Earth Science Digital Twins: A Systematic Literature Review

Jasper Albert Adriaan van den Barg
Maris B.V.
Nootdorp, The Netherlands
jaspervandenbarg@hotmail.nl

Victoria Degeler
Informatics Institute, University of Amsterdam
Amsterdam, The Netherlands
v.o.degeler@uva.nl

***Abstract*—While Digital Twins are traditionally associated with domains of human-made systems, Earth Science Digital Twins (ESDTs) have been rising in popularity. An ESDT represents a twin of a natural entity or ecosystem at a local, regional, or global scale, such as a river, a forest, an ocean, or the Earth itself. In this work, we conduct a systematic literature review of the existing ESDT systems. We identify the main ESDT domains as hydrology, geology, ecology, Earth, and atmosphere. We analyse the purposes of ESDTs, describe the software development tools that are often employed in ESDT development, and discuss the challenges and limitations that are specific to ESDTs. We conclude that data collection and computational limitations are among the most prominent, and maintaining the real-time mirror of the physical entity remains a challenge.**

***Index Terms*—digital twin, earth science, earth system, virtual mirror model, cyber-physical systems**

I. Introduction

The DT concept has been rising in popularity, reaching a market size of $18.5 billion in 2023 and is expected to grow to $156.7 billion by 2032 [62]. There have been multiple studies surveying the usage of DTs, looking at their main characteristics [63], implementation challenges [64], design methods [65, 66], industry applications [67], and the usage of ontologies in DTs [68]. However, these studies focus on domains with artificial (human-made) physical twins (PTs), like manufacturing, industrial, automotive, or smart city DTs.

More organisations are showing interest in the use of DT concepts outside of the artificial domain mentioned above, tackling topics such as climate change, weather prediction, or pollution [69–71]. These DTs used for Earth science (ES) focus on domains with natural PTs that are often at very large scales. This means they face different challenges and need other technological solutions than the typical DTs of human-made systems. Even though there is a growing interest in using DTs for ES, the holistic overview of the effort in this direction is limited, which hinders sharing of scientific knowledge, collaboration, and development of scientific research on ESDTs. Maimour *et al.* [72] present a relevant survey, but they focus specifically on the networking perspective of DTs, discussing the challenges of deploying communication networks in the natural environments.

Research performed as part of the Master's Thesis Software Engineering at the University of Amsterdam. V.Degeler is supported by the Dutch Research Council (NWO), in the scope of the DiTEC project (NWO 19454).

In this work, we provide a holistic view of the current state of DTs in ES by looking at what ES domains employ DTs, which DT definitions are used in ES, the main purposes of using DTs in ES, ESDT presentation methods, software development tools, challenges and limitations, and future directions. Our research makes the following contributions: (1) an overview of existing knowledge and findings of the current literature on ESDTs; (2) synthesis of collected data, offering new insights into ESDT research; (3) future directions for DT research within ES; (4) a DT definition tailored to ES.

II. Background

Grieves and Vickers [73] gave a formal definition of the DT as "a set of virtual information constructs that fully describes a potential or actual physical manufactured product from the micro-atomic level to the macro geometrical level. At its optimum, any information that could be obtained from inspecting a physical manufactured product can be obtained from its Digital Twin." *Earth Science* differentiates itself from other DT domains in its origin of existence. Where most common DT domains, such as manufacturing, industry, smart city, etc., deal with artificial PTs such as cars, aeroplanes, cities, or processes, ESDTs deal with PTs that are natural, non-man-made, not clearly defined, and/or measurable. We define an ESDT as *the twinning of natural physical entities that lack a (clearly) defined model or blueprint, for running simulations or observations using real-time and historical data at a local, regional or global scale*. Local scales focus on small, specific areas such as river banks or mountain slopes [74]. Regional scales focus on larger areas such as river basins, forests or the atmosphere in Europe. Global scales focus on the entire scale of the Earth, such as global sea levels or atmospheric variables. An example of an ESDT would be a DT of the North Sea, Alpine glaciers, or the atmosphere of North America.

III. Methodology

We used Google Scholar to identify primary research. We designed a broad list of keywords related to ES. These were combined into groups and used for designing the search queries as shown in Table I. Column 2 "Pages" refers to how many pages were consulted to identify research. Browsing was stopped when no more relevant results were returned.

TABLE I
SEARCH QUERIES AND BROWSED PAGES

Query	Pages
(∼earth | ∼planet | ”planet earth”) AND ”digital twin”	1-6
earth AND ”digital twin”	1-6
”earth science” AND ”digital twin”	1-6
(∼climate | ∼environment | ∼weather) AND ”digital twin”	1-6
∼climate AND ”digital twin”	1-6
(∼ocean | ∼sea) AND ”digital twin”	1-6
(maritime | marine) AND ”digital twin”	1-6
(∼natural | ∼nature) AND ”digital twin”	1-6
(∼eco | ∼ecological | ∼ecosystem) AND ”digital twin”	1-6
(∼shoreline | ∼coastal | ∼beach) AND ”digital twin”	1-6
(∼river | ∼lake) AND ”digital twin”	1-8
(∼mountain | ∼landslide | ∼volcanic | ∼valley) AND ”digital twin”	1-6
(∼hurricane | ∼tornado | ∼storm | ∼cyclone | ∼atmosphere) AND ”digital twin”	1-12
(∼disaster | ∼hazard) AND ”digital twin”	1-6
(∼desert | ∼forest | ∼grassland | ∼tundra | ∼aquatic) AND ”digital twin”	1-11
(∼flood | ∼tsunami | ∼earthquake | ∼tectonic | ∼geological) AND ”digital twin”	1-9
(∼glacier | ∼wildfire) AND ”digital twin”	1-6

To filter out irrelevant results, we designed a set of inclusion and exclusion criteria, shown in Table II, which we applied to the browsing results, titles, and abstracts of the identified research resulting from the search query. The inclusion criteria ensure the relevancy of the selected research, and the exclusion criteria ensure the quality and fairness of the selection.

Applying these inclusion and exclusion criteria to search results yielded 288 primary studies, reduced to 158 unique studies after removing duplicates. This preliminary selection still contained studies, irrelevant to the goals of our literature review. To remove them, a second set of inclusion and exclusion criteria for the subject was applied, shown in Table III. To ensure the selected literature is sufficiently developed to address our research goals, we designed our inclusion criteria to be indicative of such developments. The exclusion criteria are based on the definition of the subject and the context in which the research was conducted. To be included, a study must conform to at least one of the inclusion criteria and none of the exclusion criteria. During this process, four more sets of duplicates were discovered, and 30 references were found that should have been excluded based on the first set of exclusion criteria. This left us with 47 included, 52 excluded, and 27 doubtful studies. Doubtful studies were analysed a second time and included or excluded based on the researchers' expert opinion, resulting in 61 included studies.

TABLE II
1ST SET OF INCLUSION AND EXCLUSION CRITERIA

Inclusion Criteria
Peer-reviewed studies
The browse result reflects the goal of the search query
The abstract mentions the goal of the research and is relevant to the goal of the search query or any of the RQs
Exclusion Criteria
Secondary or tertiary studies
The study is in any language other than English
The study is from before 2018 or after May 2024. This time frame has been chosen based on the preliminary research showing significant growth in ESDT research from that period, the fast-growing pace of DT technology and the time at which this research has been conducted

TABLE III
2ND SET OF INCLUSION AND EXCLUSION CRITERIA

Inclusion Criteria
Focuses on an ESDT application
Describes an ESDT framework or development method
Describes software development tools used for ESDT development
Exclusion Criteria
Focuses on artificial DTs, e.g. smart city, mining equipment, agriculture
Scale of PT is too small to be considered an ESDT (anything smaller than local scale)
Focuses on ESDT hardware, but not on software
Do not demonstrate the ESDT actual implementation and use of software development tools (applies to inclusion criteria 3)

To help in analysing and synthesising the data from the primary studies, we created a data extraction form, shown in Table IV. Important information about each primary study in this form should include the study details such as authors, title, date, doi, and the data relevant to our research focus.

We tested the data extraction form on a semi-random set of primary studies to test the quality of the form and the consistency and to minimize the bias of the researchers. This test was deemed successful since a sufficient amount of data was extracted, the extracted data was in proportion to the inclusion criteria on which the research was filtered, and the researchers were able to extract the data consistently. Not every primary study described every item from the data extraction form. We got a final extraction ratio of 0.62 over the total data extraction sheet.

Analysis required additional pre-processing. Different papers use different wording with the same meaning or definition. This can lead to inconsistent data extraction, which decreases the data synthesis accuracy. During the pre-processing, we took all the studies that could have a similar meaning for an item of extracted data and compared their meanings. An example of different wording with similar meanings is *data detail level*, *data density*, or *data scale*, which refer to the spatial and/or temporal resolution of the data.

IV. RESULTS

61 primary studies have been used in this literature review.[1] Figure 1 shows a significant growth in ESDT literature over the last years. In terms of maturity, most ESDTs are still in the conceptual or development stage, with relatively few having reached operational maturity.

A. Earth Science Domains

We have identified five different Earth Science domains that employ ESDTs: hydrology (39.3%), ecology (21.3%), geology

[1]The full data is available at: *https://doi.org/10.5281/zenodo.15707710*

TABLE IV
DATA EXTRACTION FORM WITH AN EXAMPLE RECORD

Item	RQ
Title	Digital Twin of Atmospheric Environment: Sensory Data Fusion for High-Resolution PM2.5 Estimation and Action Policies Recommendation
DIO	10.1109/ACCESS.2023.3236414
BibLater ID	Abutalip2023DigitalRecommendation
URL	https://ieeexplore.ieee.org/document/10015739
Year	2023
Updated	21-06-2024
Use Case	PM2.5 estimation of the South coast air basin (Los Angeles, California), Delhi, Taipei
ES Domain	Atmosphere
ES Subdomain	Pollution
DT Definition	A digital replica of a living or non-living object, structure, environment, or event with a requirement of bidirectional information exchange
DT Purpose	Management, Prediction
Applies ESDT	-
Presentation Type	Data
Software Development Tool	Optuna Hyperparameter Tuning Framework
Challenges / Limitations	Cost, Data communication
Future Trends / Directions	ML, Model accuracy
Summary	-
Other Notes	-

(17%), Earth (13.1%), and atmosphere (8.2%). Figure 2 shows the literature distribution over these domains. The *hydrology* domain employs DTs the most, almost twice as much as the second largest domain, *ecology*, and almost five times as much as the smallest, *atmosphere*. Furthermore, we have identified 26 subdomains within the five major ES domains.

Hydrology domain describes all Earth's water-related processes, including subdomains: rivers, watersheds, river basins, ocean, seafloor, coasts, climate, drought, flood, typhoons, landslides, and aquatic ecology. Pillai *et al.* [1] described a DT for researching the wave-dampening effects of seagrass meadows for natural coastal protection. Qiu *et al.* [2] presented a web-based DT for watershed monitoring, visualisation, and decision support management in the Yangtze River watershed. Allen *et al.* [3] proposed a framework and prototype digital architecture for a DT for analysing and predicting coastal flooding. Wenzheng [4] proposed a concept and framework for digital river basins that provide support for developing digital river basins and digital river basin engineering.

Ecology domain entails the relationships between living organisms and includes subdomains such as forests, bogs, wildfires, climate, and pollution. Silva *et al.* [5] proposed a framework for a forest DT called Digital Amazon for researching greenhouse gases. Kwon *et al.* [6] proposed a distributed computing framework for a wildfire DT. Goleva *et al.* [7] combine real-time data from a measurement device that can be attached to drones and simulated data to predict fire spread in forest fire scenarios. Zhong *et al.* [8] described a DT fire model for improving efficiency in global wildfire prediction using the JULES-INFERNO model, where they use latent data assimilation to prevent error accumulation and adjust the bias of the prediction results. Cirulis *et al.* [9] described a DT that runs real-time simulations for bog or peatland ecosystems.

Geology describes all processes happening in the Earth's crust with subdomains such as earthquakes, volcanoes, landslides, geological structures, glaciers, ground liquefaction, slopes, mountains, and climate. Kusakabe *et al.* [10] presented and tested an earthquake simulation method, considering soil liquefaction, for DTs for high-resolution large-scale simulation. Cheverda *et al.* [11] described a 3D DT with geological faults, fractures, and cavities on which they performed tectonic simulations to validate their image scattering technique.

Earth domain describes the entire Earth or part of it, or at least two of the other four domains. Subdomains found here include: climate, coasts, ocean, the arctic, and regional segmentation. Duque and Brovelli [12] created a web application using the mediator-wrapper architecture as a proof of concept of how data integration can be used to implement DTs; it describes the interaction of land, marine, and demographics on the Italian coast. Sougioultzoglou and Cook [13] investigated the creation of an Arctic DT for overcoming the polar satellite altimetry gap by predicting sea ice concentrations and other polar variables. Ahmadi *et al.* [14] presented a global terrain and altitude map by combining a DT with deep-learning techniques on Florida's coast. The DT accurately resembles real-world locations and serves as both a physical and a simulation model.

Atmosphere focuses on Earth's atmosphere and all processes influencing it. Two subdomains are found in the atmosphere domain: climate and pollution. Zlatev and Dimov [15] used the Unified Danish Eulerian Model to create a DT named Digital Air to study ozone levels in Europe. They simulated different

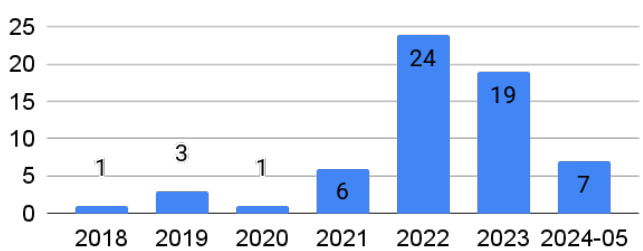

Fig. 1. Distribution of relevant publications over the years

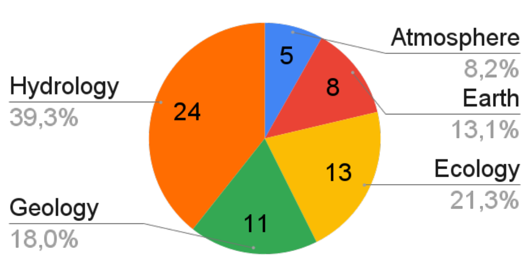

Fig. 2. Earth science domains

scenarios to predict future ozone level concentrations across Europe. Abutalip *et al.* [16] presented an atmospheric DT by fusing remote sensing and observational data. The DT is used for predicting PM2.5 concentrations and suggesting possible action policies and warnings.

B. Digital Twin Definitions Used Within Earth Science

We identified 33 DT definitions in the literature. These definitions varied per ES domain and study. We divided these definitions into five components: the virtual twin, the physical twin, the used components, functionality, and capabilities, as seen in Table V.

Chen *et al.* [17], Yuan *et al.* [18], Zhao *et al.* [19], and Li and Zhang [20] all referred to Dr Grieves's 2002 product lifecycle management presentation, where the first concept of a DT was presented as the Mirrored Spaces Model [75]. Originally developed for manufacturing, this model consists of three elements: real space, virtual space, and a linking mechanism in the form of data and information flow [75]. This can be used as a basic definition of a DT, however, as the DT has since grown in popularity more comprehensive definitions occurred such as NASA's Digital Twin Paradigm referred to by Chen *et al.* [17] and Henriksen *et al.* [21] which defined a DT as "an integrated multiphysics, multiscale, probabilistic simulation of an as-built vehicle or system that uses the best available physical models, sensor updates, fleet history, etc., to mirror the life of its corresponding flying twin" [76]. This definition is specifically tailored to NASA's manufacturing context, but can be generalised to be used for ESDTs "a DT is an integrated multiphysics, multiscale, probabilistic simulation of a physical entity or system that uses the best available physical models, sensor updates, historical data, etc., to mirror the lifecycle of its physical twin".

More specific ESDT definitions can be found. Bauer *et al.* [22], Moigne [23], and Huang *et al.* [24] defined an ESDT as "an information system that exposes the user to a digital replica of the state and temporal evolution of the Earth's system, constrained by observations and the laws of physics". Reshetova *et al.* [25] described an ESDT of a geological object as "a set of data that determines its geometric and physical properties in conjunction with the corresponding synthetic geophysical fields".

More generic ESDT definitions exist, which can be used for a broad range of ESDTs. Yun *et al.* [26] simply defined ESDTs as "virtual replicas of physical systems and natural environments", Buonocore *et al.* [27] defined a DT as "a virtual representation of real-world entities and processes, synchronized at a specified frequency and fidelity", and Fissore *et al.* [28] and Yi [29] who used a definition by El Saddik [77] that defined a DT as "a digital replication of a living or non-living entity with seamless data transmission between the physical and virtual world".

Due to the variability in existing definitions, the scientific community can benefit from a standardised definition that can consistently encompass the full scope of ESDTs. Looking at Table V, based on what most references agree on, we can define an ESDT as "a virtual replica of a physical entity that makes use of data and physics models for monitor, simulation, or prediction and is able to mirror the entire lifecycle of its physical entity". We, however, argue that this definition is not specific enough to distinguish the ES domain from other DT domains. Also, even though most references agree that "the DT should mirror the entire lifecycle of the PT", we argue that this is impossible, as a natural PT existed before the DT and will still be there after. Thus proposes to alter the definition to: *"an Earth science digital twin is a virtual replica of a physical natural entity at a local to global scale that makes use of data and physics models for monitoring, simulation, or prediction"*.

TABLE V
DEFINITIONS USED TO DESCRIBE DIGITAL TWINS IN EARTH SCIENCE

Definition	Reference
Virtual Twin	
An information system	[22–24]
A digital replica of earth	[22–24]
A digital replica of a state and the temporal evolution	[22–24]
A virtual replica	[16, 22–24, 27, 30] [28, 29], [17, 20, 31, 32]
A simulation process	[21, 33]
A comprehensive description of a system	[34, 35]
A virtual model	[18, 36–40]
Realtime	[13, 36–39, 41]
Something else	[25, 30, 37, 42]
Real Twin	
A real-world entity	[21, 27, 31, 43]
A real-world process	[13, 27, 31, 42, 43]
A (non)living object	[13, 16, 28, 29, 40]
An environment	[13, 16, 26, 44]
A physical entity	[13, 17, 18, 26, 36, 38, 41, 43, 45]
A structure or event	[16]
Digital Twin Components	
Data	[16, 21–24, 26–28, 31, 33, 34, 36, 38, 42, 44]
Physics models	[22–24, 26, 33, 38]
Artificial Intelligence (AI) / Machine Learning (ML)	[21, 26, 31]
3D visualization	[13, 20]
Internet of Things (IoT)	[13, 26, 37]
Big Data	[26]
Digital Twin Purpose	
Monitoring	[13, 21, 30]
Understanding	[21, 30]
Prediction	[13, 30, 39]
Simulation	[13, 21, 26, 38]
Updating or testing	[13]
Integrating multi-disciplinary / scale / probability / professional / physics	[33, 38, 42] / [21, 33, 38] / [33] / [42] / [21]
Digital Twin Capability	
Bi-directional information exchange	[16, 17, 46]
Being in an equal state as the physical entity	[13, 18, 20, 27, 28, 37, 39]
Mirror the entire lifecycle of the physical entity	[17, 20, 21, 31, 34, 35, 38, 42]

C. Digital Twin Purposes in Earth Science

DTs are developed for a specific reason or purpose for use in ES. We identified four main purposes: *prediction, simulation, monitoring, and management*. As can be seen in Figure 3, most ESDTs have a prediction or simulation purpose, followed by a monitoring and, lastly, a management purpose. Notably, the monitoring purpose is much more common in the Earth domain than in the other ones, while simulation and prediction are more common in geology and ecology domains. An ESDT is not limited to a singular purpose but can have a combination of multiple purposes.

An ESDT with a *management* purpose is able to autonomously respond to scenarios by executing actions or offering suggestions to the user. This replaces the human factor, in part or in full, in the process of complex data analysis and scenario-based decision making, greatly reducing the response time, which can often be the difference between life and death in hazardous situations. Ahmadi *et al.* [14] described an ESDT of the coast of Florida that segmented the terrain into seven biomes, illustrating how the use of DTs for terrain generation can be used for gaining insights into environmental and urban planning. Zhang *et al.* [33] described an ESDT for mountain geological disaster monitoring and prediction. With the use of the Internet of Things, the DT is able to send out early warnings to civilians and base stations.

ESDTs with a *monitoring* purpose allows to efficiently monitor the PT remotely through the DT, as the DT mirrors the PT in real-time. This is especially useful for larger and inhospitable monitored environments and can be beneficial for risk monitoring at large scales, such as the Duhe River basin, where Rao *et al.* [42] described an ESDT for flood monitoring and prediction, flood control dispatching, and decision comparison. Fissore *et al.* [28] proposed an ESDT for monitoring glacier aspects like distribution, velocity, temperature, and structure, to analyse how glaciers respond to climate change.

A *simulation* is an imitation of a real-world process or system that can alter the state of the simulation with the help of models. The purpose of simulation ESDTs is to experiment with different what-if scenarios and to analyse the DT output.

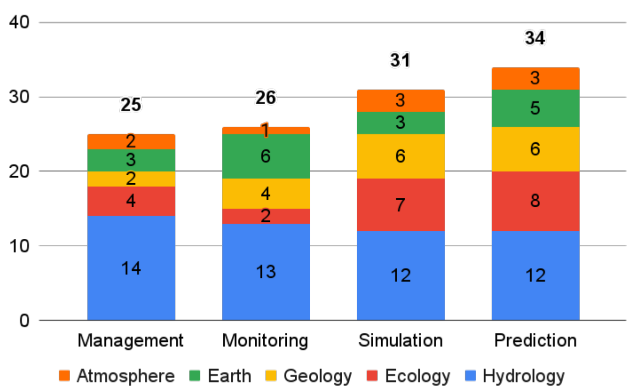

Fig. 3. Earth science digital twin purposes

In the atmosphere domain various simulated climate scenarios can be applied to, for example, analyse air pollution hazards like Todorov and Dimov [47] did with ESDT DIGITAL AIR, or in the hydrology domain where Allen *et al.* [48] described an ESDT for both real-time monitoring and disaster response for coastal flooding, and simulating flooding scenarios.

Predictions are, in essence, simulations of real-life scenarios advanced in time. The purpose of predictive ESDTs is to take early measures in anticipation of a predicted outcome. Such ESDT can be used for predicting geological environments during tunnel construction, e.g. Wu *et al.* [44] used a voxel model and volumetric b+trees to do so. Kwon *et al.* [6] described an ESDT to predict the spread of fire during wildfires by using foliage fuel models, landscape and weather data.

Table VI shows a relation matrix of ESDT purposes. Although management was the least used purpose, it is often combined with other purposes, e.g. with prediction and monitoring, both capabilities that are needed for real-time and preventive management. The simulation purpose is the least frequently paired with other purposes.

D. Real-Life Applications of Digital Twins in Earth Science

The majority of primary studies were still at the conceptual or development stage of their ESDT, with only seven clearly defined examples of real-world applications.

Zlatev and Dimov [15] looked at ozone concentrations in Europe based on five climatic scenarios over a period of 16 years (1989-2004). The first scenario was based on real meteorological and emission data, the other four were hypothetical scenarios based on IPCC reports. For some regions, the yearly maximum was exceeded in the real scenario.

Pillai *et al.* [1] used an ESDT to establish the wave, sea level and current attenuation due to seagrass nature-based solutions. They simulate different seagrass types and landscape designs for coastal protection. They analysed the wave height over a ten-year period (2010-2019) and saw a significant decrease in maximum wave height in the vegetated state.

Cheverda *et al.* [11] described a geological ESDT with faults, fracture corridors, and cavities. They simulate seismic waves to validate their scattered waves imaging approach.

Zhang *et al.* [33] described an ESDT used for warning residents and authorities through the internet and base stations by using IoT sensors and neural networks for predicting geological disasters.

TABLE VI
EARTH SCIENCE DIGITAL TWIN PURPOSE RELATION MATRIX

	Prediction	Simulation	Monitoring	Management
Management	15	9	15	25
Monitoring	12	10	26	
Simulation	12	31		
Prediction	33			

Henriksen *et al.* [21] showed the Hydrological Information and prediction (HIP) system and the DT has a number of (potential) users in Denmark, including cities, energy sector, emergency preparedness and fire and rescue services, insurance sector, and spatial planning.

Wang *et al.* [38] described an ESDT of the Weihe River Basin to address its problems in information, intelligence, and digitisation, that calculates in real-time and predicts rainfall, runoff, and water information, and can make real-time adjustments to the response measures for water conservancy project operation, emergency dispatch and personnel disaster prevention and avoidance.

Abutalip *et al.* [16] described an ESDT of the atmospheric environment of Los Angeles, Delhi, and Taipei, using 1km resolution satellite data, to provide recommendations for decreasing PM2.5 to healthier levels.

E. Ways of Presenting Earth Science Digital Twins

25 primary studies stated how their DT is presented to the user and is categorised into 6 different presentation forms: 2D, 3D, augmented reality (AR), virtual reality (VR), primary data, and graphical representation, as seen in Figure 4. Several presentation form can be used for a single DT.

2D format is a two-dimensional visual presentation. Often it visually resembles the subject. For example, a 2D representation can consist of a map interface displaying coastal regions [12] or an atmospheric overlay [15, 47, 49].

3D format adds a third-dimension depth to the 2D format. Not only does this depth provide another sense of scale, but it also allows the user to observe the subject from different angles that would not be possible in 2D. Examples are a simulation of a coastal location for user awareness [18] or a 3D grid of a geological environment [11, 25, 44].

VR enables full immersion in a virtual environment using a VR headset. Cirulis *et al.* [9] described how VR is used for the simulation of scenarios in a bog environment by cutting down vegetation using a virtual chainsaw.

AR mixes a digital overlay with the real environment. This can be done through mobile devices such as cellphones or tablets, or with AR headsets or glasses. Rao *et al.* [42] described the use of AR for the interaction and management portal of the Duhe river basin, and DTO uses AR, among others, for visualising different scenarios [17].

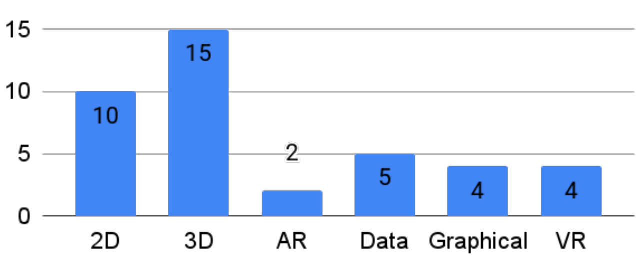

Fig. 4. Different ways in which digital twins are presented

Data can be produced by DTs through simulation, modelling, or prediction. This is often used for analysis or visualisation. Song and Jang [50] presented a DT that not only provides 3D visualisation but also offers analysis data used for other applications. Henriksen *et al.* [21] allow for API access and direct downloading of GIS files from their hydrological information and prediction system.

Graphical representation consists of graphs, such as line or bar graphs, charts, such as pie charts, and plots, such as line and box plots. Zhao *et al.* [19] describe a dashboard with different interactive graphical presentations like bar and line graphs, pie charts, and scatter plots on which the user can click/highlight data points.

F. Software Development Tools Used in Earth Science Digital Twin Development

The virtual part of a DT is a software product made from existing, modified, or built from scratch components with the help of software development tools (SDTs). We identified 74 unique SDTs, mentioned by 27 primary studies, that we grouped into seven categories as seen in Table VII.

Visual modelling, design, and simulation SDTs (mentioned in the domains: Earth, Hydrology, Ecology, and Geology) are SDTs used for modelling, designing, or simulating parts of the DT and can be used across all domains. Unity3D and Unreal-Engine are both real-time rendering applications used for gaming, simulation, or other interactive applications. 3ds Max, Revit, Dynamo, Blender, CATIA, MTree, and OpenCascade are all SDTs used for 3D modelling of objects or environments. These 3D models can be used by other software, such as Unity3D or Unreal-Engine, for visualisation, simulation, analysis etc. Autodesk and Bentley are both software companies that provide professional products for architecture, engineering, infrastructure, construction, product design, and manufacturing. 3ds Max, Revit and Dynamo are part of the Autodesk product set. Amplify shader editor and OpenVDB are visual effects tools. Finally, Insight Maker lets you create diagram models for system dynamics or agents.

Programming frameworks (Earth, Hydrology, Ecology, Geology, and Atmosphere) is a category that contains: Integrated Development Environments (IDE) that provide an interface for writing, debugging, and managing code. Visual Studio is an example of an IDE that supports multiple programming languages; Development tool suites, like Intel Parallel Studio XE, offer collection of compilers, performance libraries, and analysis tools for high-performance scientific computing; Development frameworks like WebGL, JavaEE, Django, Optuna, and Angular, that provide developers with tools, libraries, and conventions to help them develop more efficiently; And Software Development Kits (SDKs) for VR development like BNG VR SDK which is an SDK for VR user interaction, and Oculus SDK which is needed for deploying VR applications the the Oculus Quest.

Hydrodynamic modelling & simulation SDTs (Hydrology) consist of advanced models for simulating and predicting the movement, distribution, and quality of Earth's water at

TABLE VII
SOFTWARE DEVELOPMENT TOOLS USED IN EARTH SCIENCE DIGITAL TWIN DEVELOPMENT

Visual Modelling, Design, Simulation	**Scientific Computing & Data Analysis**
AutoDesk [42]	AWS [48]
3ds Max [2]	Docker [49]
Amplify [9]	FENIX [49]
Bentley [42]	ODC Jupyter [48]
Blender [19]	MongoDB [42]
CATIA [42]	MySQL [42]
OpenVDB [44]	Google Colab [8]
Dynamo [51]	PostGIS [12]
Revit [51]	eFlows4HPC [49]
MTree [9]	PDAF [49]
Insight Maker [9]	EuroHPC [49]
Open Cascade [44]	SDAP [52]
Unity3D [9, 19, 53]	PyCOMPSs [10]
Unreal-Engine [18, 32, 53]	PostgreSQL [12, 42, 48]
Hydrodynamic Modelling & Simulation	**Geographic Information Systems (GIS) & Geospatial Tools**
ADCIRC [54]	GeoDjango [12]
DELFT3D [3]	GeoSLAM [32]
FES2014 [46]	QGIS [19]
MIKE HYDRO [21]	Google Earth [14]
MIKE SHE [21]	OpenLayers [12]
Hydronet [55]	ArcGIS [45, 48]
SOCONT [43]	Terria.io [48]
SCHISM [46]	Cesium [53]
RAPID [52]	Surfer [50]
RS Minerve [43]	OpenStreetMap [12, 19]
CONTINUUM [56]	
Nasa LIS [52]	
MRI.COM-JPN [46]	
GLEAM [56]	
Programming Frameworks	**Physics Models**
Angular [12, 42]	Pro Shake [50]
Visual Studio 2012 [50]	MSM JMA [46]
Oculus SDK [9]	IMERG [56]
BNG VR SDK [9]	FALL3D [49]
WebGL [2, 42]	POWER [52]
JavaEE [42]	S3M [56]
Django [12]	TMVOCBio [34]
Optuna [16]	ERA5 [46]
Intel Parallel Studio XE 2019 [11]	SM2RAIN-ASCAT [56]
Charting Tools	
ECharts [19, 53]	
XCharts [9]	
Grapher [50]	
Stella Architect [9]	

various spatial and temporal scales. They can be split up in: oceanic models like ADCIRC, Delft3D, FES2014, SCHISM, and MRI.COM-JPN, that are capable of simulating and predicting marine and coastal systems; And watershed models that simulate terrestrial water dynamics like rainfall-runoff, groundwater flow, soil moisture dynamics, river discharge, and evaporation dynamics, and include the SDTs: MIKE HYDRO, MIKE SHE, Nasa LIS, Hydronet, SOCONT, RS Minerve, CONTINUUM, RAPID, and GLEAM.

Scientific computing & data analysis SDTs (Earth, Hydrology, Ecology, and Geology) provide ESDTS with the computational infrastructure and data management systems for processing, storing, and managing large amounts of data. They consist of: cloud infrastructure, such as AWS, Google Colab, EuroHPC, and FENIX, that provide scalable solutions for computation, data storage and management; Execution platforms, such as Docker, that enable deployment and scalable execution of applications and workflows. Database software, such as PostgreSQL, MySQL, PostGIS, and MongoDB, for storing, processing, and managing data; Scientific workflow and programming environments, like ODC Jupyter, eFlows4HPC, and PyCOMPS, integrate models, data processing and analytics, and help build and run workflows across systems; And analytical frameworks such as PDAF and SDAP, that process and analyse data.

Geographic information systems (GIS) & geospatial tools (Earth, Hydrology, Ecology, and Geology) provide collection, processing, analysis, and visualisation of spatial data. These SDTs come in the form of Web frameworks, like GeoDjango, OpenLayers, Terria.io, Cesium, and ArcGIS (web), that provide libraries and platforms for developing interactive web-based geospatial applications with tools; Geospatial tools for data preparation, spatial analysis, geospatial modelling, and interpretation of geospatial data, such as QGIS, Surfer, Google Earth, and ArcGIS (desktop); Tilling services that provide pre-rendered or dynamic map tiles that can act as contextual layers for geospatial applications or in the case of OpenStreetMap provide open source basemap tiles; And visualisation tools like GeoSLAM that consists of both hardware capable of mapping terrain and environments using lidar, and software to process the measurement data.

Charting SDTs (Hydrology, Ecology, and Geology) are SDTs that help the user visualise their data in a graphical format, such as charts, plots, heatmaps, and other types of visualisations. ECharts (JavaScript) and XCharts (Unity3D) are both 2D and 3D visualisation libraries that help develop graphical visualisations for their targeted platform. Grapher is a desktop application that is capable of a multitude of visualisation formats in both 2D and 3D. Stella Architect is a desktop application capable of 2D modelling and charting.

Physics models (Hydrology and Geology) simulate and predict natural processes like precipitation, snow and ice dynamics, biochemistry, and seismic processes. This SDT category contains: Atmospheric and precipitation models that provide atmospheric data such as temperature, humidity, solar radiation, and weather precipitation from models like MSM JMA, IMERG, SM2RAIN-ASCAT, ERA5, and POWER; Cryospheric models simulate snow and ice dynamics, S3M specifically simulates accumulation, melting, and movement of snow and ice, and can reconstruct snow and glacier evolution; TMVOCBio is a finite volume numerical simulator for multi-phase, multi-component, multi-microbe transport phenomena at a Darcy scale; ProShake is an application for one-dimensional, equivalent linear ground response analysis.

Figure 5 shows the number of SDTs mentioned across different ES domains. Even though the hydrology domain makes up 39.3%, it mentions 59.5% of the SDTs. The atmosphere domain also stands out, mentioning only 1 SDT (1.4%). The remaining domains, geology (21.6%), ecology (14.9%), and Earth (12.2%), show a more balanced relationship between

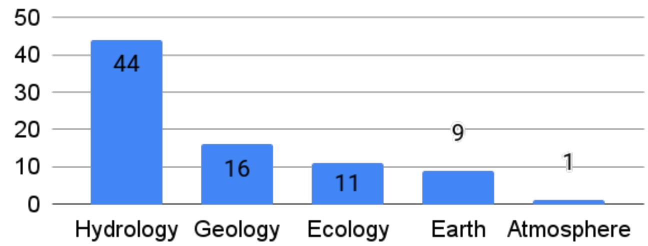

Fig. 5. Number of software mentioned per Earth science domain

their share of domain representation and the SDT proportion.

G. Challenges and Limitations Faced by Earth Science Digital Twins

ESDTs face various challenges and limitations, which we will refer to as just *challenges*. We have identified 21 different challenges that are mentioned a total of 48 times in the analysed research. We categorised these challenges into six different categories as seen in Table VIII.

The most frequently encountered challenges are *data*-related. These data challenges have to do with accuracy, how the data correctly and consistently represents real-world values or events it intends to model; coverage, spatial and temporal gaps in the data; resolution, spatial and temporal granularity of the data; and standardisation. Qiu *et al.* [58] proposed an ESDT for monitoring harmful algae blooms in Lake Chaoho, China, using real-time video monitoring, but were limited by the range, angle, and number of measurement devices, only providing data for a limited number of near-shore locations of the lake. Grossmann *et al.* [57] required a higher spatial resolution of seabed measurement for the ESDT than was currently possible with measurement devices. Whereas Fissore *et al.* [28] made use of satellites for Earth observation (EO) and were limited by how often EO satellites were able to make a successful measurement. Finally, Grossmann *et al.* [57] and Tzachor *et al.* [40] both had issues standardisation of data.

The second biggest category of challenges is related to *computational* problems. Most of these challenges are linked to data, the most predominant one is having to process enormous amounts of data [1, 13, 15, 17, 33], processing complex data [2], and memory consumption [44]. Running large and complex models and algorithms efficiently and effectively for the simulation and prediction capabilities of ESDTs [15, 53], together with the demanding visualisation of rendering virtual environments [53], are performance-related challenges faced in ESDT development.

TABLE VIII
EARTH SCIENCE DIGITAL TWIN CHALLENGES AND LIMITATIONS

Data	**Computation**
Data standardisation [40, 57]	Performance [2, 15, 53]
Data accuracy [14, 56, 57]	Data processing [28]
Data coverage [20, 35, 58]	Memory consumption [44]
Data resolution [28, 56, 57, 59]	Data amount [1, 13, 15, 17, 33]
Infrastructure	**Human Factor**
Automation [2, 41]	Cost [13, 16, 40]
Framework implementation [27]	Governance [13, 27, 40]
Cyber security [13]	Knowledge [13, 40]
Data communication [13, 16, 20, 56]	Data visualisation [53]
Model	**Environment**
Model accuracy [14, 33, 56]	Environment [33, 36]
Model consistency [56]	Adaptability [41]
Model coupling [1, 57]	

Model-related challenges are found in three forms: the accuracy of model output when compared to the real-world measurements, such as poor weather forecasting, and uncertainties therein [14, 33, 56]; the consistency of the model output in delivering accurate results [56]; and the lack of model coupling, which complicates integration with complex systems and subsystems [1, 57].

Infrastructure-related challenges mainly have to do with data communication issues such as bandwidth when communicating large amounts of data, and latency and coverage of measurement devices [13, 16, 20, 56]. Another challenge is the automation of reconstructing the PT in a virtual environment [2] and modifying tasks and relationships with other systems [41]. A lack of framework implementation in real situations also hinders the ability the validate the framework [27]. The last challenge is handling sensitive data and data vulnerability to cyber-threats [13].

Human factors also play a role in the development and research of digital twins, with challenges like: the cost in terms of resources, human capital, and time [13, 16, 40]; limited knowledge of both DT and PT aspects [13, 40]; governance obstacles such as ethical concerns about privacy and consent of inhabitants, or legal obstacles such as laws, regulations, or legal barriers [13, 27, 40]; and finding the right type of data visualisation method [53].

A challenge of twinning a natural PT, especially at a larger scale, is dealing with the PTs *environment*. So were Zhang *et al.* [33] and Liu *et al.* [36] challenged by the complexity, randomness, and regional differences in the environment, and Zohdi [41] faced issues with adapting to rapid changes in the environment during wildfires.

Figure 6 shows a distinct relationship pattern between ESDT domains and challenges. Hydrology not only stands as the most represented domain but also as the one facing the biggest number of challenges, equal to the combined total of all other domains. It is particularly dominant in the data and model category, where it accounts for more than double as many challenges as the other domains combined. The other domains face a similar number of challenges compared to each other, although there is a difference in the distribution per category.

Figure 7 shows the different ESDT purposes in relation to challenges. The management purpose, although the smallest category, deals with the most challenges, and the simulation purpose, despite being the second largest, deals with the fewest. The management purpose is often combined with other purposes and is highly dependent on them, requiring both monitoring and simulation or prediction capabilities, making it inherently more complex. It also encounters far more challenges in infrastructure and human factors compared to other purposes. In contrast, the simulation purpose is most frequently implemented as a stand-alone solution, requiring neither real-time data from monitoring systems nor integration with management functionalities, which makes it simpler to realise, often missing any human factor-related challenges.

Table IX presents a matrix illustrating how frequently different ESDT challenge categories are referenced together. Data challenges are often paired with model challenges, and infrastructure challenges are often paired with human factors. Challenges with data influence model-related challenges in two ways: first, any model relying on problematic data will face similar issues [14, 56]; second, in cases of severe data-related issues, e.g. poor-quality or missing data, DTs rely more heavily on models, thus making any model weaknesses more apparent. The relation between infrastructure and human factors can be attributed to challenges related to implementation costs [13, 16] and governance constraints [27], which are both influenced by human and organisational elements. In contrast, the environment category shows no overlap with data and human factors; and model category also shows no overlap with human factor-related challenges.

V. Research Trends and Directions

The development of ESDTs remains an emerging area of research, with a range of domain-specific challenges and limitations. To support future progress in this field, we identify key research directions that warrant further attention to drive the future evolution of ESDTs.

Data-related challenges and limitations are most prominently mentioned in the literature. On the one hand, the large amount of data that needs to be stored, managed, and analysed is a significant challenge faced in ESDT research. On the other hand, there is still a clear need for even more diverse data sources [12, 14, 22, 47], and future data collection missions [56, 58]. Implementing and responding to real-time data is what allows the DT to mirror the PT in real-time, which is necessary for monitoring [14, 17, 19] and making accurate predictions [8, 44]. To aid in efficiently processing large amounts of data, the scientific community needs to look into solutions for lossless data compression [44] and automatic data management [12]. Data must be of sufficient accuracy [39], spatial and temporal resolution and coverage [22, 56, 60]. Together with data assimilation, this can correct errors in models [46] and improve simulation results [10, 22, 44]. Finally, data standardisation is needed for fast and easy implementation of data from different sources [3, 27, 40].

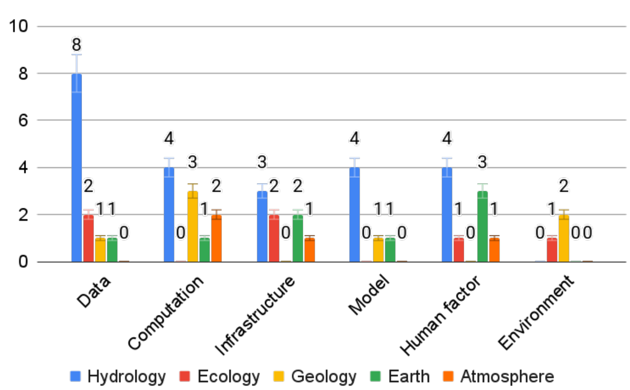

Fig. 6. Challenges and limitations in relation to ESDT domains

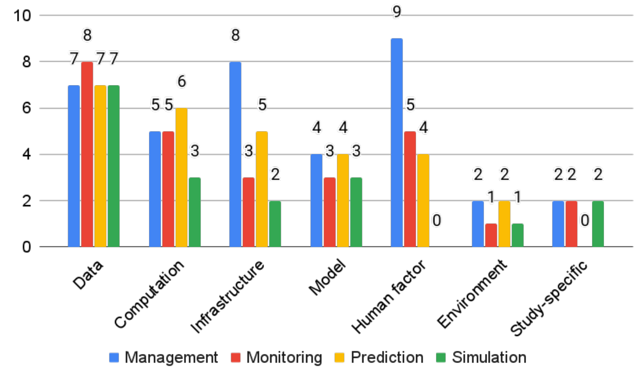

Fig. 7. Challenges and limitations in relation to ESDT purposes

Models are continuously evolving to best simulate and predict scenarios, and interpolate where data is missing. The next step in this evolution is improving the accuracy and efficiency of models [14, 16, 34]. The most widely discussed approach for achieving this is the integration of machine learning (ML) [14, 16, 61] and AI [5] with mechanistic models. ML specifically refers to algorithms that learn patterns from data to make predictions and decisions without being explicitly programmed. AI, more generally speaking, involves mimicking human functions like learning and problem-solving. Developing robust methods to couple models by aligning spatial and temporal scales, resolving feedback mechanisms and ensuring interoperability can enhance the reliability and usability across domains [1, 34, 60].

Computational processes should increase their performance and efficiency [2, 9, 14, 36, 53]. This remains an ongoing issue as both data and models become more abundant and complex. One promising direction to manage this challenge is the use

TABLE IX
ESDT challenges and limitations relation matrix

	Model	Environment	Human factor	Infrastructure	Computation	Data
Data	3	0	1	2	1	9
Computation	2	1	2	2	9	
Infrastructure	1	1	3	7		
Human Factor	0	0	5			
Environment	1	3				
Model	5					

of cloud computing [22, 39], which offers scalable resources and flexible infrastructure for handling large amounts of data and demanding computational models. Find out how to best integrate these with ESDTs can significantly enhance the responsiveness, scalability, and accessibility of ESDTs.

Lastly, because many ESDTs are still in the conceptual or development stage, the next step for ESDTs is the *implementation & application* [15, 17, 26, 48]. This is necessary to validate the effectiveness of ESDT in real-life scenarios, and confirm the accuracy of their monitoring, simulation, and predictive capabilities in real conditions [10, 47].

VI. Discussion and Conclusions

In this work, we conducted a systematic literature review (SLR) on digital twins used in earth science, based on 61 primary studies, to look into the current state of ESDTs, state-of-the-art technologies, and current challenges and limitations.

ESDTs are used in five different domains: hydrology, ecology, geology, Earth, and atmosphere. There is a big disparity in ESDT distribution across these domains, e.g. hydrology emerged as the most represented domain (39.3%), while atmosphere was the least represented one (8.2%). A potential explanation of this might be the technological maturity and adaptation of SDTs used for developing ESDTs. The availability of data for each domain, in amount, parameters, and accessibility, might be another reason. However, the investigation of the validity of this conjecture is open for future research.

We deconstructed DT definitions found in primary studies into five elements: virtual twin, physical twin, components, purpose, and capability. We used the most prominent elements to formulate a new definition for ESDTs. We highlighted two key elements that most clearly separate ESDTs from the typical DT. First, ESDTs focus on natural physical entities, such as the atmosphere, oceans, land surfaces, and ecosystems, whereas typical DTs are often concerned with artificial physical entities like machines, infrastructure, or industrial systems. Second, ESDTs represent a distinct subclass within the broader domain of natural DTs, as they are specifically focused on PTs at larger scales, capturing complex, multiscale interactions that are typically not found or insignificant in smaller PTs.

We discussed four main purposes of ESDTs: management, monitoring, simulation, and prediction. Management, which is the least frequently used purpose, is most often combines with other purposes, especially prediction and monitoring. Simulation, on the other hand, is the rarely combined with other purposes due to its low capability requirements.

ESDTs are presented in six different formats: 2D, 3D, AR, VR, data, and graphical. However, it is not clearly discussed in the literature, what some of the benefits or drawbacks are of one format over the other. For example, what is the added value of 3D over 2D in hydrological ESDTs, or if is it better to show spatial-temporal data in a graphical, 2D, or 3D format.

We identified 74 different SDTs which we categorised into seven different groups based on their functionality. These SDTs provide a good starting point for the development of ESDTs. In particular, tools such as Unity3D, Unreal-Engine, Angular, WebGL, PostgreSQL, OpenStreetMap, and ECharts have been highlighted across multiple studies for their relevance and utility in supporting visualisation, programming, and data management. Some SDTs were only mentioned once, but were thoroughly examined by the respective primary studies and were considered fit for purpose.

We looked at current ESDT challenges and limitations. In total, we identified 25 types of challenges that we grouped into 7 categories: data, computation, study-specific, infrastructure, human factor, environment, and model. Most challenges currently faced are data and computation-related. ESDTs with a management purpose face the most challenges overall. However, some primary studies lacked explanation of their challenges and limitations, leaving open questions, such as at what point data resolution becomes a limiting factor, what is the desired performance of the system, and if governing issues present a limitation. These challenges and limitations should be carefully considered in future research.

This work has some limitations. The SLR only includes primary studies written in English, which introduces a potential language bias and can lead to an over-representation of perspectives from English-speaking research communities, potentially narrowing the diversity of methodologies, use cases, and technologies. Also, despite the broad range of used keywords and search queries, relevant studies may have been missed in this SLR due to variations in terminology, index inconsistencies, or their absence in the Google Scholar database. Especially since the domain of ESDTs is still so young and constantly evolving, researchers may use overlapping terms such as virtual replica/model, cyber-physical system, or smart simulation, which are not captured in our search queries.

Several opportunities for future expansion of this SLR remain. First, a notable disparity exists in the number of primary studies across the identified domains. Future work should aim to investigate the underlying reasons for this imbalance. This may include factors such as differences in funding, data availability, technological maturity, or domain-specific challenges that influence research activity and adoption. Second, as many of the ESDTs identified are still in the conceptual or laboratory phases, it would be valuable to revisit this review in the future, once more ESDTs reach the operational stage. A follow-up study could provide insights into how these systems mature, how they are deployed in practice, and deliver value to stakeholders. Third, an important aspect that was not addressed in this study is how ESDTs leverage different types of data sources, including publicly available datasets (e.g., from governmental or scientific institutions), commercial data (e.g., from satellite providers or private sensors), and self-generated or project-specific data. It is essential to explore how ESDTs access, integrate, and manage these sources, and map future projects and missions for data measurements and collection. Lastly, a dedicated review focusing on the development and standardisation of ESDT frameworks could help identify common architectural patterns, interoperability practices, and gaps in current approaches, thereby supporting more robust and scalable implementations.

Bibliography - Primary Studies

[1] U. P. A. Pillai, N. Pinardi, J. Alessandri, I. Federico, S. Causio, S. Unguendoli, A. Valentini, and J. Staneva, "A Digital Twin modelling framework for the assessment of seagrass Nature Based Solutions against storm surges," *Science of The Total Environment*, vol. 847, p. 157 603, Nov. 2022.

[2] Y. Qiu, H. Duan, H. Xie, X. Ding, and Y. Jiao, "Design and development of a web-based interactive twin platform for watershed management," *Transactions in GIS*, vol. 26, no. 3, pp. 1299–1317, May 2022.

[3] T. R. Allen, G. McLeod, H. Richter, and A. Nielsen, "Digitally Twinning Coastal Resilience Via Multsensor Imagery, in Situ Sensors, and Geospatial Analysis," *International Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 2022-July, pp. 4739–4742, 2022.

[4] L. Wenzheng, "Digital Twin Riverbasin and Digital Riverbasin Engineering," *ICEIEC 2022 - Proceedings of 2022 IEEE 12th International Conference on Electronics Information and Emergency Communication*, pp. 107–110, 2022.

[5] J. R. Silva, P. Artaxo, and E. Vital, "Forest Digital Twin: A Digital Transformation Approach for Monitoring Greenhouse Gas Emissions," *Polytechnica 2023 6:1*, vol. 6, no. 1, pp. 1–9, Jul. 2023.

[6] J. W. Kwon, S. J. Yun, and W. T. Kim, "A Semantic Data-Based Distributed Computing Framework to Accelerate Digital Twin Services for Large-Scale Disasters," *Sensors 2022, Vol. 22, Page 6749*, vol. 22, no. 18, p. 6749, Sep. 2022.

[7] R. Goleva, A. Savov, V. Tomanov, V. Monov, Z. Koleva, R. Sokullu, H. Kostadinova, S. Mihaylov, and N. Garcia, "An Approach to Environmental Study from Observations and Sensing Towards a Digital Twin," *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering*, vol. 514, pp. 247–262, 2023.

[8] C. Zhong, S. Cheng, M. Kasoar, and R. Arcucci, "Reduced-order digital twin and latent data assimilation for global wildfire prediction," *Natural Hazards and Earth System Sciences*, vol. 23, no. 5, pp. 1755–1768, May 2023.

[9] A. Cirulis, L. Taube, and Z. Erics, "Automated Generation of Digital Twin in Virtual Reality for Interaction with Specific Nature Ecosystem," *Lecture Notes in Computer Science*, vol. 13309, pp. 187–202, 2022.

[10] R. Kusakabe, T. Ichimura, K. Fujita, M. Hori, and L. Wijerathne, "Large-Scale Stabilized Multi-physics Earthquake Simulation for Digital Twin," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 12743 LNCS, pp. 3–15, 2021.

[11] V. Cheverda *et al.*, "Digital Twin of the Seismogeological Object: Building and Application," *Communications in Computer and Information Science*, vol. 1129 CCIS, pp. 214–224, 2019.

[12] J. P. Duque and M. A. Brovelli, "BUILDING A DIGITAL TWIN OF THE ITALIAN COASTS," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLVIII-4-W1-2022, no. 4/W1-2022, pp. 127–133, Aug. 2022.

[13] F. Sougioultzoglou and E. Cook, "Digital Twins, Planes, and Drones: Bridging the Gap in Arctic Polar Altimetry Data," *1st Global Space Conference on Climate Change*, pp. 23–25, 2023.

[14] M. Ahmadi, A. Gholizadeh Lonbar, M. Nouri, A. Sharifzadeh Javidi, A. Tarlani Beris, A. Sharifi, and A. Salimi-Tarazouj, "Supervised multi-regional segmentation machine learning architecture for digital twin applications in coastal regions," *Journal of Coastal Conservation*, vol. 28, no. 2, pp. 1–15, Apr. 2024.

[15] Z. Zlatev and I. Dimov, "Using a Digital Twin to Study the Influence of Climatic Changes on High Ozone Levels in Bulgaria and Europe," *Atmosphere 2022, Vol. 13, Page 932*, vol. 13, no. 6, p. 932, Jun. 2022.

[16] K. Abutalip, A. Al-Lahham, and A. El Saddik, "Digital Twin of Atmospheric Environment: Sensory Data Fusion for High-Resolution PM2.5 Estimation and Action Policies Recommendation," *IEEE Access*, vol. 11, pp. 14 448–14 457, 2023.

[17] G. Chen, J. Yang, B. Huang, C. Ma, F. Tian, L. Ge, L. Xia, and J. Li, "Toward digital twin of the ocean: from digitalization to cloning," *Intelligent Marine Technology and Systems 2023 1:1*, vol. 1, no. 1, pp. 1–12, Sep. 2023.

[18] R. Yuan, H. Zhang, R. Xu, and L. Zhang, "Enhancing Coastal Risk Recognition: Assessing UAVs for Monitoring Accuracy and Implementation in a Digital Twin Framework," *Applied Sciences 2024, Vol. 14, Page 2879*, vol. 14, no. 7, p. 2879, Mar. 2024.

[19] Y. D. Zhao, W. Zeng, Y. Ni, P. Xia, and R. C. Tan, "Research and Design of Hydrological Data Visualization Based on Digital Twin," *Communications in Computer and Information Science*, vol. 2058 CCIS, pp. 277–289, 2024.

[20] W. Li and C. L. Zhang, "System Architecture and Core Technology, Method and Data-Driven Technology of Digital Twin Riverbasin," *Proceedings of the IEEE International Conference on Software Engineering and Service Sciences, ICSESS*, vol. 2022-October, pp. 105–109, 2022.

[21] H. J. Henriksen, R. Schneider, J. Koch, M. Ondracek, L. Troldborg, I. K. Seidenfaden, S. J. Kragh, E. Bøgh, and S. Stisen, "A New Digital Twin for Climate Change Adaptation, Water Management, and Disaster Risk Reduction (HIP Digital Twin)," *Water 2023, Vol. 15, Page 25*, vol. 15, no. 1, p. 25, Dec. 2022.

[22] P. Bauer, B. Stevens, and W. Hazeleger, "A digital twin of Earth for the green transition," *Nature Climate Change 2021 11:2*, vol. 11, no. 2, pp. 80–83, Feb. 2021.

[23] J. L. Moigne, "NASA'S Advanced Information Systems Technology (AIST): Combining New Observing Strategies and Analytics Frameworks to Build Earth System Digital Twins," *International Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 2022-July, pp. 4724–4727, 2022.

[24] T. Huang, S. Hasheminassab, O. Kalashnikova, K. Lee, and J. Roberts, "An Earth System Digital Twin for Air Quality Analysis," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 13984 LNCS, pp. 411–414, 2024.

[25] G. Reshetova, V. Cheverda, and V. Lisitsa, "Digital Twins of Geological Objects: Development and Use," *Communications in Computer and Information Science*, vol. 1437, pp. 300–311, 2021.

[26] S. J. Yun, J. W. Kwon, and W. T. Kim, "A Novel Digital Twin Architecture with Similarity-Based Hybrid Modeling for Supporting Dependable Disaster Management Systems," *Sensors 2022, Vol. 22, Page 4774*, vol. 22, no. 13, p. 4774, Jun. 2022.

[27] L. Buonocore, J. Yates, and R. Valentini, "A Proposal for a Forest Digital Twin Framework and Its Perspectives," *Forests 2022, Vol. 13, Page 498*, vol. 13, no. 4, p. 498, Mar. 2022.

[28] V. Fissore, L. Bovio, L. Perotti, P. Boccardo, and E. Borgogno-Mondino, "Towards a Digital Twin Prototype of Alpine Glaciers: Proposal for a Possible Theoretical Framework," *Remote Sensing 2023, Vol. 15, Page 2844*, vol. 15, no. 11, p. 2844, May 2023.

[29] S. Yi, "Cognitive and semantic construct for Digital Twin river basin geographic information fusion," *2023 11th International Conference on Agro-Geoinformatics, Agro-Geoinformatics 2023*, 2023.

[30] S. Albani, M. Lazzarini, P. Saameno, A. Luna, and O. Barrilero, "New Scenarios Shaping a Digital Twin Earth for Security," *International Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 2022-July, pp. 5019–5022, 2022.

[31] M. Halem *et al.*, "Towards a Dynamic Data Driven Wildfire Digital Twin (WDT): Impacts on Deforestation, Air Quality and Cardiopulmonary Disease," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 13984 LNCS, pp. 403–410, 2024.

[32] L. Hejtmánek, M. Hůla, A. Herrová, and P. Surový, "Forest digital twin as a relaxation environment: A pilot study," *Frontiers in Virtual Reality*, vol. 3, p. 1 033 708, Nov. 2022.

[33] H. Zhang, R. Wang, and C. Wang, "Monitoring and Warning for Digital Twin-driven Mountain Geological Disaster," *Proceedings of 2019 IEEE International Conference on Mechatronics and Automation, ICMA 2019*, pp. 502–507, Aug. 2019.

[34] K. Sookhak Lari, G. B. Davis, and J. L. Rayner, "Towards a digital twin for characterising natural source zone depletion: A feasibility study based on the Bemidji site," *Water Research*, vol. 208, p. 117 853, Jan. 2022.

[35] G. Sanchez-Guzman, W. Velasquez, and M. S. Alvarez-Alvarado, "Modeling a simulated Forest to get Burning Times of Tree Species using a Digital Twin," *IEEE 12th Annual Computing and Communication Workshop and Conference*, pp. 639–643, 2022.

[36] X. Liu, Y. Wang, R. C. Koo, and J. S. Kwan, "Development of a slope digital twin for predicting temporal variation of rainfall-induced slope instability using past slope performance records and monitoring data," *Engineering Geology*, vol. 308, p. 106 825, Oct. 2022.

[37] K. Hyeong-Su, K. Jin-Woo, S. Yun, and W. T. Kim, "A novel wildfire digital-twin framework using interactive wildfire spread sim-

ulator," *International Conference on Ubiquitous and Future Networks, ICUFN*, vol. 2019-July, pp. 636–638, Jul. 2019.

[38] L. Wang, R. Jiang, X. Chen, J. Xie, X. Liu, L. Tian, and M. Wang, "Design and application of digital twin platform based smart Weihe River Basin," *Int. Conf. on Smart Transportation and City Engineering*, vol. 12460, pp. 1091–1096, Dec. 2022.

[39] D. Hollenbeck and Y. Q. Chen, "A Digital Twin Framework for Environmental Sensing with sUAS," *Journal of Intelligent and Robotic Systems: Theory and Applications*, vol. 105, no. 1, pp. 1–15, 2022.

[40] A. Tzachor, O. Hendel, and C. E. Richards, "Author Correction: Digital twins: a stepping stone to achieve ocean sustainability?" *npj Ocean Sustainability 2023 2:1*, vol. 2, no. 1, pp. 1–1, Dec. 2023.

[41] T. I. Zohdi, "A machine-learning framework for rapid adaptive digital-twin based fire-propagation simulation in complex environments," *Computer Methods in Applied Mechanics and Engineering*, vol. 363, p. 112 907, May 2020.

[42] X. Rao, R. Ma, L. Zhang, and J. Liu, "Research and design of a digital twin-driven smart river basin platform," *Metaverse*, vol. 3, no. 2, pp. 1–11, 2022.

[43] C. I. Alperen, G. Artigue, B. Kurtulus, S. Pistre, and A. Johannet, "A Hydrological Digital Twin by Artificial Neural Networks for Flood Simulation in Gardon de Sainte-Croix Basin, France," *IOP Conference Series: Earth and Environmental Science*, vol. 906, no. 1, p. 012 112, Nov. 2021.

[44] H. Wu *et al.*, "Multi-level voxel representations for digital twin models of tunnel geological environment," *International Journal of Applied Earth Observation and Geoinformation*, vol. 112, p. 102 887, Aug. 2022.

[45] J. Wang, P. Li, X. Li, and H. Zhu, "Complex 3D Geological Modeling Based on Digital Twin," *IOP Conference Series: Earth and Environmental Science*, vol. 861, no. 7, Oct. 2021.

[46] J. S. Jeong and H. S. Lee, "Unstructured Grid-Based River–Coastal Ocean Circulation Modeling towards a Digital Twin of the Seto Inland Sea," *Applied Sciences 2023, Vol. 13, Page 8143*, vol. 13, no. 14, p. 8143, Jul. 2023.

[47] V. Todorov and I. Dimov, "Unveiling the Power of Stochastic Methods: Advancements in Air Pollution Sensitivity Analysis of the Digital Twin," *Atmosphere 2023, Vol. 14, Page 1078*, vol. 14, no. 7, p. 1078, Jun. 2023.

[48] T. R. Allen *et al.*, "A Digital Twin to Link Flood Models, Sensors, and Earth Observations for Coastal Resilience in Hampton Roads, Virginia, U.S.A," pp. 1388–1391, Oct. 2023.

[49] S. Cacciaguerra, A. Costa, F. Quareni, P. Papale, F. Cannavò, A. Folch, G. Macedonio, and S. Barsotti, *Digital Twin Components for Geophysical Extreme Phenomena: the example of Volcanic Hazards within the DT-GEO project — DIGITAL.CSIC*, Jun. 2023.

[50] S. J. Song and Y. G. Jang, "Construction of digital twin geotechnical resistance model for liquefaction risk evaluation," *ACM International Conference Proceeding Series*, Sep. 2018.

[51] Q. Wan, L. Niu, Y. Wang, and H. Chen, "A Digital Twin Watershed Three-dimensional Solid Topography Data Backplane Creation Method," *Proceedings - 2023 9th Annual International Conference on Network and Information Systems for Computers, ICNISC 2023*, pp. 331–339, 2023.

[52] T. Huang *et al.*, "An Earth System Digital Twin for Flood Prediction and Analysis," *International Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 2022-July, pp. 4735–4738, 2022.

[53] H. Chen, C. Fang, and X. Xiao, "Visualization of Environmental Sensing Data in the Lake-Oriented Digital Twin World: Poyang Lake as an Example," *Remote Sensing 2023, Vol. 15, Page 1193*, vol. 15, no. 5, p. 1193, Feb. 2023.

[54] H. Zhou, "Enhanced Storm Surge Forecasting Model for Digital Twin Watershed Applications of Ningbo," *2023 11th International Conference on Agro-Geoinformatics, Agro-Geoinformatics 2023*, 2023.

[55] A. Tarpanelli, B. Bonaccorsi, M. Sinagra, A. Domeneghetti, L. Brocca, and S. Barbetta, "Flooding in the Digital Twin Earth: The Case Study of the Enza River Levee Breach in December 2017," *Water 2023, Vol. 15, Page 1644*, vol. 15, no. 9, p. 1644, Apr. 2023.

[56] L. Brocca *et al.*, "A Digital Twin of the terrestrial water cycle: a glimpse into the future through high-resolution Earth observations," *Frontiers in Science*, vol. 1, p. 1 190 191, Mar. 2024.

[57] V. Grossmann, D. Nakath, M. Urlaub, N. Oppelt, R. Koch, and K. Köser, "Digital twinning in the ocean-challenges in multimodal sensing and multiscale fusion based on faithful visual models," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 5, no. 4, pp. 345–352, May 2022.

[58] Y. Qiu, H. Liu, J. Liu, D. Li, C. Liu, W. Liu, J. Wang, and Y. Jiao, "A Digital Twin Lake Framework for Monitoring and Management of Harmful Algal Blooms," *Toxins 2023, Vol. 15, Page 665*, vol. 15, no. 11, p. 665, Nov. 2023.

[59] M. Mõttus *et al.*, "A Methodology for Implementing a Digital Twin of the Earth's Forests to Match the Requirements of Different User Groups," *GI_Forum*, vol. 1, pp. 130–136, 2021.

[60] J. Hoffmann, P. Bauer, I. Sandu, N. Wedi, T. Geenen, and D. Thiemert, "Destination Earth – A digital twin in support of climate services," *Climate Services*, vol. 30, p. 100 394, Apr. 2023.

[61] H. Zhang and X. Feng, "Reliability improvement and landscape planning for renewable energy integration in smart Cities: A case study by digital twin," *Sustainable Energy Technologies and Assessments*, vol. 64, p. 103 714, Apr. 2024.

BIBLIOGRAPHY - OTHER REFERENCES

[62] Imarc, *Digital twin market: Global industry trends, share, size, growth, opportunity and forecast 2024-2032*, imarc, Mar. 2023.

[63] D. Jones, C. Snider, A. Nassehi, J. Yon, and B. Hicks, "Characterising the digital twin: A systematic literature review," *CIRP Journal of Manufacturing Science and Technology*, vol. 29, pp. 36–52, 2020.

[64] D. M. Botín-Sanabria, A.-S. Mihaita, R. E. Peimbert-García, M. A. Ramírez-Moreno, R. A. Ramírez-Mendoza, and J. d. J. Lozoya-Santos, "Digital twin technology challenges and applications: A comprehensive review," *Remote Sensing*, vol. 14, no. 6, 2022.

[65] D. Adamenko, S. Kunnen, R. Pluhnau, A. Loibl, and A. Nagarajah, "Review and comparison of the methods of designing the digital twin," *Procedia CIRP*, vol. 91, pp. 27–32, 2020.

[66] C. Lo, C. Chen, and R. Y. Zhong, "A review of digital twin in product design and development," *Advanced Engineering Informatics*, vol. 48, p. 101 297, 2021.

[67] X. Fang, H. Wang, G. Liu, X. Tian, G. Ding, and H. Zhang, "Industry application of digital twin: From concept to implementation," *The International Journal of Advanced Manufacturing Technology*, vol. 121, Jul. 2022.

[68] E. Karabulut, S. F. Pileggi, P. Groth, and V. Degeler, "Ontologies in digital twins: A systematic literature review," *Future Generation Computer Systems*, vol. 153, pp. 442–456, 2024.

[69] J. Huang, *Nvidia to build earth-2 supercomputer to see our future*, NVIDIA, Nov. 2021.

[70] *European digital twin of the ocean (dto)*, European Union, 2022.

[71] *Destination earth*, European Union, 2022.

[72] M. Maimour, A. Ahmed, and E. Rondeau, "Survey on digital twins for natural environments: A communication network perspective," *Internet of Things*, vol. 25, p. 101 070, 2024.

[73] M. Grieves and J. Vickers, "Origins of the digital twin concept," Aug. 2016.

[74] S. Appleton, "Map scale," *National Geographic*, Oct. 2023.

[75] M. Grieves, "Product lifecycle management: The new paradigm for enterprises," *International Journal of Product Development - Int J Prod Dev*, vol. 2, no. 1-2, pp. 71–84, Jan. 2005.

[76] E. Glaessgen and D. Stargel, "The digital twin paradigm for future nasa and u.s. air force vehicles," in *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*. 2012. eprint: https://arc.aiaa.org/doi/pdf/10.2514/6.2012-1818.

[77] A. El Saddik, "Digital twins: The convergence of multimedia technologies," *IEEE MultiMedia*, vol. 25, no. 2, pp. 87–92, 2018.