POLITECNICO DI MILANO

Master in

Building Information Modelling

European Master in
Building Information Modelling

Construction Digital Twin: A Framework for a General Contractor

Supervisors:
Prof. Marco Trani
Eng. Davide Simeone (Company Tutor)
Eng. Luca Sivieri (Company Tutor)

Author:
Iraj Esmaeili

Co-funded by the Erasmus+ Programme of the European Union

a.a. 2020/2021

AUTHORSHIP RIGHTS AND CONDITIONS OF USE OF THE WORK BY THIRD PARTIES

ACKNOWLEDGEMENTS

I would like to deliver my gratitude to my supervisor and tutors at WeBuild company, professor Marco Trani, Eng. Davide Simeone, and Eng. Luca Sivieri, for their kind guidance and support throughout this work.

I would also give my thanks to the BIM A+ board coordinated by professor Miguel Azenha, and the Erasmus Mundus BIM A+ consortium for providing the scholarship to pursue my studies.

Finally, I would declare my deepest gratitude to my family for their unconditional love and support, and also to my friends in BIM A+.

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration.

I further declare that I have fully acknowledged the Code of ethics and conduct of Politecnico di Milano.

SOMMARIO

Titolo: Gemello digitale delle costruzioni: un quadro per un general contractor.

L'Industria 4.0, come quarta ondata della rivoluzione industriale, offre ampi vantaggi per vari settori come la produzione, l'aerospaziale, l'ingegneria dei sistemi, il petrolio e il gas e l'edilizia. Digital Twin (DT), come uno dei principali concetti di Industria 4.0, introduce un nuovo paradigma nel settore delle costruzioni come replica virtuale in tempo reale di un bene fisico. Sebbene molti settori come quello manifatturiero, automobilistico, aeronautico e sanitario utilizzino ampiamente i concetti di Industria 4.0, l'industria delle costruzioni è in ritardo in termini di loro adozione e attuazione, nonché nella capitalizzazione dei loro benefici. Ad oggi, la maggior parte degli studi nel settore delle costruzioni si è concentrata sull'implementazione del DT nelle fasi di esercizio e manutenzione degli impianti mentre l'uso del DT nella fase di costruzione non è stato ancor affrontato a sufficienza. Una caratteristica delle attuali tecnologie di monitoraggio (ad es. telemetri, scansione laser, GPS, RFID, Wi-Fi, UWB, sensori intelligenti, ecc.) utilizzate nel settore edile è che i dati raccolti vengono generalmente utilizzati in modo isolato con un unico focus tematico, mentre sono pochissimi i casi di utilizzo integrato di più tecnologie. C'è inoltre una mancanza di chiarezza sulle potenziali tecnologie per livelli più elevati di CDT in termini di integrazione con piattaforme socio-tecniche e il relativo utilizzo di simulazioni, ottimizzazione, apprendimento e coinvolgimento dell'utente finale, principalmente a causa della mancanza di implementazione e ricerca a tale livelli di sofisticatezza. La presente ricerca ha lo scopo di studiare l'implementazione di DT durante la fase di costruzione dei progetti. Un framework Digital Twin (CDT) di costruzione è sviluppato per dimostrare i diversi livelli necessari per l'implementazione del concetto DT di supporto alle attività di costruzione. In seguito allo sviluppo del framework CDT, è stato implementato un caso di studio sulla gestione dei movimenti terra come applicazione del CDT proposto al fine di convalidare il framework e dimostrare i vantaggi dell'utilizzo dei gemelli digitali durante la fase di costruzione. La simulazione delle operazioni dei movimenti terra è stata eseguita come parte del CDT e i risultati della simulazione nella valutazione di vari scenari ipotetici per l'allocazione ottimale delle risorse e la gestione delle operazioni hanno dimostrato i vantaggi dell'utilizzo del CDT nella fase di costruzione dei progetti e per i *general contractor*. Con lo sviluppo di un framework CDT e l'implementazione di un caso di studio basato su di esso e integrato con diverse tecnologie abilitanti, questo studio cerca di colmare le lacune esistenti nell'utilizzo dei gemelli digitali nel settore delle costruzioni. L'utilizzo di tale CDT nell'implementazione di diverse applicazioni può aiutare le imprese di costruzioni in una migliore gestione dei loro processi di costruzione, colmando il divario di informazioni tra i modelli BIM as-designed in fase di progettazione e i modelli BIM as-built al momento della consegna del progetto, abilitando vari servizi attraverso l'acquisizione di dati in tempo reale, l'aggiornamento del modello e l'analisi dei dati e le previsioni per il monitoraggio dei progressi, il monitoraggio del sito, l'allocazione delle risorse, l'esecuzione di scenari ipotetici. Come risultato dell'utilizzo di un CDT nella costruzione, emergono i diversi impatti come riduzione dei costi, la migliore collaborazione e lo scambio di informazioni tra vari domini, la gestione della costruzione basata sui dati, la consapevolezza dello stato in tempo reale o quasi del progetto e delle sue prestazioni e una maggiore trasparenza durante tutto il ciclo costruttivo.

Parole chiave: BIM, Costruzione basata sui dati, Costruzione Digital Twin, Fase di costruzione, General contractor.

ABSTRACT

Industry 4.0, as the fourth wave of the industrial revolution, offers ample benefits for various industries like manufacturing, aerospace, systems engineering, oil and gas, and construction. Digital Twin (DT), as one of the main concepts of Industry 4.0, introduces a new paradigm in the construction industry as a real-time virtual replica of a physical asset. Although many industries such as manufacturing, automotive, aviation, and healthcare are extensively using the concepts of Industry 4.0, the construction industry is lagging in terms of adopting and implementing Industry 4.0 principles and taking advantage of its benefits. To date, most of the construction industry studies have focused on the implementation of DT in operation and maintenance phases of facilities, and the use of DT in the construction phase has not been addressed sufficiently. A noticeable feature of the current monitoring technologies (e.g., range finders, laser scanning, GPS, RFID, Wi-Fi, UWB, smart sensors, etc.) used in the construction industry is that the gathered data is generally used in an isolated fashion with a single subject focus where there are very few cases of integrated use of more than one technology. There is also a lack of clarity on the potential technologies for higher levels of CDT in terms of integration with socio-technical platforms and using simulation, optimization, learning, and end-user engagement, mainly due to a lack of implementation and research at such levels of sophistication. This research is motivated to study the implementation of DT during the construction phase of projects. A construction digital twin (CDT) framework is developed to demonstrate the different tiers necessary for the implementation of the DT concept in construction. Following the development of the CDT framework, a soil management case study was implemented as an application of the proposed CDT to validate the framework and demonstrate the benefits of using digital twins during construction phase of projects. Simulation of the earthwork operations was performed as part of the CDT, and results of the simulation in evaluating various what-if scenarios for optimum resource allocation and operations management proved the benefits of using CDT in the construction phase of projects and for general contractors. With the development of a CDT framework and implementation of a case study using the proposed framework and enabling technologies, this study tries to address the existing gaps in using digital twins in the construction industry. Using the developed CDT in implementing several applications can assist general contractors in better management of their construction processes, filling the information gap between the as-designed BIM models at the design phase and the as-built BIM models at the time of project hand-over, enabling various services through real-time data acquisition, model updating, and data analytics and predictions for progress tracking, site monitoring, resource allocation, performing what-if scenarios. As a result of using a CDT in construction, several impacts such as reducing costs, improved collaboration and information exchange among various domains, data-driven construction management, real-time or near real-time status awareness of the project and its performance, and improved transparency can be anticipated.

Keywords: BIM, Construction Digital Twin, Construction Phase, Data-driven Construction, General Contractor.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Although the ample benefits of Industry 4.0 have been proven and industries like automotive manufacturing and maintenance are focusing on the interaction between industry elements and IoT devices, the construction industry is lagging in technology implementation (Hasan et al. 2021). As a result of the Industry 4.0 concept that is applied to the construction industry, Construction 4.0 proposes a framework for extensive application of BIM (Building Information Modelling) in different stages of a product/asset lifecycle, including design, construction, industrial modular production, cyber-physical systems (CPS), supply chain and construction sites works monitoring, and data analytics including big data, AI, cloud computing and blockchain (Sacks et al. 2020; Turner et al. 2021). Industry 4.0 technologies are increasingly becoming appealing since construction projects' complexity is growing and coupled with the need for higher productivity (Turner et al. 2021). In addition to BIM, Digital Twin (DT), as one of the main concepts of Industry 4.0 and a subset of a CPS system, bespeaks a new paradigm in the construction industry as a real-time virtual replica of a physical asset. The federated BIM models produced during design and construction phases reflect the as-designed and as-planned states of the project, while Digital Twin captures the as-built and as-performed states by gathering real-time information of the physical asset and being updated constantly (Sacks et al. 2020). Therefore, the utilization of BIM and Digital Twin concepts ensures efficient information flow and management of the physical entity in its different stages and management aspects, including design, supply chain, construction, commissioning and handover, and maintenance. Moreover, DT deployment offers several opportunities such as asset performance management, asset risk assessment, shorter time for production planning, miscommunication and information wastage avoidance, collaborative decision making, and process automation (Wanasinghe et al. 2020).

Several industries such as manufacturing, automotive, aviation, and healthcare are extensively using the concepts of Industry 4.0, but the construction industry is in its infancy in terms of adopting and implementing Industry 4.0 principles. According to Fenn and Raskino (2008), a new technology undergoes five major stages from its growth to development, referred to Innovation Hype Cycle (see Figure 1). As it is evident from Figure 1, DT and BIM are at the “Peak of Inflated Expectation” and “Trough of Disillusionment” stages, respectively. It means that DT will experience a decline in its attractiveness to industries, and BIM is on its way to passing the decline stage. Therefore, researchers and early adopters should conduct studies to overcome the challenges of DT and BIM to reach the “Slope of Enlightenment” and “Plateau of Productivity,” respectively to expose the applications of these technologies to the real world (El Jazzar et al. 2020).

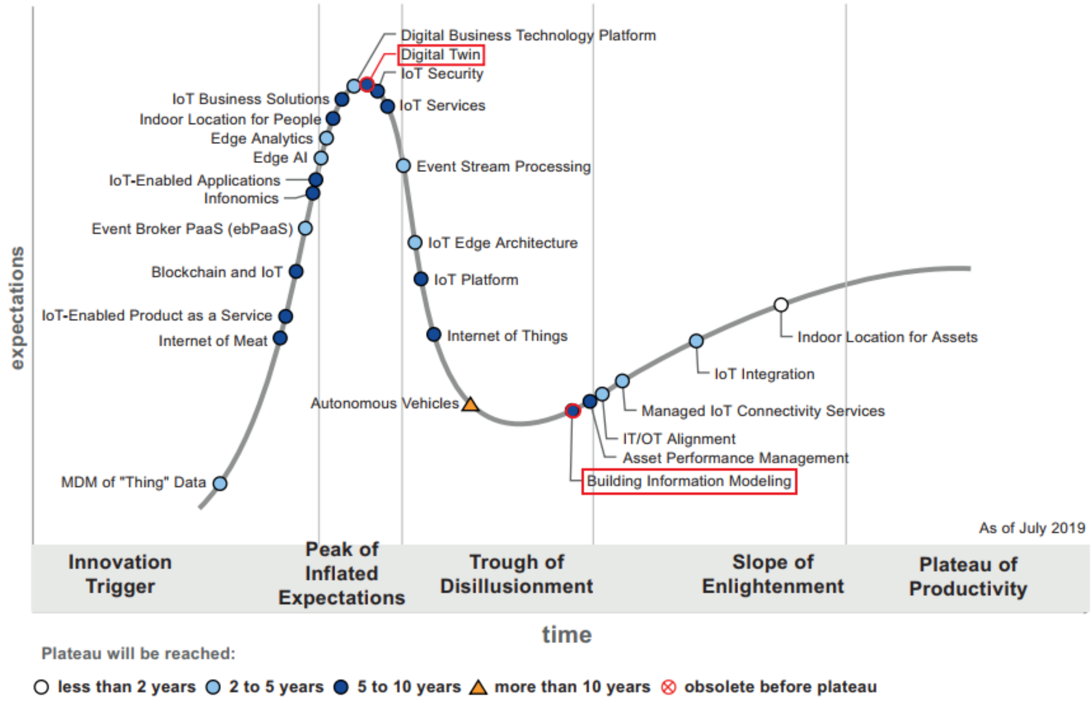


Figure 1- Hype-Cycle for Emerging technologies (Adopted from Gartner, Inc. (Walker 2018))

Moreover, most construction industry studies and practices have focused on implementing DT in operation and maintenance phases of facilities (El Jazzar et al. 2020; Lu et al. 2020), or proposing a theoretical/conceptual model of DT. Hence, there is also a need to implement it in the building and infrastructure design, construction, and supply chain phases.

The main focus of this study is multi-fold. It aims at proposing a coherent, comprehensive, and feasible construction digital twin (CDT) framework and workflow for information management throughout the construction phase. The defined conceptual schema of the construction digital twin identifies, for each component – the virtual model, the data collection systems on-site, the connecting infrastructure – their requirements, their functioning and the potential solutions available on the market, also taking into account the necessities of each discipline. Adopting predictive algorithms and data analytics to exploit the data and the information stored in the digital twin will be explored. The enabling technologies for data acquisition in a construction site will be discussed. One simulated case study in the field of civil projects will be discussed to demonstrate the potential implementation of the digital twin approach. In this case study, the digital twin solutions and data collection equipment will be explained, and modalities of predicting construction evolution and reporting to the project management team will be defined. To conclude, this study will discuss and assess potentialities and limits regarding the digital twin implementation in the construction industry, providing perspectives and requirements for further developments in the field.

This study benefits from a mixed-method approach for its research methodology. This approach consists of an extensive literature review, BIM modelling and DT implementation along with their components, data analytics and artificial intelligence techniques, and a case study conducted in a construction project.

This research project contributes to the body of knowledge by proposing a comprehensive framework for implementing DTs for the construction phase of an asset. It enhances the industry practice by better information management in the construction stages of an asset or its related components in favor of contractors' and owners’ satisfaction. The uniqueness of the proposed research project lies in its integrated approach for taking the benefits of BIM and Industry 4.0 technologies in the construction industry as a comprehensive method that is applicable in real-world projects and supporting them at the construction phase. The proposed DT approach can be highly desirable since it enables actionable knowledge and effective decision-making in the construction phase based on real-time data and performing “what-if” scenarios, and reduces the construction waste to a great extent (LaGrange 2019; Roxin et al. 2019; Sacks et al. 2020; Wanasinghe et al. 2020).

The remainder of the study discusses the digital twin as a concept of Industry 4.0. Then, construction digital twin (CDT) will be introduced along with its components and the proposed CDT framework. Next, a soil management application will be implemented as a case study in implementing the CDT framework. Finally, in the Conclusion section, the impacts and challenges of the proposed CDT framework will be evaluated.

This page is intentionally left blank

2. DIGITAL TWIN – AN INDUSTRY 4.0 CONCEPT

Industry 4.0, as an evolution from Industry 1.0 to Industry 3.0, was first coined by the German federal government in 2011 as a strategic response to the German industry (Klinc and Turk 2019). The initiative was to catch up with the industrialization of Asia, and it was expected to alternate the production and promote the computerization of manufacturing (Gokalp et al. 2017; Kan and Anumba 2019; Klinc and Turk 2019). Later on, this digital transformation, i.e. Industry 4.0, has been adopted at international levels, e.g. European Union, as a strategy to the modernization of the industry such as manufacturing and a response to their associated challenges (Klinc and Turk 2019; Negri et al. 2017). It offers the transformation of ordinary systems, including machines and production systems, to self-aware, self-predict, self-compare, self-configure, self-maintain, and self-organize devices and systems (Gokalp et al. 2017; Lee et al. 2015). This transformation enables the systems to gather real-time data of the product/process and self-diagnose to prevent or minimize the production disruptions or downtimes (Gokalp et al. 2017).

Industry 4.0 is a collective term for a number of building blocks consisting of Internet of Things (IoT), Big Data, Cloud Computing, the Internet of Services, Cyber-Physical System (CPS), Smart Factories, Advanced Manufacturing, Digital Twin (DT) etc. (Kan and Anumba 2019; Klinc and Turk 2019; Negri et al. 2017). The essential building blocks of Industry 4.0 are cyber-physical systems (CPS) that link the digital and physical worlds (Hasan et al. 2021; Klinc and Turk 2019). CPS are known as 'integrations of computation with physical processes' where computational resources monitor the physical world with a feedback loop and interact with each other through embedded computers and networks (Lee 2008). A differentiating factor between Industry 3.0 (automation) and Industry 4.0 is the human mediator that connects the digital and physical realms (Klinc and Turk 2019), while in CPS, the cyber and physical worlds are integrated through networks and computational resources. Having said that, "bi-directional communication" is a critical component of CPS for effective integration of the virtual model with the corresponding physical components (Anumba et al. 2010).

According to a guideline by Lee et al. (2015), in general, a CPS has two main functional components: (1) advanced connectivity, (2) intelligent data management, analytics and computational capability. The former establishes connectivity to gather real-time data from the physical world and information feedback from the cyber space, i.e. bi-directional communication, and the latter constructs the cyber space (Lee et al. 2015). Based on this guideline, they proposed a 5-level architecture for CPS implementation in the industry, as shown in Figure 2.

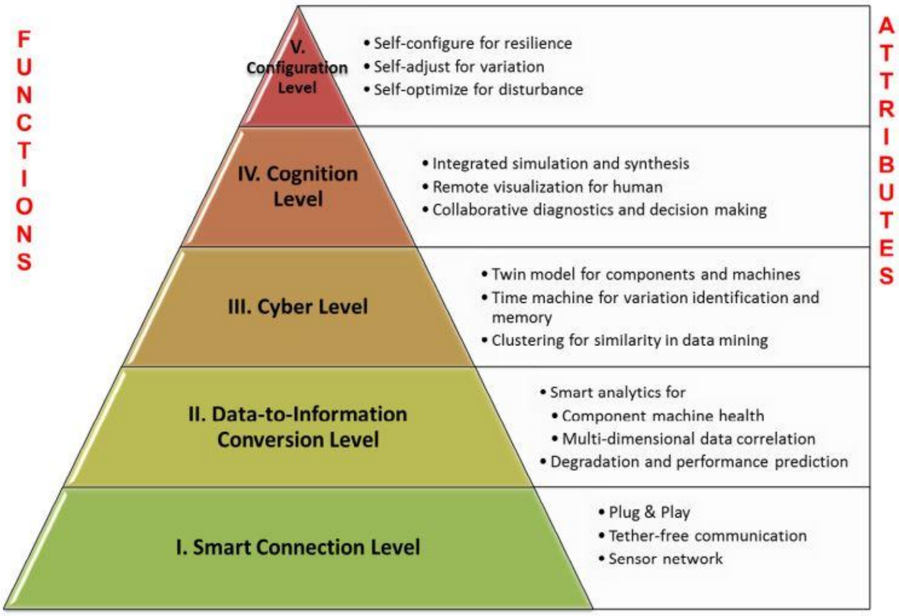


Figure 2- 5C architecture for CPS implementation (adopted from Lee et al. (2015))

2.1. Smart Connection Level

Accurate and reliable data acquisition is the first step in developing a CPS and consists of "Plug & Play" devices, communication and sensor networks. Selecting proper sensors and adopting suitable standardized protocols for data acquisition and transferring to a central server are two important factors in this level (Klinc and Turk 2019; Lee et al. 2015).

2.2. Data to Information Conversion Level

At this level, various tools and methodologies are deployed to derive information from data gathered at level 1 to be utilized in analysis and predictions in order to support decision-making (Klinc and Turk 2019; Lee et al. 2015).

2.3. Cyber Level

This level plays as a central information hub in the CPS architecture (Lee et al. 2015). Digital Twin is a key component of this level, which means: "a digital representation of an object that exists or will exist in the physical world" (Klinc and Turk 2019). Due to massive information gathered at this level, information and insights could be extracted at this level by deploying data mining and data analytics methods.

2.4. Cognition Level

This level produces comprehensive knowledge of the physical system to enable expert users with their decision-making. Hence, presentation and info-graphics methods are necessary to transfer the acquired knowledge correctly to the users (Lee et al. 2015). AI technologies can be exploited to perform machine learning and diagnostics and make advanced decisions (Klinc and Turk 2019).

2.5. Configuration Level

This level applies the corrective and preventing decisions made in cognition level to the physical system, i.e., feedback from cyber space to physical space and controlling the machines to make them self-configure and self-adaptive (Lee et al. 2015).

CPS, as a core component of Industry 4.0 requires Digital Twin for its development. As CPS aims at enhancing the production system and controlling the production in real-time, Digital Twin, as a specific form of a CPS, aims at providing a digital replica of the physical product or process in real-time or near real-time and capture all useful information throughout the product or process lifecycle (Kan and Anumba 2019; Uhlemann et al. 2017). Compared to CPS, DT focuses more on data and high fidelity models (Zheng et al. 2019). Serving as the virtual and computerized counterpart of a physical system, DT enables simulation and real-time synchronization of the sensed data from the field that is acquired via enabling technologies of Industry 4.0 such as IoT (Negri et al. 2017). DT was first developed by National Aeronautics and Space Administration (NASA) during the Apollo 13 mission when one of the oxygen tanks exploded on April 13, 1970, two days after launch, and the NASA flight control team used a simulated environment to model, simulate and test possible scenarios and successfully found a solution (Barricelli et al. 2019; NASA 2004). Though DT was first born in the aerospace field, it has been recently adopting and gaining interest in a range of industries such as manufacturing, automotive, healthcare, medicine, aviation, and terrestrial exploration (Barricelli et al. 2019; Uhlemann et al. 2017; Wanasinghe et al. 2020).

The rewarding attempts of utilizing the Industry 4.0 concept, such as CPS, from other industries encouraged the construction industry to recognize the potential benefits of such concepts. For example, energy monitoring and structural health monitoring (SHM) were among the first applications of the CPS in the construction industry (Kan et al. 2018). Although the construction industry is at the beginning of the digital transformation process, it has gained speed to become digital and seeks out innovations (Kan and Anumba 2019; Zhou et al. 2020). As such, DT is also at its infancy in the construction industry domain; therefore more research in its applications in construction and other domains would be beneficial to its maturity and adoption within the construction industry (Kan and Anumba 2019). Applying appropriate integration of automated systems and information communication technologies (ICT) in the physical construction process will result in a more controlled project delivery process (Anumba et al. 2010). According to the National Institute of Standards and Technology (NIST) report on interoperability problems in the U.S. capital facilities industry, $ 15.8 billion was quantified as the annual interoperability costs for the capital facilities industry in 2002 besides other inefficiency and lost opportunity costs associated with the inadequate interoperability (O’Connor et al. 2004). The adoption of CPS approach offer a more intelligent, sustainable, and interoperable construction process resulting in a major reduction in such losses (Anumba et al. 2010).

2.6. Digital twin definitions

Grieves introduced the concept of Digital Twin in 2003 at his university course on product lifecycle management (PLM) (Grieves 2014). NASA also provided one of the first definitions of Digital Twin as “an integrated multi-physics, multi-scale, probabilistic simulation of a vehicle or system that uses the best available physical models, sensor updates, fleet history, etc., to mirror the life of its flying twin.

The digital twin is ultra-realistic and may consider one or more important and interdependent vehicle systems, including propulsion/energy storage, avionics, life support, vehicle structure, thermal management/TPS, etc. In addition to the backbone of high-fidelity physical models, the digital twin integrates sensor data from the vehicle's on-board integrated vehicle health management (IVHM) system, maintenance history, and all available historical/fleet data obtained using data mining and text mining" (Shafto et al. 2012). According to NASA's definition, a 'digital twin' can be defined as a prototype that mirrors its corresponding physical flying twin in real-time.

DT is usually wrongly referred to as only a 3D model of the physical world (LaGrange 2019; Wanasinghe et al. 2020). Therefore, though there is no such universally accepted definition of Digital Twin due to different viewpoints, researchers and institutions have provided broader definitions of Digital Twin (Kan and Anumba 2019; Wanasinghe et al. 2020). This is also the case among AEC/FM industry stakeholders who do not share a common definition of DT (Camposano et al. 2021). Table 1 provides various definitions provided in the literature with respect to the related industry.

Table 1- Sample Definitions of a Digital Twin

Reference	Definition	Industry
Glaessgen and Stargel (2012)	"A Digital Twin is an integrated multi-physics, 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."	Aeronautics
Negri et al. (2017)	"the virtual and computerized counterpart of a physical system that can be used to simulate it for various purposes, exploiting a real-time synchronization of the sensed data coming from the field"	Manufacturing
Hasan et al. (2021)	"a representation of a physical entity via a virtual model that mimics its change in state whether it is mechanical motion or physical/dimensional changes."	Construction
Xu et al. (2018)	"a comprehensive virtual copy of the physical operating facility throughout its complete lifecycle – starting from idea and up to decommissioning and project closure "	Construction
LaGrange (2019)	"advanced 3D models, representing dynamic, cross-domain digital models that mirror the performance and operation of a physical asset or process as it moves through the lifecycle – from design, engineering, construction, commissioning, and finally, into operation."	Oil and Gas
Khajavi et al. (2019)	The digital counterpart of a physical asset designed to integrate real-time sensor readings to analyse and improve asset's	Construction

	operational efficiency and interaction with the environment and users, and to enable predictive maintenance.	
National Infrastructure Commission (2017)	“A digital twin is a virtual representation of a physical object or system across its lifecycle, using real-time data to enable understanding, learning and reasoning.”	Infrastructure
Madni et al. (2019)	A digital twin is a dynamic virtual representation of a physical system that is continually updated with its physical twin's performance, maintenance, and health status data throughout its life cycle.	Systems Engineering
Jones et al. (2020)	“a system that couples physical entities to virtual counterparts, leveraging the benefits of both the virtual and physical environments to the benefit of the entire system.”	Manufacturing
Bolton et al. (2018)	“a realistic digital representation of assets, processes or systems in the built or natural environment.”	Built Environment
Pan and Zhang (2021)	“a mirror and digital depiction of the actual production process, which can imitate all aspects of physical processes under the integration of physical products, virtual products, and relevant connection data.”	Construction
B.I.M.Dictionary (2021)	“A set of digital assets – models, documents and datasets - that mirror a physical Asset for part/whole of the Asset Life Cycle.”	Built Environment
Grieves (2014)	“a virtual, digital equivalent to a physical product … [and] rich representations of products that are virtually indistinguishable from their physical counterparts.”	Manufacturing
Boje et al. (2020)	“the digital representation of the building, enriched by the addition of sensing capabilities, big data and the Internet of Things from site to building operation.”	Construction
Barricelli et al. (2019)	“A DT is a living, intelligent and evolving model, being the virtual counterpart of a physical entity or process.”	General
Grieves and Vickers (2016)	“the Digital Twin is 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.”	Manufacturing

Regardless of the type of definition or industry, DT has three main components (Grieves and Vickers 2016):

- Physical space,
- Virtual space,
- Link for data/information flow between real/virtual spaces.

This new generation of CPS integrates the physical and virtual worlds via twinning that establishes the inter-connection between physical and virtual spaces through data collected by sensors (Kan and Anumba 2019). In this connection loop, real-world data from the physical components are collected and sent for processing, and the virtual part applies AI and engineering models to extract insights and information used for managing the physical part (Boje et al. 2020). Ideally, DT contains all information that could be obtained from various facets of a physical entity, i.e. a product, a physical system, a process, or even an organization (Grieves and Vickers 2016; Qi et al. 2021). Assuming the costs of working with a physical component and the virtual one are equal today, as it is evident from Figure 3, these costs have divergent trends where the physical costs are increasing, and the virtual costs are decreasing exponentially (Grieves and Vickers 2016). Hence, covering the entire lifecycle of the physical asset, DT's primary benefit is reducing the long-term costs of its physical twin (Boje et al. 2020).

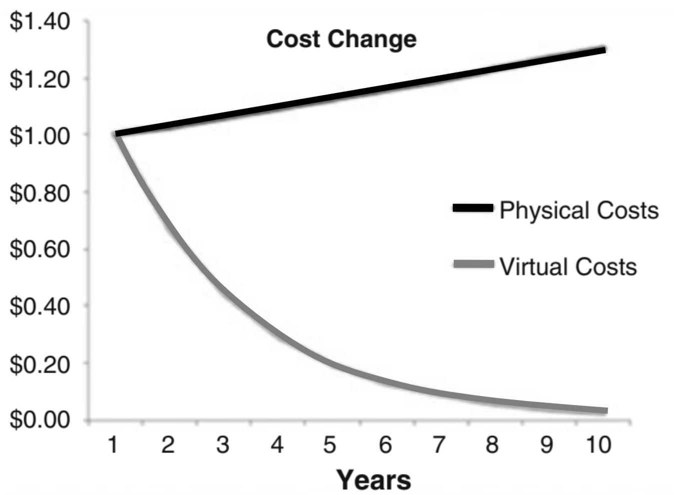


Figure 3- Real versus virtual costs (adopted from Grieves and Vickers (2016))

Collaboration between relevant stakeholders such as customers and designers is another beneficial aspect of DT. This is because DT is an updated and faithful replication of the physical product that enables fast and transparent communication between individuals (Tchana et al. 2019). With real-time data and knowledge base, early simulations of future behaviours of the designed system would be possible to evaluate the performance of the system, which in turn allows savings. DT also contributes to optimizing the operation, manufacturing, inspections, and leveraging of the product's lifetime (Tchana et al. 2019).

Considering the data flow between the physical and digital objects, three levels of digitisation exist: digital model, digital shadow, and digital twin (Ruppert and Abonyi 2020).

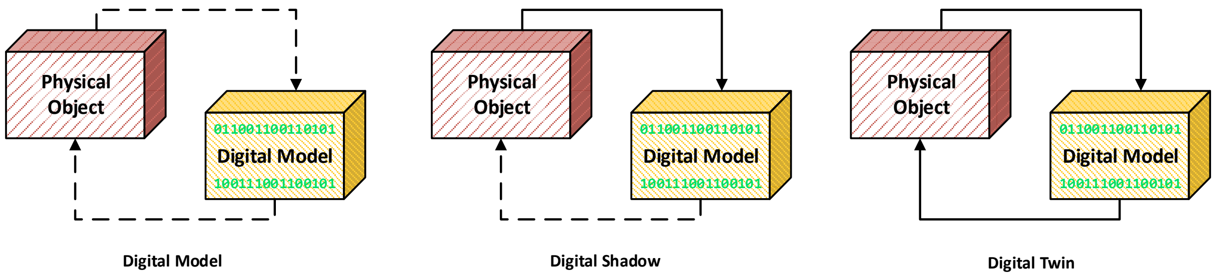


Figure 4- Levels of digitisation

On the first level, Digital Model is offline and the data flow between the physical and virtual objects is done manually (Ruppert and Abonyi 2020). It is believed that the construction industry and BIM are currently in the Digital Model subcategory (El Jazzar et al. 2020; Tchana et al. 2019). On the Digital Shadow level, the data flow is automatic in one way from the physical space to the digital space; therefore, the material world leads the change. Finally, on the top level, Digital Twin automates data flow between the physical and virtual worlds in both directions, and ensures full integration and synchronization (Ruppert and Abonyi 2020).

Considering the BIM maturity levels (Bew 2008), Jones et al. (2020) map BIM levels 1 to 3 across the Digital Model, Digital Shadow, and Digital Twin, respectively, as shown in Figure 5. Two and three-dimensional CAD models (BIM Level 1) being as Digital Models. BIM Level 1 containing data from the actual physical construction (BIM Level 2) showing the Digital Shadow. Finally, BIM Level 2 with two-way data connections (BIM Level 3) being the Digital Twin.

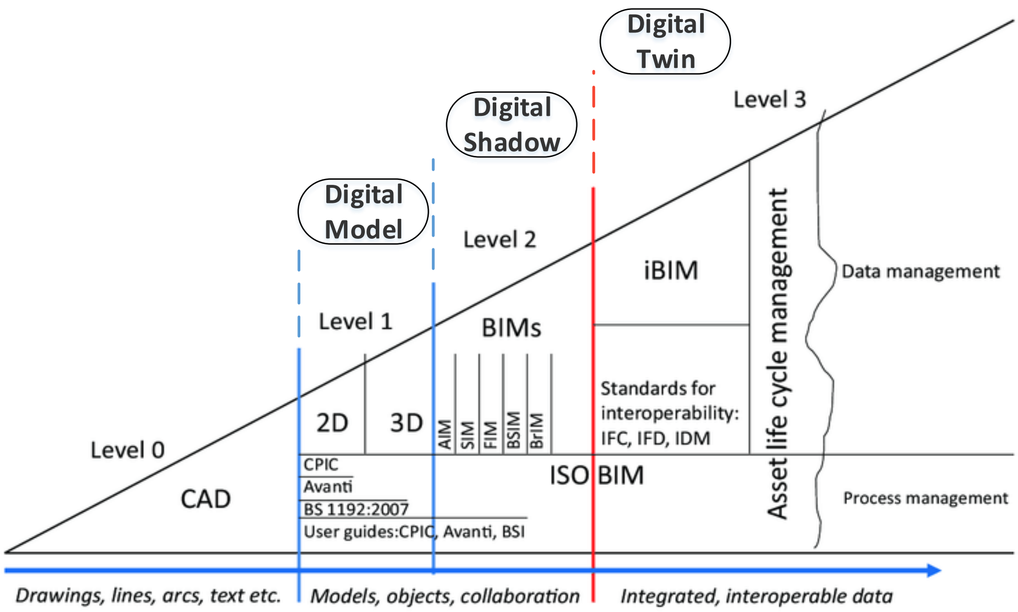


Figure 5- BIM maturity levels in relation to digitization levels

2.7. Digital Twin in the construction industry

Increased interest in the digital twin has encouraged the construction industry to follow this concept (El Jazzar et al. 2020). However, in the construction and built environment, the digital twin is not defined in a comprehensive way by academia, and it is simply considered as BIM models produced in design and construction phases (Sacks et al. 2020). Others looked at it as a digital representation of a built asset such as a building or a bridge, and for the operation and maintenance, some take it as BIM models produced during design and construction (LaGrange 2019; Wanasinghe et al. 2020). The federated BIM models produced during design and construction phases reflect the as-designed and as-planned states of the project and they are not capable of being constantly updated as the physical component's status changes (Sacks et al. 2020). Advancements in information technology such as CPS are transforming construction planning and monitoring methods (Kan et al. 2018). Unlike BIM, Digital Twin captures the as-built and as-performed states by gathering real-time information of the physical asset and being updated constantly (Sacks et al. 2020). Although as-built BIM models represent the current status of the completed work, they only contain the product information which is prepared at the time of delivery to the owner as Asset Information Models (AIM) for operation and maintenance, and they lack the process information of the product's construction history (Sacks et al. 2020).

As DT has proven its benefits in the manufacturing industry, it encouraged researchers and practitioners to make efforts in developing cyber-physical models to support digital development in the construction industry (Pan and Zhang 2021). In the construction industry, improving productivity, sustainability, safety, and achieving other organization or project's goals are the key purposes of a DT (Sepasgozar 2021). Infrastructures, built environment, and city assets can benefit from the applications of DT in monitoring, managing, and predicting an asset's current and future status (Pan and Zhang 2021). For instance, Pan and Zhang (2021) worked on building a digital twin framework with the integration of BIM, IoT, and data mining techniques for more efficient project management. They claimed that with their proposed data-driven digital twin framework, data communication and exploration will be enhanced, resulting in better understanding, predicting, and optimizing the physical construction operations. Lu et al. (2020) presented a system architecture to implement DT in building and city levels. They implemented a case study using the proposed framework in real practice with the focus on Facility Management (FM). Ham and Kim (2020) worked on DT on the city level by proposing a method for leveraging unstructured crowdsourced visual data for locating objects in urban areas that are vulnerable and have potential risks to citizens. They focused on integrating such information with a 3D virtual city model and using this model in a computer-aided virtual environment for other visualizations. Hasan et al. (2021) aimed at implementing Augmented Reality (AR) and DT in a construction context and developed a modular CPS prototype compatible with BIM platforms and smart devices to monitor the operations of the construction site machinery and interact with them. Kan et al. (2018) explored the applicability of CPS and information technology to mobile cranes on construction sites and proposed a five-layer system architecture for planning and monitoring mobile cranes operations. Zhou et al. (2020) proposed a data-driven digital process platform for the digitalization of construction sites. They focused on interoperability between different actors on construction sites and developed their approach to increase productivity and efficiency in the construction processes by real-tile monitoring and process analysis. Anumba et al. (2010) presented a CPS approach for integrating virtual models and physical construction with bi-directional communication to control the construction activities and consistency between the virtual models and the physical construction, and thus reducing the risks of undesirable project outcomes. Turner et al. (2021) discussed utilizing Industry 4.0 on the construction sites and how interconnected Industry 4.0 enablers improve productivity. They also discussed Industry 4.0 concept and its enabling technologies such as data analytics and artificial intelligence, robotics and automation, BIM, sensors and wearables, DT, and industrial connectivity and provides its relevant applications in construction industry especially in modular building components and safety domains. Akanmu and Anumba (2015) developed system architectures, scenarios and prototype systems to implement the CPS approach in construction. They demonstrated two applications of the CPS approach, namely: (1) Steel placement, and (2) Light fixture monitoring and control, and discussed the key elements of the CPS approach and enabling technologies to enhance bi-directional coordination between virtual models and the physical construction. Greif et al. (2020) explored the opportunities and potential business value of adopting DT in construction site logistics by establishing digital twins for bulk silos to continuously monitor and track silo fill levels.

Boje et al. (2020) and Wanasinghe et al. (2020) have conducted extensive literature reviews on the emergence date and distribution pattern of DT in publications in various industries as well as the construction sector. Wanasinghe et al. (2020) revealed the upward trend of DT publications from 2003 until 2020 in different industries, namely: Manufacturing, Automotive, Aviation or Aerospace, Healthcare, Retail, and Oil and Gas. According to their study, there was a big leap in DT publications in 2018 and subsequent years. A similar pattern in the construction sector was shown by Boje et al. (2020), where there was a considerable increase in DT publications in 2017 and especially in 2018, with the emergence of DT articles in 2011. They concluded that although there has been an increase in research in the area of DT, the research about DT is still scarce.

Although from industry to industry (e.g., construction compared to manufacturing), DT demands its specific models, tools, and technologies, the overall architecture, functionalities and features of a DT are considered generic (Boje et al. 2020). Sensing and monitoring the physical, data connection, and utilization of AI are among these generic functionalities of a CPS/DT which are common in cross-domain DT uses (Boje et al. 2020; Lee et al. 2015). Automatically detecting issues, performance evaluation and formulating optimized solutions to enhance the reliability and efficiency of operations increase the potential of the DT to be seen as a promising concept for the progression of the construction industry (Pan and Zhang 2021).

2.8. Digital Twin in the construction phase: Construction Digital Twin (CDT)

The design and construction phases of a project account for a considerable portion of the total project's cost (up to 40 percent). However, most construction industry studies and practices have focused on the implementation of DT in operation and maintenance phases of facilities (El Jazzar et al. 2020). In the product manufacturing stage, DT can realize real-time monitoring of products and predict their performance accurately. In addition, with the utilization of DT, the consistency of the final product with the required specifications can be evaluated in this stage (Zheng et al. 2019). The construction phase of the projects exposes its stakeholders, i.e. designers, consultants, contractors, owners, suppliers, etc., to various challenges as they generate big chunks of information about the product and construction process using their various digital tools and multiple data formats that are not usually interoperable (Sacks et al. 2020). Besides, monitoring the building components, workers, and construction equipment is challenging. Therefore, the DT approach for construction can be highly desirable since it enables actionable knowledge and effective decision-making in the construction phase based on real-time data and performing “what-if” scenarios, and reduces the construction waste to a great extent (LaGrange 2019; Roxin et al. 2019; Sacks et al. 2020; Wanasinghe et al. 2020). As an example, with the use of DT of a construction site, potential issues can be detected and predicted in the virtual space before happening in the physical space (Qi et al. 2021), which in turn prevents the rework costs or increased idle time of workforce and equipment and poor performance due to material unavailability and insufficiency.

A noticeable feature of the current monitoring technologies (e.g., range finders, laser scanning, GPS, RFID, Wi-Fi, UWB, smart sensors, etc.) used in the construction industry is that the gathered data is generally used in an isolated fashion with a single subject focus where there are very few cases of integrated use of more than one technology (Sacks et al. 2020). There is also a lack of clarity on the potential technologies for higher levels of CDT in terms of integration with socio-technical platforms and using simulation, optimization, learning, and end-user engagement, mainly due to a lack of implementation and research at such levels of sophistication (Boje et al. 2020).

Since its emergence, BIM has enhanced the collaboration and information exchange in different project stages among the involved parties. As industries are adopting the Industry 4.0 revolution and automation solutions, BIM, which was initially envisaged to facilitate the information exchange among fragmented entities, is facing new challenges in leveraging big data, IoT and AI (Boje et al. 2020). In addition, BIM tools and methods like generative design, applications for energy analysis, structural analysis, lighting etc., are designed for predictive use in the design or future performance of the built product, not for the project’s construction phase (Sacks et al. 2020). Seemingly, static BIM and its interoperability innovations, such as ISO STEP, IFC, and IfcOwl have limitations in leaping to dynamic design and construction stages where a large amount of data is produced, limiting the added benefits to the construction supply chain (Boje et al. 2020). However, efforts have been made to enhance the IfcOwl ontology, such as EXPRESS to OWL conversion for particular use in the architectural design and construction industry (Pauwels et al. 2017; Pauwels and Terkaj 2016).

Moreover, BIM object models cannot represent the construction process aspects due to the insufficiency of their schema components to reflect such processes (Sacks et al. 2020). The other limitation of BIM is that it cannot evaluate, analyse, and predict the real-time status of a product or process (Pan and Zhang 2021). Therefore, these limitations might hinder the full utilization of BIM throughout the built asset lifecycle.

As mentioned before, a DT integrates the physical and virtual worlds via real-time data from IoT and enables AI and data analytics. The bi-directional communication capability of a DT enables data to flow from the physical to the virtual and information feedback in the other direction, i.e. from the virtual to the physical (Kan et al. 2018). Therefore, a DT has great potentials in addressing the challenges of static BIM and promises a real-time and linked data paradigm and fully represented models of building products in the form of a DT (Boje et al. 2020). In the construction context, DT, as a new phenomenon (Sacks et al. 2020), aims at enhancing the construction processes and so-called nD BIMs by cyber-physical synchronicity to represent the physical construction assets in real-time (Boje et al. 2020). Various applications of BIM have been explored and applied, from design to decommissioning of a built asset in terms of architectural and structural design, quantity surveying, construction project management, and sustainability (Whitlock et al. 2018). However, there are limited studies on BIM uses during the construction phase, such as construction logistics management, and usually, 4D/5D BIM uses have been traditionally implemented during the construction phase due to increased requirements for collaboration, actors and models (Boje et al. 2020; Whitlock et al. 2018). In addition, the application of DT in buildings is mostly limited to building operation and maintenance (Sacks et al. 2020). BIM has several uses during construction, such as scheduling, clash detection, safety management, etc. Although it is becoming more common to use BIM during construction, the level of development of DT during construction is still very low (Boje et al. 2020). Considering the substantial impact of the design and construction processes of a construction asset on its operation costs (Boje et al. 2020), any benefit from DT during construction will have added value. Therefore, this study is motivated by proposing a construction digital twin (CDT) framework for general contractors to use DT in the construction phase of projects and benefit from various advantages it offers.

This page is intentionally left blank

3. CONSTRUCTION DIGITAL TWIN (CDT) – A FRAMEWORK FOR A GENERAL CONTRACTOR

Taking into account the value chain management, construction companies are seeking to adopt technologies to increase their profits and add value to their customers while reducing the costs of implementing such technologies (Boje et al. 2020). Hence, considering the implementation costs, a general contractor can benefit from a CDT during its construction activities to increase its benefits and decrease the risks, and at the delivery and handover stage to increase its customer satisfaction. The capabilities of DT bring value to building and society levels by reducing energy waste and carbon emission and developing more sustainable strategies. A CDT of a physical entity starts its life before or during the construction phase. In other words, the focus of a CDT is on the pre-construction and construction phases of a project or physical entity. The construction phase of projects takes up to 40% of the total budget and due to the existence of several problems in this phase, such as site conditions, inappropriate construction methods, mistakes during construction, poor project management, etc., (Esmaeili and Kashani 2021), projects cost overruns are inevitable and will lead to higher budget requirements. Numerous challenges occur during this phase, and the general contractor, as the main player in this phase, needs proper management and solutions to address these difficulties. As BIM models are not updated until the handover of the project and they also have limitations in real-time data synchronization and model updating, a CDT fills this gap by integrating the physical entities, virtual models, analytics and prediction capabilities, and services by real-time or near-real time data connections during the construction phase prior to the handover. Hence, while benefiting from the CDT's advantages, the general contractor plays an important role in filling this gap and meeting the owner's requirements regarding the construction history. Therefore, a framework for the general contractor to use a digital twin or a system of digital twins during the construction phase and implementing it for different applications can help the general contractor in managing construction products and processes and mitigating aforementioned gaps.

Considering the literature review from various domains and the current research landscape, this section proposes a comprehensive framework for CDT deployment. It highlights the CDT framework layers and components, allowing contractors to take advantage of DT and its uses during the construction phase. This is done by evaluating DT uses and its components and features from various domains such as manufacturing, systems engineering, oil and gas, construction etc., and superimposing them into a CDT, enabling a general contractor to deploy it in his desired areas of application. As mentioned before, the three main components of a DT are "the physical," "the virtual," and "the link," each of which has specific features and abilities. The physical senses, monitors, and actuates while the virtual simulates, predicts, optimises, and acts as an agent by embracing the data through its abilities of BIM, IoT, data linking, and knowledge storing (Boje et al. 2020). The superiority of digital twin lies in the relationship between the physical world and the virtual world and their closed-loop control system where the virtual world is updated constantly via real-time data to represent the current status of the physical world and enables the analysis, diagnosis, optimization and prediction of the real world (Pan and Zhang 2021; Rausch et al. 2020) and provides various services and applications. Hence, a DT should not be viewed as an extension of BIM that is integrated with sensing and monitoring technologies (Sacks et al. 2020). Accordingly, the proposed 3-dimensional DT in the literature is expanded in this study by adding three more components namely: data, prediction and analytics component, and service to properly serve the purpose of a CDT in terms of ingesting required data, analysing and predicting project/site evolution, and providing services.

3.1. CDT Components

3.1.1. The Physical

The physical in a CDT is a broad term for “Physical Entity,” “Physical Environment,” “Physical Process,” etc. It is where sensing, monitoring, and actuating happens. Generally, the “object layer” consists of the desired physical component of a CDT, the essential site components, and the sensing system (Kan et al. 2018). Construction sites, buildings and physical assets, human and material resources, equipment and machinery, and other physical entities related to the construction processes are samples of the physical component. The commonality in these entities lies in their physical, real-world existence. The means for sensing, monitoring, and actuating such as sensors, IoT devices, laser scanners and communication infrastructure and generally the sensing system to be leveraged onto the physical components are defined in the physical category (Kan et al. 2018; Zheng et al. 2019). The physical component can be referred to as the “physical entity” or the “physical twin” regarding it has been twinned or not, i.e., whether or not the virtual and physical states are synchronized (Jones et al. 2020). “Physical Environment” (also referred to as “real-space” and “real-world”) is the real-world space within which the physical entity exists, and “Physical processes” are those activities being performed by the physical entity in the physical environment (Jones et al. 2020).

3.1.2. The Virtual

Virtual models are the real-time digital equivalents to physical entities that reproduce their geometrical and functional aspects, i.e. geometries, properties, behaviors, and rules (Kan and Anumba 2019; Qi et al. 2021). “Virtual Entity,” “Virtual Environment,” and “Virtual process” are among the characteristics of “The Virtual.” “Model,” “cyber,” “object,” etc., are terms used in literature to mention the virtual entity. The term ‘virtual twin’ is used when the virtual entity is twinned (Jones et al. 2020). Visualization is the essence for real-time monitoring of the physical products and processes (Qi et al. 2021). “Virtual environment” mirrors the physical environment and exists within the digital domain. Similar terms such as “virtual-space” and “virtual-world” are other terms used in place of virtual environment (Jones et al. 2020). The virtual environment is used for modelling, simulating, optimizing, predicting, and visualizing the physical reality. These activities that are performed using the virtual entity within the virtual environment are called virtual processes (Jones et al. 2020). Various software environments and platforms are available and could be utilized to create virtual entities, virtual environments, and virtual processes.

3.1.3. The Link/Connection

As a central component of a DT, the link establishes the ability for data flow from the physical space to the virtual space and information flow from the virtual space to the physical space (Grieves and Vickers 2016). Data connection between different parts of a DT, e.g., physical to virtual, virtual to prediction and analytics, etc., is enabled by various technologies such as RFID, sensors, IoT, wireless sensor network, communication networks, protocol technologies (Qi et al. 2021). This connection enables physical and virtual twinning, and the data gathering that is the basis for analysis, simulations, predictions and services. According to the various forms of data (data from physical world, data from analysis, data from virtual models, etc., see section 3.1.4) different connections between the CDT components can be established, i.e. connections between the physical, the virtual, the prediction and analytics component, and the service tier. Figure 6 illustrates the connection between different elements of a CDT.

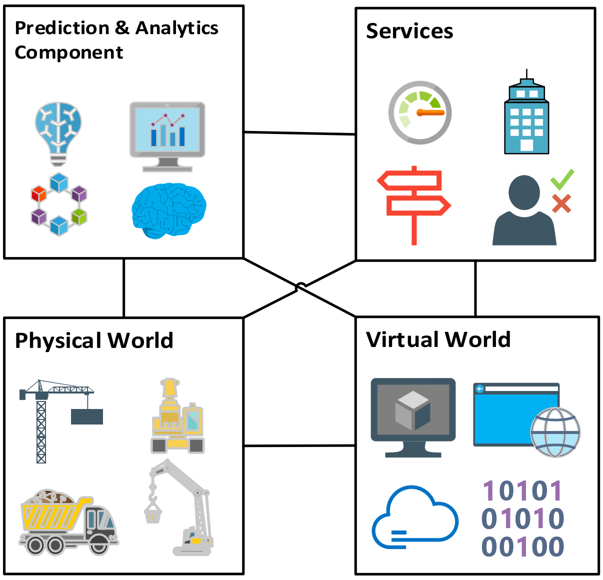


Figure 6- Connections between the CDT components

3.1.4. The Data

Data is one of the key drivers of CDT that comes from multi resources and in the form of multi-temporal and multi-dimensional data, i.e. heterogeneous data (Qi et al. 2021; Zheng et al. 2019). With the availability of low-cost sensors, a data flow from the physical space to virtual space in real-time or near real-time will be possible. Data could be obtained from physical entities or services, produced by virtual space (Zheng et al. 2019), provided as knowledge by domain experts and data analysis, or generated by data fusion (Qi et al. 2021). Based on the specific application of a CDT and enabling technology for data acquisition (sensors, IoT devices, etc.), the data can be different in types or formats. Data can be synchronized in two ways, real-time synchronization and no real-time asynchronous. The intermediate database is used to transmit data to the virtual space in no real-time asynchronous mode of data transition (Zheng et al. 2019).

Data can be acquired from or generated by the physical (entity, environment, process) or the virtual (entity, environment, process). Therefore, a CDT contains physical construction data and virtual construction data. Physical construction data includes site sensing data, site monitoring data, and lifecycle data (Boje et al. 2020). Table 2 shows samples of physical construction data types related to each activity.

Table 2- Physical construction data types and their samples

Physical Construction Data	Data Samples
Site Sensing Data	Sensors readings, location of materials and equipment, weather data, displacement data, volume data, temperature data, humidity data, load and pressure data, contact data, strain data, rotation data, laser scanning, topography, photogrammetry, videos, images, RFID
Site Monitoring Data	Schedule progress, resources availability, employee safety, construction quality, structural health, construction current status, worker and equipment performance monitoring, material tracking, temporary structure monitoring, construction component placement, site changes detection, machinery operations monitoring, non-conformities rectification
Lifecycle Data	Supply chain information, material compositions, hazardous material information, material specifications, equipment installation and operation information

Virtual construction data can include building model data, simulation data, status data, prediction data, and optimization data, each of which has specific types and samples (Boje et al. 2020). Table 3 shows samples of virtual construction data types and their corresponding samples.

Table 3- Virtual construction data types and their samples

Virtual Construction Data	Data Samples
Building Model Data	As-designed product models, As-planned process models, nD BIM models, project documents, schedules, procurement, invoices/statements and payments
Simulation Data	Clash detection, construction/deconstruction simulation, costing
Status Data	As-built product models, As-performed process models, Discrepancy analysis between as-

Virtual Construction Data	Data Samples
	designed/as-planned and as-built/as-performed products/processes
Prediction Data	Cost forecast, energy demand, traffic congestion, weather forecast, duration forecast, time/cost deviations forecast
Optimization Data	Optimized scheduling, dynamic resource allocation, cost reductions, energy usage, project duration

3.1.5. Intelligent Component

Integration of DT with big data analytics, cloud computing, simulation, AI and machine learning, mobile Internet and other technologies enables the co-evolution between the physical and virtual spaces (Qi et al. 2021). To extract information and insights (i.e., knowledge) from the massive volume of data produced during CDT usage, advanced data analytics techniques are needed (Qi et al. 2021). For instance, supervised and unsupervised learning are Machine Learning classes to make predictions or extract patterns from the existing data. In supervised learning algorithms such as Linear Regression, Logistic Regression, Convoluted Neural Network, etc., the algorithm predicts an answer, checks the predicted answer with the correct one to evaluate whether it is right or wrong and tries this process to reach the optimal state and minimize the cost function, loss, or error. In unsupervised learning such as Recursive Neural Network (RNN), the process runs recursively to recognize a pattern in data. These algorithms do not require labeled samples; hence, no prior knowledge is needed. Deep learning, a subset of machine learning, is also getting momentum since its data analysis algorithms can be performed in the cloud. The data analysis performed by this component is crucial to support monitoring, diagnosis, prediction and optimization. For data processing, Cloud, Fog, and Edge computing technologies can be deployed (Pires et al. 2019).

3.1.6. Service

The ultimate goal of a CDT is to provide value-added services and benefits. The construction phase is prone to numerous challenges of different characteristics. CDT's services such as site monitoring, optimized construction logistics, quality assessments, and safety management are among various services that can mitigate such challenges. The combination of the physical, the data, and the virtual via using the intelligent component as a CDT implemented during the construction phase has multifold benefits or "services." These services can be further divided into benefits and services during the construction stage, and those benefits manifest themselves at the delivery and handing over stage or dismissal. Table 4 enumerates some of these services.

Table 4- CDT benefits and services

Service Type	**Benefits/Services/Applications**
During Construction	Site progress monitoring (Boje et al. 2020) Clash detection (Boje et al. 2020) Site monitoring Data visualization (Boje et al. 2020) Optimizing construction logistics (Boje et al. 2020) Site planning Safety management (Kan et al. 2018) Integrated and adapted scheduling (Boje et al. 2020) Construction product quality assessment Supply chain management (Boje et al. 2020) Construction process control Sustainable building practices Tracking of changes and model updates Design office and construction site collaboration and information exchange Enhanced communication and coordination (Kan et al. 2018) Construction machinery operation (Kan et al. 2018) Interaction with the user via AR/VR technologies (Pires et al. 2019)
Handing over or Dismissal	Integrated handover to operation phase (Boje et al. 2020) As-built documentation (Boje et al. 2020) Lessons learned and historical data to experts or designers for better future design (Barricelli et al. 2019) Integration with other DTs

As long as the services are concerned, improving the integration and automation would enhance the construction services (Boje et al. 2020).

3.2. CDT Framework

Designing a Construction Digital Twin demands holistic thinking to clarify the ontological aspects of DTs and the necessary elements for the execution phase, DT's enabling technologies and their relationships, specific features and functions of each element and related technologies, and the function of the CDT as a whole (Sacks et al. 2020).

The main barrier to develop a DT is the system architecture and reliability of the DT (Liu et al. 2021). Building upon extant works on DT in various fields, e.g. Barricelli et al. (2019), Boje et al. (2020), Lu et al. (2020), Sacks et al. (2020), a CDT framework was developed in this research that serves as a roadmap to general contractors in order to implement a CDT during the execution phase of their projects and take advantage of CDT's ample benefits and applications. Figure 7 depicts a CDT's main tiers and information flow and their related enabling technologies for the CDT implementation. In the coming paragraphs, each tier will be explained in detail. Finally, a soil management case study as the application of a CDT will be implemented via the proposed framework to validate its applicability and effectiveness and detect the barriers to implementing such applications using the proposed CDT framework.

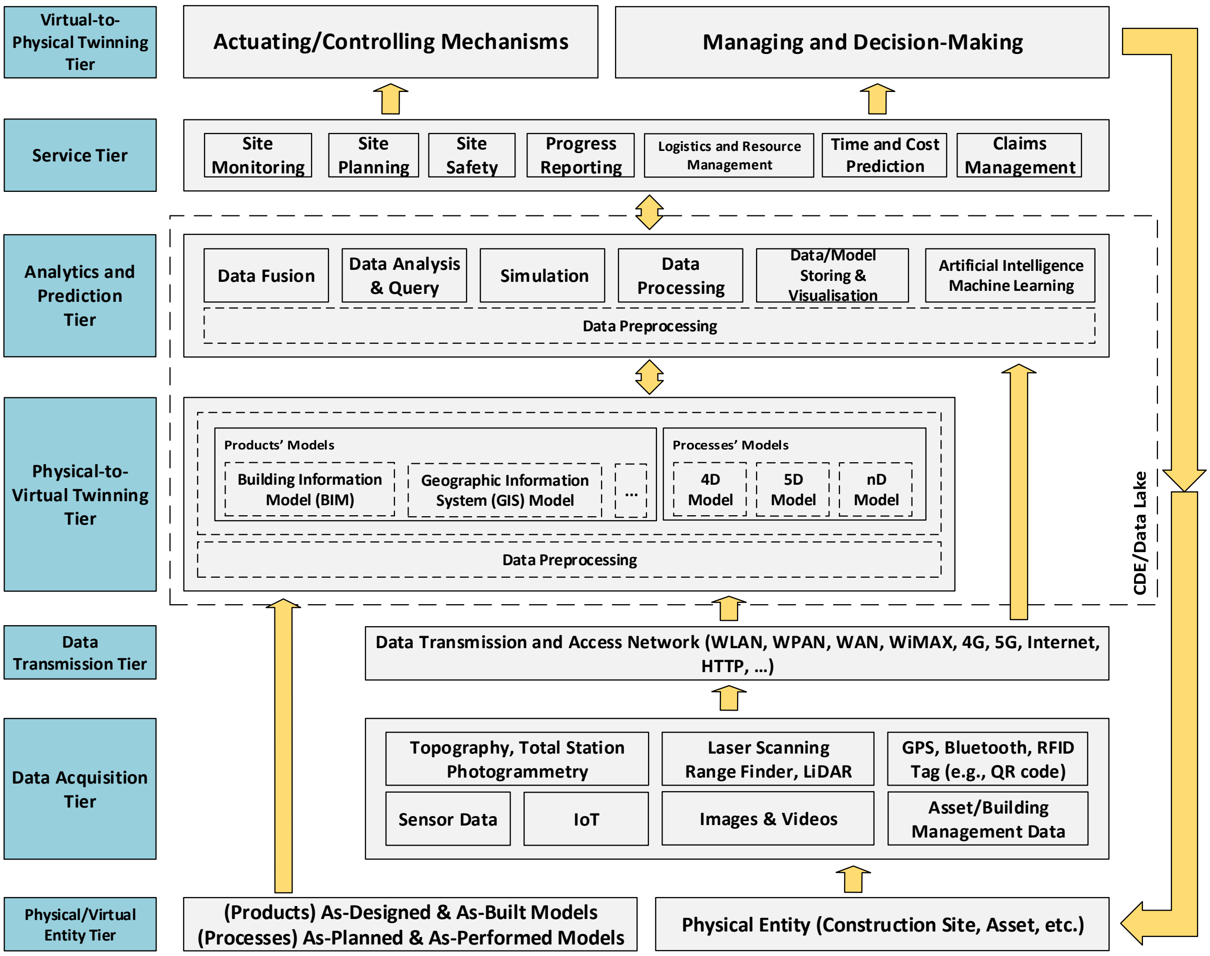


Figure 7- Construction Digital Twin (CDT) framework

3.2.1. Physical/Virtual Entity Tier

This tier represents the physical object that already exists or is going to be built and the virtual entity that will be created through digital modelling. Modelling is the basis of implementing a DT in practice (Tao et al. 2019). Digital modelling in the physical/virtual entity tier is a collective term for the various digital models of products or processes that represent physical reality. A digital model is a geometric representation of an object or set of objects, generally in 3D, made on a computer with other purposes such as analyzing, simulation etc., beyond geometrical representation (Tchana et al. 2019). According to Barricelli et al. (2019), two lifecycles can be imagined for a DT. One is when the physical object does not exist, so the design process simultaneously conceives both the physical object and its DT. In the construction and BIM terms, it can be seen as BIM or nD BIM models, which are considered as-designed and as-planned models. As-designed and as-planned models, or in short, Project Intent Information (PII), contain all information about the future state of a building that has not been built yet (Sacks et al. 2020). As the names are self-explaining, as-design models express the design of the building or parts of it, and as-planned models express the construction plans for building them (Sacks et al. 2020).

In the other lifecycle of a DT, the physical object already exists, but there is no DT (Barricelli et al. 2019). In this case, an existing as-designed or as-planned model (BIM model etc.) can be adopted and updated to represent the current status of the physical reality, or a new model can be prepared from scratch. As-built and as-performed models are Project Status Information (PSI) that manifest all information about the past state of a building. As-built and as-performed record the information about what was done and how it was done, i.e. the physical building and its construction process (Sacks et al. 2020).

3.2.2. Data Acquisition Tier

Considering the philosophy of a DT to gather data in real-time and update the virtual model of a physical component, data acquisition is a vital part of a DT. Based on the scope and focus of a specific DT/CDT, there is a range of technologies available for data gathering including, smart sensors, IoT devices, RFIDs, GPS, photogrammetry and laser scanners, etc. (Lu et al. 2020; Sacks et al. 2020). However, due to the complex nature of the construction projects and the specific configuration and purpose of a CDT, finding a suitable tool for data acquisition may be challenging. In addition, the adopted technology might not guarantee a complete autonomous gathering and transmitting of the data.

3.2.3. Data Transmission Tier

A secure, reliable, and high-speed network is required for real-time or near real-time data exchange in a DT (Sepasgozar 2021). The locally gathered data needs to be transmitted to upper layers for updating the digital model based on the status of the current physical entity. The data is also needed for data/model integration to serve various analytics and visualizations. This communication infrastructure can adopt a wide range of technologies such as WAN, WLAN, 4G/5G, Bluetooth, ZigBee, GSM, LTE-M etc., as communication and wireless technologies, and HTTP, MQTT, mDNS, DDS, CoAP, AMQP, XMPP, etc., as layer protocols applications (Liu et al. 2021; Lu et al. 2020). Each data transmission technology has its own capabilities and limitations. Security measures must be considered in developing a DT/CDT as gathering and transmitting data are suspect to hacking and viruses, especially when dealing with private, confidential and valuable information (Barricelli et al. 2019).

3.2.4. Physical-to-Virtual Twinning Tier

Unlike BIM models, DT is a living and evolving model rather than a static model, and it is connected to its physical twin and automatically updated as the physical twin changes, i.e., synchronicity between the virtual and physical twins (Sacks et al. 2020). The twinning process is bond with continuous interaction, communication, and synchronization between the virtual twin and its physical twin as well as the surrounding environment by continuous and real-time information exchange (Barricelli et al. 2019). The twinning process occurs within a frequency called Twinning Rate (Jones et al. 2020). The real-time updating, i.e., high twinning rate, enables the DT to be constantly aware of what is happening in the physical world (Barricelli et al. 2019). This capability of DT is enabled by real-time data uploading and data storage technologies (Barricelli et al. 2019). Therefore, this tier aims to continuously update the as-built or as-is and as-performed condition of the physical twin.

In this tier, the actual (PSI) and the intended (PII) could be compared. Value judgments are performed in each specific case with a threshold value to determine the degree of variation of PSI from PII (Sacks et al. 2020). The output of this "Evaluate Conformance" function is knowledge about the current status of the project termed as Project Status Knowledge (PSK), and AI methods can be adopted to enable the automated conformance evaluation (Sacks et al. 2020).

3.2.5. Analytics and Prediction Tier

The analytics and prediction tier contains the required functions for data processing, data fusion, data analytics, simulation, data storing and visualization, AI, ML, etc., to update the model based on the real-time data and also feed the upper tiers and support them in decision makings and actuating mechanisms (e.g., sounding an alarm, closing/opening a window, etc.) Enabling technologies such as smart sensors and IoT devices, communication technologies, advanced storage and cloud computing address the Industry 4.0 needs to acquire, transfer, and store data to feed this tier with real-time or near real-time data. However, the challenge is how to properly integrate technology and use the acquired data from various sources to enable handling Big data (data with high volume, high speed, and high heterogeneity) and then apply analytics, learning algorithms, and visualization techniques to detect patterns, perform predictions, and extract valuable insights and knowledge (Figueiras et al. 2021). Collecting data is a key aspect of Digital Twin. The collected historical data stored within the virtual environment can be reused for future virtual processes. This capability enables the Digital Twin to learn from its past (Jones et al. 2020). Managing raw data coming from different sources and generating valuable insights and knowledge will support improvements and optimizations in processes and enable the development of new data-driven services and applications that eventually contribute to predictability and real-time or near real-time reactivity (Figueiras et al. 2021). As shown in Figure 7, the analytics and prediction tier encompasses various components that can be common among most applications or specific to a particular use of CDT. For example, data processing tasks such as aggregating, transforming, mapping, and reducing data are done in the data processing component. Then, data analytics and visualization techniques can be performed over the processed data (Figueiras et al. 2021). Alternatively, simulation techniques are deployed to evaluate various "what-if" scenarios about the real-world system. Also, data analytics, such as knowledge discovery, machine learning, data mining, etc., aim to create value from the processed data by providing insights to related stakeholders (Figueiras et al. 2021). Sepasgozar (2021) defines different levels of capabilities for a DT, such as description, prediction, and prescription, emphasizing that a DT can learn and suggest new scenarios to develop construction processes. DT's predictive and prescriptive capabilities bring value in reducing downtime, breakdowns, costs, energy waste, and achieving Sustainable Development Goals (SDGs) (Sepasgozar 2021). Lu et al. (2020) introduce Knowledge Engines (KEs) for products and processes to describe their dynamic conditions by data integration and analysis and using intelligent functions (e.g. AI/ML). These KEs are domain and purpose-specific, for example, a pump in its operation phase, and they utilize the data integration capability of a DT to deliver better-informed services.

3.2.6. Service Tier

The service tier in the CDT architecture is the tier that interprets or visualizes the Project Status Knowledge (PSK) or knowledge from a KE and triggers the decision-making action or an actuation function. It is the ultimate tier that provides the output of a CDT for more informed decision-making or enabling the commencement of another function. Being fed by physical-to-virtual twinning and analytics and prediction tiers, this tier can provide various services such as site monitoring, site planning, claims management, etc., to the general contractor throughout the construction phase. Based on the scope and focus of a CDT, the expected outputs are different; however, a group of CDTs with different purposes could be connected together to work as a CDT system.

3.2.7. Virtual-to-Physical Twinning Tier

The virtual-to-physical twinning tier is the top tier in the CDT architecture that applies the decisions or actuates pre-defined mechanisms fed by the service tier, i.e., feedback from cyberspace to physical space. The output of this tier directly influences the physical twin that completes the closed-loop control system. Therefore, the information flow from the virtual to physical entity happens in this tier, i.e. virtual-to-physical twinning, and it completes the bi-directional communication between the virtual and physical twins. Based on the specific service, data analytics and prediction outputs, and updated models, the virtual acts on the physical and changes its status. In the next loop, the virtual catches the new changes in the physical and updates its status accordingly and through data analytics and acting on the physical twin completes the next loop. Each loop occurs according to the twinning rate that has a direct relation to the requirements of the desired application of the CDT and service. As it is shown in Figure 8, the CDT framework presented in this study, enables a closed loop control system that embeds a Plan, Do, Check, and Act (PDCA) cycle. Sacks et al. (2020) emphasized on enabling a closed loop model of construction control via DT’s ability to stream real-time information from the construction project through a PDCA cycle.

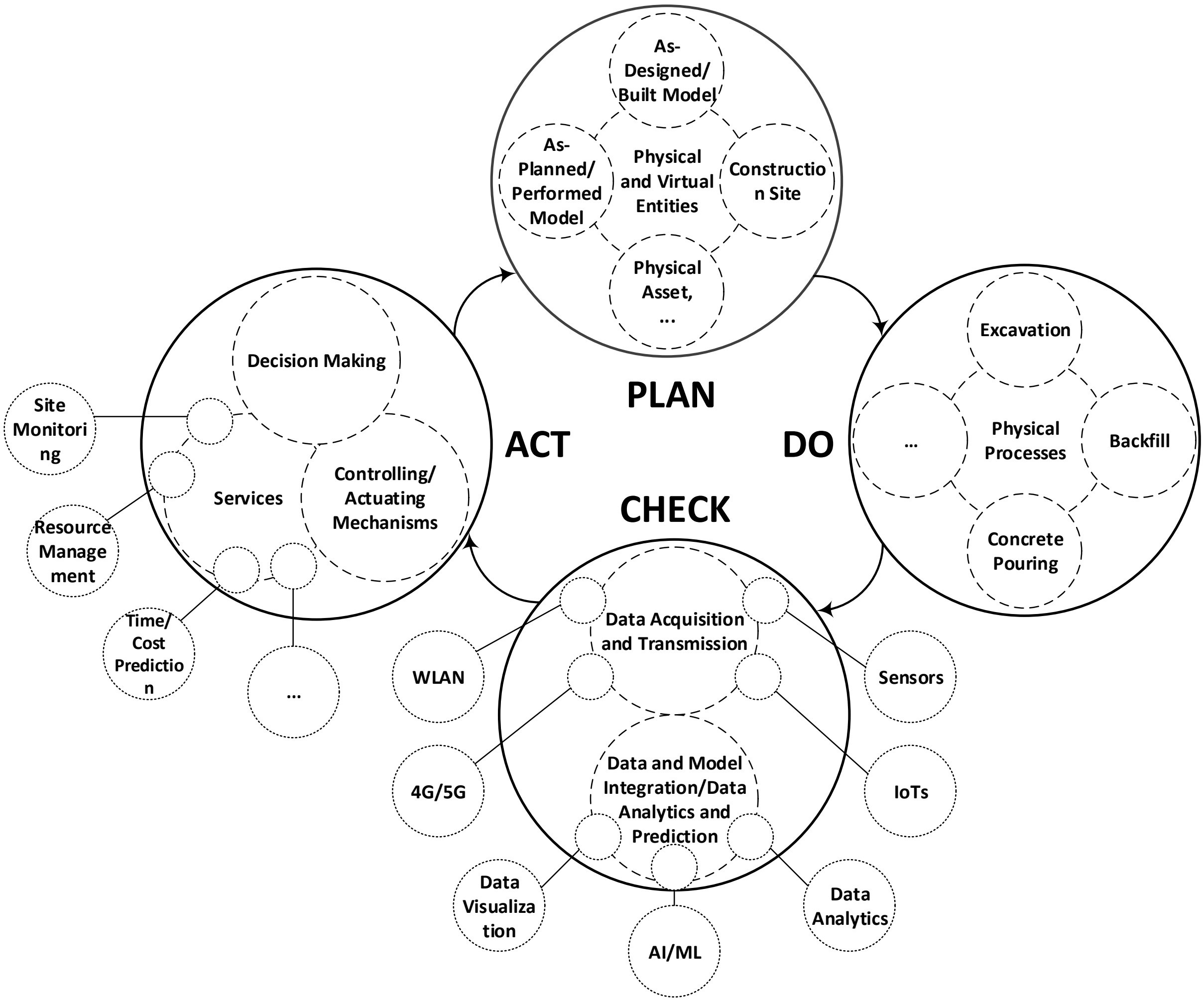


Figure 8- Enabled PDCA and closed-loop control system in the CDT framework

This page is intentionally left blank

4. CONSTRUCTION DIGITAL TWIN APPLICATIONS AND CASE STUDIES

This section introduces practical uses of a CDT to unveil the potential benefits of a CDT in managing real-world problems at the execution phase of projects. A soil management case study will be implemented using the developed CDT framework to validate the applicability and effectiveness of the framework and detect the barriers of implementing such applications using the proposed CDT framework.

One of the main issues in each construction site is managing the soil produced at the initial construction phase, where site rough grading and excavations are performed for mobilization and site preparation before the main excavations and during the construction phase for excavation (cutting) or backfill (filling) of the buildings, equipment foundations, roads, etc. Earthwork activities are pretty common in construction projects. For example, an assessment of the construction projects of 10 industrial buildings showed that the average costs of earthwork operations (excavating, backfilling, loading, transferring, and unloading soil) equals 8 percent of the total contract sum. In this assessment, the cost of supplying soil is not considered; therefore, the average cost of earthwork operations would be higher when soil supply is anticipated. Needless to say that in road projects, earthwork operations consist most of the contract amounts. In addition, from the project’s commencement, earthwork operations usually exist up to the end of the construction phase, making it not a temporary short-term process but rather a long-term process that requires proper management and time-to-time monitoring.

Even in small and medium construction projects, the amount of soil produced is considerable from the commencement of the construction phase to the project’s closeout. For example, a pump station construction project with a contract value of 2 million Euros (C contract) produces thousands of cubic meters of soil. This is the case for a variety of project types, from building and residential projects to petrochemical and refinery projects. In building projects, the soil is produced from rough grading for site preparation and excavation of buildings’ foundations. In road projects, the soil is produced when cutting is needed for establishing the designed road profile. In oil and gas projects, the soil is produced from a variety of activities such as rough grading for site preparation and piling activities, excavations for buildings and equipment’s foundations, cuttings and excavations for roads’ profiles inside the refinery, pipe trench excavations for laying pipes, etc.

There might be a specific place inside the project’s site to serve as a soil deposit for the excessive excavated/cut soil or a soil borrow for the backfills/fills. It also might be outside of the project’s site in specific places in urban areas which need specific permits and environmental considerations; hence, it demands even more attention in terms of monitoring and management. As shown in Figure 9, depending on the type of project, i.e. building, road and infrastructure, power plant, oil and gas, etc., a variety of sub-contractors might be present at a construction site and work simultaneously. Depending on the subject and scope of their contracts, they might produce extra soil, need soil, or both to perform their activities. As they are working under the umbrella of a general contractor, the general contractor needs to manage the soil in terms of defining its optimal location, monitoring its status in real-time or near real-time, predicting necessary volumes for front and upcoming activities, managing machinery (e.g., trucks and excavators), transferring the soil from the temporary location to the permanent location for abiding urban and environmental rules, and other related issues which impact the final expenses of the general contractor and the cost, time, quality, and safety of the project.

Another challenge is that a general contractor may encounter claims from its sub-contractors regarding soil and its various aspects, such as soil supply. As earthwork operations cost is a considerable portion of contracts sum, improper soil management can lead to huge expenses and non-compensable project cost overruns that the general contractor must shoulder. On the other hand, efficient soil management can help prevent such claims and facilitate the ongoing or upcoming activities that need soil deposit/borrow.

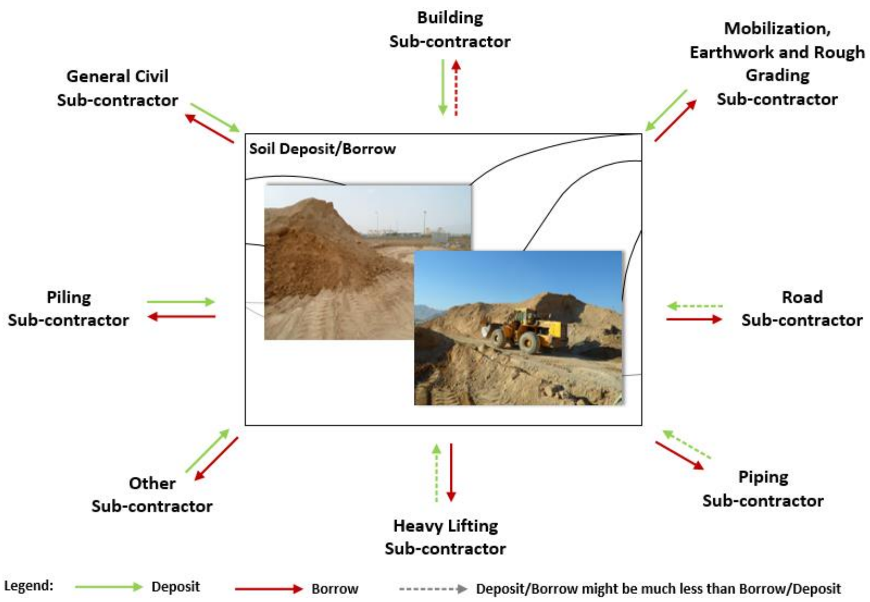


Figure 9- Probable sub-contractors at a construction site and their soil deposit/borrow activities

Following the introduction of the CDT framework, in the next sections, the case study of soil management will be implemented step by step based on the CDT framework to expand the knowledge of using digital twins during the construction phase of projects and help general contractors and practitioners in managing different aspects of their projects.

4.1. Physical/Virtual Entity Tier

As depicted in Figure 7, the first step in establishing a CDT is identifying the desired physical objects that the CDT would replicate and developing their digital models, i.e., product or process models. Since a DT is meant to be a reliable virtual replica of a physical reality, a high-fidelity virtual model is the core of any DT. Fidelity indicates the number of parameters (types of data, information, and processes), their accuracy, and the level of abstraction that are transferred between the virtual and physical twin/environment (Jones et al. 2020). Fidelity levels (low, medium, high) demonstrate how close the virtual and physical twins are aligned. A comprehensive understanding of the desired physical component is crucial in developing a high-fidelity virtual model (Qi et al. 2021). Hence, in this case study care was taken to develop a high-fidelity digital model with all necessary aspects regarding the management of the soil in terms of cut/fill volumes, location of the construction site, location of soil deposit/borrow, access routes, and distance from the construction site to the soil deposit/borrow location.

Taking into account the enabling technologies and tools for implementing a CDT, in this step, Autodesk Revit was used as a tool for digital modelling of the as-designed product and as-planned process. In addition, ArcGIS and Autodesk Infraworks are used for developing the digital models of the as-built models regarding the construction site and soil deposit/borrow locations.

The case study for demonstrating the CDT application for soil management was a hospital project in Milan, Italy. A 3D view of the as-designed model is shown in Figure 10.

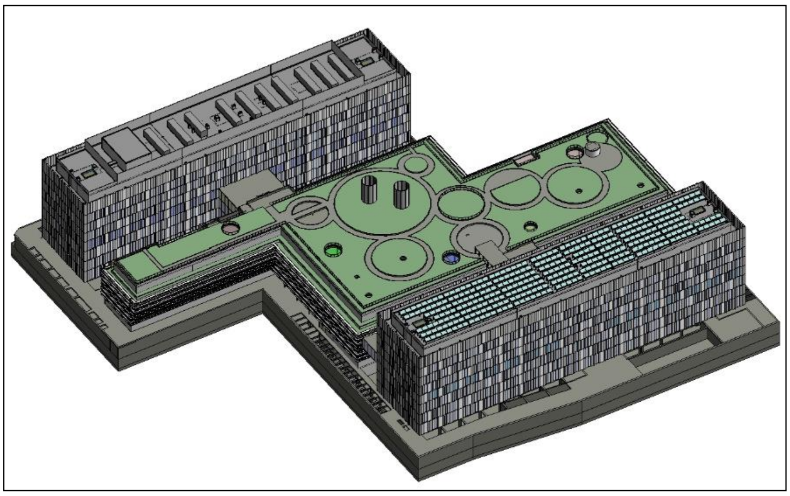


Figure 10- 3D view of the hospital's as-designed model

As it is evident from the construction site plan (see Figure 11), the project is located inside the city and surrounded by built assets in an urban area. Hence, environmental rules and sustainability issues are of great importance that the stakeholders of the project, especially the general contractor, must take into account.

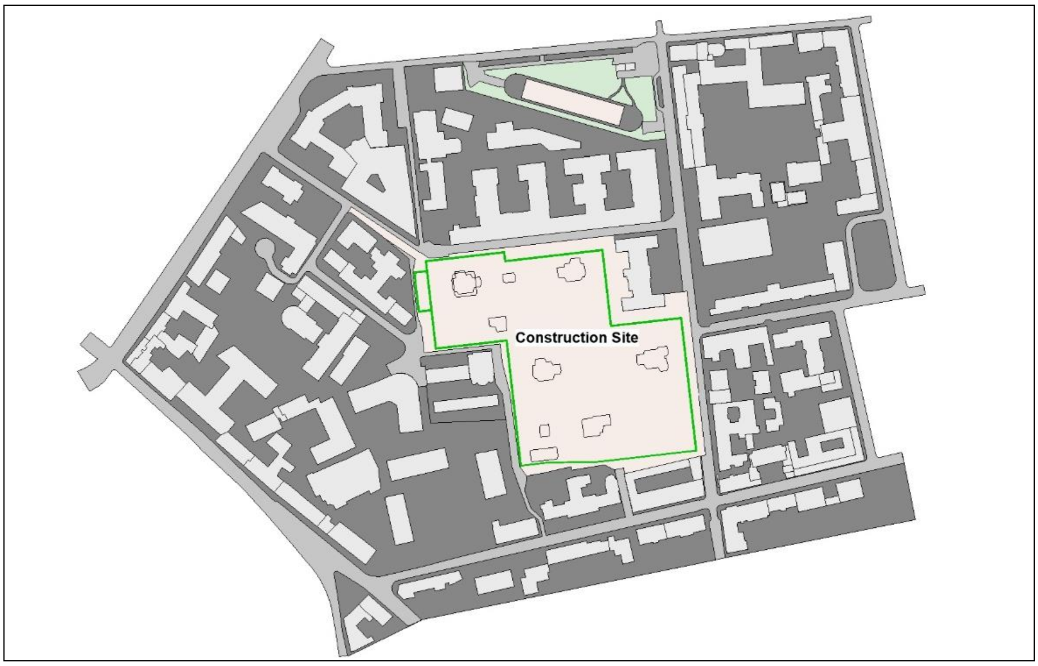


Figure 11- Construction site plan

A section of excavation is shown in the as-designed site model (see Figure 12), showing different levels and excavation depth. According to the BIM model, the earthwork operations in this project are at least 182,000 m3 of soil regardless of the fill/backfill volume and other activities. Hence, such a considerable volume of soil which is almost prevalent in every construction project, demands proper management of such materials and construction machinery.

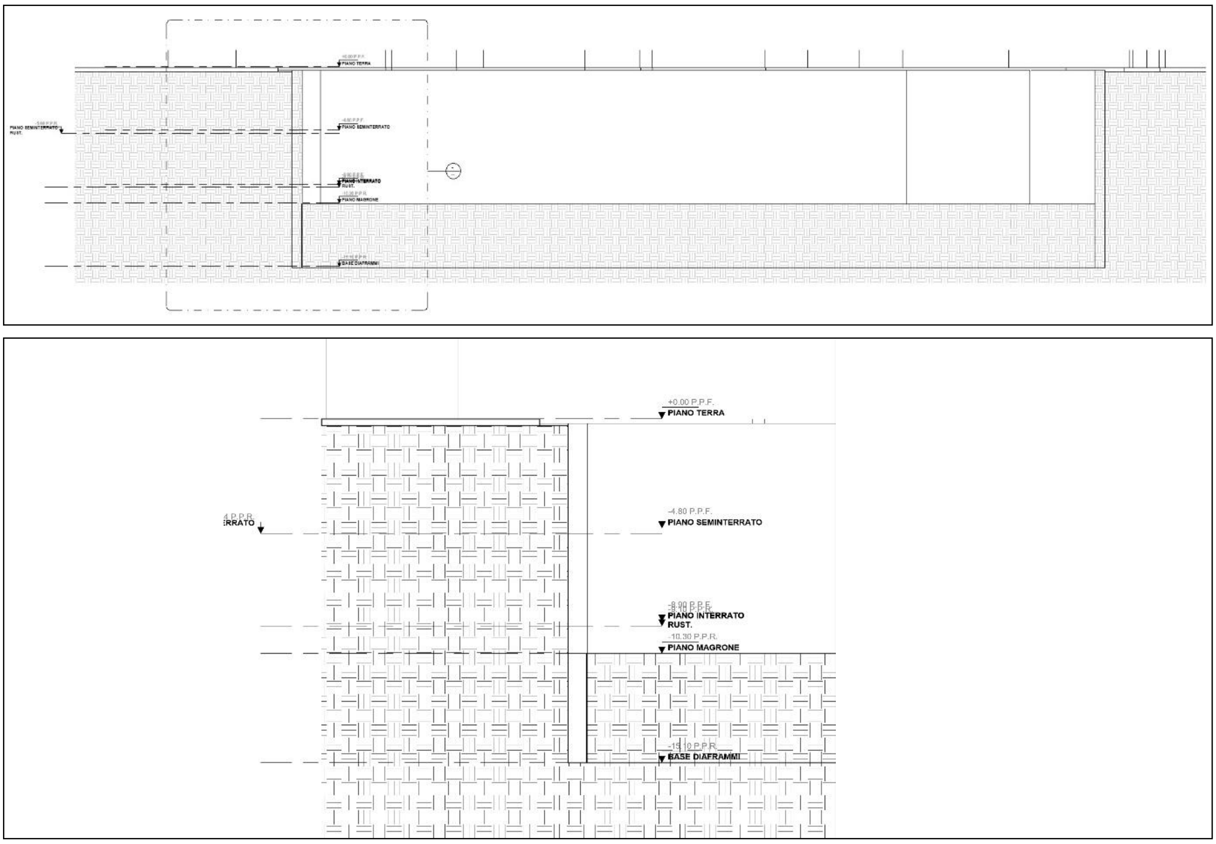


Figure 12- As-designed site model

Figure 13 shows the as-planned excavation model. Like as-designed models, this model is issued to the construction site to demonstrate the excavation zones and execution phases and sequences.

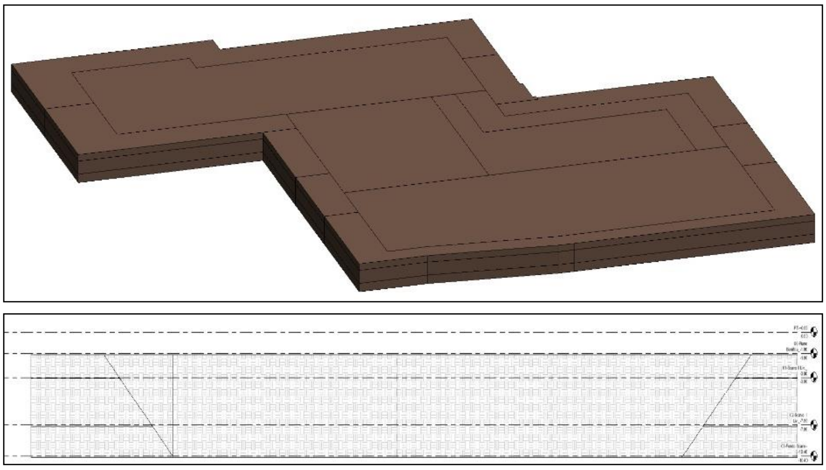


Figure 13- As-planned excavation model

In this CDT application, the physical component that is going to be virtually replicated is the soil deposit/borrow, which plays as a source of soil supply for fills and backfills, and a dumping point for the extra soil from cuttings/excavations. This component has a specific location inside or outside the construction site. Another important feature of the soil deposit/borrow is its volume, which will be monitored and calculated in real-time or near real-time to inform the decision-makers about its status.

The as-built (as-is) model of the physical twin is shown in Figure 14. Once this as-built or as-is model is created, it is continuously being updated at the “physical-to-virtual twinning” tier as data is acquired from the construction site and the physical entity.

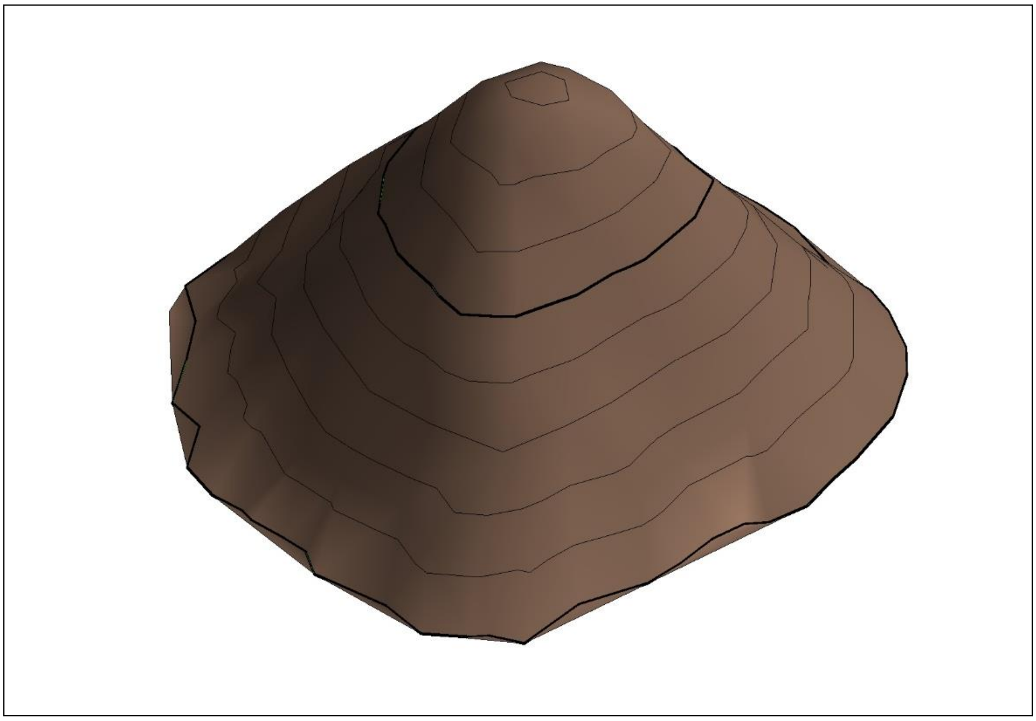


Figure 14- As-built (as-is) soil deposit/borrow model

The other physical component in this CDT is the construction site. As the earthwork operations progress, the site model can be updated and the volume of the performed excavation and remaining excavation can be calculated. This can also be used to compare the actual earthwork operations progress with the planned progress to see whether the earthwork activities are ahead of schedule or behind schedule. In this CDT application, the soil deposit/borrow was considered as the physical entity, and the construction site was only considered to reflect the location and distance of the soil deposit/borrow related to the construction site. Consideration of the construction site as the other physical entity can be the subject of future studies.

4.2. Data Acquisition Tier

Depending on the scope and focus of the CDT and the desired application, the selection of proper data acquisition tools varies. As shown in Figure 7, a range of tools and enabling technologies from smart sensors and IoT devices to most typical devices such as cameras and cell phones for images and videos can be deployed according to the CDT’s application requirements.

4.2.1. Capturing the shape and volume of the soil deposit/borrow

For the current CDT application in soil management, the first step in data acquisition is selecting the tools that capture the shape and coordinates of the soil deposit/borrow to calculate its volume. Alternatively, some sensors are able to calculate the volumes in real-time and report them back to a server or platform. For this purpose, different tools are introduced, and capabilities, capacities, advantages and disadvantages of each option are discussed in the following paragraphs.

4.2.1.1. Volume Sensor

This device enables the measurement of bulk solid materials and powders in silos or open bins. Different industries such as mining, cement and aggregates, petrochemicals, power-coal, fly ash, etc., can benefit from its features for their broad range of industrial applications. These sensors are equipped with antennas to transmit pulses and receive echoes of the transmitted pulses from the contents of the silo, bin or other containers. These antennas help to measure the time, distance, and direction of each echo. The processor embedded in these devices analyses the received signals and measures the material's level, volume, and mass. Figure 15 shows an example of these sensors called 3DLevelScanner.

Figure 15- Volume sensor (3DLevelScanner)

They can also be equipped to generate a 3D representation of the material by sending the information to an API for generating the 3D representation on a remote computer. These devices can be proper tools for material measurements in various containers, including silos, large open bins, bulk solid storage rooms, stockpiles and warehouses. In some models, based on the container’s height and diameter, the accuracy of the measured volume could be more than 95 percent. Choosing the proper location of these devices should consider every aspect of the storage and material type, angle of repose, maximum level of material, and any other considerations which may affect the scanner performance. Some PC software packages are also recommended for finding the optimal installation location and positioning. According to the installation position of the sensor and storage type (open bins, tanks, silos), one or multiple scanner systems can be used to provide accurate readings of the level and volume for the stored material based on two or more synchronized units. The system provides an analog output that represents the overall volume of the stored material. In the case of using multiple scanners, a controller is used to calculate volume and unify the visualization by integrating data from multiple scanners.

The level range or measuring range of these devices can be up to 70 m, i.e. the heights of the stored material could be up to 70 meters, and the diameter could be up to 15 meters. They can be fixed on the desired location by thread or flange. The measurement interval time starts from 2 seconds. The output can be a 4 to 20 mA signal transmitted to a remote computer using specific protocols and communications to generate the 3D model, or the volume could be displayed on an internal LCD display. Figure 16 depicts the configuration of volume sensors.

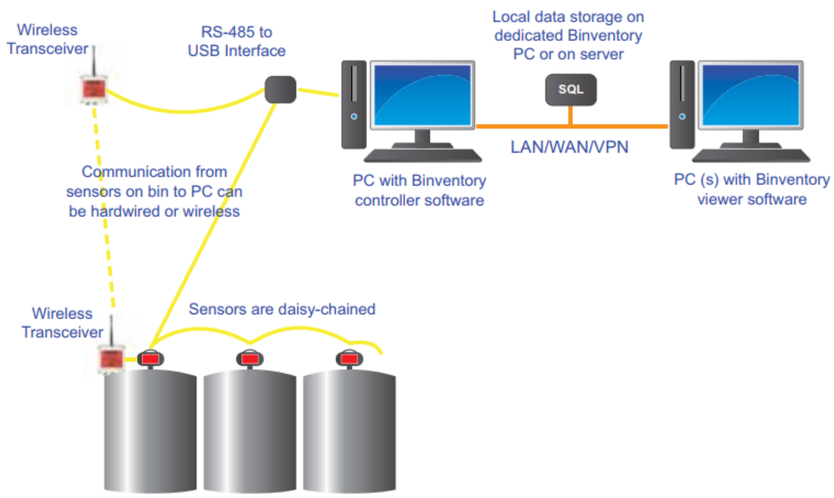


Figure 16- Volume sensor configuration

These sensors have their benefits and drawbacks. In the construction field, they are proper means for materials stored in silos, such as cement storage in concrete batching plants. However, they might not be the best alternatives for open space stockpiles such as soil deposit/borrow of the current application. Some of these advantages are enumerated below:

- Real-time and automatic data acquisition,
- Accuracy of volume measurements,
- Acceptable measuring range,
- Remote connection to existing control systems,
- Working with a variety of materials,
- Non-contact measurement suitable in contaminated or hazardous environments,
- Ability to work under harsh temperatures and dusty environments,
- Low maintenance fees.

They also have limitations and cons such as:

- Increased measurement error with increasing in stockpile diameter and height,
- Level range or measure range limitations,

- Not a suitable alternative for open space stockpiles,
- Additional costs, because they are not common means available in every construction site.

4.2.1.2. Laser Range Finder, LiDAR, and 3D Point Cloud

Laser range finders use laser technologies to calculate distance based on the flight time of transmitted pulses and the reflected pulses detected by a sensitive detector. This technology has been adopted to measure materials stockpile volumes. In these devices, a single operator needs to walk around the pile and mark instrument points that altogether will fully cover the surface. At each instrument point, all necessary points must be aimed and shot until the entire surface has been measured. Based on the gathered field data, the volume of stockpiles is calculated. The gathered data could also be transferred to a remote PC for processing. These devices can also be equipped with software and hardware technologies for topography or 3D digital representations of stockpiles.

These devices are advantageous in several ways, as enumerated below:

- No need for a reflector and crew to hold a prism, unlike traditional methods,
- Increased safety and accuracy in case of loose material where total station and prism is hard to use,
- Ability to work in different weather conditions and times of the year in handheld instruments,
- Acceptable time-efficiency (depending on the specific application),
- Easy staff training,
- Accuracy in measuring.

They also have some limitations and cons such as:

- Not being fully automated and need for an operator,
- Takes time for capturing and covering the stockpile data from different instrument points and calculating the volume,
- Using in a confined space with dangerous material may be impossible or need safety considerations,
- Need for maintenance,
- Additional costs, since they are not common means available in every construction site.

On average, the price of these devices ranges from 400 $ to 1800 $ depending on the brands and models. The necessary accessories add up to this price.

LiDAR (Light Detection And Ranging) sensors can also be used, and data gathering can be done like laser range finders by walking and scanning, or the scanner could be attached to a trolley, drone, pole, UAV or vehicle. With the LiDAR sensors, data can be acquired in real-time or near real-time, and the output could be acquired as a .csv file for further processing or sent to vendors' specific software for processing. Although it can be beneficial when dealing with large stockpiles, it might be expensive and even unavailable in severe weather conditions or specific times of the year when attached to aerial fly-over.

Three-dimensional (3D) point clouds have been recently used to achieve a more detailed spatial morphology of the stockpiles and consequently improve the measuring accuracy. 3D point cloud collects the dense 3D coordinates of the stockpile surface and its parameters such as height, floor area, etc. Then, the 3D point cloud data of the stockpile can be processed to compute its volume (Yang et al. 2020). 3D point cloud data can be obtained in many ways, such as depth camera, stereo camera, and laser scanner. Laser scanning techniques have better accuracy, robustness, and efficiency than image-based photogrammetry or manual survey methods, and they have been widely used for topographic measurements in recent years (Yang et al. 2020). Despite the speed and accuracy in volume calculation using 3D point cloud data compared with conventional methods, obtaining accurate quantitative parameters of stockpiles such as height, floor area, etc., is a challenge in this method (Yang et al. 2020). The irregularity of the stockpiles' shape, size, height, and base conditions makes it hard to measure the different stockpiles automatically and accurately (Yang et al. 2020).

4.2.1.3. Topography and Site Surveying Instruments

Topography and site surveying instruments are quite common tools in construction sites. These tools are being used as the project starts when the owner delivers the lands and surveying benchmarks to the general contractor. Site surveying instruments are necessary tools for specifying coordinates before, during, and even after performing activities. They have numerous uses to help deliver tasks and activities with coordinates and dimensions as specified in as-designed drawings or models, or to survey the as-built assets or as-is status of the site. Terrestrial geodetic methods using a total station are commonly used to obtain information about the stockpile and calculate its volume (Yang et al. 2020). In this CDT application, they serve as proper and available tools for capturing the shape and various points of a soil deposit/borrow and consequently the topography of the site and the stockpile. These devices can produce outputs in text formats that contain information about the instrument itself, measurement units and dates, surveyed points coordinates etc., or generate their specific file formats that can be handled via the provided software. The generated text output contains heterogeneous data and needs pre-processing in order to extract the surveyed points information and enable creating topography (this will be discussed in the physical-to-virtual twinning tier, see 4.4). In this method, sampling control points on the stockpile is done manually, and one needs to walk over the stockpile.

Among the benefits of these tools, some advantages are enumerated below:

- No additional costs due to availability in construction sites,
- High accuracy,
- Mobility and ability to being set up at different locations easily.

Regarding the soil management application of the CDT, they are prone to some limitations such as:

- Manual process and no real-time data acquisition,
- Possibility for decreased accuracy when working with loose material,
- Longer surveying intervals due to time and manpower-consuming processes.

4.2.1.4. Images

Images and videos are the most common sources of data (Sepasgozar 2021). Using images can be a suitable alternative for measuring the volume of stockpiles in open spaces such as soil borrow/deposit of the current CDT application. Based on the stockpile scale, drones or fixed cameras can be used (see Figure 17). With the installed camera solution, monitoring the inventory in real-time would be possible. Mobile devices applications can be used to take measurements on-demand regardless of the user's location or time preference. Some products such as Stockpile Reports offer hands-off data processing and cloud-based solutions. This service offers a minimum annual subscription of $20,000 to measure a certain number of piles of stockpiled material.

Figure 17- Using images for volume measurement (Adopted from Stockpile Reports)

This technology entails several advantages such as:

- Real-time or scheduled monitoring of the stockpiles,
- Reduced labour costs,
- Improved worksite safety,
- Remote connection.

It is also bound to limitations and disadvantages such as:

- Privacy and security issues,
- Relatively high costs of buying services,
- Need for labour when using drones,
- Need for data processing and volume calculation in case of not purchasing related services,

4.2.2. Detecting locations of soil carrying machinery

Real-time locating systems (RTLS) are used to identify the location of entities in real-time. Data collected from a real-time locating system (RTLS) could be integrated with digital twins (Ruppert and Abonyi 2020). Various studies in construction research and facility management have been conducted for resource tracking, i.e., personnel, equipment, and material. These studies proposed related technologies for indoor and outdoor environments to track the desired resources (Akhavian and Behzadan 2012). Ultra-wideband (UWB), radio frequency identification (RFID), global positioning system (GPS), ZigBee, and Bluetooth are among the technologies for tracking the location of resources. For example, Oloufa et al. (2003) used GPS and wireless communications to track construction equipment for safety management. In the current CDT application, vehicles carrying the soil can be tracked using real-time locating systems. Real-time tracking of the machinery carrying the soil and transmitting and sharing the acquired information during construction provides several advantages. Acquisition of equipment locations in soil movements during project execution can lead to higher project efficiency and cost savings. Productivity measurement and improvement, supply chain management and just-in-time delivery are among the benefits of using such data (Oloufa et al. 2003).

UWB RTLS technology is suitable for indoor localization considering its low energy consumption and accuracy compared to Bluetooth (Ruppert and Abonyi 2020). For outdoor localization, GPS can be a suitable alternative. GPS technology is a satellite-based radio-navigation system that offers several benefits and economic costs and is also usable in various areas for positioning purposes (Oloufa et al. 2003). Since the CDT application for soil management deals with outdoor localization, GPS is a suitable enabling technology to acquire real-time coordinates of the vehicles. The GPS system used for data acquisition can be as simple as a set of transceivers, GPS receivers, shields and microcontrollers, or it could be an advanced GPS platform such as GPS fleet tracking platforms that are available at various costs and rates.

4.3. Data Transmission Tier

The data transmission tier ensures the real-time synchronization of different parts of the CDT. IT technologies such as WLAN, 4G, 5G, HTTP and other communication and protocol technologies enable the real-time data and information exchange among the physical, the virtual, analytics and prediction, and services in the CDT. In addition to the IT infrastructure, a common data environment (CDE) or a data lake can be used to share digital models or store data that can be queried. The Internet and a CDE were used for data exchange and model sharing in the current soil management application.

4.4. Physical-to-Virtual Twinning Tier

In this tire, the data acquired from the physical entity is integrated with the virtual model and updates models to reflect the changes and status of the physical entity. Therefore, physical-to-virtual twinning happens in this tier as the virtual entity becomes synchronized with the physical entity. As the result of the twinning process, physical entity and virtual entity can be called physical twin and virtual twin, respectively.

Missing data and abnormal data might exist due to problems in the data collection and transmission process. To mitigate this problem, data pre-processing should be conducted for data cleaning (Kong et al. 2021). Before integrating data with the virtual model, the collected data might need preprocessing to make the integration possible. For example, when gathering data from topography and site surveying tools, the outputs of these instruments might contain additional data; hence redundant data needs to be omitted, and information about points coordinates to be extracted. Figure 18 illustrates the desired part of data that contains information and needs pre-processing and information extraction from the verbose output files of site surveying instruments.

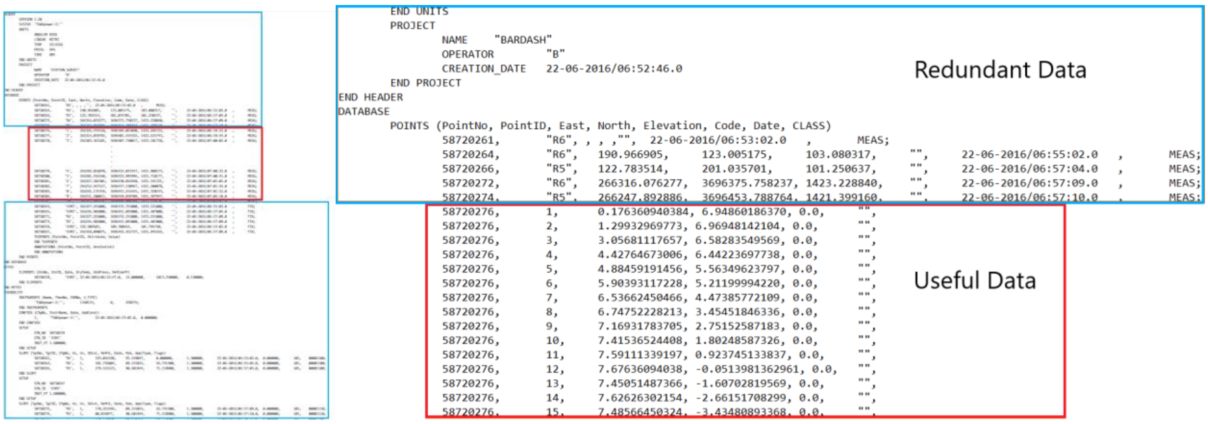


Figure 18- Pre-processing and information extraction from information containing parts

This process can be performed using scripting via a Programming Language. In this CDT application for soil management, Python Programming Language was used. Python scripting was done in Revit Dynamo environment for more consistency with the current digital models and to facilitate tasks in the ahead steps. The script for preprocessing and information extraction, along with the final result, is shown in Figure 19. Having this cleaned data, in the next steps, this data can be integrated with the virtual model to update the model and mirror the latest status of the physical entity, i.e., soil deposit/borrow status.

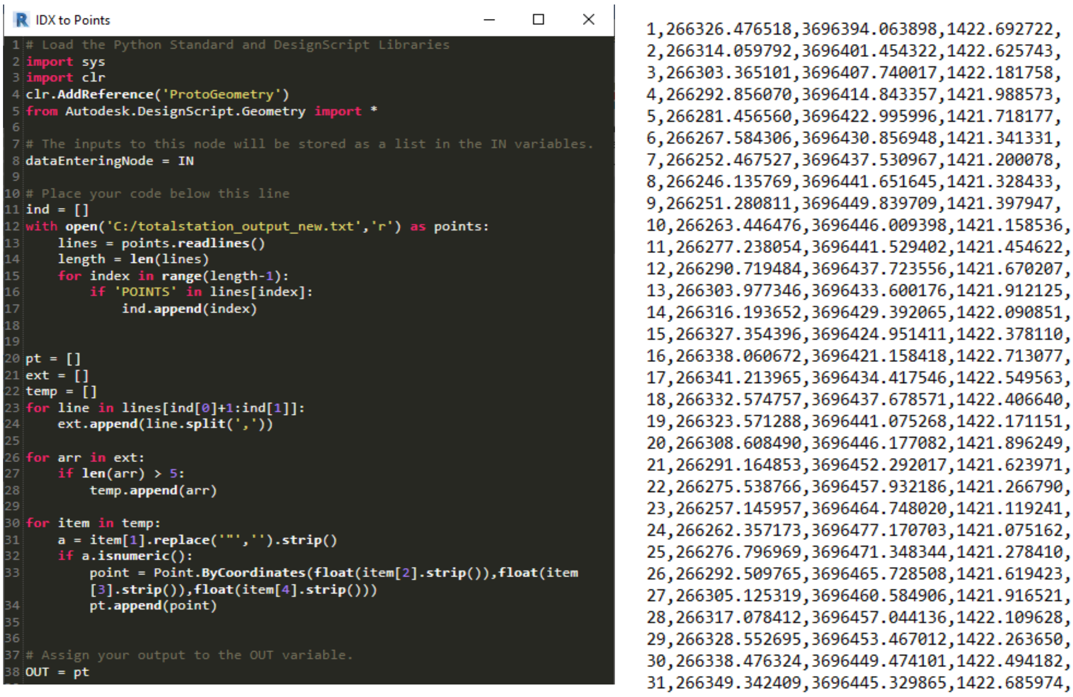


Figure 19- Scripting for data extraction and points' coordinates as the output

The extracted data points can be directly translated into a topography inside Dynamo or imported into Revit as a .csv points file to create the initial model or update the existing model. Figure 20 shows the result of updating the model based on the acquired data.

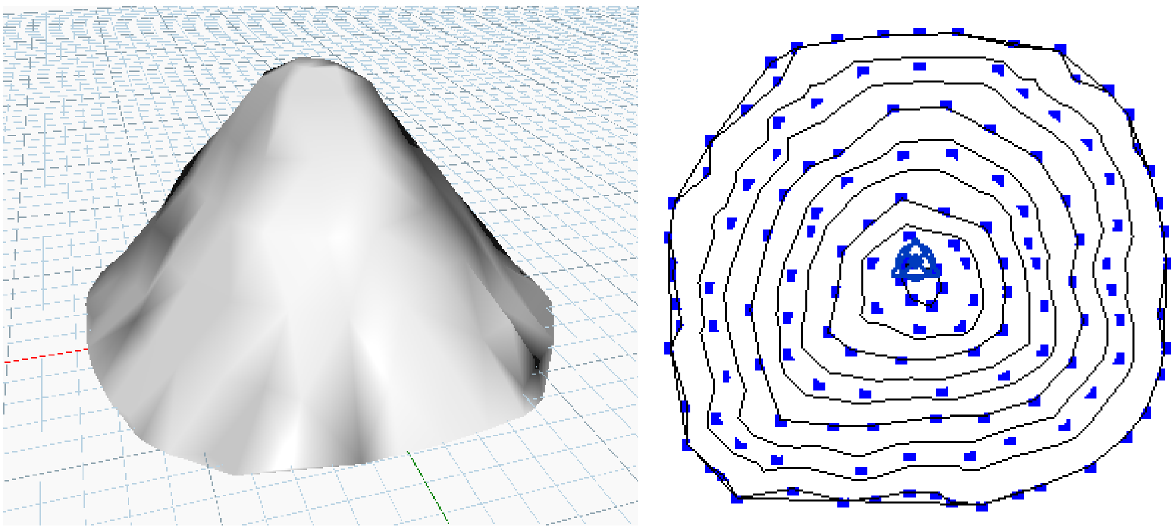


Figure 20- Updated model in Dynamo and Revit

Similarly, information from the construction site can be gathered and integrated with its virtual model to update the virtual twin of the construction site physical twin. Figure 21 shows the construction site virtual twin.

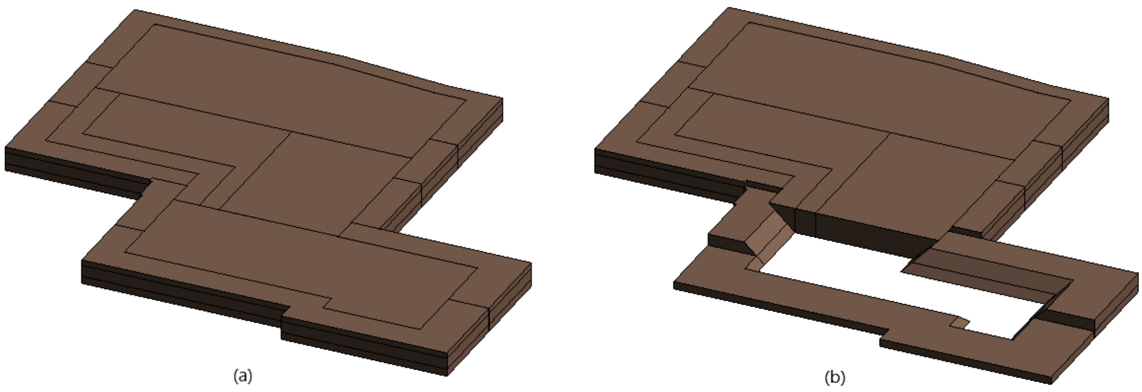


Figure 21- Construction site virtual twin, (a) before updating, (b) after updating

Following the model and data integration, depending on the data-gathering technology used, the volume of the existing soil in soil deposit/borrow will be available to relevant stakeholders for information and to make decisions and manage upcoming activities. For this purpose, a "Soil Volume" instance parameter was created to present the soil deposit/borrow volume as the model gets updated. In case that the model is updated based on the data points output from the site surveying tools, a dynamo script was developed to calculate the volume and assign the value to the Soil Volume parameter in Revit as shown in Figure 22.

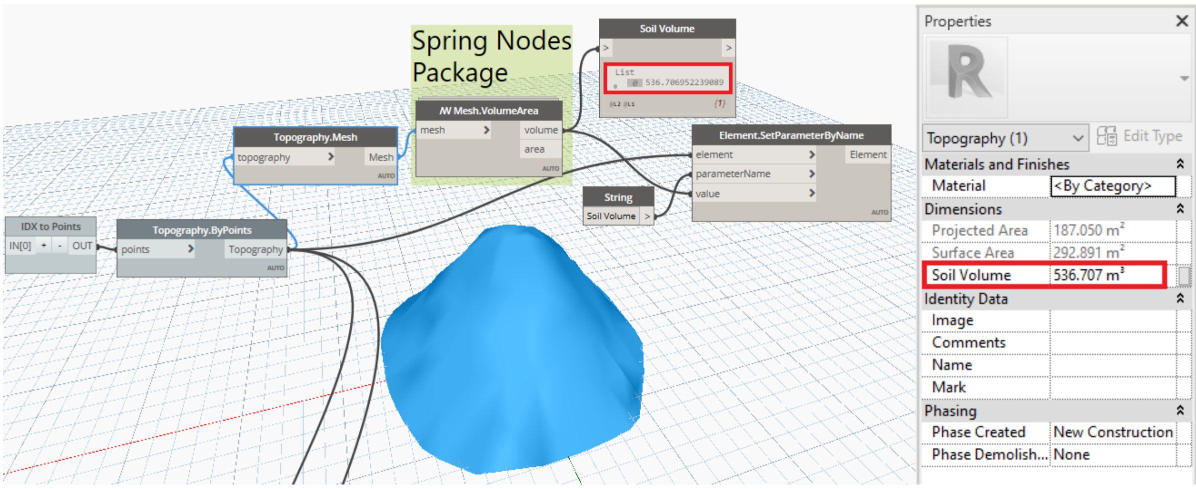


Figure 22- Calculating the soil volume and updating the model's parameter

Since in many cases the soil deposit/borrow may be located outside of the construction site due to several reasons (e.g., unavailable space, urban rules, or type of project such as road projects that the location of soil and the location of execution activities are at varying distances and changes as the project progresses), it is necessary to place the virtual models at their actual positions based on the project's coordination system, e.g., World Geodetic Reference System 84 (WGS84), European Terrestrial Reference System (ETRS89), etc. The integration between the current CDT and Geographic Information System (GIS) is advantageous in several ways for analysis and optimization studies such as minimizing the travel time, finding optimal routes, managing machinery performance and their idle time, cost minimization etc. For this purpose, ArcGIS and Autodesk InfraWorks were deployed to project the virtual models into GIS scenes.

Figure 23 shows the as-designed model at the construction site's location within the urban context in Autodesk Infraworks.

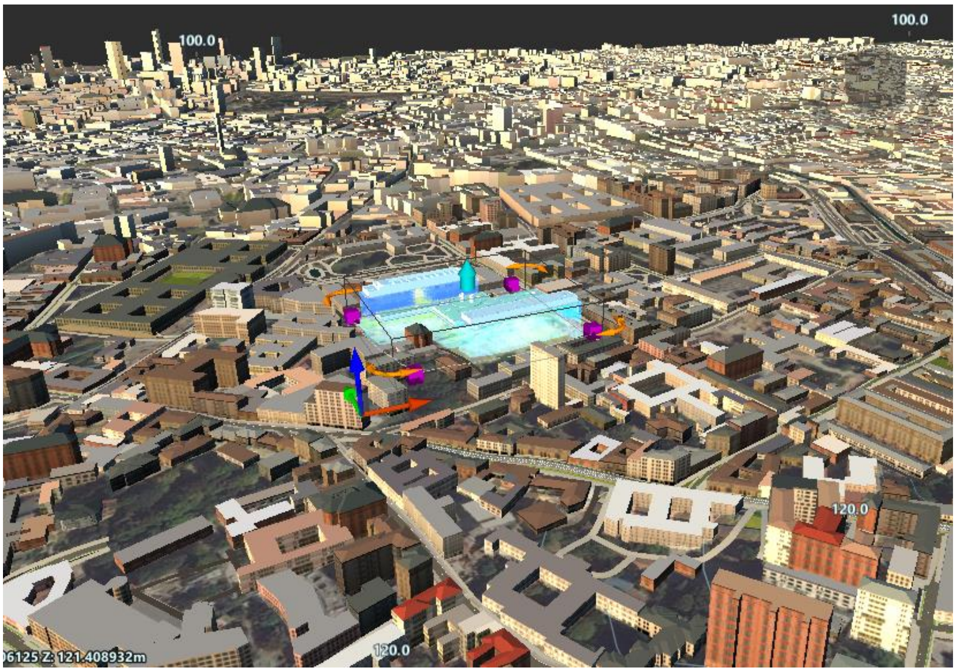


Figure 23- Hospital virtual model in the context of the city (InfraWorks)

Similarly, the soil deposit/borrow as-is model is integrated with the city model, as illustrated in Figure 24.

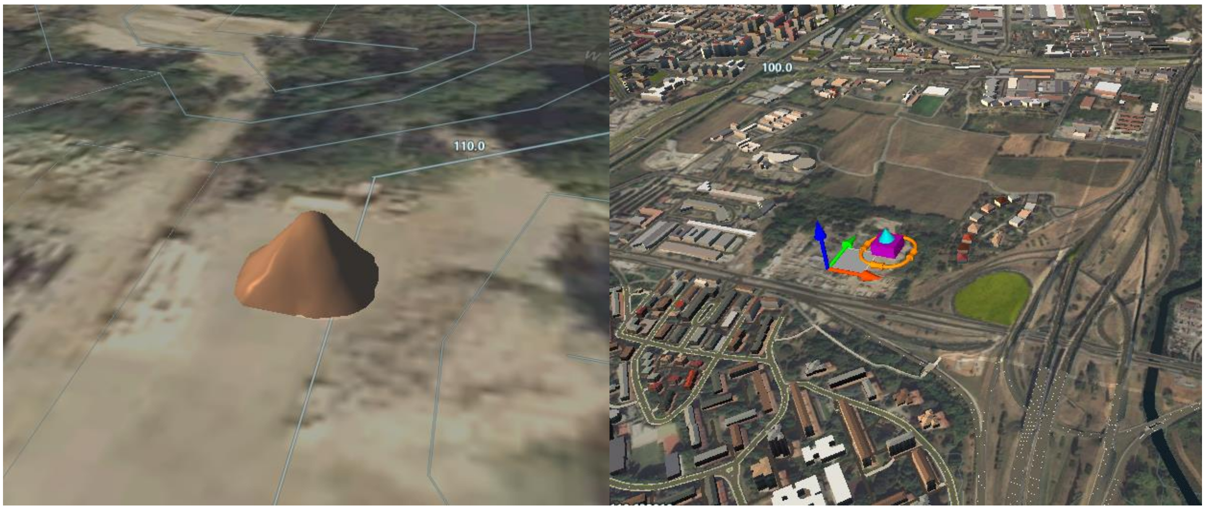


Figure 24- Soil deposit/borrow virtual model in the context of the city (InfraWorks)

Finally, both construction site and soil deposit/borrow models in the city context will provide a better understanding about their distance and relation together visually. It also enables analysis and optimizations regarding soil deposit/borrow location and access routes to its location (see Figure 25).

Figure 25- Hospital construction site and the soil deposit/borrow (InfraWorks)

Similarly, ArcGIS can also be used to map the BIM models into the GIS model of the city. Since digital models of the construction site and soil deposit/borrow need to be uploaded on InfraWorks servers after each update, the ability of ArcGIS in importing digital models locally may perform better in terms of autonomously updating BIM models in the city model as the construction site and soil deposit/borrow models evolve. Figure 26 illustrates the soil deposit/borrow virtual model in ArcGIS.

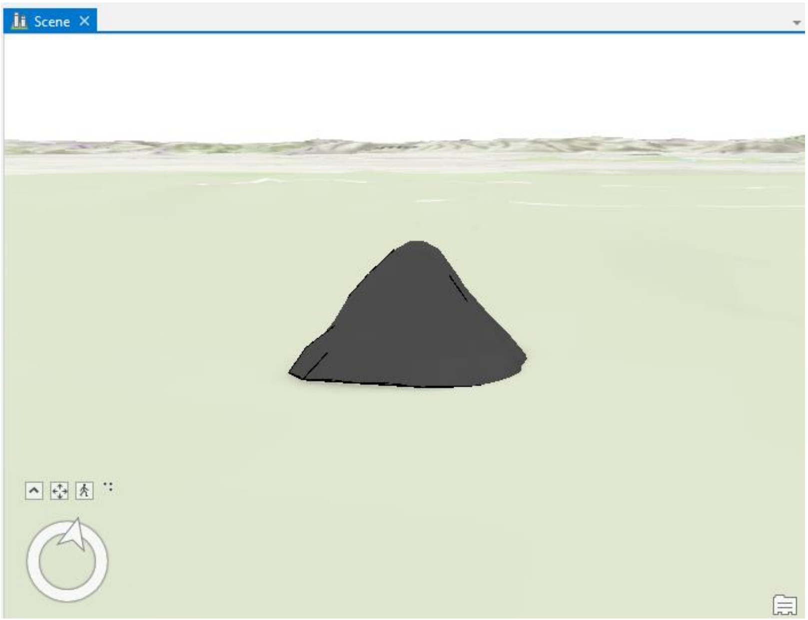


Figure 26- Soil deposit/borrow virtual model in the context of the city (ArcGIS)

Similar to the city model implemented in InfraWorks, the construction site together with soil deposit/borrow models can be integrated with the city model to better present the relation and interaction of the CDT with the surrounding environment and urban contexts, as depicted in Figure 27.

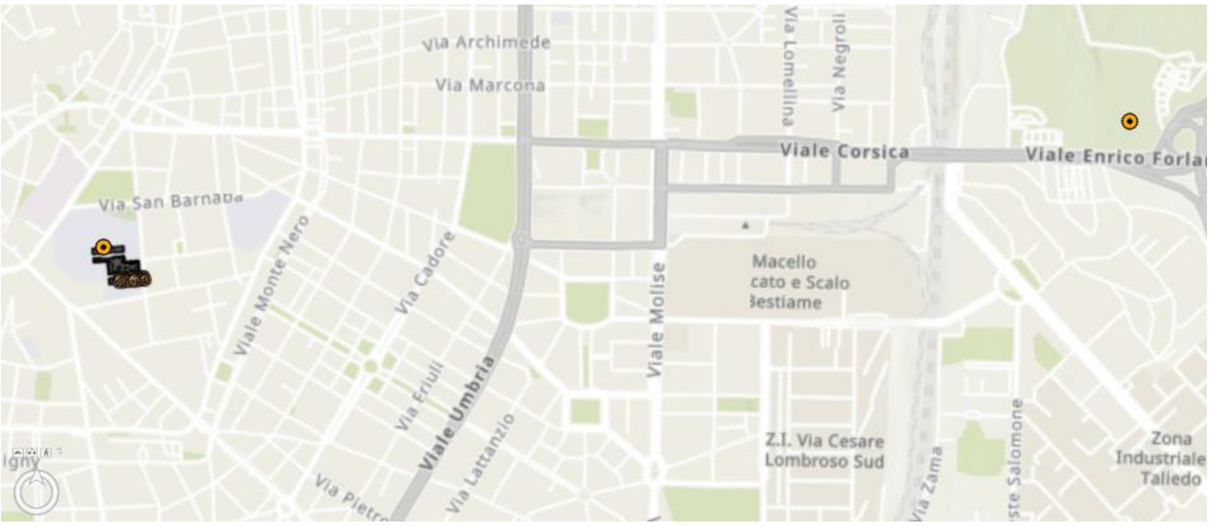


Figure 27- Hospital construction site and the soil deposit/borrow (ArcGIS)

4.5. Analytics and Prediction Tier

This tier includes various tools and techniques such as data analysis, simulation, data processing, AI/ML etc., for analytics, predictions, simulations, optimizations and other purposes. Earthwork operations are among the most important operations in construction sites. However, they are probable in nature and include repetitive work, expensive equipment, and large volumes of work performed in a highly uncertain environment (Shawki et al. 2015). Therefore, analysing the earthwork operations plays an important role in the project's productivity and final cost and time because different strategies have different impacts on the project's outcome. The planning of earthwork operations can be improved significantly using discrete-event simulations (Shawki et al. 2015). Banks et al. (2004) define simulation as "the imitation of the operation of a real-world process or system over time." Simulation enables the investigation of a wide variety of "what-if" scenarios about the real-world system (Banks et al. 2004). Simulation can be deployed to predict the performance of the new designed systems. In addition, simulation can be used as an analysis tool to predict the performance of existing systems under varying sets of changes and circumstances (Banks et al. 2004).

Simulation models can be classified into different categories based on their features. Models can be classified as mathematical or physical models. A simulation model can be discrete or continuous. In a discrete model, the state variables change only at a discrete set of points in time, while in a continuous model, the state variables change continuously over time (Banks et al. 2004). Simulation models may be further classified as static or dynamic, and deterministic or stochastic. A static model represents a system at a particular point in time, while in a dynamic simulation model, the system changes over time. Simulation models with a known set of inputs are deterministic, while stochastic simulation models have one or more random variables as inputs (Banks et al. 2004).

In the current soil management application, a discrete event simulation was deployed to perform various scenarios regarding the excavation and transportation of the excavated soil to the soil borrow/deposit location to select the proper number of equipment (trucks, excavators/loaders) based on the project schedule. In this simulation, the volume of soil that needs to be excavated and transported is acquired from the BIM model. Soil carrying trucks are entities, and excavators/loaders are resources of the simulation model. Loading, trip from the loading site (construction site) to the dumping site (soil deposit/borrow), dumping, and trip from the dumping site to the loading site are activities of the system.

Data collection for different activities can be done using historical data, direct observation or sampling, and using expert judgment. For example, in this CDT application, temporal data related to trip to dumping site and trip to loading site can be acquired from timestamped GPS data, i.e., direct observation or sampling data collection method. Next, based on the acquired data, a proper distribution (i.e., Normal, Poisson, Exponential, Beta distribution, etc.) that data fits into it needs to be identified. One method for identifying the distribution is to present data on a histogram and check whether the histogram's shape is close to the probability density function curve of any distributions mentioned above. For example, according to the historical data for loading activity (data for 120 samples is shown in Table 5), the histogram of the data is created to check the probable distribution of the data.

Table 5- Loading time (Minutes)

15.35	11.05	10.55	8.95	11.25	9.4
13.05	8.35	9.3	9.75	12.4	12.85
12.7	13.05	11.3	11.7	10.6	6.45
12.65	9.25	9.1	9.3	11.4	10.75
12.25	10.6	9.75	11.7	9.45	13.4
12.05	8	10.75	12	10.75	10.6
11.75	9.8	11.5	11.05	10.25	9.9
11.55	10.65	10.3	10.7	13.3	10.3
11.4	12.55	11.6	11.1	9.95	10.25
11.05	10.8	11.7	12.25	9.25	10.5
10.6	9.55	11.45	12.55	11.75	10.35
10.6	10.2	9.7	12.9	9.6	11.4
10.55	10.1	12.3	11.95	10.65	12.85
10.45	11.8	9	10.7	10.2	11.6
10.3	9.75	13.5	11.85	11.2	8.75
9.75	10.95	7.5	10.65	13.5	9.1
9.4	10.55	14.8	10.5	11.45	10.65
9	9.95	10.8	13.3	10.7	11
8.65	10.95	8.85	10.55	9.7	9.6
7.9	12.45	12.2	10.9	10.8	11.65

A summary description of the loading time data is provided in Table 6 along with the number of intervals and interval length for presenting data on a histogram.

Table 6- Loading time data summary

N	Min	Max	Average	Std. Dev.	Number of Intervals	Interval Length
120	6.45	15.35	10.84	1.45	11	0.809

For selecting the adequate number of intervals, a guideline is to select the number of intervals close to the square root of the number of data collected, e.g., 11 intervals for 120 data items in the loading data. Next, the frequency of occurrences within each interval is found (see Table 7) and used to create the histogram.

Table 7- Data intervals and their frequencies

Interval No.	1	2	3	4	5	6	7	8	9	10	11
Interval	6.45-7.26	7.26-8.07	8.07-8.88	8.88-9.69	9.69-10.50	10.50-11.30	11.30-12.11	12.11-12.92	12.92-13.73	13.73-14.54	14.54-15.35
Frequency	1	3	4	15	20	35	20	13	7	0	2

A histogram was developed according to the intervals and their corresponding frequencies. As shown in Figure 28, it seems that the histogram represents a Normal distribution.

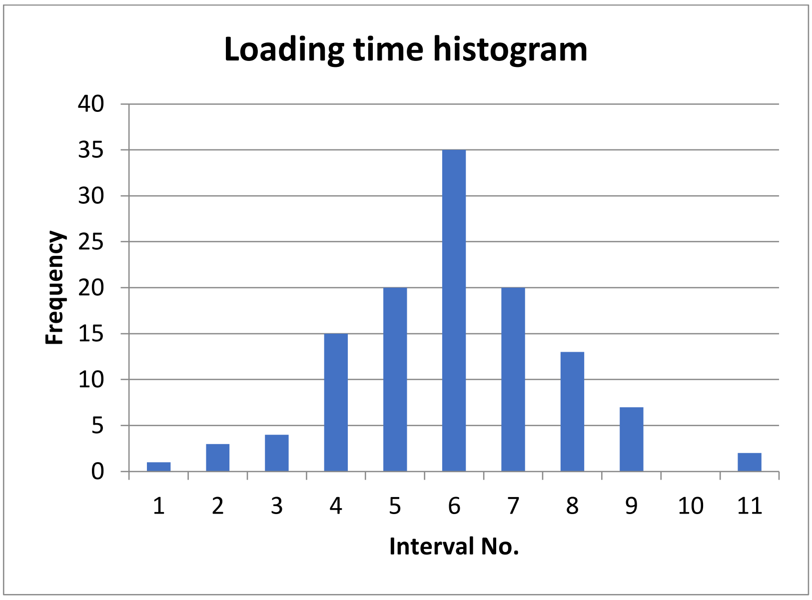


Figure 28- Loading time histogram

The main parameters of a Normal distribution, i.e., μ and σ, are calculated previously and have values of 10.84 and 1.45, respectively. Following the identification of a distribution and its parameters, the goodness of fit should be performed by comparing the fitted distribution to the empirical distribution in order to assess the quality of the fit obtained. One method for testing the goodness of fit is the Chi-Square test. The Chi-Square test is based on measuring the discrepancy between the histogram of the sample and the fitted probability density function. Chi-Square statistic χ^2 can be obtained from Equation 1:

$$\chi^2 = \sum_{j=1}^{k} \frac{(Nj - np)^2}{np} \qquad \text{Equation (1)}$$

The n observations are arranged into a set of k class intervals or cells (C1, C2, …, Ck) that each class interval has equal expected probabilities of 1/k. In Equation (1), n is the total number of observations, Nj is the number of observations in cell j and p=1/k. Based on comparing the obtained χ^2 to that from Chi-Square tables, the identified distribution can be accepted or rejected. The distribution is rejected if:

$$\chi^2_{\text{computed}} > \chi^2_{1-\alpha,(\text{k-L-1})}$$

Where α is the level of error accepted, 1-α or P-value is the level of confidence, k is the number of classes, L is the number of estimated parameters, and k-L-1 is the degree of freedom.

Now, for the loading time data with the identified Normal distribution, Nj can be calculated as shown in Table 8:

Table 8- Calculating the number of observations (Nj)

j	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
CDF	0	0.090909	0.181818	0.272727	0.363636	0.454545	0.545455	0.636364	0.727273	0.818182	0.909091	1
Cj	-1000	8.906352	9.525274	9.966016	10.33708	10.6773	11.00853	11.34876	11.71982	12.16056	12.77948	1000
Nj		8	12	13	8	17	13	7	13	7	10	12

With the obtained values for the total number of observations (Nj), n, and p, the χ^2 can be calculated from Equation (1). Therefore, $\chi^2_{\text{computed}} = 9.25$. According to the comparison of the calculated value with the Chi-Square with 8 (=11-2-1) degree of freedom and confidence level of 95% (from the Chi-Square table), the goodness of the identified Normal distribution is accepted because $\chi^2_{\text{computed}} = 9.25 < \chi^2_{95\%,8} = 15.50$. Thus, the loading activity has a Normal distribution with an average (μ) value of 10.84 and a standard deviation (σ) value of 1.45.

Similarly, this process was performed for the trip to dumping, dumping, and trip to loading activities data. The identified distributions and their corresponding parameters for all activities are shown in Table 9.

Table 9- Summary of activities' identified distributions and their parameters

Activity	Identified distribution	Distribution parameter(s) (minute)
Loading	Normal distribution	μ=10.84, σ=1.45
Trip to dumping	Normal distribution	μ=40.06, σ=3.34
Dumping	Normal distribution	μ=8.04, σ=1.66
Trip to loading	Normal distribution	μ=20.23, σ=2.39

For developing the earthmoving simulation model, Rockwell Automation Arena Version 14 was adopted. Arena as an alternative for simulating construction operations has several advantages. These advantages are summarized as follows (Shawki et al. 2015):

- The ability to model a variety of construction problems,
- Easiness in developing a simulation model,
- Easiness in learning the software,
- Effective tool for academic and commercial purposes.

The general concept of the earthmoving model is shown in Figure 29. Extracted from the BIM model, a total volume of 181,940 m3 of soil will be excavated. Assuming all of this volume needs to be transported, the simulation ends when 181,940 m3 of soil is transported.

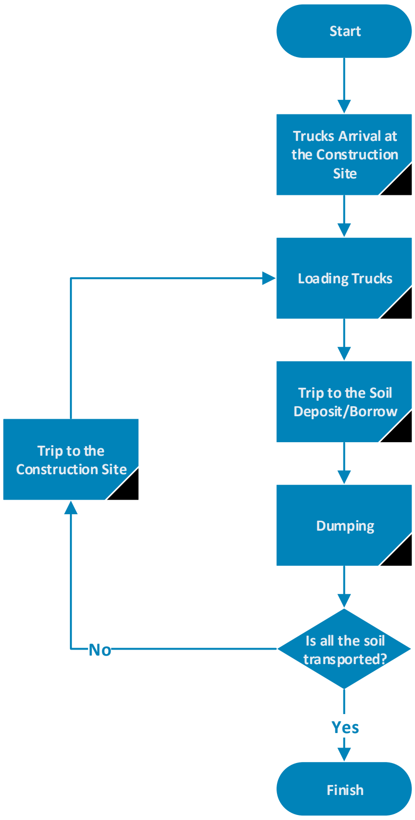


Figure 29- The general concept of the earthmoving model

According to the trucks' specifications, each truck has a capacity of 13 m3 of soil. Initially, for this amount of earthmoving volume, two excavators/loaders and eight trucks were considered. Each day is considered eight working hours. After four hours of working, a work break is considered for one hour, and after that, three hours of working until the end of the working day. At the break time, if an excavator/loader is loading a truck, the excavator/loader would stop the loading operation, and it will resume the work after the break time. In addition, it is assumed that every 960 working hours, the excavator/loader has a failure that takes 6 to 12 hours to be repaired. Also, every 72 working hours, the excavator/loader needs a check-up which follows a triangle distribution for performing the check-up with a minimum of 45 minutes, a maximum of 90 minutes, and a most likely value of 60 minutes. The developed simulation model in Arena software is shown in Figure 30.

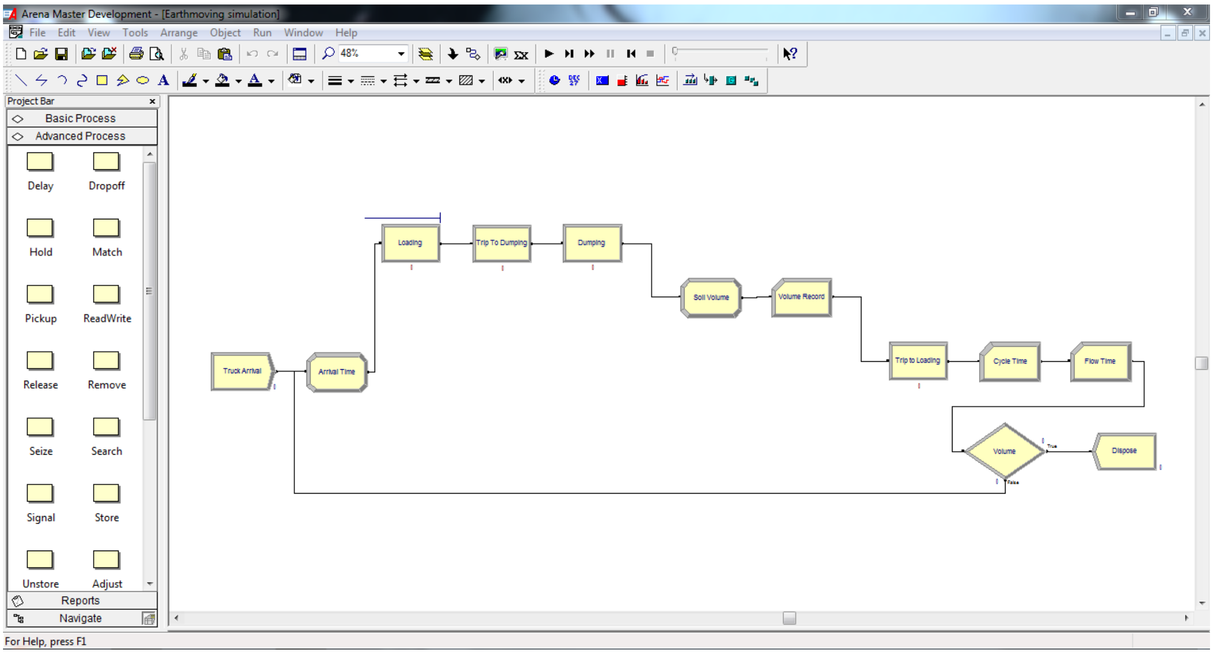


Figure 30- Earthmoving simulation model in Arena

The number of simulation runs in order to fulfil the desired confidence level needs to be determined. The minimum number of simulations can be calculated from Equation (2):

$$R \geq (\frac{Z_{\alpha/2}S_0}{\varepsilon})^2 \qquad \text{Equation (2)}$$

Where S_0 is an initial estimation of σ calculated based on an initial number of runs of R_0 (e.g., ten runs), ε is an absolute error value, α is the risk level (therefore 1- α as the confidence level), and $Z_{\alpha/2}$ is calculated based on the Normal distribution according to the confidence level. An initial run with ten replications was performed to calculate S_0 according to each replication's cycle time (how long it takes a truck to complete one cycle). Table 10 shows the cycle times for ten runs and the average and standard deviation of the cycle times:

Table 10- Cycle times for the initial runs

Run	Cycle time (min)	Average (Y..)	S_0
1	97.45	97.402	0.41
2	97.24		
3	97.09		
4	97.75		
5	96.77		
6	97.51		
7	97.27		
8	98.11		
9	96.99		
10	97.84		

Considering a 95% confidence level, and a maximum error size of 0.1 for average of truck cycle time, the minimum acceptable number of simulation runs can be calculated:

$R \geq (Z_{0.025} * S_0 / \varepsilon)^2 = (1.96 * 0.41 / 0.1)^2 = 64.5$

Therefore, a minimum number of 65 runs is required.

For the soil management case study, three scenarios were considered to perform the job in 8, 6, and 4 months. For each scenario, the proper number of construction equipment needs to be determined considering the project schedule.

In the first scenario, a duration of 8 months (240 days) was considered in the project schedule to complete earthwork operations. Following the determination of the minimum required number of runs, 65 runs were performed with the initial model setup, i.e., two excavators/loaders and eight trucks and other attributes of the system as mentioned above. After the model's run, the results of the initial strategy were reviewed to evaluate the overall performance of the system and the adopted strategy. Table 11 shows the variables and results of the simulation of the first strategy.

Table 11- Results of the first strategy for scenario #1

Variable	Result
Number of excavators/loaders	2
Number of trucks	8
Total working days	325.9
Daily production (m3/d)	558.27
Mean excavators/loaders utilization	49%
Average waiting time in loading queue (min)	8.12
Average number of waiting trucks in loading queue	0.73

The output results show the inefficiencies of the first strategy due to the following reasons:

- Project schedule is not met,
- Daily production is not sufficient,
- Utilization of construction equipment is low.

To improve the inefficiencies of the first strategy, another operation strategy was established, and four trucks were added to the existing eight trucks. The results of the second simulation run, as depicted in Table 12, show that the project duration is met and the utilization of the resources is satisfactory. However, the average waiting time in loading queue is slightly high. Reducing the number of trucks to 11 trucks might slightly decrease the average waiting time and the average number of waiting trucks while maintaining the project's schedule.

Table 12- Results of the second strategy for scenario #1

Variable	Result
Number of excavators/loaders	2
Number of trucks	12
Total working days	224.8
Daily production (m3/d)	809.34
Mean excavators/loaders utilization	70%
Average waiting time in loading queue (min)	11.73
Average number of waiting trucks in loading queue	1.52

For the second scenario, 6 months (180 days) was considered for the earthwork operations in the project schedule. Similar to the first scenario, different strategies were implemented to evaluate the best strategies. For brevity, only the final results of the implemented strategies are shown in Figure 31.

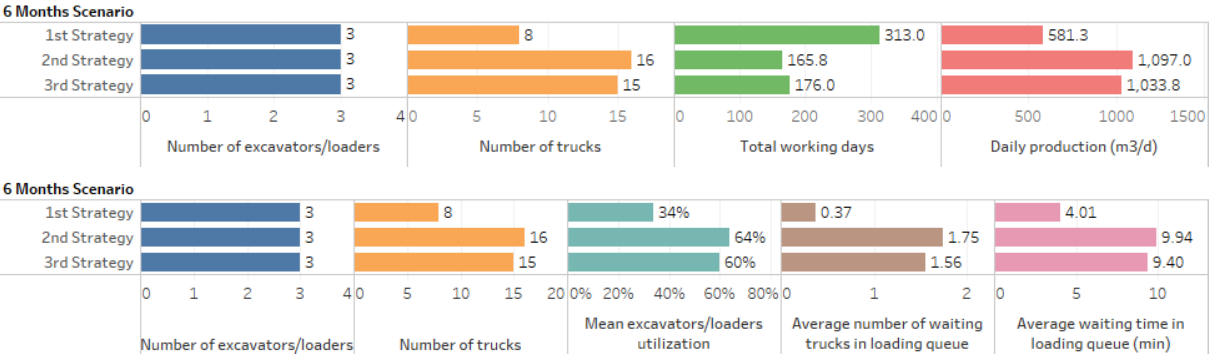


Figure 31- Results of the various strategies for scenario #2

Similarly, various simulations were performed for the third scenario with 4 months (120 days) duration to evaluate the best strategies. For brevity, only the final results of the implemented strategies are shown in Figure 32.

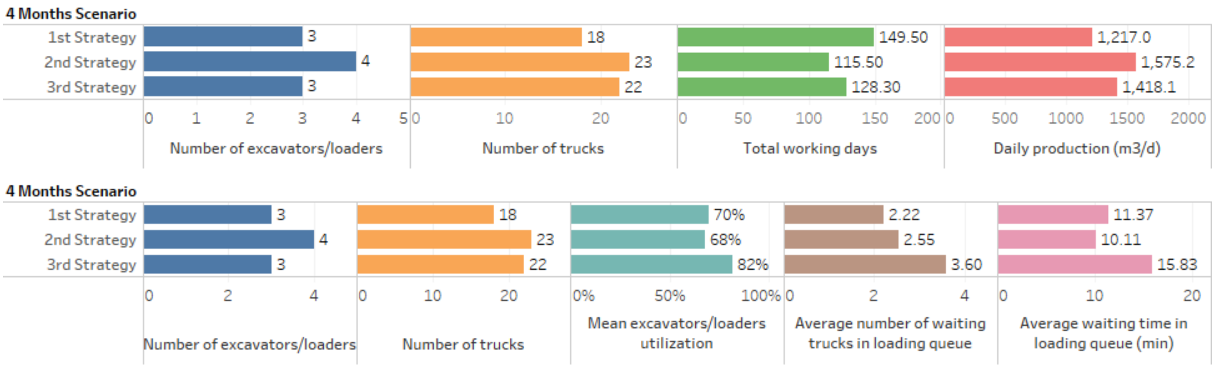


Figure 32- Results of the various strategies for scenario #3

Selection of the optimal strategy depends on various parameters such as the project's time and budget, the number of available equipment, desired equipment productivity and waiting time, etc.

4.5.1. Simulation process overview and information flow

This subsection recaps the previous simulation process and adds more details about the simulation components in Arena simulation software. The information flow of the simulation process is also presented, and the role of the BIM/GIS models stated in the previous steps of the soil management application is depicted in the simulation process.

Developing the simulation model starts with the creation of system entities. In the current earthmoving simulation model, trucks are the system entities that arrive at the construction site based on a specific time basis, e.g., constant time intervals or a distribution. Then processes or activities such as loading, trip from the construction site to the dumping site, dumping, and trip from the dumping site to the construction site are done by the entities and resources, i.e., excavators/loaders, are used in some processes. Additional components were also used to define and record various system attributes, e.g. trucks arrival time, the volume of dumped soil, cycle time, flow time etc. Several conditions, such as resources' up times and down times, working schedules, resources' hold and release conditions, etc., were modelled to develop a truthful model that mimics the real-world operations. In addition, various statistics were defined to capture the critical status of the system components, such as resources' busy and idle times and their utilization percentage, to better understand the system's performance and subsequently change the system's parameters to optimize its performance.

The simulation model needs to be fed by data. This data comes from various sources, such as the collected data from sensors, GPS data, etc., or data from models like BIM, GIS models. In the simulation model, different components of the system have separate data input/output and dataflow. The dataflow diagram for the earthmoving simulation model shown in Figure 33 depicts the simulation components, their inputs, and the flow of data.

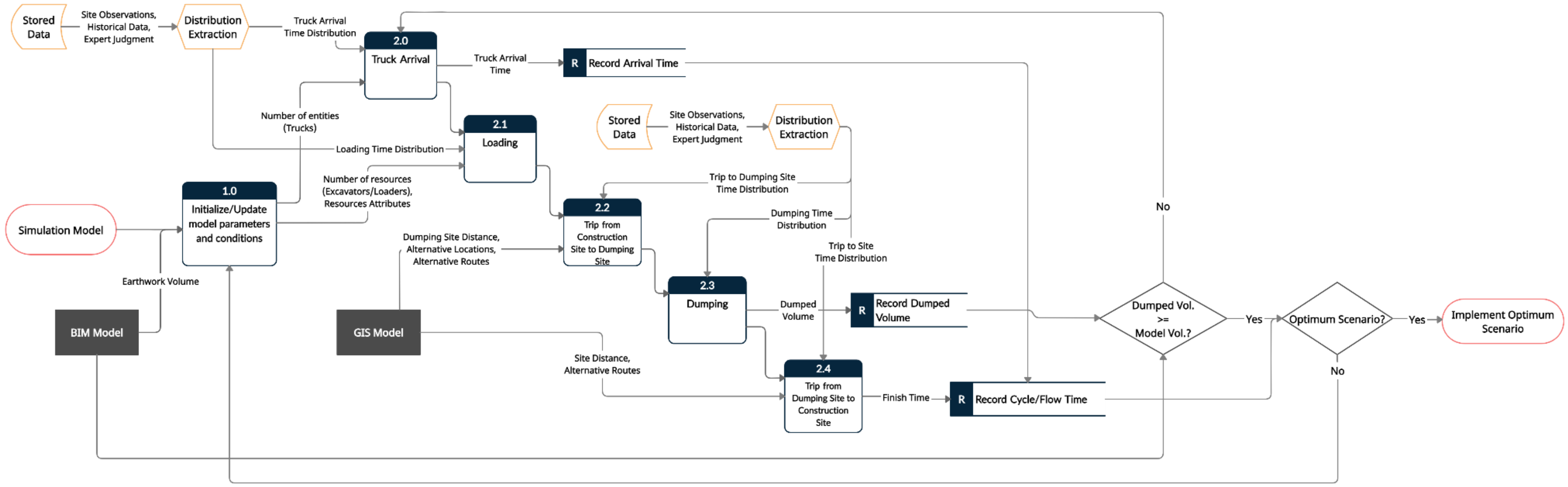


Figure 33- Dataflow Diagram for the earthmoving simulation model

Aligned with the distribution identification process mentioned previously, Arena has also a tool called "Input Analyzer" that analyses the data and identifies the best distribution that fits the imported data. To investigate the historical data of loading activity (previously identified as having a normal distribution) with Arena's Input Analyzer tool, this data was imported to the software to check the best distribution. Following the import of loading time data, Input Analyzer provides a histogram and a summary of the imported data (see Figure 34).

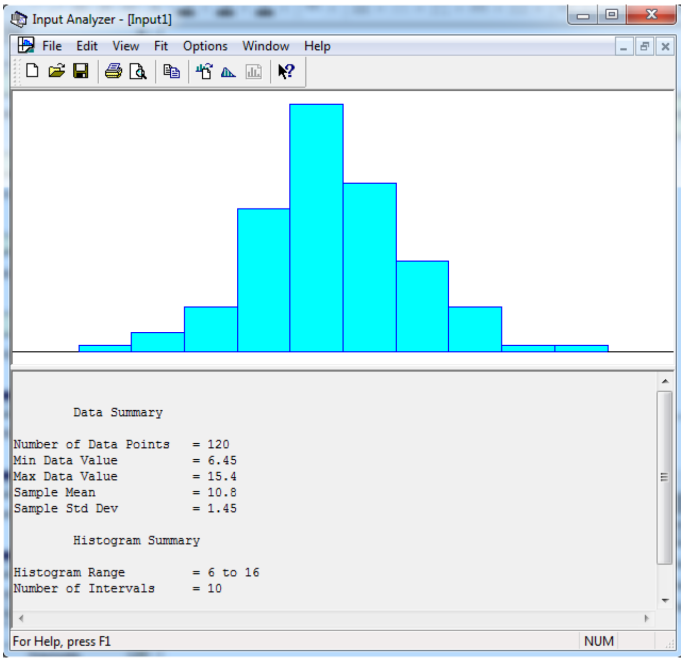


Figure 34- Histogram and summary of the imported loading time data by Arena Input Analyzer

Next, Input Analyzer evaluates various distributions (i.e., Normal, Poisson, Triangular, Uniform, Beta, etc.) to find the best distribution that fits the imported data. The results from Input Analyzer (see Figure 35) depict that loading data follows a Normal distribution with parameters values of μ= 10.8 and σ= 1.44.

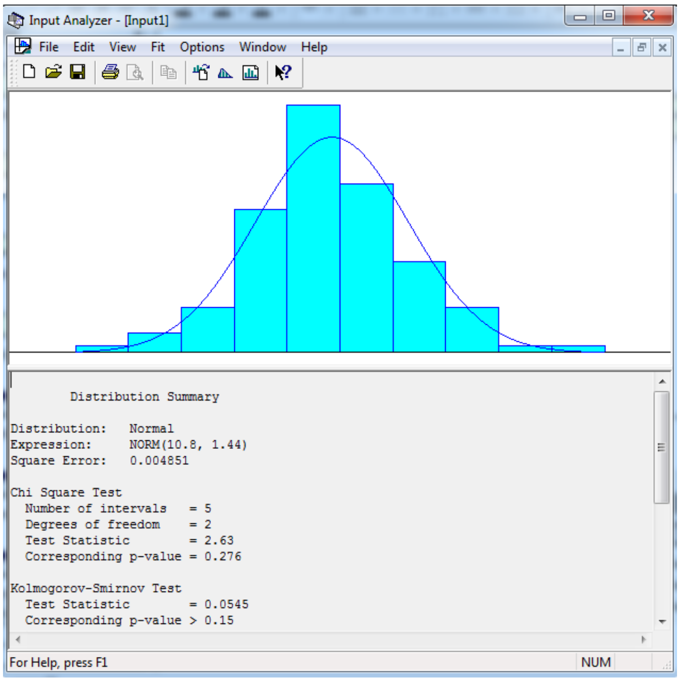


Figure 35- Identified Normal distribution for loading time data

Normal distribution was identified as the best distribution since it has the minimum Square Error compared to other distributions. The goodness of fit results for different tests (i.e., Chi-Square Test and Kolmogorov-Smirnov Test) indicate that Normal distribution is acceptable.

Model verification and validation are two important aspects in developing a simulation model. Model verification is concerned with building the model correctly, i.e., the model is correctly implemented in the simulation software and represents what it meant to represent. Model validation is concerned with building the correct model, i.e., a credible and valid model that is accurate enough to represent the real system (Banks et al. 2004). Various measures such as using expert simulation modeller, making event flowcharts, examining the model outputs, performing a manual simulation, examining the input parameters at the end of the simulation, computing the required minimum number of model runs and performing enough replications, etc., were performed to ensure the verification and validation of the earthmoving simulation model.

4.5.2. Challenges, advantages, and potential evolutions of the implemented simulation model

Data as a bridge that connects the physical world to the virtual space is a key component that drives the simulation model. Data needed for different activities and components of the simulation model comes from various sources. Acquiring data for building a truthful model that mimics the real-world system can be a challenge in implementing such a model. Building the simulation model correctly and building the correct simulation model, i.e., a verified and valid model, is another challenge. Simulation models can become very complex, which makes the correct creation of the model as well as validation of the output results challenging tasks. Interoperability and automatic data connection are other challenges that hinder the full performance of the simulation process within the implemented CDT. In this simulation model implemented in Arena software, data for various components and activities of the system, such as activities time data, data from BIM models, data from GIS models etc., were imported to the model manually.

Once the correct model is developed correctly, i.e., a verified and valid model, the simulation model offers several benefits. It enables the evaluation of several what-if scenarios to select the most optimum and practical scenario. It saves the costs of actual implementation of construction operations by simulating the operations in a virtual space. In addition, it enables simulation of complex systems that demands high computational power and hard to perform manually. Valuable insights and knowledge can be acquired from the simulation results, which will support the top-level applications of the CDT, i.e., Service Tier.

As an important aspect of the simulation process and its integration with the proposed CDT framework, automated data exchange between the tiers of the CDT, i.e., acquired data from the construction site and data from BIM/GIS models etc., could enhance the performance of the whole system. A potential solution is using APIs, Web, and Cloud-based simulations to ensure data connection and integration of different platforms. For instance, AnyLogic simulation software supports Cloud-based simulation and has its own API.

4.6. Service Tier

Being fed by lower tiers, i.e., physical-to-virtual twinning tier and data analytics and prediction tier, the service tier provides several benefits and services. In the current soil management application of the CDT, the updated virtual models of the construction site and soil deposit/borrow, simulation results, and acquired data can be used in providing various services such as site progress monitoring, logistics and material supply, site management and resource allocation, claims management etc. Informed decisions can be made based on these services, and with the predefined configurations, criteria, or thresholds, controlling or actuating mechanisms can be enabled in the virtual-to-physical twinning tier.

4.7. Virtual-to-Physical Twinning Tier

Finally, proper decisions can be made, or actuating mechanisms can be activated in the virtual-to-physical twinning tier. In this tier, the virtual entity/space acts on the physical entity/space and virtual-to-physical twinning happens. In the soil management application, a scenario with the proper number of equipment can be practically implemented at the construction site, real-time monitoring of the current status of the soil deposit/borrow and construction site would be possible, and earthwork operations progress and deviations from the project schedule can be tracked. In addition, actuation processes can be performed. As an example, an excavator digital twin can be used for operation control and monitoring. The bi-directional data connection enables the end-user to run the equipment remotely or provide a decentralised platform for collaboration (Sepasgozar 2021). A tablet with an augmented model of the excavator can also be used to transfer the changes made by an augmented model to the physical entity (Sepasgozar 2021). There are platforms such as Trimble Earthworks that can be integrated with construction fleets and provide machine operation control.

In the next section, the results and impact of this study will be presented. In addition, further developments both in theoretical and practical perspectives for future studies will be discussed.

5. CONCLUSIONS

Industry 4.0, as an evolution from Industry 1.0 to Industry 3.0, encompasses ample benefits for various industries like manufacturing, aerospace, systems engineering, oil and gas, construction, etc. Digital Twin (DT), as one of the main concepts of Industry 4.0, bespeaks a new paradigm in the construction industry as a real-time virtual replica of a physical asset. Several industries such as manufacturing, automotive, aviation, and healthcare are extensively using the concepts of Industry 4.0, but the construction industry is in its infancy in terms of adopting and implementing Industry 4.0 principles. Moreover, most construction industry studies and practices have focused on implementing DT in operation and maintenance phases of facilities to date, and the use of DT in the construction phase has not been addressed sufficiently.

5.1. CDT impacts in the construction field

This research is devoted to study the implementation of DT in the construction phase for general contractors. The impacts and benefits of the construction digital twin (CDT) framework developed in this study in the construction field are multi-fold, and various application scopes can be developed.

5.1.1. AI-assisted construction management

Following the development of the CDT framework, the soil management case study was implemented as an application of the proposed CDT to demonstrate the benefits of using digital twins during the construction phases of projects. Simulation of the earthwork operations was performed as part of the CDT. The simulation results in evaluating various what-if scenarios for optimum resource allocation and operations management proved the benefits of using CDT and subsequently data analytics and prediction in assisting management bodies in their decision-making. Machine learning and artificial intelligence can be adopted to use intelligence within the decision support system to perform autonomous decisions or back up managers with their decisions and respond to dynamic site changes on time.

5.1.2. Filling information gap in the construction phase

As an information gap exists between the as-designed BIM models at the design phase and the as-built BIM models at the time of project hand-over, i.e., information gap during the construction phase, the CDT can be implemented for the desired physical asset to fill this information gap. Although this continuous and uninterrupted flow of information from the design phase to the construction phase and then to the operation phase is beneficial to owners or is required by them, general contractors greatly benefit from a CDT by acquiring enriched data. This data is used to monitor and track the construction processes or extract knowledge and insights to use in the ongoing project that the CDT is implemented for. Equally important, this data can be used in their future projects for a more accurate time and cost estimation at the time of tendering to propose a better bid mark-up or for preparing more realistic project schedules for the construction phase.

5.1.3. Reusing the CDT and enhancing future information systems

T. and Li-Ren (2004) performed an industry-wide survey to collect project data from more than 200 capital projects to study the impact of technology usage and overall project success. In terms of project typicality regarding the overall technology usage, 85% of the projects were typical. Yang et al. (2011) studied the impact of teamwork on project performance and acquired data from more than 200 projects. According to their analysis, about 63% of the projects were typical relative to construction methods and approaches. Infrastructure projects such as roads, bridges, railways, tunnels etc., might be typical in nature in many cases and only differ in some configurations and features such as location, size, capacity, etc. Once a CDT was developed to serve a specific purpose in a project, it can be reused in future projects with some adjustments. In addition, the stored historical data of the construction processes are available to designers, domain experts, and future CDTs to facilitate and optimize the development of future CDTs and construction processes. This will reduce the cost of CDT implementation and make it a long-term benefit for the general contractor.

5.1.4. Data-driven construction management

Since the construction phase is accompanied by numerous challenges and has a high degree of complexity, using the developed CDT in implementing several applications can assist general contractors in better managing the construction processes. Equipped with BIM and digital models, real-time data acquiring and communication technologies, and data analytics and prediction, this data-driven CDT framework enables advanced construction management to better understand, predict, and optimize the physical construction processes. The data-driven loop in the CDT and the physical and virtual twinning is proven effective for advanced construction management and improving construction efficiency, collaboration, and reliability (Pan and Zhang 2021).

5.1.5. Real-time capturing of the construction status

As shown in the soil management case study, it was possible to catch and monitor the soil deposit/borrow status in real-time or near real-time with the implementation of the CDT. In addition, the progress of earthwork operations was tracked, and the models were updated accordingly. This updated status of the construction site from the CDT along with the project schedules or 4D models, can be used to monitor and calculate schedule deviations. Other studies (see, e.g., Boje et al. (2020)) have also mentioned the use of CDT in real-time site safety detection and health and safety improvements. Therefore, a CDT capability in real-time site monitoring, progress tracking, and safety detection provides ample benefits and advantages for general contractors.

5.1.6. Improved collaboration and information exchange

The CDT enables collaboration and information exchange across domains. For example, in a retaining wall CDT, as an application of a CDT during construction in deep excavations, when the wall nails are inserted and temporary shotcrete facing is poured, the construction of permanent wall facing takes time to be finished. Therefore, the CDT can monitor the temporary shotcrete facing, and any changes in the status of the temporary shotcrete facing can be reported to the designer for design modifications. In addition, safety warnings can be sent to the safety management in case of wall movements and collapsing likelihood. In another application of CDT for high-volume concrete pouring (e.g., in slabs with large areas), real-time status of the parts where concrete is poured and those parts that are remained for concrete pouring, the information exchange will facilitate the procurement and construction processes among the related parties for a more efficient supply chain.

5.1.7. Transparency and data reliability

The evolution of the physical is synchronized with the virtual in real-time or near real-time. In addition, data coming from the physical and information from the virtual are stored and can be accessed for future use. The ability of CDT to present a real-time status of the physical and retrieve historical data leads to a more transparent construction among stakeholders and also makes the CDT a reliable source of information. Claims management can be considered among the possible benefits of this transparency and data reliability. Transparency contributes to claims prevention, and recorded data is the basis for claim mitigation when a claim is raised.

5.1.8. Improved construction logistics

The lack of integration of on-site and off-site supply chain actors and construction processes compromises construction productivity (Boje et al. 2020). A CDT provides information to the related actors, and by providing the current construction status and available materials, delivery of equipment and materials for the current or upcoming processes would be facilitated.

In addition to the soil management application in this study, retaining wall monitoring application for safety issues and design purposes, monitoring high-volume concrete pouring application for better procurement and management during the construction where high volumes of concrete are needed, and several other applications can be developed to take advantage of CDT during the construction phase.

5.2. CDT challenges

Construction phase complexity and limitations regarding the enabling technologies for developing a CDT pose some challenges in fully taking benefits of the proposed CDT framework in the field. This section introduces these challenges and proposes possible solutions to overcome these challenges.

5.2.1. Data acquisition

The first challenge is that due to the complexity of construction processes, there might be a lack of proper technology for data acquisition in implementing a specific application of the proposed CDT. The existing technologies might not guarantee fully automated, real-time, human-free, or cost and time-effective solutions for data gathering. A bi-directional connection between the physical and virtual entities is a crucial part of a DT. A DT requires real-time data or data with proper acquisition intervals to meet the system's needs and top application and service layers. Variability and uncertainty of the physical environment, the different scales of the physical and virtual spaces, and the data from the different entities that is being continuously generated make the connection and integration of the physical and virtual worlds challenging (Pires et al. 2019). High-speed transmission technology is a solution to the real-time synchronization need.

5.2.2. Virtual platforms costs

Another challenge that arises due to the variety of tasks and activities in the construction processes is that more than one software might be needed for digital modelling and developing the virtual twin. This might limit the adoption of the CDT for some companies. A possible solution can be using free and open-source digital solutions. Another cost-effective solution might be using existing, state-of-the-art, and off-the-shelf technology to develop the Digital Twin that are being developed independently of the Digital Twin (Jones et al. 2020). For example, in the soil management CDT application implemented in this study, digital modelling solutions such as Autodesk Revit and Dynamo were used that are almost common in companies that are using BIM. This, on the other hand, raises a concern about whether these technologies are optimized for the purpose of Digital Twin and the applications it is designed for (Jones et al. 2020). Some platforms are oriented to build digital twins, such as Dassault 3D Experience, Autodesk Tandem, etc. However, these solutions might be costly and not affordable for every company that desires to adopt the digital twin.

5.2.3. Interoperability of the virtual environments

As more than one virtual environment might be needed for the CDT implementation, interoperability and automatic data and information exchange between different platforms, e.g., modelling environment and simulation platform, might expose another challenge. Using software that supports APIs, cloud computing, common formats and open standards, etc., can mitigate the interoperability challenges.

The concept of CPS and DT supports the bi-directional communication between the physical and virtual twins while removing the human factor in physical-to-virtual and virtual-to-physical twinning. The complexity of construction processes accompanied by the possibility of lack of proper technologies might hinder the implementation of fully human-free CDTs. In addition, the vital role of construction management bodies and relevant stakeholders in authorizing the proper decisions to be implemented in the physical world, i.e., virtual-to-physical twinning, cannot be ignored.

Considering the challenges and limitations of using CDTs, future directions to mitigate these challenges would introduce new research opportunities to enhance CDT. Quantifying the return on investment and establishing the cost-benefit approach for deploying DT (Jones et al. 2020), creation and validation of the virtual models that truthfully reflect the real processes and related variables (Pires et al. 2019), organizational structure of companies for CDT adoption, sharing the generated data and digital entity among various untrusted stakeholders or sub-contractors (Sepasgozar 2021) during the construction phase of projects, CDT security and using technologies such blockchain to mitigate data exchange security issues, implementation standards, and level of fidelity required for developing a truthful CDT, etc., are among subjects that can be studied in future researches.

6. REFERENCES

Akanmu, A., and Anumba, C. J. (2015). "Cyber-physical systems integration of building information models and the physical construction." *Engineering, Construction and Architectural Management*, Emerald Group Publishing Ltd., 22(5), 516–535.

Akhavian, R., and Behzadan, A. H. (2012). "Remote monitoring of dynamic construction processes using automated equipment tracking." *Construction Research Congress 2012: Construction Challenges in a Flat World, Proceedings of the 2012 Construction Research Congress*, American Society of Civil Engineers, 1360–1369.

Anumba, C. J., Akanmu, A., and Messner, J. (2010). "Towards a cyber-physical systems approach to construction." *Construction Research Congress 2010: Innovation for Reshaping Construction Practice - Proceedings of the 2010 Construction Research Congress*, American Society of Civil Engineers, 528–537.

B.I.M.Dictionary. (2021). "BIM Dictionary (2021), Digital Twin, English, Version 1. Retrieved from https://bimdictionary.com/en/digital-twin/1, Last accessed May 25, 2021."

Banks, J., Carson, J. S., Nelson, B. L., and Nicol, D. M. (2004). *Discrete-Event System Simulation. Prentice Hall,.*

Barricelli, B. R., Casiraghi, E., and Fogli, D. (2019). "A survey on digital twin: Definitions, characteristics, applications, and design implications." *IEEE Access*, Institute of Electrical and Electronics Engineers Inc.

Bew, M. (2008). "Bew-Richards BIM maturity model." *BuildingSMART Construct IT Autumn Members Meeting*, Brighton.

Boje, C., Guerriero, A., Kubicki, S., and Rezgui, Y. (2020). "Towards a semantic Construction Digital Twin: Directions for future research." *Automation in Construction*, 114, 103179.

Bolton, A., Blackwell, B., Dabson, I., Enzer, M., Evans, M., Fenemore, T., Harradence, F., Keaney, E., Kemp, A., Luck, A., Pawsey, N., Saville, S., Schooling, J. M., Sharp, M., Smith, T., Tennison, J., Whyte, J., and Wilson, A. (2018). *The Gemini Principles: Guiding values for the national digital twin and information management framework. Centre for Digital Built Britain.*

Camposano, J. C., Smolander, K., and Ruippo, T. (2021). "Seven Metaphors to Understand Digital Twins of Built Assets." *IEEE Access*, Institute of Electrical and Electronics Engineers Inc., 9, 27167–27181.

Esmaeili, I., and Kashani, H. (2021). "Managing Cost Risks in Oil and Gas Construction Projects: Root-causes of Cost Overruns." *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*.

Fenn, J., and Raskino, M. (2008). *Mastering the hype cycle: how to choose the right innovation at the right time*. Harvard Business Press.

Figueiras, P., Lourenco, L., Costa, R., Graca, D., Garcia, G., and Jardim-Goncalves, R. (2021). "Big Data Provision for Digital Twins in Industry 4.0 Logistics Processes." *2021 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4.0&IoT)*, IEEE, 516–521.

Glaessgen, E., and Stargel, D. (2012). "The digital twin paradigm for future NASA and US Air Force vehicles." *53rd AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics and materials* *conference 20th AIAA/ASME/AHS adaptive structures conference 14th AIAA*, 1818.

Gokalp, M. O., Kayabay, K., Akyol, M. A., Eren, P. E., and Kocyigit, A. (2017). "Big data for Industry 4.0: A conceptual framework." *Proceedings - 2016 International Conference on Computational Science and Computational Intelligence, CSCI 2016*, Institute of Electrical and Electronics Engineers Inc., 431–434.

Greif, T., Stein, N., and Flath, C. M. (2020). "Peeking into the void: Digital twins for construction site logistics." *Computers in Industry*, Elsevier B.V., 121, 103264.

Grieves, M. (2014). "Digital twin: manufacturing excellence through virtual factory replication." *White paper*, Florida Institute of Technology, 1, 1–7.

Grieves, M., and Vickers, J. (2016). "Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems." *Transdisciplinary Perspectives on Complex Systems: New Findings and Approaches*, Springer International Publishing, 85–113.

Ham, Y., and Kim, J. (2020). "Participatory Sensing and Digital Twin City: Updating Virtual City Models for Enhanced Risk-Informed Decision-Making." *Journal of Management in Engineering*, American Society of Civil Engineers (ASCE), 36(3), 04020005.

Hasan, S. M., Lee, K., Moon, D., Kwon, S., Jinwoo, S., and Lee, S. (2021). "Augmented reality and digital twin system for interaction with construction machinery." *Journal of Asian Architecture and Building Engineering*, Taylor and Francis Ltd., 1–12.

El Jazzar, M., Piskernik, M., and Nassereddine, H. (2020). "Digital Twin in construction: An Empirical Analysis." *EG-ICE 2020 Proceedings: Workshop on Intelligent Computing in Engineering*, 501–510.

Jones, D., Snider, C., Nassehi, A., Yon, J., and Hicks, B. (2020). "Characterising the Digital Twin: A systematic literature review." *CIRP Journal of Manufacturing Science and Technology*, Elsevier Ltd, 29, 36–52.

Kan, C., and Anumba, C. J. (2019). "Digital Twins as the Next Phase of Cyber-Physical Systems in Construction." *Computing in Civil Engineering 2019: Data, Sensing, and Analytics - Selected Papers from the ASCE International Conference on Computing in Civil Engineering 2019*, American Society of Civil Engineers (ASCE), 256–264.

Kan, C., Fang, Y., Anumba, C. J., and Messner, J. I. (2018). "A cyber–physical system (CPS) for planning and monitoring mobile cranes on construction sites." *Proceedings of the Institution of Civil Engineers - Management, Procurement and Law*, ICE Publishing, 171(6), 240–250.

Khajavi, S. H., Motlagh, N. H., Jaribion, A., Werner, L. C., and Holmstrom, J. (2019). "Digital Twin: Vision, benefits, boundaries, and creation for buildings." *IEEE Access*, Institute of Electrical and Electronics Engineers Inc., 7, 147406–147419.

Klinc, R., and Turk, Ž. (2019). "Construction 4.0-digital transformation of one of the oldest industries." *Economic and Business Review for Central and South-Eastern Europe*, University of Ljubljana, Faculty of Economics, 21(3), 393–496.

Kong, T., Hu, T., Zhou, T., and Ye, Y. (2021). "Data Construction Method for the Applications of Workshop Digital Twin System." *Journal of Manufacturing Systems*, Elsevier B.V., 323–328.

LaGrange, E. (2019). "Developing a digital twin: The roadmap for oil and gas optimization." *Society of Petroleum Engineers - SPE Offshore Europe Conference and Exhibition 2019, OE 2019*, Society of Petroleum Engineers.

Lee, E. A. (2008). “Cyber physical systems: Design challenges.” *Proceedings - 11th IEEE Symposium on Object/Component/Service-Oriented Real-Time Distributed Computing, ISORC 2008*, 363–369.

Lee, J., Bagheri, B., and Kao, H. A. (2015). “A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems.” *Manufacturing Letters*, Elsevier Ltd, 3, 18–23.

Liu, M., Fang, S., Dong, H., and Xu, C. (2021). “Review of digital twin about concepts, technologies, and industrial applications.” *Journal of Manufacturing Systems*, Elsevier B.V., 58, 346–361.

Lu, Q., Parlikad, A. K., Woodall, P., Don Ranasinghe, G., Xie, X., Liang, Z., Konstantinou, E., Heaton, J., and Schooling, J. (2020). “Developing a Digital Twin at Building and City Levels: Case Study of West Cambridge Campus.” *Journal of Management in Engineering*, American Society of Civil Engineers (ASCE), 36(3), 05020004.

Madni, A., Madni, C., and Lucero, S. (2019). “Leveraging Digital Twin Technology in Model-Based Systems Engineering.” *Systems*, MDPI AG, 7(1), 7.

NASA. (2004). “The Ill-Fated Space Odyssey of Apollo 13.” <https://er.jsc.nasa.gov/seh/pg13.htm>.

National Infrastructure Commission. (2017). *Data for the public good*. *National Infrastructure Commission*.

Negri, E., Fumagalli, L., and Macchi, M. (2017). “A Review of the Roles of Digital Twin in CPS-based Production Systems.” *Procedia Manufacturing*, Elsevier B.V., 11, 939–948.

O’Connor, A. C., Dettbarn, J. L., and Gilday, L. T. (2004). “Cost Analysis of Inadequate Interoperability in the US Capital Facilities Industry.”

Oloufa, A. A., Ikeda, M., and Oda, H. (2003). “Situational awareness of construction equipment using GPS, wireless and web technologies.” *Automation in Construction*, Elsevier, 737–748.

Pan, Y., and Zhang, L. (2021). “A BIM-data mining integrated digital twin framework for advanced project management.” *Automation in Construction*, Elsevier B.V., 124, 103564.

Pauwels, P., Krijnen, T., Terkaj, W., and Beetz, J. (2017). “Enhancing the ifcOWL ontology with an alternative representation for geometric data.” *Automation in Construction*, Elsevier, 80, 77–94.

Pauwels, P., and Terkaj, W. (2016). “EXPRESS to OWL for construction industry: Towards a recommendable and usable ifcOWL ontology.” *Automation in Construction*, Elsevier, 63, 100–133.

Pires, F., Cachada, A., Barbosa, J., Moreira, A. P., and Leitao, P. (2019). “Digital twin in industry 4.0: Technologies, applications and challenges.” *IEEE International Conference on Industrial Informatics (INDIN)*, Institute of Electrical and Electronics Engineers Inc., 721–726.

Qi, Q., Tao, F., Hu, T., Anwer, N., Liu, A., Wei, Y., Wang, L., and Nee, A. Y. C. (2021). “Enabling technologies and tools for digital twin.” *Journal of Manufacturing Systems*, Elsevier B.V., 58, 3–21.

Rausch, C., Sanchez, B., Esfahani, M. E., and Haas, C. (2020). “Computational Algorithms for Digital Twin Support in Construction.” *Construction Research Congress 2020: Computer Applications - Selected Papers from the Construction Research Congress 2020*, American Society of Civil Engineers (ASCE), 191–200.

Roxin, A., Abdou, W., Ginhac, D., Derigent, W., Dragomirescu, D., and Montegut, L. (2019). “Digital building twins - Contributions of the ANR McBIM project." *Proceedings - 15th International Conference on Signal Image Technology and Internet Based Systems, SISITS 2019*, Institute of Electrical and Electronics Engineers Inc., 404–410.

Ruppert, T., and Abonyi, J. (2020). "Integration of real-time locating systems into digital twins." *Journal of Industrial Information Integration*, Elsevier B.V., 20, 100174.

Sacks, R., Brilakis, I., Pikas, E., Xie, H. S., and Girolami, M. (2020). "Construction with digital twin information systems." *Data-Centric Engineering*, 1.

Sepasgozar, S. M. E. (2021). "Differentiating Digital Twin from Digital Shadow: Elucidating a Paradigm Shift to Expedite a Smart, Sustainable Built Environment." *Buildings*, MDPI AG, 11(4), 151.

Shafto, M., Conroy, M., Doyle, R., Glaessgen, E., Kemp, C., LeMoigne, J., and Wang, L. (2012). "Modeling, simulation, information technology & processing roadmap." *National Aeronautics and Space Administration*.

Shawki, K. M., Kilani, K., and Gomaa, M. A. (2015). "Analysis of earth-moving systems using discrete-event simulation." *Alexandria Engineering Journal*, Elsevier, 54(3), 533–540.

T., O. J., and Li-Ren, Y. (2004). "Project Performance versus Use of Technologies at Project and Phase Levels." *Journal of Construction Engineering and Management*, American Society of Civil Engineers, 130(3), 322–329.

Tao, F., Zhang, H., Liu, A., and Nee, A. Y. C. (2019). "Digital Twin in Industry: State-of-the-Art." *IEEE Transactions on Industrial Informatics*, IEEE Computer Society, 15(4), 2405–2415.

Tchana, Y., Ducellier, G., and Remy, S. (2019). "Designing a unique Digital Twin for linear infrastructures lifecycle management." *Procedia CIRP*, Elsevier B.V., 545–549.

Turner, C. J., Oyekan, J., Stergioulas, L., and Griffin, D. (2021). "Utilizing Industry 4.0 on the Construction Site: Challenges and Opportunities." *IEEE Transactions on Industrial Informatics*, IEEE Computer Society, 17(2), 746–756.

Uhlemann, T. H. J., Lehmann, C., and Steinhilper, R. (2017). "The Digital Twin: Realizing the Cyber-Physical Production System for Industry 4.0." *Procedia CIRP*, Elsevier B.V., 335–340.

Walker, M. (2018). "Hype cycle for emerging technologies." <www.gartner.com>.

Wanasinghe, T. R., Wroblewski, L., Petersen, B. K., Gosine, R. G., James, L. A., De Silva, O., Mann, G. K. I., and Warrian, P. J. (2020). "Digital Twin for the Oil and Gas Industry: Overview, Research Trends, Opportunities, and Challenges." *IEEE Access*, Institute of Electrical and Electronics Engineers Inc.

Whitlock, K., Abanda, F. H., Manjia, M. B., Pettang, C., and Nkeng, G. E. (2018). "BIM for Construction Site Logistics Management." *Journal of Engineering, Project, and Production Management*, Walter de Gruyter GmbH, 8(1), 47–55.

Xu, Y., Bondaletova, N. F., Kovalev, V. I., and Komrakov, A. V. (2018). "Digital Twin Concept in Managing Industrial Capital Construction Projects Life Cycle." *Proceedings of 2018 11th International Conference "Management of Large-Scale System Development", MLSD 2018*, Institute of Electrical and Electronics Engineers Inc.

Yang, L.-R., Huang, C.-F., and Wu, K.-S. (2011). "The association among project manager's leadership style, teamwork and project success." *International Journal of Project Management*, 29(3), 258–

267.

Yang, X., Huang, Y., and Zhang, Q. (2020). “Automatic Stockpile Extraction and Measurement Using 3D Point Cloud and Multi-Scale Directional Curvature.” *Remote Sensing*, MDPI AG, 12(6), 960.

Zheng, Y., Yang, S., and Cheng, H. (2019). “An application framework of digital twin and its case study.” *Journal of Ambient Intelligence and Humanized Computing*, Springer Verlag, 10(3), 1141–1153.

Zhou, H. A., Gannouni, A., Otte, T., Odenthal, J., Abdelrazeq, A., and Hees, F. (2020). “Towards a Digital Process Platform for Future Construction Sites.” *2020 ITU Kaleidoscope: Industry-Driven Digital Transformation, ITU K 2020*, Institute of Electrical and Electronics Engineers Inc.

This page is intentionally left blank

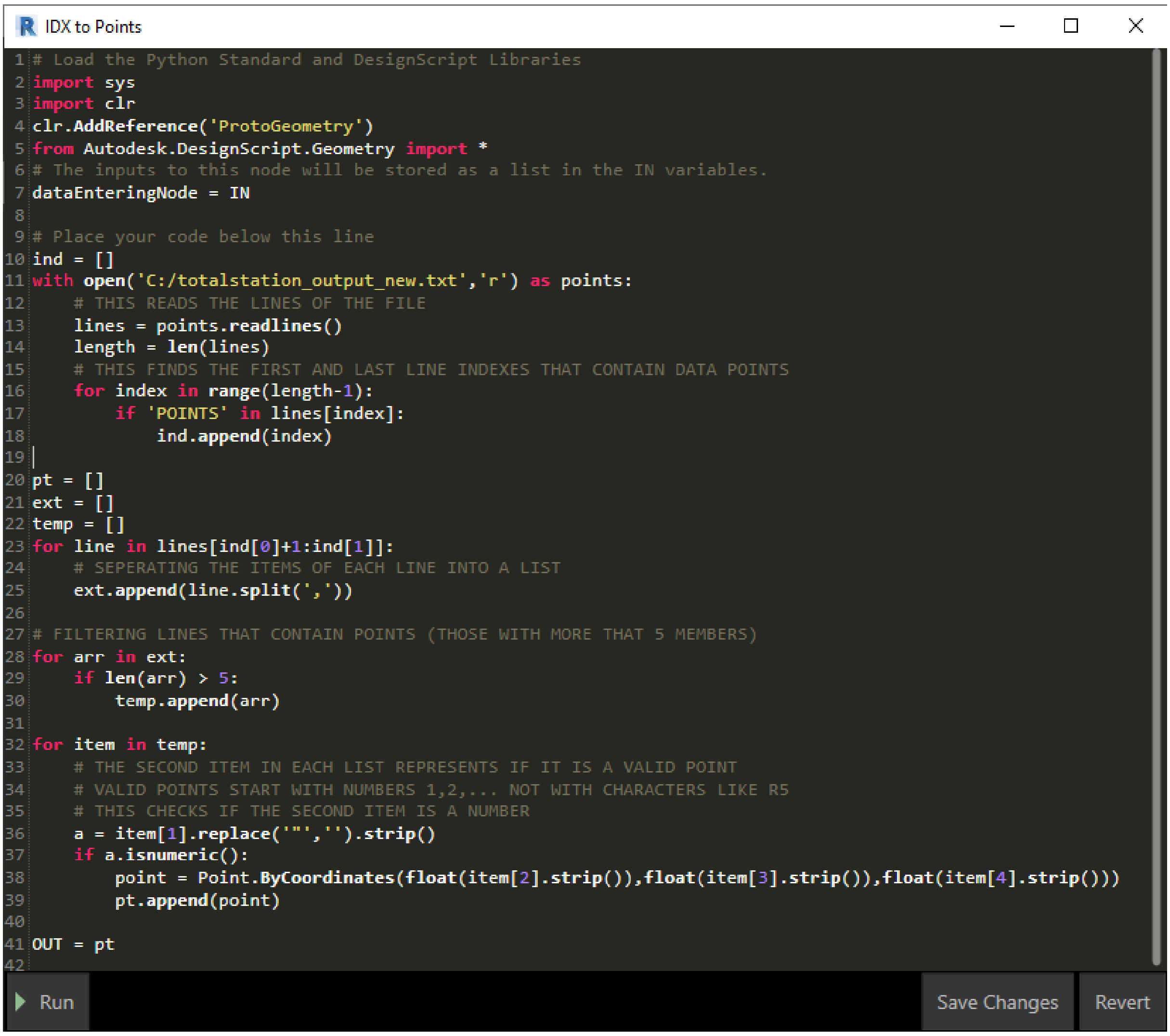
7. LIST OF ACRONYMS AND ABBREVIATIONS

3D	Three Dimensional
AI	Artificial Intelligence
AIM	Asset Information Model
AR	Augmented Reality
BIM	Building Information Modelling
CDE	Common Data Environment
CDT	Construction Digital Twin
CPS	Cyber Physical Systems
DT	Digital Twin
ETRS	European Terrestrial Reference System
FM	Facility Management
GIS	Geographic Information Systems
GPS	Global Positioning System
ICT	Information Communication Technologies
IoT	Internet of Things
IVHM	Integrated Vehicle Health Management
KE	Knowledge Engine
ML	Machine Learning
NASA	National Aeronautics and Space Administration
NIST	National Institute of Standards and Technology
PDCA	Plan, Do, Check, Act
PII	Project Intent Information
PLM	Product Lifecycle Management
PSI	Project Status Information
PSK	Project Status Knowledge
RFID	Radio Frequency Identification
RNN	Recursive Neural Network
RTLS	Real-time Locating Systems
SDG	Sustainable Development Goals
SHM	Structural Health Monitoring
UWB	Ultra-wide Band
VR	Virtual Reality
WGS84	World Geodetic Reference System 84

8. APPENDICES

APPENDIX A: PYTHON SCRIPT

The python script shown in Figure 36 is used to extract points coordinates from the output files of site surveying tools.




```python
# Load the Python Standard and DesignScript Libraries
import sys
import clr
clr.AddReference('ProtoGeometry')
from Autodesk.DesignScript.Geometry import *
# The inputs to this node will be stored as a list in the IN variables.
dataEnteringNode = IN

# Place your code below this line
ind = []
with open('C:/totalstation_output_new.txt','r') as points:
    # THIS READS THE LINES OF THE FILE
    lines = points.readlines()
    length = len(lines)
    # THIS FINDS THE FIRST AND LAST LINE INDEXES THAT CONTAIN DATA POINTS
    for index in range(length-1):
        if 'POINTS' in lines[index]:
            ind.append(index)

pt = []
ext = []
temp = []
for line in lines[ind[0]+1:ind[1]]:
    # SEPERATING THE ITEMS OF EACH LINE INTO A LIST
    ext.append(line.split(','))

# FILTERING LINES THAT CONTAIN POINTS (THOSE WITH MORE THAT 5 MEMBERS)
for arr in ext:
    if len(arr) > 5:
        temp.append(arr)

for item in temp:
    # THE SECOND ITEM IN EACH LIST REPRESENTS IF IT IS A VALID POINT
    # VALID POINTS START WITH NUMBERS 1,2,... NOT WITH CHARACTERS LIKE R5
    # THIS CHECKS IF THE SECOND ITEM IS A NUMBER
    a = item[1].replace('"','').strip()
    if a.isnumeric():
        point = Point.ByCoordinates(float(item[2].strip()),float(item[3].strip()),float(item[4].strip()))
        pt.append(point)

OUT = pt
```

Figure 36- Data points extraction from site surveying tools output

APPENDIX B: DYNAMO SCRIPT

The dynamo script shown in Figure 37 creates the topography from the extracted points of the previous script and computes the volume of the topography.

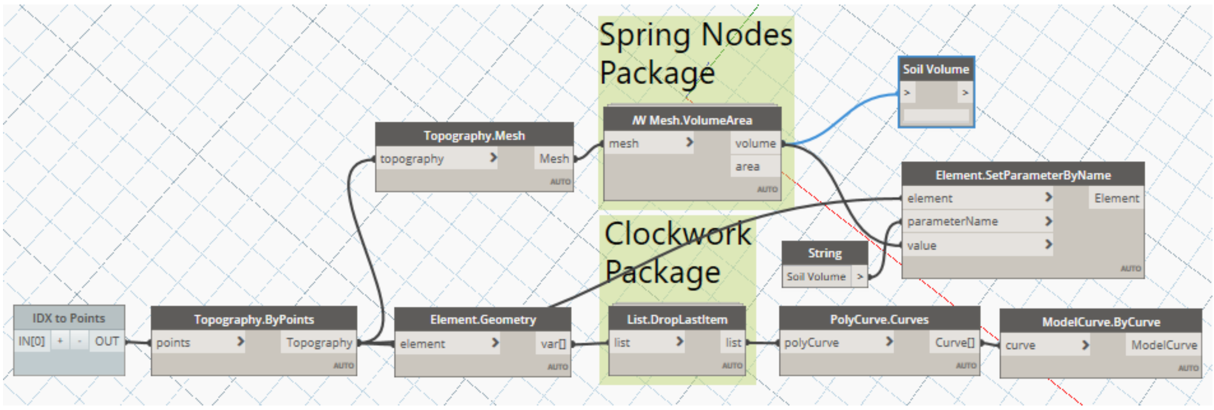


Figure 37- Creating topography from the extracted points and volume calculation

APPENDIX C: SIMULATION MODEL

Figure 38 depicts the simulation model for earthwork operations created in Arena simulation software.

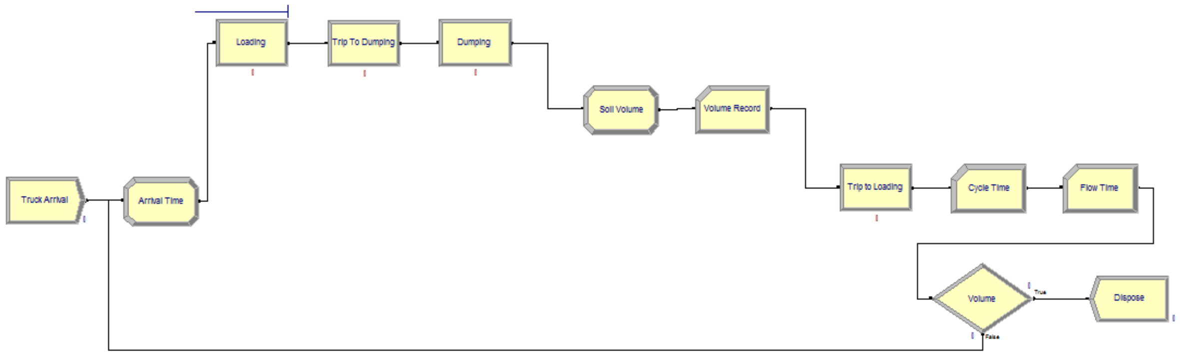


Figure 38- Earthwork simulation model in Arena