

Review

Not peer-reviewed version

---

# Digital Twin Technology in Pipeline Engineering: A Study Review of Applications, Challenges, and Future Directions

---

[Hamed Azimi](#)\*, [Rahim Shoghi](#), [Hodjat Shiri](#)

Posted Date: 11 June 2026

doi: 10.20944/preprints202606.0848.v1

Keywords: digital twin; pipeline engineering; leak detection; predictive maintenance; condition monitoring; machine learning



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC, OpenAlex.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

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.

Review

# Digital Twin Technology in Pipeline Engineering: A Study Review of Applications, Challenges, and Future Directions

Hamed Azimi \*, Rahim Shoghi and Hodjat Shiri

Civil Engineering Department, Faculty of Engineering and Applied Science, Memorial University of Newfoundland, A1B 3X5, St. John's, NL, Canada

\* Correspondence: hazimisiahch@mun.ca

## Abstract

Digital Twin (DT) technology has emerged as a transformative approach in pipeline engineering, enabling real-time monitoring, predictive analytics, and enhanced decision-making across the asset lifecycle. This review critically examines recent advancements in the application of digital twins for pipeline systems, with a particular focus on condition monitoring, leak detection, corrosion assessment, and predictive maintenance. The study synthesizes findings from a wide range of literature to identify key enabling technologies, including Internet of Things (IoT) sensors, data-driven modeling, computational fluid dynamics (CFD), and machine learning algorithms. Special attention is given to the integration of physics-based and data-driven models for improving the accuracy and reliability of digital twin frameworks. In addition, this paper proposes a unified reference architecture for pipeline digital twins, supported by a mathematical formulation of synchronization and a comparative synthesis of existing approaches. The review highlights how digital twins facilitate early fault detection and operational optimization by continuously synchronizing physical assets with their virtual counterparts. The review also emphasizes the importance of uncertainty-aware and reliability-informed digital twin frameworks for robust decision-making in safety-critical pipeline applications. Applications in subsea, oil and gas, and water distribution pipelines are explored, demonstrating the versatility of DT systems under different environmental and operational conditions. Despite significant progress, challenges remain in data integration, model validation, scalability, and cybersecurity. Furthermore, the lack of standardized architectures and interoperability frameworks limits widespread adoption. This paper concludes by outlining future research directions, including the development of hybrid modeling techniques, edge computing integration, and AI-driven autonomous decision systems. Overall, digital twin technology represents a paradigm shift in pipeline engineering, offering substantial potential to enhance safety, efficiency, and sustainability in complex infrastructure systems.

**Keywords:** digital twin; pipeline engineering; leak detection; predictive maintenance; condition monitoring; machine learning

---

## 1. Introduction

Pipeline infrastructure forms the circulatory system of modern industrialization. Spanning onshore and offshore environments, pipeline networks transport hydrocarbons, potable water, natural gas, and industrial fluids across thousands of kilometers under demanding operational and environmental conditions. The integrity, reliability, and efficiency of these systems carry profound implications for economic productivity, public safety, and environmental protection. Yet, traditional approaches to pipeline management, including periodic inspection, reactive maintenance, and simplified simulation models, increasingly fail to meet the demands of aging infrastructure, tightening regulatory environments, and the economic pressures of the energy transition. Early

conceptual foundations of digital twin technology can be traced to high-fidelity simulation practices developed by NASA during the Apollo missions in the 1960s, where physical systems were mirrored through virtual models to support mission-critical decision-making under uncertain conditions. The modern formulation of the digital twin is generally attributed to early work by Grieves in the context of product lifecycle management (PLM) in the early 2000s and was later formalized and widely disseminated in subsequent publications (e.g., Grieves and Vickers [1]). A digital twin is a dynamic, synchronized virtual representation of a physical asset or system that integrates real-time sensor data, physics-based models, and data-driven algorithms to continuously reflect the current and predicted state of its physical counterpart.

Grieves and Vickers [1] emphasized the importance of continuous interaction between physical systems and their virtual counterparts across the lifecycle, enabling improved understanding, prediction, and management of complex system behavior. They argued that undesirable emergent behaviors in complex systems originate during the creation and production lifecycle phases, manifesting only during operations, often due to human interaction. They proposed the Digital Twin, a virtual equivalent linked to the physical system, as a mitigation strategy. The study described the Digital Twin concept, its development, and its application across the product lifecycle to define and understand system behavior. The authors related the Digital Twin to Systems Engineering and addressed how it could manage human interactions leading to "normal accidents." They examined both obstacles, such as software interoperability issues, and opportunities, including system replication and front running. The chapter concluded by referencing NASA's ongoing work with the Digital Twin, emphasizing its potential to pre-emptively identify and mitigate unpredictable, undesirable emergent behavior in complex systems throughout the entire lifecycle.

Unlike static simulation models, digital twins are living entities that evolve throughout the lifecycle of the asset they represent, from design and commissioning through operation and decommissioning.

Kritzinger et al. [2] aimed to clarify the ambiguous definition of digital twin (DT) in manufacturing by conducting a categorical literature review and proposing a structured classification framework. The methodology involved a systematic analysis of existing publications across multiple disciplines, categorizing them based on the level of integration between physical and virtual systems. The authors introduced three distinct concepts: Digital Model (DM), Digital Shadow (DS), and Digital Twin (DT), representing increasing levels of data exchange and system synchronization. The findings showed that most existing studies fell within the DM and DS categories, while fully integrated digital twins were still relatively scarce in manufacturing applications. This highlighted a gap between theoretical potential and practical implementation. From a practical perspective, the study emphasized the need for standardized definitions, robust architectures, and improved real-time data integration capabilities to enable the transition toward fully functional digital twins in industrial environments.

Tao et al. [3] explored an extended five-dimensional digital twin (DT) model, which added data and services to the original three-dimensional structure to meet new application and technological requirements. The authors employed a conceptual modeling approach to define this framework and further classified DT into three application levels: unit, system, and system of systems, while discussing associated key technologies. The findings indicated that the five-dimensional model (physical entity, virtual entity, connection, data, services) offered a more comprehensive architecture, and eight guiding rules were proposed for future modeling work. Practically, this framework provided a structured foundation for implementing DT across complex systems, enabling better integration of data analytics and service functionalities for diverse applications.

Fuller et al. [4] provided a comprehensive review of Digital Twin technology, focusing on its definitions, misconceptions, enabling technologies, challenges, and open research across manufacturing, healthcare, and smart cities. The authors conducted a categorical literature review of publications from 2015 onwards, identifying 26 key sources, and analyzed papers based on integration levels (digital model, shadow, or twin) while categorizing them by research area,

technology used, and application domain. The findings revealed that many publications misidentified digital models or shadows as digital twins, highlighting definitional inconsistencies, while key enabling technologies included IoT/IIoT and data analytics, with shared challenges encompassing IT infrastructure, data quality, privacy, security, trust, and unrealistic expectations. Practically, the study emphasized that practitioners should adopt standardized modeling approaches, ensure bidirectional data integration for true digital twins, address infrastructure and security requirements early, and manage stakeholder expectations to enable successful implementation across diverse sectors.

The application of digital twin concepts to pipeline engineering has accelerated substantially over the past decade, driven by advances in the Internet of Things (IoT), edge computing, cloud infrastructure, machine learning, and high-fidelity numerical simulation. Key application areas include real-time leak detection, corrosion monitoring, structural health assessment, hydraulic performance optimization, predictive maintenance scheduling, and operator training. The convergence of these capabilities within a unified digital twin framework promises to transition pipeline management from a reactive to a fully predictive and prescriptive paradigm.

Despite this momentum, the literature reveals significant heterogeneity in how digital twins are defined, implemented, and evaluated in the context of pipeline engineering. Varying levels of model fidelity, sensor coverage, data quality, and integration architecture make direct comparison difficult. Comprehensive reviews that synthesize progress across the full breadth of pipeline engineering sub-domains remain scarce. However, existing studies are often fragmented across application domains and lack a unified framework for analysis. This paper addresses that gap.

The remainder of this paper is organized as follows. Section 2 reviews digital twin applications in oil and gas pipeline systems. Section 3 addresses hydraulic pipeline systems. Section 4 covers leak detection across gas and water pipelines. Section 5 examines subsea pipeline monitoring and structural integrity. Section 6 discusses condition monitoring and predictive maintenance. Section 7 synthesizes cross-cutting findings and introduces a unified framework for pipeline digital twins, integrating architectural, mathematical, and comparative analyses, while identifying key research gaps and future directions. Section 8 concludes the review.

To provide a structured overview of the diverse modeling paradigms used in pipeline digital twin systems, Table 1 presents a comparative analysis of the principal approaches based on key performance criteria. DT implementations in pipeline engineering span a diverse spectrum of modeling paradigms, each characterized by distinct assumptions, computational requirements, data dependencies, and operational capabilities. Broadly, these approaches can be categorized into six principal classes: physics-based models, data-driven models, hybrid approaches, reduced-order models, probabilistic digital twins, and visual or interface-oriented twins.

Physics-based models rely on first-principles formulations such as computational fluid dynamics (CFD), finite element methods (FEM), and real-time transient models (RTTM), offering high interpretability and strong extrapolation capability under varying operating conditions. However, their computational intensity often limits real-time applicability. In contrast, data-driven approaches, including artificial neural networks (ANN), support vector machines (SVM), and deep learning (DL), provide rapid predictions and adaptability but require large volumes of high-quality training data and may lack physical interpretability.

Hybrid digital twins have emerged as a dominant paradigm, integrating physics-based and data-driven components to leverage the strengths of both approaches. These models typically achieve superior predictive performance but introduce additional complexity in model coupling and validation. Reduced-order models (ROMs), including Digital Twin Reduced-Order Technology (DTROT), aim to approximate high-fidelity simulations with significantly lower computational cost, enabling real-time deployment while maintaining acceptable accuracy.

Probabilistic digital twins incorporate stochastic modeling techniques, such as Bayesian inference and Ensemble Kalman Filtering (EnKF), to explicitly account for uncertainty in measurements, model parameters, and predictions. This capability is particularly critical for risk-

informed decision-making in safety-critical pipeline operations. Finally, visual digital twins focus on real-time visualization and human-machine interaction through 3D or virtual reality (VR) environments, enhancing situational awareness but typically relying on underlying analytical models for quantitative insights.

Given the diversity of these approaches, a structured comparison is essential to evaluate their relative strengths, limitations, and suitability for different pipeline engineering applications.

**Table 1.** A comparative analysis of the principal digital twin modeling approaches based on key performance criteria, including accuracy, real-time capability, data requirements, advantages, and limitations.

Method Type	Example Techniques	Accuracy	Real-Time Capability	Data Requirement	Advantages	Limitations
Physics-Based	CFD, FEM, RTTM	High	Low	Medium	Interpretable, reliable	Computationally expensive
Data-Driven	ANN, SVM, DL	Medium-High	High	High	Fast, adaptive	Require large datasets
Hybrid	Physics + ML	Very High	Medium-High	High	Best performance	Complex integration
Reduced-Order Models	ROM, DTROT	Medium-High	Very High	Medium	Efficient	Reduced fidelity
Probabilistic DT	Bayesian, EnKF	High	Medium	High	Handles uncertainty	Complex
Visual Twins	3D/VR models	Medium	High	Medium	Operator-friendly	Not analytical alone

## 2. Oil and Gas Pipelines: Frameworks and System-Level Twins

### 2.1. Overview and Scope

The oil and gas sector was among the earliest industrial domains to embrace digital twin concepts at scale, motivated by the high capital costs of infrastructure, the severe consequences of unplanned downtime or catastrophic failure, and the complexity of integrated production systems that span wellbore, flowline, processing facility, and export infrastructure. Digital twins in this context range from isolated component-level models, addressing individual compressors, separators, or pump stations, to enterprise-wide digital twins that link reservoir inflow performance to downstream logistics.

Conejos Fuertes et al. [5] established the essential requirements and strategic framework for implementing a Digital Twin (DT) within drinking water distribution networks (WDNs) to enhance real-time decision-making. The methodology involved the development and multi-year maintenance of a large-scale DT for the water network of Valencia, Spain, serving 1.6 million inhabitants. This process integrated high-resolution hydraulic models with real-time sensor data, including pressures, flows, and hourly-metered consumption, to create a high-fidelity virtual replica. Findings revealed that a successful DT must achieve absolute topological accuracy and continuous calibration to reliably simulate "what-if" scenarios and optimize network operations. The researchers demonstrated that the DT facilitated optimal design, leak detection, and efficient pressure management, potentially yielding up to 28% in water savings. Practical considerations emphasized that developing a DT is a continuous, challenging process of learning and adjustment, requiring robust data interoperability and a shift toward "living" models that evolve alongside the physical infrastructure.

Wang et al. [6] developed a high-fidelity digital twin (DT) system for real-time pipeline condition monitoring to overcome the limitations of traditional physical-space-driven inspection methods. The methodology integrated several advanced technologies, including Internet of Things (IoT) sensing, finite element simulation, Bayesian inference, and cloud computing, to create a seamless convergence between the physical pipeline and its virtual counterpart. This framework incorporated an ensemble approach for damage detection, localization, and quantification while accounting for uncertainty

propagation through probabilistic simulation. The findings demonstrated that the digital twin accurately predicted the reliability of pipelines suffering from fatigue cracking damage, providing high-fidelity representation and real-time health state updates. Practical considerations highlighted the system's potential for industrial application, suggesting that the integration of virtual reality and automated analytics could significantly reduce operational costs and prevent catastrophic failures through predictive maintenance. These results confirmed the effectiveness of digital twins in enhancing the intelligence and long-term validity of pipeline integrity management.

Hamidishad et al. [7] examined the development and application of Digital Twin (DT) frameworks for oil and gas processing plants, specifically focusing on offshore and topside gas-processing systems like Floating Production Storage and Offloading (FPSO) units. The methodology involved a structured, systematic literature review of over 85 peer-reviewed sources and industrial frameworks to synthesize current DT architectures and modeling approaches. The findings revealed that high-fidelity process simulations, such as Aspen HYSYS, and advanced equations of state like GERG-2008 were essential for accurate virtual replicas, while hybrid machine learning models and co-simulation environments enhanced predictive capabilities. The researchers also identified significant roles for DTs in flare mitigation and ISO 50001-aligned energy optimization. Practical considerations highlighted by the review included the urgent need for industry-wide standardization, long-term validation of virtual models, and the integration of robust cybersecurity measures. Ultimately, the study provided a roadmap for engineers to deploy scalable and auditable DT solutions across the O&G value chain.

## 2.2. Component and System Integration

The architecture of digital twins in oil and gas pipeline applications typically comprises three functional layers: the physical asset layer (sensors, actuators, and pipeline hardware), the data integration layer (SCADA systems, historian databases, and communication networks), and the analytical layer (physics-based models, machine learning algorithms, and decision-support interfaces). The fidelity and coupling between these layers determine the practical utility of the twin.

Physics-based cores in oil and gas pipeline twins frequently employ multiphase flow simulators, such as OLGA, LedaFlow, or custom finite-volume solvers, to model fluid thermodynamics, phase behavior, and pressure-temperature profiles along the pipeline. These models are parameterized and calibrated using operational data and are updated dynamically as conditions evolve. Machine learning components augment the physics-based models by capturing poorly characterized phenomena such as wax deposition, emulsion viscosity, and erosion rates, and by enabling rapid anomaly detection that would be computationally prohibitive with high-fidelity simulators alone.

Enterprise-scale twins in the oil and gas sector increasingly adopt a modular, service-oriented architecture in which individual component twins are composed into system-level twins through standardized interfaces. This modularity facilitates incremental deployment, vendor-agnostic integration, and scalability across diverse asset portfolios. However, the review literature consistently identifies data governance, cybersecurity, and model validation as critical unsolved challenges for large-scale deployment [7].

## 2.3. Regulatory and Safety Considerations

In regulated industries such as oil and gas, the qualification of digital twin outputs for safety-critical decision-making requires robust validation and verification (V&V) frameworks. Regulatory bodies increasingly recognize digital twins as legitimate tools for pipeline integrity management, but clear standards for twin fidelity, uncertainty quantification, and evidence documentation remain under development. The integration of digital twin outputs with regulatory-mandated integrity management programs (IMPs) represents an active area of industry-academic collaboration.

No current standard explicitly defines digital twin fidelity requirements, real-time synchronization criteria, and acceptable uncertainty thresholds for DT decisions. This is exactly why

standards for twin fidelity, uncertainty quantification, and evidence documentation remain under development.

Pipeline integrity and safety are governed by a comprehensive set of internationally recognized standards that collectively provide the regulatory foundation for the adoption of digital twin (DT) technologies. Core pipeline standards, including API RP 1160 and ASME B31.8S, establish risk-based integrity management frameworks, emphasizing threat identification, inspection, monitoring, and documentation, while regulatory requirements such as PHMSA (49 CFR Parts 192 and 195) and CSA Z662 mandate the implementation of auditable Integrity Management Programs (IMPs). Offshore applications are further supported by DNV-ST-F101, which addresses structural reliability and verification. Complementing these, validation-oriented standards such as ASME V&V 10/40 and the NIST VVUQ framework provide guidance on model credibility, verification, validation, and uncertainty quantification, which are essential for qualifying DT outputs in safety-critical decision-making. Additionally, ISO 23247 and ISO 55000 support digital twin architecture and lifecycle asset management, respectively, while API 580/581 and ILI standards enable risk-based inspection and condition monitoring. Collectively, these standards enable the integration of digital twins into regulatory-compliant workflows, although explicit DT-specific requirements remain under development. Table 2 demonstrates the regulatory standards for DT integration in pipeline engineering.

**Table 2.** Regulatory and Supporting Standards for Digital Twin Integration in Pipeline Engineering.

Category	Standard	Scope	Key Requirements	Relevance to Digital Twins
Core Pipeline	API RP 1160	Liquid pipelines	Risk assessment, integrity verification, documentation	Enables DT integration into IMPs and continuous monitoring
Core Pipeline	ASME B31.8S	Gas pipelines	HCAAs, threat mitigation, inspection	Provides structured framework for DT-based risk modeling
Regulation	PHMSA (49 CFR 192/195)	U.S. pipelines	Mandatory IMPs, risk-based decisions	Drives adoption of DT for compliance and monitoring
Core Pipeline	CSA Z662	Canada pipelines	Design, operation, integrity	Supports DT in lifecycle monitoring and risk assessment
Offshore	DNV-ST-F101	Offshore pipelines	Structural reliability, safety factors	Enables DT for structural integrity and verification
Validation	ASME V&V 10/40	Simulation models	Model validation, credibility	Ensures DT outputs are reliable for decision-making
Validation	NIST VVUQ	Digital systems	Verification, validation, uncertainty	Supports trust and uncertainty-aware DT predictions
DT Framework	ISO 23247	Digital twin systems	Architecture, integration	Provides structural basis for DT implementation
Asset Management	ISO 55000	Asset lifecycle	Risk-based management	Aligns DT with lifecycle (digital thread) concepts
Standards Body	ISO/TC 67	Oil & gas	Global standardization	Supports harmonization of DT deployment
Inspection	API 580/581	Risk-based inspection	Probabilistic risk models	Enables DT-driven inspection prioritization
Inspection	ILI (NACE/API/ASME)	Pipeline inspection	Corrosion, defect detection	Provides data inputs for DT calibration
Design Codes	ASME B31.4/B31.8	Pipeline systems	Design, operation, maintenance	Provides baseline models for DT physics

### 3. Hydraulic Pipeline Systems

#### 3.1. Water Distribution and Hydraulic Infrastructure

Hydraulic pipeline systems, encompassing water distribution networks, hydropower penstocks, irrigation networks, and industrial cooling circuits, present distinct digital twin challenges relative to hydrocarbon pipelines. These systems are often characterized by lower operating pressures, large geographic dispersal, heterogeneous pipe materials and ages, and constrained instrumentation budgets. Despite these challenges, the consequences of failures in hydraulic infrastructure, including service disruption, water loss, and contamination, motivate substantial research investment in digital twin methodologies.

Li et al. [8] investigated the key enabling technologies and practical applications of digital twin (DT) systems in hydraulic engineering and proposed a comprehensive framework for integrating virtual-physical interactions across complex water infrastructure. The authors developed a five-dimensional digital twin architecture consisting of physical entities, digital models, data, services, and interaction mechanisms. The methodology combined multi-source data fusion, GIS-BIM integration, real-time monitoring, and numerical hydraulic simulations, and was validated through a case study of the Danjiangkou water diversion project. The findings demonstrated that the proposed framework significantly improved deformation monitoring accuracy, hydrodynamic and water quality simulation fidelity, and the prediction of geological hazards compared with traditional hydraulic monitoring approaches. The digital twin enabled continuous synchronization between the physical system and its virtual representation, supporting predictive analytics and scenario-based simulation. From a practical standpoint, the study highlighted that successful deployment depends on reliable real-time data acquisition, scalable computational infrastructure, and effective integration of heterogeneous datasets. It also emphasized that while digital twins can enhance operational efficiency, safety management, and decision-making in hydraulic projects, challenges related to data consistency, model complexity, and computational demands must be addressed for large-scale implementation.

#### 3.2. CFD-Based Digital Twins for Transient Flow

A particularly rigorous contribution to the hydraulic digital twin literature was made by Paternina-Verona et al. [9], who aimed to develop a digital twin (DT) framework based on computational fluid dynamics (CFD) to analyse transient two-phase air-water flows during pipeline filling and emptying procedures, which are critical operations associated with pressure surges and structural risks. The methodology involved integrating real-time experimental data with two-dimensional (2D) and three-dimensional (3D) CFD models, using mesh sensitivity analysis and calibration techniques to accurately simulate hydraulic behaviour. The model employed the Volume of Fluid (VoF) method to capture air-water interactions and incorporated machine learning (ML) algorithms trained on both experimental and CFD-generated datasets to enhance predictive capabilities. The findings showed that the CFD-based digital twin accurately reproduced pressure dynamics, with low relative errors compared to measured data, and effectively identified hazardous conditions such as pressure peaks and vacuum formation. ML models achieved high prediction accuracy, significantly reducing computational time while maintaining reliability. From a practical perspective, the study demonstrated that integrating CFD and ML within a digital twin enables improved decision-making for pipeline operation, particularly in optimizing valve strategies, mitigating pressure surges, and enhancing system safety.

The authors integrated CFD models into a digital twin framework and trained machine learning models using both experimental and CFD-generated data. Decision trees and ensemble classifiers achieved 100% accuracy in classifying filling process states, demonstrating that a hybrid physics-data approach can substantially reduce the computational expense of real-time CFD simulations while retaining high predictive fidelity. This work represents a methodologically exemplary approach to balancing model accuracy and computational tractability in hydraulic digital twins.

### 3.3. Smart Water Networks

The concept of smart water networks, in which distributed sensors, automated valves, and analytical platforms are integrated to optimize distribution system performance, maps naturally onto the digital twin paradigm. Recent contributions have demonstrated digital twins for real-time pressure zone management, demand-driven pump scheduling, water quality monitoring, and non-revenue water (NRW) reduction. Key technological enablers include advanced metering infrastructure (AMI), IoT-connected pressure loggers, and cloud-based hydraulic simulation platforms. Challenges specific to water utility digital twins include managing highly uncertain demand patterns, accurately characterizing distributed pipe roughness and condition, and maintaining system security in environments with constrained IT infrastructure.

Ramos et al. [27] proposed a Digital Twin (DT) model integrated with Smart Water Grid (SWG) concepts to improve monitoring, management, and efficiency in water distribution networks, specifically targeting water loss reduction. The methodology involved developing a DT for the Gaula water network in Portugal, which integrated GIS data, SCADA, and hydraulic modeling through a multi-stage process that calibrated the model with field pressure data, characterized losses via water balance, and sectorized the network into District Metered Areas (DMAs) to simulate scenarios and test interventions like pipe replacement and pressure valve optimization. The findings revealed that the existing system exhibited 80% water loss, with the DT model identifying excessive pressures and inadequate diameters as key issues, and the proposed interventions showed a potential 80% reduction in real water losses (434,273 m<sup>3</sup>), improving the infrastructure leakage index from 21.15 to near-recommended values, cutting the required supply volume by 60%, and yielding estimated savings of €165,000. Regarding practical considerations, implementation was found to rely heavily on accurate, integrated data from GIS and sensors, supporting proactive asset management and energy savings, though success required close collaboration between researchers and utility managers, with initial investment needed for sensors and model calibration, while social and environmental benefits remained unquantified monetarily.

Pandey et al. [10] developed a Machine Learning-based Digital Twin (DT) to enhance real-time leak detection and localization within Water Distribution Networks (WDNs). The methodology involved the creation of a virtual replica of the wastewater supply network at the Indian Institute of Science, Bangalore, integrated with IoT-enabled flow and pressure sensors. Researchers tested two distinct data-driven approaches: a single-stage model using Logistic Regression and Random Forest for direct classification, and a two-stage model that first predicted pressure differences via linear regression before classifying residual distributions. Findings revealed that while single-stage models provided direct detection, the two-stage approach successfully utilized posterior probabilities to improve localization accuracy under varying conditions. Practical considerations highlighted the importance of high-fidelity IoT data for real-time synchronization and the necessity of scalable ML architectures to handle the complexity of large-scale urban water infrastructure, ensuring operational resilience and reducing resource wastage.

Syed et al. [11] improved smart water management by integrating Digital Twin technology with multimodal transformer models to predict water usage and detect leakages. The researchers developed a framework that synchronizes real-time sensor data, pressure, flow, and thermal imaging, with a virtual water-network replica. Methodologically, the system utilizes the Informer transformer for long-term usage forecasting and a multimodal transformer that combines pressure-based LLM processing with Vision Transformers for leakage detection. Results show exceptional performance: the Informer achieved an  $R^2$  of 0.9995 and MSE of 2.2, while the leakage model reached 98.4% accuracy and a 0.0019 false-positive rate. Practically, the framework enables real-time monitoring, early anomaly detection, reduced water waste, and scalable deployment across urban, industrial, and future smart-city infrastructures.

Homaei et al. [12] developed an integrated DT platform for water distribution systems that leveraged IoT, AI/ML models, and cybersecurity measures to optimize operations, reduce environmental impacts, and ensure data security. The methodology involved designing the

CAUCCES platform, which aggregated historical and real-time water consumption data with meteorological variables, then implemented and compared advanced AI/ML models, including LSTM, Prophet, LightGBM, and XGBoost, after feature engineering and hyperparameter tuning, alongside formulating a Constraint Programming-based scheduling model for maintenance optimization and integrating cybersecurity measures aligned with ISO 27001. The findings demonstrated that the AI/ML models achieved a Mean Absolute Error of 5.76 and a Mean Absolute Percentage Error of 18.61% for six-month water consumption forecasting, while the Constraint Programming scheduling model achieved a 14% reduction in task completion time, a 25% reduction in delays, and 17% lower CO<sub>2</sub> emissions, with the cybersecurity implementation providing AES-128 encryption and real-time threat detection. Regarding practical considerations, the study found that successful DT implementation relied heavily on accurate, granular time-series data and robust feature engineering for model training, required careful integration with legacy systems and scalable architectures to handle diverse data sources, demanded a skilled workforce to manage advanced digital infrastructure, and necessitated comprehensive cybersecurity frameworks to protect interconnected systems while ensuring regulatory compliance with standards like GDPR and ISO 27001.

Ahanger et al. [13] addressed operational inefficiencies in wastewater management, including treatment delays, hydraulic overloads, and contamination events, by developing a DT-based real-time monitoring and forecasting framework. The framework integrated IoT-enabled sensing, EPANET–MATLAB hydraulic simulation, hybrid ANFIS-driven predictive intelligence, and a consortium blockchain for secure data auditability, and was evaluated using 80,114 samples from Kazhydromet and affiliated municipal wastewater facilities. The system achieved 89.3% precision, 88.1% sensitivity, and an F-measure of 88.7%, with near-real-time control latency of approximately 9.51 seconds and a Pearson correlation of  $r^2=0.89$ . The framework supported supervisory wastewater control through secure, asynchronous blockchain consensus that avoided disrupting time-critical decision loops, demonstrating strong potential for deployment in real-world municipal infrastructure.

## 4. Leak Detection in Gas and Water Pipelines

### 4.1. Significance and Detection Paradigms

Leak detection is perhaps the most intensively studied application of digital twin technology in pipeline engineering, reflecting the severity of potential consequences, financial losses, environmental contamination, explosion risk, and loss of life associated with undetected leakage events. Conventional leak detection methods, including volume balance, acoustic emission monitoring, negative pressure wave detection, and computational pipeline monitoring (CPM), each present trade-offs between sensitivity, specificity, localization accuracy, and operational cost. Digital twin-based approaches offer the potential to integrate multiple detection modalities within a unified model-based framework, improving overall detection performance.

Al-Ammari et al. [14] provided a robust DT model capable of overcoming the limitations of existing leak detection techniques in oil and gas pipelines, particularly their high costs and elevated false alarm rates. A comprehensive systematic review was conducted, evaluating the performance of existing leak detection methods and machine learning techniques used in DT development, which revealed that current models focused primarily on leak detection without adequately identifying leak size or location. A more comprehensive DT model was then proposed to detect pipeline abnormalities, including leaks, equipment failure, and damage, and was validated against a real-field gas pipeline case study. The results demonstrated that the developed DT model detected leaks with a zero false alarm rate and identified leak size and location with an absolute relative error of less than 3.21%. The framework served as a reference for future DT development in pipeline monitoring, offering a cost-effective and reliable approach to improving safety and flow assurance in gas pipeline infrastructure. The study identified three principal classes of digital twin architecture for leak

detection: physics-based model residual methods (which detect leaks as deviations between sensor measurements and model predictions), data-driven pattern recognition methods (which train classifiers on historical leak signatures), and hybrid methods that combine both approaches. The hybrid approach consistently demonstrated superior performance, particularly for small leak rates and noisy operating environments.

#### 4.2. *Reduced-Order Models and DTROT*

Cai and Wang [16] proposed and validated a digital twin-driven method for natural gas pipeline leakage detection. They first installed pressure sensors at the pipeline inlet and outlet, then constructed a high-fidelity virtual pipeline model. Using dynamic model order reduction techniques combined with measured pressure data, they developed a reduced-order model (ROM) that rapidly predicted outlet pressure under multiple operating conditions. An optimal leak alarm threshold was selected based on the probability of detection (POD) versus probability of false call (POFC) curve. The method was experimentally validated on a dedicated test bench at leak rates of 1%, 1.5%, 2%, and 3% under both normal and noisy conditions. Results demonstrated that the method achieved an average detection accuracy exceeding 90% for leak rates from 1% to 3%, with zero false alarms at rates of 1.5%, 2%, and 3%, while maintaining robust performance in noisy environments. This work applied Digital Twin Reduced-Order Technology (DTROT), creating virtual pipeline mappings that integrate sensor data, pressure, flow, and acoustics, and use Reduced-Order Models (ROMs) for efficient simulation and prediction of leakage events. The ROM approach preserves the physical fidelity of high-dimensional CFD or finite-element simulations while enabling real-time execution, addressing a fundamental bottleneck in the deployment of physics-based digital twins in operational settings.

#### 4.3. *Multiphase Flow and Visual Twins*

Real-world gas transmission pipelines frequently carry multiphase or impure gas streams, complicating leak detection relative to idealized single-phase scenarios. Al-Ammari et al. [15] presented a digital and visual twin system for real-time leak detection and localization in gas pipelines operating under multiphase flow conditions. The researchers leveraged experimental data from a multiphase flow-testing loop alongside synthetic data generated using OLGAs simulation software to train and optimize several machine learning models, including random forest, support vector machine, k-nearest neighbors, decision tree regression, and XGBoost. Individual models initially demonstrated moderate performance, with accuracy ranging from 42% to 57%, but significant improvements were achieved through stacking ensembles, feature engineering, and data averaging techniques. The integration of experimental and OLGAs-generated synthetic data proved essential for training robust machine learning models, ensuring high accuracy in predicting leak characteristics. The k-nearest neighbors model emerged as the most effective classifier, achieving an accuracy of 98.6% and an F1-score of 0.985, while random forest regression led regression tasks with  $R^2$  values exceeding 0.99 for both leak size and location predictions. The real-time visual twin platform effectively mapped machine learning outputs onto a three-dimensional pipeline representation using gaming engine technology, enhancing operator awareness and enabling immediate response actions. Advanced feature engineering and data averaging techniques further improved model robustness under complex multiphase conditions, setting a new benchmark for offshore pipeline monitoring systems and demonstrating strong potential for broader industrial applications.

#### 4.4. *Data-Driven and Hybrid Approaches*

Earlier foundational work by Liang et al [17], proposed a data-driven digital twin (DT) method for leak detection in natural gas pipelines, addressing challenges such as limited abnormal data and dynamic operating conditions. A high-fidelity DT pipeline was constructed using deep learning

techniques, specifically an Autoencoder-Convolutional Neural Network (AE-CNN) model, trained on normal operational data to simulate the behavior of physical pipelines under standard conditions. The DT pipeline enabled real-time monitoring, iterative updates, and anomaly detection by comparing the virtual model's output with the physical pipeline's data. A case study on a 19 km natural gas pipeline in southwest China validated the approach, demonstrating its effectiveness in detecting leaks of varying sizes. The method achieved high accuracy (97.23%), fault detection rate (90.77%), and low false alarm rate (0.88%) across all leak sizes. Performance tests showed robustness against data distortion and stable operation under production adjustments, with average accuracy exceeding 94%. The study concluded that the DT-driven method effectively enhanced leak detection and adaptability in dynamic environments, reducing false positives and improving detection rates. Future research will focus on expanding abnormal sample databases, designing dynamic threshold intervals, and exploring transfer learning techniques to adapt DT models to other pipelines.

Wang et al. [18] investigated a gas pipeline leakage identification method driven by digital twin technology, aiming to improve detection accuracy and real-time monitoring. A digital twin model was constructed to simulate pipeline behavior using pressure signals, integrating an online update system, a visualization model, and a leakage identification model based on a support vector machine (SVM). Finite element analysis and fluid dynamics simulations were performed to generate sample data, from which feature vectors were extracted to classify pipeline conditions. The system continuously mapped real-time physical data to the virtual model, enabling dynamic monitoring and rapid fault recognition. Experimental results showed that the proposed method achieved high accuracy, with simulation tests reaching about 95% classification accuracy and laboratory experiments achieving around 90.5%. Comparative analysis indicated that SVM outperformed other algorithms in balancing accuracy and efficiency. Additionally, leakage detection accuracy across multiple points averaged over 91%, demonstrating robustness. The conclusion stated that the digital twin-driven approach was effective and feasible for pipeline leakage detection, offering reliable real-time prediction and visualization. It highlighted that integrating digital twin technology with machine learning provided a strong foundation for future advancements in intelligent pipeline monitoring systems.

#### *4.5. Industry Deployment: Visual Twins and MLOps*

The practical deployment of digital twins for pipeline leak monitoring in industrial settings was documented by Hamilton et al. [19]. The authors described a visual digital twin system applying MLOps principles, continuous integration, model versioning, and performance monitoring, to both train pipeline operators and provide situational awareness for leak detection. The system combined experimental data and simulation to predict leak plumes and pipeline flow and demonstrated measurable improvements in operator response time and detection reliability relative to conventional CPM systems. The study developed a visual digital twin system designed to enhance the operation and training of data-driven leak detection for subsea pipelines. Researchers utilized gaming engine technology to create a high-fidelity virtual environment that integrated real-time data streams from physical sensors, including pressure and temperature monitors. To address the challenge of monitoring large-scale infrastructure, the system provided multiple camera views and a multi-scale interface that allowed operators to toggle between kilometer-scale overviews and localized sensor subsections. The framework incorporated machine learning-driven detection algorithms to identify anomalies and provide instantaneous visual feedback within the 3D environment. Results demonstrated that the integration of live data into a gaming engine significantly improved situational awareness and reduced operator cognitive load. Ultimately, the system established a robust platform for training personnel to respond to simulated leak scenarios while maintaining real-time oversight of offshore deployments.

## 5. Subsea Pipeline Monitoring and Structural Integrity

### 5.1. The Subsea Challenge

Subsea pipelines occupy a uniquely demanding operational environment characterized by high hydrostatic pressures, low temperatures, corrosive seawater exposure, complex seabed geomorphology, and extreme remoteness from human operators. These conditions intensify the challenge of inspection and monitoring while amplifying the consequences of failure, both economically, given the cost of deepwater repair operations, and environmentally, given the risk of hydrocarbon release in sensitive marine ecosystems. Digital twin technology is particularly compelling for subsea pipelines precisely because it offers a means of extending the reach of human engineering intelligence into environments where direct human access is prohibitively costly or hazardous.

A foundational scoping study by Chen et al. [20], systematically examined opportunities and challenges for developing digital twins of subsea pipelines. They reviewed the opportunities and challenges associated with developing digital twins for subsea pipelines across design, construction, service life, and life extension assessments. The authors described digital twins as a paradigm combining multi-physics modelling with data-driven analytics to mirror the life of physical assets. They identified key opportunities for improved integrity management, including data contextualization, standardization, automated anomaly detection, and learning through data sharing. The paper discussed the automated creation of digital twins during construction using LiDAR and sensor platforms, as well as updates based on in-line inspections and machine learning algorithms for defect detection. Predictive maintenance strategies utilizing IoT sensors and artificial neural networks were highlighted for estimating remaining fatigue life. Major challenges included data collection and interpretation, cyber-security, limited access to proprietary data, and the need for universal platforms and standardized data formats. The authors concluded that while digital twins offer significant potential for real-time monitoring and life extension, coherent research efforts are required to unify standards and develop user-friendly tools.

### 5.2. Sensor Architectures and Inspection Technologies

Wegner et al. [21] proposed a comprehensive framework for subsea pipeline monitoring that integrates fixed distributed sensor systems with mobile inspection platforms. The framework employed distributed sensors, including pressure, temperature, strain, vibration, and cathodic protection instruments, deployed along the pipeline. It was complemented by remotely operated vehicles (ROVs) and autonomous underwater vehicles (AUVs) for high-resolution, periodic inspections using multibeam sonars, laser scanners, and advanced optical imaging systems. The digital twin aggregates data from these heterogeneous sources to maintain a continuously updated structural and hydraulic model of the pipeline. The findings demonstrated that the proposed DT enabled continuous monitoring, accurate integrity assessment, and predictive maintenance through early detection of high-risk zones and failure mechanisms. From a practical perspective, the framework supported proactive decision-making, optimized inspection scheduling, enhanced situational awareness through GIS-based visualization, and shifted pipeline management from reactive to predictive strategies, thereby improving safety, regulatory compliance, and operational efficiency.

### 5.3. Corrosion Monitoring via ILI-MFL

Corrosion represents the leading cause of subsea pipeline integrity loss and is notoriously difficult to characterize non-destructively. Syuryana et al. [22] described a digital twin for subsea pipeline corrosion monitoring using In-Line Inspection based on Magnetic Flux Leakage (ILI-MFL) data combined with potentiodynamic electrochemical testing on API 5L-X42 pipeline steel. The twin integrated ILI signal interpretation, metallurgical characterization, and corrosion rate modeling

enable mapping of corrosion morphology across the pipeline and projection of remaining wall thickness over time, a capability directly relevant to fitness-for-service assessment and maintenance planning. The study utilized magnetic flux leakage (MFL) data, categorized according to Pipeline Operator Forum (POF) standards, to identify and classify anomalies such as pitting and grooving. To bridge the gap between periodic inspections and real-time degradation, the team performed potentiodynamic polarization tests on JIS S45C steel specimens in a 3.5% NaCl environment, which provided precise corrosion rate measurements ranging from 0.0068 to 0.1730 mmpy. These experimental results were then synchronized with a virtual representation of the physical asset, creating a dynamic simulation model. This integrated approach allowed for a more accurate estimation of the pipeline's remaining strength and service life compared to traditional reactive methods. Ultimately, the digital twin successfully demonstrated a proactive capability for predicting metal loss and optimizing maintenance schedules for difficult-to-access subsea infrastructure.

#### 5.4. Underwater Digital Twin Landscape

A broad systematic literature review of underwater digital twin applications by Kaarlela et al [23], identified structural health monitoring of underwater infrastructure, prediction of remaining useful life, and detection of corrosion, fatigue, and damage using data-driven approaches as the principal research themes. The review noted a rapid acceleration in publication rates since 2020 and identified significant opportunities for multi-physics simulation, digital twin standardization, and the integration of uncertainty quantification methods as priorities for future research. The study examined peer-reviewed publications across engineering, oceanography, robotics, and data-science databases, applying predefined inclusion criteria and thematic coding to classify UDT purposes, architectures, sensing modalities, and computational methods. The review found that UDT research had primarily focused on structural monitoring, subsea infrastructure inspection, autonomous underwater vehicle (AUV) coordination, and environmental modeling. Most implementations relied on physics-based simulations combined with real-time sensor fusion, though challenges in underwater communication, energy constraints, and sparse data significantly restricted model fidelity. The study also observed that machine learning integration remained limited, with few works addressing long-term predictive maintenance or adaptive control. Practical considerations highlighted the need for standardized frameworks, improved acoustic and optical sensing, scalable cloud-edge computation, and robust validation datasets. The authors concluded that advancing UDTs required interdisciplinary collaboration and investment in next-generation underwater cyber-physical systems.

## 6. Condition Monitoring and Predictive Maintenance

### 6.1. The Case for Predictive Pipeline Management

Traditional pipeline maintenance strategies, whether calendar-based or run-to-failure, are demonstrably suboptimal in both cost and safety terms. Condition-based maintenance (CBM) and its predictive extension, predictive maintenance (PdM), seek to optimize maintenance timing by acting only when and where asset condition warrants intervention. Digital twins provide the synthetic environment in which sensor data, physics-based degradation models, and statistical learning algorithms can be fused to produce actionable predictions of remaining useful life (RUL) and optimal maintenance windows.

A systematic review of digital twin-driven predictive maintenance across the industrial engineering sector conducted by Ismail et al. [24] noted that digital twin applications in predictive maintenance include using the twin to predict transmission volume in pipeline networks, then comparing the predicted against actual volume to identify anomalies indicative of leakage, blockage, or other abnormal conditions. The study analyzed prior research to trace the temporal development of digital twins from early virtual replicas to AI enabled, self learning systems, and it synthesized applications, middleware components, and technological requirements across industrial domains.

The review identified a layered digital twin architecture and proposed a taxonomy covering system types, enabling technologies, and artificial intelligence algorithms used for failure forecasting and maintenance optimization. Findings showed that increasing system complexity and the rise of IoT, machine learning, and real-time analytics had driven the need for more adaptive predictive maintenance solutions. The authors concluded that trustworthy, efficient digital twin ecosystems required standardized architectures, robust data integration, and further research into autonomous, self updating models for industrial operations. The review found that approximately 50% of digital twin predictive maintenance implementations are concentrated in manufacturing, with approximately 15% in the energy industry, suggesting significant room for accelerated adoption in pipeline-intensive oil, gas, and water utility sectors.

### 6.2. Ensemble Kalman Filter Data Assimilation

A methodologically significant contribution to pipeline condition monitoring was published by Wang et al. [6]. This study presented a digital twin system for pipeline condition monitoring, integrating key technologies such as the Internet of Things, advanced analytics, cloud computing, and augmented reality (AR). The digital twin was developed as a high-fidelity virtual replica of individual physical pipelines, enabling real-time monitoring, probabilistic simulation, damage detection, localization, quantification, and prediction. A case study focused on fatigue crack growth in S235 steel pipes under cyclic loading, utilizing guided-wave sensor networks for defect detection and Bayesian inference for real-time updates. The system incorporated physics-based models, finite element analysis, surrogate modeling, and Dirichlet process mixture models to account for uncertainties in material properties, damage detection, and operational conditions. The digital twin was validated through experimental data, demonstrating its ability to predict pipeline reliability and optimize maintenance activities. AR-based visualization using head-mounted displays facilitated intuitive interaction with the digital twin, allowing operators to monitor pipeline health without on-site inspections. Future work was proposed to expand the system's capabilities for comprehensive damage monitoring and maintenance optimization.

### 6.3. Corrosion Digital Twins and Unsupervised Learning

Corrosion monitoring via digital twin was further advanced by Bhowmik [25]. developed a digital-twin-based framework to improve corrosion monitoring in offshore pipelines by integrating sensor data with deep-learning-driven predictive models. The authors constructed a virtual replica of pipeline segments that continuously synchronized with real-time inspection and environmental data to estimate corrosion progression. Their methodology combined physics-based corrosion modeling with convolutional neural networks trained on historical inspection records, enabling automated detection of wall-thickness loss and prediction of future degradation patterns. The system was validated using offshore pipeline datasets, where the deep-learning model demonstrated high accuracy in identifying corrosion hotspots and forecasting remaining useful life under varying operational conditions. The results showed that the digital twin enhanced situational awareness, reduced manual inspection demands, and supported risk-informed maintenance planning. The study concluded that integrating deep learning into digital-twin architectures offered a scalable and cost-effective approach for offshore asset integrity management, while highlighting the need for improved data quality and long-term field validation.

Additionally, Duan et al. [26] proposed a digital twin model for real-time structural health monitoring (SHM) and damage detection in underground pipelines rehabilitated with cured-in-place pipe (CIPP) liners. The framework integrated distributed fiber optic sensing (DFOS), finite element analysis (FEA), and deep learning (DL) algorithms to monitor strain fields, predict mechanical responses, and identify structural damage. Experimental tests were conducted on steel substrate samples and liner-rehabilitated steel substrates under eccentric compression to investigate buckling behavior and damage mechanisms. DFOS enabled high-resolution, continuous strain monitoring, while FEA simulated structural responses and validated experimental results. A convolutional neural

network (CNN) was employed to correlate DFOS-measured and FEA-predicted strain fields, facilitating the development of a real-time digital twin framework. The study successfully identified and localized plastic deformation and adhesive cracks, providing insights into damage progression and rehabilitation performance. The DFOS-integrated liner demonstrated its capability to monitor strain and detect damage, supporting predictive maintenance and enhancing pipeline reliability. The findings highlighted the potential of the digital twin framework for scalable, intelligent SHM applications in underground pipeline infrastructure, with future work focusing on full-scale validation under real-world conditions.

#### 6.4. Integration with Industrial IoT and Cloud Platforms

The practical deployment of condition monitoring digital twins increasingly leverages industrial IoT (IIoT) platforms and cloud computing infrastructure. Edge computing nodes installed along pipeline corridors perform preliminary signal processing and anomaly detection, reducing bandwidth requirements for data transmission to central analytical platforms. Cloud-based twins provide the computational resources necessary for high-fidelity simulation, ensemble modeling, and long-horizon prognostics, while mobile and web-based dashboards surface actionable insights to maintenance engineers and asset managers. The integration of digital twins with enterprise asset management (EAM) and computerized maintenance management systems (CMMS) closes the loop from condition assessment to maintenance execution, enabling fully automated maintenance workflow orchestration in progressive implementations.

## 7. Cross-Cutting Themes, Challenges, and Future Directions

### 7.1. Common Methodological Themes

Across the five-pipeline engineering sub-domains reviewed, several methodological themes recur with notable consistency. First, the hybrid physics-data approach, in which first-principles simulation provides the structural backbone of the twin while machine learning components augment physical knowledge and accelerate computation, has emerged as the dominant architectural paradigm, consistently outperforming purely physics-based or purely data-driven alternatives. Second, real-time data assimilation, particularly through sequential Bayesian methods such as the Ensemble Kalman Filter, has proven essential for maintaining twin fidelity under operational variability and sensor noise. Third, reduced-order modelling has become a critical enabler for real-time execution of computationally intensive physical models. Fourth, visual twin components and human-machine interface design have emerged as important differentiators for operational adoption, reflecting recognition that technical performance alone is insufficient if operators cannot effectively interpret and act on twin outputs.

### 7.2. Proposed Reference Architecture for Pipeline Digital Twins

To address the fragmentation observed across existing digital twin (DT) implementations in pipeline engineering, this study proposes a unified reference architecture that integrates the diverse modeling paradigms, sensing strategies, and application domains discussed throughout the preceding sections. Although the literature reports a wide range of domain-specific DT solutions, a generalized structural framework is required to support scalability, interoperability, and systematic deployment across pipeline systems.

The proposed architecture is organized into six interdependent layers, each representing a fundamental functional component of a pipeline DT system: (i) physical asset layer, (ii) sensing and inspection layer, (iii) data and communication layer, (iv) model layer, (v) synchronization and assimilation layer, and (vi) decision and application layer. Together, these layers form a closed-loop cyber-physical framework capable of real-time monitoring, predictive analysis, and adaptive decision-making. As illustrated in Figure 1, the architecture captures the bidirectional relationship

between the physical pipeline system and its digital counterpart, where field measurements continuously inform the virtual model, and updated digital intelligence supports operational and maintenance actions.

The physical asset layer represents the real pipeline system, including the pipe structure, transported fluid, surrounding soil or seabed, and associated infrastructure such as valves, compressors, and supports. This layer defines the actual operating state of the system and includes key physical processes such as pressure variation, temperature change, flow behavior, mechanical loading, corrosion, and fatigue. In offshore and subsea settings, this layer also includes interaction with seabed morphology, hydrodynamic loading, and other environmental effects.

The sensing and inspection layer provides the interface between the physical system and its digital representation. It includes distributed and discrete sensing technologies such as Internet of Things (IoT) sensors, SCADA systems, distributed acoustic sensing (DAS), distributed fiber optic sensing (DFOS), and in-line inspection (ILI) tools. In subsea applications, this layer may also incorporate remotely operated vehicles (ROVs) and autonomous underwater vehicles (AUVs) equipped with sonar and imaging systems. The quality, density, and reliability of this layer directly influence system observability and, consequently, the fidelity of the digital twin.

The data and communication layer is responsible for transmitting, storing, and managing data acquired from the sensing and inspection layer. This includes edge computing units for preliminary signal processing, communication networks for real-time data transfer, and cloud-based or on-premises platforms for storage and analytics. Data integration at this level must address heterogeneity, latency, missing information, and cybersecurity concerns. Robust data governance and standardized communication protocols are therefore essential for ensuring reliable interoperability across distributed assets and operating environments.

The model layer forms the computational core of the digital twin, where physical and data-driven representations of the pipeline system are developed. This layer includes high-fidelity physics-based models such as computational fluid dynamics (CFD) and finite element methods (FEM), reduced-order models (ROMs) for real-time approximation, and machine learning models for pattern recognition, anomaly detection, and forecasting. Hybrid approaches, which combine physics-based and data-driven methods, have emerged as the dominant modeling paradigm because they balance interpretability, predictive accuracy, and computational efficiency. In pipeline applications, this layer may represent multiphase flow behavior, structural response, corrosion progression, and other degradation mechanisms.

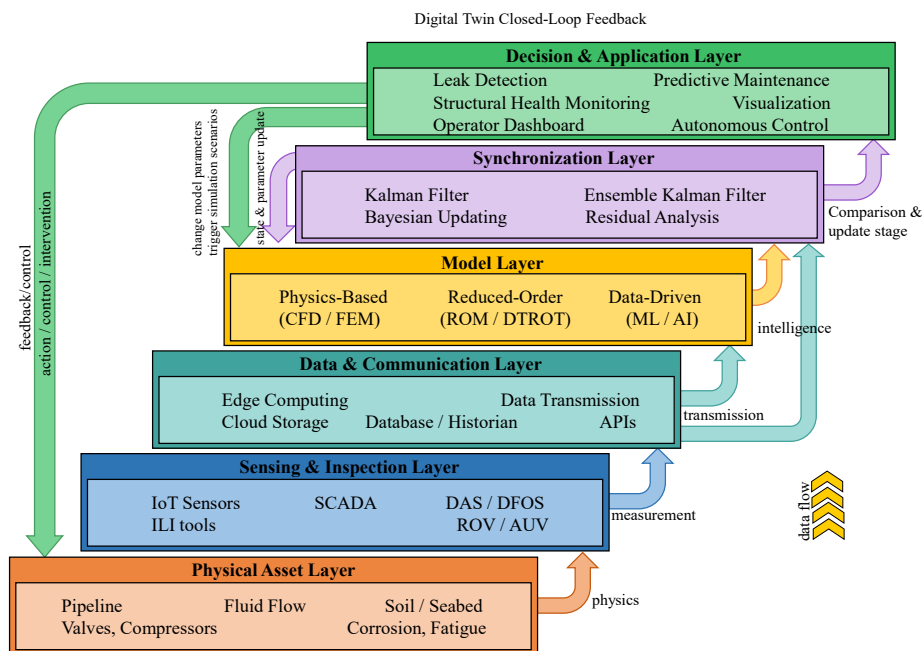
The synchronization and assimilation layer ensures continuous alignment between the physical asset and the digital model. It integrates incoming field data with model predictions using data assimilation techniques such as Kalman filtering, Ensemble Kalman filtering, and Bayesian inference. Its primary role is to update system states and model parameters in real time, thereby reducing uncertainty and maintaining model fidelity under changing operational conditions. This layer is essential because it transforms the digital twin from a static or offline model into a continuously evolving representation of the physical system.

The decision and application layer represents the functional output of the digital twin. It includes applications such as leak detection, structural health monitoring, corrosion assessment, predictive maintenance, and operational optimization. This layer also incorporates visualization environments, operator dashboards, and decision-support interfaces. In more advanced implementations, it may evolve toward prescriptive and autonomous systems capable of recommending or executing interventions based on predicted system behavior.

This layered architecture also provides a unifying lens through which the diverse studies reviewed in this paper can be interpreted. Sections 2 and 3 primarily contribute to the physical and model layers through system-level and hydraulic modeling approaches. Section 4 focuses on leak detection applications within the decision layer, supported by both physics-based and data-driven techniques. Section 5 addresses subsea sensing, inspection technologies, and structural integrity modeling, which are central to the sensing and model layers. Section 6 contributes to the

synchronization and decision layers through predictive maintenance and data assimilation strategies. Section 7, in turn, integrates these domain-specific insights into a cross-cutting systems perspective.

By structuring pipeline DT systems within this layered framework, it becomes possible to systematically evaluate performance limitations, identify integration bottlenecks, and guide the development of scalable, interoperable, and reliable digital twins for pipeline engineering. The proposed reference architecture also provides a conceptual foundation for future standardization efforts and supports the transition from isolated DT applications to integrated, system-level cyber-physical infrastructures.



**Figure 1.** Architecture for Pipeline Digital Twin Systems.

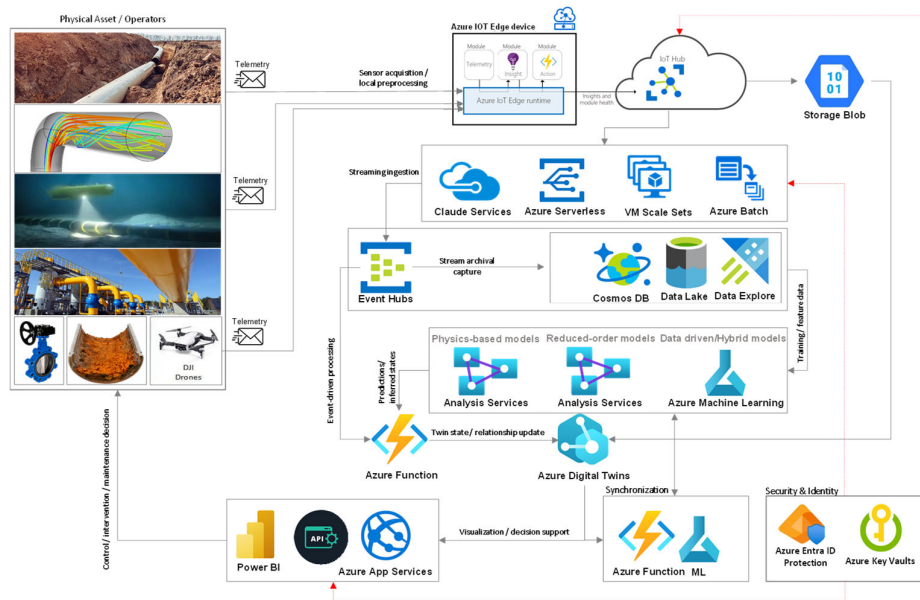
The reference architecture shown in Figure 1 is not intended to replace the domain-based organization of this review; rather, it complements it by providing a common systems-level structure that connects the diverse applications discussed throughout the paper. In this sense, the framework serves as an interpretive backbone for the remainder of Section 7, particularly when examining cross-cutting challenges related to interoperability, model fidelity, synchronization, and future DT deployment pathways. While the preceding discussion presents a platform-agnostic conceptual architecture, it is also valuable to illustrate how such a framework can be implemented using modern cloud-based infrastructures.

To complement the conceptual architecture presented in Figure 1, an implementation-oriented system architecture based on representative cloud services is illustrated in Figure 2. This figure provides a practical mapping of digital twin components to a real-world deployment framework, demonstrating how sensing, data management, modeling, and decision-making functionalities can be integrated within a cloud-based ecosystem. In this architecture, physical pipeline assets and their surrounding environment are monitored using a combination of field sensors, inspection tools, and remote platforms such as drones and subsea vehicles. These data streams are first processed at the edge to enable local filtering and preliminary analytics, before being transmitted to cloud services through IoT-based communication frameworks.

Within the cloud environment, high-throughput streaming services facilitate real-time data ingestion, while storage components support both short-term operational analysis and long-term historical data archiving. Analytical and machine learning services are then employed to extract features, train predictive models, and generate inferred system states. In parallel, physics-based and reduced-order models provide mechanistic representations of pipeline behavior, enabling

simulation-driven insight into flow dynamics, structural integrity, and degradation processes. These modeling components are integrated within a digital twin framework, where synchronization mechanisms, implemented through event-driven processing and data assimilation techniques, continuously update the virtual representation of the system.

The digital twin platform serves as a central hub that maintains the current state of the pipeline system, incorporating both measured and predicted information. Outputs from the digital twin are delivered to visualization and decision-support interfaces, enabling operators to monitor system performance, detect anomalies, and plan maintenance actions. Importantly, architecture supports a closed-loop feedback mechanism in which operational decisions, control actions, or maintenance interventions are fed back to the physical system, thereby completing the digital twin cycle. This implementation view reinforces the layered conceptual framework by demonstrating how digital twin principles can be operationalized using scalable, cloud-based technologies in modern pipeline engineering applications.



**Figure 2.** Azure-Based System-Level Architecture for Pipeline Digital Twin Implementation.

This implementation-oriented view bridges the gap between the conceptual architecture and practical deployment, demonstrating how digital twin systems can be realized using scalable cloud-based infrastructures.

### 7.3. Mathematical Formulation of Digital Twin Synchronization

To complement the proposed reference architecture, a compact mathematical formulation is presented to describe the fundamental mechanism underlying digital twin synchronization. While digital twins are often described conceptually, their core functionality can be rigorously interpreted through a state-space representation that captures the evolution of system states and their relationship with observable measurements. In a general form, the pipeline system can be expressed as:

$$x_{t+1} = f(x_t, u_t, \theta) + \omega_t \quad (1)$$

$$y_t = h(x_t) + v_t \quad (2)$$

where  $x_t$  represents the system state at time  $t$ ,  $u_t$  denotes external inputs or operating conditions,  $\theta$  is a set of model parameters,  $y_t$  is the measured output, and  $\omega_t$  and  $v_t$  represent process and measurement uncertainties, respectively.

In the context of pipeline engineering, the state vector  $x_t$  may include quantities such as pressure distribution, flow rate, temperature, structural stress, corrosion depth, or crack size, depending on the application. The input vector  $u_t$  typically includes operational and environmental variables such as inlet pressure, flow demand, valve settings, and external loads. The parameter set  $\theta$  characterizes system properties, including friction factors, material properties, corrosion growth rates, and soil or seabed stiffness. The process noise  $\omega_t$  captures model uncertainties and unmodeled dynamics, while the measurement noise  $v_t$  accounts for sensor inaccuracies and data acquisition errors.

Within this framework, the synchronization and assimilation layer introduced in Section 7.2 can be interpreted as the mechanism that reconciles model predictions with real-world observations. Specifically, the model generates a predicted state and corresponding outputs, which are then compared with measured data  $y_t$ . The discrepancy between predicted and observed quantities, often referred to as the residual, is used to update the internal state and, when necessary, model parameters. This process ensures that the digital twin remains aligned with the evolving physical system.

Data assimilation techniques such as Kalman filtering, Ensemble Kalman filtering, and Bayesian inference provide systematic approaches for performing this update. These methods estimate the most probable system state by combining prior model predictions with incoming measurements while accounting for uncertainty. As a result, the digital twin transitions from a static or open-loop model to a dynamic, closed-loop system that continuously adapts to real operating conditions.

This mathematical perspective directly supports the layered architecture shown in **Error! Reference source not found.** The model layer corresponds to the functions  $f(\cdot)$  and  $h(\cdot)$ , the data and communication layer provide the measurements  $y_t$ , and the synchronization layer performs the update of  $x_t$  and  $\theta$ . The decision and application layer then utilizes the updated state estimates to support tasks such as leak detection, predictive maintenance, and operational optimization. In this way, the mathematical formulation provides a unifying description of how information flows and evolves within a pipeline digital twin system.

#### 7.4. Comparative Analysis of Pipeline Digital Twin Studies

Despite the growing number of studies on digital twin applications in pipeline engineering, a systematic comparison of their underlying methodologies, data requirements, and operational capabilities remains limited. To address this gap, Table 3 presents a comparative analysis of representative studies across different pipeline domains, including oil and gas systems, hydraulic networks, leak detection applications, and subsea monitoring.

The table synthesizes key attributes of each study, including the adopted digital twin approach, data sources, real-time capabilities, primary application focus, and reported strengths and limitations. This structured comparison provides a clearer understanding of how different modeling paradigms, such as physics-based, data-driven, and hybrid approaches, are utilized in practice, and highlights the trade-offs associated with each.

**Table 3.** Comparative Analysis of Representative Studies across Different Pipelines.

Study	Pipeline Domain	DT Approach	Data Source	Real-Time	Primary Application	Key Strength	Main Limitation
Ahanger et al. [13]	Water network	Hybrid (Hydraulic + ANFIS + Blockchain)	IoT sensing + municipal wastewater dataset	Partial	Monitoring & forecasting (overload, contamination)	High accuracy ( $r^2=0.89$ ), secure data traceability	Requires validation across broader systems Limited real-field deployment; reliance on controlled experiments and synthetic data; computational complexity Limited real-field validation; dependence on simulated data; scalability not demonstrated
Al-Ammari et al. [14]	Gas pipeline	Hybrid (ML + visual twin + simulation)	Experimental (flow loop) + synthetic (OLGA)	Yes	Leak detection & localization (single & multiple leaks)	High accuracy (up to ~99%), robust under multiphase flow, integrated visualization platform	Requires long-term validation
Al-Ammari et al. [15]	Gas pipeline	Hybrid DT (data-driven + simulation-based)	Simulated (OLGA) + limited field/experimental validation	Partial	Leak detection, localization & diagnostics	High accuracy (<3.21% error), zero false alarms, leak size & location estimation	Requires long-term validation
Bhowmik [25]	Subsea pipeline	Hybrid (Corrosion physics + CNN)	Sensor + inspection + environmental data	Yes	Corrosion prediction & RUL	Accurate hotspot detection, scalable	Requires long-term validation
Cai & Wang [16]	Gas pipeline	Reduced-order (DTROT)	Pressure sensors + flow/acoustic data	Yes	Leak detection	>90% accuracy, real-time capability	Reduced fidelity in complex conditions Data standardization & security challenges
Chen et al. [20]	Subsea pipeline	Hybrid (Multi-physics + ML)	LiDAR + ILI + IoT sensors	Yes	Integrity & fatigue prediction	Automated DT lifecycle integration	Needs full-scale validation
Duan et al. [26]	Underground pipeline	Hybrid (FEA + CNN + DFOS)	DFOS + simulation data	Yes	Structural health monitoring	High-resolution strain monitoring Integrated visual twin with DT MLOps enabling explainable ML, experimental validation, and enhanced operator decision support	Limited to lab-scale validation; lacks real-world deployment and large-scale field validation
Hamilton et al. [19]	Subsea pipeline	Visual digital twin + ML + DT MLOps (digital thread integration)	Experimental multi-modal data (pressure, acoustic, video) + processed data streams	Near real-time	Leak detection, localization, plume prediction, and operator training	High accuracy, real-time capability	Requires high-quality data Low industrial adoption (Lack of standardized frameworks; limited real-world validation)
Homaei et al. [12]	Water network	Data-driven (LSTM, XGBoost)	IoT + historical + weather data	Yes	Demand forecasting & optimization	Operational efficiency, reduced delays	Requires high-quality data Low industrial adoption (Lack of standardized frameworks; limited real-world validation)
Ismail et al. [24]	Multi-domain	Hybrid (AI-enabled DT framework + architectural taxonomy)	Systematic review (multi-database literature)	Conceptual (real-time requirement emphasized)	Predictive maintenance & anomaly detection	Comprehensive taxonomy	Requires high-quality data Low industrial adoption (Lack of standardized frameworks; limited real-world validation)

Kaarlela et al. [23]	Subsea infrastructure	Hybrid (Simulation + sensor fusion)	Sparse underwater sensing	Yes	SHM & RUL	Comprehensive classification	Data sparsity, communication limits
Liang et al. [17]	Gas pipeline	Data-driven (AE-CNN)	Operational data	Yes	Leak detection	High accuracy (~97%)	Limited abnormal data
Pandey et al. [10]	Water network	Data-driven (ML ensemble)	IoT sensors	Yes	Leak detection & localization	Improved localization accuracy	Requires scalable ML infrastructure
Ramos et al. [27]	Water network	Hybrid (Hydraulic + GIS + SCADA)	GIS + SCADA + sensors	Partial	Water loss reduction	Significant loss reduction	Requires integrated data systems
Syed et al. [11]	Water network	Data-driven (Transformers)	Real-time sensor data (pressure, flow, thermal)	Yes	Forecasting & anomaly detection	High accuracy ( $R^2 \approx 0.999$ ) + multimodal data fusion + real-time DT integration	High system complexity; reliance on multimodal sensors and computational cost
Syuryana et al. [22]	Subsea pipeline	Hybrid (ILI-MFL + electrochemistry)	Inspection + lab testing	Partial	Corrosion prediction	Accurate degradation estimation	Not fully real-time
Wang et al. [18]	Gas pipeline	Data-driven (SVM)	Pressure + simulation data	Yes	Leak detection	Robust real-time performance	Reduced field accuracy
Wang et al. [6]	Pipeline (fatigue)	Hybrid (FEM + Bayesian + ML)	IoT + experimental data	Yes	Damage & reliability prediction	Uncertainty-aware modeling	Needs broader validation
Wegner et al. [21]	Subsea pipeline	Hybrid (physics-based + probabilistic + data-driven analytics)	Multi-source data (ROV imagery, sensors, NDT inspection, environmental & geohazard data)	Yes	Integrity management, predictive maintenance & risk-based decision support	Integrated multi-layer DT architecture + data fusion + predictive & prescriptive analytics	Conceptual framework; lacks real-world validation and quantitative performance evaluation
Conejos Fuertes et al. [5]	Water network	Hybrid (Hydraulic + real-time calibration)	Sensor + consumption data	Yes	Optimization & leak detection	Large-scale deployment	Requires continuous calibration
Li et al. [8]	Water system	Hybrid (GIS-BIM + simulation)	Multi-source + real-time data	Yes	Hydrodynamic & risk prediction	High simulation fidelity	High computational demand
Paternina-Verona et al. [9]	Hydraulic pipeline	Hybrid (CFD + ML)	Experimental + CFD data	Yes	Transient flow analysis and pressure surge prediction	High-fidelity CFD-based DT capturing air-water interactions + ML reducing computational cost with high predictive accuracy	High computational cost of CFD; validated at laboratory scale; limited direct real-time deployment without ML acceleration
Grieves & Vickers [1]	General DT concept	Hybrid conceptual	Real-time sensor data	Yes	Lifecycle management	Foundational DT concept	Interoperability challenges
Kritzinger et al. [2]	Manufacturing systems (Industry 4.0 context)	Conceptual (categorical literature review + DT integration-level classification)	Literature-based (no experimental or real-time data)	Not implemented (defines levels: DM – no flow, DS – one-way, DT – two-way real-time)	DT classification	Clear DM-DS-DT framework	Lack of real implementations, validation, and quantitative evaluation

Tao et al. [3]	General DT concept	Conceptual (five-dimensional DT modeling framework)	Multi-source (physical, virtual, service, and knowledge data)	Yes	Digital twin modeling, system representation, and lifecycle integration	Introduces five-dimensional DT model, enabling data fusion and service-oriented DT architecture	Conceptual framework without real-world implementation or quantitative validation
Fuller et al. [4]	Multi-domain (manufacturing, healthcare, smart cities)	Conceptual review (definitions, classification, enabling technologies)	Literature review (multi-domain studies)	Conceptual (supports real-time DT, not implemented)	DT definitions, applications, challenges, and enabling technologies	Comprehensive overview	Terminology inconsistency
Hamidishad et al. [7]	Oil & Gas	Multi-fidelity framework (high-fidelity simulation + hybrid ML + reduced-order + co-simulation)	Systematic literature review	Conceptual + architectural (real-time monitoring, edge-cloud DT systems)	Energy optimization, Process optimization, energy efficiency, lifecycle management, and predictive analytics in O&G processing system	Provides a modular, multi-fidelity DT architecture integrating high-fidelity simulation, hybrid ML models, and real-time data frameworks for O&G systems	Needs standardization, lack of long-term validation, and cybersecurity challenges in large-scale DT

As shown in Table 3, no single digital twin approach satisfies all performance requirements simultaneously. Physics-based models provide high interpretability and physical consistency but are often computationally intensive and difficult to deploy in real-time applications. In contrast, data-driven approaches offer fast inference and real-time capability but may lack physical transparency and generalization under unseen conditions.

Hybrid approaches, which integrate physics-based models with machine learning techniques, have emerged as a promising solution by balancing accuracy, efficiency, and adaptability. However, these methods introduce additional complexity in model integration and require high-quality data for calibration and validation. Furthermore, reduced-order models and visual twins provide practical solutions for real-time deployment, but often involve trade-offs in model fidelity.

These observations highlight a key challenge in pipeline digital twin development: achieving an optimal balance between computational efficiency, predictive accuracy, and physical interpretability. Addressing this challenge is essential for advancing digital twin systems from research prototypes to scalable industrial solutions.

### 7.5. Uncertainty Quantification and Reliability in Pipeline Digital Twins

Despite the rapid advancement of digital twin (DT) technologies in pipeline engineering, uncertainty quantification (UQ) remains an underdeveloped yet essential component for reliable and scalable deployment. Most existing implementations focus on deterministic predictions, while real-world pipeline systems are inherently uncertain due to measurement limitations, modeling approximations, and dynamic operating environments. Explicitly accounting for uncertainty is therefore essential for improving robustness, model credibility, and risk-informed decision-making of pipeline digital twins.

Uncertainty in pipeline DT systems can be broadly categorized into four main sources: measurement uncertainty, model-form uncertainty, parameter uncertainty, and environmental and operational uncertainty. Each of these sources influences the accuracy, reliability, and predictive confidence of digital twin outputs in different ways.

Measurement uncertainty arises from sensor noise, calibration errors, data loss, and the limited spatial resolution of sensing systems. For example, pressure sensors and distributed fiber optic

sensing (DFOS) systems may introduce noise or missing data, particularly in harsh subsea environments. These uncertainties directly affect the quality of input data used for model updating and anomaly detection.

Model-form uncertainty originates from simplifications and assumptions in the underlying physical or data-driven models. High-fidelity models such as computational fluid dynamics (CFD) and finite element methods (FEM) often rely on assumptions regarding flow regimes, boundary conditions, or material behavior, which may not fully capture complex real-world phenomena such as multiphase flow behavior, nonlinear soil–structure interaction, and corrosion kinetics. Similarly, data-driven models may suffer from limited generalization when exposed to conditions outside their training domain.

Parameter uncertainty is associated with imperfect knowledge of model parameters, including friction factors, material properties, corrosion rates, defect growth parameters, and soil stiffness. These parameters are typically estimated from historical data, laboratory experiments, or inverse calibration procedures and may vary over time due to aging, environmental effects, and operational changes. Inaccurate parameter estimation can lead to significant deviations in predicted system behavior.

Environmental and operational uncertainty reflects variability in external conditions and system inputs, such as fluctuating demand, temperature changes, pressure transients, multiphase flow conditions, and seabed dynamics in offshore pipelines. These uncertainties introduce variability into both system states and boundary conditions, thereby increasing prediction uncertainty, making accurate prediction more challenging.

To address these uncertainties, digital twin systems must incorporate probabilistic and data assimilation techniques within the synchronization framework described in Section 7.3, where system states and parameters are continuously updated using incoming measurements. Methods such as Bayesian inference and Ensemble Kalman filtering enable the estimation of system states and parameters while explicitly accounting for uncertainty. These approaches update probability distributions rather than single deterministic estimates, thereby providing more robust and informative predictions.

In addition, stochastic simulation techniques such as Monte Carlo analysis can be used to propagate uncertainties through the model, allowing the quantification of confidence intervals and uncertainty bounds for predicted quantities such as pressure, stress, or remaining useful life (RUL). This is particularly important in safety-critical applications such as leak detection and structural integrity assessment, where decision-making must consider the probability of failure rather than relying solely on point estimates.

From an operational perspective, incorporating uncertainty into digital twin outputs enables risk-informed and reliability-based decision-making. For example, maintenance scheduling can be based on probabilistic thresholds of failure risk, rather than fixed deterministic limits. Similarly, leak detection systems can incorporate confidence levels to reduce false alarms and improve detection reliability under noisy and uncertain conditions.

The integration of uncertainty quantification into pipeline digital twins is therefore essential for transitioning from predictive models to reliable decision-support systems. Future research should focus on developing scalable probabilistic digital twin frameworks, improving data assimilation techniques under sparse and noisy data conditions, and establishing standardized approaches for uncertainty representation, validation, and regulatory acceptance in industrial applications. Building on the architectural, mathematical, and comparative analyses presented in the preceding subsections, the following section examines the systemic limitations that currently constrain the development of pipeline digital twins.

#### *7.6. Systemic Research Gaps Across Pipeline Digital Twins*

Despite rapid advances in digital twin (DT) applications in pipeline engineering, the current body of literature continues to exhibit several systemic gaps that limit scalability, reliability, and

industrial adoption. First, the lack of standardized digital twin architectures remains a critical barrier. Existing studies employ heterogeneous frameworks with inconsistent definitions of data integration, model synchronization, and feedback mechanisms, making cross-study comparison and industrial replication difficult. The absence of unified standards for model interoperability, data exchange, and validation protocols significantly hinders the development of modular and scalable digital twin ecosystems. Second, the integration of uncertainty quantification into system-level digital twin architectures remains limited, as discussed in Section 7.5, restricting the ability to support robust and risk-informed decision-making. Most existing models provide deterministic outputs without accounting for uncertainties arising from sensor noise, model assumptions, or environmental variability. This limitation is particularly critical in safety-sensitive applications such as leak detection and structural integrity assessment, where decision-making requires probabilistic confidence bounds.

Third, limited integration across the asset lifecycle (often referred to as the digital thread) persists. While many studies focus on operational-stage twins, few integrate design, construction, operation, and decommissioning into a unified framework. This fragmentation prevents the realization of full lifecycle optimization, data continuity, and knowledge transfer across different project stages. Fourth, data limitations and sensor sparsity continue to constrain model fidelity and system observability. Many legacy pipeline systems lack adequate instrumentation, resulting in incomplete observability and forcing reliance on surrogate models or inferred data, which may degrade prediction accuracy and increase model uncertainty. Beyond these structural gaps, several persistent technical and implementation challenges continue to limit the practical deployment of digital twin systems, as discussed in the following subsection.

### 7.7. Persistent Challenges

Despite substantial progress, the literature consistently identifies a set of persistent technical and implementation challenges that constrain the maturity and generalizability of pipeline digital twin deployments. From an implementation perspective, several engineering challenges emerge when translating digital twin concepts into operational systems. Data governance, ownership, and cybersecurity introduce significant implementation barriers, particularly in multi-operator and distributed infrastructure environments. Model validation and verification (V&V), which aim to demonstrate that a twin accurately represents physical reality across a wide range of operational conditions, including rare high-consequence events, is both scientifically demanding and practically essential for regulatory acceptance. Computational cost, although partially mitigated by reduced-order modelling techniques, remains a limiting factor for full-field, high-resolution simulation in real-time applications.

Several technical challenges must be addressed to advance the state of the art:

- Real-time computational constraints: High-fidelity multiphysics simulations remain computationally expensive, limiting real-time deployment. Although reduced-order models (ROMs) alleviate this issue, maintaining accuracy under dynamic operating conditions remains challenging.
- Hybrid model integration: While hybrid physics–data approaches dominate the literature, robust coupling strategies between first-principles models and machine learning components remain challenging, particularly in ensuring stability, interpretability, and generalization.
- Model validation and verification (V&V): There is no universally accepted methodology for validating digital twins in pipeline systems. Validation is often case-specific and lacks standardized performance metrics.
- Cybersecurity risks: Integration of digital twins with Industrial IoT (IIoT) and cloud platforms introduces vulnerabilities to cyberattacks, data manipulation, and system disruption, yet this aspect remains insufficiently addressed in the literature.
- Human–machine interaction: Many DT systems lack effective visualization and decision-support interfaces, limiting their usability by operators and engineers.

Addressing these technical challenges requires targeted research directions that advance both the theoretical foundations and practical deployment of pipeline digital twin systems, as discussed in the following subsection.

### 7.8. Future Research Directions

Several research directions emerge as high priority based on the gaps identified in the reviewed literature. Multi-physics coupled digital twins, integrating structural mechanics, fluid dynamics, electrochemistry, and thermal modeling within a unified framework, are needed to represent the full degradation complexity of pipeline systems, particularly in subsea and high-temperature high-pressure environments. Explainable AI (XAI) methods are needed to make the outputs of complex machine learning components within digital twins interpretable and auditable for safety-critical decisions. As discussed in Section 7.5, uncertainty quantification frameworks that propagate measurement, parameter, and model uncertainties through the digital twin are essential for producing calibrated prediction intervals and enabling risk-informed decision-making. Digital twin interoperability standards, defining common data schemas, application programming interfaces (APIs), and model exchange formats, are a prerequisite for the composition of component twins into system-level enterprise twins. Finally, lifecycle cost-benefit methodologies that quantify the return on investment of digital twin deployment across diverse pipeline operating contexts are needed to inform asset owner decision-making.

To address the identified gaps, several research directions are proposed:

- **Multi-Physics and Multi-Scale Digital Twins:** Future digital twins should integrate fluid dynamics, structural mechanics, corrosion chemistry, and thermal effects within unified frameworks. Multi-scale modeling approaches that bridge component-level and network-level behavior are particularly important.
- **Probabilistic and Bayesian Digital Twins:** The incorporation of probabilistic methods, including Bayesian inference and stochastic modeling, is essential for uncertainty-aware predictions. This will enable risk-informed decision-making and enhance regulatory acceptance.
- **Standardization and Interoperability:** The development of standardized data models, communication protocols (e.g., OPC UA), and digital twin ontologies is critical for enabling interoperability across platforms and stakeholders.
- **Edge-Cloud Hybrid Architectures:** Future systems should adopt hybrid computing architectures that leverage edge computing for real-time processing and cloud computing for large-scale simulation and analytics.
- **Explainable and Trustworthy AI:** Explainable AI (XAI) techniques should be integrated to ensure transparency and interpretability of machine learning components, particularly in safety-critical applications.
- **Autonomous and Prescriptive Digital Twins:** The evolution from predictive to prescriptive and ultimately autonomous digital twins represents a key frontier, where systems can not only predict failures but also recommend or execute optimal interventions.
- **Economic and Lifecycle Assessment:** Comprehensive cost-benefit analyses and lifecycle performance evaluations are needed to quantify the economic value of digital twin deployment and support investment decisions.
- **Regulatory frameworks governing pipeline safety,** including API RP 1160, ASME B31.8S, PHMSA regulations, and CSA Z662, require operators to implement integrity management programs (IMPs) based on risk assessment, inspection, and continuous monitoring. Emerging standards such as ASME V&V 40, NIST VVUQ frameworks, and ISO 23247 provide foundational guidance for validation, verification, and uncertainty quantification of digital twins. However, no unified regulatory standard currently defines requirements for digital twin fidelity, real-time synchronization, or acceptable uncertainty thresholds in safety-critical decision-making. Consequently, the integration of digital twins into regulatory-compliant IMPs remains an evolving area of industry and academic collaboration.

In summary, the literature reviewed in this study demonstrates that digital twin technology is rapidly evolving from a conceptual framework into a practical and high-impact solution for pipeline engineering applications. Across oil and gas, subsea, and water distribution systems, digital twins have shown strong potential in improving visibility, enabling real-time condition monitoring, and supporting data-driven decision-making. The integration of advanced sensing technologies, physics-based simulations, and machine learning models has significantly enhanced the ability to detect anomalies, predict failures, and optimize operational performance. At the same time, the review highlights important gaps related to data interoperability, model validation, scalability, and system integration, which must be addressed to fully realize the benefits of this technology. By consolidating current developments and identifying key challenges and opportunities, this paper provides a structured foundation for future research and practical implementation, bridging the gap between emerging digital twin capabilities and their effective deployment in complex pipeline systems. Collectively, these research directions highlight the transition of pipeline digital twins from isolated analytical tools toward integrated, intelligent, and autonomous cyber-physical systems.

## 8. Conclusions

This review has comprehensively examined the current state of digital twin technology in pipeline engineering, emphasizing its growing importance in enhancing system reliability, operational efficiency, and asset integrity management. The analysis demonstrates that digital twins are increasingly being adopted as a core component of modern pipeline systems, particularly in applications such as leak detection, corrosion monitoring, predictive maintenance, and real-time system optimization. By integrating physical models with real-time data streams and advanced analytics, digital twins enable a more proactive and data-driven approach to pipeline management.

One of the key findings of this review is the effectiveness of combining physics-based models, such as computational fluid dynamics, with data-driven approaches, including machine learning and artificial intelligence. This hybrid modeling strategy significantly improves the prediction accuracy under complex and dynamic operating conditions. Additionally, the use of IoT-enabled sensing technologies has enhanced the capability of digital twins to provide continuous, high-resolution monitoring of pipeline systems across various domains, including oil and gas, subsea infrastructure, and water distribution networks.

In addition to synthesizing existing literature, this study contributes a unified reference architecture for pipeline digital twins, supported by a mathematical formulation of the synchronization process and a structured comparative analysis of existing approaches. This integrated perspective provides a foundation for understanding how physical systems, data streams, and computational models interact within a closed-loop digital twin framework.

Future research should focus on addressing these limitations by developing scalable and standardized digital twin architectures, improving data fusion techniques, advancing uncertainty-aware modeling frameworks, and leveraging emerging technologies such as edge computing and autonomous AI systems. Furthermore, greater emphasis should be placed on lifecycle integration and cross-domain applications to maximize the value of digital twins.

Ultimately, digital twin technology represents a transformative paradigm in pipeline engineering, enabling the transition from reactive operation to predictive, adaptive, and intelligent infrastructure management. Its continued advancement will be critical for achieving safer, more efficient, and resilient pipeline systems in increasingly complex operating environments.

**Author Contributions:** Conceptualization, H.A., H.S., and R.S.; methodology, H.A., R.S.; formal analysis, H.A., R.S.; investigation, H.A., R.S.; resources, H.S.; writing—original draft preparation, H.A.; writing—review and editing, R.S.; visualization, R.S.; supervision, H.S.; project administration, H.S.; funding acquisition, H.S.

**Data Availability Statement:** No data was created in this study.

**Acknowledgments:** The authors gratefully acknowledge the support provided by Mitacs and Memorial University of Newfoundland. Their support has been instrumental in facilitating this research and advancing the objectives of the project.

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

## References

1. Grieves, M.; Vickers, J. Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. In *Transdisciplinary Perspectives on Complex Systems: New Findings and Approaches*; Springer International Publishing: Cham, Switzerland, 2016; pp. 85–113.
2. Kritzinger, W.; Karner, M.; Traar, G.; Henjes, J.; Sihn, W. Digital twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnLine* 2018, 51, 1016–1022.
3. Tao, F.; Zhang, M.; Nee, A.Y.C. Five-dimension digital twin modeling and its key technologies. In *Digital Twin Driven Smart Manufacturing*; Springer: Singapore, Singapore, 2019; pp. 63–81.
4. Fuller, A.; Fan, Z.; Day, C.; Barlow, C. Digital twin: Enabling technologies, challenges and open research. *IEEE Access* 2020, 8, 108952–108971.
5. Conejos Fuertes, P.; Martínez Alzamora, F.; Hervás Carot, M.; Alonso Campos, J.C. Building and exploiting a digital twin for the management of drinking water distribution networks. *Urban Water J.* 2020, 17, 704–713.
6. Wang, T.; Feng, K.; Ling, J.; Liao, M.; Yang, C.; Neubeck, R.; Liu, Z. Pipeline condition monitoring towards digital twin system: A case study. *J. Manuf. Syst.* 2024, 73, 256–274.
7. Hamidishad, N.; Barbosa, R.S.; Allahyarzadeh-Bidgoli, A.; Yanagihara, J.I. Digital twin frameworks for oil and gas processing plants: A comprehensive literature review. *Processes* 2025, 13, 3488.
8. Li, S.; Cheng, W.; Liu, Z.; Zhang, B.; Zheng, M.; Zheng, J.; Li, Y. Key technologies and applications of digital twin hydraulic engineering. *Digital Twin* 2025, 2, 2486862.
9. Paternina-Verona, D.A.; Coronado-Hernández, O.E.; Fuertes-Miquel, V.S.; Saba, M.; Ramos, H.M. Digital twin based on CFD modelling for analysis of two-phase flows during pipeline filling–emptying procedures. *Appl. Sci.* 2025, 15, 2643.
10. Pandey, P.; Mücke, N.T.; Jain, S.; Ramachandran, P.; Bohté, S.M.; Oosterlee, C.W. Machine learning-based digital twin for water distribution network anomaly detection and localization. *Eng. Proc.* 2024, 69, 201.
11. Syed, T.A.; Muhammad, M.A.; AlShahrani, A.A.; Hammad, M.; Naqash, M.T. Smart water management with digital twins and multimodal transformers: A predictive approach to usage and leakage detection. *Water* 2024, 16, 3410.
12. Homaei, M.; Di Bartolo, A.J.; Ávila, M.; Mogollón-Gutiérrez, O.; Caro, A. Digital transformation in the water distribution system based on the digital twins concept. *arXiv* 2024, arXiv:2412.06694.
13. Ahanger, T.A.; Abdibayev, Z.; Sagnayeva, S.; Zhumadillayeva, A.; Dyussekeyev, K. Smart wastewater management in hydro-technical systems using digital twin technology. *Sci. Rep.* 2026, in press.
14. Al-Ammari, W.A.; Sleiti, A.K.; Hamilton, M.; Ferroudji, H.; Gomari, S.R.; Hassan, I.; Rahman, M.A. Development of digital/visual twin for real-time leak detection in gas pipelines under multiphase flow conditions. *Greenh. Gases Sci. Technol.* 2025, 15, 513–530.
15. Al-Ammari, W.A.; Sleiti, A.K.; Rahman, M.A.; Rezaei-Gomari, S.; Hassan, I.; Hassan, R. Digital twin for leak detection and fault diagnostics in gas pipelines: A systematic review, model development, and case study. *Alex. Eng. J.* 2025, 123, 91–111.
16. Cai, X.; Wang, J. A digital twin-driven natural gas pipeline leakage detection method. *Int. J. Press. Vessel. Pip.* 2026, article 105793.
17. Liang, J.; Ma, L.; Liang, S.; Zhang, H.; Zuo, Z.; Dai, J. Data-driven digital twin method for leak detection in natural gas pipelines. *Comput. Electr. Eng.* 2023, 110, 108833.
18. Wang, D.; Shi, S.; Lu, J.; Hu, Z.; Chen, J. Research on gas pipeline leakage model identification driven by digital twin. *Syst. Sci. Control Eng.* 2023, 11, 2180687.

19. Hamilton, M.; Ferroudji, H.; Colbourne, M.; Sheppard, S.; Barooah, A.; Sleiti, A.K.; Rahman, M.A. Digital twin for pipeline leak monitoring. In Proceedings of the Abu Dhabi International Petroleum Exhibition and Conference (ADIPEC), Abu Dhabi, United Arab Emirates, November 2025; Paper D031S102R004.
20. Chen, B.Q.; Videiro, P.M.; Guedes Soares, C. Opportunities and challenges to develop digital twins for subsea pipelines. *J. Mar. Sci. Eng.* 2022, 10, 739.
21. Wegner, D.C.; Moses, I.; Ezenwa, I.O. Digital twin framework for subsea pipeline monitoring and integrity management. *Int. J. Comput. Appl.* 2025, 187, 52.
22. Syuryana, E.P.; Kariem, M.A.; Puja, I.W.; Maulana, M.I.; Uriyana; Agus, R. Digital twin for subsea pipeline corrosion monitoring location based on ILI-MFL pipeline operator forum and potentiodynamic test. In AIP Conference Proceedings; AIP Publishing: Melville, NY, USA, 2024; Vol. 3076, p. 150002.
23. Kaarlela, T.; Padrao, P.; Boswell, K.M.; Xanthidis, M.; Duran, A.; Bobadilla, L. Underwater digital twin applications: A systematic literature review. *Digital Twin* 2025, article 2605418.
24. Ismail, L.; Abdelmoti, A.; Basu, A.; Berini, A.D.E.; Naouss, M. A systematic review of digital twin-driven predictive maintenance in industrial engineering: Taxonomy, architectural elements, and future research directions. *arXiv* 2025, arXiv:2509.24443.
25. Bhowmik, S. Digital twin for offshore pipeline corrosion monitoring: A deep learning approach. In Proceedings of the Offshore Technology Conference (OTC), Houston, TX, USA, August 2021; Paper D041S049R003.
26. Duan, J.; Wang, X.; Yan, H.; Wang, S.; Huang, Y.; Tao, C. A digital twin framework for damage detection and localization of self-sensed cured-in-place underground pipelines. *Eng. Struct.* 2026, 354, 122401.
27. Ramos, H.M.; Kuriqi, A.; Besharat, M.; Creaco, E.; Tasca, E.; Coronado-Hernández, O.E.; Iglesias-Rey, P. Smart water grids and digital twin for the management of system efficiency in water distribution networks. *Water* 2023, 15, 1129.

**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.