Metaheuristic Optimization Review

♦ **ISSN:** 3066-280X **Vol:** 07, No. 02, 2026 ♦ PP. 73–91
♦ **Article DOI:** https://doi.org/10.54216/MOR.070205

REVIEW ARTICLE

AI-Enabled Digital Twins for Water and Power Critical Infrastructure: A Cross-Sector Review of Deployment Maturity, Cascading Risk, and Trustworthy AI

Esraa Walid[1,*] **Lozan Shady**[1] **Asmaa Elsayed**[1] **Manar Mostafa**[1]

[1] Department of Communications and Electronics, Delta Higher Institute of Engineering and Technology, Mansoura 35111, Egypt

Emails: CH2200058@dhiet.edu.eg; CH2200094@dhiet.edu.eg; CH2200232@dhiet.edu.eg; ch2300285@dhiet.edu.eg

Received: June 05, 2026 Revised: August 12, 2026 Accepted: September 09, 2026 ⋆ Corresponding author

ABSTRACT

AI-enabled digital twins are moving from simulation-oriented representations toward operational decision systems for critical infrastructure, but deployment evidence remains uneven and highly sector dependent. This review synthesizes AI-enabled digital twins in water and power infrastructure using a structured narrative evidence base of 36 primary application studies, balanced between 18 water and 18 power studies, together with recent reviews, interdependency research, and governance standards. The review compares application domains, data environments, AI roles, synchronization demands, validation settings, and decision authority across both sectors. Three analytical contributions extend beyond application cataloguing: a six-level deployment-maturity model with observable criteria; a coupled water–power framework for physical, cyber/data, spatial, and organizational cascading risk; and a trustworthy-AI pathway linking explainability, uncertainty, human oversight, fallback, traceability, interoperability, and re-validation. The evidence shows that water twins are strongly shaped by heterogeneous sensing, hydraulic/process uncertainty, and site-specific transfer, whereas power twins more often face tight latency, topology change, stability margins, and protection-adjacent decision requirements. Across both sectors, the dominant unresolved gap is not isolated AI accuracy but sustained field evidence demonstrating synchronized physical linkage, uncertainty-aware decisions, governed authority, cyber resilience, and lifecycle re-validation. The review therefore provides a cross-sector basis for distinguishing promising AI models from deployment-ready digital twins and for prioritizing future infrastructure-grade validation.

Keywords: Artificial intelligence ▪ Digital twins ▪ Water infrastructure ▪ Power systems ▪ Deployment maturity ▪ Cascading risk ▪ Trustworthy AI

1. INTRODUCTION

Artificial intelligence-enabled digital twins are increasingly being investigated as dynamic computational counterparts of physical systems in which sensing, communication, virtual representation, intelligent analytics, and operational feedback are integrated within a continuously evolving physical–digital relationship. Their significance extends beyond conventional simulation because the digital representation can be updated from observations of the physical system and used to interpret current conditions, examine plausible future states, and support decisions. This transition from static modeling toward context-aware physical–virtual systems is being driven by the convergence of intelligent sensing, connected infrastructure, and data-intensive analytics [1].

The defining value of a digital twin is therefore not the existence of a detailed virtual model alone, but the preservation of a meaningful relationship between that model and a specific physical asset, process, or network. Artificial intelligence strengthens this relationship by enabling the virtual representation to infer difficult-to-measure states, identify abnormal behavior, approximate expensive simulations, and estimate future system conditions. In this setting, intelligent analysis becomes physically contextualized rather than detached from the infrastructure that generates the data [2].

The scientific role of artificial intelligence within digital twins has expanded rapidly as machine learning and data-processing methods have become more capable of extracting information from heterogeneous operational data. Such methods can support state estimation, forecasting, anomaly detection, optimization, and adaptive decision support, creating a pathway from descriptive digital representations toward intelligent cyber–physical systems. This convergence is now an important component of the broader digital-twin research landscape [3].

Critical infrastructure provides a particularly demanding environment for these technologies because analytical errors can influence systems that support public safety, economic continuity, environmental management, and essential services. Water and power infrastructures are both increasingly dependent on networked sensors, operational data, digital communication, and automated control. Their digitalization creates substantial opportunities for improved efficiency and resilience, but it simultaneously raises stronger requirements for availability, data integrity, secure interaction, and trustworthy physical–virtual synchronization [4].

The water and power sectors are especially suitable for comparative analysis because they share the need for continuous monitoring and reliable operation while exhibiting very different process dynamics. Water infrastructure combines hydraulic transport, treatment processes, water quality, distributed assets, storage, pumping, and demand uncertainty. Power infrastructure combines electrical network states, equipment behavior, topology, generation, load, protection, stability, and rapidly evolving disturbances. Consequently, a digital-twin architecture that is adequate for one sector cannot automatically be assumed to satisfy the real-time, observability, uncertainty, or safety requirements of the other [5].

Existing reviews provide important sector-specific and technology-oriented syntheses, but the literature remains fragmented between broad digital-twin surveys, water-focused reviews, power-focused reviews, and cybersecurity-oriented critical-infrastructure analyses. The present review addresses this fragmentation by concentrating specifically on artificial intelligence-enabled digital twins in water and power infrastructure and by using a common analytical framework to compare applications, data environments, intelligent methods, digital-twin functions, validation maturity, challenges, and future research priorities.

The choice to study water and power together also reflects their operational interdependence. Water treatment and distribution depend on electrical energy for pumping, aeration, desalination, monitoring, communication, and control, while power systems depend on water for several generation and cooling processes and can be affected by water-related hazards. A disruption in one sector can therefore alter the operating conditions of the other even when their digital twins are developed independently. Although the present review does not attempt to construct a joint multi-infrastructure twin, this interdependence strengthens the case for comparing design principles, validation requirements, and resilience mechanisms across the two domains.

Another important distinction concerns the difference between a digital model, a digital shadow, and an operational digital twin. A digital model can reproduce selected physical behavior without receiving continuous information from the represented asset. A digital shadow can be updated automatically from the physical system but may still provide only one-way observation. A stronger digital-twin interpretation requires a persistent physical–virtual relationship in which the virtual state supports decisions that can influence the physical system. The literature does not always use these terms consistently, so this review evaluates studies according to the functions they implement rather than relying exclusively on the label adopted by the authors.

The focus on AI-enabled twins also requires care because artificial intelligence can occupy very different roles. In some studies, AI provides only a surrogate for a computationally expensive physical model. In others, it predicts future demand, estimates hidden states, detects anomalies, selects operating actions, or adapts the digital representation as new observations arrive. These roles imply different validation burdens. A surrogate used for offline scenario analysis can tolerate different error characteristics from an AI controller that directly influences infrastructure operation. The review therefore treats AI capability and operational authority as related but distinct dimensions.

The comparative structure adopted here is intended to reveal both transferable principles and non-transferable assumptions. Water and power researchers can learn from one another about graph representations, probabilistic inference, online updating, edge deployment, lifecycle governance, and uncertainty-aware decision support. However, an algorithm should not be transferred across sectors merely because it performs well in one domain. The physical constraints, sensing architecture, failure consequences, and acceptable decision latency must remain central to the design of the twin.

Table 1 positions the present review against representative recent reviews. The comparison is intentionally conservative: rather than claiming novelty merely from the number of publications considered, it identifies the additional value created by a direct water–power comparison and a gap-driven research roadmap.

The objectives of this review are therefore to establish the technical foundations of AI-enabled digital twins, synthesize representative water-infrastructure applications, synthesize representative power-infrastructure applications, compare the two sectors using a common analytical structure, identify the principal barriers that limit infrastructure-grade deployment, and derive future research directions directly from those barriers.

Table 1. Positioning of the present review against representative recent reviews.

Review	Scope	Direct W–P comparison	Main contribution and remaining gap
[6]	Water-sector digital twins	No	Strong water synthesis; power comparison is outside scope.
[7]	Wastewater twins	No	Deep wastewater focus; no power comparison.
[8]	Water supply, ML, and control	No	Integrates water-network DT, learning, and control reviews; power comparison is outside scope.
[9]	AI and DTs for power grids	No	Detailed grid synthesis; no water transfer analysis.
[10]	Power generation/distribution	No	Power architecture and security; no water comparison.
[11]	Critical-infrastructure cybersecurity	No	Cross-sector security focus rather than application/maturity comparison.
[12]	Renewables, smart grids, storage, and V2G	No	Reviews power-grid DT applications and adoption challenges; no direct water–power maturity comparison.
Present review	Water and power critical infrastructure	Yes	Direct comparison plus maturity/readiness analysis, coupled water–power cascading-risk intelligence, trustworthy-AI/interoperability pathways, gaps, and roadmap.

2. REVIEW APPROACH

This study is a structured narrative review rather than a formal systematic review. The method is designed to make the evidence trail auditable while preserving an interpretive engineering synthesis across heterogeneous water and power applications. Accordingly, no PRISMA flow diagram, pooled effect estimate, or claim of exhaustive database capture is made. Instead, the review documents the search scope, query families, eligibility rules, coding dimensions, final evidence composition, and the distinction between primary application evidence and contextual sources.

The literature search was updated through August 2026 using ScienceDirect, SpringerLink, Wiley Online Library, MDPI publisher records, IEEE Xplore records, and cross-publisher scholarly indexing pages. The retained core application evidence primarily spans 2020–2025, while a small number of foundational sources were retained where they are necessary to define digital-twin concepts, infrastructure interdependency, or governance principles. Search terms were organized into three families: general enabling concepts (“digital twin”, “artificial intelligence”, “machine learning”, “Internet of Things”, “cyber–physical system”, “real-time monitoring”, “prediction”, “optimization”, and “control”); water terms (“water treatment”, “wastewater”, “desalination”, “water distribution”, “leak detection”, “water quality”, “predictive maintenance”, and “water cybersecurity”); and power terms (“power system”, “smart grid”, “transmission”, “distribution network”, “substation”, “transformer”, “renewable energy”, “load forecasting”, “fault detection”, “voltage stability”, and “grid cybersecurity”).

Studies were retained for the core comparative evidence when they met three conditions: they addressed a water or power infrastructure problem; the digital twin had a substantive physical–virtual, synchronization, simulation, surrogate, diagnostic, predictive, optimization, or control role rather than appearing only as terminology; and the paper reported enough methodological or validation detail to support comparative coding. Generic digitalization papers, static models with no meaningful twin relationship, papers outside the water/power scope, purely conceptual commentary without application evidence, and duplicate or superseded records were excluded from the core primary-study tables. Review papers, standards, and interdependency studies were retained separately when they supported positioning, architecture, governance, or cross-sector interpretation.

To avoid false precision, the manuscript does not retrospectively invent database-hit or deduplication counts that were not archived during the original narrative search. The auditable endpoint is instead the final source corpus and its role in the synthesis. The final revision contains 68 cited source records, including 36 primary application studies used as the evidence backbone (18 water and 18 power) and 32 contextual records comprising reviews, digital-twin background sources, water–power interdependency evidence, and standards/frameworks. Every retained primary study is listed in Supplementary Table S2, together with its validation setting, evidence indicators, decision authority, and conservative maturity assignment. Supplementary Table S1 provides reproducible Boolean search strings reconstructed from the documented search-term families to replicate the final August 2026 search scope. They are not presented as verbatim archived database logs; database-hit counts are intentionally left unreconstructed because the original narrative search did not archive a deduplicated PRISMA-style export.

Audit item	Recorded review information
Search/update period	Search updated in August 2026; retained core publications from 2020–2025, with limited foundational exceptions.
Sources searched	ScienceDirect, SpringerLink, Wiley Online Library, MDPI, IEEE Xplore, and cross-publisher scholarly indexing pages.
Initial hit count	Not retrospectively reconstructed because the original narrative search did not preserve a PRISMA-style deduplicated export; no artificial count is reported.
Eligibility assessment	Topic relevance, substantive digital-twin role, usable methodology/validation information, and water/power infrastructure scope.
Core primary evidence	36 studies: 18 water and 18 power.
Contextual evidence	32 records covering prior reviews, technical background, coupled-infrastructure evidence, standards, and forward-looking synthesis.
Major exclusions	Generic digitalization; static modeling without a meaningful twin relationship; applications outside water/power; duplicate/superseded records; insufficient evidence for comparative coding.

The retained primary studies were coded using a common evidence structure: application, AI method, digital-twin role, data and validation setting, reported result or contribution, limitation, real-time relevance, and degree of decision influence. This structure prevents a benchmark accuracy value from being treated as equivalent to deployment readiness. A simulation-only study can provide strong methodological evidence while still occupying a low deployment-maturity level, whereas a field demonstrator can contribute high operational evidence even when its predictive metrics are less spectacular.

Quality appraisal therefore focuses on four questions. First, is the physical–virtual relationship sufficiently clear for the claimed twin function? Second, is the validation setting commensurate with the operational claim? Third, are uncertainty, latency, model mismatch, or data limitations acknowledged where they materially affect decisions? Fourth, does the reported result demonstrate model performance only, or does it also show decision or infrastructure value? These questions are applied consistently to the water and power evidence tables and later operationalized in the six-level maturity model.

The review itself also has limitations. It is a structured narrative synthesis rather than a statistically exhaustive systematic review; the original deduplicated database-hit counts were not archived; heterogeneity in assets, datasets, objectives, and validation settings prevents meaningful quantitative meta-analysis; and publisher/indexing coverage can introduce retrieval bias. These constraints are made explicit so that the maturity and cross-sector conclusions are interpreted as evidence-structured analytical synthesis rather than as pooled statistical estimates.

Unless otherwise stated, all synthesis figures in this manuscript were developed by the authors from the coded evidence; no published figure is reproduced or adapted.

The evidence organization proceeds from source identification and eligibility assessment to common coding and cross-sector synthesis; the exact search strings and study-level audit are supplied in the Supplementary Material.

The resulting evidence base is deliberately representative rather than statistically exhaustive. Its strength lies in coverage across application families and validation settings, combined with explicit coding of what the digital twin actually does. This provides a defensible basis for comparing water and power without implying that heterogeneous assets, datasets, and performance metrics are directly commensurable.

3. AI-ENABLED DIGITAL TWIN BACKGROUND

An AI-enabled digital twin can be understood as a layered cyber–physical architecture rather than a single model. The physical infrastructure produces observations through sensors, meters, supervisory systems, laboratory measurements, equipment logs, and other acquisition mechanisms. These observations are transmitted to a digital environment in which their timing, physical meaning, location, and relationship to the represented infrastructure are preserved. The virtual representation organizes the relevant system state, while the AI layer transforms the synchronized information into predictions, diagnoses, uncertainty estimates, optimization results, or control-oriented recommendations [13].

Sensors and Internet of Things technologies define what the digital twin can observe. In water systems, direct measurements may include flow, pressure, tank level, conductivity, temperature, water-quality indicators, equipment condition, and treatment-process variables. In power systems, common observations include voltage, current, active and reactive power, frequency, topology, load, equipment temperature, switching events, and protection-related signals. The density, accuracy, and update rate of these measurements establish a practical boundary on the fidelity and usefulness of the virtual state [14].

Data synchronization is one of the defining properties of a digital twin. The digital representation must remain sufficiently aligned with the physical state for the intended task, but the required update frequency is application dependent. Maintenance planning may tolerate relatively slow updates, whereas fault identification, stability assessment, or active control may require a much tighter correspondence. Synchronization should therefore be evaluated relative to the physical timescale and consequence of the target decision rather than through a universal claim of "real-time" operation [15].

Artificial intelligence can contribute at several levels of the digital-twin architecture. Data-driven models can approximate computationally expensive physical simulations, estimate difficult-to-measure states, forecast future conditions, identify abnormal behavior, quantify uncertainty, optimize operating decisions, and learn control policies. The value of this intelligence depends on whether the output remains physically meaningful and whether the digital twin provides sufficient context to connect the output with the represented infrastructure [16].

The functional progression from monitoring to prediction, optimization, and control provides a useful framework for comparing digital-twin applications. Monitoring establishes awareness of the present state. Prediction estimates future conditions or risks. Optimization evaluates preferable actions under objectives and constraints. Control closes the loop by allowing virtual analysis to influence physical operation. Earlier digital-twin-driven AI frameworks demonstrated the importance of connecting model development with physically contextualized data rather than treating the learning system as an isolated analytical service [17].

Figure 1 summarizes the architecture adopted in this review. Water and power infrastructure enter a common chain of sensing, synchronization, virtual representation, AI analytics, and decision feedback, while retaining sector-specific data and operational constraints.

The application taxonomy in Fig. 2 provides a bridge between this technical background and the two sector-specific reviews that follow. It shows that similar high-level digital-twin functions can appear in different infrastructure contexts even when the underlying physical processes and decision horizons are substantially different.

The virtual-state layer can itself be implemented at different levels of physical fidelity. High-fidelity mechanistic models preserve detailed physical relationships and can support physically interpretable scenario analysis, but they may be computationally expensive and difficult to calibrate continuously. Reduced-order models retain selected dominant dynamics while lowering computational cost. Purely data-driven models can be fast and flexible but may extrapolate poorly outside the conditions represented during training. Hybrid twins combine these approaches so that physical structure constrains the learning problem while data-driven components capture residual behavior or accelerate repeated calculations.

The selection of model fidelity should therefore be decision driven. A twin used to evaluate long-term infrastructure planning can accept slower computation if the model captures important physical detail. A twin used for rapid anomaly response may require a reduced-order or learned surrogate

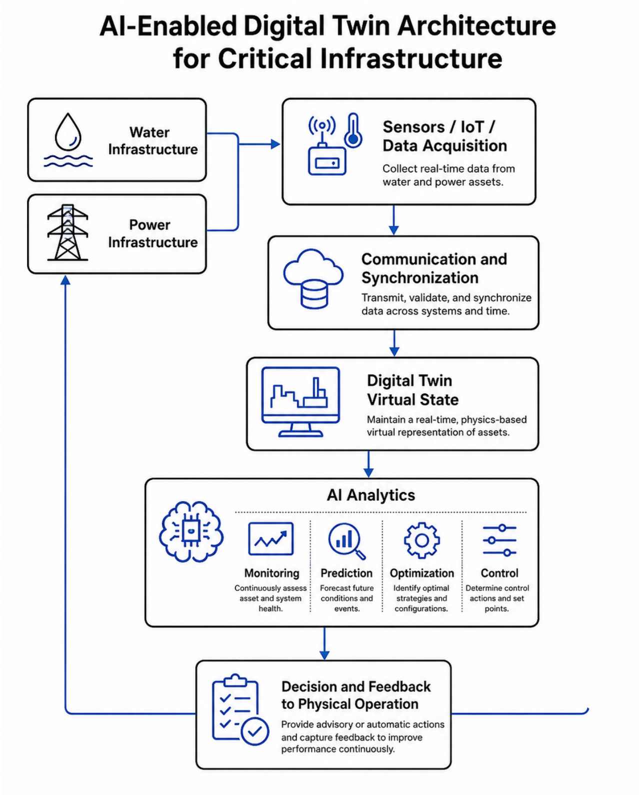

Figure 1. AI-enabled digital-twin architecture for water and power critical infrastructure.

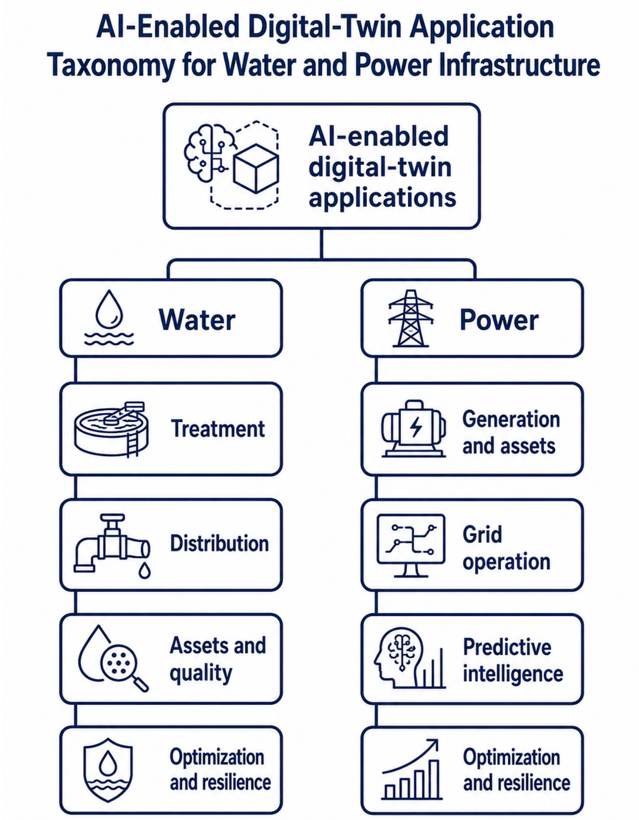

Figure 2. Taxonomy of AI-enabled digital-twin applications across water and power infrastructure.

that produces results within a strict latency budget. The relevant question is not which modeling approach is universally superior, but which representation preserves the physical relationships necessary for the target decision at an acceptable computational cost.

Data governance is another core architectural requirement. Infrastructure twins can combine operational measurements, maintenance records, simulated data, environmental information, laboratory results, configuration files, and operator annotations. These data sources differ in ownership, update rate, reliability, and sensitivity. A mature twin should retain provenance so that an important decision can be traced to the observations and model version that produced it. This is particularly important when AI models are updated over time or when different organizations contribute data to the same infrastructure environment.

Explainability should likewise be understood in relation to the decision process. An explanation is useful when it helps determine whether the model is responding to physically plausible evidence, whether an unusual operating condition is outside the training distribution, or why two apparently similar states lead to different recommendations. Generic feature-ranking graphics can support analysis, but infrastructure deployment often requires explanations tied to engineering variables, constraints, alarms, and operational procedures.

Synchronization should be interpreted across the complete sensing-to-decision chain. The physical event must first be observed, transmitted, time aligned, incorporated into the virtual state, analyzed, and converted into a decision. Any one of these stages can dominate the total response time. Consequently, reporting only the inference time of an AI model can give a misleading impression of digital-twin responsiveness when communication, state reconstruction, or optimization requires substantially longer.

This end-to-end view is particularly useful when comparing water and power infrastructure. Water-treatment and quality-management decisions frequently evolve over process timescales that permit slower synchronization, while certain distribution-control events can still require rapid updates. Power-system faults, stability events, and protection-adjacent decisions can operate on far shorter horizons. The relevant engineering question is therefore whether the complete digital-twin chain is sufficiently fast for the physical process being represented, not whether the system satisfies an arbitrary universal definition of real time.

The architecture also requires lifecycle management. Sensors can drift, process conditions can change, assets can be replaced, network topology can evolve, and AI models can lose validity as the operating distribution shifts. Infrastructure-grade twins therefore require versioning of models and data interfaces, traceability of important updates, and explicit criteria for recalibration or retraining. These lifecycle mechanisms are essential for maintaining physical–virtual correspondence over months or years rather than only during an experimental campaign.

AI methods should be interpreted according to the digital-twin function they support. Monitoring tasks emphasize state estimation and anomaly screening, predictive tasks emphasize forecasting and risk estimation, optimization evaluates alternatives under constraints, and control introduces the highest requirement for validated decision authority. This functional view prevents the review from treating all AI techniques as interchangeable simply because they appear inside a digital-twin architecture.

The distinction is important because the consequence of model error generally increases as the twin moves closer to physical action. A moderate forecast error may be acceptable when the output is used for planning, whereas the same uncertainty can be unacceptable when it directly determines an automated control action. The required level of explainability, uncertainty calibration, validation depth, and human oversight should therefore be proportional to the authority granted to the intelligent component.

Hybrid modeling offers a practical way to balance these requirements. Mechanistic models preserve conservation laws and engineering relationships, while data-driven models can capture residual behavior, accelerate expensive simulations, or adapt to observed conditions. The most useful combination is application dependent: the twin should retain enough physical structure to remain meaningful outside the training data while using AI where it creates measurable computational or predictive value.

4. DIGITAL TWINS IN WATER INFRASTRUCTURE

Water infrastructure presents a distinctive environment for digital-twin deployment because its operation combines distributed physical assets with hydraulic transport, treatment processes, water-quality constraints, environmental variability, and long asset lifetimes. The digital twin must often integrate observations generated at different spatial locations and sampling frequencies while preserving their relationship to the underlying process or network. AI becomes especially valuable when heterogeneous information must be converted into predictions, diagnostics, or operational recommendations.

Water-treatment applications use digital twins to link process models, equipment information, data-driven surrogates, and plant observations. The central value of the digital twin in this context is not simply predictive accuracy; it is the ability to associate the prediction with the current treatment configuration, recent operating history, and process constraints. This relationship can support proactive operation, maintenance planning, and scenario analysis while reducing dependence on repeated high-cost physical experimentation.

Wastewater treatment introduces additional biological and process variability. AI-enabled twins can support process-state monitoring, fault-aware representation, and faster evaluation of treatment behavior, but their reliability remains sensitive to the quality and representativeness of process data. Changes in influent characteristics, aeration behavior, microbial dynamics, and equipment performance can alter the relationship learned from historical observations, making model adaptation an important requirement.

Desalination represents another high-value application because reverse-osmosis systems combine nonlinear membrane behavior, changing feed-water conditions, energy-intensive operation, and equipment degradation. Data-driven surrogate twins can accelerate prediction of permeate flow, salinity, and specific energy consumption, but simulation-generated scenarios cannot substitute for long-duration plant evidence under changing membrane condition, feed-water chemistry, sensor quality, and maintenance interventions.

Water-distribution networks form one of the most active digital-twin application areas because network state is spatially distributed and only partially observed. Calibrated hydraulic models, graph-based representations, probabilistic inference, and optimization can support leak localization, water-quality regulation, pressure management, and what-if analysis. However, the reliability of inferred network states depends strongly on sensor placement, demand uncertainty, model calibration, and the frequency with which physical changes are reflected in the virtual model.

Leak and fault detection illustrates why uncertainty should be carried into the decision process. A deterministic alarm can identify a likely abnormal condition, but an operational digital twin should also indicate how strongly the available evidence supports a specific leak location or failure hypothesis. Probabilistic twin architectures can therefore provide more useful support by connecting localization with confidence and sensing conditions.

Predictive maintenance extends the water twin from process supervision toward asset-health intelligence. Pump stations, valves, treatment equipment, membranes, and network components can be represented using condition data and operational histories so that maintenance can be scheduled according to evolving risk rather than fixed intervals. The main difficulty is long-term transfer: an asset model that performs well under one operating regime may deteriorate as equipment ages, maintenance changes the system, or sensor behavior drifts.

Water-quality monitoring requires the twin to combine online sensing with variables that may only be available through laboratory measurements. This mixed data cadence complicates synchronization but also demonstrates why "real time" should be defined relative to the decision horizon. For some quality-management tasks, minute-level updates may be fully operationally useful even though they are much slower than the timescales required for fast electrical protection.

Optimization and cybersecurity complete the water-sector picture. Digital twins can support dosing, pumping, flow rerouting, and resilience analysis by evaluating operational decisions in the virtual environment before physical implementation. At the same time, increasing connectivity creates new attack surfaces, and malicious or corrupted observations can propagate through the twin into trusted decisions. Cyber-resilience must therefore be considered as part of the physical–virtual architecture rather than as an external security add-on.

Table 2 compares representative studies across these application families. The table separates reported contribution from validation limitation so that strong computational results are not automatically interpreted as utility-scale readiness.

The expanded 18-study water corpus spans controlled prototypes, synthetic networks, real distribution systems, treatment case studies, and a full-scale process-control demonstration. This broader validation range strengthens the distinction between algorithmic capability and deployment evidence.

Table 2. Representative primary AI-enabled digital-twin studies in water infrastructure.

Application / study	AI and twin role	Validation, contribution, and main limitation
Water treatment [18]	ML prediction for process and maintenance support	Full-scale membrane-bioreactor case integrating equipment, cloud, historical, and live data; transfer across plants remains insufficiently validated.
Drinking-water management [19]	Statistical anomaly detection and automated ML for asset/quality twins	Utility prototypes reported predictive maintenance and proactive quality support; independent multi-utility validation remains limited.
Wastewater monitoring [20]	Data-driven process monitoring and fault-aware representation	Benchmark/simulation-based wastewater-process monitoring demonstrates fault-aware twin analysis; operational field linkage is not established.
Desalination surrogate [21]	Comparative ML surrogates with explainability	18,816 generated operating scenarios for flow, salinity, and energy prediction; long-duration plant generalization is unproven.
Distribution decision support [22]	Calibration, neural/graph learning, and optimization	C-TOWN benchmark supports quality regulation and leak localization; benchmark conditions underrepresent utility uncertainty.
Probabilistic leak localization [23]	Generative deep learning with Bayesian inference	Three network test cases provide uncertainty-aware leak probabilities; utility-scale field validation is still needed.
Campus leakage twin [24]	Hydraulic calibration, optimization, and graph learning	Real campus network demonstrates end-to-end leakage analysis; campus scale remains more controlled than municipal systems.
Irrigation distribution management [25]	Deep reinforcement learning and graph-based irrigation modeling	A Moroccan district case optimizes sowing schedules under hydraulic constraints; the reported twin remains a digital model without bidirectional data integration.
Cyber-resilience [26]	Monte Carlo cyberattack simulation	Coupled treatment/distribution model evaluates attack consequence; adaptive threat and operational cyber evidence remain limited.
Drinking-water coagulation [27]	Similarity-based hybrid ML for dosing prediction	A hybrid predictor was integrated into a Texas water-purification plant twin; performance depends on data similarity, and sustained autonomous control is not established.
Treatment anomaly detection [28]	Combined anomaly-detection framework in a process-control twin	SWaT prototype provides controlled cyber/physical anomaly evidence; utility-scale coverage of attack and fault diversity is not established.
Real-time model updating [29]	Coverage-based ML update scheduling for synchronized twins	SWaT prototype reduced model-update error by up to 97%; evidence is from a controlled treatment facility.
Multimodal smart-water twin [30]	Informer forecasting and multimodal transformer anomaly detection	Sensor-synchronized experiments reported high forecasting and leakage-detection metrics; external utility replication remains absent.
Urban-river quality twin [31]	LSTM, Kriging, and 3D pollutant simulation	City-river platform integrates prediction and visualization; evidence is site-specific and decision impact is not independently validated.
Sensor placement for WDS twins [32]	Graph neural-network metamodel and optimization	Synthetic WDS proof-of-concept shows accurate pressure reconstruction; no live utility deployment is reported.
Large-scale operational WDS twin [33]	Hydraulic twin plus graph-convolution leak detection	Field data from a large campus network reported 90% leak-detection accuracy and scenario water savings; broader municipal transfer remains open.
Cross-network leak detectability [34]	CNN using AMI-pressure residual images against calibrated hydraulic twins	Evaluated across 11 smart WDSs and multiple leak sizes; methodology is broader than single-network studies but remains model/calibration dependent.
Full-scale PN/A aeration control [35]	Mechanistic digital twin with policy-iterative dynamic programming	Real-scale plant demonstration reported >95% twin accuracy and 16.7% electricity-efficiency improvement; evidence remains process/facility specific.

The evidence in Table 2 indicates that water-sector maturity is uneven across applications. Hydraulic-network studies have progressed toward uncertainty-aware diagnosis and operational decision support, while treatment studies increasingly combine physical context with data-driven surrogates. The dominant unresolved water-sector issue is longitudinal robustness under changing demand, water quality, sensing, equipment condition, network configuration, and operating policy.

The water evidence also shows why spatial context is fundamental. A pressure anomaly has different implications depending on its position relative to pumps, valves, storage tanks, elevation, and demand patterns. A water-quality observation must be interpreted together with transport time and network connectivity because the physical effect can propagate through the system. Graph and hydraulic representations are therefore not merely convenient data structures; they encode relationships that help preserve physical meaning when direct sensing is incomplete.

Treatment and desalination twins additionally face gradual performance degradation. Membrane fouling, changing feedwater composition, equipment wear, and seasonal operating changes can shift the relationship between measured variables and process outcomes. A twin that remains static after commissioning can therefore become progressively less representative even when its original validation was strong. Long-term operation requires mechanisms for detecting performance drift and distinguishing true process changes from sensor faults or data-quality problems.

The role of operators is also particularly important in water systems because maintenance and quality decisions often involve contextual information that is not fully represented in sensor streams. Inspection findings, recent repairs, weather events, customer complaints, and temporary operational constraints can materially change the interpretation of an anomaly. Human-in-the-loop twins should provide mechanisms for incorporating these observations and for preserving them as part of the historical context used in future analysis.

Across the water literature, the operational logic is consistent: heterogeneous observations update a hydraulic or process-oriented virtual state, AI functions infer difficult-to-measure behavior, and the resulting evidence supports decisions that affect physical operation. This sequence emphasizes that the intelligent component is embedded within a broader engineering workflow rather than functioning as an isolated prediction service.

A major water-sector challenge is the mismatch between spatial infrastructure extent and available sensing. Distribution networks may contain large numbers of pipes, valves, junctions, and customer demands, while only a limited subset of locations are instrumented. The twin must therefore combine measured information with physical relationships and inferred states. This makes uncertainty-aware hydraulic modeling particularly valuable because an apparently precise prediction can otherwise hide substantial ambiguity in unmeasured parts of the network.

Treatment systems create a different observability problem. Some variables are available continuously, while others are measured through laboratory procedures or indirectly inferred from process behavior. The digital twin must preserve these different data cadences without treating a slowly updated quality variable as though it were synchronized at the same frequency as pressure or flow. This difference affects both the design of the state representation and the interpretation of model freshness.

Operational usefulness also depends on how predictions are translated into action. A leakage probability, fouling forecast, or pump-health score creates value only when it informs inspection, control, maintenance, or resource allocation. Future water twins should therefore report the downstream decision enabled by the model, the time available for that decision, and the evidence showing that the resulting action improves infrastructure performance without introducing unacceptable risk.

5. DIGITAL TWINS IN POWER INFRASTRUCTURE

Power infrastructure creates a different operational environment because electrical states can change rapidly, network topology may change after faults or switching events, and system stability can be highly sensitive to disturbances. AI-enabled digital twins in this sector therefore frequently emphasize state estimation, equipment intelligence, fault diagnosis, load forecasting, voltage stability, renewable-energy integration, and controller development.

Power-generation applications use digital twins to connect equipment models with operating measurements so that thermal state, performance, degradation, and control-relevant behavior can be estimated more efficiently. Reduced-order modeling can lower the computational cost of high-fidelity representations, while data-driven updating can help the twin remain aligned with a changing physical asset.

Smart grids extend the digital-twin concept beyond individual equipment by linking network information, distributed sensing, demand, renewable generation, and intelligent analytics. The twin can function as an operational environment for state estimation, scenario testing, forecasting, and coordinated decision support. The key challenge is that increasing network complexity and decentralization make synchronization and scalability more demanding.

Transmission and distribution applications emphasize topology-aware network representation, state visibility, voltage management, and disturbance analysis. Real distribution-system demonstrators show that digital twins can support practical planning and operational studies, but they also expose the effect of measurement uncertainty and imperfect network models. This is important because field deployment reveals limitations that are often hidden in benchmark studies.

Substations represent a particularly demanding digital-twin application because diagnosis may require the fusion of waveform information, protection events, supervisory data, topology, and simulated fault scenarios. Multimodal AI methods can use the twin as a structured environment for scenario generation and fault interpretation, but the end-to-end response time depends on communication and state reconstruction as well as model inference.

Renewable-energy integration introduces variability and uncertainty that strengthen the need for forecasting, what-if analysis, and adaptive operation. Digital twins can provide a context in which weather-driven generation, network constraints, storage, and control decisions are evaluated together. However, the usefulness of the twin depends on whether forecasting uncertainty is incorporated explicitly into operational decisions. A recent renewable-energy perspective frames these roles as an integrated digital-twin opportunity rather than as proof of deployment maturity [36].

Load forecasting uses historical and contextual information to estimate future demand within the virtual environment. Although high-performing forecasting models can improve planning and grid management, model accuracy alone does not establish digital-twin maturity. The forecasting component must remain synchronized with the actual system and must adapt when user populations, consumption behavior, climate conditions, or network configurations change.

Fault detection and voltage-stability analysis illustrate the importance of latency and uncertainty. Fast intelligent surrogates can reduce computational burden and support network-scale evaluation, but a reliable digital twin must identify when the operating point lies outside the region represented during model development. Probabilistic outputs are therefore especially useful near stability boundaries or under uncertain topology and parameters.

Predictive maintenance applies similar principles at the equipment level. Transformers, generators, cables, and other assets can be monitored using temperature, loading, vibration, insulation, and condition indicators. Online learning can help maintain accuracy as the underlying equipment process evolves, yet lifecycle deployment requires version control, recalibration, and evidence that updated models remain safe.

Grid monitoring and cybersecurity complete the power-sector picture. The same connectivity that enables synchronized digital twins also creates pathways for data manipulation, communication disruption, and cyber-induced physical consequences. Digital twins can support resilience analysis and attack simulation, but they must themselves be protected as part of the critical cyber–physical architecture.

Table 3 compares representative power-sector evidence using the same structure as the water table, allowing the two sectors to be compared directly in the following section.

The expanded 18-study power corpus spans equipment experiments, cyberattack models, distribution-network cases, substation diagnostics, resilience scheduling, field measurements, and physical-testbed transfer. This broader coverage makes validation setting explicit rather than treating all AI performance results as equivalent evidence.

The evidence in Table 3 shows a stronger concentration on fast state analysis, fault diagnosis, equipment adaptation, and control-oriented use than is commonly observed in the water literature. Real demonstrators and physical testbeds further show that end-to-end latency, measurement quality, topology updates, model mismatch, and operator integration remain decisive barriers to operational use.

Power digital twins must also address the distinction between normal operating variability and rare high-consequence events. Load changes, renewable variability, and switching can alter network states continuously, while faults and protection events may occur infrequently but require rapid and reliable interpretation. Training data can therefore be highly imbalanced. Simulation-based scenario generation is valu-

Table 3. Representative primary AI-enabled digital-twin studies in power infrastructure.

Application / study	AI and twin role	Validation, contribution, and main limitation
Distribution operation [37]	State estimation and model/data integration	UK demonstrator exposes measurement/model uncertainty and operational value; results remain network- and data-dependent.
Transformer thermal management [38]	Online learning for a self-updating equipment twin	Dynamic equipment datasets show strong prediction and adaptation; fleet-wide longitudinal transfer is not established.
Plant equipment monitoring [39]	Hybrid reduced-order modeling	Field/numerical validation supports real-time generator/transformer monitoring; evidence concerns two devices in one plant context.
Substation fault diagnosis [40]	SVM anomaly detection and a learning-enabled maintenance framework	Simulation experiments use historical SCADA data from a 110-kV substation; diagnostic results support offline evaluation, not demonstrated live operational deployment.
Distribution faults [41]	Dilated convolutional multitask learning	Ultra-real-time fault identification in evaluated cases; field performance remains topology- and communication-dependent.
Voltage stability [42]	Gaussian-process modeling with Bayesian feature selection	Fast uncertainty-aware stability-margin prediction on large networks; critical-node bias remains a limitation.
Load forecasting [43]	Attention-enhanced recurrent learning	Five years of half-hourly data show lower error than compared neural baselines; transfer to changing populations requires validation.
Microgrid control [44]	Reinforcement learning with twin-to-testbed transfer	Controller transferred to a physical DC-microgrid testbed; utility deployment requires stronger safety assurance.
Smart-grid acquisition [45]	Convolutional learning, attention, and twin mapping	Comparative acquisition study reports energy-efficiency gains; independent operational replication remains needed.
Power-system stability [46]	Hybrid physics/data digital-twin simulation	Multiple stability and asset case studies demonstrate fast high-fidelity analysis; evidence is not a sustained utility field deployment.
Transformer health/life [47]	Multitask LSTM–GRU within a fog/digital-twin architecture	A 940-record transformer dataset produced $R^2 = 0.985$ in the proposed monitoring architecture; direct field deployment and longitudinal fleet validation are not demonstrated.
Demand-side smart-grid twin [48]	Hybrid digital twin with AI-enabled demand-side recommendation	Prototype/simulation study reported 36.8% reduction in net energy consumption after recommendation; sustained operational deployment is not demonstrated.
SCADA cyberattack analysis [49]	Directed-graph digital twin with label-encoded ML	Cyberattack-flow representation enables classification of SCADA attack patterns; physical grid consequence is not field validated.
Hydropower fault resilience [50]	Deep learning integrated with a hydropower digital twin	MATLAB-based evaluation reports improved fault-response and efficiency indicators; results remain simulation based.
Distribution-equipment updating [51]	IoT/optical sensing with recursive least-squares model updating	Test-case evaluation reported 0.014-s updates and 93.2% accuracy; longitudinal asset deployment remains unreported.
Transformer fault warning [52]	Sparrow-optimized XGBoost and sensor-linked digital twin	Oil/gas and temperature monitoring support fault identification and warning in an experimental transformer twin; long-duration, multi-site deployment remains unestablished.
Distribution-network operation [53]	Graph neural networks and dynamic digital twins	MATLAB/Simulink and PyTorch Geometric simulations evaluate voltage regulation and network efficiency; sustained physical-network deployment is not demonstrated.
Dynamic distribution topology [54]	CIM-based twin topology with spectral clustering and online linked updating	A demonstration-area case validates rapid topology updates from changing network states; operational decision authority and sustained pilot deployment are not demonstrated.

able for exposing AI models to rare events, but the simulated scenarios must remain physically credible and should be complemented by operational evidence whenever possible.

The increasing penetration of distributed energy resources creates an additional coordination problem. A digital twin may need to represent assets that are geographically distributed, controlled by different devices, and updated at different rates. Local decisions can interact with network-level constraints, so hierarchical or distributed twin architectures may be more appropriate than a single monolithic model. Edge intelligence can support local responsiveness, while higher-level coordination preserves system-wide context.

For stability and protection-adjacent applications, uncertainty should be coupled with operating margins. A probabilistic prediction is most useful when it helps determine how close the system is to a limit and how much confidence can be placed in that estimate. The twin should therefore support conservative action when uncertainty grows near critical boundaries rather than treating the mean prediction as sufficient evidence for automated control.

The power-sector workflow has the same physical–virtual foundation as the water workflow but places greater emphasis on dynamic network state, topology, fast diagnosis, and control-relevant outputs. The digital twin may need to integrate steady-state network information with transient measurements, asset condition, switching events, weather-dependent generation, and load variation. This multi-timescale environment makes temporal consistency as important as raw model accuracy.

For equipment-level twins, continuous adaptation is particularly important because thermal, electrical, and mechanical behavior changes with loading, ambient conditions, maintenance, and aging. Online learning can help preserve predictive performance, but unconstrained model adaptation also creates governance questions. A production twin should record model versions, validate updated behavior before granting it decision authority, and retain a recoverable baseline when new data are unreliable.

Network-level twins add topology as a dynamic source of context. Switching actions, outages, distributed generation, and restoration can change the relationships among measurements. AI models trained on one topology can therefore encounter operating states that are structurally different from those represented in training. Graph-aware or physics-guided representations are promising because they encode network relationships more explicitly, but their operational value still depends on the timeliness and accuracy of topology information.

The power literature also illustrates the difference between model latency and system latency. A fault classifier may execute in a fraction of a second, yet the useful response time also includes acquisition, communication, synchronization, state reconstruction, and interaction with protection or control systems. Field-oriented publications should therefore report end-to-end response and not only algorithmic inference speed.

6. COMPARATIVE ANALYSIS: WATER VS. POWER

The direct comparison between water and power infrastructure provides the central analytical contribution of this review. Both sectors use sensing, virtual representation, and AI-supported inference to build an operationally meaningful relationship between physical assets and computational models. However, the engineering conditions under which the twin must function differ substantially, particularly in data cadence, observability, physical timescale, and decision consequence.

Water systems typically combine hydraulic, chemical, biological, water-quality, and equipment-condition information. These variables are often asynchronous: pressure and flow may be measured continuously, laboratory measurements may be available daily, and asset-maintenance information may be event driven. Power systems typically combine electrical state measurements, equipment condition, topology, generation, load, protection events, and high-frequency disturbance information. The resulting challenge is not simply greater data volume, but the need to interpret multiple measurement rates within the relevant decision window.

The sectors also differ in the balance between physical modeling and data-driven learning. Water-distribution twins frequently retain hydraulic models because topology, mass conservation, and flow relationships provide essential structure under sparse sensing. Treatment twins similarly benefit from process knowledge because water chemistry and biological behavior can change the relationship between observations and outcomes. Power twins also depend strongly on physics, but intelligent surrogates are increasingly used to accelerate high-dimensional fault, thermal, stability, and control analysis.

Real-time requirements represent one of the clearest differences. A water-quality or treatment-management twin may remain useful when data are synchronized on the order of minutes, whereas a power fault or stability application can become ineffective if end-to-end latency exceeds the available physical response window. The correct comparison is therefore not whether one sector is "real time" and the other is not, but whether the complete sensing-to-decision chain is fast enough for the target physical process.

Digital-twin functions also differ in emphasis. Water applications commonly focus on process monitoring, hydraulic inference, water-quality support, leakage localization, maintenance, and operating optimization. Power applications more often emphasize state estimation, fast diagnosis, load or generation forecasting, voltage stability, equipment prediction, and controller development. The same high-level AI technique can therefore have very different operational meaning across the two sectors.

Implementation maturity is similarly function dependent. Water research includes real utility prototypes and physical distribution-network case studies, while power research includes distribution demonstrators, equipment field studies, substation scenarios, and physical microgrid transfer. Neither sector can be described uniformly as mature or experimental. A more defensible assessment is to compare the validation depth of individual functions.

Figure 3 summarizes these differences visually, while Table 4 presents a more detailed cross-sector matrix.

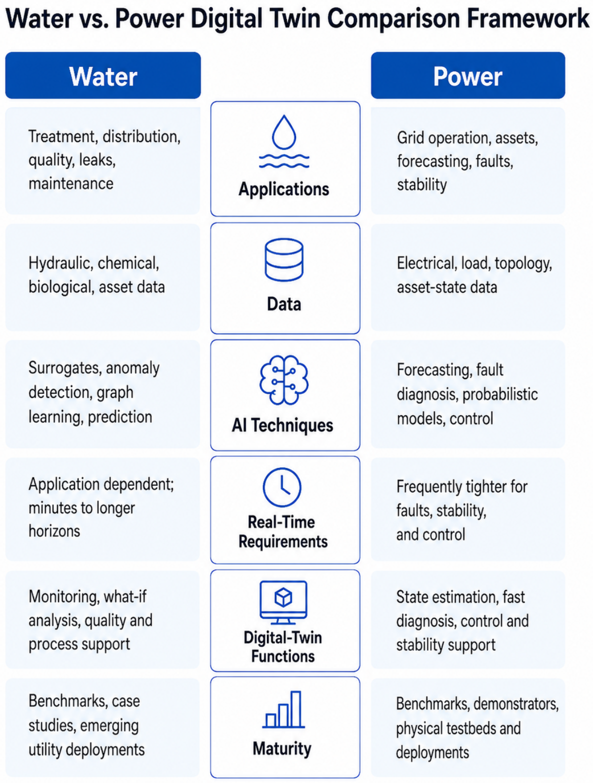

Figure 3. Water-versus-power comparison framework for AI-enabled digital twins.

The cross-sector evidence shows that the most transferable lessons are architectural and methodological rather than algorithm specific. Water research provides strong examples of physically constrained inference under sparse observability, while power research provides valuable examples of fast state analysis, uncertainty-aware stability modeling, online equipment adaptation, and virtual-to-physical controller transfer. Cross-sector learning should therefore focus on how these principles are validated and governed rather than simply transferring models without respecting the underlying physics.

A cross-sector comparison also reveals different sources of model mismatch. In water infrastructure, mismatch can arise from uncertain demand, pipe roughness, valve status, unmodeled leakage, water chemistry, or treatment-process variability. In power infrastructure, mismatch can arise from topology errors, parameter uncertainty, unobserved switching, changing generation, protection configuration, and equipment dynamics. These differences influence the type of uncertainty representation required and the measurements that provide the greatest value for state correction.

The sectors also differ in how operational benefit is measured. Water applications may emphasize reduced leakage, improved quality, lower energy use, fewer emergency repairs, or more efficient maintenance. Power applications may emphasize improved stability, reduced curtailment, faster fault localization, better load forecasting, reduced equipment stress, or more reliable control. A common evaluation framework should therefore preserve sector-specific outcome metrics

Table 4. Comparative analysis of AI-enabled digital twins in water and power infrastructure.

Dimension	Water	Power
Applications	Treatment, wastewater, desalination, distribution, leaks, quality, maintenance, optimization	Generation, smart grids, T&D, substations, renewables, forecasting, faults, maintenance, stability
Data	Hydraulic, chemical, biological, quality, demand, asset condition	Voltage, current, power, frequency, topology, load, generation, asset condition
AI techniques	Surrogates, graph learning, probabilistic inference, forecasting, RL, optimization	Deep/multimodal learning, probabilistic models, online learning, forecasting, RL, optimization
Real-time need	Application dependent; minute-to-hour horizons are acceptable for some processes	Frequently tighter for fault, stability, protection-adjacent analysis, and control
Twin functions	Process/network representation, quality support, leakage, maintenance, what-if analysis	State estimation, diagnosis, forecasting, equipment updating, stability, control development
Deployment maturity	Benchmarks, case studies, real-network demonstrations, utility prototypes	Benchmarks, equipment experiments, demonstrators, substation cases, physical testbeds
Dominant gap	Sparse sensing, heterogeneous data, model transfer, long-term model maintenance	End-to-end latency, topology change, mismatch, uncertainty, certification of automated control

while requiring shared reporting about latency, validation depth, uncertainty, and decision authority.

7. AI-ENABLED DIGITAL TWIN MATURITY AND DEPLOYMENT READINESS

A central analytical contribution of this review is the distinction between AI sophistication and deployment maturity. A complex learning architecture does not become an infrastructure-grade digital twin merely because it achieves high predictive accuracy. Maturity is instead defined through observable properties of the complete physical–virtual decision system: the persistence of the physical connection, synchronization frequency, validation environment, uncertainty treatment, decision authority, field duration, and cyber/operational assurance.

The six levels in Fig. 4 are therefore treated as evidence categories rather than as labels of technological prestige. Low-maturity L1–L2 entries are retained as DT-oriented or partial-twin implementations when they contribute relevant simulation, surrogate, synchronization, or validation evidence; the framework is explicitly used to prevent these studies from being interpreted as operational twins. Level 1 represents conceptual or simulation-only twins with no persistent live linkage. Level 2 adds repeatable benchmark or laboratory validation. Level 3 introduces site- or asset-specific case evidence using real or historically grounded data. Level 4 represents pilots and demonstrators with repeated or live updating in a realistic operational environment. Level 5 requires the twin to inform actual operational decisions under defined human authority. Level 6 requires sustained field operation with governed adaptation, re-validation, safe fallback, traceability, and explicit cyber/operational assurance.

Table 5 makes the model reproducible by defining observable criteria for every level. The classification is intentionally conservative: a study is assigned according to the strongest level for which the publication reports direct evidence. Capabilities that are proposed but not demonstrated do not raise the assigned maturity level.

The same criteria can be applied across sectors even though their physical timescales differ. A minute-scale water-quality twin can satisfy a mature synchronization requirement if that cadence is sufficient for its decision horizon, while a protection-adjacent power twin may require sub-second end-to-end latency. Maturity is therefore defined relative to the target operation, not by a universal update interval.

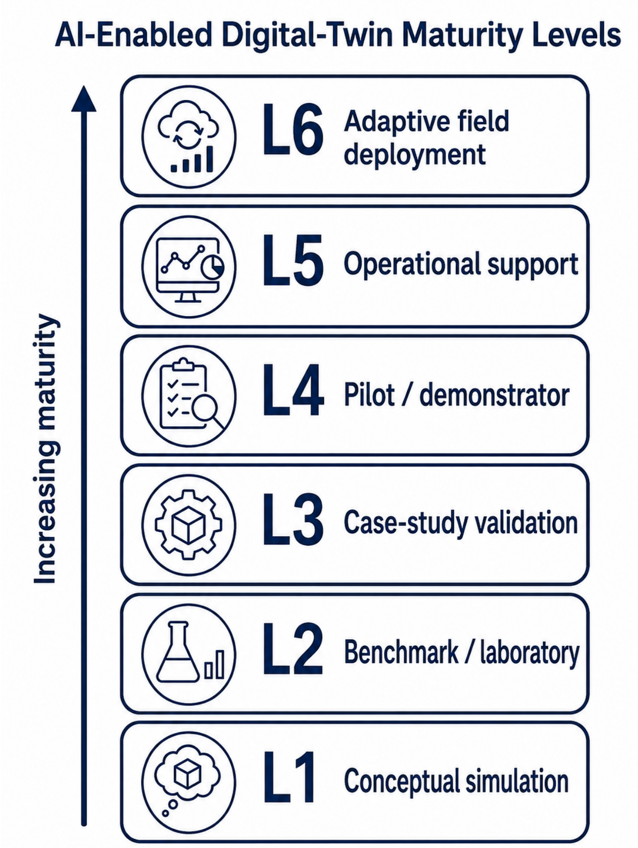

Figure 4. Six-level maturity model for AI-enabled digital twins, from conceptual simulation to adaptive field deployment.

Table 6 maps representative evidence from Tables 2 and 3 to the criteria. The mapping shows why deployment readiness is heterogeneous within each sector: real-plant demonstrations, operational demonstrators, deployed diagnostic twins, and physical-testbed transfer occupy higher levels than synthetic benchmarks, even when lower-level studies report stronger isolated AI metrics.

The maturity analysis therefore changes the interpretation of the evidence base. Under the conservative criteria in Table 5, L4 is the highest level directly demonstrated in the reviewed corpus; no study jointly establishes the sustained governance and lifecycle evidence required for L5 or L6. The principal gap is therefore long-duration evidence combining synchronization quality, uncertainty-aware decisions, controlled authority, cybersecurity, re-validation, and safe adaptation.

Table 5. Observable criteria for the six AI-enabled digital-twin maturity levels.

Level	Physical linkage / update	Validation and uncertainty	Decision authority / lifecycle evidence
L1	No persistent physical linkage; offline model or synthetic scenarios	Simulation evidence; uncertainty may be omitted	No operational authority; no field-duration or assurance evidence
L2	Controlled or scheduled data link in benchmark/lab setting	Repeatable laboratory or benchmark validation; limited uncertainty treatment	Offline or advisory analysis only; no sustained field evidence
L3	Application-specific link using real/historical data; synchronization may be periodic	Site/asset case validation; some sensitivity or uncertainty analysis	Advisory decision support; project-duration evidence, limited lifecycle governance
L4	Repeated or live updates in pilot/demonstrator	Realistic pilot validation with latency/data-quality monitoring	Bounded advisory role; partial cyber/operational controls and operator workflow demonstrated
L5	Persistent live linkage with stated update requirements	Operational validation with decision-relevant uncertainty or margins	Human-approved operational actions; sustained use with fallback, traceability, and maintenance procedures
L6	Governed continuous synchronization and adaptive updating	Repeated re-validation under drift, faults, and abnormal conditions	Bounded automation/authority; long-duration field evidence, safe fallback, cyber assurance, and lifecycle governance

Table 6. Representative maturity classification using the observable criteria in Table 5. A single conservative level is assigned according to the strongest directly demonstrated evidence.

Representative evidence	Validation setting	Decision authority	Assigned level
Water: desalination ML surrogate	Generated operating scenarios	Offline predictive support	L1
Water: GNN sensor-placement twin	Synthetic distribution network; no specific DT implementation	Design support	L1
Water: probabilistic leak localization	Controlled network test cases	Diagnostic advisory	L2
Water: real campus leakage twin	Field data and calibrated network	Operator-informed diagnosis	L3
Water: drinking-water utility prototypes	Operational utility prototypes	Operational advisory	L4
Water: full-scale PN/A aeration control	Real plant demonstration	Bounded automated process action	L4
Power: hybrid stability twin	Multi-case hybrid simulation	Offline engineering analysis	L2
Power: transformer online-learning twin	Dynamic equipment data	Maintenance support	L3
Power: distribution demonstrator	Real network demonstrator	Operational decision support	L4
Power: substation fault diagnosis	Simulation with historical SCADA records	Offline diagnostic evaluation	L2
Power: physical microgrid transfer	Physical DC-microgrid testbed	Supervised action transfer	L4
Overall corpus-level deployment gap	No reviewed study demonstrates all L5 or L6 criteria	Sustained fallback, traceability, re-validation, and governed authority are not jointly established	Not achieved (L5–L6 gap)

8. AI-ENABLED COUPLED WATER–POWER DIGITAL TWINS AND CASCADING RISK INTELLIGENCE

Water and power digital twins are usually developed as sector-specific systems, yet the physical infrastructures are interdependent. Water treatment, pumping, desalination, sensing, and control require electrical energy, while power generation and grid resilience can depend on water availability, cooling, hydropower, and water-related operating constraints. The analytical value of a coupled twin is therefore not obtained by merely placing two sector models side by side; it comes from representing the exchange points through which state changes and failures propagate between them.

The literature supports four dependency classes that are especially relevant to AI-enabled coupled twins. Table 7 distinguishes physical, cyber/data, geographic/spatial, and organizational/operational dependency. This distinction also clarifies the strength of evidence. Physical and spatial cascading pathways have direct water–power evidence, whereas cyber/data coupling is supported more broadly by critical-infrastructure interdependency research and remains less mature as a specifically coupled water–power digital-twin implementation.

Within this framework, AI has five specific roles. Early-warning models can combine weak precursors from both sectors; anomaly-correlation models can distinguish a local sensor problem from an interdependent disturbance; cascading-risk predictors can estimate downstream service consequences; coordinated-response models can compare actions across sector boundaries; and cyber-risk models can identify whether apparently physical anomalies are consistent with manipulated data or communication failure. The output should remain dependency aware: the same pump failure has different significance depending on storage, backup power, demand, and the state of the supplying electrical network.

Figure 5 represents this relationship as two sector-specific twins linked by an AI intelligence layer. The figure deliberately preserves separate water and power twins because their physics, operators, and decision responsibilities remain distinct. Coupling occurs through validated exchange variables and risk intelligence rather than through an undifferentiated monolithic model.

The review therefore treats coupled water–power digital twins as an evidence-grounded extension of sector twins, while also identifying a clear limitation: current literature is stronger on conceptual interdependency and modeled cascading pathways than on long-duration, jointly operated AI-enabled water–power twins. This boundary is important because it prevents the proposed framework from being presented as a level of deployment that the literature has not yet demonstrated.

9. TRUSTWORTHY AI, INTEROPERABILITY, AND CERTIFICATION PATHWAYS

Trustworthy AI, interoperability, and certification should not be treated as a single evidence category. This review separates three layers explicitly. The first is *implemented evidence*: capabilities actually demonstrated in the primary water and power studies, such as probabilistic leak localization, online model updating, explainable transformer-health prediction, real demonstrators, and supervised transfer from a twin to a physical microgrid. The second is *standards and governance guidance*: requirements or recommendations for AI risk management, lifecycle control, cybersecurity, asset management,

Table 7. Evidence basis for coupled water–power dependency classes and their digital-twin implications.

Dependency class	What crosses the sector boundary	Representative evidence	Coupled-twin implication
Physical	Electricity for pumping/treatment; water for generation/cooling and related services	Water–power resilience literature documents bidirectional service dependency and cascading effects [55]	Model exchange points, service thresholds, backup capacity, and restoration interactions
Cyber/data	Measurements, communications, control information, and digital service availability	General interdependent-infrastructure evidence shows failures can propagate through multiple dependency mechanisms [56]	Correlate cross-domain anomalies while treating cyber/data coupling as an evidence gap requiring dedicated water–power validation
Geographic / spatial	Co-located hazards, shared corridors, and spatially coupled damage	Coupled synthetic Phoenix power and water networks reveal cascading outage patterns after transmission-line failures; site-specific prediction requires real-network validation [57]	Explore cross-network disruption patterns and prioritize protective buffers
Organizational / operational	Restoration priorities, dispatch, emergency actions, and co-managed resources	A water–power digital-twin energy-hub framework demonstrates real-time vulnerability and coordinated resilience logic [58]	Support coordinated response while preserving sector-specific authority and operating constraints

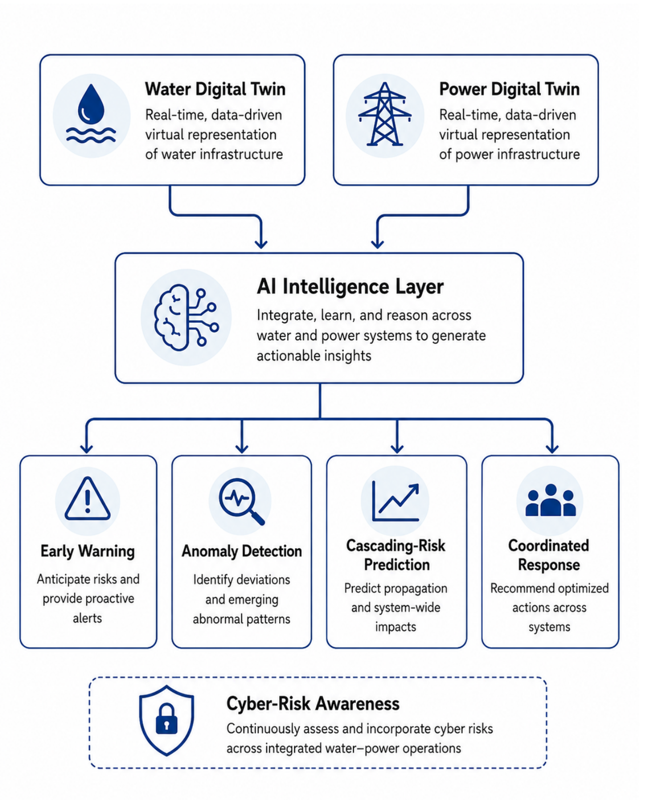

Figure 5. Coupled water–power digital-twin framework in which an AI intelligence layer supports cross-sector early warning, anomaly detection, cascading-risk prediction, coordinated response, and cyber-risk awareness.

and information exchange. The third is the *deployment pathway proposed by this review*: a structured way to combine evidence from the first two layers into decision-authority gates, re-validation triggers, fallback behavior, and an auditable readiness case. Keeping these layers separate prevents standards from being misrepresented as empirical deployment evidence and prevents research proposals from being described as already implemented practice.

Implemented evidence shows that individual trust mechanisms are technically feasible but fragmented. Some primary studies quantify uncertainty, some update models online, some add explainability, and some demonstrate operator or physical-testbed interaction. Very few demonstrate all of these capabilities within one long-duration field twin. The immediate implication is that trustworthy deployment should be assessed as a system property rather than inferred from one model metric.

Explainability is useful when it supports an operational check: why an alert was raised, which evidence drove a recommendation, whether a prediction is physically plausible, and whether the model is operating outside its validated region. Uncertainty has a complementary function. It should not be reported merely as an interval; it should change what the system is allowed to do. Low-confidence monitoring may trigger inspection, while low-confidence high-consequence control should reduce automated authority, request additional observations, or return control to a conservative fallback policy.

Human involvement is therefore represented as a graded authority structure rather than a binary human-versus-autonomy choice. Figure 6 separates observation, recommendation, explanation, approval, conditional automation, and autonomous action. Movement upward requires stronger evidence, clearer constraints, and more demanding assurance.

Model traceability and re-validation are equally important because the digital twin changes over time. Training data, calibration parameters, sensor configurations, software versions, model weights, and operating assumptions should be traceable to the deployed decision output. A significant change in asset condition, topology, sensing, data distribution, or AI model version should trigger a defined re-validation process rather than silent continuation of the previous assurance claim.

Figure 7 frames this requirement as lifecycle governance. Validation is not a one-time pre-deployment test; it is connected to monitoring, controlled updating, and retirement. Safe fallback provides the boundary condition: if synchronization fails, uncertainty exceeds the validated range, or a cyber event compromises evidence integrity, the twin should degrade to a defined safe mode instead of preserving the same automation level.

Interoperability determines whether these assurance mechanisms can operate across the complete twin. Water systems may need to connect process historians, supervisory control, laboratory results, hydraulic simulators, maintenance systems, and quality databases. Power twins may connect state estimation, asset systems, substation communication, topology models, planning tools, and distributed-energy resources. Interface compatibility alone is insufficient: exchanged data must retain physical meaning, units, time semantics, provenance, and quality status.

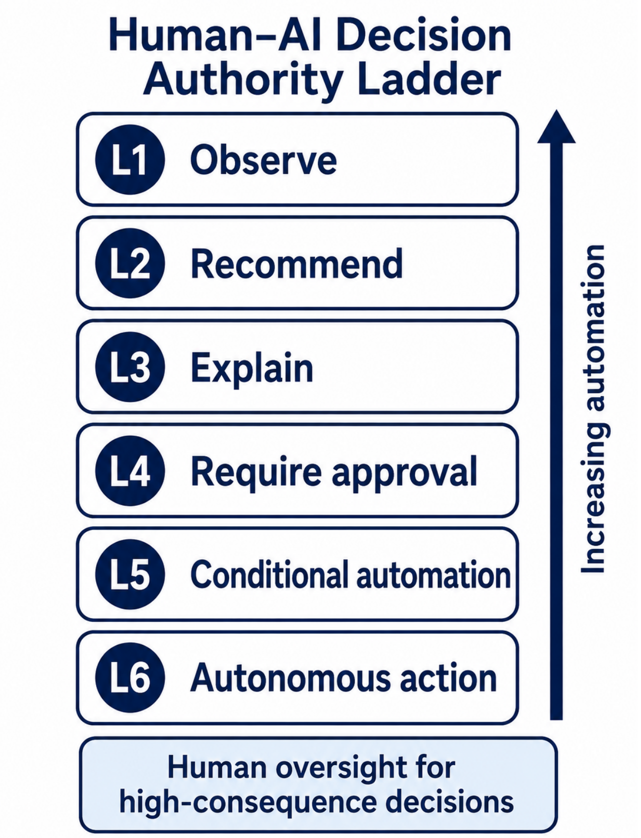

Figure 6. Human–AI decision-authority ladder for progressively higher levels of automated action.

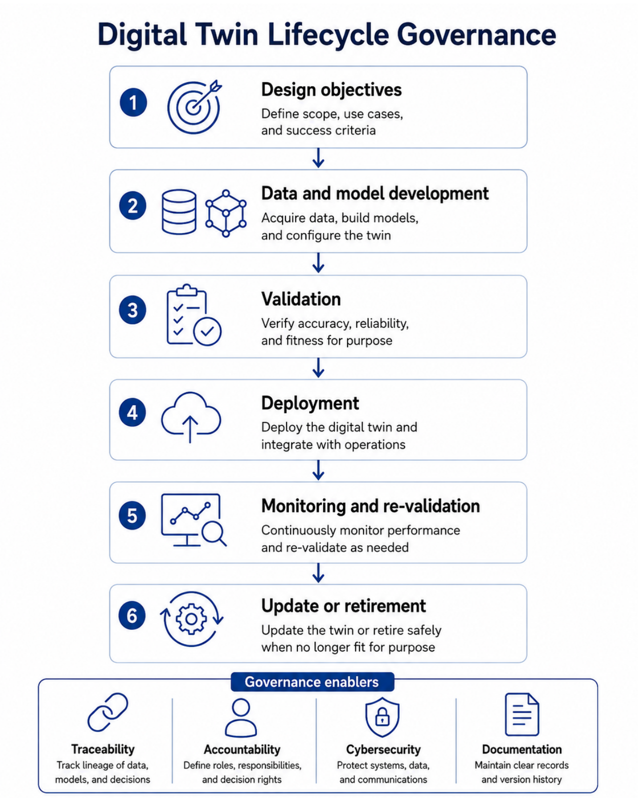

Figure 7. Lifecycle governance for AI-enabled digital twins, linking development and deployment to traceability, re-validation, controlled updating, and retirement.

10. CHALLENGES AND RESEARCH GAPS

The preceding sections separate deployment maturity and assurance governance from the remaining technical research gaps. This section therefore focuses on unresolved empirical and engineering barriers that continue to limit the fidelity, responsiveness, and generalizability of AI-enabled digital twins after a use case and authority model have already been defined.

Data quality and observability remain fundamental because a synchronized twin can still be wrong when sensors drift, values are missing, or important state variables are unobserved. The research need is decision-aware evidence quality: the twin should indicate when the available observations are insufficient for the requested inference rather than silently imputing a high-confidence state.

Synchronization is a separate issue from model accuracy. The relevant metric is the age and consistency of the virtual state at the time of decision. Water applications may tolerate slower updates for some quality or planning tasks, whereas power faults and stability decisions can require much tighter latency. Reporting should therefore decompose acquisition, communication, state updating, inference, optimization, and actuation instead of describing a system simply as “real time.”

Computational cost becomes a research constraint when high-fidelity models cannot run within the decision horizon. Surrogates, reduced-order models, edge inference, and distributed computation can reduce latency, but simplification should be tested under disturbances and operating shifts to ensure that the physical relationships needed for the decision are preserved.

Scalability is similarly multidimensional. A component twin may scale computationally yet fail when extended to a network with changing topology, sparse sensing, heterogeneous assets, and multiple data owners. Research should therefore evaluate how state estimation, communication, model updating, and data management behave as network size and heterogeneity increase.

Cybersecurity remains an empirical challenge even when governance controls are defined. AI-enabled twins must distinguish physical faults, sensor failures, communication disturbances, and malicious manipulation because these conditions can produce similar residuals while requiring different responses. Figure 8 treats this as a closed evidence-recovery problem rather than an isolated classifier.

The evidence gap that cuts across all of these issues is validation under change. Most studies validate under a bounded dataset, fixed topology, controlled testbed, or limited field period. Far fewer expose the twin to seasonal drift, sensor replacement, maintenance, topology changes, rare events, communication outages, and adversarial conditions over a sustained period. Table 8 summarizes the principal research gaps without repeating the maturity and certification criteria defined earlier.

These empirical gaps are carried directly into the research priorities in Section 11, without introducing additional maturity criteria.

Table 8. Unresolved technical challenges and research gaps in AI-enabled water and power digital twins.

Challenge	Operational effect	Current technical responses	Remaining research gap
Data quality / observability	Incorrect or weakly supported virtual state	Calibration, fusion, imputation, probabilistic estimation	Decision-aware tests of whether evidence is sufficient for the intended inference
Synchronization	Correct information can arrive too late for the physical event	Streaming, online updating, event-triggered synchronization	Application-specific bounds linking state age and end-to-end latency to decision validity
Computation	High-fidelity simulation can violate response-time constraints	Surrogates, reduced-order models, edge/distributed computing	Demonstrated speed–fidelity trade-offs under disturbances and distribution shift
Scalability	Component-level methods degrade on large heterogeneous networks	Hierarchical twins, graph models, modular services	Infrastructure-scale tests with changing topology, sparse sensing, and multiple owners
Uncertainty	Point estimates conceal model, sensor, and operating risk	Bayesian models, ensembles, Gaussian processes, calibrated intervals	Evidence that uncertainty remains calibrated under regime change and rare events
Cyber/physical ambiguity	Faults, bad data, and attacks can produce similar signatures	Anomaly detection, attack simulation, residual analysis	Causal or multi-source diagnosis that identifies the source of abnormal evidence
Long-duration validation	Short studies do not reveal drift, maintenance, aging, or recovery behavior	Pilots, demonstrators, hardware-in-the-loop, field cases	Sustained studies covering non-nominal operation, adaptation, and failure recovery

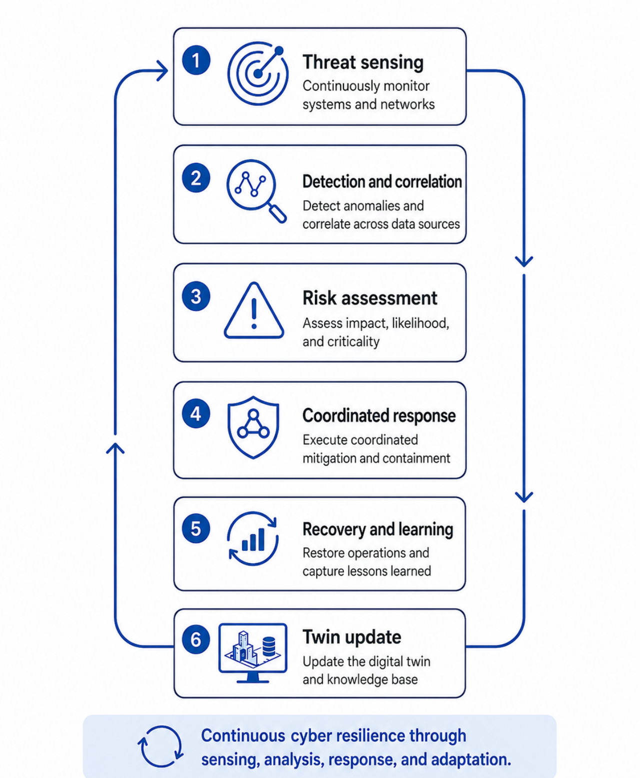

Figure 8. Cyber-resilient response loop connecting threat sensing, detection, risk assessment, response, recovery, and twin updating.

11. FUTURE RESEARCH DIRECTIONS

Future research should target evidence that directly closes the technical gaps above rather than simply adding more complex AI architectures. Five immediate research priorities are especially important.

First, end-to-end timing should become a standard experimental variable. Studies should report acquisition delay, communication delay, synchronization/update time, AI inference time, optimization time, and actuation delay. This makes it possible to determine whether a twin is fast enough for its specific physical decision rather than relying on the ambiguous label "real time."

Second, anomaly detection should evolve toward causal diagnosis. Future twins should distinguish among equipment faults, sensor drift, communication failure, cyber manipulation, and legitimate regime changes by combining physical constraints, cross-sensor consistency, and temporal context. For coupled infrastructure, the same principle should extend across validated water–power exchange variables.

Third, scalable architectures should be evaluated under network growth and heterogeneity. Edge AI can reduce communication dependence for pumps, treatment units, substations, and distributed energy resources, but the research question is how local models remain consistent with higher-level state and how synchronization cost grows as assets are added. Comparative studies should therefore report computational and communication scaling, not only prediction accuracy.

Fourth, field trials should be designed around change rather than nominal operation. Multi-season experiments should include maintenance, sensor replacement, model drift, topology changes, abnormal events, and communication loss. Such trials provide the evidence needed to determine when recalibration or re-training is required and whether a twin recovers reliably after disruption.

Fifth, benchmark suites should progress from static AI tests to staged digital-twin tests. Early benchmark levels can assess state fidelity and prediction; later levels should inject missing sensors, delay, drift, topology change, cyber events, human intervention, and closed-loop decision consequences. This would align benchmark difficulty with the maturity criteria developed in this review.

The power-grid review by Al-Shetwi et al. summarized in Table 1 highlights real-time monitoring, predictive maintenance, and interoperability as priorities for reliable DT adoption. Building on these priorities, future evaluation should test disturbance propagation and recovery in water systems and coupled water–power settings where hazards affect shared dependencies.

Figure 9 summarizes these priorities as a progression from current gaps to research enablers and operational evidence.

Evaluation should also move beyond a single prediction metric. Figure 10 summarizes six complementary dimensions used in the revised framework: model fidelity, data quality, synchronization, latency, robustness, and operational value. These dimensions are intentionally framed as evaluation outputs rather than maturity labels. They allow two twins at the same maturity level to be compared on technical performance without conflating performance with decision authority.

A useful benchmark report should therefore state which variables are measured or estimated, how prediction uncertainty

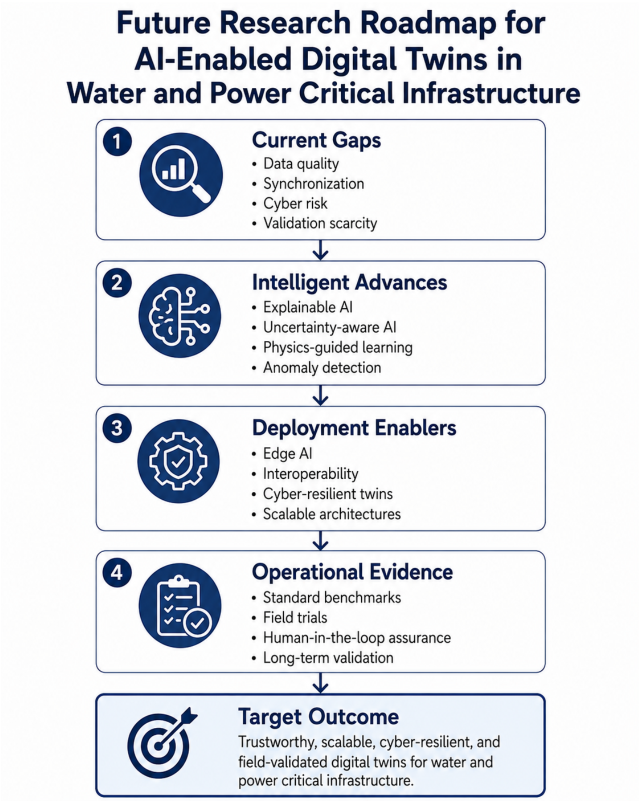

Figure 9. Future research roadmap for AI-enabled digital twins in water and power critical infrastructure.

Figure 10. Recommended evaluation dimensions for infrastructure-grade AI-enabled digital twins.

is calibrated, how latency is measured end to end, how robustness is tested under missing/corrupted data, and what decision outcome changes when the twin output is used. The central future-research objective is not merely a more accurate virtual model, but evidence that the twin remains useful when the infrastructure, data, and operating context change.

12. CONCLUSION

This review establishes that AI-enabled digital twins in water and power share a common physical–virtual intelligence architecture but differ materially in observability, data cadence, latency, topology, process dynamics, and decision consequence. The cross-sector evidence therefore supports sector-specific implementation rather than direct algorithm transfer.

The review's main contribution is an evidence-structured comparison that links application studies to deployment maturity, coupled water–power cascading risk, and trustworthy-AI governance. The six-level maturity model separates simulation and partial-twin evidence from pilots, operational support, and field deployment, while the coupled framework identifies physical, cyber/data, spatial, and organizational dependencies that isolated sector twins cannot represent adequately.

The dominant unresolved gap is sustained field evidence. Under the conservative maturity criteria used here, no reviewed study fully satisfies L5 or L6 because continuous synchronization, uncertainty-aware decisions, governed authority, cyber assurance, safe fallback, traceable adaptation, and repeated re-validation are not jointly demonstrated over long operating periods. Progress toward infrastructure-grade digital twins should therefore be judged primarily by the quality of this lifecycle evidence rather than by isolated model accuracy.

REFERENCES

[1] Z. Zhang, F. Wen, Z. Sun, X. Guo, T. He, and C. Lee, "Artificial Intelligence-Enabled Sensing Technologies in the 5G/Internet of Things Era: From Virtual Reality/Augmented Reality to the Digital Twin," *Advanced Intelligent Systems*, vol. 4, no. 7, p. 2100228, 2022.

[2] M. Groshev, C. Guimarães, J. Martín-Pérez, and A. de la Oliva, "Toward intelligent cyber-physical systems: Digital twin meets artificial intelligence," *IEEE Communications Magazine*, vol. 59, no. 8, pp. 14–20, 2021.

[3] M. M. Rathore, S. A. Shah, D. Shukla, E. Bentafat, and S. Bakiras, "The role of AI, machine learning, and big data in digital twinning: A systematic literature review, challenges, and opportunities," *IEEE Access*, vol. 9, pp. 32 030–32 052, 2021.

[4] P. Radanliev, D. De Roure, R. Nicolescu, M. Huth, and O. Santos, "Digital twins: artificial intelligence and the IoT cyber-physical systems in Industry 4.0," *International Journal of Intelligent Robotics and Applications*, vol. 6, no. 1, pp. 171–185, 2022.

[5] P. Rakshit, N. Saha, S. Nandi, and P. Gupta, "Artificial Intelligence in Digital Twins for Sustainable Future," in *Transforming Industry using Digital Twin Technology*, A. Mishra, M. El Barachi, and M. Kumar, Eds. Cham: Springer Nature Switzerland, pp. 19–44, 2024.

[6] P. Ghorbani Bam, N. Rezaei, A. Roubanis, D. Austin, E. Austin, B. Tarroja, I. Takacs, K. Villez, and D. Rosso, "Digital twin applications in the water sector: A review," *Water*, vol. 17, no. 20, p. 2957, 2025.

[7] A.-J. Wang, H. Li, Z. He, Y. Tao, H. Wang, M. Yang, D. Savic, G. T. Daigger, and N. Ren, "Digital twins for wastewater treatment: A technical review," *Engineering*, vol. 36, pp. 21–35, 2024.

[8] T. C. Pereira, A. G. Andrade-Campos, and R. Vilanova Arbos, "The synergy between digital twin, machine learning, and control technologies in water supply systems: A critical literature review," *IEEE Access*, vol. 13, pp. 208 227–208 261, 2025.

[9] A. Ranawaka, D. Alahakoon, Y. Sun, and K. Hewapathirana, "Leveraging the synergy of digital twins and artificial intelligence for sustainable power grids: A scoping review," *Energies*, vol. 17, no. 21, p. 5342, 2024.

[10] J. B. Heluany and V. Gkioulos, "A review on digital twins for power generation and distribution," *International Journal of Information Security*, vol. 23, pp. 1171–1195, 2024.

[11] T. Itäpelto, M. Elhajj, and M. van Sinderen, "Digital twin application in lifecycle security of critical infrastructures: A systematic literature review," *International Journal of Critical Infrastructure Protection*, vol. 50, p. 100783, 2025.

[12] A. Q. Al-Shetwi, I. E. Atawi, M. A. El-Hameed, and A. Abuelrub, "Digital twin technology for renewable energy, smart grids, energy storage and vehicle-to-grid integration: Advancements, applications, key players, challenges and future perspectives in modernising sustainable grids," *IET Smart Grid*, vol. 8, no. 1, p. e70026, 2025.

[13] S. M. Zayed, G. M. Attiya, A. El-Sayed, and E. E.-D. Hemdan, "A review study on digital twins with artificial intelligence and internet of things: concepts, opportunities, challenges, tools and future scope," *Multimedia Tools and Applications*, vol. 82, no. 30, pp. 47 081–47 107, 2023.

[14] Z. Lv and S. Xie, "Artificial intelligence in the digital twins: State of the art, challenges, and future research topics: [version 2; peer review: 2 approved]," *Digital Twin*, vol. 1, no. 1, p. 12, 2022.

[15] S. M. M. Sajadieh and S. D. Noh, "From Simulation to Autonomy: Reviews of the Integration of Artificial Intelligence and Digital Twins," *International Journal of Precision Engineering and Manufacturing-Green Technology*, vol. 12, no. 5, pp. 1597–1628, 2025.

[16] J. Zhu, Y. Yang, M. Xi, S. Ji, L. Jia, and T. Hu, "The next-generation digital twin: from advanced sensing towards artificial intelligence-assisted physical-virtual system," *Journal of Industrial Information Integration*, vol. 48, p. 100942, 2025.

[17] K. Alexopoulos, N. Nikolakis, and G. Chryssolouris, "Digital twin-driven supervised machine learning for the development of artificial intelligence applications in manufacturing," *International Journal of Computer Integrated Manufacturing*, vol. 33, no. 5, pp. 429–439, 2020.

[18] Z. Ma, Y. Zhu, C. Chen, T. Li, Y. Li, X. Li, Y. Wang, T. D. Waite, and J. Guan, "Towards the digitalization of water treatment facilities: A case study on machine learning-enabled digital twins," *Journal of Water Process Engineering*, vol. 77, p. 108316, 2025.

[19] S. Baena-Miret, M. Alet Puig, R. Bardisa Rodes, L. Bonastre Farran, S. Durán, M. Ganzer Martí, E. Martínez-Gomariz, and A. Carrasco Valverde, "Enhancing efficiency and quality control: The impact of digital twins in drinking water networks," *Water Environment Research*, vol. 96, no. 10, p. e11139, 2024.

[20] W. Liu, S. He, J. Mou, T. Xue, H. Chen, and W. Xiong, "Digital twins-based process monitoring for wastewater treatment processes," *Reliability Engineering & System Safety*, vol. 238, p. 109416, 2023.

[21] A. Daaboub, L. Echeverria Rovira, and E. Rubion Soler, "Enhancing seawater reverse osmosis desalination efficiency using digital twins and machine learning," in *Artificial Intelligence Research and Development: Proceedings of the 26th International Conference of the Catalan Association for Artificial Intelligence, CCIA 2024, Barcelona, Spain, 2–4 October 2024*, ser. Frontiers in Artificial Intelligence and Applications, T. Alsinet, X. Vilasís-Cardona, D. García, and E. Álvarez, Eds., vol. 390. IOS Press, pp. 2–5, 2024.

[22] P. Brahmbhatt, A. Maheshwari, and R. D. Gudi, "Digital twin assisted decision support system for quality regulation and leak localization task in large-scale water distribution networks," *Digital Chemical Engineering*, vol. 9, p. 100127, 2023.

[23] N. T. Mücke, P. Pandey, S. Jain, S. M. Bohté, and C. W. Oosterlee, "A probabilistic digital twin for leak localization in water distribution networks using generative deep learning," *Sensors*, vol. 23, no. 13, p. 6179, 2023.

[24] A. Singh, A. Maheshwari, and S. Singh, "Digital twin framework for leakages detection in large-scale water distribution systems: A case study of IIT-Jodhpur campus," *IFAC-PapersOnLine*, vol. 57, pp. 280–285, 2024.

[25] C. El Hachimi, S. Belaqziz, S. Khabba, H. Karjoun, M. H. Kharrou, B. Ait Hssaine, S. Er-Raki, A. Daccache, A. Chehbouni, and Q. Weng, "Towards collective intelligence in agriculture: Deep reinforcement learning and digital twins for efficient management of collective irrigation water distribution systems," *Energy Nexus*, vol. 20, p. 100599, 2025.

[26] R. Patriarca, F. Simone, and G. Di Gravio, "Modelling cyber resilience in a water treatment and distribution system," *Reliability Engineering & System Safety*, vol. 226, p. 108653, 2022.

[27] J. Zhang, N. Tylor, W. Zhao, R. Rui, and C. He, "Enhancing coagulation prediction in water treatment using a similarity score-based piecewise machine learning model," *Journal of Environmental Engineering*, vol. 151, no. 6, p. 04025023, 2025.

[28] Y. Wei, A. W.-K. Law, C. Yang, and D. Tang, "Combined anomaly detection framework for digital twins of water treatment facilities," *Water*, vol. 14, no. 7, p. 1001, 2022.

[29] Y. Wei, A. W.-K. Law, and C. Yang, "Real-time data-processing framework with model updating for digital twins of water treatment facilities," *Water*, vol. 14, no. 22, p. 3591, 2022.

[30] T. A. Syed, M. A. Muhammad, A. A. AlShahrani, M. Hammad, and M. T. Naqash, "Smart water management with digital twins and multimodal transformers: A predictive approach to usage and leakage detection," *Water*, vol. 16, no. 23, p. 3410, 2024.

[31] Y. Xu, M. Hui, and H. Qu, "Design of a 3D platform for the evaluation of water quality in urban rivers based on a digital twin model," *Water*, vol. 16, no. 24, p. 3668, 2024.

[32] A. Menapace, A. Zanfei, M. Herrera, and B. Brentan, "Graph neural networks for sensor placement: A proof of concept towards a digital twin of water distribution systems," *Water*, vol. 16, no. 13, p. 1835, 2024.

[33] A. Singh, S. Singh, and A. Maheshwari, "A proof-of-concept study towards developing digital twins for operational excellence in large-scale water distribution networks," *Urban Water Journal*, vol. 22, no. 5, pp. 507–525, 2025.

[34] S. Jun and D. Jung, "Exploration of deep learning leak detection model across multiple smart water distribution systems: Detectable leak sizes with AMI meters," *Water Research X*, vol. 29, p. 100332, 2025.

[35] S. K. Heo, T. Oh, T. Y. Woo, S. Y. Kim, Y. Choi, M. Park, J. Kim, and C. K. Yoo, "Real-scale demonstration of digital twins-based aeration control policy optimization in partial nitritation/anammox process: Policy iterative dynamic programming approach," *Desalination*, vol. 593, p. 118235, 2025.

[36] A. Kavousi-Fard, M. Dabbaghjamanesh, M. Jafari, M. Fotuhi-Firuzabad, Z. Y. Dong, and T. Jin, "Digital twin for mitigating solar energy resources challenges: A perspective," *Solar Energy*, vol. 274, p. 112561, 2024.

[37] M. Deakin, M. Vanin, Z. Fan, and D. Van Hertem, "Smart energy network digital twins: Findings from a uk-based demonstrator project," *International Journal of Electrical Power & Energy Systems*, vol. 162, p. 110302, 2024.

[38] T. Wu, F. Yang, U. Farooq, X. Li, and J. Jiang, "An online learning method for constructing self-update digital twin model of power transformer temperature prediction," *Applied Thermal Engineering*, vol. 237, p. 121728, 2024.

[39] M. Haida, M. Stebel, P. Lasek, R. Fingas, R. Krok, J. Bodys, M. Palacz, J. Smolka, P. Jachymek, and W. Adamczyk, "A hybrid reduced-order model of a large-scale generator and power transformer applied in an artificial intelligence-supported power plant control system," *Energy Conversion and Management*, vol. 333, p. 119788, 2025.

[40] L. Lang, J. Du, J. Guo, and G. Xue, "A digital twin-based method for monitoring the whole process of automated equipment operation and maintenance," in *Design Studies and Intelligence Engineering: Proceedings of DSIE 2024*. IOS Press, pp. 154–162, 2025.

[41] Y. Yin, H. Chen, X. Meng, and H. Xie, "Digital twin-driven identification of fault situation in distribution networks connected to distributed wind power," *International Journal of Electrical Power & Energy Systems*, vol. 155, p. 109415, 2024.

[42] D. K. Panda and S. Das, "Gaussian process digital twin for voltage stability analysis of complex power networks under perturbation," *Expert Systems with Applications*, vol. 267, p. 126151, 2025.

[43] Z. Wang, R. Cong, R. Wang, and Z. Wang, "Digital twin for power load forecasting," in *Proceedings of the 2023 International Conference on Wireless Communications, Networking and Applications*, ser. Lecture Notes in Electrical Engineering, vol. 1361. Springer Nature, pp. 366–374, 2025.

[44] M. Seyedi, S. Ebrahimi, K. Sheida, and F. Ferdowsi, "Digital twin-enabled transfer learning of RL-driven control for DC microgrids," *Digital Twins and Applications*, vol. 2, no. 1, p. e70012, 2025.

[45] Z. Jing, Q. Wang, Z. Chen, T. Cao, and K. Zhang, "Optimization of energy acquisition system in smart grid based on artificial intelligence and digital twin technology," *Energy Informatics*, vol. 7, no. 1, p. 121, 2024.

[46] A. G. Abo-Khalil, "Digital twin real-time hybrid simulation platform for power system stability," *Case Studies in Thermal Engineering*, vol. 49, p. 103237, 2023.

[47] N. El-Rashidy, Y. A. Sultan, and Z. H. Ali, "Predecting power transformer health index and life expectation based on digital twins and multitask LSTM-GRU model," *Scientific Reports*, vol. 15, p. 1359, 2025.

[48] A. E. Onile, E. Petlenkov, Y. Levron, and J. Belikov, "Smartgrid-based hybrid digital twins framework for demand side recommendation service provision in distributed power systems," *Future Generation Computer Systems*, vol. 156, pp. 142–156, 2024.

[49] N. Al-Qirim, A. Bani-Hani, M. Majdalawieh, H. Al Hamadi, and M. K. Hasan, "Digraph-enabled digital twin and label-encoding machine learning for SCADA network's cyber attack analysis in industry 5.0," *IEEE Open Journal of the Communications Society*, vol. 6, pp. 3404–3416, 2025.

[50] J. Tan, R. M. Radhi, K. Shirini, S. S. Gharehveran, Z. Parisooz, M. Khosravi, and H. Azarinfar, "Innovative framework for fault detection and system resilience in hydropower operations using digital twins and deep learning," *Scientific Reports*, vol. 15, p. 15669, 2025.

[51] J. Shen, L. Hu, Y. Yang, Y. Li, and P. Lou, "Real-time update algorithms for digital twin models of distribution network equipment under internet of things and optical imaging technology," *Scientific Reports*, vol. 15, p. 5910, 2025.

[52] L. Jiang, D. Wang, W. Sun, S. Bao, H. Liu, and S. Chang, "Research on transformer fault diagnosis method based on digital twin," *Journal of System Simulation*, vol. 37, no. 3, pp. 775–790, 2025.

[53] G. He, M. Lei, L. Han, P. Shan, and R. Chen, "Energy-efficient smart grid operations through dynamic digital twin models and deep learning," *Sustainable Computing: Informatics and Systems*, vol. 48, p. 101200, 2025.

[54] Z. He, H. Yu, L. Peng, A. Zhou, H. Wang, and J. Xu, "Digital twin topology modelling method of new-type distribution network based on CIM specifications and spectral clustering," *Global Energy Interconnection*, vol. 8, no. 6, pp. 947–958, 2025.

[55] K. Oikonomou, K. Mongird, J. S. Rice, and J. S. Homer, "Resilience of interdependent water and power systems: A literature review and conceptual modeling framework," *Water*, vol. 13, no. 20, p. 2846, 2021.

[56] M. Ouyang, "Review on modeling and simulation of interdependent critical infrastructure systems," *Reliability Engineering & System Safety*, vol. 121, pp. 43–60, 2014.

[57] R. Hoff, R. Sparks, M. Chester, A. Mustafa, N. Johnson, A. Birchfield, T. McPhearson, R. Li, N. Ahmad, and I. Searles, "Cascading failure propagation and perfect storms in interdependent infrastructures," *ASCE OPEN: Multidisciplinary Journal of Civil Engineering*, vol. 3, no. 1, pp. 1–15, 2025.

[58] G. Nasiri and A. Kavousi-Fard, "A digital twin-based system to manage the energy hub and enhance the electrical grid resiliency," *Machines*, vol. 11, no. 3, p. 392, 2023.