Digital Twin in Aerospace Industry:
A Gentle Introduction

LUNING LI, SOHAIB ASLAM, ANDREW WILEMAN, AND SURESH PERINPANAYAGAM
IVHM Centre, Cranfield University, Bedfordshire MK43 0AL, U.K.
Corresponding author: Luning Li (luning.li@cranfield.ac.uk)

ABSTRACT Digital twin (DT), primarily a virtual replica of any conceivable physical entity, is a highly transformative technology with profound implications. Whether it be product development, design optimisation, performance improvement, or predictive maintenance, digital twins are changing the ways work is undertaken in various industries with multifarious business applications. Aerospace industry, including its manufacturing base, is one such keen adopter of digital twins with an unprecedented interest in their bespoke design, development, and implementation across wider operations and critical functions. This, however, comes with some misconceptions about the digital twin technology and lack of understanding with respect to its optimal implementation. For instance, equating a digital twin to an intelligent model while ignoring the essential components of data acquisition and visualisation, misleads the creators into building digital shadow or digital models, instead of the actual digital twin. This paper unfolds such intricacies of digital twin technology for the aerospace community in particular and others in general so as to remove the fallacies that affect their effective realisation for safety-critical systems. It comprises a comprehensive survey of digital twins and their constituent elements. Elaborating their characteristic state-of-the-art composition along with corresponding limitations, three dimensions of the future digital twins for the aerospace sector, termed as aero-Digital Twins (aero-DTs), are proposed as an outcome of this survey. These include the interactive, standardisation, and cognitive dimensions of digital twins, which if leveraged diligently could help the aero-DT research and development community quadruple the efficiency of existing and future aerospace systems as well as their associated processes.

INDEX TERMS Digital twins, aircraft operation and maintenance, aerospace manufacturing.

I. INTRODUCTION
The aviation industry is regarded as technology-intensive by virtue of the gradual digital transformation that has taken place over several decades within the entire aerospace industry. This has helped aircraft systems become safer and more efficient [1]. However, the integration of digital avionic systems, vehicle health management, and sensors within the aircraft has also increased complexity in terms of system configuration, maintainability, and data enormity [2]. This implies the need for effective, rapid and accurate data management and analysis to ensure sustainable reliability and safety of the aircraft platform over its complete lifecycle [3].

A digital twin, herein referred to as DT, is an emerging technology, which can provide a real-time, high-fidelity virtual model for its aviation counterparts. With an ability to collect, collate, store, analyse and feedback data, a DT can be made to provide continuous evaluation of its physical entity. Since the sensors can capture and continually update the system’s DT throughout its operational life, aircraft manufacturers and operators can hold a live window inside the physical system at all times [4]. Gradually, but assuredly, DTs and their technologies are being applied to space systems, UAV, military and commercial aviation, encompassing functions from manufacturing to anomaly detection, and asset as well as fleet management. The implications of DT implementation are profound. For instance, it is possible to carry out real-time system assessments, diagnostics and prognostics more precisely than with traditional health management methods. Repairs could be executed instantly, and innovation could be much faster, economical, and more disruptive.

DTs have shown significant potential value (as discussed in Section 3) in asset-intensive industries, triggering an upsurge in DT R&D. The upward trend of the Aero-DT is attributable to advancements in the technologies and processes related to AI, Big Data, cloud and IoT [2]. From academia, there...
were initially less than ten DT-related articles published before 2015. Since 2017, the number of DT-related journal papers has risen to over hundred in 2020 [5]–[7]. Whereas, from the perspective of aerospace industries (including aerospace manufacturing), the number of articles published from 2017 to 2020 has doubled each year, fuelling great interest in DTs among the aerospace community. NASA and the U.S. Air Force dominated the early research direction of DTs, and have been predominant in aero-DT based maintenance and applicability [8]. The application of DTs has been widely discussed in terms of airframe, avionics, crack detection, and fleet level health management, whereas aerospace OEMs, e.g. GE, Boeing and Airbus, have included DTs in their future layout strategy [2].

This, however, comes with issues such as imprudent research and application, misapplying the DT concept, insufficient understanding and over-interpretation of the DTs that eventually place constraints on the development of highly efficacious DTs for aircraft systems. For example, equating a DT to an intelligent model while ignoring data collection and visualisation leads to the creation of a digital shadow or digital model instead of a DT. This is attributable to the lack of knowledge and clarity about the enabling tools or components of DTs, thereby demanding a comprehensive survey and discussion of the state-of-the-art research, encompassing the architecture and enabling technologies of aero-DT (such as critiquing the available modelling approaches and highlighting their pros and cons for aero-DTs).

Accordingly, this paper provides a detailed survey of DTs, from their conceptual, value-rendering, and technology-enabling perspectives, with the following main contributions:

1) An elaborated account of DT development with impact and implementation across several key industries, such as manufacturing (Industry 4.0), healthcare, and smart cities. This helps in mapping various DT attributes to aerospace systems’ environment and assessing their limitations and implementation challenges.

2) Highlighting technical infirmities that are hampering the smooth adoption and optimal implementation of DTs across the wider aerospace industry. Issues and challenges pertaining to real-time data collection, synchronisation, and processing for high-fidelity DTs have been put forth to inform aero-DT developers and implementers of various opportunities that can be carved out to further improve DT modelling, visualisation, and infrastructure.

3) A roadmap for steering the existing aero-DT technologies and associated processes into an era of DTs that are standardised, highly interactive, and exhibit cognitive capabilities to expand technological and business horizons of aerospace applications.

The rest of this paper is organised as follows. Section 2 provides a detailed account of the concept and composition of DTs so as to remove the scepticism that surrounds them. Section 3 starts with a brief historical account of the origin and evolution of DTs, and then discusses the significance of aero-DT for the aerospace and its manufacturing industry. Section 4 provides an in-depth analysis of DTs from a modelling, visualisation and infrastructure perspective. Post analytical survey of key enablers in DT development. Section 5 presents a roadmap for aero-DT researchers and developers. Finally, various significant conclusions are drawn in Section 6.

II. INTRODUCTION TO DIGITAL TWINS: CONCEPT AND COMPOSITION

A. CONCEPT

With the ongoing expansion of the DT concepts, the way academia looks at DTs from different angles results in the DT definition being continuously enriched. This section attempts to discuss the definitions of DTs from different angles so as to help readers distinguish DTs from other similar concepts. Initially, DTs are recognised as an information system. The motivation of creating DTs is to develop a life-long asset information visualising technique for asset-intensive industries. This is why a DT is described, in many research articles, as the virtual information integration of a physical asset. For instance, Professor Grieves [9], the first person to put forward the DT, defined the DT as:

- “The Digital Twin is a set of virtual information constructs that fully describes a potential or actual physical manufactured product from the micro atomic level to the macro geometrical level. At its optimum, any information that could be obtained from inspecting a physical manufactured product can be obtained from its Digital Twin.” [9]

There is no strict definition of the scale, property and complexity of physical assets, so it can be a single component, a system of components, or a system of systems, for example, pump, engine, human body, power plant or even a city. The virtual replica could reflect every information that can be obtained from the physical entity, which is recognised as the basic function of a DT. From this angle, we understand the basic elements of a DT, i.e. a ‘physical entity’, a ‘virtual model’, and a description of the connection between the two.

With a gradual progression of research on different perspectives of DTs, they are expected to present processed and valuable information rather than high-volume and low-value unprocessed data. Also, the DT community realises the important role of the analytical model for information processing. This is evident from many research articles that highlight the information analysis capacity of DTs. For example, NASA [10] defined the DT as:

- “A Digital Twin is 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.” [10]

From this angle, DTs not only emphasise building end-to-end connections, but how to process information during the end-to-end process.
Following the enrichment of analytic models, some scholars point out the property of DTs as an integration of models with multiple functions and multiple scales. These models jointly describe the behaviour of physical entities through internal connections or rules. For example, [11] defined the DT as:

- “Newer increases in the computational power that is easily and widely available have enabled more complex simulations that integrate previously separate models of various aspects of structural design in order to accurately simulate the behavior of the system as whole. These ultra-high-fidelity simulations are commonly called a digital twin with respect to the system they model.” [11]

In a relatively complex system, single function models cannot meet the need for information processing. Instead, multi-functional, multi-scale, and inner-regional models collaborating with each other and following certain rules are essential to realising DTs.

Combining the definitions above, Figure 1 shows the position of DTs from different angles. In summary, a DT is a high-fidelity and up-to-date representation of an actual physical asset in operation that reflects the current asset condition and includes relevant historical data about the asset. DTs can be used to evaluate the current condition of the asset, and more importantly, predict future behaviour, refine the control, and optimise operation.

**B. MAJOR COMPONENTS OF DTs**

Since the 2000s, the available components of DTs have been continuously expanded along with the innovation of technologies. Some novel technologies enable DTs to be applied in new domains, such as biochips for human body DTs [12], [13], 5G and edge computing in aircraft DTs [14]. The major components of DTs can be divided into three categories based on the function of components:

1) PHYSICAL SIDE: SENSORS AND FUNCTIONAL INFRASTRUCTURE

Components on the physical side are mainly to support data collection and computing. Sensors are the ‘must-have’ components in every form of DT. Coupled with data transmission technologies, they guarantee that DTs can realise real-time data collection and synchronisation. Various sensors enrich the types of data that can be collected, from text, audio, hyper-spectral images, video, temperature, pressure etc. to behaviour and biological characteristics. A DT, hence, can prove more efficient and realistically reflect the dynamic state of its counterpart. On the other hand, high performance computing hardware is necessary. HPC (high-performance computing) has now become a ‘must-have’ infrastructure for enterprises that have already been successful in realising digitalization, with high computing capacity ensuring the success of big data processing and analysis. Similarly, cloud computing is an emerging technology and a ‘good-to-have’ infrastructure to meet requirements pertaining to growing complexity and scale of DTs. Cloud computing can break up the limitations of local computing devices whilst improving the overall computing capacity. Additionally, the physical side should have the functional infrastructure to enable and perform specific functions (e.g. immersive sim for VR, actuator for control, etc.) depending on the requirements of DT customers.

Compared with the virtual side, the enabling technologies on the physical side are relatively mature. Nowadays a number of hardware suppliers are committed to providing various physical side solutions. For example, IBM, GE, Siemens and Oracle can provide a complete set of DT hardware solutions. Companies such as Amazon Web Services (AWS), Alibaba Cloud and Microsoft Azure provide IoT suites for IoT and SAP cloud platforms as well as for handling Big data.

2) VIRTUAL SIDE: ANALYTICAL MODELS AND AI

The main function of the virtual side is to gather, process and analyse the data. In general, the virtual side is made up of sets of models with different functions. The model is mainly divided into two forms: traditional modelling methods
which are mostly physics-based, dominating most of the earlier practical cases of DTs and data-driven modelling. Since the 2010s, with the continuous improvement of computing power, developing data-driven models has become more efficient and in line with the trend of Big Data and IoT. The other reason for the popularity of data-driven models is adaptability. Current techniques in machine learning (ML) allow engineers to model in a relatively short time without experimentation and prior knowledge. By contrast, physics-based models, including experimental models, multi-dimensional models and high fidelity numerical models, have certain requirements on engineers’ professional and mathematical knowledge. There is no consensus of which modelling approach is better (see details in Section 4) because the choice of model also depends on the expected function of DTs and heterogeneous data whilst considering factors such as co-operation and compatibility between models.

Although the application of AI is quite common in DT related research, AI is actually not a ‘must-have’ component. However, AI is the key to making DTs smart and automatic. Most AI applications on the virtual side reflect on using machine learning to create the data-driven analytical model. These models are widely applied in diagnosis, manufacturing, and decision-making due to their proven performance in classification, clustering, and generation tasks.

3) CONNECTION: DATA TRANSMISSION AND HUMAN-MACHINE INTERFACE

The essence of connections lies in the data communication technologies. In this era of Big Data and IoT, communication technologies are becoming abundant, such as 5G and LoRa (Long Range – low power wide area network modulation technique) for large-scale networks, Wi-Fi for small and medium workshops, and SatCom for aircraft data transmission, etc. Users can choose according to the specific connection requirement.

In any kind of DT, there are generally three types of connections. The first is the connection between the physical entity and the virtual side, which, more specifically, is from sensors to service models. It is the basis and necessary connection for realising ‘twins’. The second is the connection between the virtual side to humans. This is about using the visualisation techniques (e.g. VR, AR and 3D simulation models) to present virtual information to its operator. This connection is the foundation for realising high-fidelity models or human-machine interaction. However, it is not necessary (e.g. high-fidelity models are usually ignored in some simple DTs). The last is the connection from the human to the physical side. It involves conveying operator commands to the physical entities through, for example, controllers or human-behaviour capturing devices. In most diagnosis and prediction purposed DTs, this connection is not necessary.

Nowadays, DTs are becoming more flexible. As the ‘must-have’ components continue to mature and costs continue to decrease, the options for components and enabling technologies will grow. For different domains, available DT technique combinations are continuously explored and it will become more and more difficult to accurately define the standard components of DTs.

C. THE APPLICATION OF DIGITAL TWINS IN INDUSTRIES

Over the past decade, digital twins have been widely used in many industries. These industries, such as automotive surgery and aero-engine manufacturing, often have a high requirement for asset controllability and reliability. They rely on the DT providing a high-fidelity model to improve the visibility and transparency of assets. There are also industries that need large-scale data and asset management, such as urban or factory management. Digital twins could provide an integrated model with data processing and analysis functions in these applications. As frontline digital twin applications, three areas have been widely discussed in literature, namely smart cities, healthcare and smart manufacturing.

1) SMART CITIES

In smart cities, DTs are the bridge between the physical city and virtual worlds, which maps an original two-dimensional urban information system to three dimensions, and static urban information to dynamic. Figure 3 describes how digital twins work in the context of smart cities. Ideally, the digital twin for a smart city could simulate the entire urban system in order to co-ordinate and manage the city as a whole. It centralises the urban information for integrated urban management while keeping a certain decentralising autonomy for each urban sub-system. At this stage, DTs used in the urban sub-systems are quite mature. For example, Siemens has practised a water treatment digital twin in the Middle East’s largest desalination plant. [15] presents an urban road and traffic DT and British Petroleum has applied digital twins for the monitoring and maintenance of their oil and gas facilities [16]. By contrast, an entire city-level DT is still in the exploration stage; the challenges are not only on the technical side, such as how to collect and process various heterogeneous city information but also in the consideration of operating costs, government support, transition plans, and other non-technical issues. Some researchers have described
the conceptual architecture of the city-level DT, including the Digital geoTwin Vienna proposed in [17] and the DT for west Cambridge campus in [18].

2) HEALTHCARE

Digital twins in healthcare is a relatively-novel concept. The emergence of healthcare DTs corresponds with the development of smart wearable devices and biosensors, which enable data collection from human bodies. The value of healthcare DTs is mainly embodied in improving visibility by creating high-fidelity digital bodies so as to support healthcare monitoring, digital surgery, remote surgical assistance, etc. (as shown in Figure 4). [19], [20] believe that once the digital twin is realised, people will observe the changes of the body more intuitively, which may change the human understanding of the human body, medical treatment and surgery. For example, [13] uses body-on-a-chip (BOC) to understand the effects of certain drugs on patients. [21]–[23] apply VR technologies to synchronise surgical procedures so as to realise digital surgery.

3) MANUFACTURING

DTs were originally conceived and created for manufacturing. The primary motivation of creating DTs, according to the founder Professor Grieves, is to improve the product life cycle management in manufacturing industries. Nowadays, DTs are widely applied in design, manufacturing process and assembly to collaborate with or replace original CAD/CAE/CAM tools. DTs and some similar concepts, such as cyber physical systems (CPS), are the inevitable post-digitalisation trend of manufacturing industries. From a historical perspective, digitalisation has gradually moved manufacturing to the intelligent manufacturing era with practices like ‘information centralisation’ and ‘regional autonomy’. This explains why DTs have fast become the cornerstone of Industry 4.0 (also known as smart manufacturing and Factory of the Future). A number of established, prolific OEMs and technical suppliers have joined the DT research and innovation platform, such as Airbus and Boeing [24], GE [25] and IBM [26], along with newer companies, such as Tesla, making significant progress in the application of DTs. They have created DTs for each product to collect individual data generated from manufacturing process to in-service life cycle, and then use the data to feedback the new product development process, manufacturing, after-sales service, and maintenance [27], [28].

III. DIGITAL TWIN IN AEROSPACE: HISTORY AND VALUE

A. THE HISTORY OF DIGITAL TWINS IN THE AEROSPACE INDUSTRY

1) 1970s. PHYSICAL TWINS IN AEROSPACE

The first use of the ‘twin’ concept (or at least a similar technology) can be traced to 1970 when NASA launched Apollo 13. NASA built an almost identical Apollo 13 physical model on the ground [2]. The initial intention of NASA was to find an approach to monitor the spacecraft operation, manage risks and respond to emergencies, thereby reducing the burden on astronauts. Although the establishment of a physical twin is extremely costly, this approach has undoubtedly proven to be effective and successful during the Apollo 13 space mission. Figure 5 shows the Apollo’s ‘physical twin’ in NASA’s workshop. After the Apollo 13 mission, the ‘physical twin’ was still adopted continuously in many space programmes for many years.

Physical twins, from the current point of view, have a limited reference value for the profit-oriented manufacturing industry or civil aviation industry due to its construction process and cost. However, the early physical twin did show the value of the twin concept in health management, diagnosis and prognosis.

2) 2000s. THE EMERGENCE OF DIGITAL TWINS

In 2002, Professor Michael Grieves from University of Michigan proposed an idealised lifecycle management approach for intelligent manufacturing and future factories.
in his presentation and named it “Digital Twin” for the first time. However, the concept of ‘digital twin’ was very rarely mentioned specifically in research between 2002 and 2010 [6]. However, another similar concept, ‘Digital Shadow’ has been highlighted. Currently, Digital Shadow is viewed as more akin to a prototype version of a DT with limited and immature communication technologies. The Digital Shadow only has one-way data flow, from physical entity to the virtual side, whilst not realising the closed-loop of data which involves using virtual data to affect the physical entity [29].

3) 2010-2014. THE INCUBATION OF DIGITAL TWINS
In the 2010s, with the maturity and industrialisation of IoT and big data technologies, the DT was ready to be applied in practice. NASA and the U.S. Air Force made outstanding contributions to the R&D of aerospace DTs (aero-DTs) in this period. In 2010, NASA first defined DTs in an aerospace context and laid out a roadmap for DT development [30] clarifying the strategical value of DTs for both U.S. space science and the Air Force, whilst NASA set the goal of developing adaptive and full-mission spacecraft DTs by 2035. The U.S. Air Force has contributed a series of novel research outcomes of DTs in terms of feasibility analysis, fleet management and in-flight diagnosis and prognosis [8].

Meanwhile, in the 2010s, DT-driven smart manufacturing has become a popular direction in Industry 4.0 [31]. Large-scale aviation OEMs, such as Boeing, Airbus and GE Company, started to develop their own DT programmes. The world leading aviation OEMs expect that DTs can dynamically optimise design manufacturing processes to further improve the products’ quality and reliability while reducing cost and saving time.

4) 2015-PRESENT. THE BURGEON OF DIGITAL TWINS
From 2015 onwards, DT research has surged. According to the statistics of [6], published articles with the topic of DTs have grown by more than 10 per year. In the aerospace industry, DTs have been widely applied in space science, security and defence, commercial aircraft, and aerospace manufacturing from design to the product launch. DT research for aerospace covers the single component to system levels and even fleet levels [8], [32]. DTs and their relevant technologies have greatly boosted the aerospace industry in manufacturing, operation & maintenance (O&M).

B. THE VALUE OF DIGITAL TWINS FOR THE AEROSPACE INDUSTRY
Due to the extremely long period of the aerospace product lifecycle (more than 40 years—including production, manufacturing, and in-service), the entire aerospace industry is committed to improving product lifecycle management for decades. Even so, the cost and profit issues from the production and service cycles are still placing unprecedented pressure on the industry under the gamut of innovation, international competition and risk management. For example, Figure 6 from [33] describes the circumstances of a New Product Development (NPD) process that the aerospace industry faces. Compared with the automotive and integrated circuits industries, the NPD period of the aerospace systems continues to grow with the increasing product complexity, which forms an opposite trend with the others. On the other hand, both aviation manufacturers and commercial aviation companies are sensitive to uncertainties impacting the global economy and various emergencies (e.g. the COVID-19 pandemic) that hit them both financially and technically. This necessitates a careful management approach to ensure the overall efficiency of manufacturing processes and/or aircraft operations. For example, optimising Maintenance, Repair and Overhaul (MRO) and further improvement to aircraft reliability in order to reduce unscheduled maintenance, improve the scheduled maintenance efficiency and reduce per-flight costs [33], [34].

Because of the urgent need for the entire aviation industry to adapt effective life cycle management tools, the DT is able to demonstrate its value to the aerospace industry. In the entire life cycle management, DT provides an iterative closed-loop process that integrates all the nodes of aerospace products from design, manufacturing, to O&M until retirement (see Figure 7). Compared to the traditional PLM method, DTs have brought the following changes to the industry.
1) DTs FOR AEROSPACE DATA FLOW: THE NEW GENERATION OF DATA CARRIER

Data collection and storage is a crucial function of DTs. Initial misunderstandings within research cite DTs as a specific data collection tool. In fact, DTs are more likely to be the carrier of data flow. As shown in Figure 8, DTs provide two-dimensional support for data flow. Horizontally, DTs provide lifelong data support across every node of the product including product updates and iterations. Vertically, for each node, DTs further provide data mining and analysis functions on the basis of collecting and storing data. Moreover, the DT itself can also generate data based on the results from product simulation, diagnosis and prediction. Together, these data and product data collectively form a bi-directional data flow to help product optimisation and decision-making [35].

FIGURE 8. Role of DTs in data science.

FIGURE 9. Role of DT in NPD process [36].

2) DTs IN SMART MANUFACTURING: THE ENABLER OF NPD EVOLUTION

The overall value of DTs for NPD is to further improve the flexibility and efficiency while reducing the overall cost and period of the NPD process. Since the 1990s, when Boeing first employed CAD/CAE/CAM in its Boeing-777 NPD process [37], the aerospace NPD has undergone an evolution from traditional drawing- and machine-dominated platforms to digitalization. DTs, in the NPD process, can be seen as the upgraded integration of CAD, CAE and CAM. DTs has been proven its effectiveness in many application cases, such as, aircraft engine design [38]–[41], CNC and robot arm aided machine process [42]–[45] and final assembly [46]–[49]. Moreover, by virtue of DTs’ information integration and high-fidelity simulation characteristics, digital factory twins are also on the R&D agenda. The factory twin not only dynamically simulates the factory layout, production line and track job progress [50], [51] but also provides market analysis and demand prediction [52], [53], supply chain management [54], thereby bridging the management system to the production system dynamically [46].

Compared to the prevalent CAD/CAE/CAM methodology, DTs can improve the products’ quality the very first time and accelerate decision-making by high-fidelity presentation of valuable data to the customer. Figure 9 shows how DTs can benefit and improve the NPD cycle in the virtual space through constant data iterations, thereby optimising and verifying the previous stages [36]. According to the GE company, DTs can effectively reduce the NPD cycle by 10%-75% [55]. DTs also allow the originally expensive and time-consuming physical tests (e.g. material test, wind tunnel test, flight test and ground test) to move to the virtual space. [56] pointed out that the application of DTs has reduced the time originally spent on aerospace material testing and verification by 80% and 25% respectively.

A popular topic of discussion is how technologies of manufacturing processes have benefited since the emergence of the Cyber-Physical System (CPS) and Hardware-in-the-Loop (HiL) in the manufacturing industry. DT, CPS and HiL are similar in many ways, as they all rely on the tetra-drivers of innovation (i.e. IoT, Big data, cloud and AI) and all create a virtual model for the physical entity in order to support the manufacturing process. However, CPS and HiL attend more to machine process control, while DT focuses on information management and process, and it is unlikely that CPS and HiL are used at the design or test stage. Therefore, the application of DT is broader, but also relatively difficult to achieve [57]. In this case, there is interest in research on combining DTs and CPS [58], [59] so as to create an integrated system. In the integrated system, CPS is in charge of the machine control and links cyberspace with the digital production line, while DTs are used to improve the manufacturing process visualisation and data collection, storage and analysis.

3) DTs IN AIRCRAFT OPERATION AND MAINTENANCE (O&M): MAXIMISE ASSET VALUE

Ultra-reliability is and will always be the most crucial criterion in the aerospace industry. Arguably, the features of DTs are closely coherent with the modern O&M paradigm. In the modern O&M paradigm (i.e. predictive maintenance and condition-based maintenance), real-time data and forward-looking data analysis are undoubtedly important. On this basis, the owner can realise effective fault diagnosis and prediction, maintenance scheduling and management. This explains the basic and obvious value of DTs to aerospace O&M. The world-leading aerospace institutions including NASA, EASA [60], U.S. Air Force [8] and Royal Canadian Air Force (RCAF) [61] are all working on applying DTs to O&M.

In the daily O&M of the aerospace industry, the real value of DTs is mainly reflected in its model. Models determine the functionality and effectiveness of DTs in O&M, and with the
support of an appropriate model, DTs can be competent in most of the tasks in O&M. DTs can realise aircraft health management and in-flight optimisation. DTs can integrate dynamic information such as the environment and aircraft status to provide optimisation recommendations for aircraft operations, such as fuel optimisation, flight route recommendations, etc. DTs can also predict the RUL of components to achieve predictive maintenance and reduce downtime. Additionally, DTs provide support for the entire maintenance chain, including inventory management, maintenance plan formulation, and maintenance process tracking. The entire maintenance chain can provide end of life decision aid for DTs’ support customers.

Beyond that, DTs have the potential to unlock more functions which have not been considered in the modern O&M paradigm. For example, consider the more comprehensive maintenance scheduling and health management at fleet-level [8], [62], instead of focusing on individual aircraft. Geographical restrictions are removed and eventually they help realise more efficient virtual-based remote maintenance [63], etc. These features will further help the aviation industry to reduce costs and optimise O&M activities.

C. BUSINESS CHALLENGES

It is important to understand various challenges before introducing DTs into production. One of the most disputed issues is the cost [64] taking the U.S. Air Force’s existing system as an example. Based on the estimation of coding, hardware and software requirements, it is predicted that the overall U.S. Air Force system DT budget will reach between $1 and $2 trillion, and the development time will be a hundred years or even longer! The user’s computing and data process capacity also need to be considered because the DTs could be computationally expensive. Additionally, Chartered Engineer training and cyber security are also some of the key restrictions.

However, [35] found that DTs reduced the cost of F-22 wind tunnel activities by 8 million dollars when developing CFD model-based DTs for the U.S. Air Force. [65] has reported that the US Navy was able to reduce 25% of the NPD period for their large aircraft programme by successfully using DTs and related technologies.

IV. MODELLING, VISUALISATION AND INFRASTRUCTURE

It is equally important to determine specific technologies and methods for developing a DT, targeted at some specific system/process/product. This section, highlights three key enablers of DT, i.e. modelling (including data processing), visualisation and infrastructure, so as to help the aerospace systems’ designers and maintainers gain an understanding of how to build DTs aimed at their specific requirements.

A. MODELLING

1) DATA COLLECTION AND PRE-PROCESS

Nowadays, the big data era brings massive amounts of information and has profoundly influenced the way people acquire hidden value and information from data. In big data, the value of historical data has been further explored [66], and the data flow becomes flexible and highly traceable (e.g. reverse assembly engineering in manufacturing [47]), and provides a platform for heterogeneous, multi-dimensional and multi-source data integration and utilisation. Data mining and management in the context of big data have released many original information restrictions, especially for complex systems. The aerospace industry is one of the most data-intensive industries. In the Airbus A380 manufacturing, Airbus and its contractor need to process information from four million components. Further to this, a single flight test of a Boeing 787 collected data from 200,000 multi-modal sensors. In service, an Airbus 380 comprises 25,000 sensors, and a Boeing 787 engine generates 1 Terabyte sensor data per 24 hours [45], [67], [68]. This amount of data is only set to increase in the future. For the aerospace industries, the foremost consideration is to ‘put the correct sensor to the correct location and pass the correct information on to the correct analytic model’. Apart from hardware requirements, collecting high quality from multiple heterogeneous data sources will significantly impact the following realisation of DTs’ functions.

![FIGURE 10. Data pre-process.](14.png)

Data collection does not simply transform the data from the sensor to models but needs to fully consider every risk and uncertainty [69], e.g. sensor fault or extreme operating environment. Figure 10 shows a standard data pre-process procedure. After the real-time data has been collected, the second step is to ensure data reliability. Data from a single source is generally to be considered unreliable data. The common solution is to add additional sensors and data sampling points. Another solution is to use a more reliable external data source. There are three primary ways to obtain aircraft data, i.e. the simulation or synthetic data (generated by simulation software), experimental data (e.g. wind tunnel test) and in-flight data. Among the above data sources, the cost of acquiring experimental and in-flight data is staggeringly high, and irrespective of whether it is for defence or commercial aircraft, manufacturing or flight data. By contrast,
simulation data is cheaper, but the practical value is limited. Both experimental data and simulation data can be used as the external data sources to support the in-flight real-time data merging the data from different sources to improve data reliability. Simple fusion methods include averaging or weighting, while more complex data fusion methods may consider using the Bayesian model or finite element models (FEM) etc. The third step is to rebuild the missing data and remove the noise. Often noise and missing data issues usually can be solved during the data fusion process. Even so, in extreme conditions, if the fused data is still unusable, data reconstruction may be employed. The process of data reconstruction can consider inputting the real environment (including flight status, instructions, etc.) into the reconstruction model to retrieve the data. The popular reconstruction tools such as Kalman filter and its derivative method (e.g. extended Kalman filter [70]; unscented Kalman [71] filter and particle filter [72]), can also be used in data fusion to generate simulation data. The fourth step is data standardisation or normalisation. Data standardisation or normalisation can be seen as a process of removing heterogeneity to building a uniform information environment, because different types and formats of data will directly impact the data exchange. DT structure coherence and co-operation between models in different DT units. A feasible approach is to establish a unified data format. The most popular data mapping technology is XML/XPDL [73], [74]. Others available technologies are AutomationML (AML) [48], [75] and Siemens JT. The unification of the format is also conducive to data storage and query. [76] proposed an approach to establish a single information environment by creating the multi-dimensional database in order to realise data standardisation. Eventually, processed data will transform into an analytic model as the input.

\( a: \) SURROGATE MODEL

A surrogate model is an engineering method used when data cannot be directly obtained or measured from the target. In data processing, the surrogate model can be directly used to eliminate heterogeneous data to achieve the unification of data types or to generate ideal data to combine with real data to achieve data fusion, eliminate noise or rebuild missing data etc. Surrogate models are widely used in aerospace modelling due to the restrictions of sensors in aircraft (e.g. weight, work environment and limited space). For example, aero-engine condition monitoring is one of the representative domains using surrogate models to assist in the construction of physics-based models or data-driven models. It can trace to the early aero-engine model, piecewise linear (PWL) [77], and including nearest physics-based models, such as STORM [78] and IHKF [79]. Another example of using surrogate models is the PGCA model which Royal Canadian Air Force widely used in their airframe digital twins (ADT) [61]. In manufacturing, the surrogate model is also widely applied in complex assembly tasks [66], [80]. However, surrogate models are mostly physics-based, which in the case of insufficient professional knowledge may lead to invalid surrogate models. Therefore, when using the surrogate model in data processing, the necessity, complexity, and accuracy must be considered.

2) PHYSICS-BASED MODEL

Physics-based model (PBM) (also referred to as model-driven) is defined as using mathematical equations to transform the engineering science knowledge building an experimental or numerical model that can reflect the relationship between physics phenomena. The essence of the PBM is to describe the relationship between variables, and hence, it is not sensitive to the input and output.

The research and application of PBMs have been available for decades. Now, the PBMs are quite mature and applied almost everywhere. Arguably, systems or components that can be described by mathematical equations can build PBMs. The last evolution of PBMs comes from the addition of high fidelity simulation technologies. In the aerospace industry, Boeing took the lead in introducing computer-aided physical based 3D models (CAD/CAE/CAM) into the design and achieved huge success in both cost and reputation. The aviation industry has benefited from these 3D PBMs for more than 20 years. Until now, the common PBMs are mostly driven by high-fidelity software (e.g. CAD, Catia, ANSYS, Siemens Nx, Autodesk etc.).

Thus far, most analytic models in DTs are physics-based. [72], provides some current examples of the application of PBMs in aero-DTs. [81] used a set of PBMs to represent the dynamics of the system and its degradation. The models can generate random scenarios and compare to the measurable data in order to identify the fitted scenarios for crack detection and remaining useful life (RUL) estimation. [82] creates a non-linear, touchdown wear response surface model through empirical simulation combined with the slip wear rate equation to serve as the DT of the aircraft touchdown system. [83] uses probabilistic damage tolerance analysis to achieve crack prediction.

Recently, using DTs to simulate a process (e.g. machine process, assembly process) has become popular. In this ‘digital process twin’, PBM has obvious advantages compared to data-driven models (DDMs). A process twin generally contains three kinds of models, machine behaviour model, material model and process simulation model [39], [42], [45], [84]–[86]. The process simulation model is used to simulate the process that the operator can monitor and control. The machine behaviour model represents the behaviour of the CNC or robot that will execute the task, with the material model used for supervising and verification of the effect of the process. A process twin usually needs a clear presentation for the process behaviour, and therefore, PBMs is thus far have performed better than DDMs.

Compared to the data-driven model, the advantage of PBMs is that there is high interpretability and transparency. All mathematical equations used to represent the logic between physics are traceable and explainable and follow
specific physics principles. Therefore, the error can be easily located and optimised when the model does not work well. Meanwhile, PBMs easily transferrable, when two different tasks have similar physics characteristics and logic. Additionally, PBMs do not consider bias. Although PBMs for the same system established from different domain knowledge may have differences in the effectiveness of the models, the difference comes from the model property instead of the target system and the data it provides.

However, the restriction of the PBM is obvious as well. The design and effectiveness of the model largely depend on prior knowledge from engineering experience and experiments. In systems with complex working conditions where engineers have very limited ability to describe system dynamics, such as aero-engines, the establishment of PBMs will be difficult [87]. There are some discrete methods that can mitigate this difficulty and help the PBMs be as close to the real physical situation as possible [88], such as Finite Element Methods (FEM) [89]–[92], the boundary integral method and finite differences. However, these still cannot hide the limitations of PBMs. Additionally, the model itself cannot process data (such as eliminating noise and environmental interference), and therefore the model performance is extremely sensitive to data stability [72].

3) DATA-DRIVEN MODEL

The data-driven model (DDM) represents the capacity of using statistics and modern computing power to analyse data and dig out its potential value. Although the application of data science can be traced back to fifty years ago [67], the popularity of the DDM has taken off in this decade. There are four trends that led to the popularity of data-driven, 1) The significant improvement of sensor, data transmission and data storage technologies. 2) Computing capacity is dramatically increasing. 3) The application of Machine Learning (ML). 4) Open-source ML architecture and open access ML tools reduce the cost and barriers of establishing a DDM. Figure 11 gives a general classification of DDMs. It should be noted that ML is just one of the efficient approaches to realise DDMs but is not necessary. ML is more strictly a tool rather than a modelling method. There are still some statistical method-based DDMs excluded from the existing ML architecture, such as empirical likelihood method, and some data-driven modelling choose not to use ML [93]–[95].

DDMs provide the opportunity for the aerospace industry to deal with high-dimensional, non-convex, and constrained, multi-objective problems, such as aero-engine diagnosis [87] and composite fabrication [67]. This is because DDMs can mine the potential connections between variables from the data without any prior physics knowledge. The essence of DDMs is actually a hypothesis which assumes the change of variables (data) can potentially describe the system behaviour. The logic between PBMs and DDMs is essentially different. The PBM is to build rules based on natural laws, and then leave the variable running naturally under the rules.

The DDM, on the other hand, is focused on the variables, and use them to reason the rules.

a: MACHINE LEARNING

ML is a subfield of AI. From a modelling perspective, ML can be regarded as DDMs supported by higher performance computer technologies, in which engineers use a set of basic coding frameworks based on statistics and mathematics to build DDMs completing a series of general tasks, such as classification, decision-making and generation.

The ML methods can be divided into three categories, supervised learning, semi-supervised and unsupervised and reinforcement learning (Figure 11). Each category corresponds to a different function. Supervised learning is mainly to realise classification, regression and estimation tasks. It usually requires sufficient amounts of labelled data for the algorithm training and testing before it can be used in practice. [90] uses a support vector machine to estimate the aircraft fuel consumption. [97] proposed a DDM using a support vector machine for aero-engine condition monitoring. [98] compares the application performance of common unsupervised learning methods in avionics fault diagnosis.

Unsupervised learning usually applies in clustering and predicting tasks. Compared to supervised learning, unsupervised learning is less dependent on labelled data. [99] proposed an unsupervised ML using a deep auto-encoder to detect the in-flight failure data. [100] evaluates the performance of common unsupervised learning methods in the early diagnosis of faults. Reinforcement learning is a relatively new approach in ML. Scheduling and decision-making are the two most widely used domains of reinforcement learning. Reinforcement learning does not require any labelled data. Instead, it uses the reward mechanism to achieve autonomous analysis in a certain task environment. [101]–[103] show the application of reinforcement learning in the aircraft maintenance schedules.

Recently, the emergence of deep neural networks (DNNs) has, once again increased the popularity of machine learning. DNNs are a kind of ML structure that comes from the combination of deep learning (DL) and artificial neural networks (ANN). By definition, a neural network with more...
than three layers can be called a DNN. DNNs deepen the complexity of the machine learning model but also greatly improve the accuracy of the model. Convolutional neural network (CNN) is a kind of supervised DNN currently popular in the image processing domain. CNN has an excellent performance in diagnostic tasks that use pictures as input data. [104] uses CNN combined with 4K cameras to detect airframe dents. [105] applies CNN with thermal imaging to detect mechanical failure. The other popular DNN is a recurrent neural network (RNN). RNN is a semi-supervised DNN that can handle sequential related prediction tasks. [106] proposed an LSTM (long short-term memory) RNN to predict the aircraft boarding time. [107] developed a LSTM RNN for aircraft trajectory prediction. Similarly, generative adversarial networks (GAN) are also commonly used in aerospace prediction tasks with image-based input [108], [109]. GAN is an unsupervised generating learning model with a self-supervised learning capacity. GAN can generate simulation data which is close to the real data through self-confrontation and is considered to be the key technology for autonomous maintenance in the future [96]. Nowadays, machine learning (ML) has gained tremendous attention in almost every domain in the aerospace industry. In aerospace manufacturing, ML-based DDMs are used to support production line robotics and automation [110]; product design [111]; non-destructive inspection [112]; assembly including inspection, processing and verification [113]; and management including supply and logistic chain management [114], manufacturing scheduling [115], strategical decision-making and recommendation [67]. In in-service M&O, ML-based DDMs are leveraged to airside support including anomaly detection [116]; airframe injury and damage diagnosis [117]; dynamic state estimation and reasoning [118], [119] and maintenance scheduling [120].

The initial application of ML in aerospace DT model and data analytics is mainly to assist or combine PBMs to identify system status [98]. It is also used to build the surrogate model to replace the part that for which is difficult to create PBMs, such as aero-engines [87] and aerofoil [121]. However, the performance of DDMs largely depends on the validity and quality of the selected data [119]. The effectiveness of DDMs, therefore, has a certain uncertainty and instability between testing and practice. Furthermore, because the rules in DDMs operate in the black-box, when the model malfunctions, it is very likely that it cannot be corrected effectively in a short time.

4) HYBRID MODEL

Consequent upon weighing the pros and cons of both PBM and DDM, a community of researchers has put forth a niche concept of ‘Hybrid Models’. By devising an orderly combination of the two, they have been able to eliminate limitations (such as the stringent requirements of large amount of data and knowledge) experienced when PBMs and DDMs are used individually.

One of the hybrid modelling approaches is called the physics-informed neural network (PINN). PINN essentially is still a DDM. The distinguishing feature of PINN is to incorporate the physics principles and governing laws into the neural network. The neural network thus gains the ability to generate data and resist noisy data through conducting physics principles [122]. Compared to traditional DDMs, PINN has the following advantages: 1) low requirement for data and the ability for efficient self-training; 2) improve the versatility of the model, even if there is a deviation between the training data and the real data; 3) improve the transparency and interpretability of the model due to the addition of the principles and laws. [122]–[124]. [125] proposed a PINN for monitoring airframe corrosion fatigue. The proposed PINN is a cumulative damage model consisting of physics-informed layers for modelling the well-understood part, and data-driven layers for the hard model part. Similarly, [126] developed a PI-RNN for airframe corrosion fatigue detection.

The other form of the hybrid modelling approach is called the data-driven physics-based model (DDPBM). DDPBM is actually an integrated PBM in order to realise co-operation between models in complex simulations (e.g. DTs). In DDPBM, the physics model serves as the basic model or sub-model for simulation, while the data model serves as a navigator to connect the sub-models to eventually form as a whole. [127], [128] use such a method to develop a digital twin library for UAV. In the research, the UAV DT is made up of a set of PBMs. The PBMs as a replaceable part can be customised according to specific missions or customer requirements to assemble different UAV DTs. The DDMs are used to determine how to combine and operate based on the requirements.

From the perspective of the current DT trend, the hybrid model may be a necessary modelling approach in DT architecture, because on one hand, as the modelling of DTs becomes more complicated, it is difficult to find a ‘one type of model to fit all’, modelling scenarios, while the complementarity of PBMs and DDMs will be more advantageous and increase modelling flexibility. On the other hand, the hybrid model will be an efficient solution for the current DTs’ inner multi-scale and inter-regional models’ co-operation issues.

B. SIMULATION

Simulation is a topic with a vaguely defined boundary. Generally, DTs themselves can be seen as a simulation method, while not all simulations are DTs.

The approaches used to implement DT simulation today can be roughly divided into three levels based on the degree of visualisation. The first level is abstract models. It constitutes an abstract simulation through simple icons and lines, for example, the topology model and Matlab/Simulink model. This type of model highlights potential physics principles driven by the physical entity rather than presenting the shape and details. In DTs, one of the important functions of abstract models is to provide simulation data for analytic models as
surrogate models, in particular, for the physics-based analytic model which also relies on physical principles. Abstract models, especially Matlab/Simulink models, are very common in current DT applications. The advantage of the abstract model is the low-cost visualisation of physical entities. However, the disadvantage is obvious, users/customers without certain background knowledge cannot understand the models, especially as they are mostly non-3D models.

The second level is the 3D model. Compared to abstract models, 3D models are more intuitive to present the details of physical entities. Nowadays, 3D modelling is quite mature. A series of 3D modelling software, such as Nx [129], ANSYS [130], CFD [131] and CAD [132], can provide an efficient platform and toolkit for visualisation of DTs.

The third level is immersed simulation, which mainly refers to VR and AR. VR is a technology that creates an immersive virtual environment that realises the interaction between the virtual and reality, while AR is a combination of the real environment and the virtual environment providing virtual support to a real scenario. Nowadays, they already play an important role in some specific tasks, for example, VR for UAV tasks in open terrain [133]. AR for remote aircraft maintenance tasks [63]. Meanwhile, in addition to being an interactive visual interface [134], state-of-the-art research in AR and VR are also trying to achieve analytical reasoning [135]. This trend means that in the future immersive technology will have the opportunity to completely replace the existing simulation technology. However, VR and AR are still immature. A crucial challenge is the quality of the human-machine interface [136]. It involves how to realise and apply some specific interactive technologies, such as gaze-tracking and hand-tracking. Additionally, some general challenges such as cost and enabling technologies are restricting its development.

C. INFRASTRUCTURE

A basic DT should have three kinds of infrastructure, i.e., data collection, data transformation, and data computing equipment. Additionally, to expand the functionality of DTs, other infrastructure also needs to be considered depending on the requirements of different situations, such as data gathering, visualisation equipment and data management system. Sensors are the main approach for the aerospace industry to collect data. With the variety of data (e.g. torque, pressure, speed, vibration, voltage, current, temperature, voice and thermal imaging) that can be collected, the types of sensors to choose from have also become diverse. Meanwhile, new sensor types are emerging and already prove their value in specific tasks, such as RFID [58], [66], intelligent sensors (e.g. light sensor monitoring label, current monitoring label and industrial sensor), laser sensor [80] and camera based visual sensor [137], [138]. However, different from manufacturing and ground maintenance, on-flight aircraft has stricter requirement for sensor selection. High temperature, high pressure, and the reliability of the sensor in the extreme environment must all be considered [139]. [140] proposed four criterions for aircraft sensor selection, i.e. the detectability of sensor, can all levels of fault be detected and how early after the fault; the diagnosability within certain restrictions (e.g. weight, cost); the reliability and robustness; the observability (e.g. what kind of information from the sensor can be observed). Meanwhile, the adaptability of sensors in the network and the chosen IoT platform software need to be considered.

Data transmission has made great progress in the past decade. The rapid development and variety of communication technologies facilitate the maturity of real-time data transmission. Building IoT networks have more choices on the sensor side. Manufacturing is the one of the biggest beneficiaries. The IoT communication technologies available in the manufacturing industry can be generally divided into two categories: short-range networks and low power wide area networks (LPWA) [141]. Short-range networks are the more traditional application that are already widely applied in the manufacturing industry and even across our life. The representative technologies of short-range networks are Bluetooth, Legacy wireless local area networks (WLANs), ZigBee, Z-Wave and wirelessHART. LPWA is a relatively novel technology that has gained significant attention this decade for IoT application. Some mature LPWA technologies are 4G/5G, LoRa, NB-IoT and SigFox. Principally, LPWA is designed for long distance and lower cost and energy consumption. However, in practice, both short-range networks and LPWA have their own specialist domain. For example, both 4G/5G and WLAN have good versatility and enough fast transmission speed but much higher energy consumption. LoRa and Bluetooth have lower energy consumption, but the former cannot continuously transmit data, and the latter has a smaller coverage area. Generally, a series of factors needs to be balanced when choosing communication technologies for manufacturing, such as transmit rate, distance, security, cost, energy consumption, portability, etc.

Nowadays, the data transmission methods of aircraft are mainly divided into three types, HF (High Frequency), VHF (Very High Frequency) and SatCom (Satellite Communication). HF and VHF are both radiowave technologies. The former is widely used in long-distance communication while the latter is suitable for short distance. Compared with popular data transmission technology, radio technology has appeared for decades. However, modern radio technology has undergone significant promotion and improvement such as the application of advanced automated HF communications management (AHFCM) systems and high-performance data modems [142], and its transmission efficiency is not lagging. Compared with the other two, SatCom has obvious advantages in performance and has been widely used in many new aircraft, but satellite technology entails much higher costs. Recently, researchers have begun to explore the possibility of applying 5G to air-to-ground communications. [143] discussed the possibility of 5G air-to-ground from the perspective of ground base station layout and cost. [144] discussed the challenges and future direction for applying 5G New
Radio (NR) in Air-to-Ground communication systems. 5G undoubtedly has advantages in data transmission efficiency, but more practice is still needed to measure the commercial value between costs and margin. In the urban environment, 5G is already sufficient to cover the UAV airspace and provides faster and stable remote data transmission and control [145], [146].

Computing capacity determines the upper limit of DT infrastructure. For the construction of the high dimensional and complex DT with huge data, high-performance computers will be necessary [8], [131] points out that the HPC community has started to deal with DT-related issues i.e., handling large and noisy data, quantify uncertainty and huge number of equations. HPC with cloud computing technology will maximise the utilisation of computing power to ensure the operation and interaction of DT components. Additionally, edge computing technology becomes increasingly important for alleviating the data pressure of aircraft under restricted conditions. For example, Boeing’s 787 generates 5 GB data per second, the data transmitted pressure will be given to satellites or base stations for central process and computing [147]. Edge computing will efficiently shorten the distance between data collection and processing while speeding up the response time.

D. PRACTICAL CHALLENGES
Even considering all the above aspects, there are still some common practical challenges during the implementation. The first common challenge is the accuracy of the initial data. The inaccuracy of initial data may come from 1) actual aircraft manufacturing error which is the difference between manufacturing and design, such as material properties and wrong punching. It may cause deviations in modelling (especially PBMs); 2) the MRO-caused error. These errors will increase the uncertainty of the DT in simulation and condition-monitoring, which will or may result in the DT failure. The second common challenge is noisy data with ambiguous features. Although theoretically, data processing can alleviate these issues, in practice, a series of external factors (such as sensor accuracy and degradation) will slowly but continue to produce uncertainties to affect the data collection quality. The third common challenge is model verification and validation (V&V). The model V&V in DTs is a long-term requirement, because as time progresses, it will become difficult to maintain consistency between DTs and physical entities permanently. Eventually, according to NASA [10], a self-V&V model may become necessary. Additionally, there are some other common challenges such as trust issue and cyber security issues need to be considered as well during practice.

V. DTs ROADMAP: A PATH TOWARDS THE FUTURE
It has been almost 20 years since the first proposed DT concept. For the whole 2000s, DTs were still at the highly conceptualised stage. Until around 2015, the definition of basic DTs was further clarified. According to Grieves [148], a DT should at least have three-dimensional parts: physical products in real space, virtual products in virtual space, and the connection of information linked the virtual and real. The concept formed the most popular and classic shape of DTs today. Meanwhile, research is also trying to explore the next generation of DTs. This section introduces the limitation of the basic DT and summarises three directions of research for the next stages of DT Development.

A. BASIC DT
Basic DTs are the currently most widely-used DT. A virtual replica can be called a basic DT when it has the following three elements: the ability of simulation, real-time state synchronisation, and data collection from its physical entity [149]. Therefore, a basic DT usually contains a single function model and a simulation model (usually a non-3D model) to realise such basic functions as monitoring and simple analytical tasks. In some ways, a basic DT is more like a simple integration of the tetra-driven technologies of innovation. It has a similar ideology with DT, but it cannot realise many core functions. For example, 1) The basic DT only covers a certain stage or task of its counterpart (such as design, manufacturing, etc.), and cannot realise a closed loop of the whole lifecycle. 2) A basic DT cannot evolve with its counterpart when the external or internal environment changes. 3) The basic DT can meet the function of synchronisation with its counterpart but cannot be stand-alone as an independent part to develop its own lifecycle. 4) A basic DT usually weakens the functions of visualisation and human-computer interaction experience.

The missing function in basic DT also reveals two fundamental issues in DT development. 1) DTs’ definition and functionality are still unclear and imprecise. Some researchers believe that DTs should be process-oriented. The ‘counterpart’ in this context should be the behaviour of multiple objects in this process. Others argue that DT should be component-oriented. Its counterpart should be the behaviour of one component in multiple processes. Whether a high-fidelity model is necessary, what is a complete DT architecture, etc. are not uniformly defined; 2) The enabling technologies of DTs are still immature or vague. There is still no substantive unified way of how to twin the whole life cycle, and the existing technology is facing problems of high cost and high barriers to entry. Based on the issues above, there are three research directions for the future DT.

B. INTERACTIVE DT
For a long period of time, DTs research has only focused on creating data flow and analytic data models while ignoring the value of DTs that can work as an independent system. Interactive DT aims further explore the value of the virtual side. [150] adds two dimensions of service and DT information to the original three-dimensional DT paradigm, thus forming a new five-dimensional DT paradigm. The two newly added dimensions are mainly used to reflect the connection from the virtual side to the customer side and the virtual side to the physical side. Service mainly represents
a series of human-machine interactions for DT customers, such as detection, diagnosis and prediction services through functional models, high-fidelity simulation services through visualisation technology, the services for customers to call and view product historical, real-time and predict data. DT information refers to the data generated when DTs are running as an independent system. For example, [149] points out that DT should have its own identifiable ID and management system, including management of DT version and information. Some information can be used to, directly and indirectly, influence the development of the physical side. Besides, DTs should reserve interfaces for tools, new components and other DTs.

Figure 13 shows a schematic diagram of a five-dimensional DTs and its data flow. In the diagram, a data layer with multiple sensors and information models is built for data collection. The application layer is the core of the interactive DT, which consists of sets of analytical models and other models, systems or components corresponding to the services based on requirements. The presentation layer is for the processed data visualisation, storage and management, including data from both the physical asset and DT self-generated data.

C. STANDARDISE DT

DT standardisation means unifying the DT architecture, data format and modelling method. It aims to offer practical guidance for the user to answer questions such as ‘where can we begin with digital twin’, ‘how can we build a DT’ etc. Especially when considering the different formats, protocols and standards in enabling technologies, a uniform DT architecture and tools become necessary [151]. The business community will be the direct stakeholder in DT standardisation. It will greatly shorten DTs’ R&D period and
reduce uncertainties. Customers do not need to worry about ‘which models are better’ or ‘inconsistent information environments’. DT standardisation can further expand DTs’ commercial value.

The most direct work on DT standardisation is the development of DT toolkits and platforms. There are a number of high-tech companies already working in this field which have proposed excellent DT toolkit. Oracle Internet of Things Cloud Service offers a platform for DT implementation including a number of toolkit combinations for different solutions such as predictive twin, project twin, etc. [157], IBM Digital Twin Exchange is also a DT platform, which focuses more on ERP and related system [26]. Additionally, Microsoft Azure digital twin, GE Predix OPM, Siemens Digital Enterprise Suite, Cisco Kinetic IoT platform, Dassault Systems, etc., all provide R&D platforms for DTs. In the UK, the British government released the Gemini Programme to unify DT R&D toolkit to build digital cities [153].

Recently, establishing a prototype model through transfer learning to support DT standardisation is considered to have high potential commercial value. Transfer learning can thereby re-use the model to significantly reduce the modelling cost. For example, [154] applies a prototype model for estimating runway occupancy time to both Vienna Airport and Barcelona Airport, and claim the prototype can potentially apply to the worldwide airport. [155] developed a CNN model to help the airport detect pneumonia. The prototype of the model comes from the open-source pre-training model.

D. INTELLIGENT DT

The majority of the DT research emphasises the phrase twinning, an example of which could be, how to build the high-fidelity twin and collect synchronisations information. However, only twinning is not sufficient for a DT to realise higher-level information gathering, learning and reasoning in order to complete a complex mission. This explains why the DT community is committed to making DTs smart (i.e. intelligent DT). NASA, in their roadmap published in 2010, mentioned that they would build an intelligent and adaptive DT to support their space mission around 2025 [10]. Being intelligent and adaptive means that a DT should not be limited to simple system monitoring and task analysis. According to the original DT concept [156], it should not only describe the state of the system, but also derive solutions for the real system. [157] pointed out that DTs should have cognitive ability to evolve with its counterpart through the whole life cycle.

Cognitive Digital Twin (CDT) is one particular kind of intelligent DT. The creation of CDT comes from the combination of cognitive science and machine learning, and it is deliberately designed to give normal DTs more powerful reasoning and learning capabilities. CDT is an adaptive and evolvable DT. Compared to the basic DT, CDT should additionally have the abilities of perception, attention, memory, reasoning and problem solving. CDT can synchronise with its counterparts, but it can also operate independently, it aims to realise the complex information process and correlation to support decision-making and recommendation. The core of the CDT is a knowledge/rule-based information map and its reasoning algorithm, from the whole system level (factory, production line, etc), gathering and categorising all sources of information into the map, and arranging and merging information through algorithms to respond to queries or navigation.

VI. CONCLUSION

Inevitably, the tetra-drivers of innovation, namely cloud, big data, Artificial Intelligence (AI), and Internet of Things (IoT), have supported and catapulted the growth of digital replicas of physical entities, the DTs, across a wider range of industries. This paper has presented an up-to-date survey and highlighted the very concept and composition of DTs inter alia highlighting its potential value as enablers for aerospace and its vital manufacturing segment. It has been shown how DTs originated and made progress to industrial digitalization. Their gradual adoption by the aerospace industry is elaborated with reference to NPD and O&M. Undoubtedly, with the high density of sensors onboard aircraft and intra-connectivity of a number of systems, we are witnessing complex system configurations and generation of massive amounts of heterogeneous data. This has brought about a number of challenges in optimising massive data management (in terms of transferability, processing and analysis) to build high-fidelity Aero-DTs for different vital aircraft systems (such as propulsion, landing gear, avionics, etc.). Aero-DTs, if optimised responsibly, will respond positively to the challenges and provide comprehensive technical support to improve product lifecycle management.

By virtue of the growing value of Aero-DTs in the aerospace industry, organisations and researchers are actively involved in DT R&D. However, many practitioners are not clear about the types, pros and cons of enabling technologies of DTs and do not know how to get started. This paper provides an in-depth review of the state-of-the-art of aero-DTs, classifies, introduces and analyses key enabling technologies in order to give the readers a clear overview. At present, there is no optimal solution for Aero-DT enabling technology (e.g. the literature suggests that there is a tendency amongst DT developers to blindly choose ML models without considering the physics involved). It is, therefore, important that DT developers should design and build Aero-DTs using relevant optimal DT components exactly related to the specific operational and functional requirement. Various limitations and challenges in the optimisation of DTs for the aerospace industry are highlighted: lack of relevant DT design knowledge, real-time data acquisition and interface issues being the most prominent. The survey puts forth a solution to these challenges by proposing a roadmap that encompasses three main elements (standardisation, interactivity, and cogntiveness), which if leveraged sensibly, could help the aerospace industry transform their systems’ performance and useability to higher levels of operational readiness.
It is further argued that as DTs gradually augment or totally replace the CAD/CAE/CAM-dominated manufacturing design, machining, and assembly processes, high-fidelity aero-DTs will essentially improve the existing under-performing condition-monitoring and PHM (prognostics and health management) technologies with enhanced visualisation, faster data computation, and accurate predictive analysis.

REFERENCES


ANDREW WILEMAN received the B.Eng. degree (Hons.) in electronic engineering from the University of Nottingham, U.K., and the M.Sc. degree in digital communication systems from Loughborough University, U.K. In 2012, he joined the Integrated Vehicle Health Management (IVHM) Centre, Cranfield University, as a Ph.D. Researcher in prognostics for electronics/defense systems, sponsored by Thales, U.K., and Meggitt. He worked with the U.K. Ministry of Defence and other engineering companies within the U.K. He has lectured in electronics/communication engineering with the Further and Higher Education Institutions for a number of years. After completing his Ph.D., he is currently working as a Research Fellow in power electronics reliability, AI systems and applications, digital twins, and data-centric engineering. His research interests include the development of IVHM solutions for power electronics systems in relation to more-electric aircraft (MEA), prognostic algorithms, and the application of digital twins for improving diagnostics and maintenance in aircraft and autonomous systems.

SURESH PERINPANAYAGAM received the M.Eng. degree in aeronautical engineering from the Imperial College of Science, Technology and Medicine, London, and the Ph.D. degree from the Rolls-Royce University Technology Centre, Imperial College, under the auspices of the European FP7 Program. During this tenure, he worked with major engine manufacturers in Europe (Rolls-Royce, U.K.; MTU, Germany; and SNECMA, France). He is currently a Senior Lecturer in intelligent systems with the Integrated Vehicle Health Management (IVHM) Centre, Cranfield University, U.K. He was one of the pioneering researchers who was awarded $1 million project, funded by Boeing, EADS, and Singapore’s Science and Engineering Research Council (SERC), under the auspices of the SERC Aerospace Research Program. The project looked at developing novel sensors and signal-processing techniques for the detection of defects on composites of thickness of more than one inch, such as those on the Boeing 787 skin. He has published one book chapter on sensors, instrumentation and signal processing in the SAE International book on Integrated Vehicle Health Management and numerous research articles. He, as a Principal Investigator, has recently received more than half a million pounds in funding from Innovate U.K., and Airbus for exploring data-centric engineering and digital twin concepts for future landing gear platforms. His research interests include data-centric engineering and digital twins for intelligent systems, sensory systems, advanced reasoners, artificial intelligence, predictive capabilities, and decision-making for intelligent systems which demonstrate self-health awareness, operational resilience, failure recovery, system re-configurability, and end-of-life prediction. He is also a member of the Editorial Board of Microelectronics Reliability, Special Issue on Reliability of Power Electronics. He is an Associate Editor of the IEEE TRANSACTIONS ON POWER ELECTRONICS, Special Issue on Robust Design and Reliability in Power Electronics.

* * *
Digital twin in aerospace industry: a gentle introduction

Li, Luning

2021-12-20
Attribution 4.0 International

https://doi.org/10.1109/ACCESS.2021.3136458
Downloaded from CERES Research Repository, Cranfield University