring methods, have moderate computational requirements and may be suitable for supervisory monitoring workflows under controlled operating conditions.

Nevertheless, practical deployment of adaptive updating mechanisms would require additional safety-oriented control procedures, including monitoring of model drift, validation of recalibrated parameters before deployment, rollback mechanisms, operator supervision, and uncertainty-aware evaluation, in order to reduce the risk of unstable monitoring behavior or increased false-alarm rates under changing operational conditions.

Additional investigation of computational efficiency, uncertainty calibration, explainability integration, and long-term operational stability also remains necessary before the proposed architecture can be considered a deployment-ready monitoring solution for safety-critical reactor environments. This includes maintaining an appropriate balance between anomaly sensitivity and false-alarm burden in order to support reliable operator decision-making and effective human-machine collaboration in safety-critical reactor environments.

7. Conclusions

- (1) The present study proposes a modular digital twin-oriented architecture for reactor process analysis, anomaly detection, and analytical support for nuclear reactor monitoring under multivariate operating conditions. The proposed framework integrates multivariate reactor-data analysis, temporal analysis, reconstruction-based monitoring, anomaly-score generation, and supervisory analytical integration, while adaptive recalibration, uncertainty-aware analysis, explainability support, and physics-informed integration are defined as architectural extensions for future development.
- (2) The experimental validation was conducted using the publicly available PWR Abnormality Dataset containing multivariate reactor-monitoring time-series measurements. The implemented analytical workflow included temporal partitioning, sliding-window sequence generation, controlled synthetic perturbation procedures, anomaly-score normalization, temporal smoothing, threshold calibration, and comparative evaluation of several complementary analytical approaches, including temporal residual analysis, robust multivariate monitoring, PCA-based reconstruction, anomaly-score fusion, and score-level Gradient Boosting classification. The experimental validation additionally included repeated evaluation runs and comparison with a conventional Isolation Forest baseline in order to improve robustness assessment of the implemented analytical monitoring approaches.
- (3) The obtained results indicate that reconstruction-based and score-fusion approaches demonstrate sensitivity to deviations in multivariate reactor-monitoring data under controlled experimental conditions. PCA-Based Reconstruction achieved the highest F1-score, while score-level Gradient Boosting demonstrated the highest ROC-AUC discrimination capability between normal and abnormal operating conditions. The experiments further demonstrate that complementary analytical approaches can be integrated within a common analytical monitoring workflow for reactor condition assessment and anomaly analysis. The results additionally highlight the importance

- of multivariate temporal analysis for identification of gradually developing deviations associated with changing relationships among monitored process variables. A limited proof-of-concept adaptive recalibration experiment further demonstrated that threshold adaptation under simulated operating drift can reduce false alarms while preserving high anomaly-detection sensitivity, providing an initial experimental indication of adaptive monitoring behavior relevant to future digital twin deployment.
- (4) The present study should be interpreted primarily as a comparative analytical validation conducted within a modular reactor-monitoring architecture, rather than a fully operational adaptive digital twin implementation. The conducted experiments are limited to offline evaluation using controlled perturbation scenarios and do not include experimentally validated online adaptive updating, uncertainty-estimation mechanisms, explainability support, or coupling with physics-based reactor simulations. Future work should therefore include physically validated transient scenarios, long-term operational datasets, sensor-drift evaluation, adaptive online updating and recalibration, uncertainty-aware monitoring, explainability integration, and additional validation under realistic operating environments while preserving the essential roles of operator oversight and engineering expertise.

Funding: This research was funded by the Ministry of Education and Science of the Republic of Bulgaria, through National Program D01-99: Qualification improvement in the field of nuclear technologies and nuclear engineering.

Data Availability Statement: The publicly available pressurized water reactor (PWR) abnormality dataset used in this study is available from the Kaggle repository: <https://www.kaggle.com/datasets/avibagul80/pressurized-water-reactor> (accessed on 9 July 2026). The processed analytical datasets, preprocessing procedures, temporal sequence generation workflow, and controlled synthetic perturbation procedures are described in Sections 5.2–5.4 of the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Mondal, K.; Pal, S.; Sleiti, A.; Attaran, M. Advanced Manufacturing and Digital Twin Technology for Nuclear Energy Systems: A Review. *Front. Energy Res.* **2024**, *12*, 1339836. [CrossRef]
2. Kochunas, B.; Huan, X.; Pevey, L.; Halsey, W.; Mohamed, A.; Heifetz, A.; Vilim, R.; Boring, R. Digital Twin Concepts with Uncertainty for Nuclear Power Systems. *Energies* **2021**, *14*, 4235. [CrossRef]
3. Sandhu, R.; Yadav, V.; Tiwari, A.; Jain, V. A Future with Machine Learning: Review of Condition Assessment Frameworks for Nuclear Power Plants. *Energies* **2023**, *16*, 2628. [CrossRef]
4. Wang, M.D.; Lin, T.H.; Jhan, K.C.; Wu, S.C. Abnormal Event Detection, Identification and Isolation in Nuclear Systems. *Prog. Nucl. Energy* **2021**, *140*, 103928. [CrossRef]
5. Chen, Z. Research on Fault Prediction of Nuclear Safety-Class Signal Conditioning Module. *Energies* **2024**, *17*, 4063. [CrossRef]
6. Ioannou, G.; Tagaris, T.; Alexandridis, G.; Stafylopatis, A. Intelligent Techniques for Anomaly Detection in Research Reactor Monitoring Systems. In *PHYSOR 2020: Transition to a Scalable Nuclear Future*; EPJ Web of Conferences: Les Ulis, France, 2021; Volume 247, p. 21011. [CrossRef]
7. Chaudhary, A.; Han, J.; Kim, S.; Kim, A.; Choi, S. Anomaly Detection and Analysis in Nuclear Power Plants. *Electronics* **2024**, *13*, 4428. [CrossRef]
8. Daniell, J.; Kobayashi, K.; Alajo, A.; Alam, S.B. Digital Twin-Centered Hybrid Data-Driven Multi-Stage Deep Learning Framework for Enhanced Nuclear Reactor Power Prediction. *Energy AI* **2025**, *19*, 100450. [CrossRef]
9. Kobayashi, K.; Alam, S.B. Deep Neural Operator-Driven Real-Time Inference to Enable Digital Twins for Reactor Power Prediction. *Sci. Rep.* **2024**, *14*, 2101. [CrossRef] [PubMed]
10. Bei, X.; Cheng, M.; Zuo, X.; Yu, K.; Dai, Y. Surrogate Models Based on Back-Propagation Neural Network for Parameters Prediction of the PWR Core. In *Proceedings of the 23rd Pacific Basin Nuclear Conference*; Springer: Singapore, 2023; Volume 284, pp. 1207–1217. [CrossRef]
11. Hossain, R.B.; Ahmed, F.; Kobayashi, K.; Koric, S.; Abueidda, D.; Alam, S.B. Virtual Sensing-Enabled Digital Twin Framework for Real-Time Monitoring of Nuclear Systems Leveraging Deep Neural Operators. *arXiv* **2024**, arXiv:2410.13762.

12. Lu, L.; Jin, P.; Pang, G.; Zhang, Z.; Karniadakis, G.E. Learning Nonlinear Operators via DeepONet Based on the Universal Approximation Theorem of Operators. *Nat. Mach. Intell.* **2021**, *3*, 218–229. [\[CrossRef\]](#)
13. 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. [\[CrossRef\]](#)
14. Gal, Y.; Ghahramani, Z. Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. In Proceedings of the 33rd International Conference on Machine Learning, New York, NY, USA, 19–24 June 2016; pp. 1050–1059.
15. International Atomic Energy Agency (IAEA). *Considerations for Deploying Artificial Intelligence Technologies in Nuclear Power Plants*; IAEA Nuclear Energy Series No. NR-T-3.2; International Atomic Energy Agency: Vienna, Austria, 2024.
16. Hashemian, H.M. Aging management of instrumentation and control sensors in nuclear power plants. *Nucl. Eng. Des.* **2010**, *240*, 3781–3790. [\[CrossRef\]](#)
17. IAEA. *Safety of Nuclear Power Plants: Design, Safety Standards Series No. SSR-2/1 (Rev. 1)*; IAEA: Vienna, Austria, 2016.
18. IAEA. *On-Line Monitoring for Improving Performance of Nuclear Power Plants Part 2: Process and Component Condition Monitoring and Diagnostics*, IAEA Nuclear Energy Series No. NP-T-1.2; IAEA: Vienna, Austria, 2008.
19. IAEA. *Implementation Strategies and Tools for Condition Based Maintenance at Nuclear Power Plants*, IAEA-TECDOC-1551; IAEA: Vienna, Austria, 2007.
20. Huang, Q.; Liu, Y.; Zhang, H.; Wang, J. A Review of the Application of Artificial Intelligence to Nuclear Reactors: Where We Are and What's Next. *Heliyon* **2023**, *9*, e13883. [\[CrossRef\]](#) [\[PubMed\]](#)
21. International Atomic Energy Agency (IAEA). *Considerations for Deploying Artificial Intelligence Applications in the Nuclear Power Industry*; IAEA Technical Document; IAEA: Vienna, Austria, 2025.
22. Jendoubi, C.; Al Rashdan, A.; Aljamaan, H. A Survey of Artificial Intelligence Applications in Nuclear Power Plants. *Computation* **2024**, *12*, 30. [\[CrossRef\]](#)
23. Saad, M.H.; Said, A. Machine Learning-Based Fault Diagnosis for Research Nuclear Reactor Medium Voltage Power Cables in Fraction Fourier Domain. *Electr. Eng.* **2023**, *105*, 25–42. [\[CrossRef\]](#)
24. Kautz, E.; Hagen, A.; Johns, J.; Burkes, D. A Machine Learning Approach to Thermal Conductivity Modeling: A Case Study on Irradiated Uranium-Molybdenum Nuclear Fuels. *Comput. Mater. Sci.* **2019**, *161*, 107–118. [\[CrossRef\]](#)
25. Guo, Z.; Wu, Z.; Liu, S.; Ma, X.; Wang, C.; Yan, D.; Niu, F. Defect Detection of Nuclear Fuel Assembly Based on Deep Neural Network. *Ann. Nucl. Energy* **2020**, *136*, 107078. [\[CrossRef\]](#)
26. He, C.; Ge, D.; Yang, M.; Yong, N.; Wang, J.; Yu, J. A Data-Driven Adaptive Fault Diagnosis Methodology Based on NSGAIL-CNN. *Ann. Nucl. Energy* **2021**, *158*, 108326. [\[CrossRef\]](#)
27. Mendoza, H.; Yadav, V.; Alam, S.B.; Eskins, D.; Iyengar, R. Advances in Digital Twins and AI/ML for Condition Monitoring in Nuclear Applications. *Front. Energy Res.* **2026**, *14*, 1716514. [\[CrossRef\]](#)
28. Hu, M.; He, Y.; Lin, X.; Lu, Z.; Jiang, Z.; Ma, B. Digital twin model of gas turbine and its application in warning of performance fault. *Chin. J. Aeronaut.* **2023**, *36*, 449–470. [\[CrossRef\]](#)
29. Zhu, E.; Li, T.; Xiong, J.; Chai, X.; Zhang, T.; Liu, X. A digital twin framework for real-time operation monitoring of space nuclear power systems. *Reliab. Eng. Syst. Saf.* **2026**, *267*, 111843. [\[CrossRef\]](#)
30. Lakshminarayanan, B.; Pritzel, A.; Blundell, C. Simple and Scalable Predictive Uncertainty Estimation Using Deep Ensembles. In Proceedings of the Advances in Neural Information Processing Systems 30: NIPS 2017, Long Beach, CA, USA, 4–9 December 2017; pp. 6405–6416.
31. Ribeiro, M.T.; Singh, S.; Guestrin, C. “Why Should I Trust You?”: Explaining the Predictions of Any Classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 1135–1144. [\[CrossRef\]](#)
32. OECD Nuclear Energy Agency (NEA). *Benchmark on Artificial Intelligence and Machine Learning for Scientific Computing in Nuclear Engineering*; OECD NEA Working Paper; OECD NEA: Paris, France, 2024.
33. Idaho National Laboratory. *Explainable Artificial Intelligence for Nuclear Power Applications*; Idaho National Laboratory: Idaho Falls, ID, USA, 2023.
34. Lundberg, S.M.; Lee, S.-I. A Unified Approach to Interpreting Model Predictions. In Proceedings of the Advances in Neural Information Processing Systems 30: NIPS 2017, Long Beach, CA, USA, 4–9 December 2017; pp. 4768–4777.
35. Pressurized Water Reactor (PWR) Abnormality Dataset. Kaggle Dataset Repository. Available online: <https://www.kaggle.com/datasets/avibagul80/pressurized-water-reactor> (accessed on 6 June 2026).

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.