DIGITAL TWINS
IN MECHANISED MINING

PROGRAMME	MECHANISED MINING SYSTEMS
PROGRAMME MANAGER	M PRETORIUS
RESEARCH AGENCY	UNIVERSITY OF PRETORIA
RESEARCH CENTRE	RESEARCH ACTIVITY IN MECHANISED MINING SYSTEMS (RAMMS)
DATE	01-04-2024
REVISION	1.0
AUTHOR(S)	LD VAN EYK, BRIAN ELLIS, PS HEYNS
PROJECT TEAM	LD VAN EYK, BRIAN ELLIS, PS HEYNS

MANDELA MINING PRECINCT
MINDS FOR MINES

AD DESTINATUM PERSEQUOR

This report clarifies the concept of digital twins, often misunderstood due to variations influenced by industry, lifecycle phases, and maturity levels. We propose a framework tailored to the South African context, emphasizing five dimensions and eight properties. Through case studies in mining, we demonstrate the framework's value and alignment with real-world applications. Using this framework, we further identify potential digital twinning projects within the RAMMS program.

1 INTRODUCTION

South African mining companies are actively embracing digitalization as they strive to leverage the promised benefits of the fourth industrial revolution (4IR) (PwC, 2023). This transformative era is characterized by rapid technological advancements and large-scale data collection, with digital twins (DTs) emerging as a key enabling technology of the 4IR (Ghobakhloo, 2020). DTs offer the potential for mining companies to effectively utilize the vast amount of data available to them through 4IR, transforming it into actionable and interpretable insights (Tao et al., 2019). Interpretable, actionable insights are current challenges faced by the mining industry and literature alike (PwC, 2023), emphasizing the significance of digital twins in this context.

Yao et al. (2023) indicate that the last few years of DT literature have been indicative of industry penetration stage, with global players introducing policies surrounding DTs. In addition, standardisation organisations such as ISO, IEEE, IEC and ITU are introducing industry standards around DTs. Standardized DTs hold promise for addressing interoperability issues in mining by integrating disparate datasets, models, and systems into a single framework.

DTs are not only gaining traction in mining but also across various sectors. In 2017, (Tao et al., 2017) accurately foresaw that DT popularity would surge in the coming years. From 2017 to 2019, Gartner Consulting listed DTs as one of the top 10 strategic technologies and the German Information Technology and New Media Association (BITKOM) indicate that the expected value of DTs in the manufacturing market alone will exceed 78 billion euros by 2025 (Wu et al., 2020).

DTs were previously largely ignored by researchers due to a limitation of the Internet of Things (IoT) and data processing technologies. However, with these technologies becoming ever more mature, DT research is set to enter a rapid growth stage (Yao et al., 2023). Therefore, it is paramount that the South African mining industry stay cognisant of this fact and reap the benefits of this technological advancement from its earliest stages. Given the varying perceptions of DTs within the mining industry and literature, our objective is to establish a clear understanding of DTs, tailor them to the South African mining context, and maximize the value derived from the digitalisation effort in South African mining.

The primary contribution of this work is therefore the development of a DT framework, which:

- Establishes a common understanding of DTs which:
    - Reduces ambiguity.
    - Aligns industry and research perspectives.
    - Avoids duplication of work.
    - Creates a perspective of where various research works fit into the bigger picture of DTs.
- Facilitates the adoption of DT technologies by clearly articulating their value.
- Identifies opportunities for leveraging newfound data.

To address the first point, we start by highlighting current DT definitions, frameworks and properties in section 2 in an attempt to understand the nuance associated with the term, identify common traits and ultimately understand which components are crucial for developing a DT. In section 3, we try to incorporate these findings into a common framework, specifically tailored to South African mining operations.

Section 4 showcases various mining DT case studies in the context of the proposed framework, addressing the second point. These case studies not only illustrate the value derived from DTs, but also show the value of the proposed framework for identifying the core components of each case study.

The final point is addressed in section 5, which further illustrates the value of the proposed framework by using this tool to identify future potential DT applications in South African mining. These applications are derived from a mechanised mining systems (MMS) perspective and are by no means exhaustive.

Section 6 closes with concluding remarks, highlighting the contents of the report. Ultimately, the aim of this work is to equip readers with a comprehensive understanding of DTs and supply them with tools to identify DT opportunities within their respective domains.

2 WHAT IS A DIGITAL TWIN?

Defining a digital twin (DT) presents a challenge due to its fluid nature, influenced by factors such as time (Yao et al., 2023), lifecycle phase (Thelen et al., 2022b), industry (Schweiger & Barth, 2023), and maturity (Autodesk, 2022; Kritzinger et al., 2018). Schweiger and Barth (2023) emphasize that DTs have various properties and characteristics and therefore each DT is essentially unique, explaining why a single definition remain elusive. Tao (2017) notes that the absence of a universal DT definition has prompted researchers to simplify DTs into broad concepts, to accommodate varied perspectives. It is therefore essential to explore these perspectives comprehensively to grasp the essence of DTs in their entirety.

This section begins by examining the evolving landscape of DT definitions, considering a diverse range of influencing factors. The subsequent section reviews existing frameworks aimed at capturing the essence of DTs. The final section discusses the most important properties of a DT, offering insights beyond conventional definitions and frameworks.

2.1 DIGITAL TWIN DEFINITIONS

The evolution of digital twin (DT) definitions has been influenced by various factors, with time playing a significant role. While rudimentary forms of DTs date back to their use by the National Aeronautics and Space Administration (NASA) in 1970, it was Professor Michael Grieves who formalized the concept in 2003 during his course on product lifecycle management (Schweiger & Barth, 2023; Yao et al., 2023). Professor Grieves' definition encompassed three components: a physical product, a virtual product, and their interconnections (Tao et al., 2019; Yao et al., 2023). In 2011, Grieves and John Vickers from NASA coined the term "digital twin," which was subsequently defined by NASA in 2012, primarily within the aerospace industry context. They define a DT as an "integrated multiphysics, multiscale, probabilistic simulation of an as-built vehicle or system that uses the best available physical models, sensor updates, fleet history etc., to mirror the life of its corresponding flying twin." (Glaessgen & Stargel, 2019). (Thelen et al., 2022b) indicates widespread acceptance of this definition, cited in various works like Tao et al. (2019) and Schweiger and Barth (2023) but often applied beyond aerospace contexts.

As more researchers entered the field, updated DT definitions proliferated. Between 2012 and 2015, emphasis shifted towards mapping physical reality to virtual space and simulating physical phenomena within this virtual environment. By 2015, researchers recognized the expansive nature of DTs, comprising a multitude of digital components requiring integrated architectures. In 2017, Professor Grieves highlighted fidelity as a crucial aspect, advocating for mapping products at micro-atomic levels alongside macro-geometric levels. Concurrently, definitions began incorporating the notion of control, emphasizing DTs' role in not just modeling but also influencing physical reality. Subsequent definitions from 2018 onwards emphasized modeling at various scales and over the entire asset lifecycle, employing multi-dimensional, multi-physical, and multi-temporal approaches (Yao et al., 2023). The evolution of DT definitions over time is mainly attributed to technological advancements which advance the capabilities of the DTs (Grieves, 2014). Grieves and Vickers (2017) indicate that in 2003, most information that was collected was paper-based and manual. By contrasting this reality with today's digitalizing mining environment (PwC, 2023), it becomes apparent that the capabilities, and hence definitions, of DTs have evolved alongside technology.

Moreover, research by Tao et al. (2019) and Thelen et al. (2022b) suggest that the evolution of DT definitions is not solely time-dependent but also influenced by the lifecycle stage of the physical entity being twinned. Figure 1 illustrates various DT models and their associated communication technologies, demonstrating how DT functionalities and hence definitions may vary depending on the entity's lifecycle stage. Typically, the design phase is associated with twins which reconfigure or optimize the entity towards some goal before it is manufactured. The real-time requirement of a twin is less strict in this phase, as the entity is yet to be manufactured. The manufacturing stage is concerned with the equipment manufacturing the entity being twinned. In this phase, DTs are mostly used for to support decisions around process control, production planning and maintenance of manufacturing equipment. In the service phase, the focus is on the entity itself, optimizing towards certain goals such as path planning or minimizing downtime through predictive maintenance strategies. These examples clearly illustrate that the twinned element may form part of multiple DTs throughout its lifecycle.

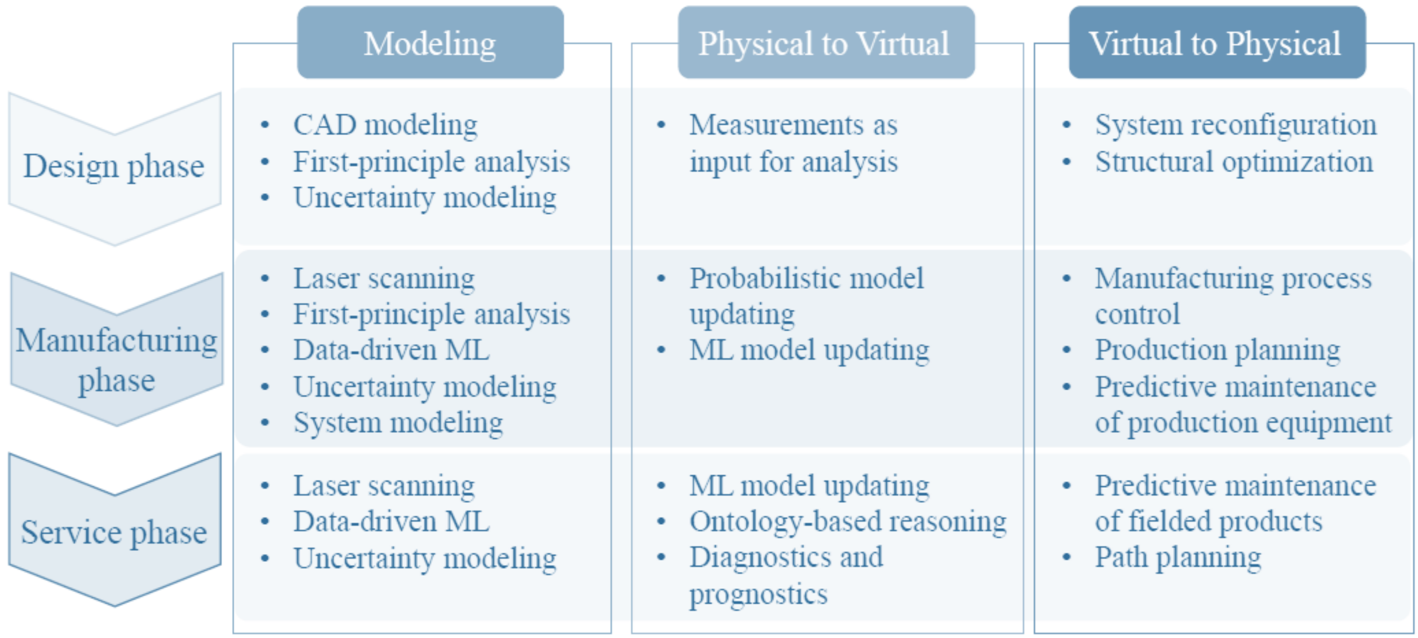


Figure 1 Illustration of various digital twin functionalities as a function of the physical entity's lifecycle, adapted from Thelen et al. (2022b). The entity is shown to move from design phase, through a manufacturing phase to the service phase. Throughout all three phases, the entity may form part of multiple digital twins, each associated with each lifecycle phase.

Another factor which influences DT definitions, is the context in which they are applied. DTs are dependent on their individual use cases, leading to various configurations in practice. It is therefore difficult to identify common characteristics of a digital twin because each context is unique (Schweiger & Barth, 2023). By extension, each industry's definition of a digital twin is therefore also expected to be unique.

A final consideration which complicates a universal DT definition, is model maturity. DTs are not necessarily fully mature at inception and therefore it is useful to look at definitions based on maturity levels of DTs. Many definitions mention a digital counterpart to the physical reality. (Kritzinger et al., 2018) argue that a misinterpretation of the relationship between these elements has caused various related models to be labelled as digital twins, when in fact they are not. They distinguish between digital models, digital shadows and digital twins, which are related, but ultimately differ by their levels of data integration and hence maturity.

Figure 2 showcases the difference between the three types of digital counterparts. In their framework, the sole property that defines a DT, is the level of data interconnectedness. Automatic data flow refers to data flow that does not require human-intervention. For example, data capturing systems constitute manual data flow, whereas sensor readings constitute automatic data flow.

A DM is a digital representation of an existing or planned physical object, but does not use any form of automated data exchange. Changes in the physical state are not automatically reflected in the digital object, and vice versa. For example, CAD models are seen as DMs, because they are manually updated as new information becomes available. Furthermore, after being updated, they do not impact the physical asset automatically, but rather, manual communication is used to re-design or change the physical asset.

Digital shadows are characterised by automated one-way data flow from the physical entity to the digital object. A change in physical state is automatically reflected in the digital object, however, the opposite is not true. Digital shadows are regarded as forerunners to DTs. Consider a dashboard that showcases health states of various equipment. This dashboard acts as a digital shadow because the digital health-predicting model is continually updated with new sensor data, however, this model never influences the physical object.

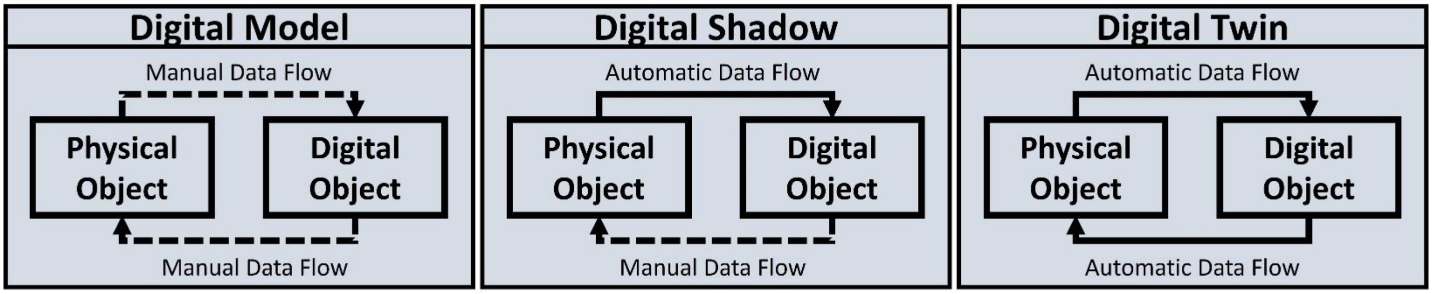


Figure 2 Illustration of a digital model, digital shadow and digital twin. A digital model is characterized by manual data flow from the physical and to the physical object. A digital shadow has improved data flow, with automatically flowing data from the physical to digital object. A digital twin ultimately has automatic data flow in both directions.

A DT is characterised by bi-directional data flow. In this case, the physical object state is automatically reflected in the digital object, but more importantly, the digital object may act as a controlling instance of the physical object. For example, a drilling machine may measure vibrations, temperatures, rotation speeds etc and use this data to determine the degradation of the drill bit. Based on this conclusion, the model could then lower or increase the drilling speed to decrease wear rate. In this example, the data not only flows to the digital object (sensor readings), but communication returns to the physical object to control its behaviour.

Kritzinger et al. (2018) showcased that based on this definition of a DT, only 18% of the reviewed papers truly entailed bi-directional data flow. That is to say, 82% of research papers classify DMs or shadows as DTs. This large body of mis-labelled works illustrates one of the many reasons the DT concept varies so widely.

Autodesk takes a different approach to labelling DTs according to maturity, defining five levels of DTs (https://www.autodesk.com/design-make/articles/what-is-a-digital-twin).

Level 1 twins are called a descriptive twins and are editable visual replicas of the physical reality. This definition is similar to the one originally proposed by Professor Grieves, which describes a DT as "a digital informational construct", which should "optimally include all information concerning the system asset that could be potentially obtained from its thorough inspection in the real world" (Kritzinger et al., 2018). Considering the definitions outlined by Kritzinger et al. (2018), a level 1 twin would be classified as a digital model, not a DT.

Level 2 twins are called informative twins and are integrated with sensors, allowing insights at any time. This definition would be classified as a digital shadow, according to Kritzinger et al. (2018). Level 3 twins are called predictive twins, capturing real-time data (as in level 2), but further aim to generate insights to identify potential issues. Kritzinger et al. (2018) would likely still identify this delineation as a digital shadow, as data is only flowing in one direction. Level 4 twins are called comprehensive twins, leveraging advanced modelling and simulation technologies to predict future scenarios and make recommendations. The important difference from a level 3 twin, is the ability to forecast. Although the flow of data feeds back to the physical model through actionable insights, they are ultimately not automatically executed and will therefore still fall under a digital shadow definition according to the proposed framework from Kritzinger et al. (2018).

Level 5 twins are called autonomous twins and can gather insights and act on these insights automatically. This is the first definition which has automatic dataflow in both directions, implying it would constitute as a digital twin within the framework proposed by Kritzinger et al. (2018).

An interesting distinction appears when comparing the DT definitions of Kritzinger et al. (2018) and Autodesk (2022). Kritzinger et al. (2018) emphasize the data interconnectedness as the defining property of a DT whereas Autodesk (2022) adopt a more fluent definition of a DT, rather distinguishing between the maturity of a DT. A combined outlook is perhaps more useful, where it is understood that digital models and digital shadows are stepping stones on a trajectory towards DTs.

In this first sub-section, the complications associated with DT definitions have been highlighted, with the term developing over time, lifecycle stage, industry of application, and maturity. In response to these difficulties, multiple works have proposed general frameworks to describe DTs. The following section outlines some of the recent frameworks and highlight useful insights from each.

2.2 DIGITAL TWIN FRAMEWORKS

In response to the evolving landscape of DT definitions, various authors have developed generalized frameworks to elucidate the concept. These frameworks offer diverse perspectives and understandings of DTs, laying the groundwork for our proposed framework tailored to South African mining operations.

Professor Grieves, credited with coining the term "digital twin," introduced a foundational three-dimensional framework (Grieves, 2014) depicted in Figure 3. This framework delineates a real space housing physical products, a virtual space housing multiple virtual products, and the connections bridging these domains. Building upon this concept, Tao et al. (2018) expanded the framework by introducing two additional dimensions pertaining to services and DT data (Figure 4). While comprehensive, this framework primarily focuses on twinning physical entities, potentially overlooking processes or systems.

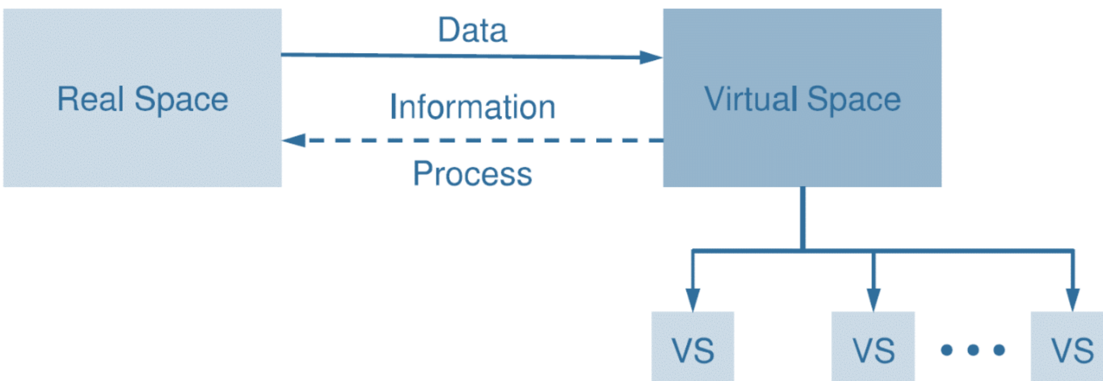


Figure 3 Three-dimensional digital twin proposed by Grieves (2014). Image adapted from Sjarov et al. (2020). In the real space, one has a physical product that is twinned. In the virtual space, a plurality of virtual products exists, each representing a facet of the physical product. These two spaces are joined by a connection, with regular data and information flow between the two spaces.

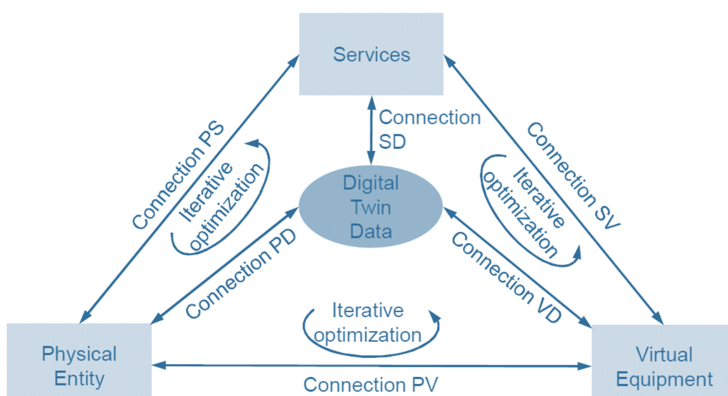


Figure 4 Five-dimensional digital twin proposed by Tao et al. (2018). Image adapted from Sjarov et al. (2020). The five dimensions depict the physical entity, virtual equipment, services, digital twin data and connections. The physical entity refers to a technical system which is equipped with sensors. Virtual equipment refers to multiple distinct models describing the geometry, physical properties, behaviour and rules of the entity. Services refer to functions that can operate and inform both physical entities and virtual equipment. The digital twin data component refers to the element in the digital twin which stores, fuses and enriches physical and virtual data. Finally, connections exist between these four components, exchanging data in a bi-directional manner where necessary.

Stark and Damerau (2019) developed a more expansive framework delineating eight crucial dimensions of a DT, each with varying levels (see Figure 5). These dimensions encompass integration breadth, connectivity mode, update frequency, cyber-physical system (CPS) intelligence, simulation capabilities, digital model (DM) richness, human interaction mode, and product lifecycle stage. For dimensions 1,2,7 and 8, the differing levels refer to the degree of enrichment. For the remaining dimensions, the levels describe different categories that are applicable to the specific dimension. These elements will be discussed in greater detail in the following sub-section.

Another noteworthy framework is proposed by Thelen et al. (2022b), building upon the foundational dimensions introduced by Grieves (2014). Illustrated in Figure 6, this framework adds a two-dimensional communication aspect and an additional dimension for model optimization to the existing three dimensions of Grieves (2014). The first dimension delineates the physical entity with a set of states, while the second dimension pertains to the digital system, also with a set of states. These digital states are related but distinct from those of the physical entity. The third dimension encompasses the process of estimating the digital state from physical state measurements, while the fourth dimension captures decision-making algorithms facilitating the reflection of updated digital system output in the physical entity. Lastly, a fifth dimension describes optimization processes across all twin dimensions, depicting offline and online DT optimization methods.

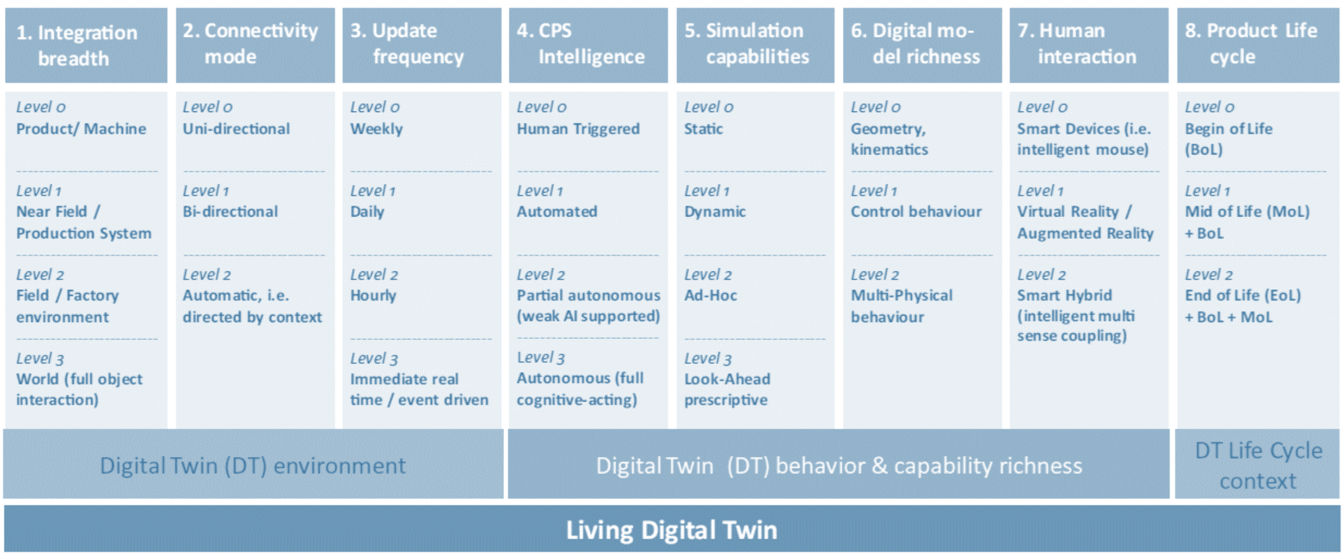


Figure 5 Eight-dimensional digital twin proposed by Stark and Damerau (2019). Image adapted from Stark and Damerau (2019). Integration breadth describes the breadth and scope of the physical reality being twinned. Connectivity mode refers to the connectivity capabilities which will result in a successful digital twin. Update frequency refers to the interval between twin updates and are dependent on the specific application. Cyber Physical System (CPS) intelligence refers to the maturity of the mechanism which causes the interaction with the physical reality. Simulation capabilities describe the use-case for the digital twin at varying levels of maturity. Digital model richness describes which characteristics of the physical reality are mapped to the digital twin. The human interaction dimension delineates the digital twin interface. Finally, product life cycle describes the phase of the physical product's life during which it is twinned.

2.3 PROPERTIES OF DIGITAL TWINS

While DT definitions and frameworks offer valuable insights, they often oversimplify the complexities of the concept. Therefore, breaking down DTs into their fundamental properties provides a more complete understanding of the concept. This section highlights some of the properties prevalent throughout recent literature, although the specifics may vary between works. These properties encompass three major areas: the physical reality, digital counterpart, and the connections between them.

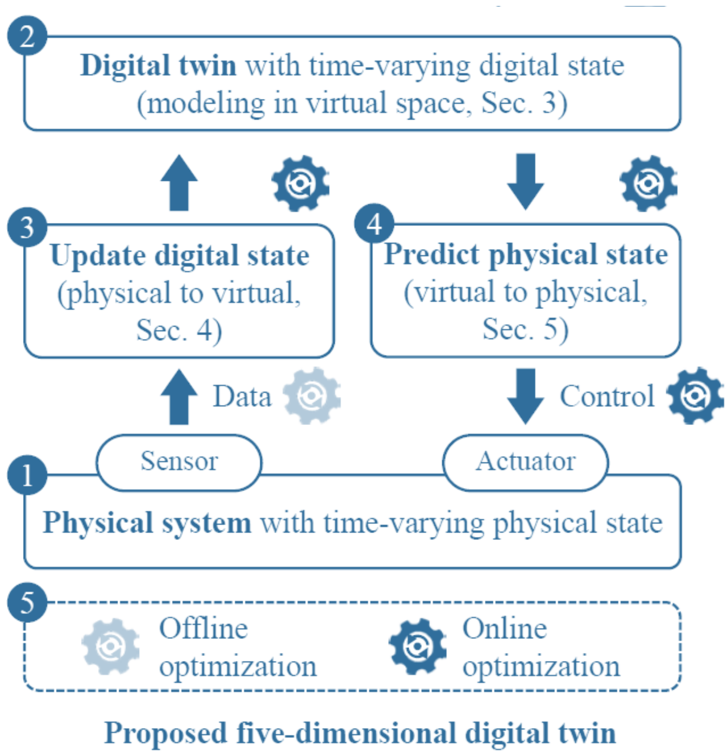


Figure 6 Five-dimensional digital twin proposed by Thelen et al. (2022b). Image adapted from Thelen et al. (2022b). The first dimension refers to the physical system, which is the entity to be twinned. The second dimension is assigned to the digital system, which acts as the digital model that generates actionable insights. The connection from the physical system to the digital system describes the third dimension, with the reverse direction being attributed to the fourth. The final dimension encapsulates all other dimensions and pertains to offline and online optimization of digital twin elements.

2.3.1 PHYSICAL REALITY

Counterpart: DTs typically involve a physical counterpart, ranging from tangible objects to abstract systems or processes. Earlier definitions focused on physical products (Grieves, 2014), but recent insights recognize the broader scope, encompassing non-physical entities with a recognisably distinct existence and relevance for creating value. Consequently, a distinction between physical, non-physical or any distinct value-generating entity may be made (Schweiger & Barth, 2023).

Thelen et al. (2022b) indicate that a single DT may not be sufficient to describe a single component, with multiple failure modes being modelled by different twins. Therefore, the physical counterpart may be a single failure mechanism being modelled on a single component.

Stark and Damerau (2019) distinguish the physical counterpart according to integration breadth, referring to product/machine level, near field/production system level, field/factory environment level and world level. Therefore, in this case the emphasis is on application scope.

Data source: A fundamental component of a DT is data collection. These data sources vary and do not originate from sensors alone. Schweiger and Barth (2023) indicate that data may also be extracted from systems related to product data management (PDM), enterprise resource planning (ERP) and customer relationship management (CRM) as well as programs like computer-aided design (CAD). Furthermore, data may originate from third-party systems. Collectively, Schweiger and Barth (2023) identify three distinct sources of data, namely sensors, internal systems, and external systems.

2.3.2 DIGITAL COUNTERPART

Fidelity: In many instances, particularly the first definitions of DTs, authors such as Grieves (2014) mention a high-fidelity model as a requirement for a DT. However, Schweiger and Barth (2023) highlight that for many industrial applications this is impractical or altogether impossible. Therefore, they relax this requirement and note that some DTs can have a "sufficient" fidelity. This term arises from the Industrial Internet Consortium definition: "[a] digital representation, sufficient to meet the requirements of a set of use cases". Stark and Damerau (2019) highlight a similar property, indicating that DM richness may be confined to geometry and kinematics, control behaviour and multi-physical behaviour. Modern requirements of digital model fidelity are therefore more relaxed and are ultimately determined by the use case.

Capability: Schweiger and Barth (2023) identify seven distinct capabilities of DTs, mainly based on DT function, namely simulation, optimisation, prediction, detection, prevention and automation. These capabilities do not need to operate in isolation and can occur in combination. They also point out that in other frameworks, such as the one proposed by Tao et al. (2018), these capabilities are labelled as "services".

Stark and Damerau (2019) define capabilities according to the DT's reaction to data. They delineate four levels, namely static, dynamic, ad-hoc and look-ahead prescriptive. Static refers to scenarios where the simulation model is not time-dependent (for example CAD). Dynamic capabilities refer to models where the parameters are time-dependent (for example CFD). Ad hoc describes real-time DTs, which act on data as it becomes available. Look-ahead prescriptive models are equipped with predictive simulation capabilities and aim to predict future events as new data becomes available.

Therefore, the conclusion is that digital twins have a wide range of capabilities, and constraining the definition to a single aspect is limiting.

Purpose: According to Schweiger and Barth (2023), the ultimate purpose of DTs is to optimise the efficiency of operations. Based on this requirement, they use this definition to tie the purpose of a DT to increasing the overall equipment effectiveness (OEE), which is a function of performance, availability and quality. Therefore, the purpose of a DT can be said to positively impact one or more of these items.

Life cycle: DTs are useful at various stages of a product life cycle. For this reason, both Stark and Damerau (2019) and Schweiger and Barth (2023) indicate that a DT can be applied at the beginning, middle or end of life. Thelen et al. (2022b) follow a similar delineation and classify this property according to the physical counterpart's product lifecycle phase, discerning between design, manufacturing and service phase.

2.3.3 CONNECTION BETWEEN PHYSICAL AND DIGITAL WORLDS

Synchronisation: A crucial requirement for a DT is synchronisation between the physical and digital world, with many definitions specifically stating "real-time" synchronisation. However, Schweiger and Barth (2023) highlight that the synchronisation frequency is dependent on the use case, and in some cases near real-time or periodic synchronization may be sufficient. Schweiger and Barth (2023) further highlight that in some applications a combination of synchronization periods may be required depending on the use case. For example, battery power may become a limiting factor, dynamically requiring the synchronization interval to change. Schweiger and Barth (2023) ultimately distinguish between real-time, near real-time and periodic synchronisation. Stark and Damerau (2019) propose a more rigid delineation with four levels of synchronisation, namely weekly, daily, hourly and immediate or event driven, which presumably is any DT acting on the sub-hour scale.

Data link: The data link refers to the communication between the DM and the physical reality. Most works indicate that a bi-directional connection is a requirement for a DT. That is, information flows between the physical reality and the digital model continuously. Recall the discussion on digital models, shadows and twins in section 2.1, where only digital twins possessed the property of bi-directional, automatic data flow.

Schweiger and Barth (2023) argue that in many industrial applications, uni-directional communication from the physical entity to the DM is sufficient, given that the output of the DM is actionable by a human. (Stark & Damerau, 2019) make a similar distinction between uni- and bi-directional connections and include both options as valid properties of a DT. A uni-directional data link is therefore characterised by automatic updating of the DM, but manual updating of the physical reality. This is in stark contrast to the earlier work from Kritzinger et al. (2018), which explicitly classifies the uni-directional data link as a property of digital shadows. Thelen et al. (2022b) contribute a different perspective, outlining certain scenarios in which a uni-directional link may be sufficient for DT applications. For control applications, bi-directional flow is necessary, as automatic decision making is required. However, where DTs are used for support decision making, for example predictive maintenance, one usually finds a human in the loop that carries out said decision. This implies that a uni-directional connection is sufficient for the DT's purpose in this scenario. Thelen et al. (2022b) therefore argue that the context determines whether a uni-directional connection constitutes a digital shadow or a DT.

Interface: The interface refers to the action-mechanism from the DM to the physical reality. Schweiger and Barth (2023) delineate two options, namely human-to-machine or machine-to-machine interaction. This property ties in strongly with the data link property because the interface will determine which type of data link is present in the DT. Stark and Damerau (2019) highlight related properties called "Cyber-physical intelligence" and "human interaction". Cyber-physical intelligence refers to the intelligence of the DT, with the lowest level using humans, and the highest levels indicating full autonomous behaviour. The human interaction property relates to human interfaces with the DT and outline technologies such as smart devices, mixed reality and smart models that combine smart devices and mixed reality. Therefore, recent works acknowledge that a digital twin does not necessarily have to rely on electronic devices, but may require human intervention to close the communication loop.

2.4 CONCLUDING REMARKS

Section 2 has provided a comprehensive examination of digital twins (DTs), encompassing their evolving definitions, various frameworks and essential properties that characterize them. We have explored the nuanced landscape of DT definitions, acknowledging the influence of factors such as time, lifecycle stage, industry context, and maturity. Furthermore, the discussion on DT frameworks has underscored the diverse approaches taken by researchers to conceptualize and describe DTs, each offering unique insights into their structure and functionality. Finally, the exploration of DT properties has shed light on critical aspects related to the physical reality, digital counterpart, and the interconnectedness between the two realms. As we move forward into section 3, we aim to harness this knowledge to propose a tailored DT framework specifically designed to empower South African miners in leveraging their data effectively within the DT paradigm.

3 A DIGITAL TWIN FRAMEWORK FOR SOUTH AFRICAN MINING

In this section, we propose a digital twin (DT) framework tailored specifically for the mining industry, aiming to harness the immense potential of digitalization within this sector. Drawing inspiration from the DT frameworks in section 2.2 we have identified the need for a balanced yet comprehensive approach. While simplistic frameworks (Grieves, 2014) may overlook crucial nuances, overly complex frameworks (Stark & Damerau, 2019) may hinder practical implementation. The five-dimensional variant proposed by Thelen et al. (2022b) (see Figure 6) strikes a good balance between complexity and comprehensiveness.

In this section, we expand this framework by incorporating important findings from section 2 into the proposed framework. We therefore propose that a DT consists of five fundamental dimensions:

$$\mathrm{DT} = f\ (\mathrm{PE}, \mathrm{DM}, \mathrm{P2D}, \mathrm{D2P}, \mathrm{OPT})$$

The proposed five-dimensional DT model consists of a physical entity (PE), digital model (DM), a physical to digital (P2D) updating engine, a digital to physical (D2P) prediction engine, and an optimization (OPT) of the previous dimensions. Seamless integration of these five dimensions enables DTs to support optimal decision-making, in real-time. In recognition of important insights in section 2.3, we extend the five-dimensional framework to include eight properties, which describe the physical counterpart being twinned, data source, synchronisation frequency, twin acting interface, model fidelity, domain of application, capability and period in the lifecycle application point. The proposed framework is illustrated in Figure 7. By utilising the five dimensions and eight properties, we construct a digital twin definition which is well-aligned with Figure 7. Italic words in the definition indicate each of the eight proposed digital twin properties.

A digital twin is a pivotal tool for driving optimization across a physical entity. Comprising a digital model, physical *counterpart*, and intricate connections between the two, it embodies a *synchronized* flow of *data* between the physical entity and the digital model, facilitating continual estimation of the entity's state. This digital model, characterized by its *fidelity* and optimization *capabilities*, serves as the cornerstone for enhancing various application *domains* throughout the physical entity's *lifecycle*. Crucially, the digital twin's defining feature lies in its bidirectional connection, enabling *actionable interventions* to enforce the proposed optimisation actions.

The selected dimensions for this framework are tailored for the mining environment, where the likely counterpart of many digital twins are physical machines. Therefore, we reference “physical entities” as opposed to “real entities”, which include a broader spectrum of non-physical and abstract entities. The eight selected properties are also carefully selected with mining considerations in mind. The South African mining environment faces unique challenges and we believe the proposed eight properties capture many of the characteristics that will be common, but different between digital twins.

In the subsequent sections, the proposed framework is expanded to detail different considerations and model types that are used for each of the digital twin dimensions. The final sub-section showcases useful process diagrams for determining which components may be best suited to certain digital twin use cases, further aiding the adoption of this framework.

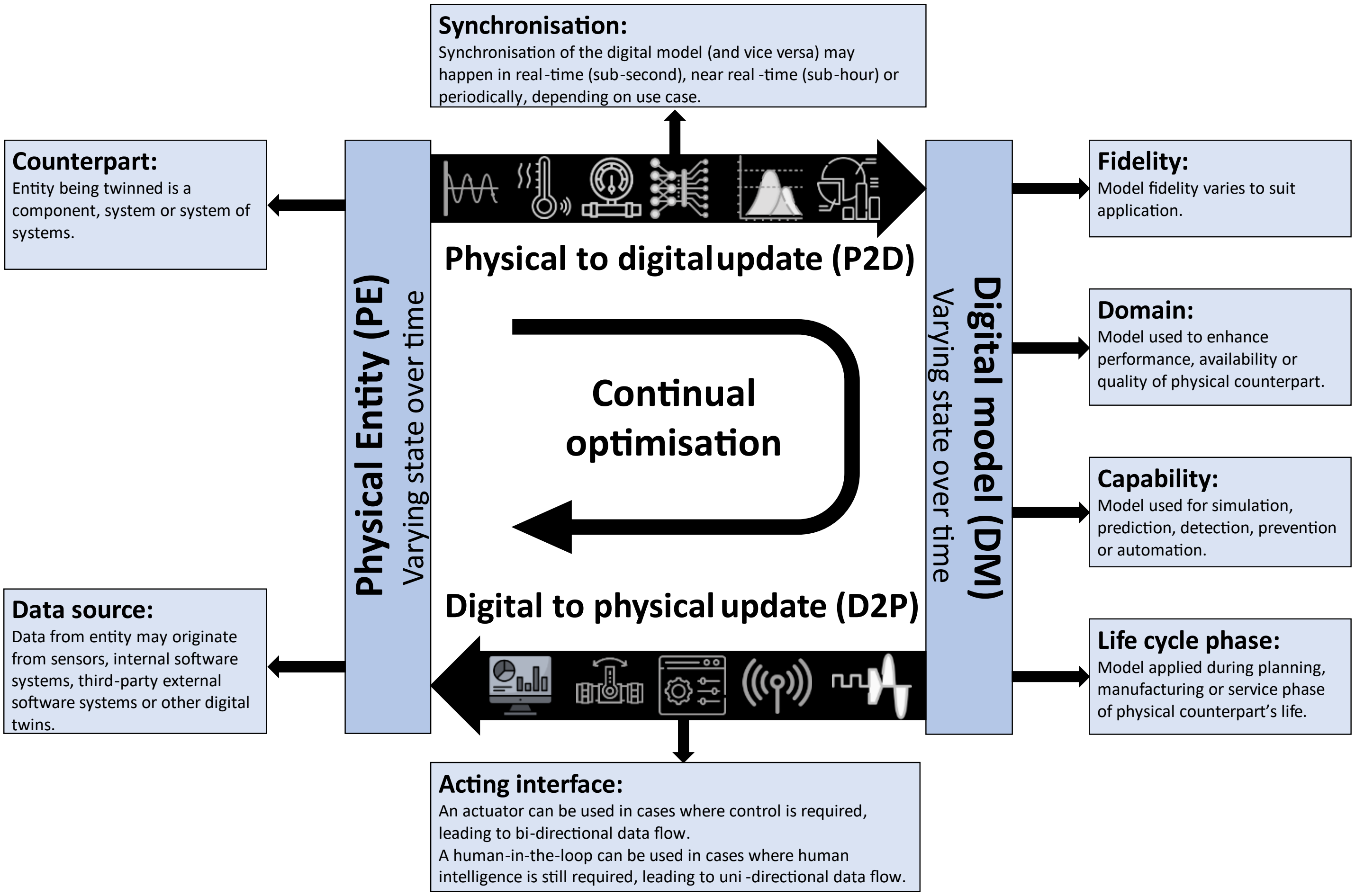


Figure 7 The physical entity refers to the physical counterpart (a component, machine, or system) that is being twinned. The physical entity generates data which is used by the P2D updating engine to modify the state of the digital model. The digital model captures the physical entity's state, optimising some aspect of the physical entity and generating valuable insights that can be acted upon. After updating the digital model's state, an updated prediction of the physical entity's optimal state is obtained through the D2P prediction engine. These predictions can be acted on though actuators or by a human-in-the-loop (HITL). Optimisation (OPT) affects all dimensions of the digital twin, being used to optimise data acquisition of the physical entity, structure of the digital model, state estimation and decision making. OPT can further be split into offline and online elements.

3.1 PHYSICAL ENTITY (PE)

Schweiger and Barth (2023) highlight that "a digital twin can represent anything in the real world of interest to an application as long as this real-world counterpart can be defined as an item with a recognizably distinct existence." They refer to the digital model's counterpart as the "real entity". Although this term is all-encompassing, it may be too broad for a South African mining context, where the likely candidates for twinning are physical assets. Therefore, in the proposed framework, we explicitly refer to the digital model's counterpart as the physical entity (PE). The PE refers to the component, machine or system of machines that is being twinned. The PE has a time-varying state (Thelen et al., 2022b), and the purpose of the DT is to estimate this state, optimise some variable(s) in relation to this state, and enforce control to alter this state.

A DT does not necessarily need to capture a single entity but may capture a set of entities, termed a system (Juarez et al., 2021), or at a larger scale, a system of systems (Yao et al., 2023). These concepts are illustrated in Figure 8. Wu et al. (2023) indicate that twin systems and networks tend to emerge as applications of DTs become increasingly complex. Consider for example, twinning multiple components of a mine vehicle. These twins themselves can feed into a larger twin of the vehicle as a whole (a system), creating a systems twin. On a grand scale, a twin may be constructed which captures the interaction between the various mine vehicles, creating a DT of a system of systems.

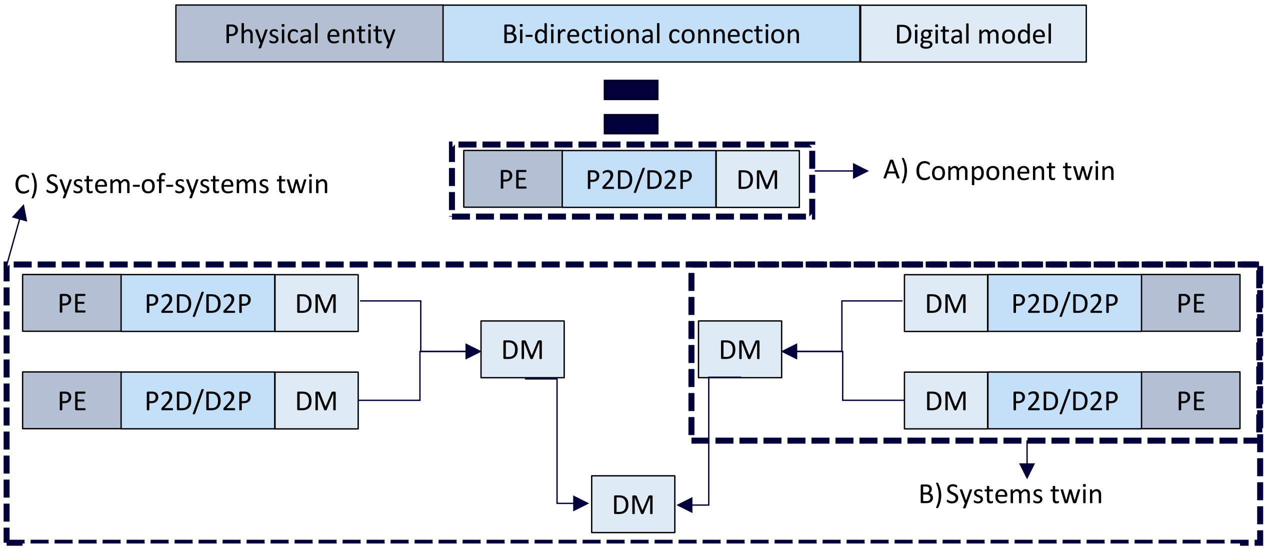


Figure 8 Illustration of the scale of digital twins. A) A twin of a single physical entity, shown with a simplified version below. B) A twin of multiple physical entities - a twin of a system. C) A twin of multiple systems - a system of systems twin.

The PE is associated with two properties, namely that of a counterpart and data source (see Figure 7). The counterpart property defines the type of physical entity being twinning, namely a component, system or system-of systems. The data source property highlights that the entity of interest generates data that originates from various sources. Most commonly, sensors are used (Thelen et al., 2022b), however, other sources of data may also be utilised to inform digital twin decision-making. These sources include internal software systems (for example a resource planner), external third-party systems (for example weather data) or outputs from other related DTs, as seen in Figure 8.

Considerations regarding the collection, transmission, storage, processing and security of data are important to any digital twin application. Tao et al. (2018) place a large emphasis on the data aspect of a digital twin, placing it at the centre of their digital twin framework. However, within the context of the research activity in mechanized mining systems (RAMMS), while these considerations remain crucial, they fall beyond the primary scope. Instead, such data-related considerations align more closely with the research activities of the real-time information management systems (RTIMS) program. For further details on RTIMS, please refer to their website (https://mandelaminingprecinct.org.za/real-time-information-management-systems-rtims/). Consequently, only select aspects concerning data will be briefly touched upon below.

3.1.1 SENSORS AND IOT DEVICES

Sensors and IoT devices play a fundamental role in a DT, capturing real-time data from physical assets (Juarez et al., 2021). The quality of a DT is linked to the quality of sensor inputs (Yao et al., 2023). Furthermore, DTs are limited by the level of technology available at any given time. These devices collect a wide array of information, including temperature, pressure, vibration, and more, providing the foundational data necessary for DTs to accurately represent their physical counterparts. The integration of IoT technology ensures continuous data flow, essential for the dynamic updating of DTs (Khalid et al., 2023).

A concept closely related to the sensors and IoT devices, is that of data fusion. Data fusion involves three processes, namely data preprocessing, data mining and data optimization. The massive number of sensors are expected to create large amounts of data and therefore this three part process is necessary to distill the data into manageable sizes for DTs (Tao et al., 2019).

3.1.2 CLOUD COMPUTING AND EDGE COMPUTING

Computing devices serve as the hosts for DTs, allowing large amounts of data to be processed and valuable insights to be drawn (Yao et al., 2023). Cloud computing offers scalable, on-demand computing resources for storing and processing the vast amounts of data generated by sensors and IoT devices. It enables powerful analytics and machine learning processes that form the brain of DTs, supporting complex simulations and predictive modeling (Macchi et al., 2018). Edge computing complements cloud computing by processing data closer to its source, reducing latency and bandwidth requirements. This is particularly important for time-sensitive applications, where real-time decisions can prevent failures and optimize performance.

The distinction between cloud and edge computing may not be a fixed decision beforehand (Wu et al., 2023). Recent works have started incorporating network latency and computational speed into optimization algorithms to determine whether it would be better to:

- Run the digital twin locally, on the edge.
- Transfer data to a cloud, compute on a server and send the results back.

These works highlight that this conclusion may change as network resources become available or as the physical entity moves closer to a server, reducing network latency. This concept is known as twin migration. Therefore, the decision between whether to compute on the edge or cloud is dynamic.

3.1.3 COMMUNICATION NETWORKS

Robust communication networks are essential for transmitting data between sensors/IoT devices and the DT platform. These networks must be reliable and secure, as they carry sensitive operational data. The implementation of 5G technology in industrial settings promises to revolutionize this aspect, offering high-speed, low-latency connections that can support the real-time data needs of DTs (Yassin et al., 2023).

3.1.4 CYBERSECURITY MEASURES

As DTs involve the transfer and processing of critical operational data, cybersecurity measures are paramount (Yao2023). Ensuring the integrity and confidentiality of data prevents malicious attacks that could lead to operational disruptions or safety hazards. Security technologies, including encryption, secure access controls, and anomaly detection systems, protect DTs from cyber threats (Somers et al., 2023).

3.1.5 INTEGRATION PLATFORMS

Integration platforms facilitate the seamless combination of different technologies and systems that make up a DT. They ensure that data flows smoothly between sensors, analytics platforms, simulation tools, and user interfaces. Open standards and interoperable solutions are crucial for this, allowing equipment and software from different vendors to work together effectively. Yao et al. (2023) highlight poor interoperability due to varying encoding formats of sensors being a cause for concern when constructing a DT. They further highlight that many current shortcomings of DTs are due to lacking interdisciplinary cooperation between various engineering fields. Integration platforms are therefore a good first step towards integrating various engineering fields in a common space.

3.1.6 CONCLUDING REMARKS

The proposed framework assigns two attributes to the physical entity, namely a counterpart type and data source. The counterpart comprises a component, system or system-of-systems, with the data from these entities originating from sensors, internal systems, externals systems or even other digital twins. The following section describes the digital model, which is responsible for the optimization of the physical entity, in digital space.

3.2 DIGITAL MODEL (DM)

The digital model (DM) is a virtual replica of the physical counterpart and is the core component of any DT (Wu et al., 2023). The digital model contains a set of internal states and parameters, which are used to estimate the true state of the physical system (Thelen et al., 2022b). These states and parameters are optimised to achieve a certain goal in the real domain.

In the proposed framework, we emphasize four key properties associated with the DM: model fidelity, domain of application, model capability, and life cycle. Model fidelity refers to the level of complexity employed in modelling the physical entity, recognizing that computational constraints may necessitate variations in fidelity levels (Thelen et al., 2022b). Thus, the proposed framework acknowledges the flexibility of DM design to accommodate diverse fidelity requirements dictated by specific applications and computational constraints.

Moreover, the DM is tailored for a specific domain of application, aligning with the overarching goal of optimizing operational effectiveness (OEE) as delineated by Schweiger and Barth (2023). Accordingly, a digital twin is geared towards enhancing facets of performance, availability, or quality – the elements that drive OEE.

The DM may be designed with distinct capabilities to achieve the OEE optimisation objective. Drawing inspiration from Schweiger and Barth (2023), we delineate five core DM capabilities: simulation, prediction, detection, prevention and automation.

The application of DTs transcends various phases of the physical entity's life cycle. Schweiger and Barth (2023) and (Stark & Damerau, 2019) define these life-stages as beginning of life, middle of life and end of life. Thelen et al. (2022b's) more clearly distinguish DM utility across planning, manufacturing/development, and service phases. Therefore, the proposed framework uses the latter definition.

Expanding beyond mere properties, the proposed framework delineates distinct DM types utilized to achieve the overarching objectives of the DT. These encompass geometric models, physics-driven models, and data-driven models, with hybrid models—often referred to as physics-informed machine learning (Karniadakis et al., 2021) – serving as an amalgamation of data-driven and physics-driven approaches. These model types are inspired by the taxonomy established by Thelen et al. (2022b). By systematically categorizing these DM variants in the following sub-section, we provide a comprehensive understanding of the diverse modelling methodologies available when developing a DT.

3.2.1 GEOMETRIC MODELS

Geometric models refer to virtual representations of reality in some geometric form. These models include solid models, laser scans, virtual reality (VR), augmented reality (AR) and mixed reality (MR) models (Thelen et al., 2022b).

Solid models are generally constructed through Computer Aided Design (CAD) software. Solid models are useful for spatial planning and may further be used as inputs to simulations (see Figure 9A). Laser scans generate surface-level features of a part or environment and are similarly used for spatial planning (see Figure 9B) but are also useful as input to many safety systems. VR, AR and MR models can be used to design models within 3D space, but can also be used to augment 3D space, displaying a virtual world over the physical world (see Figure 9C).

3.2.2 PHYSICS-BASED MODELS

Physics-based models fundamentally try to model the laws governing the real world. Properties from a physical entity as well as experimental calibrations are used as input to some mathematical model which aims to approximate the relationship between these physical properties and desired outputs. Partial differential equations (PDEs) are often used to describe these laws, and therefore many of the physics-based models currently in use fundamentally rely on PDEs (Thelen et al., 2022b).

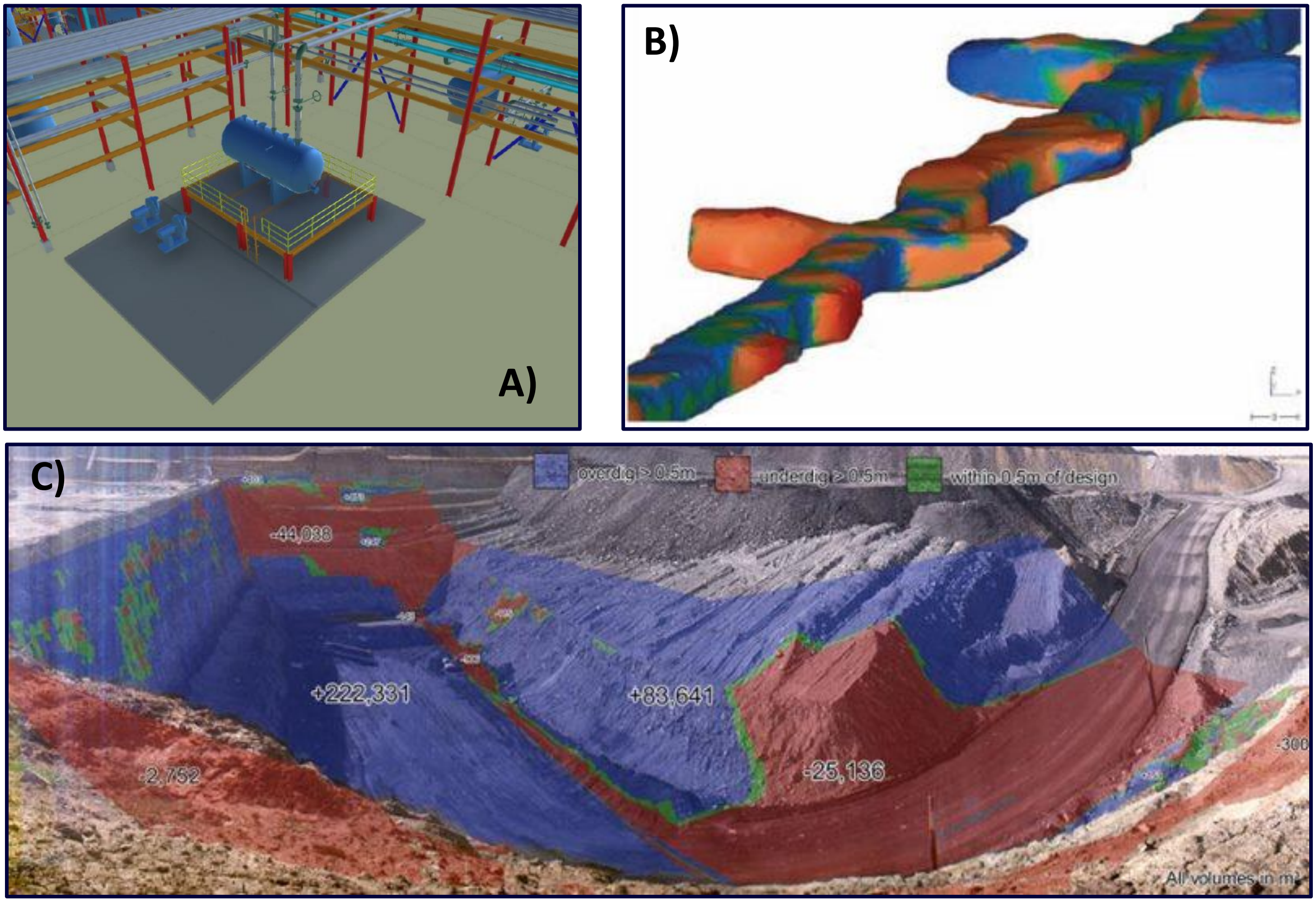


Figure 9 Illustration of various geometric models. A) CAD model of a planned vessel being evaluated in the virtual environment (Image courtesy of UniversalPlantViewer by CAXperts GmbH). B) Laser scan of an underground mine (https://shorturl.at/bgxG9). C) Augmented reality display of open-pit mine (https://shorturl.at/gjKW5).

Although many variants of physics-based models exist, they can broadly be grouped into solid body models, thermal and fluid flow models, kinematic and dynamic models and multiphysics models. Solid body models are typically used for stress and strain analyses of structures, and commonly use finite element analysis (FEA) for this purpose (see Figure 10A). Thermal and fluid-flow models are used in multiple engineering applications and often analyse thermal and fluid effects (see Figure 10B) through computational fluid dynamics (CFD). Kinematic and dynamic models are used to analyse physical interactions between objects (see Figure 10C), usually using multi-body dynamics (MBD). Multiphysics simulations are useful in simultaneously simulating coupled physical phenomena (see Figure 10D) by sharing information between individual physics-based simulations (Thelen et al., 2022b).

3.2.3 DATA-DRIVEN MODELLING

Data-driven models are designed to capture the underlying physical relationship between a set of inputs and outputs. In the context of DTs, these models are useful at efficiently predicting digital states for real-time control and optimisation (Thelen et al., 2022a). These models are often used instead of physics-based approaches for two reasons. Firstly, if the physics is too complicated or not well understood, data-driven models can still be used to capture the input-output relationships of the PE. Secondly, although the physics may be understood, the scale or complexity of a physics-based model may not be practically applicable in a real-time DT environment. In both these instances, data-driven models are often used, assuming relevant input-output data pairs are available.

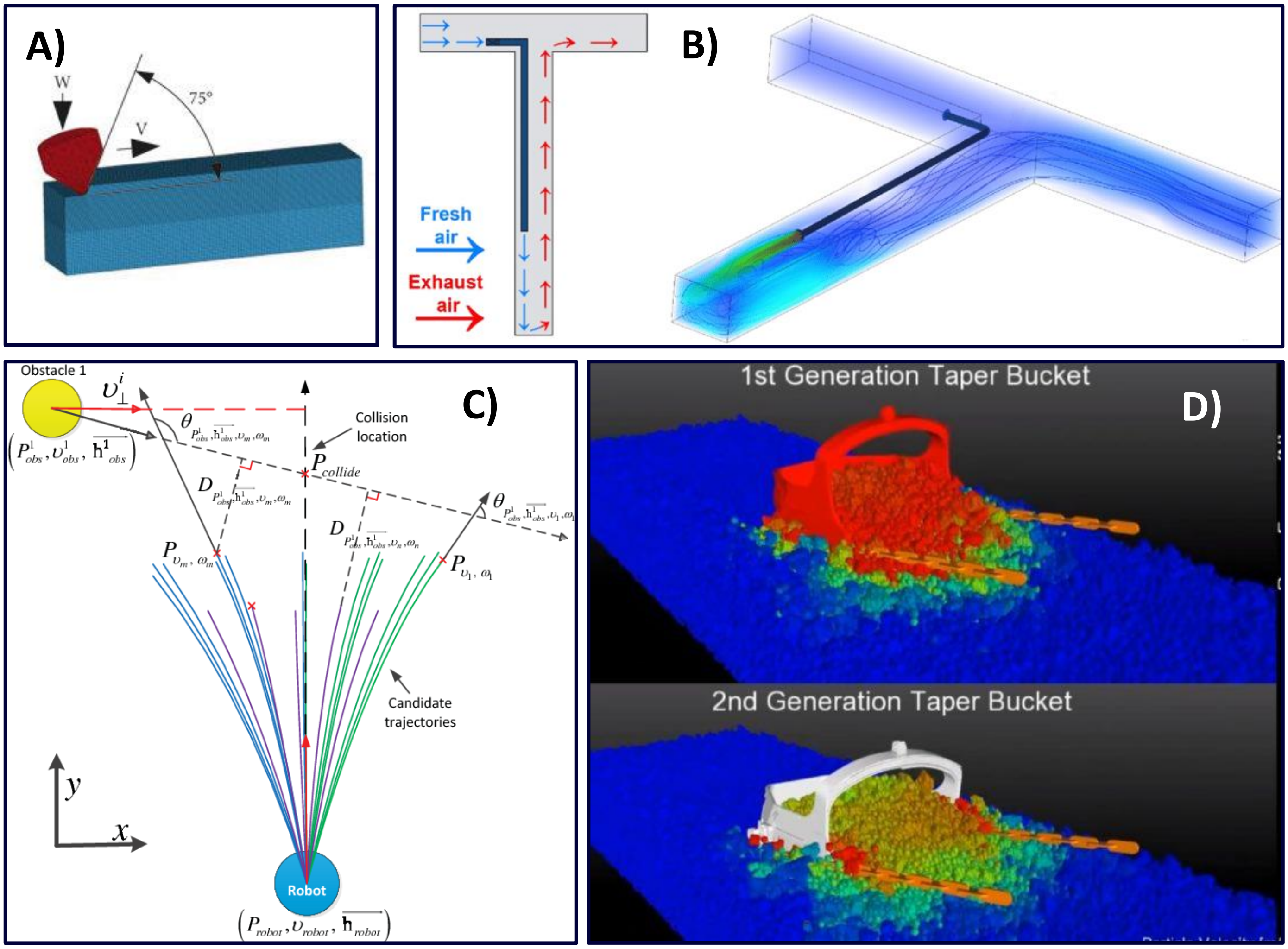


Figure 10 Illustration of various physics-based models. A) A rock cutting tool simulation using FEM (https://shorturl.at/elwC2). B) A CFD simulation of ventilation in an underground mine (https://shorturl.at/jsDR0). C) Multi-body dynamics simulation of collision avoidance between two moving vehicles (Guan et al., 2018). D) A multi-physics simulation to optimize a taper bucket design using FEM, DEM and multi-body dynamics (https://rb.gy/ap0c6p).

Data-driven models are applied in three broad domains as indicated by Thelen et al. (2022b), namely degradation modelling, surrogate modelling and dynamic system identification. Degradation models aim to model the deterioration of a component or system over time, often using statistical methods to model degradation behaviour and predict future performance. Surrogate models aim to approximate complex systems with good accuracy and are usually employed when physics-based modelling would be too computationally expensive. Dynamic system identification refers to methods which aim to extract the underlying physics from a system by looking at data. All three domains utilise some form of machine learning (ML) or statistical modelling. Figure 11 showcases various statistical and ML techniques, and how they are typically applied to various data-driven problems.

Statistical models typically extract features that are designed from expert human knowledge. These models find relationships within data by using “pre-programmed” algorithms. Machine learning models aim to automatically learn important features from data and without the need for expert knowledge (Lei et al., 2020). Therefore, machine learning models mainly differ from statistical counterparts, in that they automatically extract important features from data. These features are often highly abstract, non-linear and more effective at finding expressive features from data. These methods are becoming more relevant, with authors such as Yao et al. (2023) emphasizing the need to develop efficient big-data processing techniques for modern DTs.

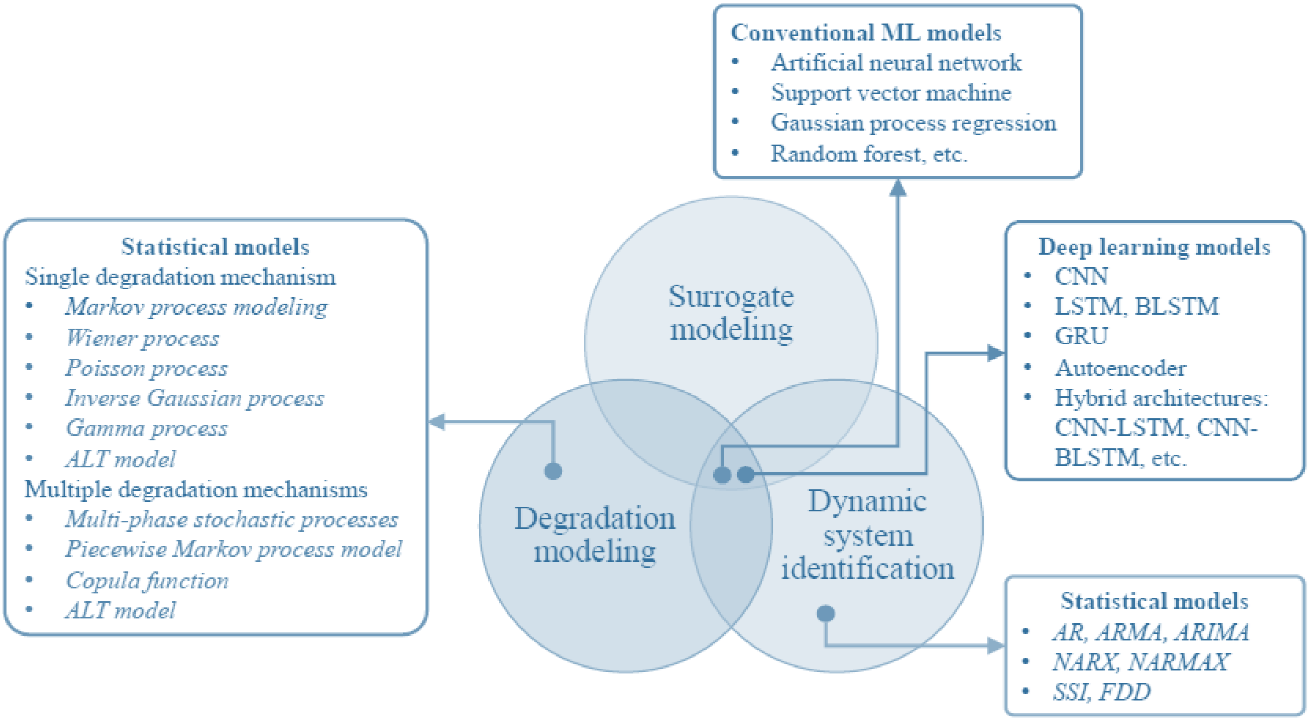


Figure 11 Illustration of the application of data-driven models to surrogate modelling, degradation modelling and dynamic system identification. Statistical, machine learning or deep learning techniques are often used for these purposes. Image courtesy of Thelen et al. (2022b).

A current shortcoming of many ML methods is that due to the complex, non-linear nature of features, they are often uninterpretable and not physically meaningful. Therefore, machine learning models are often referred to as “black box” models.

Deep learning models are a sub-set of machine learning models with model architectures allowing for much richer features to be extracted. This is achieved through adding more learning layers, deepening the architecture (Jia et al., 2016) – hence the term deep learning. These models have an added advantage that they not only extract rich features, but usually link these features to a certain output label. They therefore act as end-to-end models, automatically extracting features from data and making consequent conclusions from those features (Lei et al., 2020). These models have seen great success in domains where large bodies of data are available, however, they also suffer from the interpretability challenge and are therefore also labelled as black boxes.

3.2.4 HYBRID MODELLING

In certain scenarios where both physics-based knowledge and empirical data from the physical entity are available, hybrid models, also known as "physics-informed machine learning" according to Thelen et al. (2022b) and Karniadakis et al. (2021), emerge as a compelling solution. Figure 13 offers an intuitive depiction of the advantage a hybrid approach offers.

The PE may exhibit known and unknown physics. The known physics are modelled using physics-based models discussed in section 3.2.2. If supplemental data are available from the machine, we would expect to capture both known and unknown physics within this data. The aim, therefore, is to encapsulate as much of the unknown physics as possible from the data. By integrating the data-driven model with the physics-based model, a more robust hybrid model is obtained. This synergy is further illustrated in Figure 12.

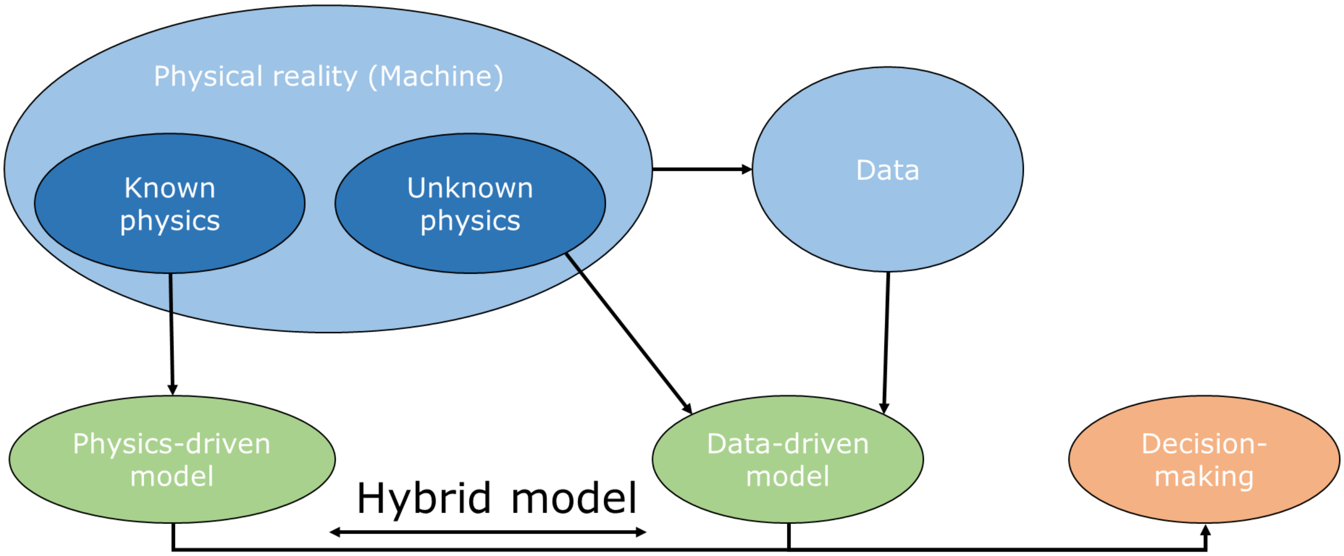


Figure 13 The physical reality contains both known and unknown physics. Physics-driven models capture the known physics from the physical reality. However, due to unknown unknowns or certain constraints, some of the physics will not be captured in the physics-driven model (unknown physics). When measuring data from the physical reality, the assumption is that both known and unknown physics are measured, implying that a data-driven model may be used to extract the unknown portion to compliment the known portion from the physics-driven model. This results in a hybrid approach.

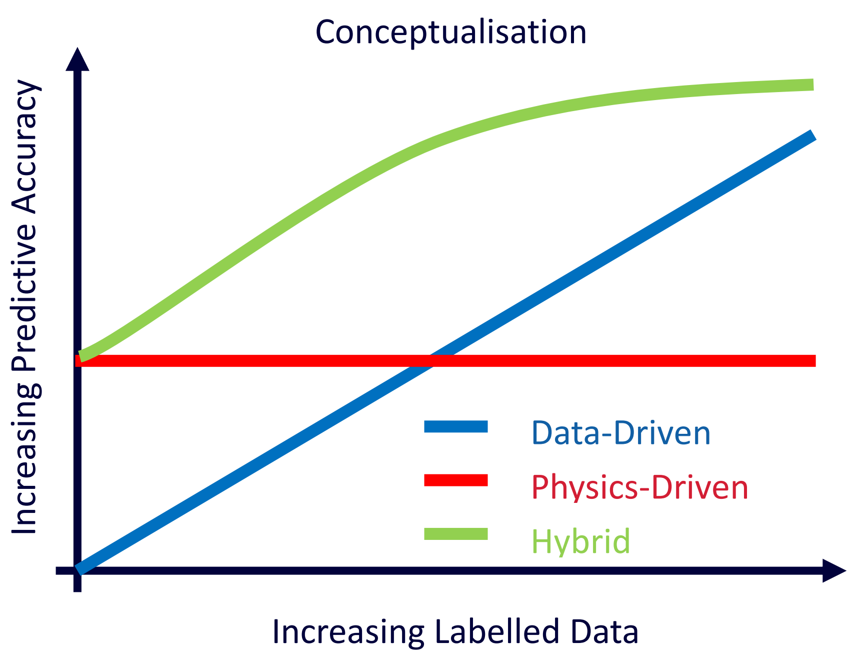


Figure 12 With a physics-based model, we expect to derive some set baseline knowledge about the PE. Regardless of how much labelled data we collect, we are capped at a set predictive accuracy. A data-driven model, however, becomes better at the predictive task as more training data is introduced. This requires many datapoints and when the available data is minimal, these methods are expected to perform poorly. By using a hybrid approach, we aim to start with some baseline accuracy that will improve as more data becomes available, ultimately obtaining a model that outperforms either a physics-driven or data-driven approach (Kundu et al., 2020).

The versatility of hybrid models extends beyond an amalgamation of approaches. Rooted in physical principles, they offer enhanced generalization to unseen data, expedited training times owing to constrained solution space dictated by physics, and improved interpretability stemming from their inherent grounding in physics (Thelen et al., 2022b).

Thelen et al. (2022b) delineate six distinct hybrid modelling methods, each offering a unique approach to leveraging both physics-based knowledge and empirical data. These methods, illustrated in Figure 14 include:

1. Physics-informed machine learning
2. Data augmentation
3. Transfer learning
4. Delta learning for missing physics
5. Delta learning for residual prediction
6. Machine learning for parameter calibration

The six hybrid approaches outlined above are not an exhaustive list, and other configurations are possible. However, they do present typical approaches and are therefore useful in this framework's context. Thelen et al. (2022b) do not discuss the applicability of these approaches, therefore, we have broadly summarised their applicability according to the maturity of physical understanding and availability of data, as illustrated in Figure 15.

Physics-informed (or constrained) machine learning is fairly similar to conventional data-driven methods, but with an additional loss term that is enriched by physical constraints, constraining model solutions to comply with physics. Physics-informed methods generally predict some variables of interest (displacement, temperature, pressure), whilst simultaneously ensuring all relationships between these variables adhere to physical constraints. These methods are capable of effectively capturing differential equations, which form the basis for many known physical laws. In an in-depth review, Karniadakis et al. (2021) highlight physics-informed machine learning as a technique to not only simulate multi-physics problems where noisy, high dimensional data is present but also as a method to effectively solve general inverse problems. These methods are applicable in cases where the relevant differential quantities of the problem are understood, but the detailed relationship between these quantities are not. Therefore, these models are labelled as having a low maturity of physical understanding and requiring large amounts of data.

Data augmentation is a technique of using simulated, first-principles data together with empirical data to form an augmented, richer dataset used to train a data-driven model. The philosophy of this approach is to augment the measured data with simulated data of scenarios that are not present in the measured data, therefore enhancing the predictive capability of a machine learning model. This approach assumes that the simulated data is representative of reality, hence why augmentation is possible. Therefore, these techniques are likely applicable in cases where good agreement between simulated and real data exist, requiring a higher level of maturity in terms of physical understanding, and hence less empirical data.

Transfer learning uses the same principle of generating first-principles data alongside measured data, with the exception that simulated and measured data are not combined into one large dataset. Instead, a data-driven model is initially trained on simulated data. After training, the machine learning model is fine-tuned on empirical data. The logic behind this approach is that simulated data may contain useful physical knowledge, but may not be fully representative of the physical entity. Therefore, one trains a machine learning model to extract the useful physical knowledge from the simulated data, and simply calibrate the model to work on empirical data. In this approach, the assumption is made that the physical knowledge is incomplete, hence the requirement for calibration. This approach is therefore more applicable when partially understood/modelled physics and some measured data is present.

Delta learning for missing physics utilises machine learning to capture unmodelled physics. When creating a physics-driven model, it is not guaranteed that all possible physics is understood and therefore might not be modelled. In delta learning, the difference (delta) between the modelled physics and measured data is fed to a machine learning model. The machine learning model aims to recover any unmodelled physics from this difference, allowing a user to update their predictions by combining the known physics and learned physics into a single model. This approach inherently assumes missing physics and therefore a less mature physics-based understanding is required for application. Therefore, a relatively large amount of data is required to use this technique.

Delta learning for residual prediction involves a machine learning model that is trained to make predictions on data generated from a physics-based model. These predictions are likely to be different from real measured data. A second machine learning model is then trained to learn the discrepancy (delta) between the first machine learning model's predictions and the measured data. This results in corrected output predictions. The goal of this approach is to train the first machine learning model on general, well-understood physics and simply calibrate the output of this model to match reality. Although this approach assumes a difference between modelled and measured physics, it still assumes that the majority of underlying physics was captured, and therefore this approach is applicable for models with mature physical understanding, requiring less data than the other approaches.

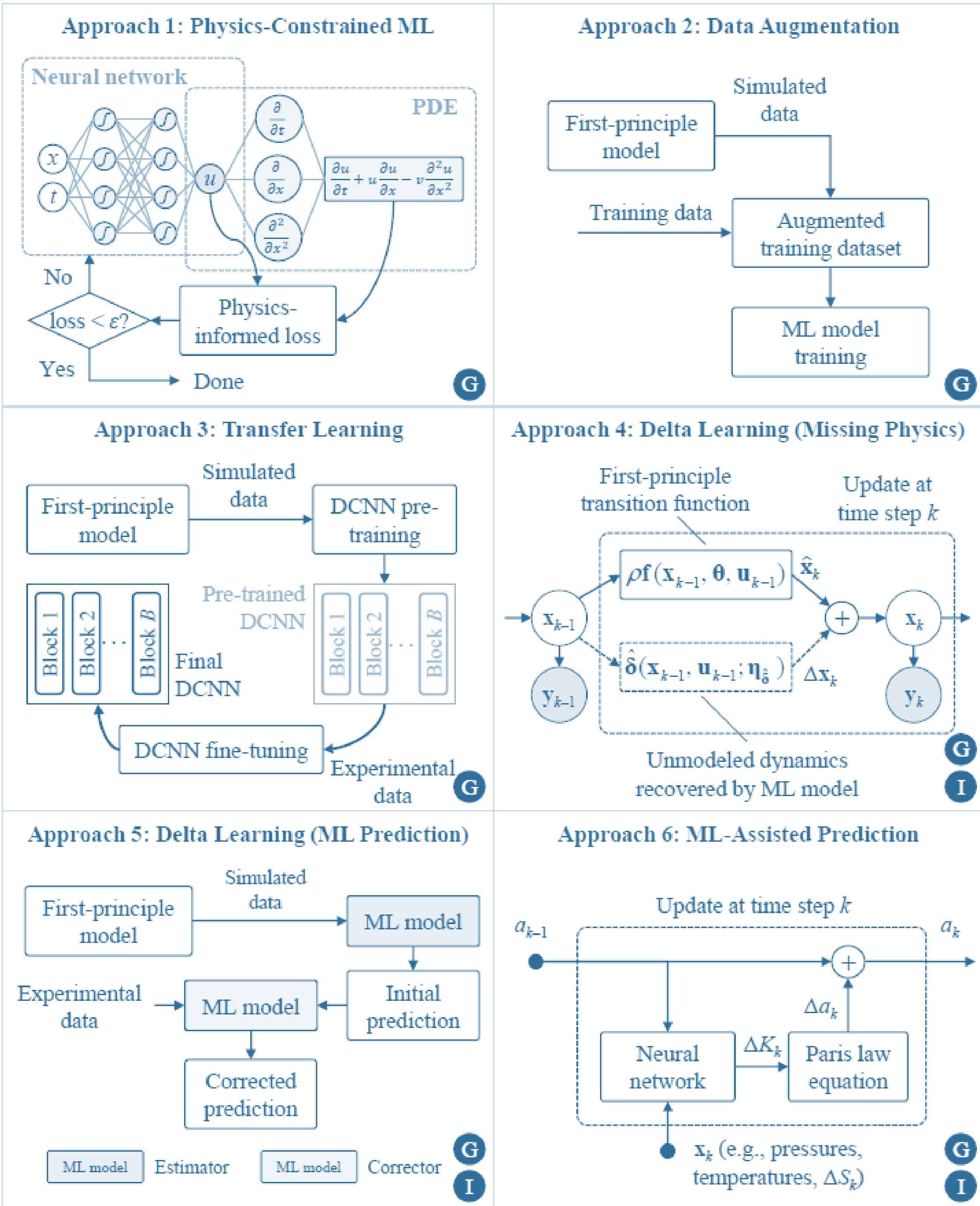


Figure 14 Technique generalizability and interpretability is indicated by G and I on the figure. **Approach 1)** Physics-informed machine learning integrates physical constraints into conventional data-driven models, enhancing predictive accuracy while ensuring compliance with underlying physics laws. **Approach 2)** Data augmentation enriches real data with simulated, first-principles data to form a more informative dataset, thereby enhancing the predictive capability of machine learning models. **Approach 3)** Transfer learning trains a machine learning model on simulated data to extract useful physical knowledge, then fine-tunes the model on real data to calibrate its performance. **Approach 4)** Delta learning for missing physics utilizes machine learning to capture unmodeled physics by analyzing the difference (delta) between modeled and measured physics, updating predictions accordingly. **Approach 5)** Delta learning for residual prediction trains a machine learning model to learn the discrepancy (delta) between predictions from a physics-based model and real measured data, correcting output predictions accordingly. **Approach 6)** Machine learning for parameter calibration directly optimizes the parameters of a physics-based model using machine learning algorithms to generate outputs more closely aligned with real measurements.

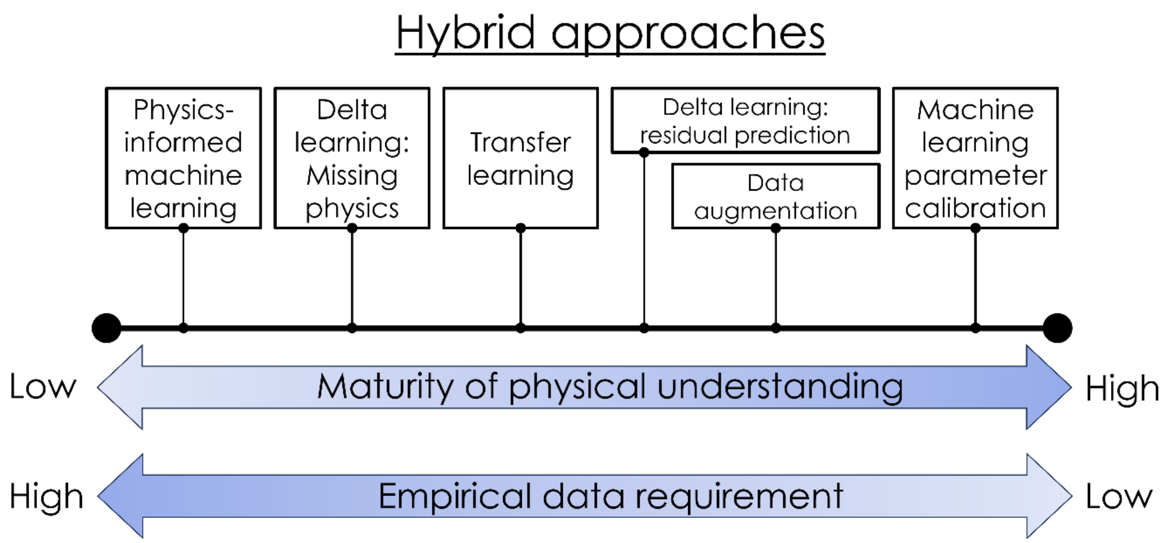


Figure 15 The applicability of different hybrid approaches based on the required physical understanding and consequent data requirement for each. The left side of the spectrum involves models which do not have strong physical understanding, and therefore likely require more data. On the opposite side, the physical knowledge is high, and less data is required to train the model. This illustration is merely a simplification and may not hold true in every scenario.

The final approach uses machine learning to directly optimise the physics-based model's parameters. In some instances, the exact parameters used in the physics-based model may be initial estimates, and may cause discrepancies between the physics model output and real measurements. In this case, a machine learning algorithm is tasked with learning the correct parameters that will generate realistic outputs more like those that were measured. This technique may suffer from the identifiability problem, where multiple combinations of input parameters may lead to the same target output. This approach assumes the most mature physics-based model, as the only use for machine learning is to calibrate model parameters, with no assumption about missing physics being present. Therefore, minimal data should suffice to calibrate the model.

While these methods offer diverse approaches to hybrid modelling, they all prioritize generalisability, addressing limitations inherent in data- and physics-driven models. When solely employing data-driven or physics-based models, overfitting or discrepancies may occur, respectively. However, by combining both methods, these hybrid approaches yield more robust and interpretable solutions, as highlighted by Thelen et al. (2022b).

3.2.5 CONCLUDING REMARKS

In this section, we showcased four properties that typically accompany digital models. These properties describe model fidelity, application domain, model capability and ultimately which part of the physical entity's lifecycle the model is being applied to. Alongside these properties, four types of DMs were outlined, namely geometric, physics-driven, data-driven and hybrid models. Each one of these models have further sub-divisions which were outlined and will differ based on the digital twinning application.

Currently, we have discussed the physical entity, and its digital model counterpart. The digital model, as indicated in the introduction of this section, is responsible for estimating the state of the physical entity. Moving forward, the focus shifts to the vital connection bridging the physical entity with its digital twin counterpart, enabling state estimation across various applications.

3.3 PHYSICAL TO DIGITAL (P2D) UPDATING

Ensuring that the digital model accurately reflects the state of the physical entity is paramount for effective decision-making in a digital twin. This necessitates a robust physical to digital (P2D) updating process, where data from the PE is integrated into the DM. The primary objective of P2D updating is to adjust the state variables within the DM, which are pivotal for estimating the PE's state and subsequently optimizing parameters to achieve the DT's objectives.

A common requirement for a DT, is that it is synchronised, usually in real-time. However, Schweiger and Barth (2023) highlight that synchronisation period is dependent on the use case, and in some cases near real-time or periodic synchronization is sufficient. The synchronisation period may not be an arbitrary choice, as in some instances, human-in-the-loop DTs are possible (Schweiger & Barth, 2023). This implies that in such scenarios, synchronisation on the millisecond or second level is infeasible. Consequently, the proposed framework (Figure 7) acknowledges that synchronisation frequency is an important property of P2D updating and varies based on the specific application, ranging from sub-second updates to hourly, shift-end, or even weekly intervals. Various external factors such as unexpected downtime of data acquisition systems or humans in the loop of a DT application will cause the synchronisation frequency of the DT to vary. Essentially, the synchronisation frequency of a P2D connection is limited by the slowest element in the DT optimisation cycle. The synchronisation property types are grouped accordingly:

- Real-time, on the order of milliseconds to seconds, where a control system is likely required.
- Near real-time, on the order of seconds, minutes, or hours, where control systems or humans are involved.
- Periodic, on orders longer than hours, where humans are likely involved.

It is important to consider that in certain applications, the same DT may operate at multiple synchronisation periods based on external factors (Schweiger & Barth, 2023) such as varying computational limitations (edge vs server processing), communication limitations (cloud connectivity) or battery level. Therefore, a digital twin need not have a fixed synchronization period and can dynamically change according to the use case.

This section delineates common P2D updating techniques inspired Thelen et al. (2022b). While not exhaustive, these methods offer a comprehensive overview of prevalent practices. Thelen et al. (2022b) identify five main P2D techniques, namely physical measurements, probabilistic model updating, machine-learning model updating, fault diagnostics and prognostics and ontology-based reasoning. In the proposed framework, fault diagnostics and prognostics are considered as special cases of the first three approaches. Ontology-based reasoning is further excluded from the proposed framework, as it is currently deemed out of scope. Consequently, the subsequent sections delve into the intricacies of physical measurements, probabilistic model updating, and machine learning model updating, shedding light on their implementation within the digital twin framework.

3.3.1 PHYSICAL MEASUREMENTS

Direct physical measurements serve as a cornerstone for updating the digital state within a digital twin. Leveraging an array of sensors and communication technologies collectively known as the Industrial Internet of Things (IIoT), these measurements provide real-time data essential for maintaining an accurate representation of the physical entity.

There are two primary purposes for utilizing physical measurements in digital twin updating. Firstly, these measurements facilitate the tracking and comparison of the digital model's movement or position with that of the physical entity, enabling anomaly detection and performance monitoring. For instance, MineRP's tracking system (Figure 16A) visualizes the location of assets within an underground mine, allowing operators to monitor drill rigs, load haul tracks, and other equipment across different levels.

Secondly, physical measurements play a crucial role in physics-based analysis or design, where they are used to directly update or calibrate parameters within physics-based models. This calibration is exemplified by tools like Deswik's reconciliation tool (Figure 16B), which compares planned mining geometries to actual mined geometries, updating both geometric digital models and associated simulations accordingly.

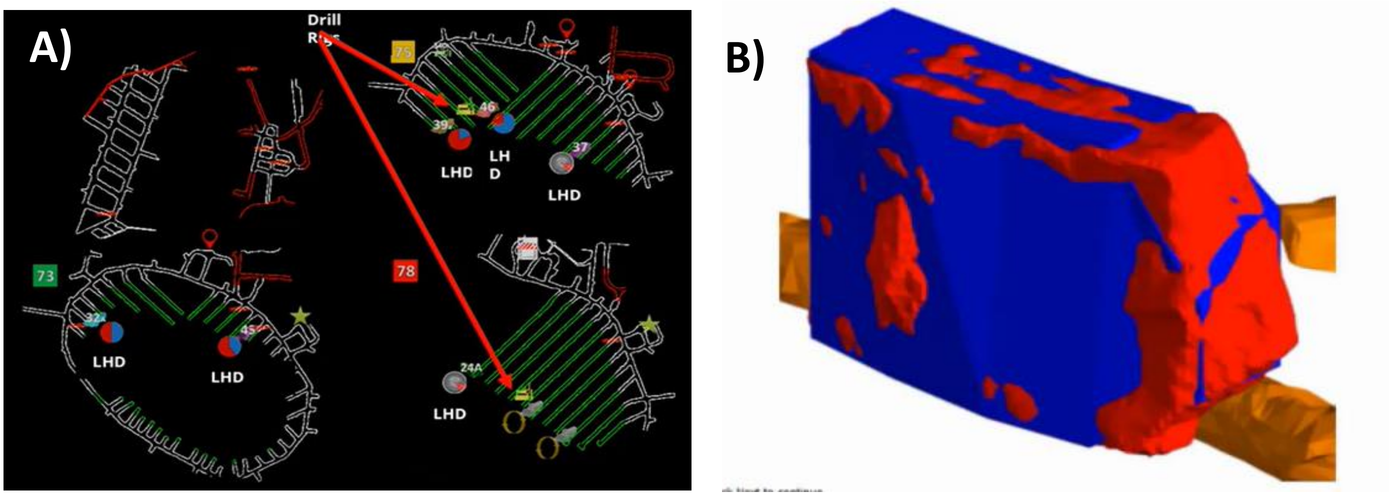


Figure 16 Two examples of using physical sensor updates to inform a digital model. A) MineRP's vehicle tracking dashboard, showcasing vehicle type, location and progress within various levels of an underground mine. B) Deswik's mine reconciliation tool, which shows actual (red) vs. planned (blue) mining results. This tool allows operators to measure planned and mined, unplanned and mined and planned and unmined.

It's important to recognize that the effectiveness of these techniques depends not only on the sophistication of the digital models but also on the quality and reliability of the sensors and communication networks comprising the IIoT. Challenges such as data constraints and communication limitations, particularly prevalent in environments like underground mining, underscore the broader scope of digital twins beyond their constituent digital models, encompassing the entire IIoT infrastructure.

3.3.2 PROBABILISTIC MODEL UPDATING

In some instances, direct measurements of a DM state are not possible. For example, a "health state" or "state of charge" cannot directly be measured, but rather indirectly inferred through noisy measurements. In these instances, the digital state is expected to vary as a function of time and measurements are required to calibrate the digital state. These methods will likely be used in conjunction with statistical DMs, as opposed to machine- or deep-learning approaches, which have different updating methodologies.

State estimation results are necessary to inform decision-making algorithms in the DM. Depending on the timeline of decision making, different state estimation models may be used. Different applications may require offline or online state estimation. For example, predicted material properties are not expected to vary rapidly, and may be estimated offline. However, trajectory planning for a mining vehicle will require near real-time calibration. In instances where online state estimation is required, computational power and speed need to be considered to models with acceptable response times.

Model updating can be achieved by a variety of algorithms, all of which aim to propagate the digital state over time, based on measurements. Bayesian filtering methods, including Kalman filters, extended Kalman filters, unscented Kalman filters, and particle filters, are commonly employed for model updating. For systems with nonlinear dynamics and arbitrary distributions, dynamic Bayesian networks provide a suitable framework, while population-based optimization algorithms like particle swarm and genetic algorithms are preferred for high-dimensional and nonlinear systems. (Thelen et al., 2022b). An example of practical work which use various Kalman filters (extended and unscented), for state estimation is given in Figure 17. In this work, a mine vehicle's path is planned, whilst simultaneously avoiding obstacles. This work showcases a practical example where a statistical DM is used alongside probabilistic model updating for path planning.

The choice of model updating algorithm depends on factors such as the level of nonlinearity captured, computational resources, and the timescale involved in state estimation tasks. The digital state within a DM comprises hidden states and parameters. In instances where the hidden state does not change rapidly over time (for example a degradation state), the task is called a parameter estimation task. However, if hidden states change rapidly (for example fuel flow rate), the task is called a state estimation task. Various algorithms, such as Kalman filters, gradient descent, particle swarm, Markov chain Monte Carlo, Kennedy and O'Hagan (KOH) framework, etc. have been developed to address these state estimation tasks and the specific use case determines which algorithm is best suited.

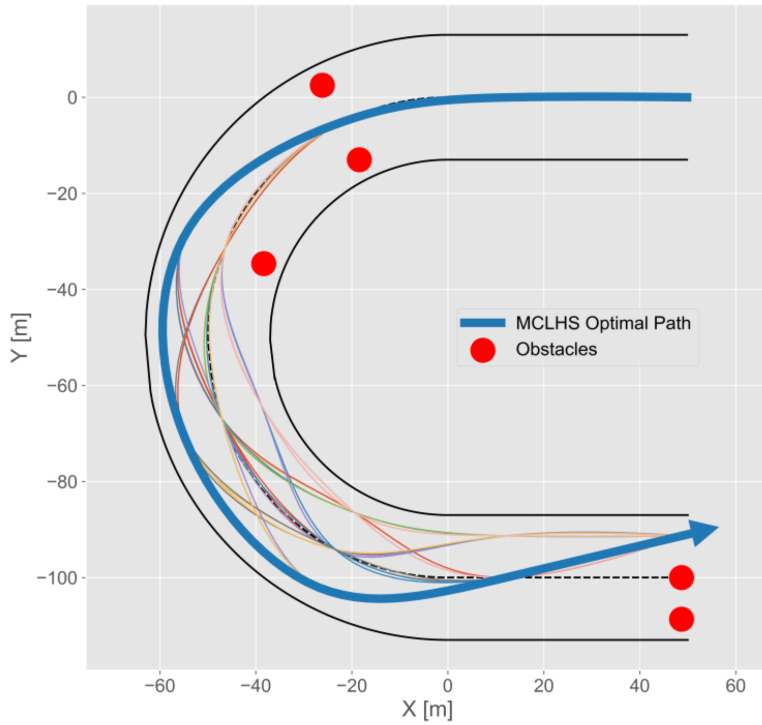


Figure 17 Illustration of various planned paths and identification of the most optimal path for the task (Declercq et al., 2021).

3.3.3 MACHINE LEARNING MODEL UPDATING

Machine learning models are employed in environments where the input data may vary significantly due to periodic changes in the physical world. A variety of factors contribute to this variance, such as changes to processes, removal or addition of sensors, changes in equipment's operational settings etc. If certain operational conditions are not captured in the training data, the machine learning model will encounter unseen data, called out-of-distribution data.

Machine learning models are generally trained to fit a wide range of operating conditions present in the training data. Therefore, the resulting machine-learning based DM does not necessarily have to be updated with every sensor update, because the training process produces a model which can adapt to a range of known varying inputs. However, if out-of-distribution data is presented, these models usually showcase degraded performance. The concept of degraded performance is known as model drift or model decay. Two main factors contribute to model drift, namely data drift and concept drift. Both forms of drift are illustrated in Figure 18.

Data drift occurs when the statistical properties of the training data substantially differ from those of the testing data. Figure 18 (left) illustrates that when training data is not representative of the testing data, the learnt model performs poorly. Concept drift occurs when the underlying relationship between the training data and predicted output changes. For example, a change in operational state of a machine would inherently change the relationship between variables, changing the relationship between input and output data. Figure 18 (right) illustrates this type of drift, where the learnt relationship between x and y changes, causing the ML model to under-predict the true value of y.

Safeguarding machine learning models against these sources of drift is paramount to a successful DT. A common approach is to measure model performance variation using a drift indicator (Thelen et al., 2022b). If the indicated drift is large enough, the ML model is updated. Drift indicators are application specific, but some general approaches are used. Data drift may directly be measured by comparing the distribution between the training and test data through statistical metrics such as Kullback-Leibler or Jensen-Shannon divergence. Another approach involves trending model predictive quality over time to indicate when model quality has reduced by a sufficient amount to require updating. Alternatively, model predicted output and real output can be compared and when deviations become large, updating may be required. Methods such as sanity checks and energy balances are also used to identify data-drift (Collins & Balshaw, 2024).

3.3.4 CONCLUDING REMARKS

This section has elucidated common methods for updating the digital state using measured data from the physical entity.

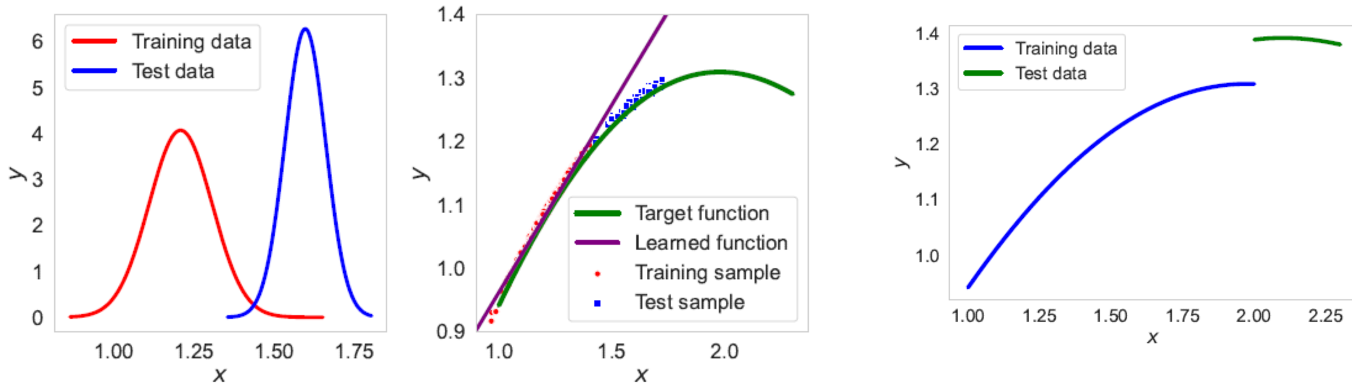


Figure 18 Illustration of data-drift (left) and concept drift (right) (Thelen et al., 2022b). Data-drift occurs when the training and testing data differ. Concept drift occurs when the underlying relationship between input and output changes, causing the model to make incorrect predictions.

The frequency and methodology of updates vary based on the digital twin's specific requirements, ranging from milliseconds to weeks. The choice of updating model depends on the nature of the digital model, with geometric or physics-driven models often relying on direct sensor measurements, while certain data-driven models, such as statistics-based models, will require advanced updating methodologies outlined in the probabilistic model updating section. If the model of choice is a machine- or deep-learning model, machine learning model updates are rather required. Ultimately, the objective of all updating models is to maintain the digital model's accuracy at a sufficient update interval to achieve the digital twin's optimization goals.

3.4 DIGITAL TO PHYSICAL (D2P) UPDATING

The essence of a DT lies in its ability to translate digital insights into actionable changes in the physical world, a process referred to as Digital to Physical (D2P) updating. The PE's inferred state is captured by the DM, which in turn determines an optimal future PE state. Therefore, D2P is responsible for turning the digital state back into a physical state. For this reason, the proposed framework assigns the property of "acting interface" to the D2P connection (see Figure 7).

The acting interface defines how insights from the digital model are implemented in the physical entity. Traditionally, automatic actuators within control systems facilitated this process. However, recent literature argues that even human intervention qualifies a system as a digital twin, as long as the optimization objective is achieved (Schweiger & Barth, 2023). This distinction gives rise to two categories: bi-directionally connected DTs, where automated systems complete the synchronization loop, and uni-directionally connected DTs, where human intervention is involved in implementing insights. This property is tightly linked with the P2D dimension's property, namely synchronisation. Where humans are used as the acting mechanism, synchronisation is expected to be slow, happening periodically. If a control system is employed, sub-second action is likely possible, implying real-time synchronisation.

Although it is possible to highlight various acting interfaces for the D2P updating mechanism, it is more cumbersome to find a meaningful classification of various D2P updating types, as in previous sections. D2P updating may be achieved through a control system, lights on a machine, a dashboard on a computer screen, augmented reality and a plethora of other possible acting interfaces. Therefore, the proposed framework does not create different classifications of D2P updating methods. Rather, inspired by the work of Thelen et al. (2022b) we highlight five typical D2P methods based on their application, namely:

- System reconfiguration
- Production planning

- Path planning
- Process control
- Maintenance scheduling

System reconfiguration is a broad term encompassing a reconfiguration in scheduling, planning or layout of a system to optimise an output (Mo et al., 2023). Related to this concept, is production planning which entails concepts such as availability prediction and production disturbance detection (Macchi et al., 2023) to optimise the production output of a system. Reconfiguration and production planning activities can be optimised offline, before production starts, or online during production, ultimately reacting to external factors impacting production. The underlying goal is to reduce variability in production outputs by more readily reacting to varying factors that impact production inputs (Groeneveld et al., 2019). This method can likely be applied at various synchronisation periods depending on the type of reconfiguration method being used. To demonstrate the benefit of a reconfiguration D2P model, consider the following scenario. A mine may draw data from existing assets (P2D), feed the information into a DM, and determine asset location, speed, capacity, fuel level etc. Based on these parameters, a DM can determine which asset is best suited for the task and schedule the asset (D2P) to attend to the task. In effect, the mine has been reconfigured to increase production output.

Work from Groeneveld et al. (2019) showcase another example related to practical production planning in an open pit mine. A production model is developed to simulate various production scenarios, taking mining and milling capacity, capital and operating costs, cut-off grade and mine production schedules into account (see Figure 19A). By incorporating uncertainties around these processes into the model, Groeneveld et al. (2019) were able to predict scenarios which would lead to greater profit in the long run, allowing the mine to adjust the production plan as new information is obtained.

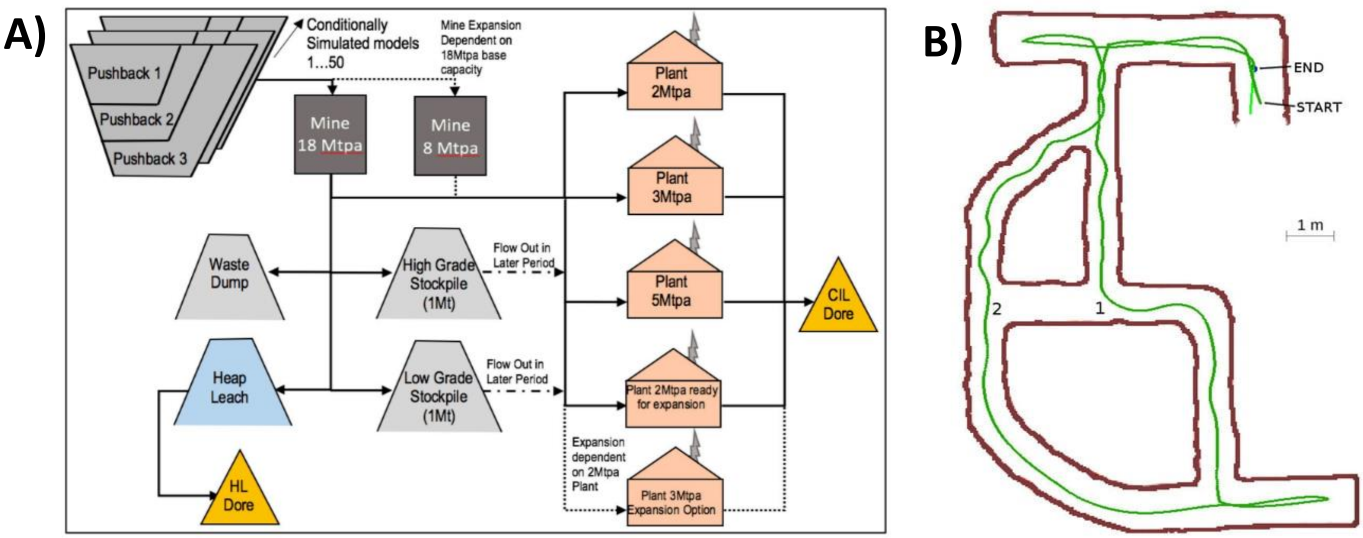
Path planning describes a field of study that aims to identify the most efficient, collision-free route from a starting point to a destination. Path planning is often applied to mobile, autonomous robots (or vehicles). Path planning algorithms cater for static (Wenna et al., 2022) or dynamic (Rafai et al., 2022) situations. For example, a stationary piece of equipment with a movable arm may need to plan an arm trajectory as to not damage the surroundings. Similarly, a dynamic piece of equipment may need to move around without hitting obstacles or other pieces of dynamic equipment (see Figure 19B). Path planning algorithms try to optimise these movement paths subject to a certain loss function, such as smoothest path, lowest energy expenditure, lowest risk and lowest speed. These methods will likely prove most useful in aiding the transition of South African mines to comply with collision avoidance regulations.

Maintenance activities can greatly benefit from DT technologies, enabling the use of predictive methods. Corrective maintenance optimises a component or machine’s life, but at the cost of unexpected failures, increased downtimes and lost production. Preventive maintenance tries to capture an average life of a component or machine, replacing said component or machine before failure, eliminating a certain amount of unexpected downside. The drawback of this method, however, is the potential underutilisation of components, which may still be perfectly healthy at the time of replacement. Predictive maintenance delivers the best-case scenario, where a component or machine is run, whilst simultaneously inferring its health. When health is sufficiently degraded, the component or machine may be replaced, knowing the maximum life was extracted without incurring unexpected downtimes. It is therefore clear that a predictive maintenance strategy is preferred.

Through DT technologies, predictive maintenance becomes more feasible, as more real-time data is available. The goal of a predictive maintenance strategy is to identify potential anomalous signals from sensor data, predicting when a machine or component may fail and consequently schedule maintenance of the machine or component during a known maintenance window. Ultimately, maintenance schedules may be optimised (D2P) based on the latest health state prediction of a component or machine.

Predictive maintenance often requires a large amount of fault data to develop sufficiently mature predictive models. Such failure data is not always readily available, and as such, predictive maintenance technologies often benefit from the use of hybrid models, where simulated data is used to train models to identify and quantify real-life damage. In a practical example, Patil et al. (2021) showcased a model which estimates engine, transmission and brake health for underground trucks and loaders (see Figure 19C). Furthermore, the model predicts whether failure is likely to occur in the next “t” days, allowing maintenance engineers to schedule maintenance for the applicable part. Finally, the model outputs an explanation of anomalous events to help the maintenance manager to more easily act on the insights. These outputs are clear examples of D2P connections, which generate actionable insight to maintenance managers.

The D2P technologies discussed up to this point, are actionable, but may require some form of human intervention to complete. Process control, however, does not require a human-in-the-loop and can directly use actuators to complete the DT cycle. For this reason, control is one of the most widely used D2P methods (Thelen et al., 2022b). Model predictive control (MPC) refer to DMs that aim to predict future behaviour of processes and determine optimal control measures within a set of constraints. MPC models have seen widespread adoption across multiple industries, being largely enabled through the rise in IIoT devices. Recent MPC models utilise machine learning to improve prediction accuracy, control performance and speed. A very tangible example of MPC is found in EMESRT level 9 collision avoidance machinery. These machines, after sensing a potential collision, apply brakes (D2P) to avoid said collision. Figure 19D illustrates a graphic from Newtrax, where proximity detection results are used to determine when to apply the brakes (red zone).



C) Critical Health Indicator

Equipment Name	Component Name	Health Score	Status in Next t Days	EXPLANATION
Equipment 1	engine	**46%**		Few instances of high Engine Intake Manifold Pressure, few instances of low Engine Cooler Pump Pressure and increasing instances of low Engine Fuel Rate
Equipment 1	transmission	**68%**		
Equipment 1	brake	**73%**		

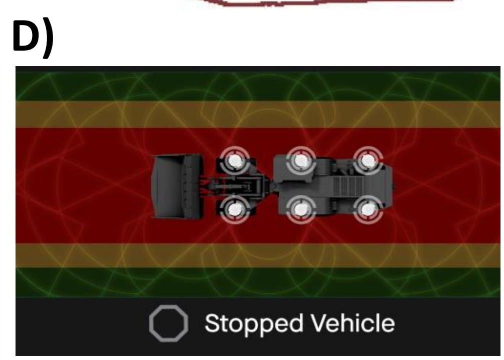


Figure 19 A) Production planning optimisation example from Groeneveld et al. (2019) B) Path planning algorithm from Rubio-Sierra et al. (2020). C) Predictive maintenance system from Patil et al. (2021). D) Collision avoidance system visualisation from Newtrax (https://newtrax.com/cas).

D2P updating plays a pivotal role in realizing the full potential of digital twins, translating digital insights into tangible actions in the physical world. By considering the actioning mechanism, synchronization period, and specific D2P mechanisms, organizations can implement effective D2P strategies tailored to their optimization goals and operational requirements.

3.5 OPTIMISATION (OPT) AND UNCERTAINTY QUANTIFICATION (UQ)

Optimisation and uncertainty quantification (UQ) are integral components of DT frameworks, influencing the P2D, DM and D2P dimensions (Thelen et al., 2022a). Optimisation enables the DM and connecting technologies (P2D, D2P) to stay calibrated, where UQ enables informed decision-making under uncertainty. Both elements are important in ensuring seamless synthesis between P2D, DM and D2P, ultimately enabling the DT to provide optimal value in uncertain environments. The discussion surrounding UQ and optimisation is broad and therefore the roles of each are discussed in their own sub-sections to follow.

3.5.1 UNCERTAINTY QUANTIFICATION

Probabilistic models, as discussed in section 3.3.2, play a crucial role in quantifying uncertainties associated with the real world. These uncertainties primarily fall into two categories:

1. Aleatory Uncertainty: Arising from natural variability, aleatory uncertainty includes factors like sensor measurement errors, material properties variations, and system load fluctuations. These uncertainties are inherent and cannot be eliminated.

2. Epistemic Uncertainty: Stemming from limited data, lack of knowledge, or model simplifications, epistemic uncertainty can be reduced with more data or improved understanding of the system. This type of uncertainty can be further divided into model parameter uncertainty and model form uncertainty.

Machine learning models are used as an alternative mechanism for P2D updating, and are also exposed to sources of uncertainty in their training data (model input). This is compounded by the fact that machine learning models are often described as "black-box" models (model output). We would therefore like to quantify our confidence in these models for certain predictions. Most approaches focus on modelling epistemic uncertainty, as aleatory uncertainty is learnt directly from the training data. Two tasks are generally relevant to the UQ of machine learning models.

The first task aims to quantify the uncertainty around model predictions from the training and testing data. The goal is to obtain a model that not only predicts a value, but the accompanying uncertainty of that prediction. The second task aims to identify samples which are unlike samples that the machine learning model was trained on. This is often called out-of-distribution detection. Data-driven models generally perform poorly on data that is dissimilar to training data and therefore it is important to identify these out-of-distribution data samples. If a data sample is deemed to be out-of-distribution, it should not be passed to the machine learning model, as the accompanying prediction will likely have a prediction error that is too large to be of any use.

Thelen et al. (2022a) outline a few approaches for both machine learning UQ tasks, namely Gaussian process regression, Bayesian neural networks, ensemble neural networks, Monte Carlo dropout and deterministic approaches. They summarise the benefits and drawbacks of each method in Thelen et al. (2022a).

Another set of models that often require UQ, are dynamic system models. The main uncertainty here, arises due to model parameter uncertainty and model form uncertainty. Model parameter uncertainty arises due to both aleatory and epistemic uncertainty. For example, we may not have knowledge about a stiffness coefficient of a component, implying epistemic uncertainty. However, stiffness may also vary from component to component, implying an underlying aleatory uncertainty as well. Model form uncertainty arises from imperfect modelling due to simplification or lack of understanding of the full physics.

Generally, UQ of dynamic system models includes two main tasks, namely UQ for measurement and UQ for state transition equations. The former focuses on capturing the uncertainty from model inputs (P2D), whereas the latter focuses on capturing uncertainty when updating the digital state (DM) and applying control based on the state (D2P). Both approaches utilise some form of parameter calibration, bias correction or make use of the KOH framework.

Table 1 A comparison of four different approaches for uncertainty quantification of machine-learning models (Thelen et al., 2022a). Gaussian process regression approaches have high computational efficiency, are accurate at quantifying uncertainty and have a strong ability to detect out of distribution examples. However, they have poor scalability to higher dimensions. Deterministic approaches have similar advantages, with slightly worse uncertainty quantification accuracy, but much better scalability to high dimensions. Neural network ensembles and Monte Carlo dropout have similar advantages in scaling to higher dimensions, but are poor at detecting out of distribution samples and are not as computationally efficient as Gaussian or deterministic approaches. Neural network ensembles have good accuracy for quantifying uncertainty, with Monte Carlo dropout performing worse on this task. Ultimately, the appropriate method is dependent on the problem at hand, level of complexity the user is willing to undertake in constructing these models, and property of interest (accuracy, scalability, speed etc.).

Quantity of interest	Gaussian process regression	Bayesian neural networks	Ensembles of neural networks	Monte Carlo dropout	Deterministic approaches
Accuracy in UQ	High	High-medium[a]	High	Medium	High-medium
Computational efficiency (test time)	High[b]	Low[c]	Medium-low	Medium-low	High
Ability to detect OOD samples	Strong	Weak (may estimate low uncertainty)	Weak (may estimate low uncertainty)	Weak (may estimate low uncertainty)	Strong
Scalability to high dimensions	Low	High	High	High	High

3.5.2 OPTIMSATION

Optimisation affects different facets of a DT across P2D, DM, and D2P dimensions, and can be applied either offline or online. Offline optimisation involves the initial calibration and optimisation of DT elements, prior to the DT's application. Offline optimisation is applied to multiple facets, such as sensors and experiments (P2D), model parameters (DM) and model validation (D2P).

Sensor optimisation is an important consideration in minimising aleatory uncertainty with sensor type, placement and number of sensors all playing a role (Thelen et al., 2022a). Related to sensor optimisation, is the concept of design of experiments. Experiments need to be conducted to gather initial data for DT calibration. Experimentation may be expensive and time-consuming and therefore one would like to extract the optimal amount of information from the experiments as possible. Therefore, fields of study exist to optimise the experimental input settings to extract the most informative data for model calibration and validation (discussed shortly).

DM parameters are calibrated to bridge the gap between the DM and PM. Bayesian approaches such as maximum a posteriori estimates, Markov chain Monte Carlo, Gibbs sampling, the KOH framework or even Bayesian filters may be used to perform the initial calibration. However, optimisation-based approaches may also suffice when some metric is optimised. The most used metrics are least squares and maximum likelihood (Thelen et al., 2022a). Finally, after obtaining the optimal parameters for the given data, various statistical metrics may be used to measure the agreement between the obtained DM and the real-world data. These metrics act to validate or invalidate the DM.

Online optimisation occurs when the DT has been deployed and is operational. In this setting, real-time data is used to calibrate and optimise the DT, taking UQ, as mentioned in the previous sub-section, into consideration. It is important to reiterate that real-time differs depending on the application. For certain DTs, millisecond response times are required, however, for others, such as maintenance-based DTs, days or weeks are sufficient. Ultimately, the application of the DT determines what is "real-time". The timescale of interest also determines the type of optimisation which may be applicable. For short time scales, feedback-based optimisations are possible. A practical example is mission planning, where the planned movement or behaviour of a vehicle is continually updated. For long-term timescales, open loop optimisation may be required. A practical example involves predictive maintenance scheduling, where the maintenance action may only occur after days or weeks.

Thelen et al. (2022a) highlight some of the most common optimisation algorithms used to solve the various optimisation domains outlined above. The benefits and drawbacks of various methods are evaluated in Table 2 Summary of optimization methods in digital twins.

Table 2 Summary of optimization methods in digital twins (Thelen et al., 2022a). Evolutionary optimisation methods are slower, but tend to converge to global optima. On the opposite side of the spectrum, greedy algorithms converge on local minima, but at a much quicker rate. Reinforcement learning-based methods are useful when the optimisation problem is difficult to define, but are complex to construct. Bayesian optimisation methods offer globally optimum solutions, however, they suffer from the curse of dimensionality, meaning that the more variables one optimises for, the slower the algorithm becomes. Finally, algorithms like A*, RRT* and the Dijkstra algorithm have been developed that are fast and find near-optimal results, but are mostly applicable to mission planning applications.

Method		**Evolutionary optimization methods**	**Greedy algorithms**	**Reinforcement learning-based methods**	**Bayesian optimization**	**A*, RRT, RRT*, Dijkstra algorithm**
Online/offline		Online/Offline	Offline	Online/Offline	Online/Offline	Online
Application	Sensor placement optimization	✓	✓	✓	✓	✗
	Physical system modeling	✓	✗	✗	✓	✗
	Process control	✓	✓	✓	✓	✗
	Mission planning	✓	✗	✓	✗	✓
	Maintenance	✓	✓	✓	✗	✗
Pros		High-dimensional optimization; Global optimization;	Relatively low number of function evaluations	Global optimization with long-term rewards	Global optimization	Real-time global optimization
Cons		Large number of function evaluations	Local optimal	Data-intensive and high complexity	Curse of dimensionality	Suboptimal solution

3.5.3 CONCLUSION

In conclusion, the integration of optimisation and uncertainty quantification within the framework of digital twins is indispensable for enhancing decision-making capabilities and ensuring robust performance in dynamic environments. By effectively managing aleatory and epistemic uncertainty, digital twins can provide more reliable insights and predictions.

Simultaneously, optimisation methods enable the calibration and refinement of digital twin elements across different dimensions, from sensor placement to model parameter tuning, facilitating adaptive decision-making processes. Whether applied offline for initial calibration or online for continuous adaptation, optimisation algorithms play a crucial role in maximizing the utility and responsiveness of digital twins to evolving conditions.

Together, these methodologies empower digital twins to navigate complex scenarios, mitigate risks associated with uncertainty, and drive towards optimal outcomes across various domains and applications.

3.6 IDENTIFYING SUB-COMPONENTS OF A DIGITAL TWIN

The proposed digital twin framework has been outlined in five dimensions. This section aims to serve as an aid in designing a digital twin by creating simplified selection processes for the DM, P2D and D2P dimensions outlined in Figure 7. A summary of the models and mechanisms associated with the DM, P2D and D2P dimensions is illustrated in Figure 20. The proposed framework has a further eight accompanying properties, also seen in Figure 7. The selection of these properties may also be simplified by showing the relationship between various properties, aiding in the property selection process for a digital twin.

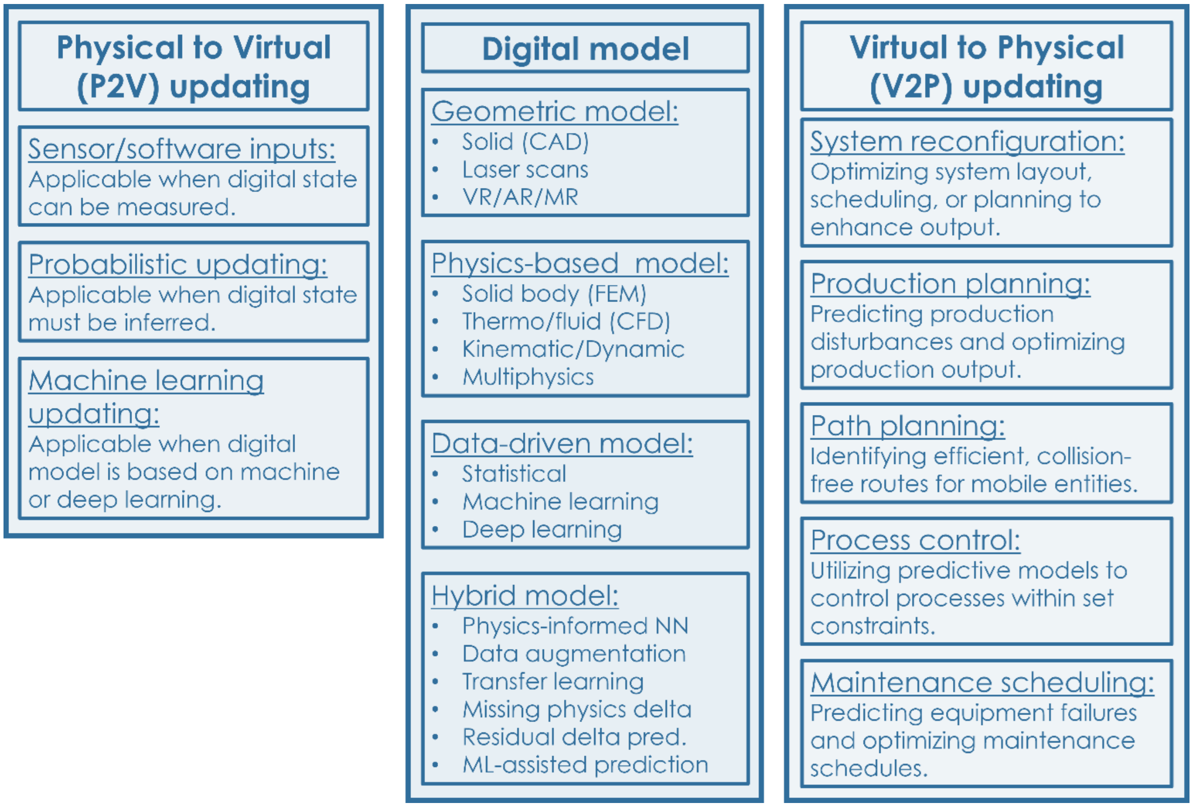


Figure 20 A summary of various model types and mechanisms that are present in a digital twin. The physical to virtual updating process may rely on raw sensor measurements, probabilistic model updating or machine learning model updating techniques to handle input data to the digital model. Digital models can broadly be classified as geometric, physics-based, data-driven or hybrid. These models optimise the physical entity they are modelling via V2P mechanisms, some of which include system reconfiguration, production planning, path planning, process control and maintenance scheduling.

These tools merely serve as guidelines and are simplifications of a complex decision-making process. They are therefore likely useful as a first iteration in DT creation, but ultimately expert opinions are required to create the best possible DT model for a certain application.

3.6.1 CHOOSING A DIGITAL MODEL

Choosing the correct DM for a DT application requires various considerations. These considerations are outlined in Figure 21. If the user requires a visual solution, a geometric model should suffice. However, should an output be required which serves some non-visual function, for example health state prediction, three possible model types are possible, namely data-driven, physics-driven or hybrid models. If both data and physical knowledge is available, a hybrid model is sensible, and the applicable hybrid model can be chosen based on the amount of physical knowledge and available data. If only physics knowledge is available, a physics-based model should be used. Depending on the quantity of interest being optimised, different methods such as multi-physics, CFD, FEM/DEM and multi-body dynamic models are possible. If only data is available, a data-driven method is applicable. The main decision is whether to employ a statistical, machine learning or deep learning model. Statistical models are useful in cases where the expected relationship within the data is simple. These methods tend to be more interpretable, have a lower computational burden but are less accurate. Deep learning models are more applicable in instances where the dataset is assumed to have complex relationships. These models are less interpretable, being called "black box" models, have high computational requirements but are generally more accurate than statistical models. Machine learning models are applicable when a middle-ground between these two extremes is required.

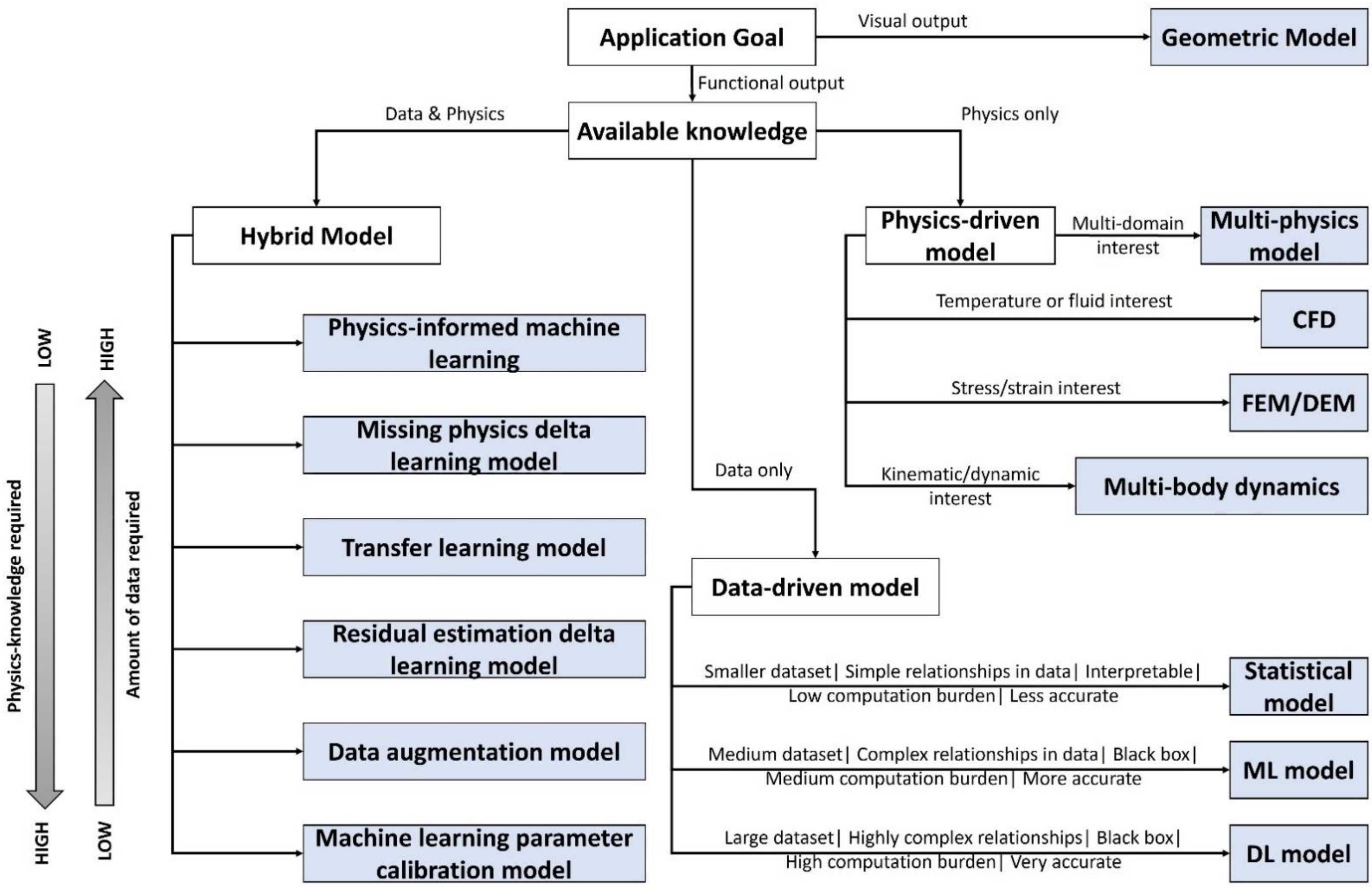


Figure 21 A decision diagram to aid in choosing the correct digital model for a digital twinning application. Four broad domains of models are possible, namely geometric, physics-driven, data-driven or hybrid models. The main distinction between each domain is the purpose (visual or functional), and the amount of physical knowledge and available data for the use case.

3.6.2 CHOOSING A PHYSICAL TO DIGITAL UPDATING MODEL

The selection of a P2D updating model is highly dependent on the DM choice. If the chosen DM states are physically meaningful (temperature, pressure, stress, strain etc.), they can be inferred directly from sensors, and simple sensor measurements should suffice. However, if the states are more abstract (health state, state of charge), statistical or machine learning models will be required. Depending on whether the chosen DM utilises statistical or machine learning (deep learning included) models, probabilistic or machine learning updating models are best suited, respectively. The selection process for the potential P2D is illustrated in Figure 22.

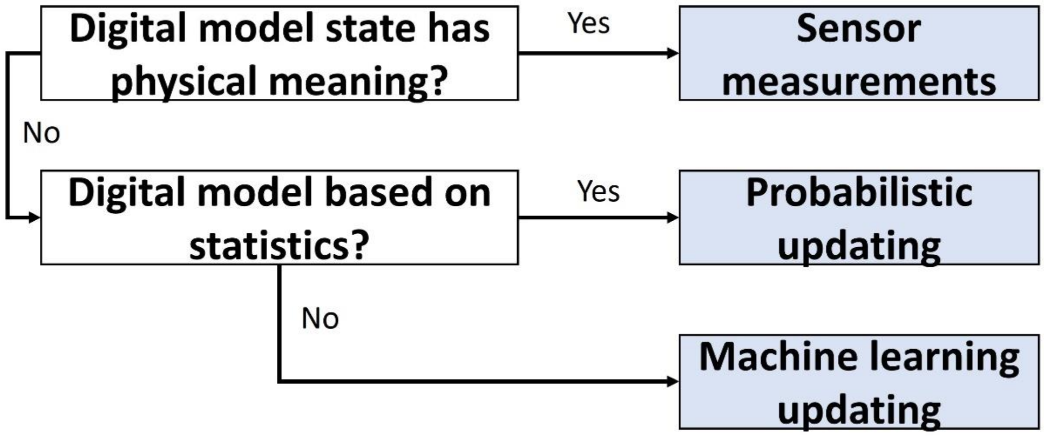


Figure 22 Selection criteria to determine whether sensor measurements, probabilistic updating or machine learning updating is required for the digital twin application. Sensor measurements are likely when the digital model contains physically measurable state variables such as temperature, pressure, strain etc. If the digital model is based on a statistical model, and has abstract state variables, probabilistic updating methods are likely. If the digital model is based on machine learning, machine learning updating methods are more applicable.

3.6.3 CHOOSING A DIGITAL TO PHYSICAL UPDATING MODEL

The choice of D2P updating is less dependent on the DM, but rather on the DT application. Although Figure 20 outlines five examples of D2P applications, this is not a selectable parameter when building a digital twin. The application will automatically determine which type of D2P method is applicable. For example, if traffic management is required, one will likely use path planning D2P methods for this purpose.

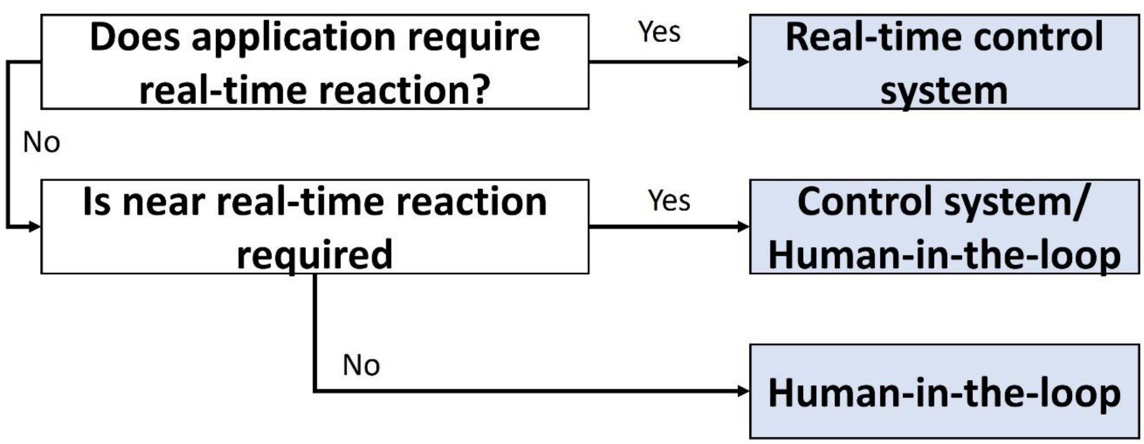


Figure 23 Selection criteria to determine the applicable digital to physical acting interface. For sub-second applications, control systems should be used. For near real-time applications, either control system or a human-in-the-loop is possible. For periodic updating applications a human-in-the-loop is likely the most sensible choice.

It is therefore more sensible to define the D2P selection criteria according to the acting interface, which is a function of the synchronisation period. The selection criteria are therefore outlined in Figure 23. If real-time (sub-second) reaction is required by the digital twin, the applicable model to action the DM insight is likely a control system. If the application requires sub-hour intervention, either control system or human may be used to execute the action proposed by the DM. If larger time-scales are involved, the need for a control system is likely unnecessary, and human-in-the-loop intervention is sufficient.

3.6.4 IDENTIFYING THE RELEVANT PROPERTIES OF A DIGITAL TWIN

The proposed framework associates a set of properties for the P2D, DM and D2P dimensions that define a DT. Here we outline which properties are likely to be used together, and which properties may be mutually exclusive. The inter-relationships between properties are illustrated in Figure 24.

Starting with the physical entity, one must identify which counterpart is being twinned. One may start as small as component level (a single bearing) but may also choose to model a system (a wheel) or perhaps a system-of-systems (a full vehicle). Depending on the chosen counterpart, different sources of data may be relevant. For example, component level twins will likely strongly rely on sensors, but will have little use for internal or external systems. Conversely, system-of-system twins will likely incorporate a large array of data sources.

After defining the properties of the physical entity, the connections between this entity and the digital model need to be specified. The synchronisation period is application specific, with real-time applications likely using direct sensor measurements and perhaps inputs from other digital twins. Internal and external software systems may be too slow for real-time applications. However, for longer synchronisation periods, these software systems may be applicable. Depending on the synchronisation frequency, two actioning interfaces are relevant. For real-time decision-making, actuators are the only sensible option and will therefore complete the digital twin loop, creating a bi-directional automatic data transfer link within the DT. However, for applications where periodic updating is sufficient (like condition-based maintenance), a human will likely be the applicable actioning mechanism to complete the cycle, leading to uni-directional automatic data link in the DT. For near real-time applications either option should be sufficient.

Finally, four digital twin properties need to be specified. In cases where model fidelity is high, one would likely not be able to have a real-time, control system action and therefore this link is not shown. For these applications models of lower but sufficient fidelity are likely required. For slower, human-in-the-loop applications, both high or sufficient fidelity models may suffice. The application domain of the DM is not affected by fidelity considerations and furthermore can be used with any capability of the digital twin. However, not all capabilities are necessarily applicable to the life cycle phase of the physical entity. For the planning phase, simulation models will most likely be used to optimise the twinned object before manufacturing. For the manufacturing phase, we expect any capability to be applicable, except prevention. For the service phase, any capability of the DM is relevant.

3.6.5 CONCLUDING REMARKS

The proposed DT framework comprises various model types and properties. This section showcased model selection and property inter-relationship diagrams which may aid in identifying the relevant models and properties for a DT application. These diagrams serve as guidelines but should by no means be used as the sole source of information to design a DT and experts in the field of the DT application should still be consulted.

The following section outlines various digital twins that have been applied to mining, with two main goals. Firstly, we wish to convince the reader that there is immense value to be unlocked with digital twins. Secondly, we wish to showcase these various case studies in the context of the proposed digital twin framework, validating the usefulness of this framework for more readily identifying and constructing future digital twins.

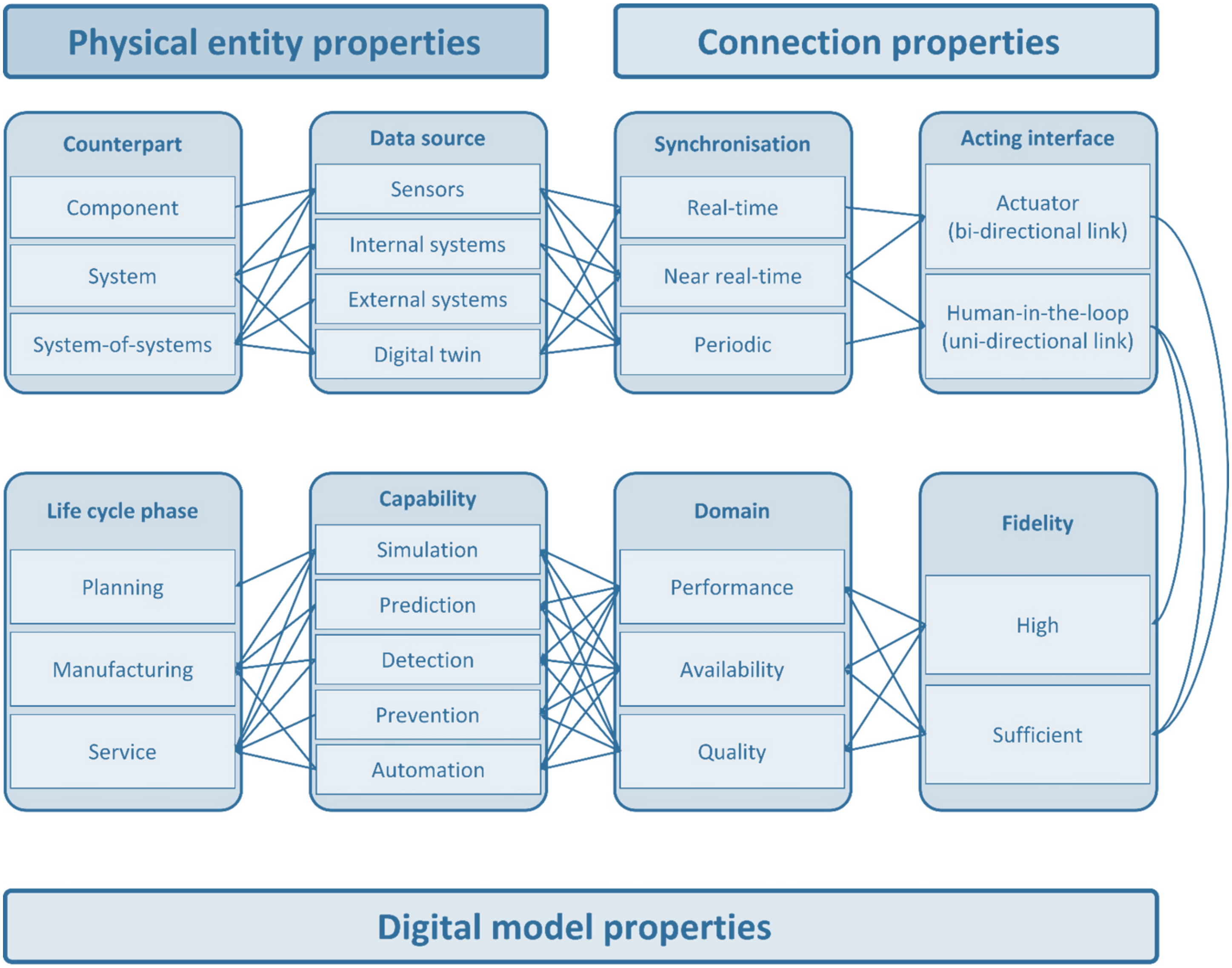


Figure 24 The inter-relationship between various properties of the proposed digital twin. The data source associated with the physical entity is mainly dependent on the counterpart being twinned. The counterpart, in turn, affects the type of data sources that will be applicable to the digital twin application. These data sources limit the synchronization frequency within the digital twin, which determine which acting interface is most useful for the digital twin application. The acting interface determines whether high-fidelity models are possible, or whether simpler, sufficient fidelity models need to be used. These models are applicable to any domain of application, and can utilize any capability to fulfill the digital twin goal. These capabilities are largely restricted by the life cycle phase of the twinned entity.

4 CASE STUDIES: DIGITAL TWINS IN MINING

The widespread adoption of digital twins has yet to materialize, but initial efforts within the mining sector have demonstrated their efficacy. This report focuses specifically on the extraction phase of mining (see Figure 25), excluding any downstream processing, as this is deemed beyond the scope of RAMMS.

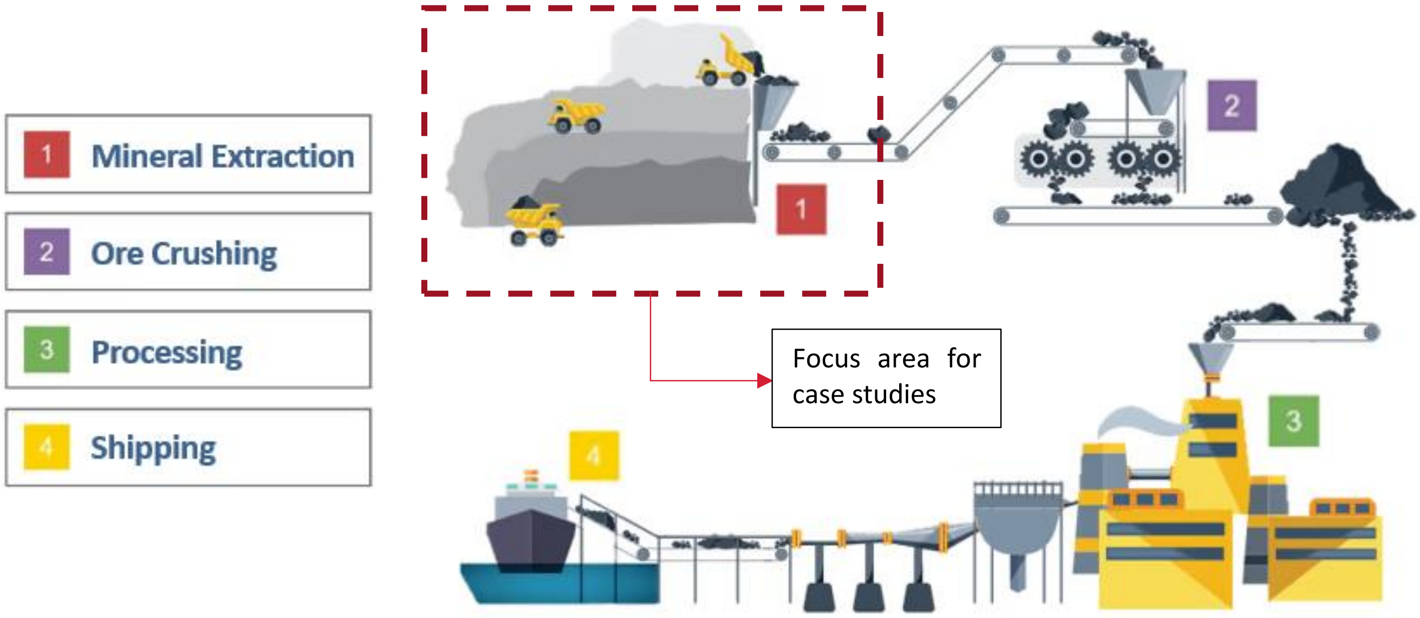


Figure 25 A simplified overview of the "pit to port" mine cycle. The specific focus of this work lies in the mineral extraction portion of the process. Image courtesy of https://shorturl.at/cxW46.

The mineral extraction process can further be broken down into multiple components, depending on the mining type. RAMMS focus mainly lies within underground mining, and therefore a simplified overview of underground mineral extraction is given in Figure 26. This illustration breaks down the mineral extraction task into rock breaking and transport operations. Rock-breaking can be achieved by various means, with drill and blast and rock cutting operations the most popular. Transport operations refer to activities which move the ore from the face to above ground for processing. Therefore, machines such as load and haulage trucks and conveyor are shown.

This section highlights five case studies within the mineral extraction portion of the underground mining process. This has a twofold goal. Firstly, the goal is to highlight state-of-the-art digital twin models in the underground mining environment, showcasing the immense value of each twin in their respective domains. Secondly, we wish to position these within the proposed digital twin framework to demonstrate its capacity to effectively describe diverse digital twins.

The five digital twin case studies investigate the following underground mining machines/components:

1. A road header used for rock cutting.
2. Hydraulic supports used in longwall mining.
3. A ventilation system in an underground mine.
4. Shovel, haul trucks and workshop configuration for improved mine operations and maintenance.
5. Conveyor belt motor.

The following sub-section outline each of these digital twin case studies in more detail, firstly summarising their value and secondly showcasing their description in the proposed digital twin framework.

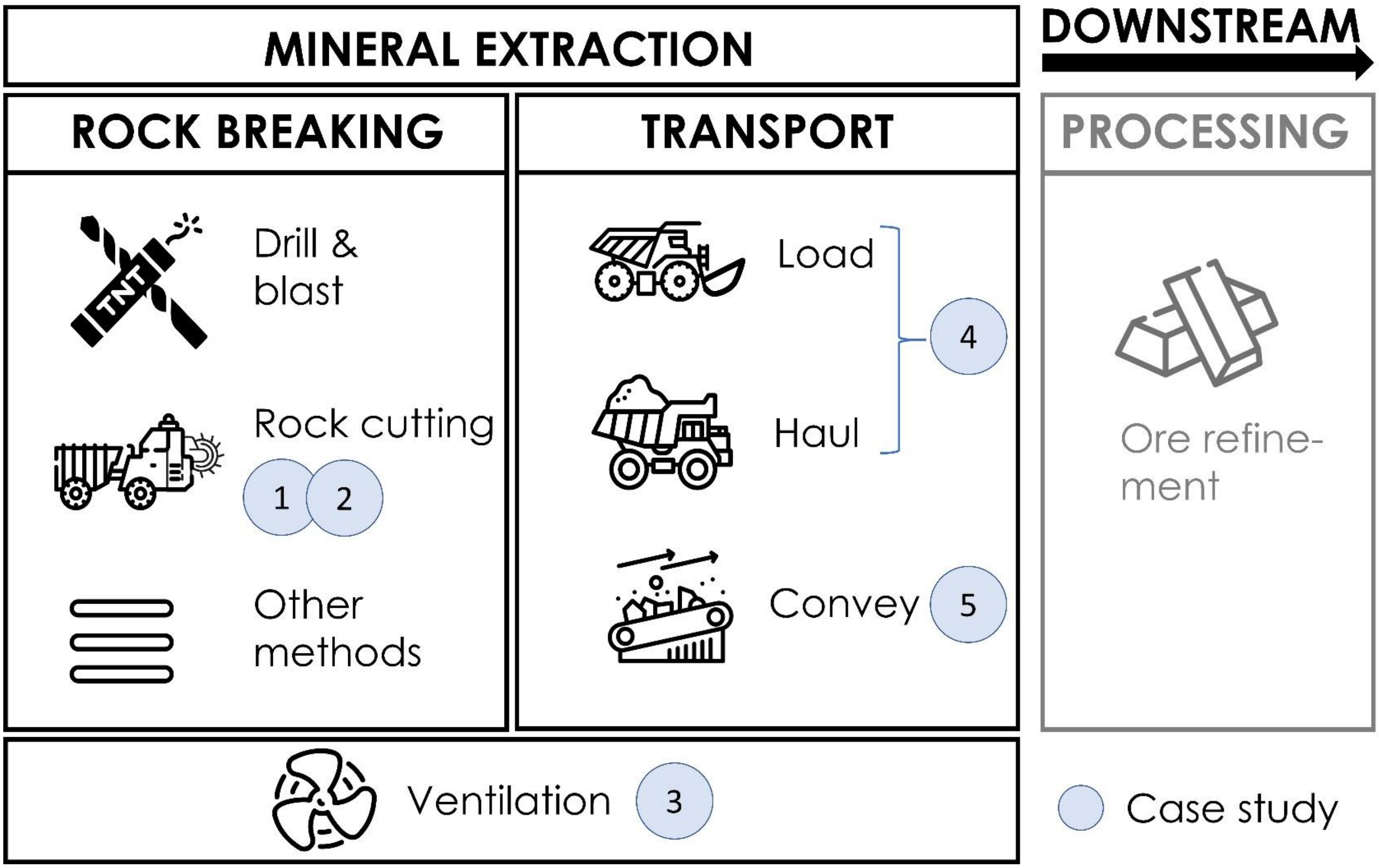


Figure 26 Mineral extraction process in mining. Five case studies are outlined according to their domain of applications. Case studies 1 and 2 address rock cutting. Case study 3 addresses mine ventilation, which is affected by both rock breaking and transport domains. Case study 4 deals with mine reconfiguration, which will likely affect loading and hauling operations. Case study 5 showcases conveyor belt motor condition monitoring.

4.1 CASE STUDY 1: ROAD HEADER DIGITAL TWIN

The first case study shows a digital twin developed by Zhang et al. (2024) for real-time monitoring and control of the mechanical performance of road headers. By inputting multi-sensor data into the digital twin, the system can predict the stress of the road header in real time. This stress prediction is used to adjust the swing speed of the cutting arm based on the predicted stress of the road header. By controlling the swing speed, the system can prevent overload conditions and optimize the cutting efficiency of the road header.

The digital twin system further utilizes physical rendering engines, databases, and 3D modelling tools to visualize and store the movement, performance, and control parameters of the road header. This provides a comprehensive view of the road header's state in real-time. An illustration of the rendered model is shown in Figure 27.

Overall, the digital twin system for boom-type road headers provides a valuable tool for real-time monitoring, adaptive control, and comprehensive understanding of the mechanical performance of road headers. Consequently, the efficiency of road headers is improved and failure rates are reduced.

A summary of the road header digital twin is seen in Figure 28. The digital twin used a deep learning model as the digital model. This deep learning model takes eight sensor inputs and predicts the stress on the cutting head at multiple locations. The input sensors are:

- Cutting arm vibration (acceleration)
- Lifting piston rod displacement

- Swinging piston rod displacement
- Pressure of left swing cylinder
- Pressure of right swing cylinder
- Pressure of lifting cylinder
- Cutting head speed
- Hydraulic motor torque

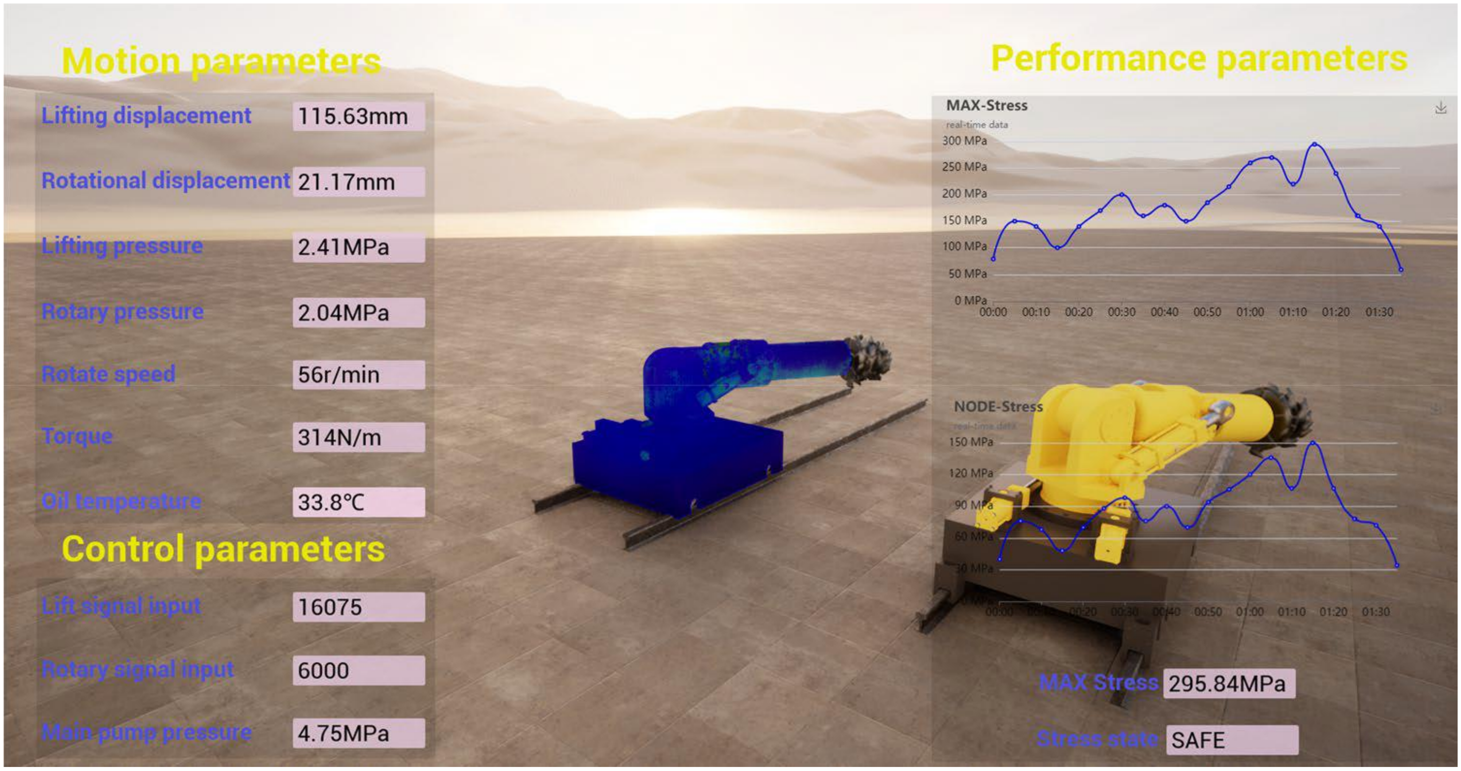


Figure 27 Road header digital twin input and output parameter reporting, along with visualisation of the physical road-header's state.

To pre-train the deep learning model (offline optimisation), a numerical dataset was compiled from FEM simulations of the road header, which were informed from a real-life run of the road-header. The goal of the FEM in this DT is to numerically estimate the stress on the entire road header, an otherwise impossible task from experimentation alone. This allowed the authors to compile a numerical database of 900 samples comprising eight sensor inputs and 278747 stress points outputs. The deep learning model could therefore be trained to convert the eight input sensors to a full-body stress estimate on the road-header. An additional benefit of the digital model is that it can help users visualise the stress on the road header in real time.

The deep learning model stress estimate is passed to a model predictive control algorithm, which also uses a neural network for the control task. The control system uses the predicted stress conditions to determine the head cutting speed, based on a preset stress threshold. If the threshold is crossed, the cutter speed is reduced.

Comparing the study to the proposed framework, we may identify the various elements corresponding to the dimensions of the proposed model. These are seen in Figure 29, Figure 30 and
Table 3. Figure 29 emphasizes the digital twin dimensions, where Figure 30 emphasizes the digital twin properties.
Table 3 summarises both the dimension and property aspects of the digital twin. By referencing the aforementioned figures and table, we can more readily place the road header digital twin in the context of the proposed digital twin framework:

The physical entity comprises a physical road header and a corresponding hydraulic system. The counterpart property is thus selected as a system. Various sensors measure the state of the physical entity, and are directly used as inputs to the digital model. The data source property is therefore sensor inputs. The digital model in this case study is a hybrid, transfer learning model. This model is a lower fidelity estimate of the high-fidelity FEM, hence the online DT is said to have sufficient fidelity. The digital model optimises performance (domain property), through predicting (capability property) stress, allowing the road-header to operate automatically (capability property). This twin is applied during the service phase of the physical entity's life cycle (life cycle phase property). The DT has a response time below 1 second, and therefore has a real-time synchronisation property. The action of the digital twin is enforced through a control system with hydraulic actuation, therefore the digital twin has an actuating interface property.

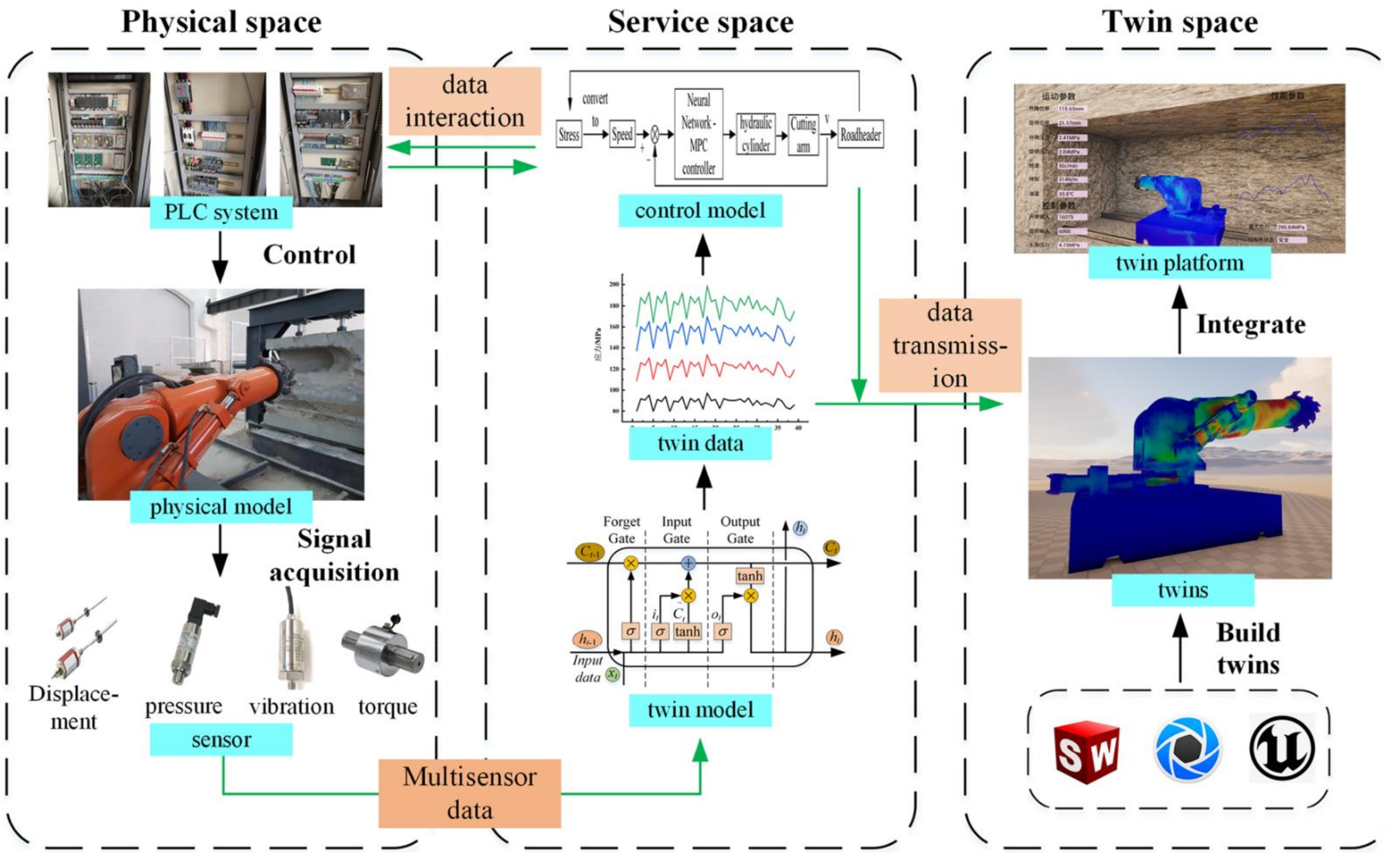


Figure 28 Integrated digital twin system developed by Zhang et al. (2024).

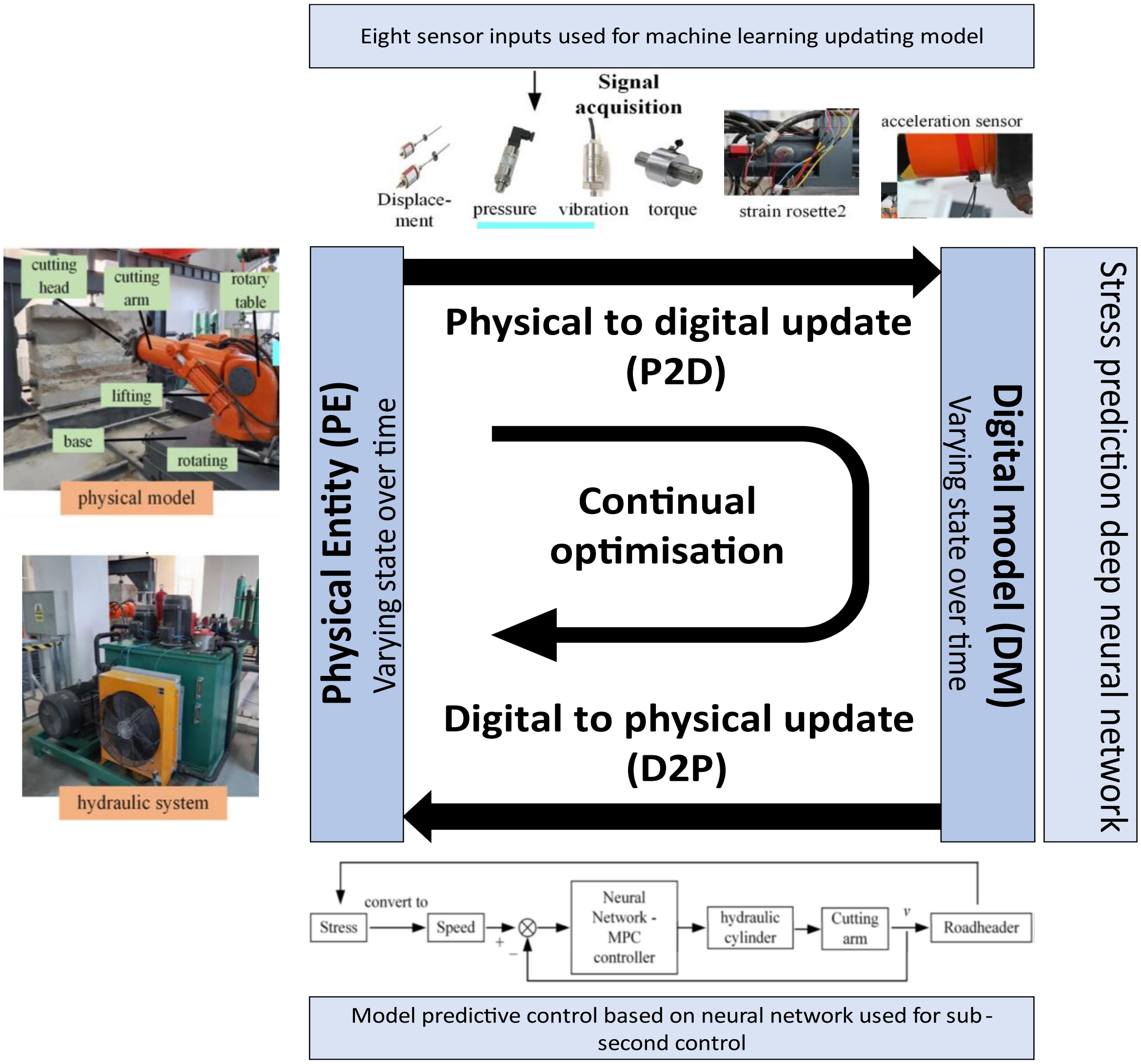


Figure 29 Road header digital twin as described by the proposed digital framework.

Table 3 A summary of the physical entity, digital model and communication interfaces along with their properties for case study 1.

Case study 1 summary according to proposed framework types and properties	
Physical entity	
Type	Road header
Counterpart	System
Data source	• Cutting arm vibration (acceleration) • Lifting piston rod displacement • Swinging piston rod displacement • Pressure of left swing cylinder • Pressure of right swing cylinder • Pressure of lifting cylinder • Cutting head speed • Hydraulic motor torque
Digital model	
Type	Hybrid – Transfer learning
Fidelity	Training: High (FEM), Application: Sufficient (NN)
Domain	Performance
Capability	Prediction and automation

Life cycle phase	Service phase
Physical to digital updating	
Type	Machine learning model updating
Synchronisation	Real-time (0.3s)
Digital to physical updating	
Type	Process control (Model predictive control)
Acting interface	Control system – Hydraulic pressure to enforce speed

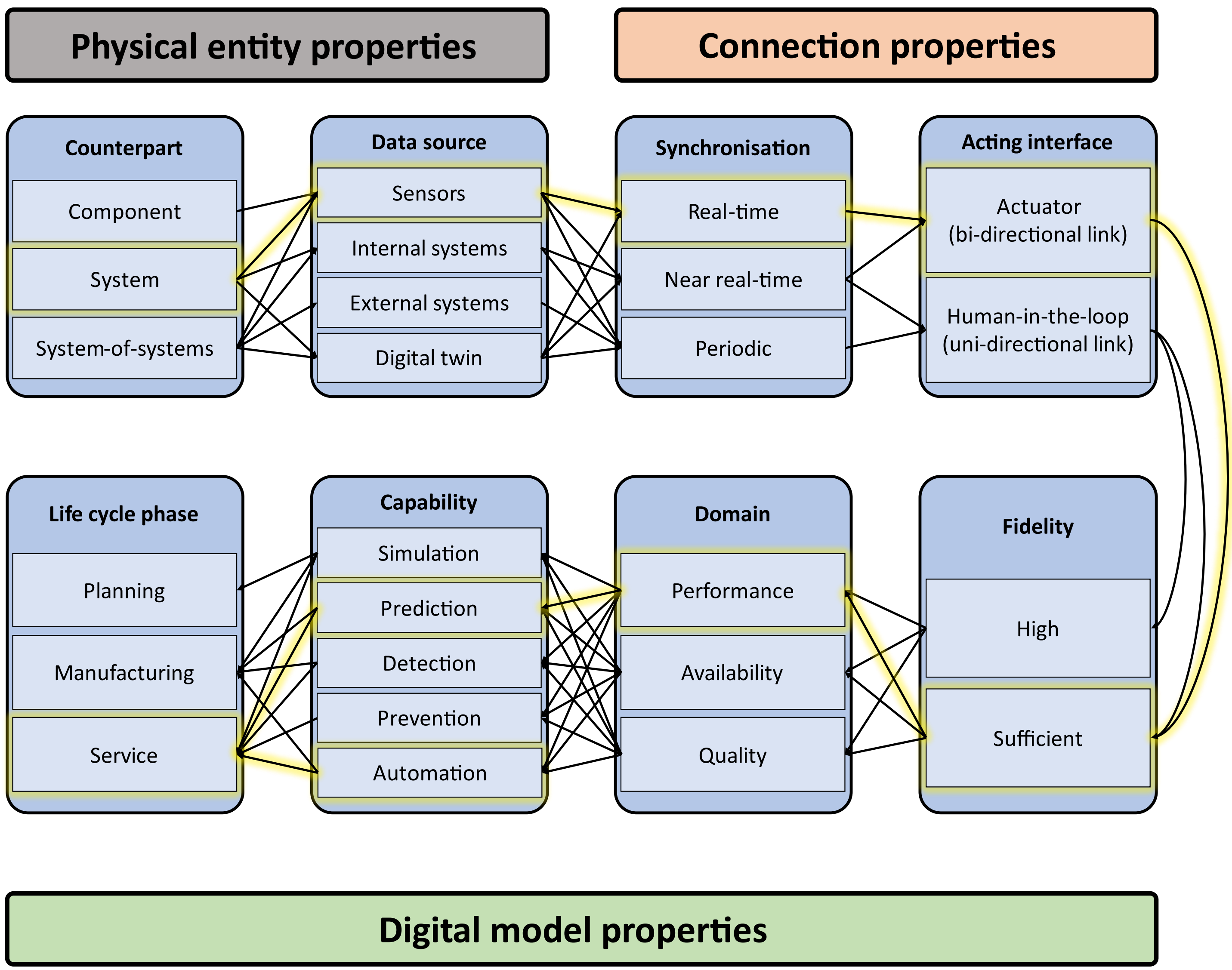


Figure 30 Property inter-relationship for the road-header digital twin

4.2 CASE STUDY 2: HYDRAULIC SUPPORT DIGITAL TWIN

The second case study is based on a hydraulic support digital twin from Jiao et al. (2022). The research work addresses the problem of position and attitude self-adjustment of hydraulic support groups in fully-mechanized mining. During the propulsion process, these hydraulic supports become skewed, requiring manual intervention to re-align them. The existing manual adjustment methods based on experience limit the efficiency and intelligence of the support system. The researchers propose an intelligent decision method for the self-adjustment of hydraulic support groups driven by a digital twin system. This method involves reconstructing the position and attitude of the support in a virtual space, comparing the estimated position and attitude to the desired position and attitude, completing virtual decision-making and execution, and optimizing the adjustment behaviour. The optimal control command derived from the virtual decision-making guides the real support to achieve self-adjustment. Figure 31 illustrates a completed pose correction on a physical system.

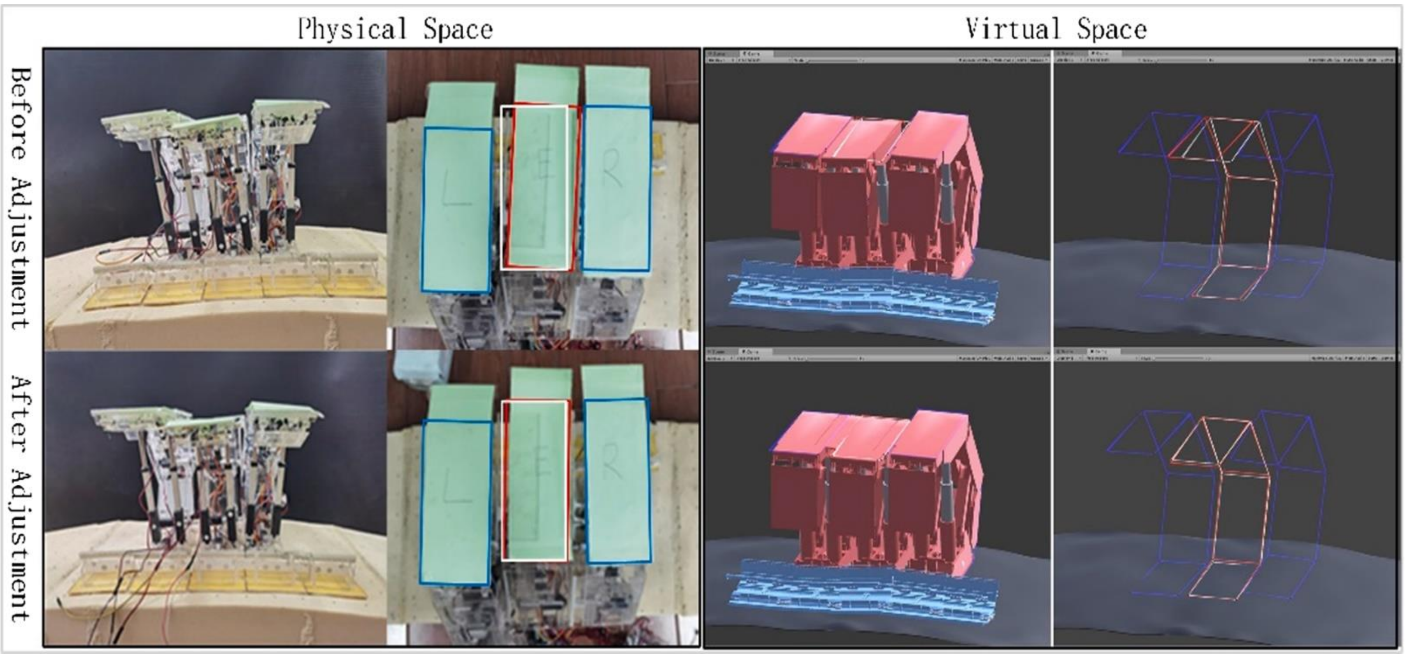


Figure 31 Support poses of the physical system before and after adjustment. The optimal adjustment parameters are simulated in digital space and control commands are sent back to the physical entity.

The value of this research lies in its contribution to the intelligent control of fully-mechanized mining equipment. By using a digital twin system and virtual reality technology, the proposed method enables the autonomous control and adjustment of hydraulic support posture. This method improves the stability and safety of the mining operation by effectively addressing the abnormal poses of the supports. An additional benefit from this work is that the estimated pose of the hydraulic support can easily be visualised in real-time, allowing operational centres to see the estimated state of the hydraulic supports in real-time.

The digital twin uses kinematic relationships, i.e. a physics-driven digital model. The physics-driven model uses a range of sensors (see Figure 33), namely:

- Three tilt sensors
- A 9-axis attitude sensor
- An ultrasonic ranging module

These sensors possess the necessary information to accurately estimate the position and attitude of the hydraulic support. The digital model is updated, and optimisations are run to determine which actuators need to be controlled, and by what amount. These conclusions are communicated to the physical entity through a control system, which adjusts electric push rods on the physical entity. In the scaled-up mining environment, these would likely be hydraulic cylinders, but the principle holds. The entire process is illustrated in Figure 32.

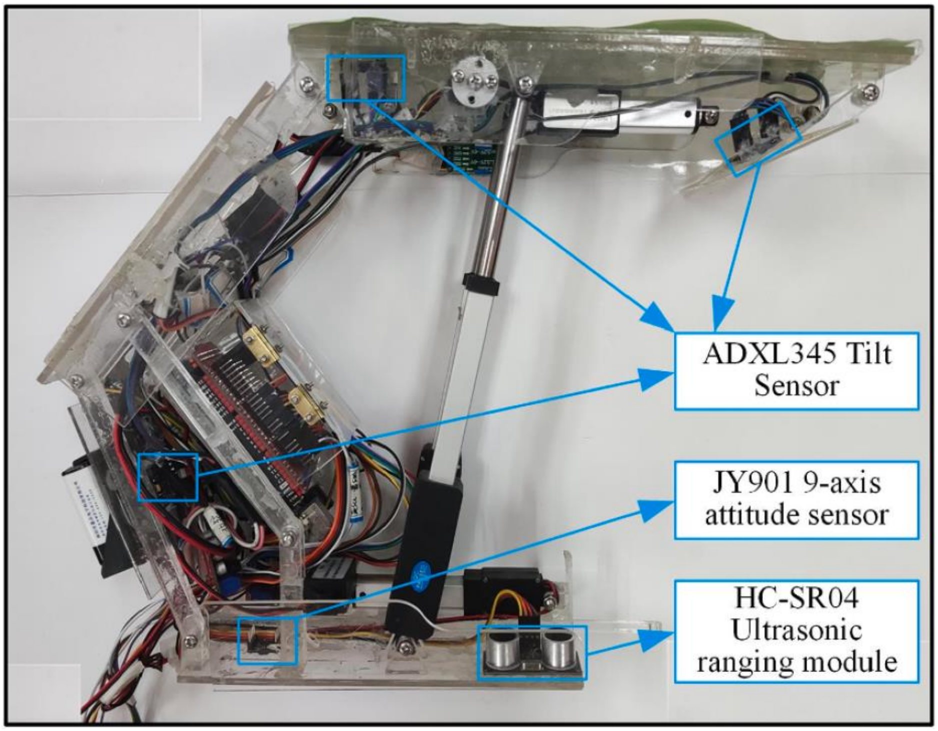


Figure 33 Illustration of the position of sensors on the physical entity.

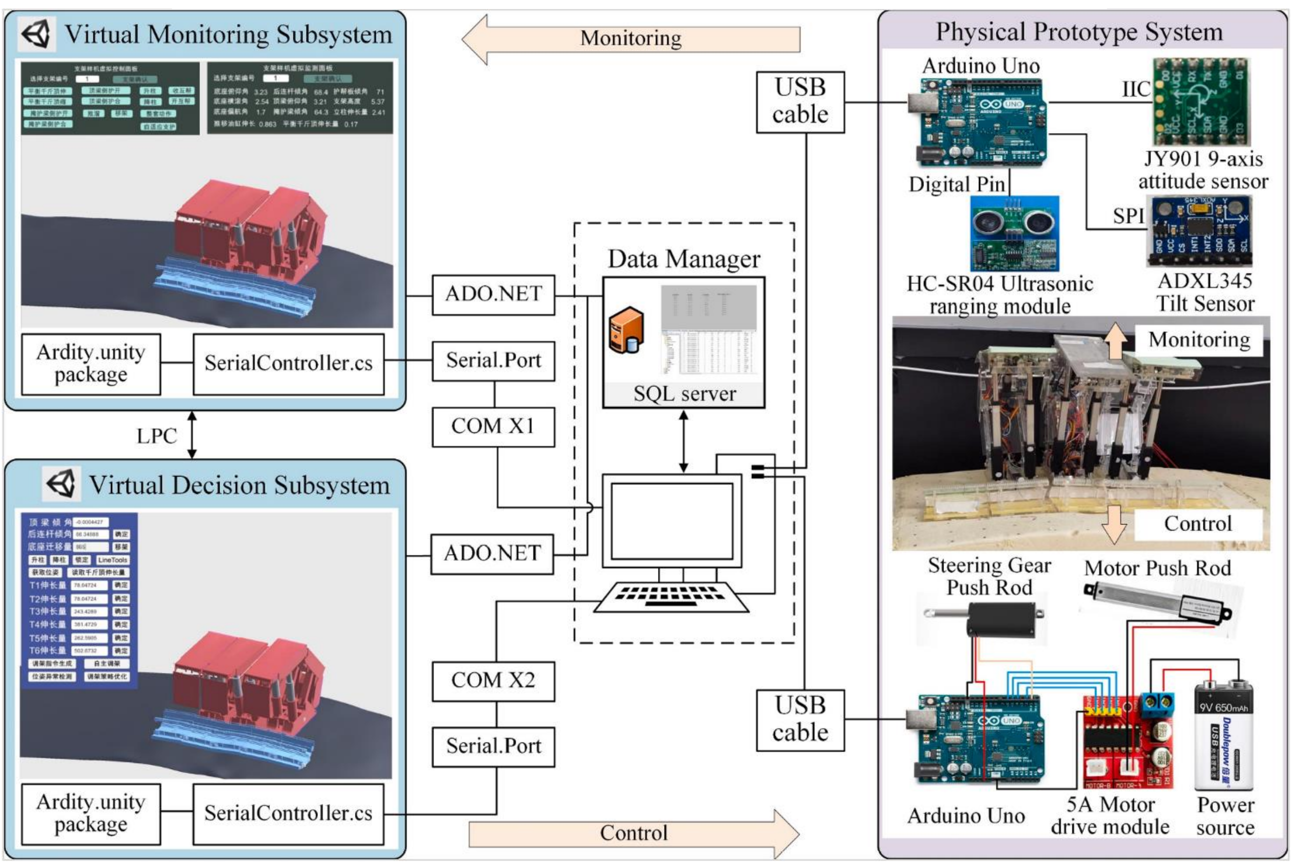


Figure 32 The digital twin setup used by Jiao et al. (2022) during their experiments.

Comparing the study to the proposed framework, we may identify the various elements corresponding to the dimensions of the proposed model. These are seen in Figure 34, Figure 35 and Table 4. Figure 34 emphasizes the digital twin dimensions, where Figure 35 emphasizes the digital twin properties. Table 4 summarises both the dimension and property aspects of the digital twin. By referencing the figures and table, we can more readily place the hydraulic support digital twin in the context of the proposed digital twin framework:

The physical entity comprises a physical (although scaled down) hydraulic support. The counterpart property is thus selected as a system. Various sensors measure the state of the physical entity, and are directly used as inputs to the digital model. The data source property is therefore sensor inputs. The digital model in this case study is a physics-driven, kinematic model of the hydraulic support. This model is a sufficient fidelity estimate of the hydraulic support, capturing the necessary sensory input to estimate the support's position and attitude. The digital model optimises performance (domain property), through simulating (capability property) optimal re-positioning of the hydraulic support's elements, allowing automatic (capability property) repositioning. This twin is applied during the service phase of the physical entity's life cycle (life cycle phase property). The DT does not need to react in real-time, but will likely reposition shortly after simulation, implying the synchronisation property is likely near real-time. The action of the digital twin is enforced through a control system with electric push rod actuation; therefore, the digital twin has an actuating interface property.

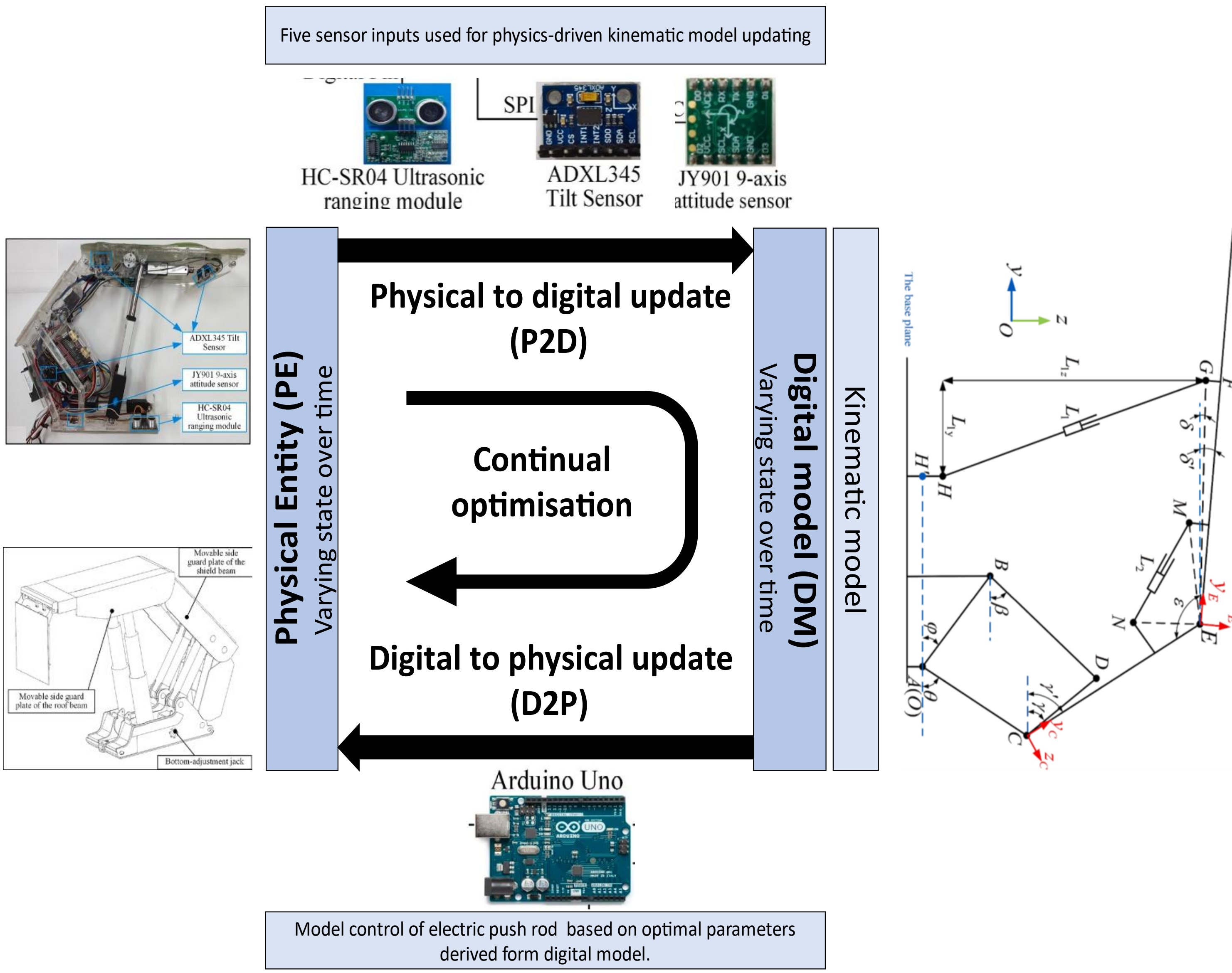


Figure 34 Hydraulic support digital twin as described by the proposed digital framework.

Table 4 A summary of the physical entity, digital model and communication interfaces along with their properties for case study 2.

Case study 2 summary according to proposed framework types and properties	
Physical entity	
Type	Hydraulic support
Counterpart	System
Data source	• Three tilt sensors • A 9-axis attitude sensor • An ultrasonic ranging module
Digital model	
Type	Physics-driven – Kinematic model
Fidelity	Sufficient fidelity kinematic model
Domain	Performance
Capability	Simulation, automation
Life cycle phase	Service phase
Physical to digital updating	
Type	Direct sensor measurements
Synchronisation	Near real-time
Digital to physical updating	
Type	Process control
Acting interface	Control system – Electric push rod to adjust position

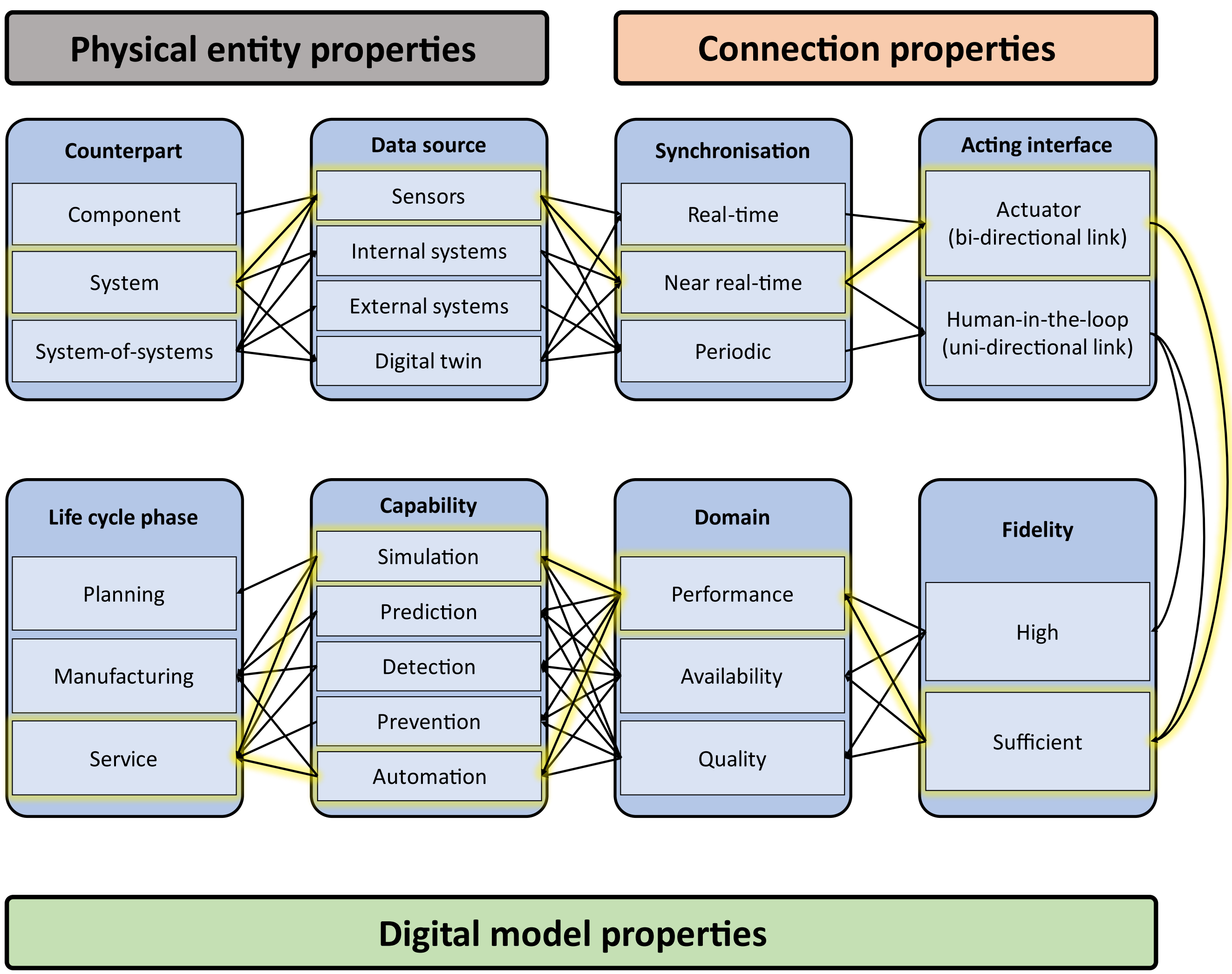


Figure 35 Property inter-relationship for the hydraulic support digital twin.

4.3 CASE STUDY 3: MINE VENTILATION DIGITAL TWIN

The third case study is based on a mine ventilation digital twin from Kychkin and Nikolaev (2020). The digital twin implements of a dynamic ventilation system in underground mining operations, which is crucial for ensuring safety and reducing energy consumption. Ventilation systems consume between 30% to 50% of the energy in underground mines (Kychkin & Nikolaev, 2020). By accurately modelling and simulating the ventilation system, the digital twin enables the optimization of air distribution and the selection of the most efficient mode of operation for the main fan unit. An illustration of the ventilation model is shown in Figure 36.

Existing methods of ventilation control often do not ensure rational energy consumption as they do not consider the dynamics of air distribution and changes in environmental parameters. The digital twin, by incorporating real-time data from energy meters, controllers, and air environment parameters, allows for dynamic calculation of control signals on the fan unit, accounting for the impact of the general natural draught and inertia within the ventilation system. This ensures more efficient energy usage in the ventilation system.

The digital twin also provides a comprehensive graphical view of the ventilation system. It enables the monitoring of energy consumption, the prediction of changes in parameters, and the continuous regulation of the ventilation system based on real-time and historical data.

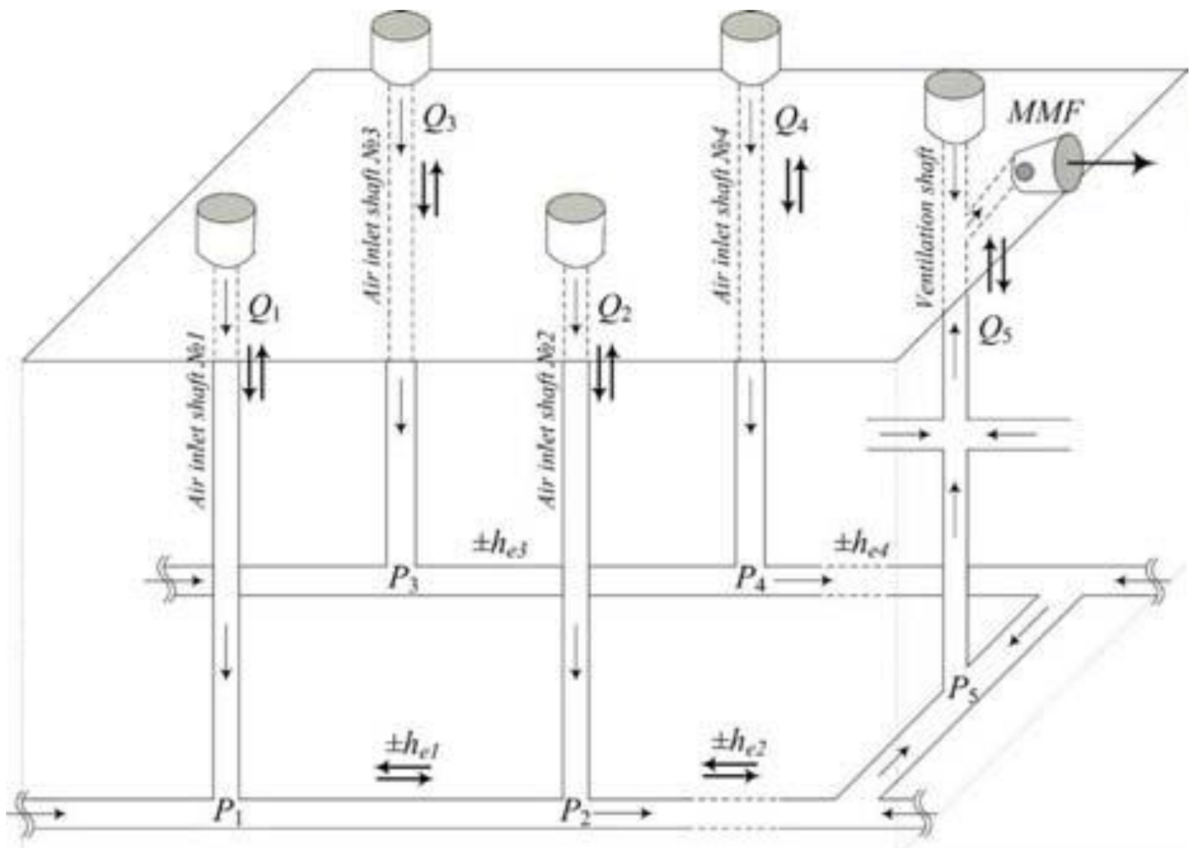


Figure 36 Digital model of the mine shaft layout, along with the important variables for the digital model optimization.

The digital model is rooted in physics, describing thermodynamic relationships between various shafts. Therefore, this digital twin uses a physics-driven digital model. The physics-driven model uses a range of sensors which originate from an internal software platform, InfluxData. The sensor data from this database reports:

- Temperatures
- Air pressure
- Airflow

These sensors possess the necessary information to estimate the thermodynamic quantities associated with the thermodynamic digital model. This information updates the digital model, allowing for calculation of optimal ventilation requirements. These calculations account for natural draughts between mine shafts as well as the inertia of the ventilation system. Model actions are sent to the ventilation and heating unit controllers. A summary of the digital twin system is illustrated in Figure 37.

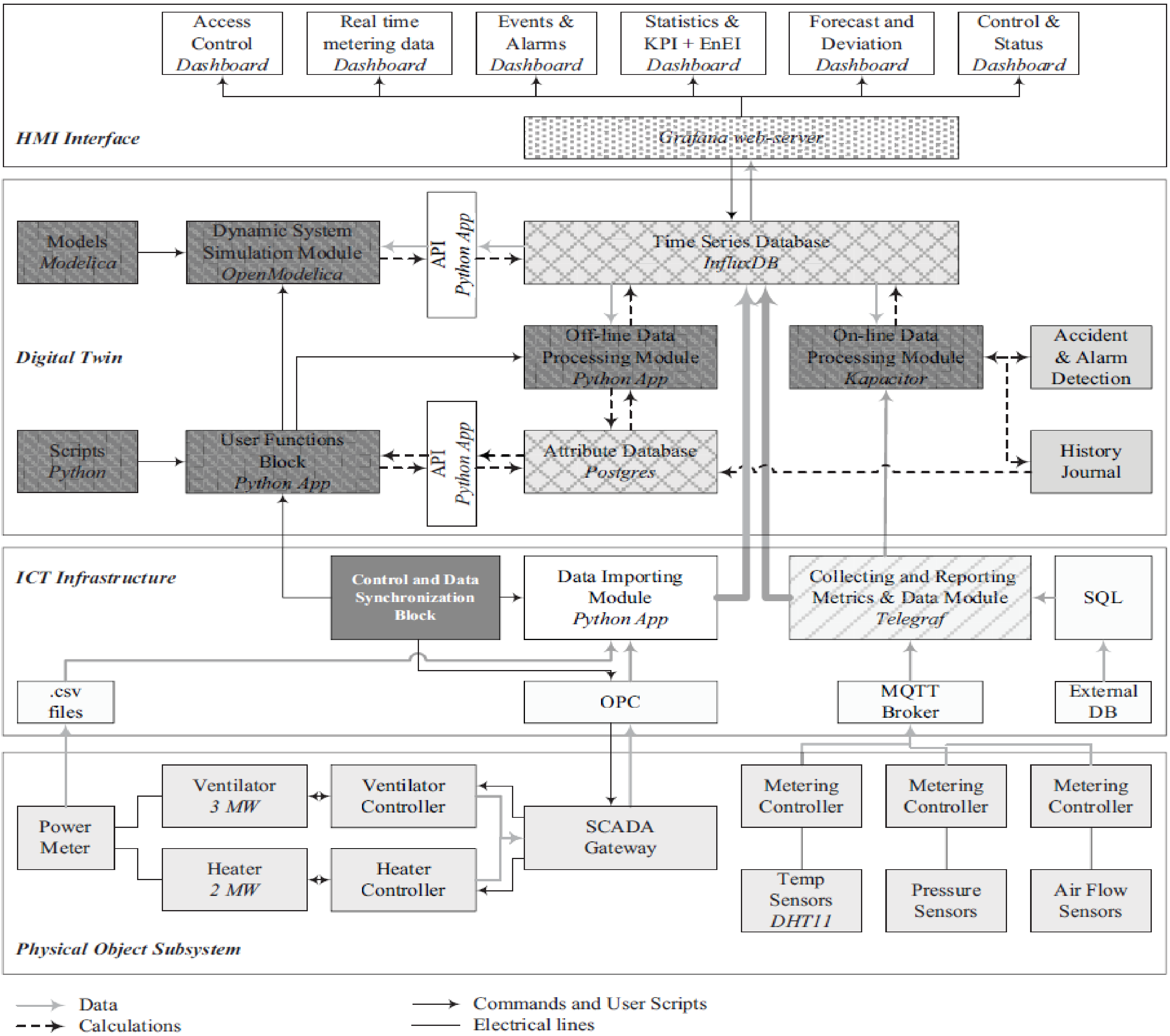


Figure 37 The digital twin setup used by Kychkin and Nikolaev (2020).

Figure 38, Figure 39 and Table 5 showcase case study 3 in relation to the proposed framework. Figure 38 describes the digital twin dimensions, with Figure 39 and Table 5 analysing and summarising the digital twin properties, respectively. By referencing the figures and table, we can more readily place the mine ventilation digital twin in the context of the proposed digital twin framework:

The physical entity may be defined as the mine shafts, a ventilator and a heater. Therefore, the counterpart may be seen as a system-of-systems digital twin because multiple distinct entities are being modelled simultaneously. Various software inputs are used by the digital model to summarise the state of the physical entity. Therefore, the data source property is internal software systems. The digital model in this case study is a physics-driven, thermodynamic model of the mine shaft. The model is a simplification of a much more complex reality, and hence the fidelity is described as being sufficient. The domain of application for the digital model is ventilation performance. The digital model is capable of prediction, and if correctly extended, automation. The twin is implemented during the service phase of the physical entity.

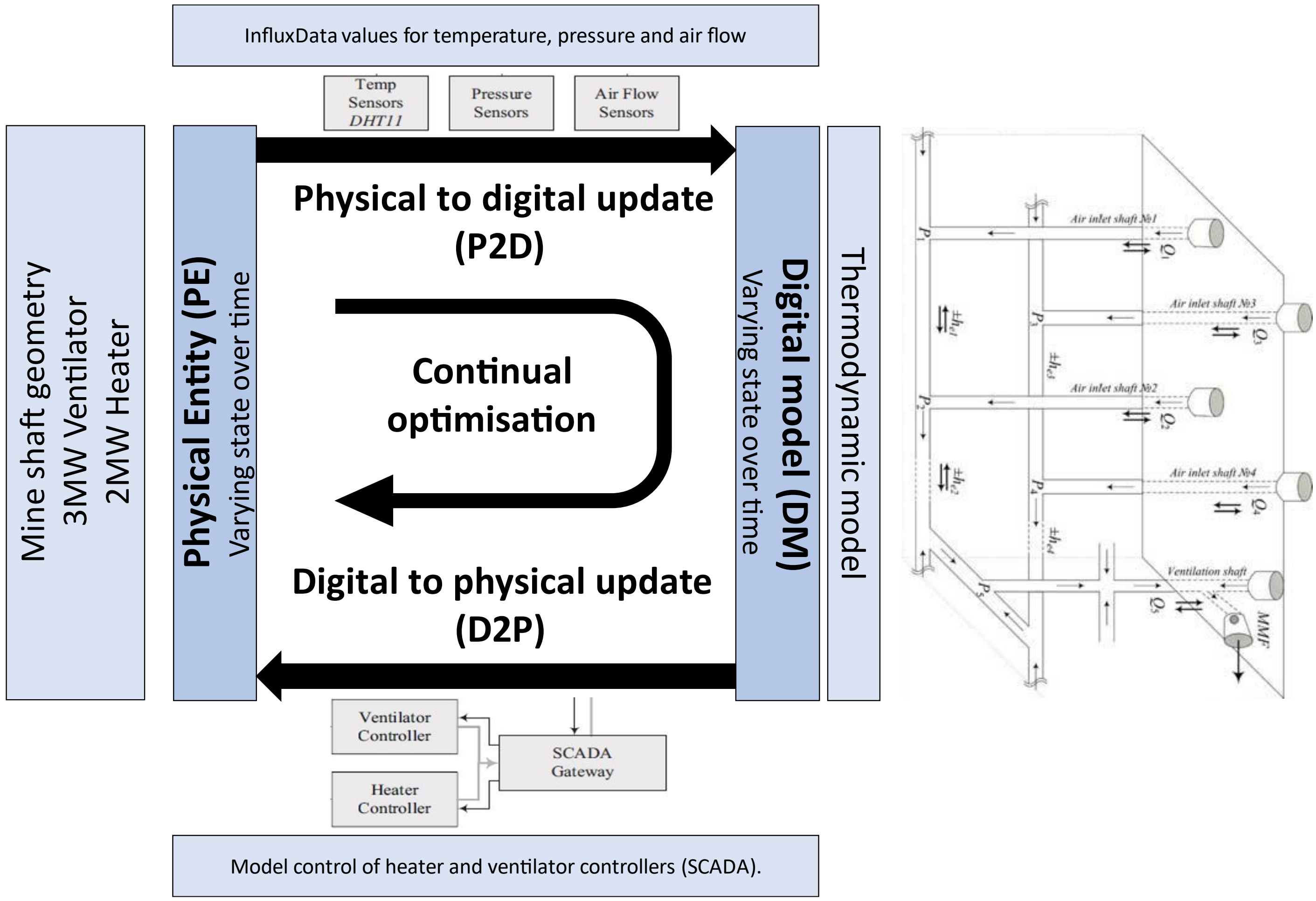


Figure 38 Mine ventilation digital twin as described by the proposed digital framework.

The DT does cannot react in real-time, because there is inherent inertia present due to the scale of ventilation within the mine. However, near real-time reaction is expected. The ventilation requirements are enacted through SCADA control systems, which control the ventilator and heater controllers in the mine. Therefore, actuating control is present in this digital twin application.

Table 5 A summary of the physical entity, digital model and communication interfaces along with their properties for case study 3.

Case study 3 summary according to proposed framework types and properties	
Physical entity	
Type	Mine shaft, ventilator and heater.
Counterpart	System-of-systems
Data source	Originating from an internal software system: • Temperature sensors • Pressure sensors • Air flow sensors
Digital model	

Type	Physics-driven – Fluid dynamic model
Fidelity	Sufficient fidelity fluid dynamic model
Domain	Performance
Capability	Prediction, automation
Life cycle phase	Service phase
Physical to digital updating	
Type	Internal software systems (InfluxData platform)
Synchronisation	Near real-time
Digital to physical updating	
Type	Process control (SCADA)
Acting interface	Control system – Electric push rod to adjust position

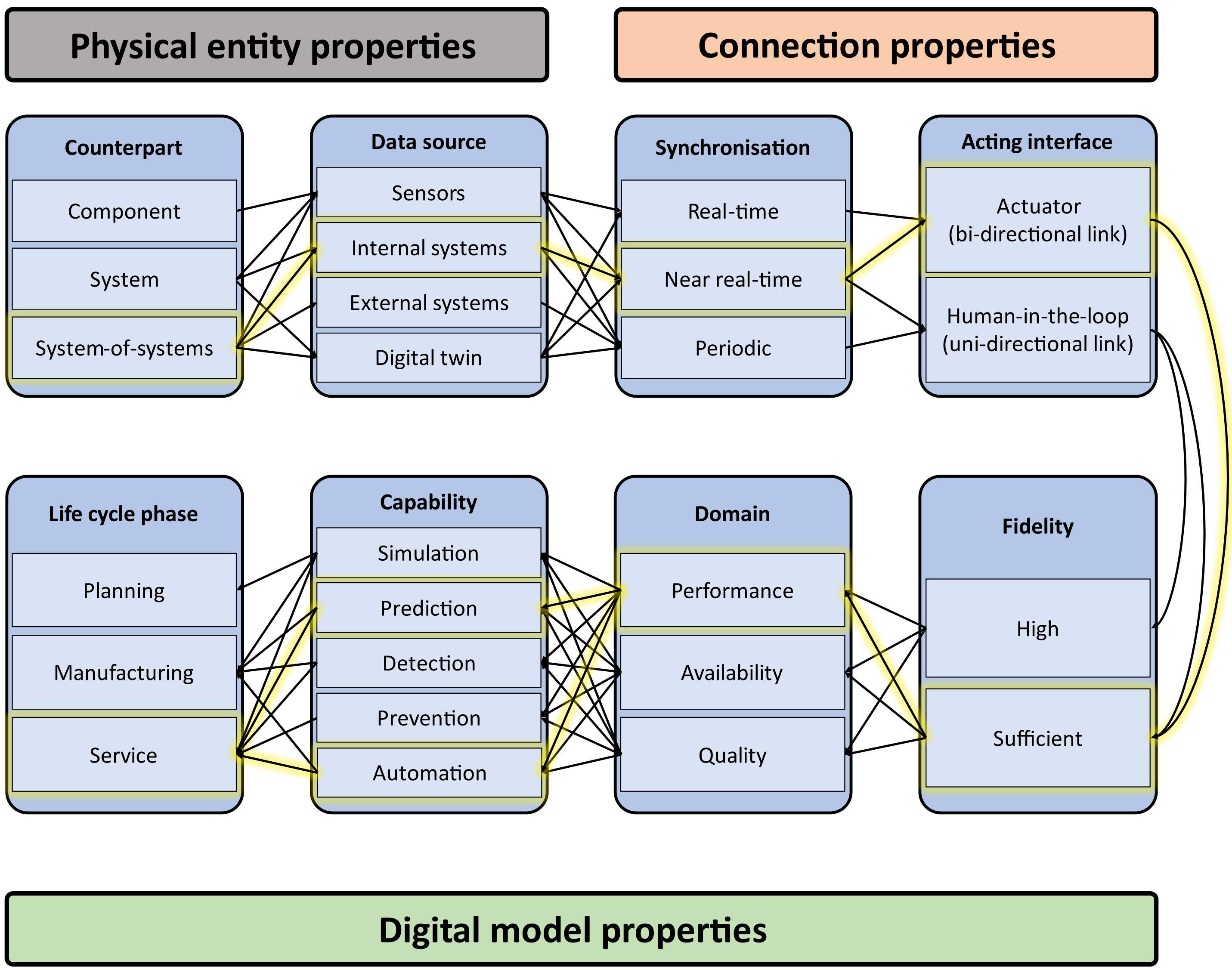


Figure 39 Property inter-relationship for the mine ventilation digital twin.

4.4 CASE STUDY 4: MINE OPERATIONS AND MAINTENANCE DIGITAL TWIN

The fourth case study is based on a digital twin for mine operations and maintenance (O&M) optimisation from Savolainen and Urbani (2021). The motivation for an O&M digital twin is indicated to be maintenance optimisation, as maintenance can contribute up to 60% to the production costs of a mine (Savolainen & Urbani, 2021).

Digital twin optimisations are achieved by simulating various mine fleet combinations and selectin the combination that maximises the mine's cashflow. Multiple facets of the mine are simulated simultaneously. Firstly, an O&M module is used to simulate the interactions between shovels, trucks, dump sites and workshops. Shovels extract ore, which needs to be hauled by trucks to dump sites. Shovels and trucks are subject to breakdowns and therefore each item has a maintenance policy within the simulation model. The second facet is a cashflow module, which accounts for factors such as the current commodity price and various operational costs associated with maintenance labour, leasing of machines etc. Figure 40 illustrates the digital twin structure and interrelationship between various models.

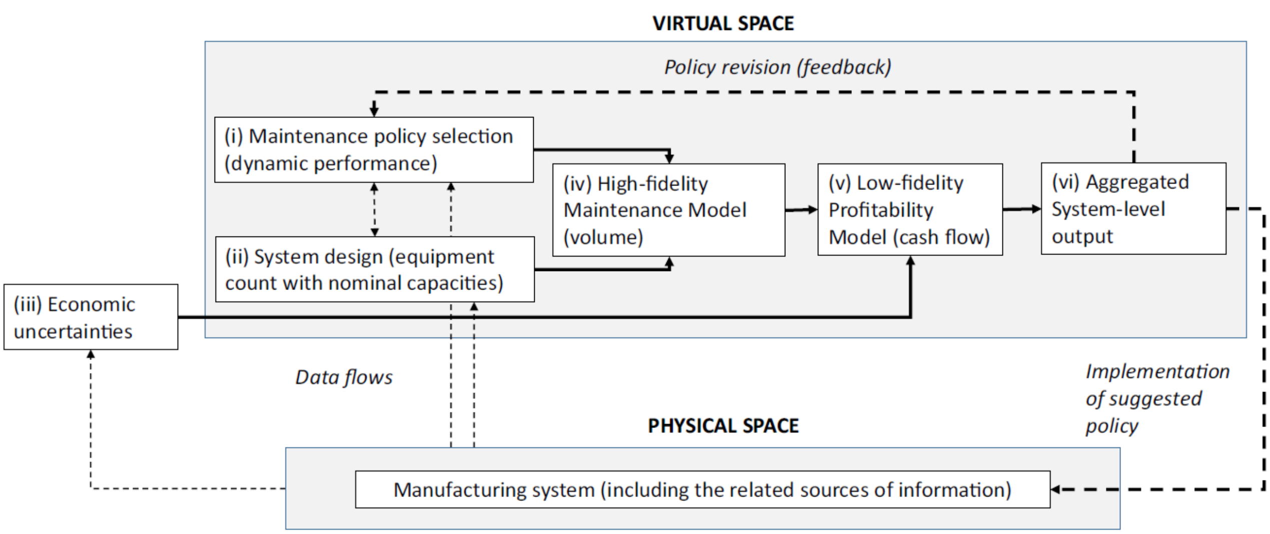


Figure 40 The digital twin setup used by Kychkin and Nikolaev (2020).

The simulation model has multiple variables which may be changed to evaluate optimal fleet selections. For example, predictive maintenance schedule, number of trucks, number of shovels, number of workshops, effect of commodity price etc. may all be evaluated to determine the optimal fleet for the mine. These optimal conditions are reported to management for human decision making. Figure 41 illustrate three different simulation possibilities of the DT.

Figure 42, Figure 43 and Table 6 showcase case study 4 in relation to the proposed framework. Figure 42 describes the digital twin dimensions, with Figure 43 and Table 6 analysing and summarising the digital twin properties, respectively. By referencing the figures and table, we can more readily place the mine ventilation digital twin in the context of the proposed digital twin framework:

For this twin, there are multiple entities. These include shovels, dump sites, workshops and trucks that may travel between them. Therefore, the counterpart property is a system-of-systems digital twin. The data source for the digital twin originates from both internal and external systems. Internal systems are used to calibrate the maintenance digital model, whereas external systems will be used to calibrate the cashflow model, which is dependent on the commodity pricing. The digital twin is useful for mine planning, implying digital model updates will likely only occur on a periodic basis. The digital model contains two distinct models, namely an operations and maintenance module and a cashflow module. Both modules use statistical modelling techniques to capture the behaviour of the physical entities. Monte Carlo simulations are used to derive interpretable results from the statistical models.

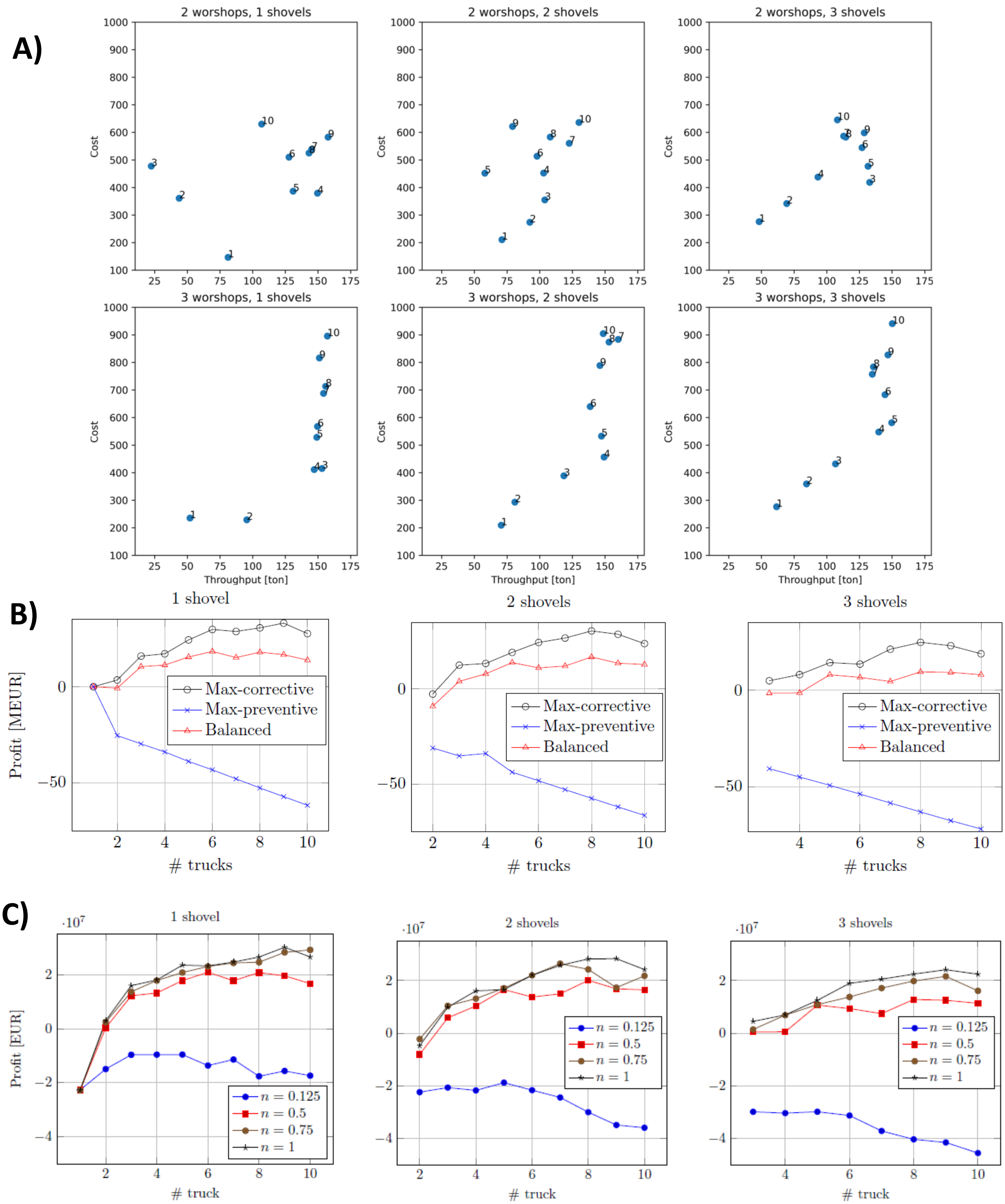


Figure 41 Various simulation capabilities of the digital twin. A) Identification of lowest cost per throughput for various fleet layouts. B) Optimisation of profit based on the maintenance philosophy. C) Optimisation of profit based on regularity of preventive maintenance.

The digital model simplifies various cost and time-based quantities, implying that it is of sufficient fidelity. The digital twin may be applied to optimise performance and availability domains, which are realised through simulation capabilities. The twin is likely to be applied during the planning phase of a mine, although it is certainly possible, to a lesser extent, to apply the model during service phase. At this stage, various equipment will already exist, and workshops and dumpsites cannot necessarily be extended or reduced. Finally, outputs from the digital model are not suitable for control systems and are rather useful for applications such as system reconfiguration, production planning and maintenance scheduling, all of which are actionable by humans.

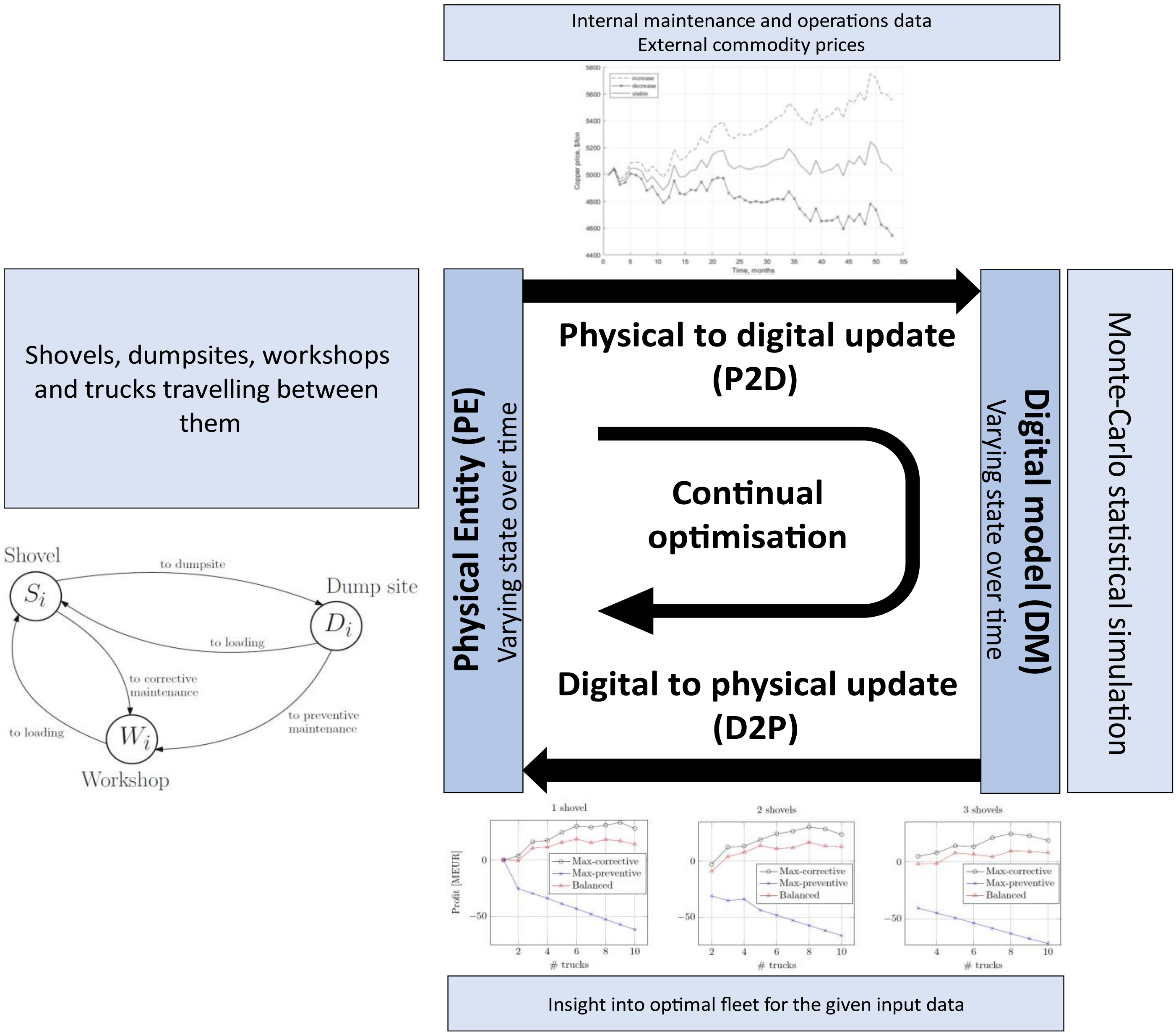


Figure 42 Mine operations and maintenance digital twin as described by the proposed digital framework.

Table 6 A summary of the physical entity, digital model and communication interfaces along with their properties for case study 4.

Case study 4 summary according to proposed framework types and properties	
Physical entity	
Type	Fleet: Truck, shovel, workshop, dumpsite
Counterpart	System-of-systems
Data source	Internal software systems: • Maintenance systems (prices, times) • Operations systems (prices, times)

	External software systems: • Commodity price
Digital model	
Type	Data-driven: Statistical model
Fidelity	Sufficient fidelity (simplification of reality)
Domain	Performance and Availability
Capability	Simulation
Life cycle phase	Planning or service phase
Physical to digital updating	
Type	Probabilistic model updating (from software systems)
Synchronisation	Periodic
Digital to physical updating	
Type	System reconfiguration, production planning, maintenance scheduling
Acting interface	Human in the loop – Act on optimised insights

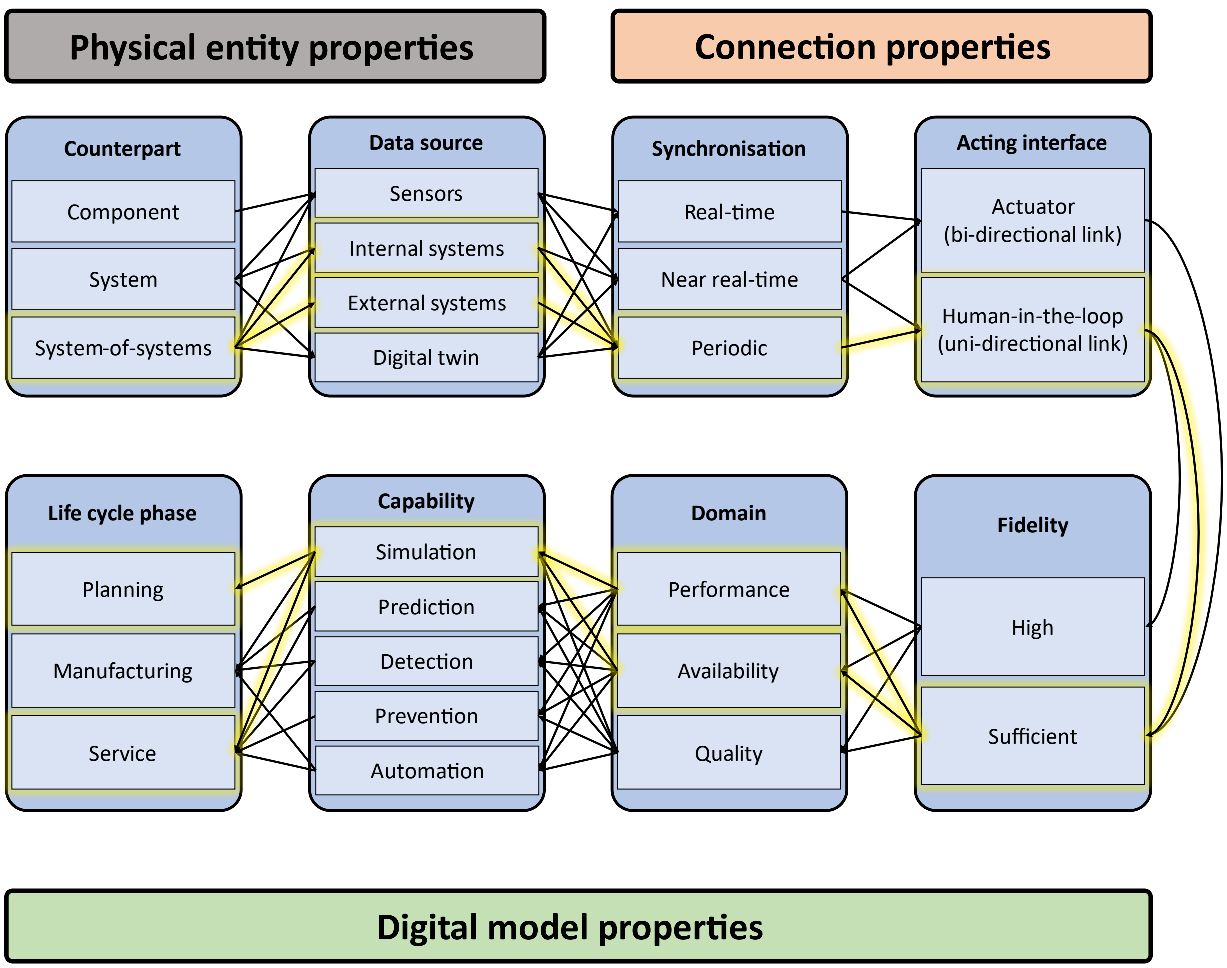


Figure 43 Property inter-relationship for the mine operations and maintenance digital twin.

4.5 CASE STUDY 5: CONVEYOR MOTOR DIGITAL TWIN

The fifth case study is based on a mine conveyor belt motor digital twin from (Huang et al., 2022). In the digital twin, a multi-layer approach is used to build the motor digital twin (see Figure 44). The service layer acts as a human-machine interface that distils the results from the digital twin. The model layer contains a "digital twin model" (see Figure 45) which comprises various sub-models. The only functional digital model in this work is the fault diagnosis model, which is a machine learning model trained to predict motor faults and intensities. The digital twin model is a first step in what will become a much larger model, extending to the monitoring of the conveyor itself. Transmission and perception layers outline how the digital model receives and distributes communication.

The value of this digital twin lies in its strong ability to diagnose fault types and intensities on the motor. The authors were able to train a machine learning model to identify over 98% of faults at various intensities on the motor. This approach allows the user to visualise the health state of the motor with great confidence, and track fault intensity levels. When the fault intensity is too high, preventive maintenance may be applied.

This work illustrates one example of a much larger field of condition monitoring within a digital twin environment. The approach followed within this work can easily be extended to various critical mine components.

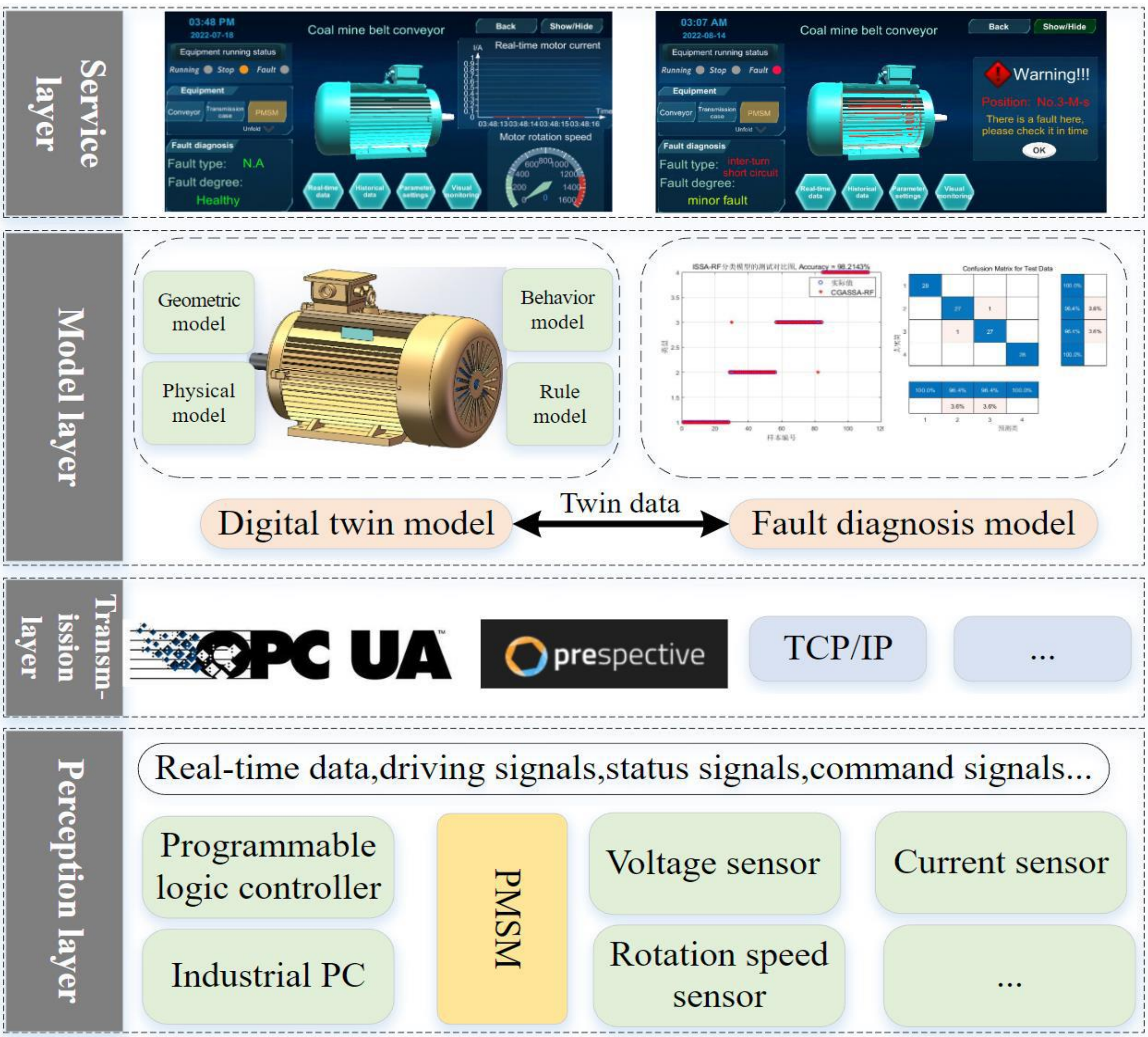


Figure 44 The digital twin setup used by Huang et al. (2022).

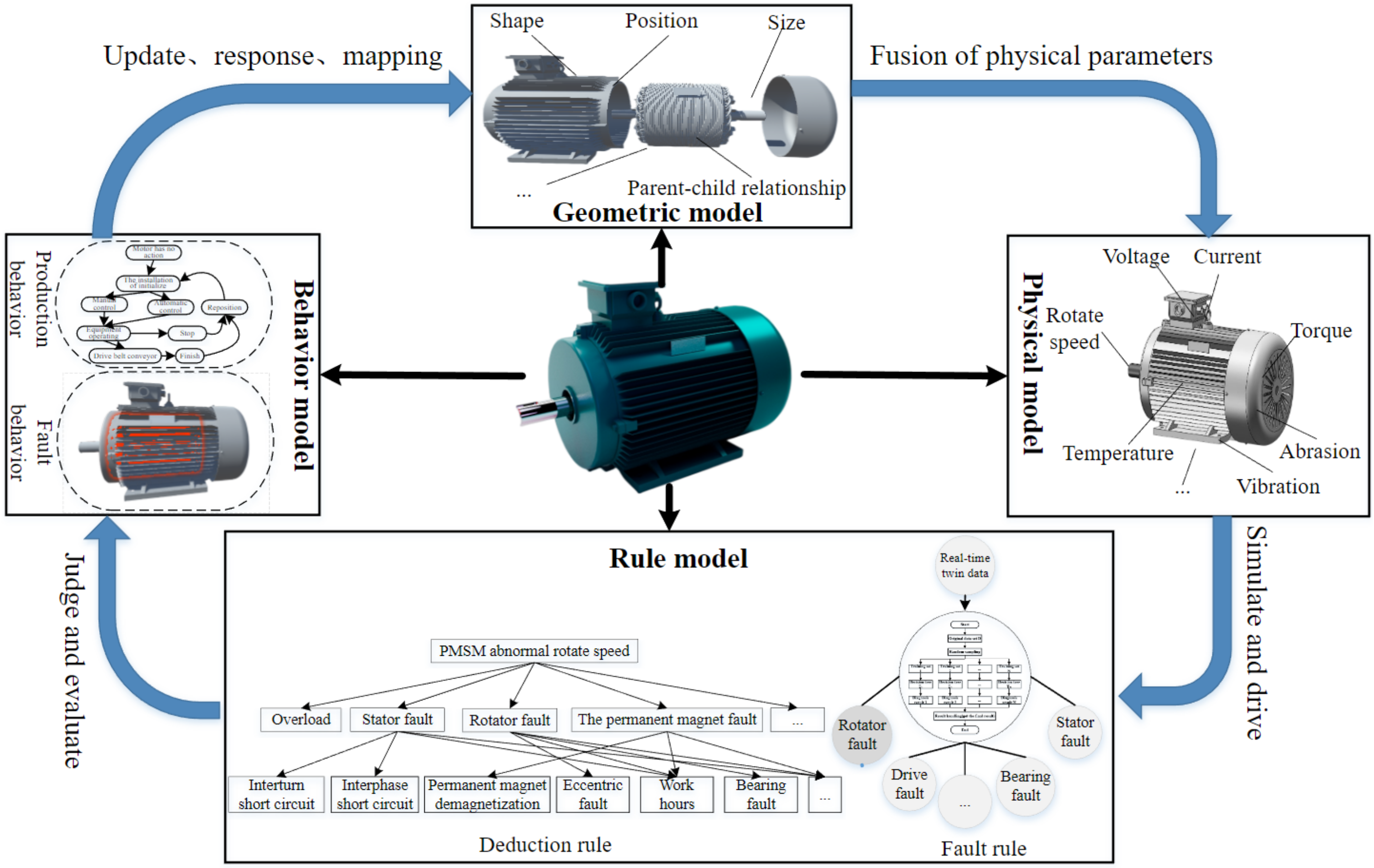


Figure 45 Digital model structure for the conveyor belt motor case study. The geometric model is used to determine the relevant entities for modelling and describe the relationships between these entities. The physical model contains various physics-based equations based on sensor data to model the relationships between the entities outlined in the geometric model. Ultimately this model estimates the physical entity's state. The rule model is used to convert estimated motor states into health states. Finally, the behaviour model serves as a visual aid to humans to quickly visualize the real-time behaviour of the physical entity.

Figure 46, Figure 47 and Table 7 showcase case study 5 in relation to the proposed framework. Figure 46 describes the digital twin dimensions, with Figure 47 and Table 7 analysing and summarising the digital twin properties, respectively. By referencing the figures and table, we can more readily place the conveyor belt motor digital twin in the context of the proposed digital twin framework:

The physical entity is quite simply the motor which drives the conveyor belt. However, this digital twin has logical extensions to any application where motors are used. The counterpart property is a component, as this is likely the lowest level of physical entity twinning which is still meaningful. The data source property is dependent on the digital twin application. This work describes a theoretical framework, where multiple digital models work together. In that instance, multiple current, temperature, voltage and speed sensors will likely be integrated. However, in the actual application of this digital twin, only current sensors are used, therefore the data source is said to be sensors. These sensors inform the digital model on a regular basis, and are therefore said to have high synchronization frequencies. However, this digital twin model is not used for real-time control, and therefore the synchronisation period is most suitably defined as near real-time. The digital model in the applied digital twin is a machine learning model. This model is pre-trained on physics-based data from a high-fidelity model. Therefore, the digital model is a hybrid, transfer learning model. The training model is therefore said to be high-fidelity, but the applied model is of sufficient fidelity. The model is used for fault diagnosis, and therefore it is used to enhance the domain of availability. The model achieves availability optimisation through its fault detection property. The model is applied during the service phase of the physical entity. Digital model insights are communicated to a visual dashboard, which is easily interpreted by a human. Therefore, the digital to physical updating occurs through the action of a human, which may decide that a certain level of motor degradation is sufficient for inspection or replacement.

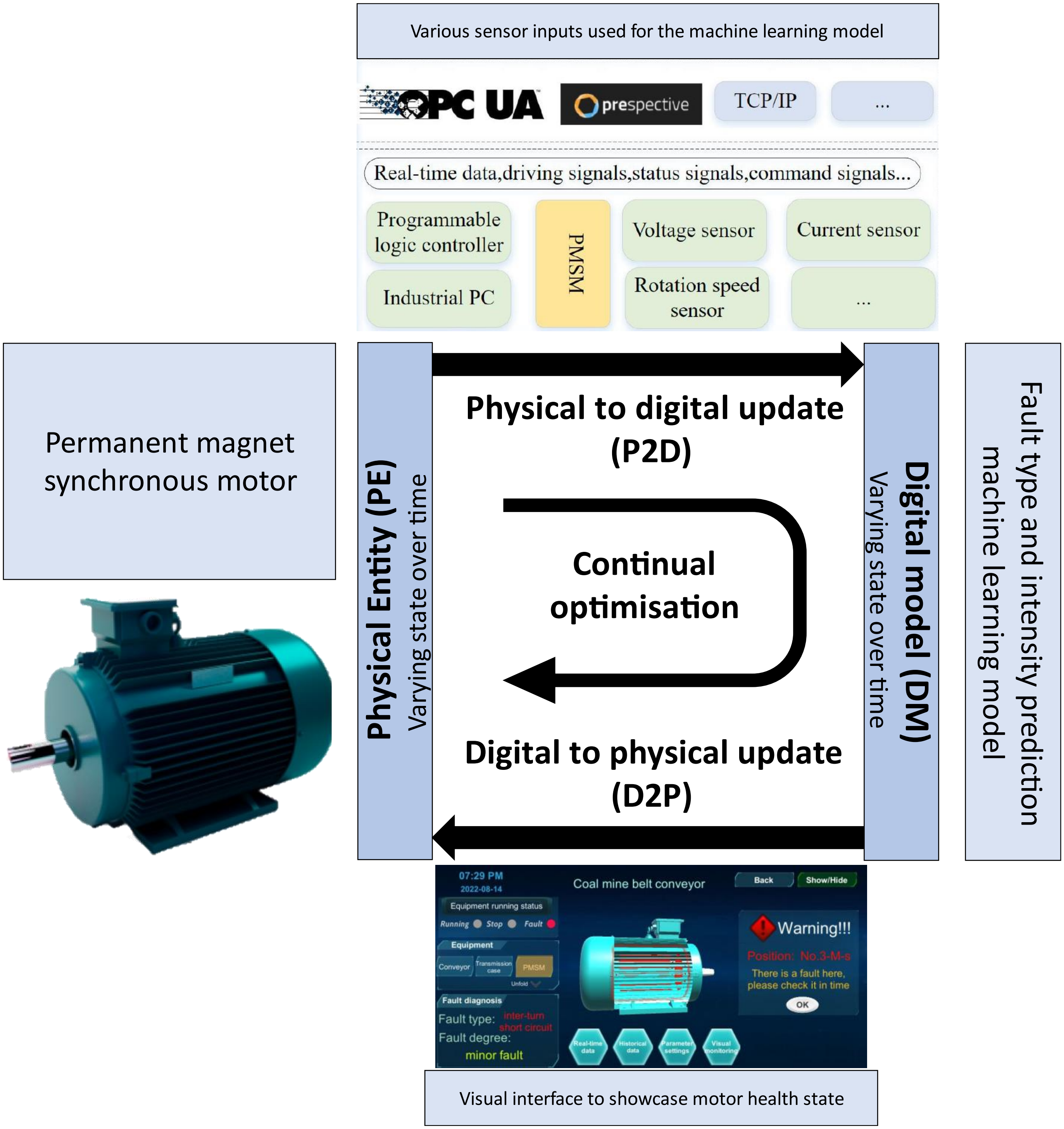


Figure 46 Mine conveyor belt digital twin as described by the proposed digital framework.

Table 7 A summary of the physical entity, digital model and communication interfaces along with their properties for case study 5.

Case study 5 summary according to proposed framework types and properties	
Physical entity	
Type	Permanent magnet synchronous motor
Counterpart	Component
Data source	Direct sensor values: • Time • A-phase current

	• B-phase current • C-phase current
Digital model	
Type	Hybrid: Transfer learning (train machine learning model on physics-driven simulation data)
Fidelity	Training uses high-fidelity physics, but implementation uses sufficient fidelity machine learning model
Domain	Availability
Capability	Detection
Life cycle phase	Service phase
Physical to digital updating	
Type	Direct sensor measurements
Synchronisation	Near real-time
Digital to physical updating	
Type	Maintenance scheduling: Use insights for maintenance
Acting interface	Visualisation dashboard, implying human-in-the-loop

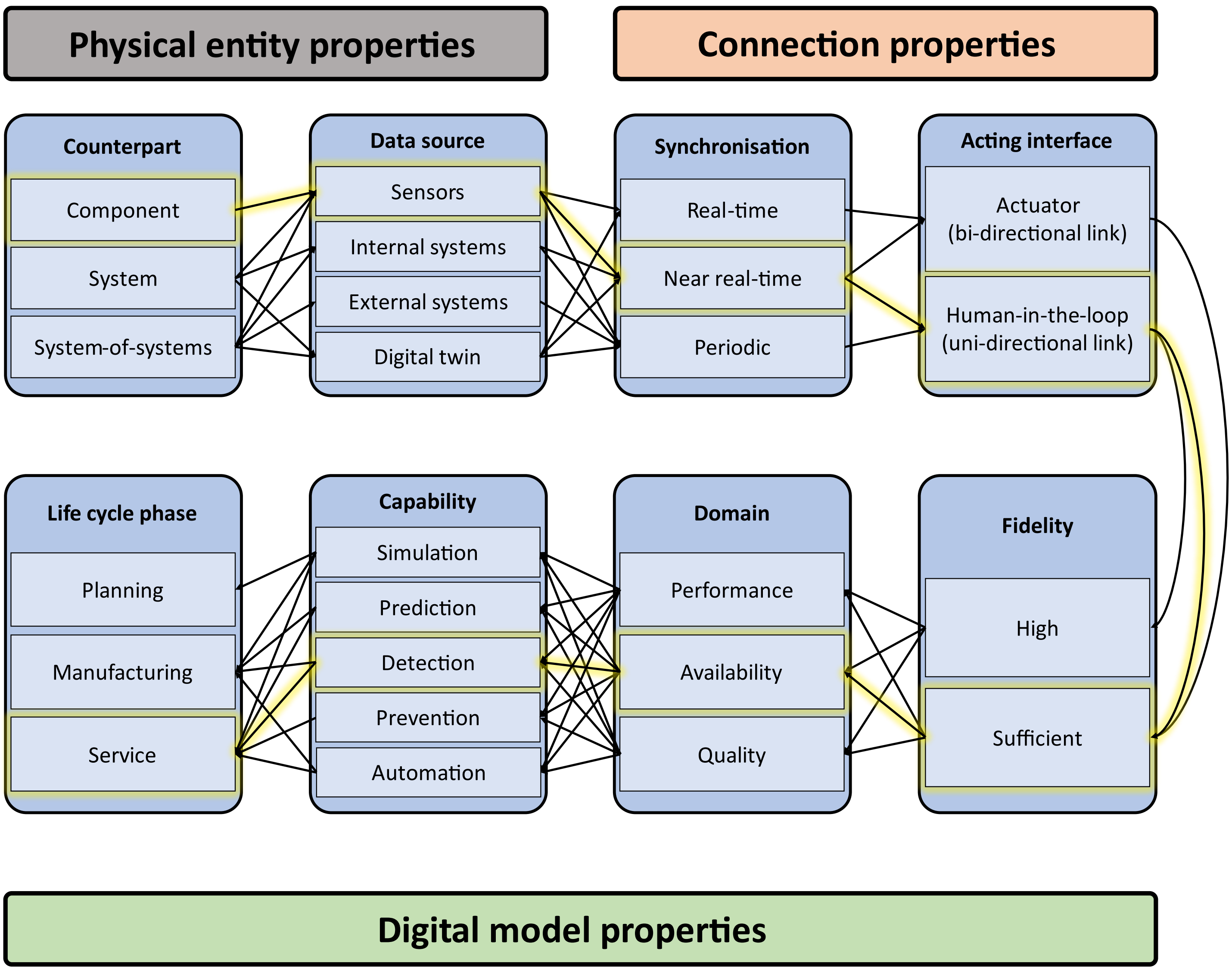


Figure 47 Property inter-relationship for the conveyor belt motor digital twin.

4.6 CONCLUDING REMARKS

The presented case studies have illuminated state-of-the-art advancements in digital twinning within the underground mining environment, showcasing twins across various aspects of the mineral extraction phase. These twins offer substantial value within their respective domains, underscoring the utility of employing twins in a digitalized context. Notably, many digital twin endeavours have concentrated on optimisation of functional parameters (e.g., cutting speed, hydraulic positioning, motor health), but as a byproduct, produced visual interfaces for human inspection. This insight underscores the importance of developing function-based digital twins over visual ones: While functional twins necessitate a broader technical knowledge base, they inherently encapsulate the information required for relaying the physical asset's state. Consequently, a functional digital twin naturally enables a visual interface, whereas the converse is not necessarily true.

Each twin has been dissected into its fundamental dimensions and associated properties within the proposed digital twin framework. This underscores the framework's value in creating a comprehensive yet descriptive summary to assess digital twins. The proposed framework therefore facilitates a clear understanding of the digital twin concept and succinctly highlights the most important properties of a digital twin.

During the search for mining case studies, two pivotal observations emerged. Firstly, literature pertaining to the processing phase of mining notably exceeded that of the extraction phase. This discrepancy suggests a lag in extraction-focused research, presenting a ripe opportunity for RAMMS to make a meaningful contribution on the digital twin front. Secondly, notable works on the drill and blast rock breaking method are absent. While this may reflect the prevalence of more mature rock breaking methodologies (rock cutting) elsewhere, South African mining's continued reliance on drill and blast warrants the development of digital twinning capabilities in this domain.

5 DIGITAL TWIN PROPECTS WITHIN THE RESEARCH ACTIVITY IN MECHANISED MINING SYSTEMS

The proposed digital twin framework has demonstrated its compatibility with existing case studies within the mining sector, offering concise but detailed identification of twinning prospects. Consequently, it is prudent to delve into the research activity in Mechanized Mining Systems (RAMMS) to ascertain how these efforts can be integrated into the proposed framework. RAMMS primarily focuses on three key research thrusts:

- Cutting and drilling hard rock
- Productivity and maintenance optimization
- Utilization, performance, and condition monitoring of mining equipment

In the ensuing sections, we leverage the existing research roadmaps for these thrusts to pinpoint digital twin opportunities. These opportunities are delineated and subsequently situated within the proposed digital framework, serving as a blueprint for forthcoming research endeavours. Given the intricacies involved, generating a digital twin framework image for each case is impractical. Therefore, we opt to summarize each digital twin in a tabular format, accentuating model types and properties for clarity and conciseness.

5.1 CUTTING AND DRILLING HARD ROCK

The efforts within this research thrust focus on understanding and optimizing the mining machine and the way it interacts with the rock face. An illustration of the research roadmap, along with identified digital twin opportunities are shown in Figure 48. The aim is to enhance mining operations by focusing on hard rock cutting and drilling. Currently, rock cutting hard rock represents an unsolved problem (Vogt, 2016), necessitating extensive research to optimize this procedure. While drilling and blasting remain relevant, the direct involvement of the MMP Advanced Orebody Knowledge (AOK) program in drilling, guides our attention in RAMMS towards advancing capabilities in rock cutting rather than drilling. AOK's drilling activities however presents a clear opportunity for collaboration between MMS and AOK.

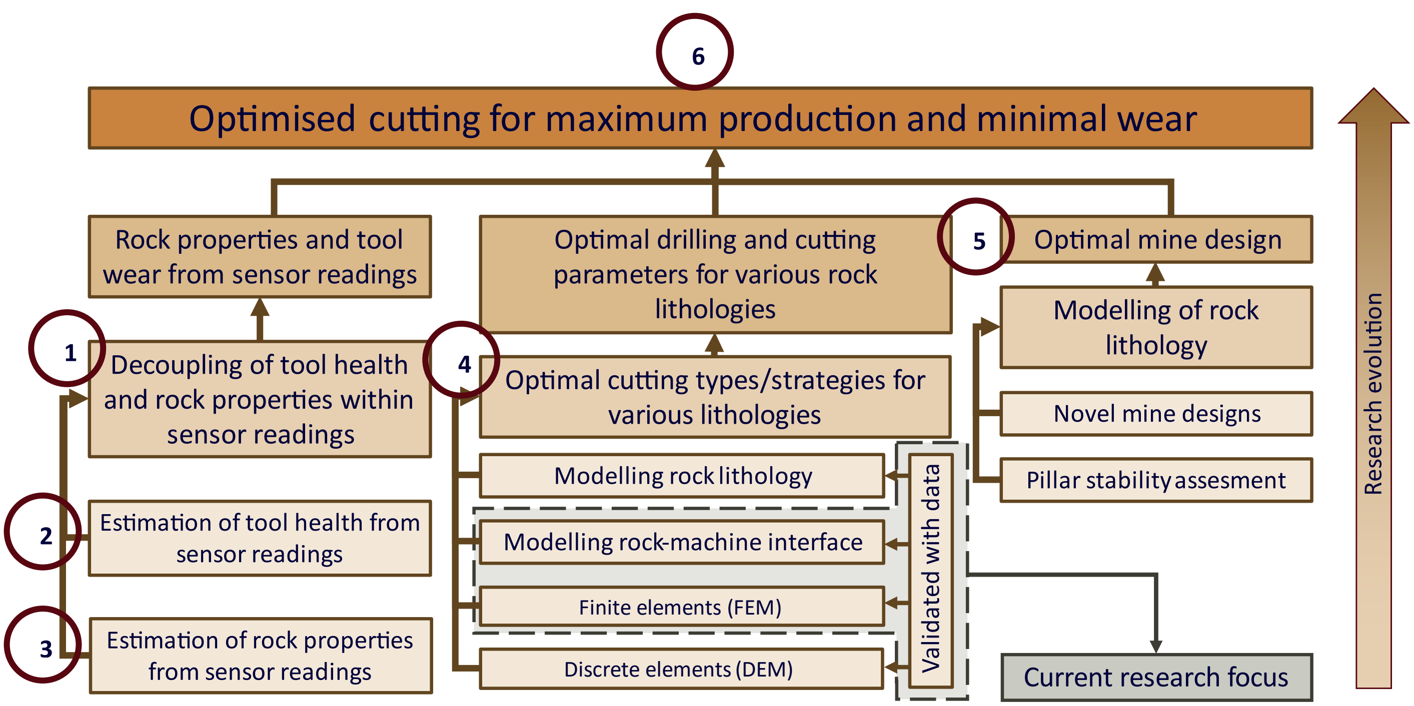


Figure 48 Research roadmap for thrust 1: Cutting and drilling hard rock. The six labelled circles indicate the six proposed digital twin models for thrust 1.

Understanding the cutter parameters, rock properties, and their interaction is crucial. This understanding enables us to optimize a cutting machine for reduced wear and increased productivity. We aim to develop a digital twin based on these insights, allowing us to predict optimal cutting parameters for specific rock lithologies. Comprehending rock lithology provides an alternative and equally influential method to optimize mining operations. By understanding the rock properties, we can adjust pillar sizes and extract more ore, potentially leading to innovative mining layouts.

The ultimate goal is to build a digital twin that allows us to adjust our machine cutting parameters optimally in real time, but also forecast what lies ahead using the latest data (link to AOK). Another twin could also update the mine layout in real-time as new rock property data becomes available. The real-time data from such a digital twin will naturally dovetail with Real-Time Information Management Systems (RTIMS). Thus, this research thrust aligns with the wider aim of creating a more efficient, interconnected mining environment.

The first identified digital twin has the purpose of separating measured signals according to their contribution from tool health and rock properties. Both elements are expected to contribute to the measurements during tool use and the ability to separate these elements gives one the ability to do better tool diagnosis and prognosis and rock lithology estimation. The output of the digital twin is envisioned to be reprocessed sensor data or state variables that correspond to each contributing factor. For example, the measured tool vibration may be split into components due to wear and components due to rock lithology. These outputs are therefore not directly actionable, but are extremely valuable to downstream digital twins.

1. Wear and rock property separation digital twin

Physical entity	
Type	Drill or cutting bit
Counterpart	Component
Data source	Direct sensor values: • Vibration • Temperature • Rotation speed • Torque Internal software systems: • Known rock lithology
Digital model	
Type	Hybrid: Delta learning – Use physics to define expected tool sensor readings, and assign delta to machine learning model for rock property correlation
Fidelity	Training uses high-fidelity physics, but implementation uses sufficient fidelity machine learning model
Domain	Performance
Capability	Prediction
Life cycle phase	Service phase
Physical to digital updating	
Type	Machine learning model updating
Synchronisation	Real-time
Digital to physical updating	
Type	Production planning and maintenance scheduling
Acting interface	Human-in-the-loop: Dashboard showing estimated properties

The second identified digital twin is ideally developed in conjunction with digital twin 1, so that the measured signals during use of the tool are sensibly separated into the components which originate from the tool and not the surrounding rock. The digital twin will use sensor data alongside insights from digital twin 1. These inputs are fed to a machine learning model to diagnose any tool problems and estimate remaining useful life (RUL). The machine learning model will also utilise physics-based simulated data for training to ensure the physics of the problem are captured by the machine learning model.

2. Tool diagnosis and prognosis digital twin	
Physical entity	
Type	Drill or cutting bit
Counterpart	Component
Data source	Direct sensor values: • Vibration • Temperature • Rotation speed • Torque Digital twin: • Digital twin 1 separated states
Digital model	
Type	Hybrid: Transfer learning – Use physics-generated data alongside real data to build a machine learning model for fault diagnosis and prognosis.
Fidelity	Physics-generated data requires high-fidelity simulation, but implemented machine learning model is sufficient fidelity.
Domain	Availability
Capability	Detection and Prediction
Life cycle phase	Service phase
Physical to digital updating	
Type	Machine learning model updating
Synchronisation	Periodic
Digital to physical updating	
Type	Maintenance scheduling
Acting interface	Human-in-the-loop: Predictive maintenance action to be taken by human

The third identified digital twin is ideally developed in conjunction with digital twin 2, so that the measured signals during use of the tool are sensibly separated into the components which originate from the tool and the surrounding rock. The digital twin will use sensor data alongside insights from digital twin 1. These inputs are fed to a machine learning model to estimate the rock lithology. The output of the model is not directly actionable by the tool system, but is very valuable to inform production teams of the rock properties. Therefore, the output is likely reported to a human, which will use the information in other systems.

3. Rock property estimation digital twin	
Physical entity	
Type	Drill or cutting bit
Counterpart	Component
Data source	Direct sensor values: • Vibration • Temperature • Rotation speed • Torque Digital twin: • Digital twin 1 separated states
Digital model	

Type	Data-driven: Machine learning model to learn complex correlations between measured signals and rock properties.
Fidelity	Sufficient fidelity machine learning model.
Domain	Performance
Capability	Prediction
Life cycle phase	Service phase
Physical to digital updating	
Type	Machine learning model updating
Synchronisation	Periodic
Digital to physical updating	
Type	Production planning
Acting interface	Human-in-the-loop: Rock properties likely used in external software system.

The fourth identified digital twin should ideally use the outputs from digital twin 2 (machine wear) and digital twin 3 (rock properties). Tool properties should also be measured (angles and positions of cutting head). Firstly, a high-fidelity FEM/DEM model may be developed to simulate various cutting scenarios based on various tool properties, rock properties and tool health state. These simulations are used to create a rich database of stresses on the cutting head. Secondly, a machine learning model may be trained on this simulated data to correlate the tool and rock properties (which are measurable and inferred from other digital twins) to certain stress conditions. These estimations may be used to control the cutting process directly, to maximise cutting performance whilst minimising tool wear.

4. Optimal cutting parameters digital twin	
Physical entity	
Type	Drill or cutting bit
Counterpart	Component
Data source	Direct sensor values: • Vibration • Temperature • Rotation speed • Torque Digital twin: • Digital twin 2 and 3 outputs
Digital model	
Type	Hybrid: Transfer learning – Use physics-driven model to create rich database. Train a ML model on the rich database for application.
Fidelity	Physics-generated data requires high-fidelity simulation, but implemented machine learning model is sufficient fidelity.
Domain	Performance and availability
Capability	Process control
Life cycle phase	Service phase
Physical to digital updating	
Type	Machine learning model updating
Synchronisation	Real-time
Digital to physical updating	
Type	Process control
Acting interface	Control systems that alter machine cutting speed/torque

The fifth identified digital twin is ideally developed to optimise the mine layout. The digital twin will use geological data alongside insights from digital twin 3 (rock properties digital twin). These inputs are fed to a high-fidelity physics-driven model which simulates the safety of various mine layouts. Uncertainty around the rock properties may lead to high safety factors when designing a mine. However, if some of this uncertainty can be reduced during cutting, by directly measuring and estimating rock properties (twin 3), it may become possible to be less conservative with safety factors, saving operational cost whilst increasing confidence in the surrounding rock structure. This digital twin is therefore designed for simulations, which are informed by live data from the mine.

5. Optimal mine layout digital twin	
Physical entity	
Type	Mine layout (Rock structure)
Counterpart	System (multiple cutters' information used)
Data source	Digital twin inputs (rock properties)
Digital model	
Type	Physics-driven: High fidelity simulation model (FEM/DEM) of mine structure, based on rock properties.
Fidelity	High fidelity
Domain	Production
Capability	Simulation
Life cycle phase	Planning and Service
Physical to digital updating	
Type	Probabilistic model updating
Synchronisation	Periodic
Digital to physical updating	
Type	Reconfiguration and production planning
Acting interface	Human-in-the-loop: Optimal mine design requires human input to enforce (communicate to equipment operators)

The sixth identified digital twin is developed for intelligent mine development. The goal is to automate many of the mine cutting operations based on the insights garnered from the previously described digital twins. The digital twin would use inputs from twins 2 (tool wear), 3 (rock properties), 4 (optimal cutting parameters) and 5 (optimal mine layout) to determine which actions a cutter should take next. This twin will likely require a physics-based kinematic model to describe the cutting motion of a cutting head. This motion will be determined based on the tool wear, rock properties, optimal cutting parameters and optimal mine layout. The physics-based model will determine which cutting parameters are most optimal to achieve the desired optimal mine layout. Therefore, the output of this model is control instructions for the cutting head, which take into account a large array of factors.

6. Intelligent mine development digital twin	
Physical entity	
Type	Cutting tool
Counterpart	System-of-systems (incorporating multiple digital twin systems)
Data source	Digital twins: • Digital twin 2 • Digital twin 3 • Digital twin 4 • Digital twin 5
Digital model	
Type	Physics: Kinematic model
Fidelity	Sufficient fidelity
Domain	Production and availability
Capability	Automation

Life cycle phase	Service phase
Physical to digital updating	
Type	Probabilistic model updating
Synchronisation	Real time
Digital to physical updating	
Type	Process control
Acting interface	Control system – Cutting parameters

5.2 PRODUCTIVITY AND MAINTENANCE OPTIMISATION

Mechanization is seen as a necessity for survival across conventional PGM mines, regardless of size. Mechanization offers two critical advantages: predictable machine and mining operations control and access to real-time operational data. Therefore, research thrust 2 positions itself to understand the "zoomed out" view of mining, focusing on modelling machines within a larger network-centric mine. The research roadmap for this thrust along with two identified twinning opportunities are shown in Figure 49.

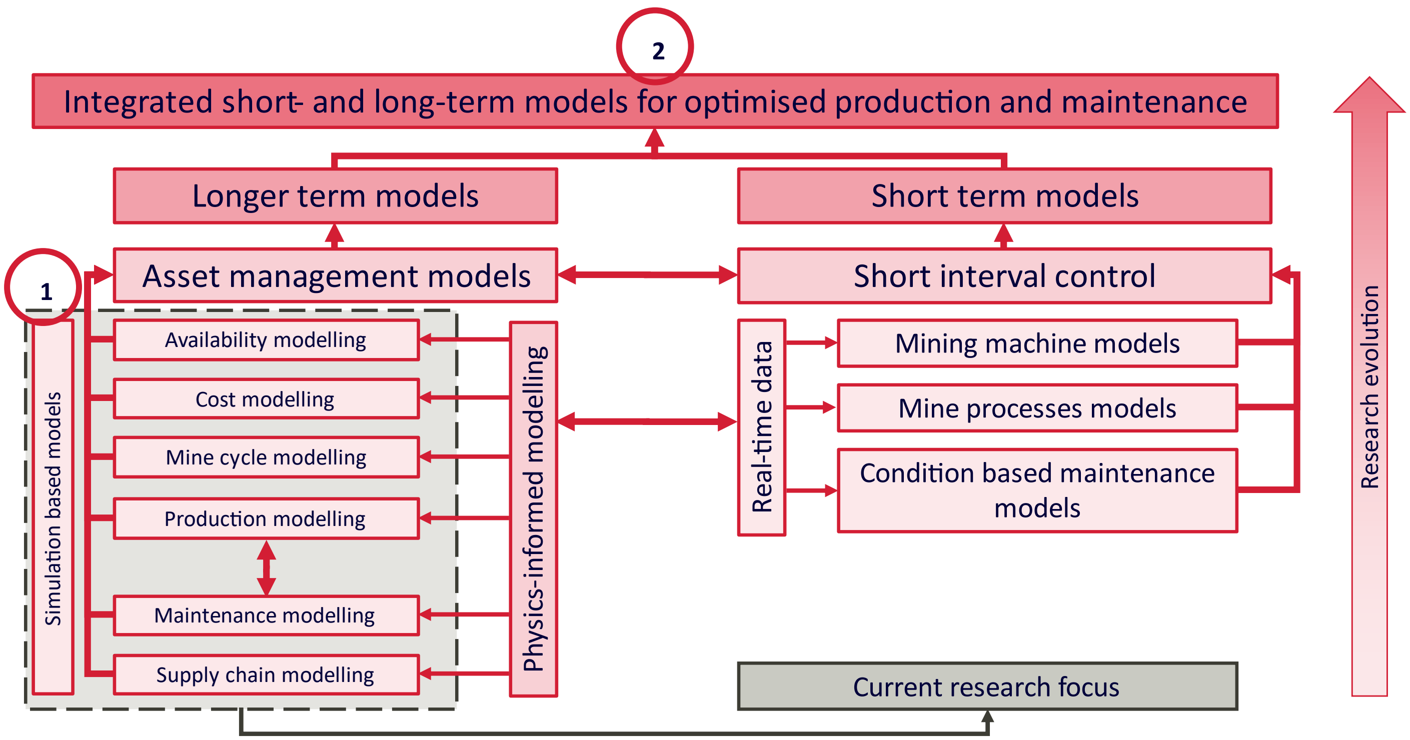


Figure 49 Research roadmap for thrust 2: Productivity and maintenance optimisation of mining equipment. The two labelled circles indicate the two proposed digital twin models for thrust 2.

The goal is to develop forecasting simulation capabilities around key factors such as availability, cost, production, mine cycles, maintenance, and the supply chain. This can significantly affect a mine's total operating cost. These simulation models are crucial for understanding and planning the mine's life.

Mechanisation yields more instrumented machines that produce real-time data (in RTIMS domain), presenting a clear collaboration opportunity between RAMMS and RTIMS programs. Currently, there exists a persisting gap between long-term mine planning and short interval control. This gap may be bridged by utilizing real-time data, tying long-term forecasting models to short-term operational decisions. Fusing enhanced forecasting models with real-time data enables the creation of a continuously updated digital twin. This digital twin not only provides valuable forecasts for future mining but also offers real-time insights into the mine's most optimal path, enabling automated decision-making. Therefore, this thrust aims to develop forecasting capabilities which draw on physical knowledge and real-time data.

The first identified digital twin does not necessarily refer to a single digital twin, but rather a type of digital twin. These types of twins are used to optimise production and maintenance systems within a mine. These models use statistical models or machine-learning models to model the behaviour of mine processes based on historical data. They are mainly useful for simulations, aiding mine production and maintenance staff to optimise their operations. These twins mainly use historical data, and physical knowledge about the process to build simulation models. The outputs of these models are used by humans to inform future decisions in the mine.

1. Production and maintenance planning digital twins	
Physical entity	
Type	Various mining equipment
Counterpart	System or system-of-systems
Data source	Historical data based on the optimisation goal. Generally internal and external software systems.
Digital model	
Type	Data-driven: Statistical or machine learning models
Fidelity	Sufficient fidelity of operations, with individual components usually being simulated as probability distributions.
Domain	Performance and availability
Capability	Simulation
Life cycle phase	Planning and service phase
Physical to digital updating	
Type	Statistical or machine learning model updating
Synchronisation	Periodic
Digital to physical updating	
Type	System reconfiguration and production planning
Acting interface	Human-in-the-loop: Insights from digital twin are only actionable by humans.

The second identified digital twin type is more advanced, and acts in a more pro-active manner, likely on the near real-time scale. These twins use live sensor data throughout the mine to inform simulation models. Therefore, these digital twins can pro-actively adapt to changing mine conditions, rather than relying on historical data to develop probability distributions. For example, instead of modelling travel time from a face to dump site with a probability distribution, live data can be used to better estimate this quantity. These digital twins are likely to output their insights to humans, however, the insights will likely be more useful to operators themselves as opposed to operational centres.

2. Intelligent production or maintenance digital twin	
Physical entity	
Type	Various mine equipment
Counterpart	System or system-of-systems
Data source	• Direct sensor values • Type 1 digital twin simulation model outputs
Digital model	
Type	Data-driven: Statistical or machine learning models
Fidelity	Sufficient fidelity of operations, with individual component states usually being measured directly from sensors.
Domain	Performance and availability
Capability	Simulation, prevention
Life cycle phase	Service phase (live data is required)
Physical to digital updating	

Type	Statistical or machine learning model updating
Synchronisation	Near real-time
Digital to physical updating	
Type	Reconfiguration, production planning, path planning and maintenance scheduling.
Acting interface	Human-in-the-loop: Insights from digital twin are only actionable by humans.

5.3 UTILISATION, PERFORMANCE AND CONDITION MONITORING

The third research thrust centres on the utilization, performance, and condition monitoring of mechanised mining equipment. The primary objective is to understand the workings of specialised mining equipment and improve its efficiency in an underground mining environment. This is particularly critical in the unique context of underground mining where constant connectivity cannot always be guaranteed, making the need for edge decision-making capabilities crucial to mechanised machines. An illustration of the research roadmap, along with three identified digital twin applications are shown in Figure 50.

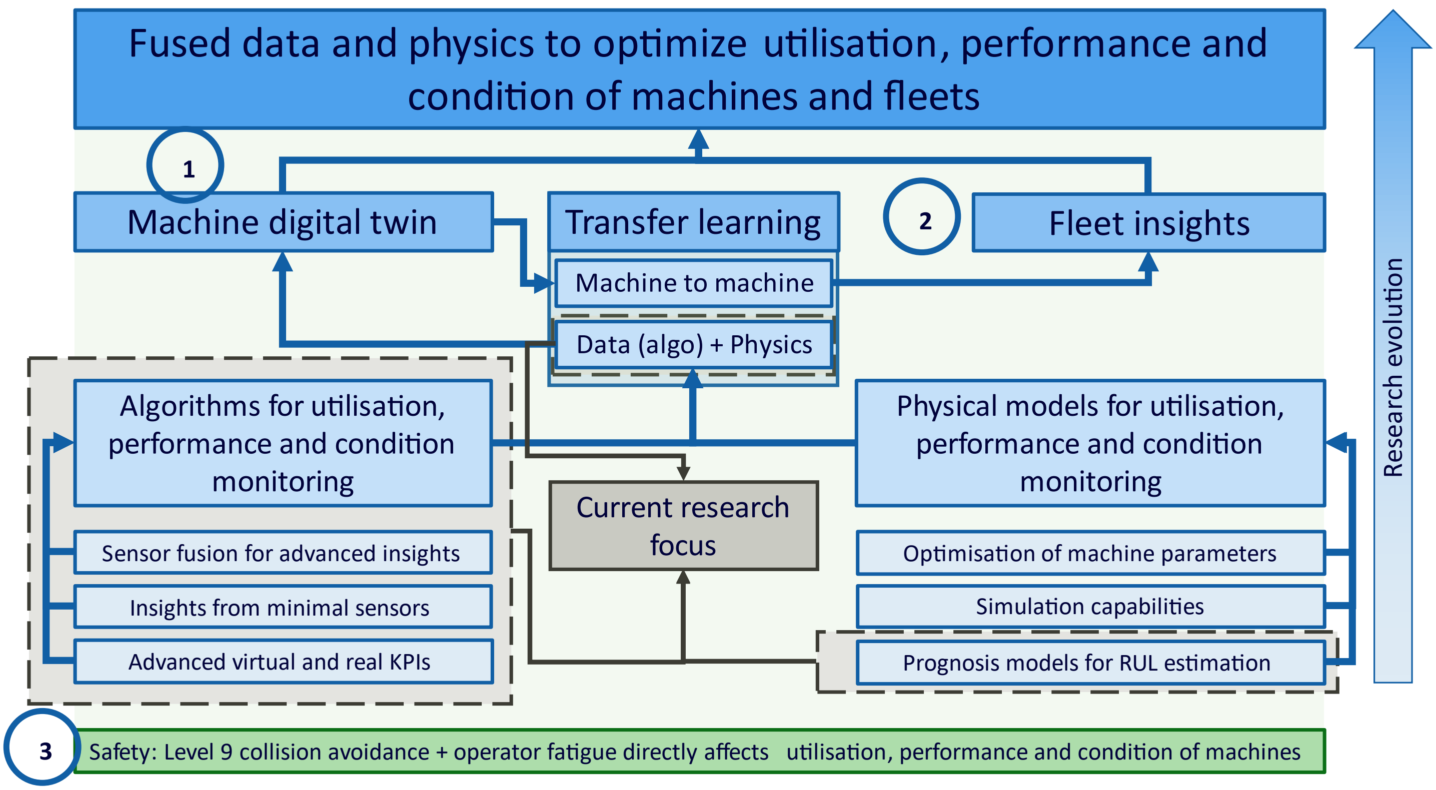


Figure 50 Research roadmap for thrust 3: Utilisation, performance and condition monitoring of mining equipment. The three labelled circles indicate the three proposed digital twin models for thrust 3.

Within this context, the research focuses on the development of algorithms that can aid decision-making on the edge. These algorithms produce leading KPIs, which are immediately actionable and beneficial to the equipment operator. An additional advantage of edge-processing algorithms is their ability to relay only the algorithm outputs to the network when in range, reducing the bandwidth used by each equipment piece. Low-cost algorithms are of interest to network-centric mines, making the development of such technologies impactful.

This thrust also involves developing physics-based models of the machines. These models serve multiple purposes: enriching the edge algorithms, optimizing machinery, and providing forecasting capabilities, such as estimating the remaining useful life of equipment. These physics-driven models can further be enriched by utilizing the on-board real-time data, presenting a clear opportunity to interface with the RTIMS program.

The goal is to combine these algorithms, physics-driven models, and real-time data into digital twins. These twins can optimise not only the performance and utilization of the machine but also estimate the machine's condition and the remaining life of critical components.

Digital twining enables this thrust to extend past the individual machine. The creation of digital twins for multiple machines opens the door to a broader examination of fleet applications. If digital twins are applied across a fleet of machines, one can start to develop algorithms that generate insights from the fleet collectively.

Considering recent developments, it is important to consider the dimension of safety. Safety-related issues have started to significantly impact the performance, utilization, and condition of machines. Level 9 collision avoidance systems, for example, are causing excess braking events, negatively affecting production and causing excess wear on brakes. The operator's fatigue also directly affects the machine's utilization, performance, and condition. Thus, by addressing safety challenges, we can directly improve the performance, utilization, and condition of mining machines. It is, therefore, within the scope to consider projects that enhance safety systems to increase machine performance.

The first identified digital twin, as for thrust 2, does not necessarily refer to a single digital twin, but rather aims to highlight the type of digital twins being built in the left-hand side of Figure 50. This digital twin type uses live sensor data from components and machines (systems) to optimise the utilisation, performance or condition of the physical entity. Safety systems (such as collision avoidance systems) are also considered in this thrust, as they directly impact performance, utilisation and condition of a mechanised machine.

In RAMMS, the focus is to develop hybrid models, which aim to combine physical knowledge of these physical entities with the measured data. Therefore, the digital models in these twins will likely be hybrid models, although physics-driven or data-driven models are certainly possible. These digital twins will output insights that are both actionable by humans or machines at various time scales. For example, gearbox condition monitoring of an LHD will generate diagnostic or prognostic information that needs to be acted upon by humans. However, machine vision for automatic charging will require real-time reaction of the machine. These types of twins are therefore very broad, and it would be impossible to exhaustively outline each possibility. Therefore, we simply summarise the likely digital twin dimensions and properties.

1. Performance, utilisation and condition monitoring digital twins	
Physical entity	
Type	Various mechanised mine equipment (LHD, haul truck, conveyor belts, ventilation fans etc.)
Counterpart	Component or system
Data source	Sensors
Digital model	
Type	Data-driven: Likely machine learning models. Physics-driven: Likely FEM/DEM or kinematic models. Hybrid: All six hybrid methods are likely depending on the problem's available data and known physics.
Fidelity	For performance, utilisation and safety considerations, sufficient fidelity. For condition monitoring, high fidelity.
Domain	Performance (performance and utilisation) and availability (condition monitoring)
Capability	Prediction, detection, prevention and automation.
Life cycle phase	Service phase
Physical to digital updating	
Type	Likely sensor or machine learning updating

Synchronisation	For performance, utilisation and safety considerations, real-time or near real-time. For condition monitoring, periodic.
Digital to physical updating	
Type	Path planning, process control and maintenance scheduling.
Acting interface	For performance, utilisation and safety considerations, likely control systems. For condition monitoring, human-in-the-loop.

The second identified digital twin also refers to a type of digital twin, as potential applications are endless. These digital twins are designed to derive insights beyond a single component or machine, rather extending to a set of components or a fleet of vehicles. For example, a brake wear model, when combined with an LHD load profile, could better predict the wear on the brakes based on the loaded and unloaded states of the LHD. Conversely, brake condition may inform the maximum load an LHD may haul before maintaining the brakes. The insights that are possible from these digital twins are much richer and will add a large amount of value to mining operations. These twins will likely use sensor data, but also draw on insights from the first type of digital twin model. These twins will use physical knowledge of the relationships between components or machines to better model joint behaviour. These twins will operate at similar updating frequencies and fidelities as the first type of digital twins.

2. Multi-component and fleet digital twins	
Physical entity	
Type	Various mechanised mine equipment (LHD, haul truck, conveyor belts, ventilation fans etc.)
Counterpart	System or system-of-systems
Data source	• Sensors • Component/machine digital twin input
Digital model	
Type	Data-driven: Likely machine learning models. Physics-driven: Multi-physics models. Hybrid: All six hybrid methods are likely depending on the problem's available data and known physics.
Fidelity	For performance, utilisation and safety considerations, sufficient fidelity. For condition monitoring, high fidelity.
Domain	Performance (performance and utilisation) and availability (condition monitoring)
Capability	Prediction, detection, prevention and automation.
Life cycle phase	Service phase
Physical to digital updating	
Type	Likely sensor or machine learning updating
Synchronisation	For performance, utilisation and safety considerations, real-time or near real-time. For condition monitoring, periodic.
Digital to physical updating	
Type	Path planning, process control and maintenance scheduling.
Acting interface	For performance, utilisation and safety considerations, likely control systems. For condition monitoring, human-in-the-loop.

The third identified digital twin is centred on safety, specifically identifying an opportunity for collision avoidance systems. South Africa faces unique challenges around collision avoidance, having to solve the complex problem of developing collision avoidance technology among unclear legislation and a lack of standards. Digital twins are likely to be invaluable in South Africa's transition to level 9 collision avoidance systems. To address safety concerns, we identify a digital twin that will utilise machine data as input. This data will be used to infer that state of the machine in real-time. A digital model of the machine can be developed that has predictive capabilities. These predictions can determine safe braking policies depending on the weight of the machine (which may vary for LHDs and haul trucks), current slope, brake wear state etc. These models may also be used to estimate minimum braking distance and hence continually update the operator warning and intervention distances based on the machine state. These insights will likely be implemented at the control level in real-time, however, operator dashboards are also useful, requiring near real-time technology.

3. Collision avoidance digital twins	
Physical entity	
Type	Various trackless mechanised mine equipment (LHDs, haul trucks etc.)
Counterpart	System (a machine)
Data source	• Sensors • Component/machine digital twin input
Digital model	
Type	Hybrid: All six hybrid methods are likely depending on the problem's available data and known physics.
Fidelity	Sufficient fidelity to ensure real-time and near real-time response.
Domain	Performance and availability, as both are impacted by safety.
Capability	Prediction
Life cycle phase	Service phase
Physical to digital updating	
Type	Likely sensor or machine learning updating
Synchronisation	For machine control: Real-time For operator dashboard: Near real-time
Digital to physical updating	
Type	Process control and path planning.
Acting interface	Mostly control systems, but human-in-the-loop applications are possible where operators are involved.

5.4 CONCLUDING REMARKS

The research thrusts within RAMMS have been analysed and multiple digital twinning opportunities have been highlighted. The proposed digital twin framework has significantly simplified this process, allowing clear, commonly understood digital twining projects to be communicated under a common language. These opportunities clearly highlight the types of models that need to be developed, whilst identifying the most important properties that these models need to have. Although these simplified views of potential digital twins do not capture every detail, we believe the most important factors are captured, and the details may be refined on a case-by-case basis.

6 CONCLUSION AND RESEARCH RECOMMENDATIONS

In summary, this report aimed to elucidate the concept of digital twins, which has exhibited variations influenced by temporal factors, lifecycle phases, industry type, and maturity levels. Various frameworks have emerged to encapsulate the essence of digital twins, typically emphasizing the presence of a physical entity, its digital counterpart, and bi-directional data exchange between the two, often coupled with continual optimization.

To address the need for a unified understanding of digital twins, a framework comprising five dimensions and eight properties was proposed, tailored to the South African mining context. These dimensions encompass the physical entity, digital model, updating processes, and optimization. The associated properties delineate aspects such as the twinned entity, data sources, synchronization frequency, fidelity, domain, lifecycle phase, and acting interface. A set of decision diagrams accompany the framework, assisting users in selecting suitable communication and digital models for their twinning applications.

The value of digital twins in mining was underscored through five case studies focusing on mineral extraction within underground mining, demonstrating tangible benefits and alignment with the proposed framework. This exercise highlighted the framework's efficacy in distilling complex digital twins into essential components, enhancing comprehension of the concept.

The framework was further utilized to identify potential digital twinning applications within the context of RAMMS, offering a practical tool for identifying and describing twin projects. The project identification process should further serve as an aid for other users of the framework to comprehensively understand digital twins and identify opportunities within their own diverse domains.

Noteworthy observations during the search for mining case studies highlighted a research gap in extraction-focused literature, signalling an opportunity for RAMMS to contribute significantly to this area. Additionally, the absence of works on drill and blast rock breaking methods underscores the relevance of developing digital twinning capabilities in this domain, given South African mining's continued reliance on such techniques.

Considering these observations, recommendations are made for future RAMMS research to prioritize expanding digital twin capabilities in the extraction phase of mining and to address the scarcity of literature on drilling and blasting digital twins, which holds potential for impactful contributions to both academia and the South African industry.

7 REFERENCES

Autodesk. (2022). *What Is a digital twin and how does it work*. Retrieved March 2024 from https://www.autodesk.com/design-make/articles/what-is-a-digital-twin

Collins, B., & Balshaw, R. (2024, 27 March). *Digital twin framework discussion with Kriterion - Deep digital twins* [Interview]. 121 Boshoff St, Nieuw Muckleneuk,;

Declercq, J., Botha, T., & Hamersma, H. A. (2021). Collision prediction and optimal trajectory generation for collision avoidance systems in trackless mobile machines. *MATEC Web of Conferences*, *347*. https://doi.org/10.1051/matecconf/202134700032

Ghobakhloo, M. (2020). Industry 4.0, digitization, and opportunities for sustainability. *Journal of Cleaner Production*, *252*. https://doi.org/10.1016/j.jclepro.2019.119869

Glaessgen, E. H., & Stargel, D. S. (2019). *The digital twin paradigm for future NASA and US air force vehicles* 53rd Structures, Structural Dynamics, and Materials Conference, https://doi.org/https://doi.org/10.2514/6.2012-1818

Grieves, M. (2014). *Digital twin: Manufacturing excellence through virtual factory replication*.

Grieves, M., & Vickers, J. (2017). Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. In *Transdisciplinary Perspectives on Complex Systems* (pp. 85-113). https://doi.org/10.1007/978-3-319-38756-7_4

Groeneveld, B., Topal, E., & Leenders, B. (2019). Examining system configuration in an open pit mine design. *Resources Policy*, *63*. https://doi.org/10.1016/j.resourpol.2019.101438

Guan, M., Wen, C., Wei, Z., Ng, C.-L., & Zou, Y. (2018). A dynamic window approach with collision suppression cone for avoidance of moving obstacles. 2018 IEEE 16th International Conference on Industrial Informatics (INDIN), https://doi.org/10.1109/INDIN.2018.8472029

Huang, Y., Yuan, B., Xu, S., & Han, T. (2022). Fault Diagnosis of Permanent Magnet Synchronous Motor of Coal Mine Belt Conveyor Based on Digital Twin and ISSA-RF. *Processes*, *10*(9). https://doi.org/10.3390/pr10091679

Jia, F., Lei, Y., Lin, J., Zhou, X., & Lu, N. (2016). Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data. *Mechanical Systems and Signal Processing*, *72-73*, 303-315. https://doi.org/10.1016/j.ymssp.2015.10.025

Jiao, X., Xie, J., Wang, X., Yan, Z., Hao, Z., & Wang, X. (2022). Intelligent decision method for the position and attitude self-adjustment of hydraulic support groups driven by a digital twin system. *Measurement*, *202*. https://doi.org/10.1016/j.measurement.2022.111722

Juarez, M. G., Botti, V. J., & Giret, A. S. (2021). Digital twins: Review and challenges. *Journal of Computing and Information Science in Engineering*, *21*(3). https://doi.org/10.1115/1.4050244

Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L. (2021). Physics-informed machine learning. *Nature Reviews Physics*, *3*(6), 422-440. https://doi.org/10.1038/s42254-021-00314-5

Khalid, S., Song, J., Raouf, I., & Kim, H. S. (2023). Advances in Fault Detection and Diagnosis for Thermal Power Plants: A Review of Intelligent Techniques. *Mathematics*, *11*(8). https://doi.org/10.3390/math11081767

Kritzinger, W., Karner, M., Traar, G., Henjes, J., & Sihn, W. (2018). Digital twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnLine*, *51*(11). https://doi.org/info:doi/10.1016/j.ifacol.2018.08.474

Kychkin, A., & Nikolaev, A. (2020, 2020). IoT-based mine ventilation control system architecture with digital twin. International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), https://doi.org/10.1109/icieam48468.2020.9111995

Lei, Y., Yang, B., Jiang, X., Jia, F., Li, N., & Nandi, A. K. (2020). Applications of machine learning to machine fault diagnosis: A review and roadmap. *Mechanical Systems and Signal Processing*, *138*. https://doi.org/10.1016/j.ymssp.2019.106587

Macchi, M., Ragazzini, L., & Negri, E. (2023). A conceptual framework for Digital twins in production scheduling and control. *IFAC-PapersOnLine*, *56*(2). https://doi.org/info:doi/10.1016/j.ifacol.2023.10.491

Macchi, M., Roda, I., Negri, E., & Fumagalli, L. (2018). *Exploring the role of digital twin for asset lifecycle management* IFAC (International Federation of Automatic Control) PapersOnLine, https://doi.org/10.1016/j.ifacol.2018.08.415

Mo, F., Rehman, H. U., Monetti, F. M., Chaplin, J. C., Sanderson, D., Popov, A., Maffei, A., & Ratchev, S. (2023). A framework for manufacturing system reconfiguration and optimisation utilising digital twins and modular artificial intelligence. *Robotics and Computer-Integrated Manufacturing*, *82*. https://doi.org/10.1016/j.rcim.2022.102524

Patil, S. D., Mitra, A., Tuggali Katarikonda, K., & Wansink, J.-D. (2021). Predictive asset availability optimization for underground trucks and loaders in the mining industry. *Opsearch*, *58*(3), 751-772. https://doi.org/10.1007/s12597-020-00502-4

PwC. (2023). *Ten insights into 4IR 2023* (The state of digital transformation in the South African mining industry, Issue.

Rafai, A. N. A., Adzhar, N., Jaini, N. I., & Ding, B. (2022). A review on path planning and obstacle avoidance algorithms for autonomous mobile robots. *Journal of Robotics*, *2022*, 1-14. https://doi.org/10.1155/2022/2538220

Rubio-Sierra, C., Dominguez, D., Gonzalo, J., & Escapa, A. (2020). Path planner for autonomous exploration of underground mines by aerial vehicles. *Sensors (Basel)*, *20*(15). https://doi.org/10.3390/s20154259

Savolainen, J., & Urbani, M. (2021). Maintenance optimization for a multi-unit system with digital twin simulation. *Journal of Intelligent Manufacturing*, *32*(7), 1953-1973. https://doi.org/10.1007/s10845-021-01740-z

Schweiger, L., & Barth, L. (2023). Properties and characteristics of digital twins: Review of industrial definitions. *SN Computer Science*, *4*(5). https://doi.org/10.1007/s42979-023-01937-4

Sjarov, M., Lechler, T., Fuchs, J., Brossog, M., Selmaier, A., Faltus, F., Donhauser, T., & Franke, J. (2020, 2020). The digital twin concept in industry – A review and systematization. https://doi.org/10.1109/etfa46521.2020.9212089

Somers, R. J., Douthwaite, J. A., Wagg, D. J., Walkinshaw, N., & Hierons, R. M. (2023). Digital-twin-based testing for cyber–physical systems: A systematic literature review. *Information and Software Technology*, *156*. https://doi.org/10.1016/j.infsof.2022.107145

Stark, R., & Damerau, T. (2019). Digital twin. In *CIRP Encyclopedia of Production Engineering* (pp. 1-8). https://doi.org/10.1007/978-3-642-35950-7_16870-1

Tao, F., Cheng, J., Qi, Q., Zhang, M., Zhang, H., & Sui, F. (2017). Digital twin-driven product design, manufacturing and service with big data. *The International Journal of Advanced Manufacturing Technology*, *94*(9-12), 3563-3576. https://doi.org/10.1007/s00170-017-0233-1

Tao, F., Zhang, H., Liu, A., & Nee, A. Y. C. (2019). Digital twin in industry: State-of-the-art. *IEEE Transactions on Industrial Informatics*, *15*(4), 2405-2415. https://doi.org/10.1109/tii.2018.2873186

Tao, F., Zhang, M., Liu, Y., & Nee, A. Y. C. (2018). Digital twin driven prognostics and health management for complex equipment. *CIRP Annals*, *67*(1), 169-172. https://doi.org/10.1016/j.cirp.2018.04.055

Thelen, A., Zhang, X., Fink, O., Lu, Y., Ghosh, S., Youn, B. D., Todd, M. D., Mahadevan, S., Hu, C., & Hu, Z. (2022a). A comprehensive review of digital twin - Part 2: Roles of uncertainty quantification and optimization, a battery digital twin, and perspectives. *Structural and Multidisciplinary Optimization*, *66*(1). https://doi.org/10.1007/s00158-022-03410-x

Thelen, A., Zhang, X., Fink, O., Lu, Y., Ghosh, S., Youn, B. D., Todd, M. D., Mahadevan, S., Hu, C., & Hu, Z. (2022b). A comprehensive review of digital twin - Part 1: Modeling and twinning enabling technologies. *Structural and Multidisciplinary Optimization*, *65*(12). https://doi.org/10.1007/s00158-022-03425-4

Vogt, D. (2016). A review of rock cutting for underground mining: past, present, and future. *Journal of the Southern African Institute of Mining and Metallurgy*, *116*(11), 1011-1026. https://doi.org/10.17159/2411-9717/2016/v116n11a3

Wenna, W., Weili, D., Changchun, H., Heng, Z., Haibing, F., & Yao, Y. (2022). A digital twin for 3D path planning of large-span curved-arm gantry robot. *Robotics and Computer-Integrated Manufacturing*, *76*. https://doi.org/10.1016/j.rcim.2022.102330

Wu, H., Ji, P., Ma, H., & Xing, L. (2023). A comprehensive review of digital twin from the perspective of total process: Data, models, networks and applications. *Sensors (Basel)*, *23*(19). https://doi.org/10.3390/s23198306

Wu, J., Yang, Y., Cheng, X. U. N., Zuo, H., & Cheng, Z. (2020). *The development of digital twin technology review* 2020 Chinese Automation Congress (CAC), https://doi.org/10.1109/CAC51589.2020.9327756

Yao, J. F., Yang, Y., Wang, X. C., & Zhang, X. P. (2023). Systematic review of digital twin technology and applications. *Vis Comput Ind Biomed Art*, *6*(1), 10. https://doi.org/10.1186/s42492-023-00137-4

Yassin, M. A. M., Shrestha, A., & Rabie, S. (2023). Digital twin in power system research and development: Principle, scope, and challenges. *Energy Reviews*, *2*(3). https://doi.org/10.1016/j.enrev.2023.100039

Zhang, J., Wan, C., Wang, J., Chen, C., Wang, T., Zhang, R., & Guo, H. (2024). Research on the "shape-performance-control" integrated digital twin system for boom-type roadheaders. *Sci Rep*, *14*(1), 5780. https://doi.org/10.1038/s41598-024-56539-8