hydrology

MDPI

Systematic Review

Digital Twins for Sustainable Groundwater Resources Management: From Monitoring and Prediction to Governance and Resilience—A Review

Iolanda Borzì

Department of Mathematics, Computer Sciences, Physics and Earth Sciences, The University of Messina, 98166 Messina, Italy; iolanda.borzi@unime.it

Abstract

This article presents a scoping review of the literature on digital twins (DTs) for sustainable groundwater resources management, which constitutes a very recent and rapidly expanding research field, with literature moving quickly from conceptual frameworks to application-oriented systems. The literature, selected through the PRISMA 2020 methodology, is organized into seven sub-topics: AI and ML foundations, digital twin architectures and frameworks, aquifer-scale DT applications, agricultural and water–energy–food (WEF) nexus DTs, basin and urban water DTs, sensing, monitoring and IoT infrastructures, and governance, resilience and socio-hydrology. This structure shows how the field is shifting from monitoring and prediction toward integrated decision support, where process-based models, machine learning surrogates, real-time sensing and optimization are combined to support drought mitigation, saltwater intrusion control, irrigation management, climate adaptation and basin-scale planning. Across the reviewed studies, the most recurrent contributions are the construction of hybrid model architectures, the use of DTs to close the loop between observation and control, and the growing recognition that groundwater management must incorporate governance, stakeholder decision-making and socio-hydrological feedbacks. At the same time, the literature still faces key limitations, especially uncertainty quantification, interoperability between models and data streams, transferability to data-scarce settings and limited validation under real operational conditions. Future research should therefore focus on physics-informed and explainable AI, federated and scalable DT architectures, stronger coupling with socio-hydrological and governance frameworks, and more field-tested implementations that can demonstrate robust performance across diverse hydrogeological and institutional contexts.

Keywords: artificial intelligence; aquifer sustainability; decision support; digital twin; groundwater resources management; hydroinformatics; machine learning; real-time monitoring; socio-hydrology; water governance

Check for updates

Academic Editor: Minxue He

Received: 16 July 2026
Revised: 16 August 2026
Accepted: 24 August 2026
Published: 31 August 2026

Copyright: © 2026 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

1.1. Digital Twins in Water Resources Management

Digital twin (DT) technology has rapidly evolved from its industrial origins into a broader environmental systems paradigm centered on the creation of dynamic digital counterparts of physical assets and processes [1,2]. In water resources management, DTs are increasingly understood not simply as numerical models or visualization platforms, but as integrated cyber-physical systems that combine monitoring, simulation, prediction, data

assimilation, and decision support in a continuously updated operational environment [3,4]. This evolution is especially relevant to hydrology, where water systems are inherently dynamic, multi-scalar, and shaped by strong interactions among climate, infrastructure, ecosystems, and human management [3,5,6].

Within the broader water sector, DT applications have expanded across a wide variety of domains. At the river-basin scale, recent studies describe DTs as instruments for sustainable allocation, flood forecasting, disaster mitigation, and climate adaptation, with the capacity to integrate observations, hydrological models, and policy-oriented scenario analysis in a single analytical framework [3,5,7–9]. In parallel, urban water applications have advanced rapidly, including DTs for drinking-water distribution systems, sewer and drainage networks, smart pumping operations, and water-quality control. In many of these urban cases, hydraulic engines such as EPANET provide the computational basis for representing network behavior, while IoT sensing, optimization algorithms, and visualization interfaces convert those models into operational twins capable of supporting leakage detection, pressure management, energy savings, and predictive maintenance [9–15].

This broadening of DT applications reflects a deeper change in how water systems are conceptualized and governed. Conventional modeling approaches have long played a central role in hydrology and water engineering, but they often remain detached from real-time observations and from the operational logic of management institutions [3,4]. In contrast, DTs seek to establish a tighter coupling between the physical system and its digital representation, allowing managers not only to describe current conditions but also to test alternative interventions, anticipate risks, and update decisions as new data becomes available [5,6]. In this sense, DTs represent a transition from static or episodic analysis toward adaptive and increasingly intelligent water management infrastructures [3].

1.2. Growing Relevance and Why Groundwater Is Emerging as a Distinct DT Frontier

The growing importance of DTs in water management is closely linked to the intensifying pressures affecting contemporary hydrological systems. Climate variability and climate change are altering recharge, runoff generation, flood frequency, drought severity, and water demand patterns, while urban expansion, agricultural intensification, infrastructure aging, and competing sectoral uses are increasing the complexity of water governance [7,9,11–15]. Under these conditions, traditional planning tools, although still indispensable, are often too static, too slow, or too compartmentalized to support timely and integrated decisions [3,6].

Recent reviews of DTs in the water sector consistently emphasize several advantages that explain the rapid rise in the field [3,5,7]. First, DTs can support near-real-time system awareness by integrating sensor observations, remote sensing, and model outputs into a unified representation of system state [3,7,16]. Second, they can improve forecasting and scenario analysis by combining physically based models with machine learning, surrogate models, or optimization algorithms, thereby enabling faster exploration of operational and strategic alternatives [4,5,14]. Third, DTs offer a framework for closing the loop between diagnosis and action: rather than merely identifying problems, they can also support proactive interventions such as dynamic control, anomaly detection, adaptive pumping, or risk-based allocation [3,9,10].

These advantages have made DTs increasingly visible in both research and practice over the last few years. The literature now includes sector-wide reviews, river-basin perspectives, urban digital twin studies, and hydrological position papers that collectively portray DTs as a promising foundation for more resilient, efficient, and equitable water management [3,5,7]. Importantly, this growth is not only quantitative but also qualitative: DTs are evolving from isolated prototypes toward more integrated architectures that

combine data infrastructures, interoperable services, advanced analytics, and stakeholder-facing interfaces.

Against this wider background, groundwater has emerged as a particularly important frontier for DT development [17,18]. Groundwater systems underpin irrigation, public water supply, industrial production, drought buffering, and ecological baseflow in many regions of the world, yet they remain among the most difficult water resources to monitor and manage effectively. Their hidden nature, strong spatial heterogeneity, delayed responses, and frequent data scarcity create major challenges for sustainable management, especially under climate change and rising abstraction pressures [19].

These same characteristics make groundwater an especially suitable domain for DT innovation. A groundwater DT has the potential to integrate piezometric observations, pumping records, hydroclimatic drivers, geospatial datasets, and hydrogeological simulation models into a dynamic decision-support environment capable of tracking aquifer states, anticipating threshold exceedance, and evaluating management actions before they are implemented [17,18]. In urban contexts, prototype DTs are already being explored for groundwater-table monitoring and proactive pump control to mitigate flooding or desiccation risks. In other settings, groundwater DTs are being proposed or developed for aquifer resilience, sustainable abstraction, salinity intrusion management, drought preparedness, and integrated basin planning [10,15,17,20,21].

What makes this topic especially compelling is that its specificity is not a weakness but a strength. The literature on groundwater DTs is not yet massive, but it has grown perceptibly in recent years and is clearly transitioning from conceptual exploration to early implementation. This creates an important scholarly moment: the field is mature enough to support synthesis, but still young enough that a well-structured review can meaningfully shape its conceptual boundaries, terminology, and future directions. In other words, this is precisely the stage at which a focused review can have high value for both researchers and practitioners.

1.3. The Need for a Dedicated Review on Digital Twins for Sustainable Groundwater Management

Despite the broader expansion of water-sector DT research, the groundwater-specific literature remains scattered across hydrogeology, hydrology, environmental engineering, smart agriculture, urban drainage, data science, and water governance. Existing reviews usually address DTs in the water sector as a whole, or they focus on water sustainability, river basins, or utility-scale applications without giving groundwater the analytical depth it requires [3,5,7]. As a result, there is still limited synthesis of how DT concepts are being translated into groundwater applications, which methodological strategies are being adopted, what types of management problems are targeted, and what barriers continue to hinder operational uptake [17,19].

This gap is particularly significant from a sustainability perspective. Groundwater management is not merely a matter of simulating heads or storage variations; it also involves questions of safe yield, recharge enhancement, drought resilience, salinity control, land-use interactions, agricultural demand, infrastructure operation, and institutional governance [19]. Consequently, a meaningful review must extend beyond technical architecture and consider how DTs intersect with resilience, socio-hydrology, and long-term sustainability goals. A focused review is therefore needed to assess whether current groundwater DT studies are truly moving toward sustainable resource management or whether they remain concentrated on narrow proof-of-concept demonstrations.

1.4. Novelties and Contributions of This Review

The novelty of the present review lies in its explicit and exclusive focus on digital twins for sustainable groundwater resources management. While broader water-sector reviews provide essential context [3], they do not fully capture the scientific and managerial specificities of groundwater systems—such as delayed aquifer responses, sparse monitoring networks, strong heterogeneity, recharge uncertainty, groundwater–surface water interactions, and the difficulty of translating subsurface information into actionable operational decisions—characteristics that strongly affect groundwater resource sustainability [19]. By concentrating specifically on groundwater, this review can examine in greater depth how DTs are being used—or proposed—to address abstraction control, aquifer resilience, saltwater intrusion, irrigation support, groundwater-table monitoring, and basin-scale sustainability problems.

A second contribution of this review is its integrative perspective. Rather than treating DTs as purely technical artifacts, the paper considers them as socio-technical systems whose effectiveness depends on the interplay among sensing infrastructures, process-based and data-driven models, optimization frameworks, governance arrangements, and decision contexts. This perspective is especially important for groundwater, where sustainable management depends not only on predictive accuracy but also on issues such as monitoring coverage, institutional coordination, transparency, and user trust.

By organizing the literature across these interacting dimensions, the review aims to identify both the state of the art and the key research gaps that must be addressed if groundwater DTs are to become credible tools for operational and strategic management.

1.5. Scope and Organization of the Literature Review

Considering the above, this review aims to provide a comprehensive and critical synthesis of recent research on digital twins for sustainable groundwater resources management, with emphasis on their hydrological foundations, enabling technologies, application domains, and governance implications. The review first situates groundwater DTs within the broader evolution of DTs in water resources management, then examines the literature through thematic lenses including modeling paradigms, architectural frameworks, sensing and IoT infrastructures, agricultural and water–energy–food applications, basin and urban water integrations, and governance and socio-hydrological dimensions. Finally, the review identifies major research gaps and outlines future directions for building groundwater DTs that are not only technically sophisticated but also scalable, interpretable, and aligned with long-term sustainability objectives.

Such framing is particularly useful because it connects recent digital innovation to core hydrological concerns: observation, process understanding, prediction, uncertainty, and management under change. The objective of this review is therefore both to synthesize recent research on digital twins for sustainable groundwater resources management, with emphasis on their hydrological foundations, enabling technologies, application domains, and governance implications; and also to help consolidate a coherent research field around a topic that is likely to become increasingly important for groundwater science and sustainable water governance in the coming years.

2. Materials and Methods

2.1. Literature Search and Selection Protocol

The search and selection of studies were conducted following the PRISMA 2020 framework [22], which provides a transparent structure for identifying, screening, assessing eligibility, and including records in systematic reviews. Details on the PRISMA 2020 Checklist can be found in the Supplementary Materials of this review.

The literature search was conducted using the Scopus and Web of Science databases. These databases were selected because together they provide broad multidisciplinary coverage across hydrology, hydrogeology, environmental engineering, computer science, artificial intelligence, agricultural water management, and sustainability research. This coverage was considered appropriate for the interdisciplinary scope of the present review, which addresses digital twins as socio-technical systems rather than solely as hydrogeological models.

Scopus was used as the primary database for the initial retrieval because it provides structured searching across title, abstract, and keyword fields and enables filtering by document type. Web of Science was searched using the same conceptual scope and related search terms. The Web of Science search did not substantially expand the retrieved literature because the records identified there largely overlapped with those obtained from Scopus. Duplicate records from the two databases were therefore removed before screening and synthesis. All database searches, manual searching and reference-chain screening were conducted up to May 2026.

The initial database search focused on the terms digital twin, "groundwater", "aquifer", "sustainability", and "management". The Scopus database was queried using the search string TITLE-ABS-KEY (digital twin AND (groundwater OR aquifer) AND (sustainable OR sustainability OR management)) AND (LIMIT-TO (DOCTYPE, "ar") OR LIMIT-TO (DOCTYPE, "cp") OR LIMIT-TO (DOCTYPE, "re")). This query was applied to the title, abstract and keywords fields ("TITLE-ABS-KEY"), ensuring that the retrieved records explicitly referred to digital twins and groundwater/aquifer systems in a management or sustainability context, rather than mentioning these concepts only tangentially.

The document type filter was restricted to journal articles (ar), conference papers (cp) and review articles (re) to focus on peer-reviewed scientific outputs and substantial conference contributions.

In addition to the database searches, manual searching and reference-chain screening were performed to reduce the risk of overlooking studies that address digital twin functions without explicitly using the term "digital twin". In this supplementary phase, relevant literature was sought using overlapping concepts, including smart and intelligent groundwater management, cyber-physical groundwater or aquifer systems, IoT-based and real-time groundwater monitoring, autonomous subsurface sensing, integrated groundwater observation systems, data-assimilative and real-time groundwater modeling, model–data fusion, AI-enabled groundwater prediction and decision support, hybrid and physics-informed groundwater modeling, adaptive and optimization-based groundwater management, integrated groundwater information systems, hydroinformatics platforms, and digital water management platforms.

A study was considered relevant when its framework performed, or was explicitly intended to perform, a combination of system monitoring, digital representation or modeling, data assimilation or updating, prediction or scenario analysis, and management or decision support. Studies that used related terminology without the label "digital twin" were therefore considered as DT-related contributions when they satisfied these functional criteria. They were distinguished from explicit DT studies during the synthesis to avoid equating every smart-monitoring or AI application with a complete digital twin.

A focused selection was subsequently applied to ensure consistency with the specific aim of the review. Studies were included when they addressed groundwater or aquifer systems and presented, evaluated, or substantially discussed an explicit digital twin, a closely related digital twin implementation, or a functionally equivalent digital, cyber-physical, intelligent, monitoring–modeling, or decision-support framework relevant to groundwater resources management. Studies were excluded when they focused exclusively

on surface water, wastewater, or unrelated digital twin applications; did not address groundwater management or aquifer processes; were not written in English; or did not have an accessible full text.

Screening of titles and abstracts, assessment of full-text eligibility, and data extraction were performed by the author. No automation tools were used in the selection process.

A total of 40 records were identified, including 22 articles, 12 conference papers, and 6 reviews (Figure 1). After abstract screening, 7 records were excluded because they were not written in English, 9 were excluded due to lack of relevance to the scope of the review, and 6 were removed because the full text was not found or accessible. After applying these criteria, 18 sources were retained for detailed analysis, consisting of 15 articles, 2 conference papers, and 3 reviews.

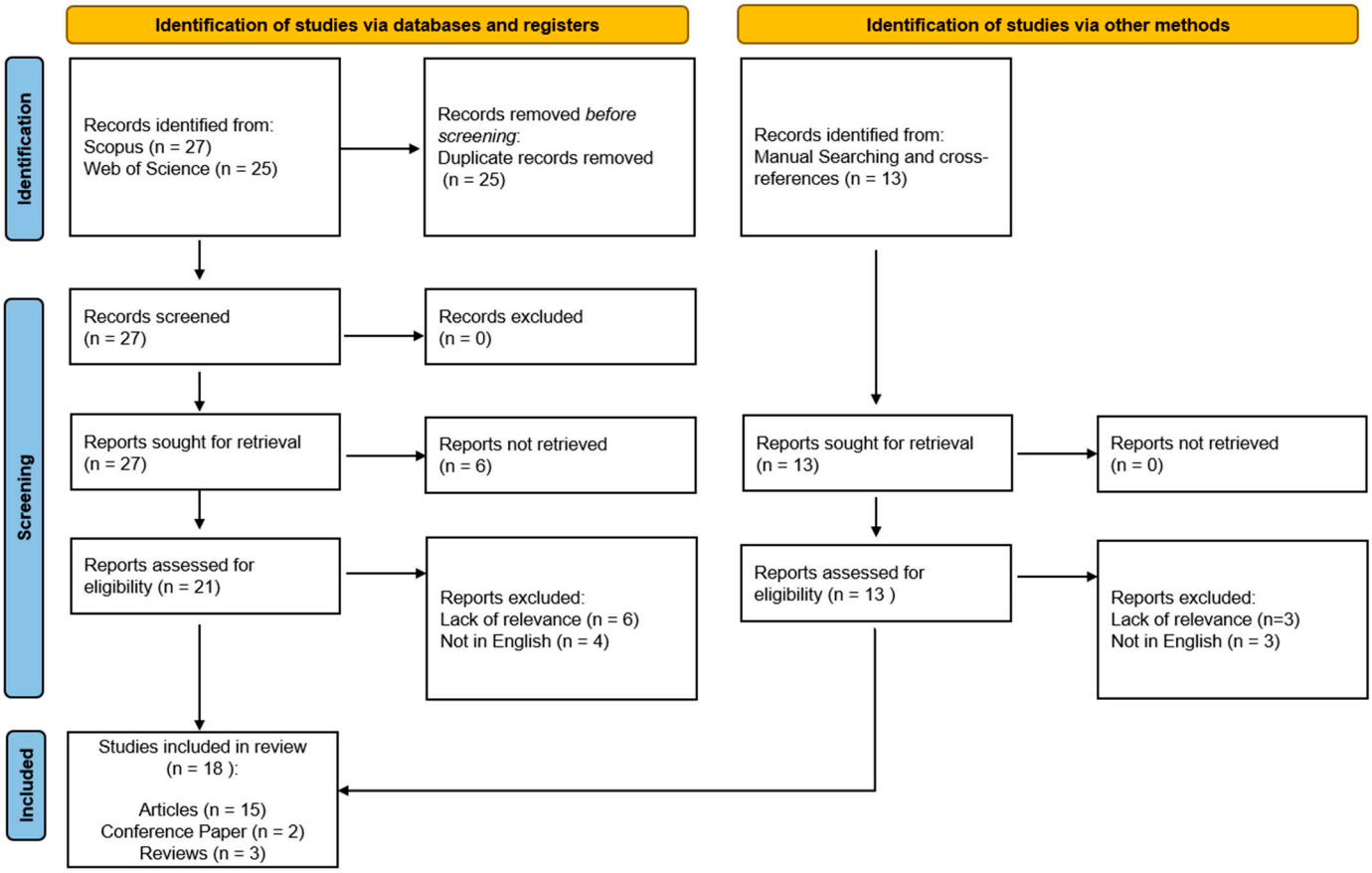


Figure 1. PRISMA [22] flowchart describing different stages of the literature review process.

Despite searching Scopus and Web of Science and conducting supplementary manual screening, some relevant studies may have been overlooked because terminology in this emerging field remains inconsistent. The final corpus should therefore be understood as a focused, function-based synthesis of groundwater digital twin research rather than an exhaustive review of all digitally enabled groundwater management studies.

2.2. Database Description and Sub-Topics Identification

The final database comprised 18 records, including 15 research articles, 2 conference papers, and 3 review papers, retained after the screening phase. Although numerically limited, this corpus is highly concentrated in most recent years, as shown in Figure 2, which highlights a clear increase in publications over the last four years. This temporal clustering suggests that digital twin applications to groundwater and aquifer systems

are an emerging and rapidly expanding research area, with growing scientific attention reflecting its increasing relevance for sustainable water resources management.

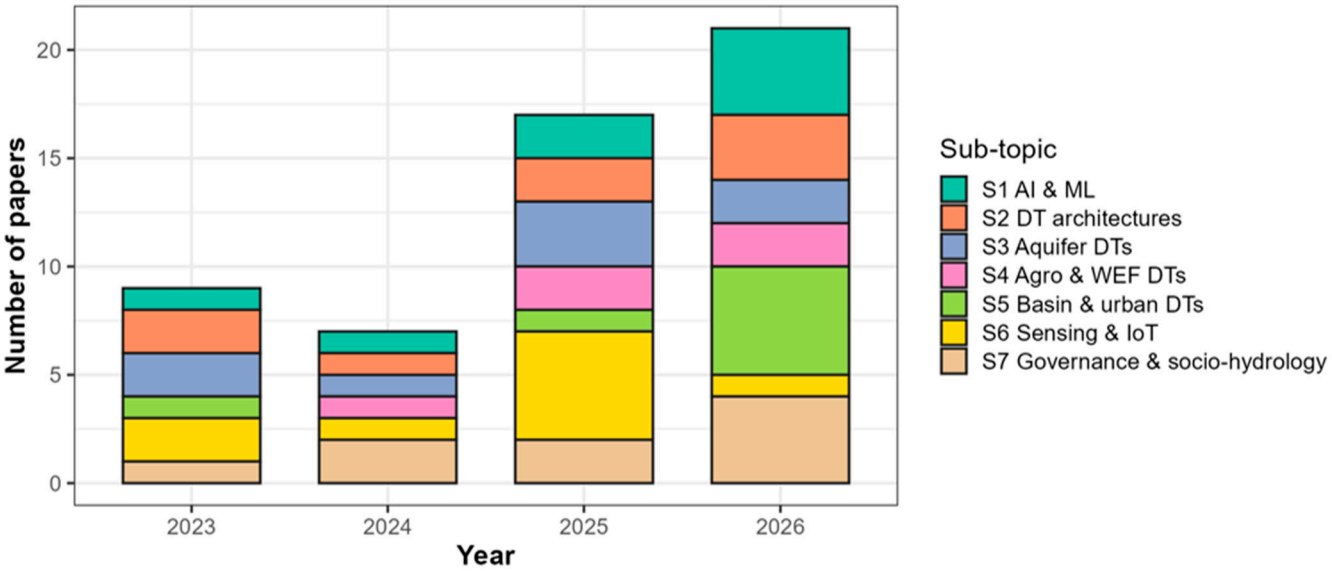


Figure 2. Temporal distribution of the records included in the review database, together with the thematic sub-topics identified in the literature, illustrating both the recent growth of publications and the evolving structure of the field.

To synthesize the heterogeneous set of 18 studies, a structured classification was applied rather than a purely narrative grouping. Each paper was characterised along multiple dimensions and data items: (i) methodological focus (e.g., AI & ML, process-based modeling, optimization, system dynamics), (ii) system scale (aquifer, irrigation district, basin, national, continental, urban water cycle), (iii) digital twin architecture and enabling technologies (process-based models, AI & ML components, IoT/sensing infrastructure), and (iv) management lens (e.g., drought mitigation, saltwater intrusion control, water–energy–food nexus, climate adaptation, governance). Papers could be assigned to more than one theme when they genuinely spanned several aspects (for example, a basin-scale DT that both advances AI & ML methodology and implements an operational decision-support platform). Additional variables included study location, aquifer type (confined/unconfined, coastal/inland), and the presence of governance or socio-hydrological components. Where information was unclear or not explicitly reported, no assumptions were made. Formal risk-of-bias assessment was not performed, consistent with the scoping nature of this review.

Through iterative comparison of these attributes across the 18 records, seven recurring and analytically useful sub-topics were identified:

- *S1—AI & ML foundations*: papers where the primary contribution lies in AI/ML paradigms for groundwater and water-system modeling (including physics-informed, explainable and surrogate models) and in how these support digital twin design and operation.
- *S2—Digital twin architectures and frameworks:* works that focus on conceptual or system architectures for DTs in natural resource and water systems, including layered designs, data/model interfaces, and integration of AI services and external tools.
- *S3—Aquifer-scale DT applications:* studies that implement or closely approximate digital twins for specific groundwater bodies (confined/unconfined aquifers, groundwater reservoirs, coastal aquifers), often with explicit management or resilience objectives.
- *S4—Agricultural and WEF-nexus DTs:* applications in irrigated agriculture and saline agroecosystems where DTs are used to support water–energy–food (or water–energy–

grain) nexus optimization, intelligent irrigation–drainage control, or groundwater-dependent crop management.

- *S5—Basin and urban water DTs:* digital twins that integrate groundwater within larger hydrological or infrastructural systems (river basins, national hydrological platforms, urban water cycles), emphasizing multi-scale interactions between aquifers, rivers and networks.
- *S6—Sensing, monitoring and IoT infrastructures:* contributions where the main innovation is in sensing and data infrastructures enabling DTs (smart sensors, IoT networks, autonomous or in situ monitoring, data fusion pipelines for the subsurface).
- *S7—Governance, resilience and socio-hydrology:* papers foregrounding socio-technical and governance aspects of groundwater DTs, such as resilience-based management, stakeholder-oriented decision support, smart-village frameworks and integrated urban water governance.

This seven-sub-topic framework was chosen for two reasons. First, it reflects the technical stack of groundwater digital twins, moving from modeling paradigms (S1), through architectures (S2), to concrete applications at different scales (S3–S5) and the enabling data infrastructure (S6). Second, it aligns with the socio-hydrological perspective required for sustainable groundwater management, which emphasizes the co-evolution and feedbacks between human and hydrological systems as a basis for long-term water sustainability [23–27], explicitly separating governance and resilience issues (S7) rather than subsuming them under purely technical categories. Studies were assigned to one or more of seven sub-topics (S1–S7) based on their methodological focus, system scale, digital twin architecture, and management lens. Each sub-topic was then synthesized narratively, highlighting main contributions, patterns, and gaps. No data conversions or imputation of missing summary statistics were required. Studies were synthesized narratively and thematically according to the seven sub-topics. The final mapping table (Section 4.1) assigns each of the 18 papers to one or more of these sub-topics, and the Sections 3 and 4 are structured accordingly, allowing the reader to trace how methods, architectures, applications and governance issues co-evolve within the emerging field of digital twins for sustainable groundwater resources management.

The final classification of the 18 papers into these seven sub-topics is reported in Section 4.1 and forms the structural backbone of the review. In the next section, each sub-topic (S1–S7) is examined in turn, highlighting the main contributions of the selected literature and the emerging patterns and gaps in digital twin research for groundwater sustainability.

3. Results: Thematic Analysis by Sub-Topic

The results of the literature review are reported according to the seven sub-topics defined in Section 2.2. For each sub-topic, its scope is restated, the main contributions of the associated papers are synthesized, and the principal trends and research gaps are identified. This organization follows a logical progression from methodological foundations (S1, S2), through application domains at multiple spatial scales (S3–S5), to enabling infrastructures (S6) and governance and socio-hydrological dimensions (S7).

3.1. S1: Artificial Intelligence (AI) and Machine Learning (ML) Foundations

3.1.1. Artificial Intelligence and Machine Learning Paradigms for Groundwater and Digital Twins

Over the last decade, artificial intelligence (AI) has evolved from isolated proof-of-concept models into a central component of groundwater science, supporting prediction, quality assessment, optimization and real-time decision support [28]. Recent reviews document how machine learning (ML) and deep learning frequently outperform traditional

statistical or purely process-based models for groundwater-level forecasting, contaminant detection and pumping optimization, while also stressing challenges related to data scarcity, interpretability and institutional uptake [28,29]. ML has been applied to groundwater-level, storage and recharge prediction; water-quality and contaminant classification; resource optimization (abstraction, recharge, cropping); and the fusion of remote sensing and IoT data for smart monitoring, with digital twins (DTs) increasingly framed as the natural "host environment" for such models [28,29].

Within this context, recent work positions groundwater ML within a decision framework distinguishing mechanistic, data-driven and hybrid models and articulating their trade-offs [29,30]. Ma et al. propose selecting model classes based on data availability and quality, strength of physical theory and deployment requirements (real-time operation, interpretability, regulatory acceptance) [29]. Secci et al. compare process-based, data-driven and system-dynamics approaches in the Konya Closed Basin, showing how each addresses different management questions and operates at distinct spatial and temporal scales [30]. Process-based models (PBMs), typically MODFLOW- or FEFLOW-type partial differential equation models [31,32], remain the backbone for high-fidelity representation of hydrogeological structure and flow/transport, and serve as "truth models" for scenario analysis and surrogate training [17,30,33]. In the CONUS DT, ParFlow-CLM simulations at 1 km resolution provide the physical backbone used to train an adjusted Random Forest model that downscales and bias-corrects daily water-table depth to 1-arcsecond (~30 m) resolution [17]. However, PBMs are computationally demanding and difficult to calibrate for large, heterogeneous systems with limited observations, constraining their direct use in real-time DTs [30,34].

Data-driven models (DDMs) range from classical regression and tree-based methods to deep neural networks (DNNs), sequence models such as LSTMs and generative adversarial networks (GANs) [29,33]. They capture non-linear relationships between climate, pumping, land-use and sensor data and outputs such as head, quality indicators or yields without explicit governing equations and run orders of magnitude faster than PBMs once trained [17,33]. In the CONUS DT, a physics-guided Random Forest surrogate generates full-domain daily hyper-resolution water-table maps in about 35 min of GPU time, correcting biases in ParFlow-CLM and achieving high Pearson and Spearman correlations ($\rho \geq 0.5$ for a large fraction of previously unseen grid cells) [17]. In groundwater applications, DDMs have been used for level forecasting, surrogate modeling of complex three-dimensional flow systems and estimation of effective parameters or recharge at regional scale [28,30,33]. Their limitations include sensitivity to training data representativeness, extrapolation risk and "black-box" perceptions that hinder regulatory or stakeholder acceptance [28,29].

Hybrid and system-dynamics approaches aim to exploit complementary strengths of PBMs and DDMs [29,30]. Hybrid models couple mechanistic cores with ML components that learn closures, sub-grid processes or bias corrections or impose physical constraints directly in ML architectures through physics-informed loss functions and design [28,29]. System-dynamics models, as applied in Konya, link simplified representations of groundwater stocks with socio-economic feedbacks (crop choice, investment in wells, policy interventions), creating a bridge between aquifer behavior and stakeholder decision-making [30]. Secci et al. show that DDM surrogates and PBMs can supply boundary conditions and performance metrics to system-dynamics models, which then support exploration of long-term policy pathways, making the triad of PBM–DDM–system dynamics complementary rather than competing [30].

Rahman et al. describe this evolution as a shift toward "intelligent aquifer systems" in which dense IoT and satellite observation networks feed AI models that update forecasts, detect anomalies and propose control actions within a governance framework [28]. They

identify four paradigms particularly relevant for DTs: physics-informed ML to embed conservation laws and hydrologic constraints; explainable AI (XAI) to clarify controlling drivers and build trust; surrogate modeling for real-time decision support; and reinforcement learning and evolutionary optimization for adaptive pumping and recharge strategies [28]. Ma et al. similarly argue for a transition from "black-box prediction" to "physics-informed water intelligence", with DTs as the operational environment in which prediction is systematically coupled to action [29].

Surrogate modeling emerges as a central theme, especially where DTs must support intensive scenario exploration and rapid optimization [30,33,34]. Payne et al. train DNN, elastic-net and GAN surrogates on a FEFLOW model of a small carbonate island (Barbados), achieving $R^2 \approx 0.95$–0.96 and RMSE $\approx$ 2.6–2.8 m for steady-state heads, with evaluation times reduced from ~45 min to seconds [33]. Ma et al. demonstrate that physics-guided Random Forest surrogates can support a continental-scale DT that downscales ParFlow-CLM outputs, assimilates observations and delivers daily hyper-resolution water-table depth estimates with substantially improved skill over the parent model [17]. In Konya, an artificial neural network trained on PBM outputs predicts groundwater levels at 30 control points under varying climate and cropping scenarios and is deployed via a user-oriented interface to generate basin-wide piezometric maps, effectively acting as a DT computational core [30].

Model selection and complementarity are therefore central to DT design [29,30]. A recurring pattern is the use of PBMs as "digital laboratories" to generate synthetic data, conduct stress tests and delineate safe operating envelopes; ML surrogates to approximate PBMs for real-time prediction and sensitivity analysis; and system-dynamics or rule-based components to represent human decisions and policy levers [28–30,33]. The CONUS DT exemplifies this layered architecture at hyper-resolution, with ParFlow-CLM providing the physical envelope, a physics-guided Random Forest surrogate downscaling and bias-correcting water-table fields, and the resulting DT enabling GPU-accelerated scenario exploration and monitoring across the contiguous United States [17]. Such architectures match DT requirements to maintain physical grounding, respond rapidly to new data and expose interpretable indicators and controls to stakeholders.

Finally, the literature emphasizes that trustworthiness—not only accuracy—is essential for AI-enabled groundwater DTs [28,29]. Ma et al. call for transparent model design, cross-basin validation and XAI techniques (e.g., feature importance, SHAP) to diagnose failure modes and avoid spurious correlations, particularly in high-stakes decisions [29]. Rahman et al. highlight ethical data governance, open datasets and capacity-building in water agencies to ensure that AI supports equitable groundwater governance rather than reinforcing asymmetries in knowledge and power [28]. AI and ML foundations for DTs are thus as much about socio-institutional embedding as about algorithms.

3.1.2. Digital Twin Case Studies with Strong ML Components

The transition from abstract AI paradigms to operational digital twins is exemplified by case studies that couple machine learning engines with sensing, visualization and optimization layers around groundwater systems [17,33–36]. These implementations show how surrogates, sequence models and evolutionary algorithms can be orchestrated within DT architectures for drought-risk mitigation, irrigation management and water–energy–food (WEF) nexus optimization, while also revealing current limitations.

Chen et al. develop a simulation-driven DT for drought-risk mitigation in the Tailan River groundwater reservoir, using a five-dimensional architecture aligned with a four-stage proactive risk-management cycle [34]. A dual-scale Random Forest prediction engine, trained on a high-fidelity groundwater model, reproduces annual storage and monthly

multi-point heads with $R^2 \approx 0.95$ and mean absolute errors below 0.05 m, providing near real-time updates without repeatedly solving the full model [34]. A Hierarchical Collaborative Optimization (HCO) module combining linear regression, Differential Evolution and Gradient Descent generates multi-objective mitigation strategies that respect both long-term storage constraints and short-term operational limits [34].

At irrigation-district scale, Huang et al. propose a DT-enabled framework for collaborative optimization of the water–energy–grain nexus in Hebei, where 72% of groundwater zones are over-exploited and pumping accounts for ~60% of energy use [35]. An enhanced NSGA-III algorithm is coupled with a four-dimensional nexus model and a real-time DT verification loop, producing strategies that achieve ~18.5% water savings, reduce energy consumption by ~1285 kWh ha^{-1} and yield ~0.07 m groundwater recovery during droughts [35]. A dense sensor network (32 meteorological stations, 48 soil sensors) and a cyber-physical platform verify groundwater-level predictions with ~0.03 m accuracy and ~186 ms decision latency, enabling near real-time adjustment of irrigation schedules; scalability is demonstrated by extension to the Hexi Corridor [35].

Qin et al. present a DT-enabled intelligent irrigation–drainage system for precision water–salt management in saline agroecosystems, integrating IoT sensing, deep learning and closed-loop control [36]. A closed-loop infrastructure with subsurface drains, multi-source reuse pipelines and photovoltaic-powered brackish-water treatment is instrumented with high-frequency sensors of soil moisture, temperature, salinity and groundwater depth, while an LSTM model predicts soil electrical conductivity ($R^2 \approx 0.97$) and water content ($R^2 \approx 0.92$) under varying schedules [36]. Feature-importance analysis identifies soil moisture, temperature and groundwater depth as dominant drivers, and the DT uses LSTM forecasts to coordinate irrigation and drainage, reducing soil salinity in permeable profiles without compromising water-use efficiency [36].

Payne et al. consider ML surrogates as precursors to a DT for a small carbonate island, training DNN, elastic-net and GAN emulators on a FEFLOW model of Barbados' karst aquifer [33]. The surrogates achieve $R^2 \approx 0.95$–0.96 while reducing runtimes from tens of minutes to seconds, enabling rapid exploration of pumping and recharge scenarios relevant to climate-induced recharge changes and sea-level-driven salinization [33]. Future extensions envisaged by the authors include physics-informed neural networks to enforce groundwater flow equations in the loss function and improve robustness [33].

At continental scale, Ma et al. develop a physics-guided groundwater DT for the contiguous United States, training an adjusted Random Forest model on 1 km ParFlow-CLM simulations and water-table observations to downscale and bias-correct daily water-table depth to 1-arcsecond (~30 m) resolution [17]. The model produces hyper-resolution maps with markedly improved temporal skill (median Spearman $\rho \approx 0.66$ in the training year and increased by ~0.13 in a future period), and multi-GPU computing enables generation of each daily map in ~35 min while quantifying predictive uncertainty from tree-level outputs [17].

Across these case studies, several common design patterns emerge. All rely on close coupling between physically based simulations and ML surrogates—RF in Tailan, DNN/elastic-net/GANs in Barbados, LSTMs in saline agro-systems, PBM-trained ANNs in the Konya basin and a physics-guided RF in the CONUS twin—ensuring that computational speed does not compromise physical plausibility [17,30,33,34,36]. Optimization or control layers (HCO, NSGA-III, rule-based irrigation–drainage scheduling) transform predictions into actionable variables such as pumping rates, irrigation allocations and reservoir releases, while dense sensing and IoT infrastructures continuously update model states and validate forecasts [34–36].

At the same time, important limitations remain. The Tailan DT is heavily simulation-driven with limited field validation, underscoring the need for long-term observations to constrain biases [34]. The Hebei framework depends on high infrastructure density, raising questions about transferability to data-scarce or institutionally fragmented regions and the need for simpler surrogates and stronger uncertainty quantification [28,35]. The saline agroecosystems DT is sensitive to non-stationary climate and management, motivating physics-informed and explainable AI approaches as discussed by Rahman et al. and Ma et al. [28,29]. Finally, the reviewed systems remain largely technology-centric; integration with system-dynamics representations of behavior, policy feedbacks and economic incentives, as explored for the Konya Closed Basin, is needed to capture the full socio-hydrological complexity of sustainable groundwater management [28,30,33].

3.2. S2: Digital Twin Architectures and Frameworks

DT architectures for groundwater and coupled water systems are converging toward a set of common principles: layered designs that separate data, models and services; open, interoperable interfaces; and explicit slots where AI and optimization components can be plugged in as the system matures [20,29,34,37]. These principles are being instantiated at multiple scales—from regional natural-resource twins and smart-village governance frameworks to urban and national water-cycle platforms—each emphasizing different aspects of modularity, stakeholder interaction and hydrologic detail [38–42].

3.2.1. Layered Architectures and Data–Model Interfaces

Iglesias et al. propose an explicit, generalizable DT architecture for natural resource management, using the ENVRI Reference Model to structure an AI-assisted DT for the Centre-Val de Loire region, with the Beauce aquifer as a key groundwater example [37]. The DT is decomposed into five interacting layers—Science, Information, Computational, Engineering and Technology—with domain concepts and questions in the Science layer, standardized data models and provenance in the Information layer, analytical and ML services in the Computational layer, workflow orchestration in the Engineering layer, and hardware/containers/middleware in the Technology layer [37]. A FIWARE-based proof-of-concept ingests daily data from 53 piezometric stations and meteorological networks (≈2.8 million records), exposing piezometric series, rainfall and derived indicators via NGSI-compliant interfaces to dashboards, AI components and external clients, so that new ML services can be added without re-engineering ingestion or storage pipelines [37].

Duggal et al. similarly outline a three-layer DT for water-table monitoring in Punjab, separating a Data layer (sensor discovery, acquisition, cleaning and fusion of piezometer, EC, pH, rainfall, CSV, GIS and satellite data), a Modeling layer (numerical and data-driven models mapping recharge, pumping and soil properties to water-table responses) and an Application layer (dashboards and decision-support tools) [41]. Their "Data Collection and Fusion" workflow explicitly defines the data path from sensor identification to unified datasets, emphasizing extensibility and well-defined interfaces between layers [41].

At national scale, Henriksen et al. conceptualize Denmark's Hydrological Information and Prediction (HIP) system as a de facto DT for climate adaptation, water management and disaster risk reduction [20]. HIP couples the fully distributed DK-model with nation-wide groundwater, soil-moisture and streamflow networks and meteorological forcing, all served through a web portal; while initially primarily "one-way", the envisioned evolution includes real-time updating and plug-in local twins (e.g., basin or urban DTs) connected to the national backbone [20]. Architecturally, HIP mirrors layered designs but emphasizes federation and nesting: the national DK-model provides boundary conditions and large-scale context, while local DTs contribute higher-resolution sub-models, sensing networks

and decision rules, with interoperability ensured by shared data standards, calibration protocols and common definitions of exposure, vulnerability, resilience and water-security indicators [20].

Ma et al. complement these structural views by framing DT architectures from the perspective of ML deployment and trustworthy AI [29]. They argue that water-system DTs should be built around three classes of ML components (data-driven, physics-informed, hybrid) and must provide hooks for embedding physical constraints (e.g., physics-informed loss functions, graph representations), interpreting decisions via XAI modules and safely integrating reinforcement learning for closed-loop control. In this view, DTs become "AI-ready" platforms in which sensor and remote sensing streams feed standardized feature schemas, models publish predictions with uncertainty and explanations, and control policies are executed through actuation interfaces constrained by mechanistic models [29].

Collectively, these works outline a reference pattern for hydrological DT architectures: layered, standard-compliant data infrastructure; modular, containerized mechanistic and ML-based models interacting via APIs; and service layers that deliver analytics, visualization and control to human and automated decision-makers [20,29,37,41].

3.2.2. Smart-Village Governance and Watershed-Scale Digital Twins

At local scales, digital twins are increasingly conceptualized as governance frameworks for village-level groundwater sustainability and watershed management, with architectures that are simultaneously technical and institutional [38,42]. Singh and Sharma propose a DT-based smart-village framework for semi-arid rural India (e.g., Satara district, Maharashtra), integrating AI, GIS and DT concepts to connect multi-source hydrogeological data (well logs, pumping, rainfall, land use) with participatory decision-making structures for communities facing over-extraction and climate variability [38]. The framework comprises a data-acquisition block (IoT sensors, manual measurements, remote sensing), a GIS-based hydrogeological analysis block (3-D aquifer and recharge mapping), a predictive modeling block (AI & ML forecasts of water-table trajectories and drought risk under alternative pumping/recharge strategies), and a community-governance block that exposes results to stakeholders (Gram Panchayats, farmer groups) and supports rule-making on well drilling, abstraction caps and recharge measures [38].

In a related contribution, Singh and Sharma detail a Digital Twin of Randullabad Watershed and Groundwater (DToGW), targeting SDG-6 in a specific village [42]. The architecture includes a physical layer (watershed, supply network, groundwater system), a data layer (IoT sensors, satellite products, GIS layers), a model layer coupling open-source hydrological models (e.g., FREEWAT [43], MIKE SHE [44]) with AI predictors, and an application layer providing dashboards, alerts and scenario-exploration tools for village meetings [42]. Key features are bi-directional data flow between physical and digital spaces, modular integration of existing authority models with new ML components, and explicit human-in-the-loop interfaces where local knowledge and preferences shape model configurations and scenario selection; DEM-derived drainage maps, groundwater-potential maps and socio-economic layers are ingested to align interventions with both hydrogeologic feasibility and livelihood constraints [38,42].

Duggal et al.'s Punjab architecture, although more technically oriented, also has governance implications: robust data-quality checks, scalable cloud storage and standardized formats are designed to provide transparent, shareable groundwater information for farmers, planners and regulators, turning the DT into a common information substrate for negotiating pumping limits and crop choices [38,41,42].

Across these smart-village and watershed-scale DTs, three architectural themes emerge: (i) tight coupling of GIS and DT layers, treating high-resolution spatial data (DEMs, land

use, aquifer boundaries) as core objects for spatially explicit reasoning on recharge and vulnerability [38,42]; (ii) accommodation of low-cost, heterogeneous sensing (IoT and manual observations) to match rural resource constraints and maintain functionality under limited connectivity [41,42]; and (iii) participatory interfaces, with application layers tailored to non-experts through intuitive dashboards and scenario tools, embedding DTs in socio-hydrological feedbacks rather than as purely technical artefacts [30,38,42].

3.2.3. Urban and National Water-Cycle Digital Twins

At urban and national scales, digital twins shift focus from individual aquifers to the full river–aquifer–pipe network continuum, integrating natural hydrological systems, built infrastructure and institutional governance [20,39,40]. Marcoie et al. frame an Integrated Urban Water Management paradigm in which an urban water DT acts as a diagnostic engine within a three-stage Diagnose–Prescribe–Optimize framework, using coupled river, groundwater, water-quality and network-hydraulic models driven by a shared data infrastructure [40]. Diagnosis quantifies how channelization, urbanization and pumping fragment the water cycle (e.g., in the Gherăești-Bacău wellfield), prescription prioritizes nature-based measures such as floodplain restoration and recharge zones, and optimization uses smart, real-time control to fine-tune demand and pumping [40].

Davis et al. complement this with a vision of "autonomous" subsurface sensing and prediction for contaminated sites, in which dense in situ sensor networks (gas, temperature, moisture, passive samplers) stream data to a subsurface DT that continuously estimates degradation rates, plume dynamics and risk [39]. Their layered concept—underground sensors, edge devices, data manipulation pipelines and adaptive computational platforms—emphasizes tight sensor–model coupling and bi-directional feedbacks where model outputs guide sensor deployment, principles directly transferable to quantity- and quality-oriented groundwater DTs [39].

Henriksen et al. extend DT concepts to national scale via Denmark's Hydrological Information and Prediction (HIP) system, which couples a fully distributed DK-model (groundwater and surface water) with nationwide sensor networks and meteorological forcing, disseminated through a web portal [20]. HIP is evolving from batch modeling to a real-time DT where model states are continuously updated and local plug-in twins (e.g., basin or urban DTs) "dock" into the national backbone, ingesting boundary conditions and forecasts and returning aggregated indicators such as flood probabilities or groundwater-deficit indices [20]. This federated architecture depends on shared data formats, joint calibration frameworks and common hydrological and socio-economic indicators, enabling multi-scale water-security assessments [20].

Ma et al. further argue that DTs for networked water systems should exploit graph neural networks (GNNs) to natively represent river networks, distribution pipes and monitoring networks and to learn flow and transport over these graphs, while federated learning allows utilities or regions to train shared models without centralizing sensitive data [29]. In such architectures, local twins maintain their own data and models but periodically synchronize parameters, rather than raw data, with central or peer coordinators, combining privacy with broader learning [29].

Overall, urban and national water-cycle DTs are characterized by: (i) multi-model integration across rivers, aquifers and infrastructure, treating groundwater models as integral components [20,40]; (ii) autonomous or semi-autonomous sensing networks tightly coupled to computational platforms for near real-time state updating and forecasting [20,39]; and (iii) federated, nested structures in which local DTs plug into regional or national backbones while retaining local specificity [20,29].

3.2.4. Cross-Cutting Architectural Insights for Groundwater DTs

Across regional natural-resource twins, smart-village governance frameworks and urban/national platforms, several cross-cutting architectural insights emerge that are directly relevant for digital twins targeting sustainable groundwater management.

First, layered architectures with clear separation of concerns—data ingestion, storage, modeling/analytics, and application/decision support—appear to be a prerequisite for modularity and long-term evolvability [20,37,41,42]. This separation allows new AI components (e.g., physics-informed surrogates, XAI modules) to be integrated without disrupting existing data pipelines, and conversely permits gradual upgrades of the sensing infrastructure without requiring wholesale re-engineering of models.

Second, interoperability through common data and model interfaces is essential for harnessing complementarity between process-based, data-driven and system-dynamics models in a DT context [20,29,40]. Whether at the scale of the Beauce aquifer, the Randullabad watershed or the entire Danish hydrological system, architectures that rely on open standards (e.g., NGSI, OGC, well-defined REST APIs) and shared ontologies (for aquifers, wells, recharge, demands) are more likely to support cross-scale analyses and coordinated management than bespoke, siloed platforms [20,37,42].

Third, AI integration is not an afterthought but a design driver: platforms that expose feature-level data streams, uncertainty channels, and actuation interfaces are inherently better suited to host ML components that must operate in real time, explain their decisions and respect physical and regulatory constraints [20,29,39]. This is evident in the way Ma et al. define AI-ready DTs, in the formation of dual-scale RF engines within the Tailan reservoir DT, and in the NSGA-III-driven Hebei irrigation DT, even though the latter two are more fully discussed under the S1 topic [29,34,35].

Finally, all of the representative architectures indicate that digital twins for groundwater cannot remain purely technical artifacts: they must be coupled to governance processes, whether via smart-village interfaces, urban IUWM frameworks or national risk-knowledge systems, if they are to effectively support sustainable and equitable management under climate and socio-economic change [20,38,40,42].

3.3. S3: Artificial Intelligence (AI) and Machine Learning (ML) Foundations

Aquifer-scale digital twins (DTs) sit at the interface between detailed hydrogeological understanding and operational decision support, wrapping process-based models, surrogate ML components and sensing infrastructures into cyber-physical systems focused on specific groundwater bodies [21,33,34,45]. In the current literature, the most advanced examples target (i) drought-risk mitigation in managed groundwater reservoirs, (ii) control of saltwater intrusion in coastal and island aquifers, (iii) resilience-based management of confined aquifers under stress, and (iv) climate-change stress testing combined with managed aquifer recharge (MAR), all of which are natural focal points for sustainable groundwater digital twins [21,30,34,45,46].

3.3.1. Simulation-Driven Digital Twins for Groundwater Reservoirs and Drought Mitigation

Chen et al. present an explicit groundwater-reservoir DT for drought-risk mitigation in the Tailan River reservoir (arid northwest China), where a karez-type system is managed to maintain subsurface storage above critical thresholds under agricultural and ecological demands [34]. A high-fidelity numerical model is embedded in a layered DT architecture and wrapped with a dual-scale Random Forest (RF) prediction engine: one surrogate predicts annual total storage, the other monthly multi-point heads, achieving $R^2 \approx 0.95$ and MAE < 0.05 m while enabling near real-time updates at a fraction of the MODFLOW-

type runtime [34]. A Hierarchical Collaborative Optimization (HCO) module—combining multiple linear regression, Differential Evolution and Gradient Descent—uses the RF surrogates to search for pumping and recharge strategies that minimize drought risk subject to supply-reliability and cost constraints, with risk indicators exposed through a Services layer [34].

Henriksen et al.'s HIP DT, although national in scope, similarly supports aquifer- and reservoir-scale drought assessment: the fully distributed DK-model, driven by real-time data, provides groundwater levels and storage across Denmark and allows plug-in local twins (river basins, wellfields) to use national boundary conditions for scenario-based drought analysis, abstraction limits and MAR design at recharge-area and capture-zone scales [20]. Secci et al. complement these reservoir-oriented systems with a multi-model framework for the Konya Closed Basin, where process-based, ANN surrogate and system-dynamics models jointly represent three-dimensional flow, basin-wide groundwater levels at 30 control points ($R^2 > 0.96$) and farmer behavior/cropping choices [30]. Although not formulated as a DT, this PBM–ANN–system-dynamics triad anticipates aquifer-scale twins that must support both hydrogeologic realism and socio-economic decision layers [30].

3.3.2. Digital Twins for Coastal and Island Aquifers and Saltwater Intrusion Control

Coastal and island aquifers pose DT design challenges due to density-driven flow and saltwater intrusion (SWI), which create non-linear feedbacks between pumping, recharge and quality [21,33]. Sharan et al. develop a DT framework for a coastal island aquifer in Port Vila (Vanuatu), coupling a three-dimensional SEAWAT model with machine learning surrogates and a simulation–optimization (S–O) module to manage SWI under current and future pumping [21]. SEAWAT simulations provide training data for surrogates that approximate the relationship between production/barrier-well pumping and salinity distributions; the S–O module then identifies pumping configurations that keep salinity at supply wells within targets (0.45–1.15 kg m^{-3}) while minimizing barrier-well pumping, thus reducing costs and preserving freshwater reserves [21,47].

Payne et al. contribute an island-scale surrogate framework for Barbados, where a karstic unconfined aquifer is vulnerable to over-abstraction and salinization [33]. A FEFLOW model of groundwater flow [32] is used to train DNN, elastic-net and GAN surrogates for steady-state heads, yielding an RMSE $\approx$ 2.6–2.8 m and an $R^2 \approx$ 0.95–0.96, with runtimes reduced from ~45 min to seconds [33]. Elastic-net models provide variable selection (highlighting recharge and specific wells), and GAN outputs match the physics-based model with ~95% accuracy [33]. These surrogates are explicitly positioned as computational cores of a future Barbados DT able to evaluate large ensembles of pumping and climate/sea-level scenarios and SWI risk [21,33]. Together, the Vanuatu and Barbados studies show how high-fidelity density-dependent models, ML surrogates and optimization can be combined in coastal and island aquifer DTs to jointly manage quantity and quality [21,33].

3.3.3. Resilience-Based Management of Confined Aquifers and MAR Under Climate Change

Confined aquifers under intense abstraction and climate stress are natural candidates for resilience-oriented DTs [30,45,46]. Cohen-Manrique et al. develop a DT of the Morroa confined aquifer (Colombia) for resilience-based management, combining historical and real-time hydrogeological data (water level, pH, dissolved oxygen, temperature via ESP32-based devices) with a numerical model and heuristic optimization [45]. A resilience-based control model defines a target dynamic water level (e.g., 58.5 m) and uses grid search, genetic algorithms and Particle Swarm Optimization (PSO) to determine extraction rates that minimize RMSE between simulated and target heads under sustainability constraints,

with PSO converging to ~8.3 L s^{-1} [45]. The DT demonstrates that current extraction is unsustainable even in wet years, motivating revised pumping and MAR/demand-side measures, and is deployed as a decision-support tool for environmental authorities [45].

Szanyi et al. address resilience via climate-change adaptation in the Nyírség shallow groundwater system (northeastern Hungary), combining 1970–2022 monitoring, CHIRPS/ERA5-Land climate products and Processing MODFLOW Pro v8, a MODFLOW-based groundwater-flow model [31,48] to separate climate-driven from pumping-driven impacts [46]. Mid-century scenarios show climate-induced recharge reductions dominating basin-scale declines, while increased pumping intensifies local drawdown [46]. Tested MAR options (infiltration basins, streambed recharge, injection wells) reveal that subsurface injection is most effective locally but MAR largely remains a local mitigation in the face of widespread recharge loss; this configuration is readily embeddable in an aquifer-scale DT for climate adaptation [20,46]. Secci et al. further demonstrate, for Konya, how climate scenarios (RCP4.5/8.5), irrigation changes and policy choices can be combined in PBM, ANN and system-dynamics models to stress-test resilience pathways, suggesting DT architectures where climate forcing, cropping decisions and policy instruments are treated as time-varying drivers of aquifer resilience [30,45,46].

3.3.4. Surrogates, Monitoring Frameworks and DT Readiness at the Aquifer Scale

Several works focus on sensing and surrogate frameworks that are DT-ready even if not fully implemented as twins. Duggal et al. propose a multilayer architecture for Punjab that is directly applicable to over-exploited unconfined aquifers: a Data layer standardizes/fuses time series from piezometers, EC/pH sensors, rainfall gauges and satellite products; a Modeling layer hosts groundwater models and ML predictors; and an Application layer supports monitoring, alerting and planning [41].

Henriksen et al.'s HIP system provides complementary monitoring–model integration, with dense online networks feeding a coupled surface–subsurface model that can be downscaled or nested for specific aquifers [20]. Their emphasis on open data, shared calibration guidelines and best practices (multi-variable optimization, uncertainty analysis) is directly relevant to aquifer-scale DTs seeking scientific credibility and operational utility [20,30]. Finally, surrogate-based approaches such as the Barbados DNN/elastic-net/GAN models and the Konya ANN predictors bridge the computational gap between high-resolution hydrogeology and DT needs for fast, repeated evaluations, enabling extensive scenario exploration without sacrificing accuracy [30,33].

3.3.5. Cross-Cutting Lessons for Aquifer-Scale Groundwater Digital Twins

Across these diverse applications, several cross-cutting lessons emerge for the design of aquifer-scale digital twins focused on sustainable and resilient groundwater management:

- DTs are inherently hybrid systems: the most advanced examples combine process-based numerical models (e.g., SEAWAT, MODFLOW, FEFLOW), ML surrogates (RF, DNN, LSTM, ANN, GAN) and heuristic or evolutionary optimization algorithms (PSO, DE, GA, NSGA-III) within a single architecture [21,30,33,34,45]. This hybridity allows DTs to preserve hydrophysical realism while achieving the responsiveness required for real-time decision support.
- Saltwater intrusion and drought risk are natural focal points: coastal and island aquifers (Vanuatu, Barbados) and inland groundwater reservoirs (Tailan, Konya, Nyírség) all face forms of "hydrological depletion"—either through SWI or chronic water-table declines—making them prime candidates for DT-enabled optimization and resilience planning [21,30,33,34,46].

- Resilience metrics and targets must be made explicit: the Morroa DT defines a target dynamic water level and optimizes extraction around it; the Tailan DT builds a drought-risk mitigation cycle; and the Nyírség and Konya studies quantify climate-induced groundwater deficits under different adaptation scenarios [30,34,45,46]. Aquifer-scale DTs should similarly encode resilience metrics (e.g., minimum permissible head, safe SWI front position, acceptable depletion rates) as core design elements.
- Monitoring and data fusion are as critical as modeling: whether in Punjab, Denmark or Colombia, successful DT-like systems depend on robust sensing, rigorous data cleaning and harmonization and transparent data access, without which surrogate training, data assimilation and real-time control become unreliable [20,41,45,46].
- Socio-hydrological integration is essential for long-term sustainability: system-dynamics models in Konya, governance-focused architectures in smart-village DTs, and national risk-knowledge frameworks in Denmark all illustrate that aquifer-scale DTs must eventually incorporate human behavior, policy instruments and economic incentives, not just hydrogeology and hydraulics [20,30,42,45].

Together, the representative works demonstrate that aquifer-scale DT applications are no longer hypothetical: they are emerging across climatic zones and hydrogeological settings, providing concrete templates for drought-risk mitigation, SWI control, resilience-based management and climate adaptation in groundwater systems [20,21,30,33,34,41,45,46].

3.4. S4: Agricultural and WEF-Nexus DTs

Digital twins in irrigated agriculture sit at the heart of the water–energy–food (WEF) nexus, because they explicitly couple groundwater dynamics, pumping energy use and crop production outcomes in closed-loop decision frameworks [34,35]. Work shows how DTs can (i) drive multi-objective optimization in large irrigation districts, (ii) support intelligent irrigation–drainage and water reuse in saline agroecosystems, and (iii) provide climate-aware decision support for aligning irrigation practices and managed aquifer recharge (MAR) with groundwater sustainability constraints [34–36,46].

3.4.1. Digital Twins for WEF/Water–Energy–Grain Optimization in Irrigation Districts

Huang et al. develop a WEF-oriented DT for Hebei irrigation areas, where 72% of groundwater zones are over-exploited and pumping accounts for ~60% of total energy consumption [35]. The architecture couples: (i) an enhanced NSGA-III multi-objective evolutionary algorithm with dynamic reference points and hybrid constraints; (ii) a four-dimensional nexus model (groundwater dynamics, pump energy lifecycle, economic returns, climate-resilient yields); and (iii) a real-time DT verification platform that streams sensor data back into the virtual model [35]. Calibrated with data from 32 meteorological stations and 48 soil sensors, the nexus model attains a determination coefficient of 0.913 for economic outputs and a crop-yield error of ~0.47%, supporting seasonal decisions at district scale [35]. NSGA-III identifies strategies delivering ~18.5% water savings and 1285 kWh ha^{-1} energy reduction, while achieving ~0.07 m groundwater recovery during droughts [35].

A PID-based verification loop compares predicted and observed groundwater levels and crop responses, achieving ~0.03 m head-prediction accuracy and ~186 ms decision latency; deviations trigger re-optimization, implementing an "optimization–verification–iteration" cycle at 15 min resolution [35]. Application in the Hexi Corridor yields 15.3% water savings and 77% acceptance of recommended decisions, with implementation costs (~475,000 USD) offset in ~6.2 years via ~19% pumping-cost reductions, indicating technical and economic feasibility [35].

Chen et al.'s simulation-driven DT for the Tailan River reservoir, although framed as drought-risk mitigation, also embodies a WEF logic: the dual-scale RF prediction engine and HCO implicitly balance irrigation water availability against maintaining storage above drought-risk thresholds, enabling exploration of pumping schedules that minimize low-storage probability while meeting irrigation demand [34]. Szanyi et al. similarly propose using their Nyírség MODFLOW framework as an updatable groundwater DT that periodically recalibrates with new levels, pumping and climate data to derive adaptive abstraction caps linked directly to irrigation planning and crop choices, thereby aligning the WEF nexus with groundwater sustainability [46]. Collectively, these works indicate that WEF-oriented agricultural DTs require multi-objective optimization (NSGA-III, HCO), groundwater models capable of quantifying aquifer responses under climate and pumping scenarios, and cyber-physical loops where real-time sensing verifies model predictions and triggers re-optimization [34,35,46].

3.4.2. DT-Enabled Intelligent Irrigation–Drainage and Water Reuse in Saline Agroecosystems

In salt-affected agroecosystems, DTs must jointly manage water quantity, energy use and root-zone salinity [36]. Qin et al. develop a DT-enabled intelligent irrigation–drainage system for saline fields in north-western China, integrating IoT sensing, LSTM-based prediction and closed-loop water-reuse infrastructure [36]. The physical setup comprises Yellow River supply canals, pipelines for reusing saline ditch and purified brackish water, subsurface drains and photovoltaic-powered brackish-water treatment, deployed over a 6.67 ha experimental site with heterogeneous soils [36]. High-frequency measurements of soil moisture, temperature and EC at multiple depths, together with groundwater depth and drainage-network quality, feed the DT.

The LSTM core forecasts soil EC and volumetric water content under alternative irrigation and drainage schedules, achieving $R^2 \approx 0.97$ for EC and ≈0.92 for water content [36]. Feature-importance analysis identifies soil moisture, temperature and groundwater depth as dominant drivers of water–salt interactions, providing interpretable guidance for control rules. Real-time data are ingested into the model, which predicts short-term impacts of candidate operations on salinity profiles and root-zone water status; control logic then adjusts irrigation and drainage to enhance leaching when thresholds are approached, while avoiding wasteful over-irrigation [36]. Field experiments show that DT-guided, coordinated irrigation–drainage reduces salinity in permeable horizons without increasing deeper salinity or compromising water-use efficiency, and that closed-loop reuse reduces dependence on fresh Yellow River water and the energy footprint of long-distance transfers. From a WEF perspective, this system exemplifies a fine-scale agricultural DT where groundwater depth, soil water, salinity and multi-source irrigation water are jointly managed using predictive ML to pre-empt salinity stress and optimize water applications [36].

3.4.3. Climate-Aware DTs Linking Groundwater, Irrigation and MAR

Szanyi et al.'s Nyírség study illustrates how climate-aware numerical frameworks can evolve into DTs that inform irrigation and MAR at regional scale [46]. A ~3.1 °C rise in mean annual temperature over 40 years (1.8 °C in the last 15 years) and marked snow-storage reduction increase evaporative demand and reduce cold-season recharge; calibrated MODFLOW simulations reproduce observed groundwater declines of ~0.5–1.5 m since 2010, especially in elevated recharge zones with high irrigation potential [46]. Scenario analysis for 2050 shows that an "extreme drought" case (−20% infiltration, +25% evapotranspiration) causes widespread declines (>0.25 m across most of the domain, up to 1.25 m in recharge areas), with additional pumping intensifying localized drawdown near major wellfields [46].

MAR options (channel/lake infiltration vs. subsurface injection) indicate that injection provides ~30% more effective replenishment and a larger radius of influence (~1–1.25 km for a 5 cm rise) than surface infiltration. Szanyi et al. propose institutionalizing their framework as an updatable DT that continuously ingests groundwater, pumping and hydroclimate data, periodically recalibrates the model, computes threshold-based indicators (e.g., depth-to-water, storage deficits) at management-unit scale and converts these into adaptive extraction caps and cropping recommendations [46]. Combined with the multi-objective optimization and verification loops demonstrated in Hebei and Tailan, this points toward climate-aware agricultural DTs that jointly internalize climate risks, groundwater dynamics, energy costs and food-production targets at basin scale [34,35,46].

3.4.4. Cross-Cutting Traits of Agricultural and WEF-Nexus DTs

Across these studies, agricultural and WEF-nexus DTs share several traits. They are multi-objective and multi-scalar, balancing groundwater sustainability, energy consumption and yield stability from field plots to irrigation districts and basins [34,35,46]. They rely on tight cyber-physical integration, with dense sensor networks feeding predictive models that drive automated or semi-automated control of pumps, valves and drainage infrastructure [34–36]. Advanced ML methods (RF, LSTM, ANN surrogates) are used mainly as surrogates or forecasters embedded in physically informed architectures, rather than as stand-alone black boxes, helping maintain hydrological plausibility and stakeholder trust [34–36].

Furthermore, climate change is explicitly recognized as reshaping the feasible operating space for agricultural groundwater use, pushing DTs toward dynamic, threshold-based governance in which extraction caps, irrigation strategies and MAR deployments are periodically recalibrated as hydroclimatic and socio-economic conditions evolve [34,35,46]. In this sense, agricultural and WEF-nexus DTs act as a bridge between aquifer-scale twins and broader socio-hydrological governance, ensuring that high-resolution subsurface insights are translated into actionable, climate-robust decisions on water allocation, energy use and food production.

3.5. S5: Basin and Urban Water DTs

Basin and city-scale digital twins (DTs) extend beyond single aquifers to represent the full continuum from climate forcing through rivers and aquifers to distribution networks and irrigation schemes, aiming to support adaptation, risk reduction and resource optimization [17,20,35,40,46]. Architecturally, these systems combine distributed hydrological models, monitoring networks, optimization engines and, increasingly, AI components, and they are being positioned as templates for broader environmental DTs that can be nested or federated across scales [17,20,34,35,46].

3.5.1. Regional Natural-Resource and Basin-Scale Digital Twins

At regional scale, Iglesias et al. propose an AI-assisted DT architecture for natural resource management in Centre-Val de Loire, using the Beauce aquifer as a key groundwater example [37]. The DT is decomposed into Science, Information, Computational, Engineering and Technology layers, separating hydrological concepts and indicators from data models, analytics and deployment, and enabling modular integration of groundwater models, AI services and external tools; it is explicitly intended as a reusable blueprint for other basins and environmental themes [37].

Szanyi et al. provide a basin-scale framework for the Nyírség recharge–discharge system (northeastern Hungary), combining 1970–2022 shallow-groundwater monitoring, CHIRPS and ERA5-Land climate data and a calibrated MODFLOW model to quantify the effects of warming, declining snow storage and abstraction on recharge and storage [46].

Observations and simulations show intensified post-2010 groundwater decline (up to ~3 m in individual wells and ~0.5–1.5 m basin-average), and mid-century "extreme drought" scenarios (−20% infiltration, +25% evapotranspiration) lead to >0.25 m declines over most of the domain and ~1.25 m in recharge zones, while additional pumping mainly creates localized drawdown hotspots [46]. MAR scenario tests indicate limited influence of channel/lake infiltration (hundreds of meters) compared with drainage-injection wells (~1–1.25 km radius, ~30% more effective replenishment), motivating a DT-like workflow in which continuous monitoring, periodic recalibration and threshold-based indicators support adaptive abstraction caps and targeted MAR [46].

Huang et al. extend DT concepts to irrigated basins in Hebei, where per capita water availability is ~one-seventh of the national average, 72% of groundwater zones are over-exploited and pumping accounts for ~60% of energy use [35]. An enhanced NSGA-III algorithm is coupled with a four-dimensional nexus model (groundwater dynamics, pump energy lifecycle, economic returns, climate-resilient yields) and a real-time DT verification loop (32 meteorological stations, 48 soil sensors), yielding strategies with ~18.5% water savings, 1285 kWh ha^{-1} energy reduction and ~0.07 m groundwater recovery during droughts, with head-prediction errors ~0.03 m and decision latency ~186 ms [35]. The framework effectively functions as a regional water resources DT linking basin-scale hydrogeologic responses to district-scale water and energy operations [35].

Chen et al.'s Tailan River groundwater-reservoir DT also operates at basin scale, embedding a reservoir-focused sub-twin in a larger alluvial-fan system that integrates mountain runoff, alluvial recharge and agricultural withdrawals in Xinjiang [34]. A nested-grid MODFLOW configuration (500 m for the basin, 50 m for the reservoir) provides basin-wide groundwater dynamics that feed RF surrogates and hierarchical optimization, mirroring the nested structure envisaged in other regional DT platforms [20,34]. Henriksen et al. go further by implementing the national-scale DK-model and HIP portal for Denmark, simulating coupled surface–groundwater processes and streaming outputs and observations via a web platform; local, higher-resolution DTs (river basins, wetlands, wellfields) can plug into this backbone using its boundary conditions and climate scenarios while supplying refined local information and rules back to the national risk knowledge base [20].

At continental scale, Ma et al. construct a hyper-resolution groundwater DT for the contiguous United States (CONUS) by training a physics-guided adjusted Random Forest on 1 km ParFlow-CLM simulations and water-table observations to downscale and bias-correct daily WTD to 1-arcsecond (~30 m) [17]. The DT produces continental maps with improved temporal skill (median Spearman $\rho \approx 0.66$ in the test year and increased by ~0.13 in a future period), and multi-GPU computing generates each daily 1-arcsec map in ~35 min while quantifying uncertainty from tree-level outputs, demonstrating how physically based modeling and ML surrogates can support operational groundwater decision-making across an entire continent [17]. Collectively, these basin-scale, national and continental DTs show how climate- and pumping-driven pressures on groundwater, and the effectiveness of MAR and demand-side measures, can be monitored and managed through adaptive, threshold-based guidance on extraction and infrastructure operations [17,20,34,35,46].

3.5.2. Urban Water-Cycle DTs and the River–Aquifer–Pipe Continuum

Marcoie et al. address DTs in the urban water cycle, arguing that cities must be managed as a "river–aquifer–pipe network continuum" rather than as separate stormwater, supply and river-engineering subsystems [40]. Case studies from eastern Romania (Prut, Olt, Bahlui basins and the Iași network) illustrate how channelization, unplanned urbanization and aging infrastructure convert stormwater from a recharge source into flood hazard, degrade rivers and stress aquifers and networks through over-extraction and leakage [40].

Hydrogeological modeling of the Gherăești-Bacău wellfield shows that intensive pumping can shift river–aquifer interactions from gaining to losing and reduce sustainable yield, while analysis of the Iași distribution network reveals significant leakage feeding an "urban aquifer" that both recharges and potentially contaminates shallow groundwater [40].

To operationalize the continuum concept, they propose a three-stage Diagnose–Prescribe–Optimize (DPO) framework anchored in an integrated DT. In Diagnose, hydrodynamic flood, hydrogeological and hydraulic network models are combined into a system-wide DT to map vulnerabilities and river–aquifer–network feedbacks; in Prescribe, this DT supports nature-based and structural interventions (e.g., restoration of >8000 ha of Trifești–Sculeni wetlands) to restore attenuation and water quality and indirectly support aquifer recharge; and in Optimize, smart technologies such as active pressure management reduce leakage, energy use and mechanical stress within hydrological and ecological constraints [40]. Although conceptual, the framework emphasizes that urban DTs must integrate river, aquifer and network models in adaptive loops where monitoring feeds back into the twin, enabling iterative refinement of nature-based measures and smart controls and embedding groundwater within full-cycle urban management under climate and urban-growth pressures.

3.5.3. Cross-Scale Integration and Templates for Environmental DTs

Taken together, these basin and urban DTs provide templates for broader environmental twins that link climate drivers, hydrology, infrastructure and human decisions across scales [17,20,35,37,40,46]. Conceptually, Iglesias et al.'s ENVRI-inspired layered architecture shows how groundwater and other resource domains can be organized within a common DT framework separating science concepts, data standards, computational services and deployment technologies, thereby facilitating cross-domain interoperability [37]. Operationally, HIP's federated national DT, Nyírség's climate-aware basin model, Hebei's WEF-nexus twin and the CONUS groundwater DT demonstrate implementation with real- or near-real-time data, calibrated models and optimization/ML engines serving diverse users [17,20,35,46].

A key theme is nesting and federation: national backbones like HIP and continental twins like CONUS provide boundary conditions and climate scenarios for regional and local twins, while basin and urban DTs (Tailan, Nyírség, Hebei, Iași) can feed aggregated indicators (groundwater-storage deficits, risk indices, leakage-induced recharge) back to coarser-scale twins informing national or transboundary policy [17,20,34,35,40,46]. Another is integration of optimization and governance: Hebei's NSGA-III engine and Iași's DPO framework illustrate how DTs move from passive diagnostics to active, multi-objective management of water, energy and ecological constraints in basins and cities [35,40]. Finally, climate change drives the need for explicitly adaptive DTs with periodic recalibration, threshold-based early-warning indicators and dynamic adjustment of abstraction limits, MAR deployment and infrastructure operations, making these basin-, national- and continental-scale twins both practical tools for groundwater-inclusive water management and methodological templates for future environmental DTs under coupled hydroclimatic conditions.

3.6. S6: Sensing, Monitoring and IoT Infrastructures

Robust sensing and data infrastructures are the backbone of groundwater-centric digital twins, because they determine how faithfully subsurface states and fluxes are represented in the virtual environment and how quickly DTs can react to evolving hydroclimatic and operational conditions [20,34,35,39]. Across current applications, three complementary strands can be distinguished: (i) conceptual visions of "autonomous" subsurface sensing

tightly coupled to computational platforms, (ii) practical DT prototypes with embedded well instrumentation and MAR/abstraction monitoring, and (iii) IoT-enabled field deployments in agricultural and saline environments, supported by multilayer data fusion frameworks for water-table monitoring [20,21,34–36,39,41,42,45].

3.6.1. From Conceptual "Autonomous" Sensing to Operational DT-Ready Networks

Davis et al. outline an ambitious vision for "autonomous" sensing and prediction of subsurface contamination, framed around a digital twin of contaminated sites [39]. They argue that subsurface characterization is still largely manual and episodic, and propose dense networks of in situ, often low-power sensors (temperature, soil moisture, gas composition) deployed throughout vadose and saturated zones, telemetering data via IoT links to an intelligent database and computational platform that continuously updates a process-based digital representation [39]. In this concept, coupled flow-and-transport and degradation models assimilate streaming data to provide near-real-time estimates of mass loss and risk, replacing intermittent investigations with self-updating site models [39]. A petroleum-site case study with online gas and temperature sensors illustrates this approach: multi-parameter, real-time data, interpreted within a high-quality conceptual site model and simulations, yield spatially and temporally resolved NSZD rates and plume evolution beyond what sporadic sampling can provide [39].

At the other end of the spectrum, Henriksen et al.'s HIP DT for Denmark shows how nationwide monitoring networks (groundwater, soil moisture, streamflow, meteorology) can support a continuously running coupled surface–groundwater model for climate adaptation and risk reduction [20]. Although HIP focuses on quantity rather than contamination, its architecture—sensor networks feeding a continuously updated national model—closely matches the "self-adjusting computational platform" envisaged by Davis et al., demonstrating how large-scale sensing–modeling infrastructures can underpin groundwater-inclusive DTs [20,39].

3.6.2. Well-Scale Instrumentation and Embedded DT Prototypes

At well and confined-aquifer scale, Cohen-Manrique et al. implement a DT prototype for the Morroa aquifer (Colombia) based on embedded instrumentation and resilience-oriented control [45]. An ESP32 microcontroller (Espressif Systems, Shanghai, China) with low-cost sensors (DS18B20, Maxim Integrated, San Jose, CA, USA; SKU-SEN0161, SKU-DFR0300, SEN0237-A, DFRobot, Shanghai, China) measures temperature, pH, dissolved oxygen and water level at a key production well, transmitting data to a server where they are merged with historical and climatic information and used to calibrate a numerical model [45]. Heuristic optimization (grid search, genetic algorithms, PSO) then identifies extraction rates that keep dynamic water levels near a target (e.g., 58.5 m), revealing that current pumping is unsustainable even in wet years and providing an extensible template for a multi-well DT [45].

Chen et al.'s Tailan River groundwater-reservoir DT is similarly grounded in multi-source monitoring [34]. Their data platform integrates hydrogeological characterization, operational data from eight IoT-equipped wells (levels, pumping, artificial recharge) and meteorological/hydrological drivers, which are aggregated, quality-checked and gap-filled to create a coherent dataset for RF surrogate training [34]. In both Morroa and Tailan, sensor types, locations and sampling frequencies are co-designed with model structure: water-level sensors in hydrologically representative wells, operational sensors on control variables, and climate sensors on key recharge drivers, underscoring that DT informativeness depends critically on process-aware sensor network design [34,45].

3.6.3. IoT-Enabled Field Deployments in Agricultural and Saline Systems

In irrigated and saline agroecosystems, sensing must resolve groundwater heads and soil–water–salinity dynamics at plot scale [35,36]. Qin et al. implement a DT-enabled intelligent irrigation–drainage system over 6.67 ha in a saline field in north-western China, combining IoT-based sensing with LSTM prediction and closed-loop water-reuse infrastructure [36]. Soil water content, temperature and EC are measured at multiple depths, together with groundwater depth and drainage-network quality, and streamed to an LSTM model that forecasts soil EC and water content under alternative irrigation–drainage schedules [36]. This allows the DT to anticipate when root-zone salinity approaches crop-specific thresholds and to adjust quotas, timing and drainage to enhance leaching while conserving water [36]. The tightly instrumented physical system—closed-loop channels, reuse pipelines, subsurface drains, PV-powered treatment units—enables tracking of the full water–salt circuit and demonstrates that dense, stratified sensing can reduce salinity in permeable horizons without compromising water-use efficiency [36].

Huang et al.'s Hebei WEF-nexus DT likewise relies on distributed sensing: 32 meteorological stations and 48 soil sensors provide spatially resolved rainfall, temperature and soil-moisture data to a quadruple-objective nexus model and NSGA-III optimizer, while groundwater-level sensors at representative points supply the verification loop with head data (prediction errors ~0.03 m) [35]. Coastal and island DTs for SWI management, such as Sharan et al.'s Port Vila system, build on observation-well networks measuring levels and salinity at depth to calibrate and validate SEAWAT models and surrogates, forming the empirical basis for long-term DTs that track the saltwater interface under changing pumping and recharge [21].

3.6.4. Multilayer Frameworks for Water-Table Monitoring and Data Fusion

Digital twins require not only sensors but also data fusion and management frameworks that integrate heterogeneous observations into standardized, model-ready datasets [20,34,35,41,42]. Duggal et al. address this for over-exploited aquifers in Punjab via a three-layer architecture—Data, Modeling, Application [41]. The Data layer performs sensor/source identification, acquisition, cleaning, unit harmonization and spatio-temporal alignment of piezometer, EC, pH, rainfall, satellite and GIS data, using methods such as interpolation and Kalman/Bayesian fusion; outputs are stored in centralized repositories (e.g., CSV, NetCDF, HDF5) as quality-controlled inputs for groundwater and ML models in the Modeling layer, while the Application layer delivers dashboards and decision-support tools [41].

Singh and Sharma's Randullabad watershed DT (DToGW) adopts a similar multilayer scheme, with a GIS-centric data layer ingesting IoT readings, satellite land-use/vegetation indices, topography and socio-economic information, feeding coupled hydrological–groundwater models and a governance-oriented application layer that produces spatial water-table maps, potential zonation and scenarios for village co-design of abstraction and recharge interventions [42]. Iglesias et al. and Henriksen et al. extend these ideas with ENVRI/FIWARE and HIP data standards: NGSI APIs and containerized services in Centre-Val de Loire, and shared hydrological data models, calibration protocols and open web services in Denmark, enabling local DTs to plug into regional/national backbones with consistent semantics [20,37].

3.6.5. Cross-Cutting Insights for Sensing and IoT in Groundwater DTs

The reviewed studies yield several design principles for sensing and IoT in groundwater DTs. First, hydro-process-aware sensor design is essential: networks must be configured around dominant processes, whether NSZD in contaminated soils, drought-driven storage

depletion, climate-induced recharge changes or root-zone salt accumulation [30,34,39,45,46]. Second, multi-parameter, multi-depth monitoring (e.g., vertical profiling of heads, EC, pH, temperature, gases and soil moisture) is critical for reconstructing subsurface states that cannot be inferred from single-variable series [21,36,39,41,45].

Third, sensor networks must be tightly integrated with specific models and optimization engines—MODFLOW, SEAWAT, surrogates, LSTMs, PSO, NSGA-III, HCO—whose resolution and error requirements drive sensing design; DTs should thus be conceived as co-evolving sensor–model systems [21,34,35,45]. Fourth, robust data engineering and standardization, as in Punjab, Randullabad, ENVRI/FIWARE and HIP, are indispensable for moving from ad hoc monitoring to continuous DT-grade information flows [20,35,37,41,42].

Finally, scalability from well to nation is achievable: ESP32-based prototypes (Morroa), eight-well IoT networks (Tailan), field-scale deployments (saline fields), basin-scale arrays (Hebei) and national networks (Denmark) all apply the same core principles—continuous measurement, automated transfer, standardized storage, and tight model coupling [20–34,45,46]. Sensing, monitoring and IoT infrastructures therefore emerge as central design elements of groundwater DTs, determining what can be known, how quickly systems can respond and how confidently decisions can be made under uncertainty [20,21,34,35,39,41,42,45].

3.7. S7: Governance, Resilience and Socio-Hydrology

The recent wave of groundwater-focused digital twins (DTs) makes clear that technical advances in sensing and modeling are only part of the story: without careful attention to governance, resilience goals and socio-hydrological feedbacks, DTs risk becoming sophisticated dashboards with little effect on actual abstraction practices, crop choices or infrastructure decisions [20,28,30]. The representative works in this sub-topic collectively argue for DTs as socio-technical systems: they must embed resilience-based control logics, open and multi-level governance frameworks, smart-village institutions and participatory modeling processes that explicitly reflect stakeholder behavior, policy levers and long-term sustainability targets [20,28–30,38,40,42,45,46].

3.7.1. AI-Enabled Governance and "Intelligent Aquifer Systems"

Rahman et al. frame the evolution of AI in groundwater science as a trajectory from isolated proof-of-concept models toward a new generation of "intelligent aquifer systems", defined as dynamic platforms that combine continuous monitoring, adaptive learning and participatory decision-making": dynamic platforms that couple dense IoT and satellite monitoring, continuously retrained ML models and optimization engines (e.g., reinforcement learning, evolutionary algorithms) to adjust pumping and recharge within sustainability thresholds [28]. They stress that such systems must be embedded in ethical, inclusive governance based on open datasets, transparent model designs, cross-regional validation, explainable AI (XAI) and institutional capacity-building, so that agencies and communities can scrutinize and steer AI-driven recommendations [28,29].

Ma et al. similarly argue for "physics-informed water intelligence", in which DT architectures provide hooks for physics-informed models, XAI modules and safe reinforcement-learning policies [29]. Control decisions (e.g., pumping adjustments) should thus be physically defensible, accompanied by uncertainty estimates and traceable rationales that can be examined in regulatory and public forums, making DTs decision-support instruments rather than opaque oracles [28,29].

3.7.2. Smart-Village Groundwater Governance and Socio-Hydrology

At local scale, Singh and Sharma propose a smart-village framework for the semi-arid Satara (Maharashtra) that integrates multi-source data, GIS-based hydrogeological analysis,

AI-driven forecasts and a village-scale aquifer DT within a governance structure emphasizing transparency and shared access among rural stakeholders [38]. The twin provides water-table maps, forecasts and "what-if" scenarios that Gram Panchayats and farmer collectives use to negotiate well-spacing rules, abstraction limits and recharge interventions, explicitly drawing on socio-hydrology and participatory management literature [30,38].

Their Randullabad Digital Twin of Watershed and Groundwater (DToRWG) refines this concept using a specific architecture, distinguishing full twins (bi-directional data and control), digital shadows and static models, and operating under challenging conditions of recurrent drought, seasonal scarcity and poor groundwater quality [42]. By fusing IoT, remote sensing and historical data, the DToRWG identifies recharge-potential and safe-yield zones and prioritizes climate-resilient interventions, positioning the twin as both an SDG-6 tool and a template for rural water-governance reform [38,42]. These smart-village twins act as boundary objects linking hydrogeology and local socio-political processes, provided they are explicitly designed for participatory use and local knowledge integration [30,38,42].

3.7.3. National and Basin-Scale Resilience Governance

At national scale, Henriksen et al. present Denmark's HIP system and portal as a de facto DT for climate adaptation, water management and disaster risk reduction [20]. Built around the fully distributed DK-model and national monitoring networks, HIP acts as a shared risk-knowledge system providing consistent groundwater, streamflow and water-security indicators to municipalities, utilities, emergency managers and national agencies [20]. Its federated design allows local or basin-scale DTs to plug into the national backbone, using DK-model outputs as boundary conditions and feeding back aggregated indicators, with open data, shared calibration rules and common resilience metrics as prerequisites for policy uptake [20].

Szanyi et al. complement this with a basin-scale view for Nyírség, proposing their MODFLOW framework as an updatable DT that periodically recalibrates with new monitoring data to define depth-to-water thresholds, storage-deficit indicators and adaptive extraction caps at management-unit scale [46]. These caps are then intended to link directly to cropping, irrigation and MAR decisions, aligning basin-level resilience targets with local water-use practices [46]. Together, HIP and Nyírség show how concepts such as early-warning thresholds and safe operating spaces can be operationalized in DTs and connected to formal and informal governance at multiple scales [20,46].

3.7.4. Integrated Urban Water Governance and the River–Aquifer–Pipe Continuum

Marcoie et al. address urban governance through the lens of a "river–aquifer–pipe network continuum", showing how channelization, urbanization and high-pressure distribution networks jointly degrade rivers, alter river–aquifer exchanges and create leakage-fed "urban aquifers" [40]. They propose a Diagnose–Prescribe–Optimize (DPO) framework in which an integrated DT—combining hydrodynamic river, hydrogeological and network-hydraulic models—serves as a diagnostic engine to map dysfunctions, quantify feedbacks and support scenario analysis [40].

In DPO, DT-based diagnosis informs nature-based and structural measures (e.g., Trifești–Sculeni floodplain restoration) in the Prescribe phase, while the Optimize phase deploys smart technologies such as active pressure management to reduce leakage, energy use and asset stress [40]. The framework is explicitly cast as a technological governance workflow: diagnosis must be DT-driven, prescription must prioritize nature-based solutions and optimization must use smart sensing and control, providing a concrete roadmap for embedding DTs into integrated urban water management [40].

3.7.5. Resilience-Based Control, Decision Support and Multi-Paradigm Socio-Hydrological Modeling

Cohen-Manrique et al.'s Morroa DT exemplifies resilience-based control and stakeholder-oriented decision support at confined-aquifer scale [45]. A target dynamic water level at a key well defines the resilience objective; heuristic optimization (grid search, GA, PSO) identifies extraction rates that keep simulated heads near this target under physical and institutional constraints, and an ESP32-based sensor suite provides real-time data to support decision-making by environmental authorities [45].

Secci et al. extend governance and socio-hydrology to basin scale in the Konya Closed Basin, integrating process-based models, ANN surrogates and system-dynamics representations of farmers' decisions, irrigation demand, well investments and policy levers, developed through a "living lab" stakeholder process [30]. Process-based models define hydrological constraints, ANNs provide fast groundwater predictions and system-dynamics components capture socio-economic feedbacks, enabling stakeholders to explore policy scenarios over decades in terms of both groundwater storage and economic outcomes [30]. Rahman et al. and Ma et al. argue that this multi-paradigm triad—mechanistic, data-driven and socio-dynamic—is essential for DTs that aim to inform not only "what the aquifer will do" but also "what actors and institutions will do in response" [28–30]. Smart-village twins such as Randullabad further underline that DTs should empower local actors via granular, place-based knowledge and intuitive tools for community-driven rules aligned with long-term sustainability [30,38,42,46].

3.7.6. Emerging Principles for DT-Enabled Groundwater Governance

Synthesizing across these contributions, several governance and socio-hydrological principles for groundwater digital twins emerge:

- Resilience as a design objective: Whether expressed as target water-level thresholds (Morroa, Nyírség), safe storage envelopes (Tailan, Konya) or multi-objective WEF goals (Hebei), resilience metrics must be explicitly encoded into DT architectures and optimization problems, not treated as ex post evaluation criteria [20,30,34,35,45,46].
- Multi-level, participatory governance: From smart villages in Satara to national HIP backbones in Denmark, effective DTs are designed for use by multiple actor groups, with open data policies, accessible interfaces and participatory modeling processes that treat local knowledge and social capital as integral to groundwater management [20,30,38,40,42].
- Multi-paradigm socio-hydrological modeling: The complementarity of process-based, data-driven and system-dynamics models is not only a technical consideration but a governance necessity, enabling DTs to represent both the physics of aquifers and the behavior of actors and institutions that drive abstraction and recharge decisions [28–30].
- Transparency, explainability and ethics: AI components within DTs must be transparent and interpretable, with clear documentation of training data, physical constraints and uncertainty, so that their outputs can be interrogated and contested in governance processes rather than treated as opaque "truths" [28,29].
- Climate-aware adaptivity: Climate change is reconfiguring recharge, demand and risk, so DTs must be designed as adaptive platforms with periodic recalibration, threshold-based alerts and mechanisms to adjust extraction caps, MAR deployment and cropping strategies as conditions evolve [20,28,30,46].

Under these principles, groundwater digital twins become not just technical artifacts but governance instruments that can help reconcile competing demands on groundwater, support equitable and climate-resilient policies and anchor socio-hydrological

learning processes across scales—from individual wells and villages to basins and nations [20,28–30,38,40,42,45,46].

4. Discussion

4.1. Cross-Cutting Patterns and Gaps Across Sub-Topics

The thematic classification (Table 1) shows that the emerging DT–groundwater literature is highly cross-linked rather than neatly partitioned into disjoint themes. Works such as [17,20,30,34,35,46] sit at the intersection of three or more sub-topics, indicating that AI & ML foundations (S1), DT architectures (S2), aquifer-scale applications (S3), WEF-nexus systems (S4), basin/urban DTs (S5), sensing infrastructures (S6) and governance/socio-hydrology (S7) are already being combined in practice. For example, Refs. [34,35] simultaneously advance surrogate modeling, aquifer-scale applications and WEF optimization, while [17,20] operate at basin–national–continental scales and link DT architectures, sensing, and governance.

Table 1. Thematic classification and cross-cutting patterns in the reviewed articles among the seven identified sub-topics. The checkmark (✓) indicates that a reference pertains to a specific sub-topic.

Reference	Year	S1 AI & ML	S2 DT Architectures	S3 Aquifer-Scale DTs	S4 Agro & WEF DTs	S5 Basin & Urban DTs	S6 Sensing & IoT	S7 Governance & Socio-Hydrology
[28]	2026	✓						✓
[38]	2025		✓					✓
[34]	2026	✓		✓	✓	✓	✓	
[37]	2026		✓			✓		
[35]	2025	✓			✓	✓	✓	
[46]	2026			✓	✓	✓		✓
[33]	2023	✓		✓				
[21]	2025			✓			✓	
[39]	2023		✓				✓	
[40]	2026		✓			✓		✓
[20]	2023		✓	✓		✓	✓	✓
[45]	2025			✓			✓	✓
[29]	2026	✓	✓					✓
[36]	2025	✓			✓		✓	
[30]	2024	✓		✓	✓			✓
[17]	2026	✓				✓		
[41]	2025		✓	✓			✓	
[42]	2024		✓				✓	✓

Figure 3 makes these overlaps explicit by highlighting a small set of "bridging" contributions that connect otherwise sparsely populated thematic areas. In particular, S1 (AI and ML foundations) is tightly coupled to S3 and S4 through simulation-assisted surrogates and optimization in Tailan, Hebei, Konya and Barbados [17,30,33–35], while S2 (architectures) and S6 (sensing/IoT) form a backbone that underpins most operational or near-operational twins [21,34–37,39,41,42,45]. S7 (governance and socio-hydrology) appears less densely populated but is linked to several technical themes via a few key works: conceptual AI/governance framing [28,29], smart-village and watershed twins [38,42], socio-hydrological and system-dynamics modeling [30] and national risk-knowledge systems [20].

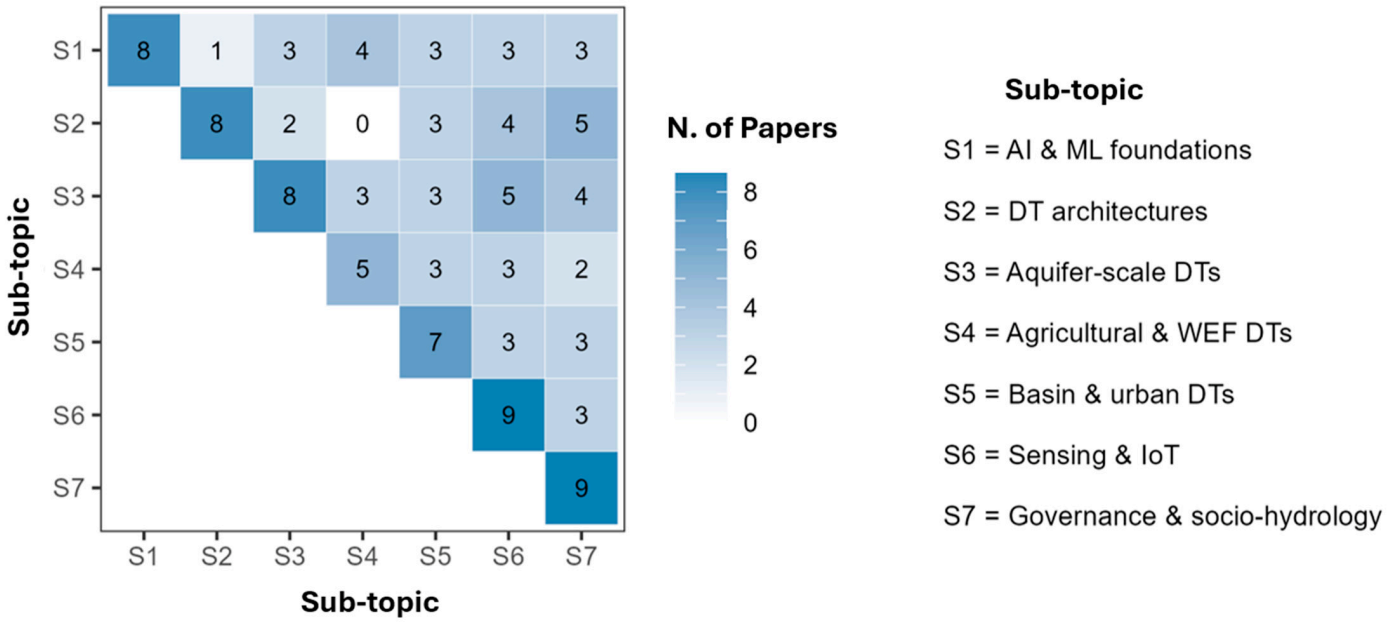


Figure 3. Sub-topic overlaps among literature review contributions.

The second pattern is temporal and conceptual maturity. The majority of technical DT implementations—Tailan, Morroa, Hebei, Port Vila, saline fields, Punjab—are recent (2023–2026), strongly case-specific and heavily reliant on numerical models and local sensor networks [21–34,41,45,46]. Conceptual and review-type contributions such as [28–30,39,40] play a disproportionately large role in articulating cross-cutting ideas: "intelligent aquifer systems", autonomous sensing, river–aquifer–pipe continua, trustworthy ML paradigms and multi-paradigm modeling. Figure 3 therefore also illustrates a maturity gradient: the best-developed overlaps are between S1/S3/S4 (AI-augmented aquifer and WEF applications) and S2/S5/S6 (architectures, multi-scale hydrology and sensing), whereas connections with S7 (socio-hydrology and governance) remain thin and often conceptual.

Finally, there is a clear spatial and sectoral bias. Most case studies arise from water-stressed regions in China (Tailan, Hebei, Hexi Corridor, saline fields), Turkey (Konya), Hungary (Nyírség), Colombia (Morroa), Barbados and Vanuatu [21,30,33–36,45,46]. Urban DTs for the river–aquifer–pipe continuum [40] and national/continental backbones [17,20] are still rare. Applications to groundwater quality beyond salinity and contamination (e.g., nutrients, emerging pollutants), managed ecosystems, and transboundary aquifers are essentially absent, signaling important thematic gaps despite the broad conceptual ambition visible in S1, S2 and S7.

4.2. Methodological and Data Limitations of the Current DT–Groundwater Literature

Across sub-topics, the most pervasive limitation is the heavy reliance on model-generated training data with limited multi-decadal, multi-site observational validation. Surrogate models in Tailan, Barbados and Konya are trained primarily on outputs from high-fidelity PBMs (MODFLOW/ParFlow-CLM/FEFLOW) rather than on long records of heads, fluxes or quality [17,30,33,34]. This raises the risk of "in silico echo chambers" where DTs reproduce structural biases of the parent model and perform poorly under hydroclimatic regimes or management patterns not represented in the synthetic training set, a concern explicitly raised in AI reviews [28,29] and in the Nyírség and HIP calibration discussions [20,46].

Second, sensing networks are often sparse, ad hoc and site-specific. Many implementations rely on a handful of instrumented wells (Tailan, Morroa, Port Vila) or a single

experimental field (saline agroecosystem), with limited analysis of optimal sensor density, placement or variable choice [21,34,36,39,45]. In contrast, Denmark's HIP twin and Hebei's WEF twin demonstrate the benefits of dense national or district-scale networks [20,35], yet even these do not systematically explore sensor–model co-design. Under-instrumented systems limit the twin's ability to resolve key processes such as SWI dynamics, NSZD or water–salt interactions and can foster misplaced confidence in DT outputs.

Third, multi-scale and cross-sector integration remain conceptually rich but methodologically thin. HIP, Hebei, Nyírség and the CONUS twin all hint at nested or federated configurations where local twins plug into basin, national or continental backbones [17,20,35,42,46]. However, formal up/downscaling schemes for states, fluxes, decisions and indicators are rarely specified. Similarly, WEF-nexus and urban continuum DTs consider groundwater within broader water–energy–grain and river–aquifer–pipe systems [35,40], but cross-sector feedbacks (energy pricing, crop markets, ecosystem services, regulatory constraints) are only partially represented.

Fourth, most DTs remain technology-centric. Only a few explicitly model behavior, institutions or co-production processes: the Konya living-lab system-dynamics model, smart-village and Randullabad twins, and Denmark's national risk-knowledge framework [20,30,38,40,42]. Governance failures—power asymmetries, enforcement gaps, local norms—are central drivers of unsustainable abstraction [28,30] yet they are largely external to current DT architectures.

Finally, AI & ML components are often proof-of-concept and only partially aligned with best practices for robustness, explainability and safety. Physics-informed constraints are still emerging, XAI is applied only in a subset of studies, and formal uncertainty quantification and safe reinforcement-learning policies are largely absent, despite being emphasized in [28,29]. ML surrogates are therefore vulnerable to regime shifts and may be difficult to defend in regulatory contexts if their limitations are not made explicit.

At the review level, limitations include reliance on two bibliographic databases (Scopus and Web of Science), the exclusion of non-English publications, and the impossibility of capturing all relevant studies due to terminology instability in this emerging field.

4.3. Implications for Practice and Policy

For practitioners and policymakers, the current DT–groundwater literature demonstrates both significant promise and clear constraints. On the positive side, case studies show that well-designed twins can materially improve groundwater management outcomes. In Hebei, NSGA-III coupled with a calibrated nexus model and a DT verification loop delivers double-digit water and energy savings with measurable groundwater recovery at acceptable payback times [35]. In Morroa, resilience-based control reveals unsustainable extraction even in wet years and supports revised pumping strategies [45]. In Port Vila and saline agroecosystems, DTs enable the design of pumping and irrigation–drainage strategies that jointly protect water quality and maintain yields [21,36].

If deployed without acknowledging data and model limitations. Tailan- and Barbados-style surrogates cannot be treated as general-purpose decision tools beyond the hydrogeological and management envelopes for which they were trained [33,34]. Nyírség demonstrates that climate-driven recharge reductions may dominate pumping effects at basin scale, implying that local MAR or efficiency measures, though beneficial, cannot fully offset regional climate signals [46]. HIP and CONUS show that national and continental twins can provide coherent risk-knowledge backbones [17,20], but their use for local licensing or enforcement must be mediated through local calibration, uncertainty analysis and stakeholder engagement.

For regulators, a critical implication is that DT outputs should be treated as structured evidence—transparent, interrogable and uncertain—rather than as definitive "answers". Works on AI trustworthiness and governance [28,29] and on socio-hydrological modeling [30] suggest that regulatory acceptance will depend on:

- The openness of data and model assumptions;
- Documented validation against multi-decadal, multi-site observations where available;
- Clear communication of uncertainty and sensitivity; and
- Procedures for contestation and revision when DT recommendations clash with local knowledge or equity concerns.

Smart-village and watershed twins highlight another implication: DTs can democratize groundwater information if they are deliberately designed to empower local users [38,42]. Village-scale maps of safe-yield zones, recharge potential and vulnerability hotspots—if coupled with participatory interfaces—can support community-driven rules such as bans on new boreholes in critical areas or collective investment in recharge structures [30,38,42,46]. Conversely, if DT infrastructures remain in the hands of a few agencies or private providers, they risk reinforcing existing asymmetries in knowledge and power, a concern explicitly raised in [28].

Finally, the WEF-nexus and urban continuum DTs demonstrate that groundwater policy cannot be divorced from energy and food policies or from urban infrastructure planning [35,40]. Pumping-cost reductions, leakage control, floodplain restoration and crop-mix changes appear repeatedly as co-levers in DT-based scenarios; policymakers will need governance arrangements that can coordinate across ministries and agencies if these integrated levers are to be implemented in practice.

4.4. Future Research Directions

Table 2 synthesizes main research gaps and several corresponding priority research directions that emerge from this review, with examples and links to literature.

Table 2. Future research directions emerging from this literature review, corresponding main research gaps and examples.

Thematic Area	Research Gap (with Examples)	Why It Matters for Practice	Future Research Directions/ Example Questions
Data, validation and generalizability	DTs and ML surrogates are often trained almost entirely on outputs from numerical models (e.g., Tailan RF surrogates, Barbados DNN/GAN, Konya ANN), with limited long-term, multi-site observational validation [30,33,34].	Risk of "in silico echo chambers" where DTs reproduce structural/model biases and fail under unseen hydroclimatic or management conditions, undermining trust and regulatory acceptance [28,29].	- Build open, multi-basin benchmark datasets for groundwater DTs (levels, pumping, climate, quality) to test transferability. - Design joint PBM–ML calibration strategies against multi-decadal data (as in Nyírség, HIP) rather than purely synthetic training sets. - Embed standardized uncertainty and sensitivity analyses into DT workflows (e.g., how sensitive are drought-risk prescriptions to surrogate structural choices?) [20,34,35,46].

Table 2. *Cont.*

Thematic Area	Research Gap (with Examples)	Why It Matters for Practice	Future Research Directions/ Example Questions
Sensing, monitoring and IoT	Existing DTs use relatively sparse, site-specific sensor networks (e.g., 8 wells in Tailan, a single ESP32-instrumented well in Morroa, limited wells in Port Vila, plot-scale IoT in saline fields), and there is little rigorous work on optimal sensor density, placement and modality [21,34,36,39,45].	Under-instrumented or poorly designed sensor networks constrain what the twin can reliably infer (e.g., about SWI fronts, NSZD rates or water–salt dynamics) and increase the risk of false confidence in DT outputs [20,39].	- Develop sensor–model co-design frameworks that jointly optimize sensor location, depth, variables and sampling frequency for specified DTs performance metrics. - Explore adaptive/mobile sensing guided by DT-identified hotspots and uncertainties (e.g., plume edges, SWI interfaces). - Study low-cost, robust IoT designs and maintenance models (including community-based monitoring) for long-term, DT-grade data provision in resource-constrained regions [34,36,38,39,41,42,45].
Multi-scale and cross-sector integration	Methodologies for nesting and federating DTs across local–basin–national scales (e.g., village aquifers → basins → HIP backbone) and across sectors (WEF nexus, river–aquifer–pipe continuum) remain ad hoc and largely conceptual [20,35,40,42,46].	Without consistent up/downscaling of states, fluxes and decision rules, local DTs risk being misaligned with basin or national policies; conversely, national backbones may not capture local constraints and opportunities (e.g., for MAR or local recharge structures) [20,46].	- Formalize multi-resolution coupling schemes for nested DTs (e.g., how should village-scale recharge and abstractions modify HIP-scale water-security indicators?). - Extend WEF and urban continuum DTs to fully represent cross-sector feedbacks (energy pricing, crop markets, ecosystem services). - Develop open, interoperable ontologies and APIs for hydrological and socio-economic objects to allow plug-and-play DT components across basins and countries [20,29,35,37,40].
Socio-hydrology, behavior and governance	Most DTs remain technology-centric; only a few explicitly represent farmer behavior, institutional rules or co-production processes (e.g., Konya living-lab system dynamics, smart-village and Randullabad concepts) [20,30,38,40,42].	Governance failures, not just technical constraints, drive unsustainable abstraction; DTs that ignore behavioral responses, power asymmetries and local norms risk producing technically optimal but socially infeasible or inequitable recommendations [28,30].	- Systematically embed group model building, living labs and citizen dashboards into DT design, extending the Konya and smart-village approaches to different socio-economic and cultural settings. - Develop evaluation frameworks that track not only hydrological outcomes but also equity, legitimacy, conflict reduction and compliance when DTs are deployed. - Explore institutional and legal frameworks for DT-informed rules (e.g., threshold-based caps, DT-driven MAR permits) [20,28,30,38,42,45].

Table 2. *Cont.*

Thematic Area	Research Gap (with Examples)	Why It Matters for Practice	Future Research Directions/ Example Questions
AI & ML trustworthiness and robustness	Many groundwater and water-system ML models are still proof-of-concept, trained on siloed datasets with limited use of physics-informed constraints, XAI, uncertainty quantification or safe reinforcement-learning policies [28,29,33,34].	Opaque or brittle ML components embedded in DTs can undermine stakeholder trust and may fail catastrophically under regime shifts (e.g., unprecedented droughts, policy shocks), leading to poor or unsafe management decisions [28,29].	- Advance physics-informed ML for groundwater, SWI and water–salt systems, to improve extrapolation and data efficiency. - Integrate XAI (e.g., SHAP, LIME, feature-importance diagnostics) and uncertainty metrics into all DT ML pipelines, so recommendations are accompanied by interpretable rationales. - Explore safe RL and hybrid control schemes that operate within physically and legally defined safety envelopes and can be audited by regulators. - Use federated learning to train shared models across utilities/basins without centralizing sensitive data [28,33–35,38].

Building on that synthesis, four broad agendas can be outlined:

(i) *Data, validation and generalizability*

There is a need to move from DTs trained almost exclusively on model outputs to twins anchored in long-term, multi-basin observational data. Open benchmark datasets combining groundwater levels, pumping, climate and quality for diverse hydrogeological and socio-economic contexts would allow systematic testing of surrogate transferability and robustness [20,28,29]. Methodologically, joint PBM–ML calibration against multi-decadal records, as glimpsed in Nyírség and HIP [20,46], should replace purely synthetic training regimes, and standardized uncertainty and sensitivity analyses should be embedded in DT workflows to assess, for example, how drought-risk prescriptions depend on surrogate structure or parameterization.

(ii) *Sensing, monitoring and IoT*

Future work should treat sensor networks and DT models as co-designed systems. Formal frameworks are needed to jointly optimize sensor locations, depths, measured variables and sampling frequencies for specified DT performance metrics (e.g., accuracy of SWI front position, NSZD rate, water–salt budget) in settings such as Port Vila, Morroa, Tailan, Hebei and saline fields [21,34,36,39,45]. Adaptive or mobile sensing guided by DT-identified hotspots and uncertainties (e.g., plume edges, SWI interfaces, recharge hotspots) is another promising direction. Research on low-cost, robust IoT designs and maintenance models—including community-based monitoring in rural smart-village settings—will be essential for long-term DT-grade data provision outside well-resourced utilities [34,36,38,39,41,42,45].

(iii) *Multi-scale and cross-sector integration*

Methodologies for nesting and federating DTs across local–basin–national scales and across sectors remain underdeveloped [20,35,40,42,46]. Research is needed on multi-resolution coupling schemes—for example, how village-scale recharge and abstraction (Satara, Randullabad) should modify basin-level resilience indicators (Nyírség, Konya) and

national water-security metrics (HIP), and how information should flow back down as constraints or incentives [20,29,35,37,40,46]. Future WEF and urban continuum DTs should more fully represent cross-sector feedbacks, including energy pricing, crop markets and ecosystem services, rather than treating them as exogenous scenario drivers. Developing open, interoperable ontologies and APIs for hydrological and socio-economic entities would facilitate plug-and-play DT components across basins and countries, building on ENVRI/FIWARE and HIP experiences [20,29,37].

(iv) *Socio-hydrology, behavior, governance and trustworthy AI*

Finally, DTs research needs to systematically integrate socio-hydrological processes and governance. Extending living-lab and group-model-building approaches from Konya and smart-village pilots to a broader set of contexts could help co-design DTs that reflect local values, power relations and institutions, rather than imposing purely technical solutions [20,30,38,42,45]. Evaluation frameworks should track not only hydrological metrics but also equity, legitimacy, conflict reduction and compliance when DT-based rules (e.g., threshold-based caps, DT-informed MAR permits) are implemented [28,30,45].

On the AI side, advancing physics-informed ML for groundwater flow, SWI and water–salt systems; integrating XAI and uncertainty metrics into all DT ML pipelines'; and exploring safe reinforcement-learning or hybrid control schemes within legally and physically defined safety envelopes are immediate priorities [28,29,33–35]. Federated learning architectures, as proposed in [29] and hinted at by multi-utility and multi-basin contexts, could allow shared models to be trained across utilities or regions without centralizing sensitive data. Under these directions, DTs can evolve from promising prototypes into robust, trustworthy socio-technical systems that support climate-resilient and equitable groundwater governance.

5. Conclusions

This review shows that digital twins for sustainable groundwater resources management have moved in a few years from conceptual visions to technically sophisticated, if still fragmented, implementations [20,28,34,35]. By structuring the literature into seven subtopics—AI & ML foundations, DT architectures, aquifer-scale applications, agricultural and WEF-nexus systems, basin/urban DTs, sensing/IoT infrastructures and governance/socio-hydrology—it becomes clear that digital twins are emerging as integrative socio-technical platforms rather than as isolated modeling tools [20,30,34,35,37,40]. Across these themes, the most advanced contributions combine process-based models, machine learning surrogates, dense sensing networks and optimization/control modules to support drought-risk mitigation, saltwater intrusion control, irrigation and salinity management, and climate adaptation in diverse hydrogeological settings [21,30,33–36,45,46].

A first overarching conclusion is that AI and ML are no longer peripheral utilities but core building blocks of groundwater digital twins. Hybrid architectures that use PBMs as "digital laboratories" and ML surrogates for fast prediction are now demonstrated from island and coastal aquifers (Barbados, Port Vila) to continental-scale systems (CONUS) and large irrigated basins (Tailan, Hebei, Konya) [17,21,30,33–35]. These implementations confirm that carefully trained surrogates can preserve hydrophysical realism while enabling real-time scenario exploration, multi-objective optimization and closed-loop control. At the same time, the analysis highlights a strong dependence on model-generated training data and often limited use of long-term, multi-site observations, underscoring the need for more systematic validation, uncertainty analysis and attention to transferability beyond the calibration envelope [17,20,28,29,33,34,46].

A second conclusion concerns architectures and data infrastructures. Layered and modular designs—such as the ENVRI-inspired natural-resource DT, the Punjab and Ran-

dullabad frameworks, the Hebei WEF twin, and the HIP national backbone—demonstrate that robust data-model decoupling, standardized interfaces and containerized services are critical for scalability and interoperability [20–38,40–42]. These cases collectively show that effective twins are built as co-designed sensor–model systems: hydro-process-aware monitoring networks, rigorous data fusion pipelines and open data standards are at least as important as the choice of numerical solver or ML architecture [20–34,39,41,42,45]. Yet, optimal sensor density and placement, formal nesting across local–basin–national scales and cross-sector coupling (e.g., between WEF and urban continuum DTs) remain methodologically underdeveloped, limiting the ability to propagate information and decisions consistently across scales [17,20,35,40,46].

Third, the review makes clear that governance and socio-hydrology are still the weakest links in the DT–groundwater chain. Only a subset of studies explicitly represent farmer behavior, institutions or co-production processes (Konya living lab, smart-village and Randullabad twins, national HIP risk-knowledge system), and even fewer embed resilience metrics and equity considerations directly into DT optimization problems [20,30,38,40,42,45,46]. Conceptual contributions to "intelligent aquifer systems" and physics-informed, trustworthy water intelligence emphasize that AI-enabled twins must be transparent, explainable and ethically governed if they are to gain regulatory acceptance and avoid reinforcing existing power asymmetries [28–30]. The synthesis therefore confirms that progress on algorithms and architectures must be matched by advances in participatory design, multi-level governance and evaluation frameworks that consider legitimacy, conflict reduction and compliance alongside hydrological performance [20,28,30,38,42,45].

Finally, this review points to a coherent agenda for the next generation of groundwater digital twins. Priority directions include: (i) building open, multi-basin benchmark datasets and joint PBM–ML calibration workflows to avoid "in silico echo chambers" and to test generalizability under climate extremes; (ii) developing sensor–model co-design and adaptive sensing strategies that optimize information gain for DT objectives in both data-rich and data-scarce regions; (iii) formalizing multi-resolution coupling schemes and interoperable ontologies that allow village, basin, national and continental twins to exchange states, indicators and decision rules; and (iv) systematically embedding physics-informed ML, XAI, safe reinforcement-learning and socio-dynamic models into DT architectures so that recommendations are robust, interpretable and socially feasible [17,20,29,30,34,35,46]. Under these conditions, digital twins can evolve from promising prototypes to operational socio-technical infrastructures that support climate-resilient, equitable and adaptive groundwater governance from the field and village scale to basins, nations and beyond [17,20,28,30,35,38,40,45,46].

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/hydrology13090236/s1, PRISMA 2020 Checklist.

Funding: This research received no external funding.

Data Availability Statement: The data supporting the findings of this review are available from the corresponding author upon reasonable request.

Acknowledgments: No AI tools were used to generate the scientific content, arguments, or conclusions of this manuscript. Only basic language-support tools (Word dictionary, Grammarly, and web searches) were used for grammar, spelling, and wording.

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

References

1. Bolton, A.; Butler, L.; Dabson, I.; Enzer, M.; Evans, M.; Fenemore, T.; Harradence, F.; Keaney, E.; Kemp, A.; Luck, A.; et al. The Gemini Principles. In *Guiding Values for the National Digital Twin and Information-Management Framework*; Centre for Digital Built Britain: Cambridge, UK, 2018. [CrossRef]
2. Tao, F.; Qi, Q. Make more digital twins. *Nature* **2019**, *573*, 490–491. [CrossRef] [PubMed]
3. Ghorbani Bam, P.; Rezaei, N.; Roubanis, A.; Austin, D.; Austin, E.; Tarroja, B.; Takacs, I.; Villez, K.; Rosso, D. Digital Twin Applications in the Water Sector: A Review. *Water* **2025**, *17*, 2957. [CrossRef]
4. Qureshi, M.; Feng, Z.; Kumar, V.; Lovreglio, R.; Abdelmegid, M. A Review of the Role of Digital Twin Applications for Water Sustainability. In Proceedings of the 32nd Annual Conference of the International Group for Lean Construction (IGLC 32), Auckland, New Zealand, 1–7 July 2024; pp. 731–742. [CrossRef] [PubMed]
5. Yang, Y.; Xie, C.; Fan, Z.; Xu, Z.; Melville, B.W.; Liu, G.; Hong, L. Digital twinning of river basins towards full-scale, sustainable and equitable water management and disaster mitigation. *npj Nat. Hazards* **2024**, *1*, 43. [CrossRef]
6. Brocca, L.; Barbetta, S.; Camici, S.; Ciabatta, L.; Dari, J.; Filippucci, P.; Massari, C.; Modanesi, S.; Tarpanelli, A.; Bonaccorsi, B.; et al. Digital Twin of the terrestrial water cycle: A glimpse into the future through high-resolution Earth observations. *Front. Sci.* **2024**, *1*, 1190191. [CrossRef]
7. Botai, J.O.; Ghosh, S.; Matheswaran, K.; Dickens, C.; Langa, N.; Garcia Andarcia, M. Options for digital twin application in developing country river basin management: A review. In *Colombo, Sri Lanka: International Water Management Institute (IWMI)*; CGIAR Initiative on Digital Innovation: Montpellier, France, 2023; 20p. Available online: https://hdl.handle.net/10568/134763 (accessed on 8 June 2026).
8. Li, Q. Digital Twins for Water Utilities: Architectures, Calibration Workflows and Measured Benefits. *Int. J. Mech. Electr. Eng.* **2025**, *6*, 34–42. [CrossRef]
9. Sweetapple, C.; Salomons, E.; Le Gall, F.; Abid, A.; Vamvakeridou-Lyroudia, L.; Chen, A.S.; van den Broeke, J. Integrating Epanet and FIWARE for Development of Water Distribution System Digital Twins. In *Springer Water*; Springer Nature: Berlin/Heidelberg, Germany, 2022; pp. 1081–1086. [CrossRef]
10. Borzì, I. Evaluating Sustainability Improvement of Pressure Regime in Water Distribution Systems Due to Network Partitioning. *Water* **2022**, *14*, 1787. [CrossRef]
11. Park, J.-Y.; Kim, K.-J.; Jeung, M.; Jang, I.-S.; Yu, J.-W.; Kang, M.-S.; Bae, H.-S.; Jeong, C.; Baek, S.-S. Digital-twin tool for a drinking water distribution system using augmented reality and EPANET. *Environ. Model. Softw.* **2026**, *197*, 106829. [CrossRef]
12. Kusuma, B.; Isnaini, K.; Hamdi, A. Dynamic IoT–PID Control for Energy-Efficient Water Distribution: EPANET-Based Digital Twin Validation in Varied Geographical Terrains. *J. Appl. Data Sci.* **2026**, *7*, 773–785. [CrossRef]
13. Borzì, I. Vulnerability assessment of water supply infrastructures through multiple indicator methodology. *J. Water Clim. Change* **2023**, *14*, 2110–2123. [CrossRef]
14. Berglund, E.Z.; Shafiee, M.E.; Xing, L.; Wen, J. Digital twins for water distribution systems. *J. Water Resour. Plan. Manag.* **2023**, *149*, 02523001. [CrossRef]
15. Borzì, I. Assessing impacts of climate crisis on water distribution systems though service inefficiency indicator. *J. Water Clim. Change* **2025**, *16*, 307–321. [CrossRef]
16. Chang, Y. AI-Enabled Digital Twin Framework for Real-Time Water Resource Management Using Multi-Source Remote Sensing Data. *J. Environ. Earth Sci.* **2008**, *4*, 203–229. [CrossRef]
17. Ma, Y.; Tijerina-Kreuzer, D.; Defnet, A.; Artavanis, G.; Condon, L.E.; Maxwell, R.M. A prototype hyper-resolution groundwater digital twin for the contiguous United States: Integrating physics-based modeling, machine learning, and observations. *J. Hydrol.* **2026**, *671*, 135189. [CrossRef]
18. Morales Ortega, L.R. A Digital Twin for Ground Water Table Monitoring. Master's Thesis, University of Twente, Enschede, The Netherlands, 2023. Available online: https://purl.utwente.nl/essays/97168 (accessed on 8 June 2026).
19. Borzì, I. Modeling Groundwater Resources in Data-Scarce Regions for Sustainable Management: Methodologies and Limits. *Hydrology* **2025**, *12*, 11. [CrossRef]
20. Henriksen, H.J.; Schneider, R.; Koch, J.; Ondracek, M.; Troldborg, L.; Seidenfaden, I.K.; Kragh, S.J.; Bøgh, E.; Stisen, S. A New Digital Twin for Climate Change Adaptation, Water Management, and Disaster Risk Reduction (HIP Digital Twin). *Water* **2023**, *15*, 25. [CrossRef]
21. Sharan, A.; Datta, B.; Roy, D.K.; Lal, A. Conceptual development and implementation of a digital twin model for managing saltwater intrusion of an island coastal aquifer. *Environ. Monit. Assess.* **2025**, *197*, 1148. [CrossRef] [PubMed] [PubMed Central]
22. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* **2021**, *202*, n71. [CrossRef] [PubMed]
23. Sivapalan, M.; Savenije, H.H.G.; Blöschl, G. Socio-hydrology: A new science of people and water. *Hydrol. Process.* **2012**, *26*, 1270–1276. [CrossRef]

24. Sivapalan, M.; Konar, M.; Srinivasan, V.; Chhatre, A.; Wutich, A.; Scott, C.A.; Wescoat, J.L.; Rodríguez-Iturbe, I. Socio-hydrology: Use-inspired water sustainability science for the Anthropocene. *Earths Future* **2014**, *2*, 225–230. [CrossRef]
25. Herrera-Franco, G.; Montalván-Burbano, N.; Carrión-Mero, P.; Bravo-Montero, L. Worldwide Research on Socio-Hydrology: A Bibliometric Analysis. *Water* **2021**, *13*, 1283. [CrossRef]
26. Kumar, P.; Avtar, R.; Dasgupta, R.; Johnson, B.A.; Mukherjee, A.; Ahsan, M.N.; Nguyen, D.C.H.; Nguyen, H.Q.; Shaw, R.; Mishra, B.K. Socio-hydrology: A Key Approach for Adaptation to Water Scarcity and Achieving Human Well-Being in Large Riverine Islands. *Prog. Disaster Sci.* **2020**, *8*, 100134. [CrossRef]
27. Khan, K.F.; Bagtzoglou, A.C.; Dokou, Z.; Anagnostou, E. A Socio-Hydrological Investigation with Groundwater Models to Assess Farmer's Perception on Water Management Fairness. *J. Hydrol.* **2023**, *620*, 129481. [CrossRef]
28. Rahman, M.; Raihan, A.; Rahman, S.M.; Hossain, M.A.; Benaafi, M.; Aljundi, I.H. Groundwater science in the age of AI: Emerging paradigms and challenges. *Adv. Space Res.* **2026**, *77*, 4184–4207. [CrossRef]
29. Ma, R.; Li, J.; Zhang, Z.; Liu, Y.; Xu, H.; Yan, W. Machine learning paradigms in natural and engineered water systems: From proof-of-concept to trustworthy deployment. *Water Res.* **2026**, *300*, 125932. [CrossRef] [PubMed]
30. Secci, D.; Saysel, A.K.; Uygur, İ.; Yoloğlu, O.C.; Zanini, A.; Copty, N.K. Modeling for sustainable groundwater management: Interdependence and potential complementarity of process-based, data-driven and system dynamics approaches. *Sci. Total Environ.* **2024**, *951*, 175491. [CrossRef] [PubMed]
31. Harbaugh, A.W.; Banta, E.R.; Hill, M.C.; McDonald, M.G. *MODFLOW-2000, The U.S. Geological Survey Modular Ground-Water Model: User Guide to Modularization Concepts and the Ground-Water Flow Process*; Open-File Report 00-92; U.S. Geological Survey: Reston, VA, USA, 2000; 121p. Available online: https://pubs.er.usgs.gov/ (accessed on 12 June 2026).
32. Diersch, H.-J.G. *FEFLOW Software—Finite Element Subsurface Flow and Transport Simulation System—Reference Manual*; WASY Institute for Water Resources Planning and Systems Research: Berlin, Germany, 2005.
33. Payne, K.; Chami, P.; Odle, I.; Yawson, D.O.; Paul, J.; Maharaj-Jagdip, A.; Cashman, A. Machine Learning for Surrogate Groundwater Modelling of a Small Carbonate Island. *Hydrology* **2023**, *10*, 2. [CrossRef]
34. Chen, T.; Wu, B.; Chen, D.; Du, M.; Wang, C.; Liu, K. Simulation-driven digital twin framework for drought-risk mitigation in groundwater reservoirs via hierarchical optimization. *Geomat. Nat. Hazards Risk* **2026**, *17*, 2652586. [CrossRef]
35. Huang, P.; Jia, M.; Gu, F. Multi-objective collaborative optimization of water resources in Hebei irrigation areas: Maximizing the benefits of the water-energy-grain nexus driven by the NSGA-III algorithm and verified by digital twins. *Front. Sustain. Food Syst.* **2025**, *9*, 1701718. [CrossRef]
36. Qin, M.; Zhang, C.; Ma, G.; Ma, Y.; Feng, X.; Li, P.; Feng, K. Digital twin-enabled intelligent irrigation-drainage system for precision water-salt management in saline agroecosystems. *Agric. Water Manag.* **2025**, *322*, 109957. [CrossRef]
37. Iglesias, F.; Ros, F.; Thuy, L.H.V.; Gourcy, L.; Moquet, J.-S.; Daële, V.; Dupraz, S. A conceptual architecture for AI-assisted Digital Twins in natural resource management. *Ecol. Inform.* **2026**, *94*, 103635. [CrossRef]
38. Singh, D.; Sharma, V. Proposing a digital twin-based sustainable water governance system for rural Indian villages. *Int. J. Inf. Technol.* **2025**, *17*, 1777–1783. [CrossRef]
39. Davis, G.B.; Rayner, J.L.; Donn, M.J. Advancing "Autonomous" sensing and prediction of the subsurface environment: A review and exploration of the challenges for soil and groundwater contamination. *Environ. Sci. Pollut. Res.* **2023**, *30*, 19520–19535. [CrossRef] [PubMed]
40. Marcoie, N.; Iliesi, E.; Barta, A.-I.; Raboșapca, I.; Toma, D.; Boboc, V.; Balan, C.-D.; Tofănică, B.-M. Integrated Management of the Urban Water Cycle: A Synthesis of Impacts and Solutions from Source to Tap. *Urban Sci.* **2026**, *10*, 175. [CrossRef]
41. Duggal, S.; Sharma, K.N.; Shilpa. Monitoring Water Tables in Punjab: A Digital Twin Approach. In Proceedings of the 2025 IEEE 4th International Conference for Advancement in Technology, ICONAT 2025, Goa, India, 19–21 September 2025. [CrossRef]
42. Singh, D.K.; Sharma, V. Smart Water Management for a Sustainable Tomorrow: A Digital Twin Approach for Rural India. In Proceedings of the 2024 3rd Edition of IEEE Delhi Section Flagship Conference, DELCON 2024, New Delhi, India, 21–23 November 2024. [CrossRef]
43. Rossetto, R.; De Filippis, G.; Borsi, I.; Foglia, L.; Cannata, M.; Criollo, R.; Vázquez-Suñé, E. Integrating free and open source tools and distributed modelling codes in GIS environment for data-based groundwater management. *Environ. Model. Softw.* **2018**, *107*, 210–230. [CrossRef]
44. Hughes, J.D.; Liu, J. MIKE SHE: Software for integrated surface water/ground water modeling. *Groundwater* **2008**, *46*, 797–802. [CrossRef]
45. Cohen-Manrique, C.S.; Villa-Ramírez, J.L.; Camacho-León, S.; Solano-Correa, Y.T.; Alvarez-Month, A.A.; Coronado-Hernández, O.E. Simulation and Optimisation Using a Digital Twin for Resilience-Based Management of Confined Aquifers. *Water* **2025**, *17*, 1973. [CrossRef]
46. Szanyi, J.; Abdulhaq, H.A.; Hegyi, R.; Gál, T.; Szabó, É.; Lossos, L.; Tóth, E. Assessment of Changes in Groundwater Resources Due to Climate Change for the Purpose of Sustainable Water Management in Hungary. *Water* **2026**, *18*, 724. [CrossRef]

47. Guo, W.; Langevin, C.D. User's Guide to SEAWAT: A Computer Program for Simulation of Three-Dimensional Variable-Density Ground-Water Flow. In *Techniques of Water-Resources Investigations Book 6*; U.S. Geological Survey: Denver, CO, USA, 2002; Chapter A7, 77p. [CrossRef]
48. Harbaugh, A.W. *MODFLOW-2005: The U.S. Geological Survey Modular Ground-Water Model—The Ground-Water Flow Process*; Techniques and Methods 6-A16; U.S. Geological Survey: Reston, VA, USA, 2005. [CrossRef]

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