Toward a concept of digital twin for monitoring assembly and disassembly processes

Towards a concept of digital twin for monitoring assembly and disassembly processes

Elisa Verna, Stefano Puttero, Gianfranco Genta, Maurizio Galetto*

Politecnico di Torino, Department of Management and Production Engineering, Corso Duca degli Abruzzi 24, 10129 Torino, Italy

*Corresponding author (mail *maurizio.galetto@polito.it*)

Abstract

Quality is one of the key factors in the customer's selection process between competing products. Producing high-quality, defect-free products that meet consumer expectations is crucial for manufacturing companies to gain a competitive advantage. Accordingly, developing appropriate defect generation models is essential in modern manufacturing companies to predict defects and plan efficient quality control and production. On the other hand, with its ability to support new business models and decision support systems, digital twin technology is one of the new technologies emerging to support digital transformation. Faster optimisation algorithms, more powerful computers and a massive increase in available data are just some of the features of digital twins that can be used to advance simulation towards real-time quality control and optimisation of products and production systems. This paper aims to model the generation of defects of product variants in assembly and disassembly processes and evaluate their integration within a digital twin system to prevent the occurrence of defects and ensure product quality. The proposed strategy is expected to improve the optimization, monitoring and diagnostic capabilities of complex product variants' assembly and disassembly systems, realising an upgrade from a single physical implementation to a combination of physical and digital.

Keywords: quality control, prediction model, defect generation, digital twin.

1 Introduction

Digitalisation is accelerating significantly due to the development of information technologies such as the Internet of Things (IoT), cloud computing, big data analytics, and artificial intelligence (AI). Together, these technologies enable the creation of virtual versions of physical systems and entire environments. Digitalization is a major driver of innovation across all industries as the physical and virtual worlds collide. The concept of digital twins (DTs), introduced by Professor Grieves in 2002 and refined in subsequent studies (Grieves and Vickers 2017; Grieves 2014), provides a framework for integrating the physical and virtual worlds. DTs are consistent with the idea of digital transformation, which aims to create business models that reflect the value derived from data (VanDerHorn and Mahadevan 2021). DTs serve as a bridge between the physical and digital worlds. According to Grieves et al. (2017), DTs consist of three main elements: the real world, the digital world (a virtual representation of the real world), and the bi-directional data connections. Physical-to-Virtual (P2V) connections transfer data collected by sensors in the real world to the virtual world, providing input to create the virtual environment. Sensors such as spectroscopic and 3D vision systems accurately represent the product shape without damaging the physical system. Virtual-to-Physical (V2P) connections enable information and processes to flow from the virtual to the real world, allowing for physical changes based on virtual models. This bidirectional connection allows for physical experimentation data to improve the virtual representation. On the other hand, a Digital Shadow is defined as a virtual model that only replicates the physical model without this flow of information (Kaewunruen, Rungskunroch, and Welsh 2018). Thus, DT is distinct from conventional simulation and modelling activities, where analysis is typically done offline and maintains a constant connection between the physical and virtual worlds (Jones et al. 2020).

The modern industrial sector has started the process of full digitalization. In the past five years, the need for increased productivity, quality, and performance has led to growing interest in DTs. DTs create a communication path between the physical and digital worlds, a critical aspect in transitioning to Industry 4.0 (Guo and Lv 2022). Many industries, including healthcare, manufacturing, agriculture, automotive, and smart cities, have effectively deployed DTs. However, literature shows that DTs often neglect quality control measurements in manufacturing processes (Modoni, Stampone, and Trotta 2022; Liu et al. 2021; De Ketelaere et al. 2022). Section 2 provides a more detailed account of the use of DTs in quality inspections.

Efficient and cost-effective inspection procedures are crucial in modern industry for reducing quality-related costs. It is a key component for establishing market competitiveness (Franceschini et al. 2018) because defects can significantly impact a product's quality and price. Predictive models can be used to monitor the production process in real time, predict quality fluctuation trends, and give early warnings, resulting in reduced wastage of production resources, optimized product yield rates, and reduced losses (He et al. 2022). Moreover, due to customers' growing need for customization, companies now need to produce small batch sizes, and appropriate defect generation models are essential for manufacturing industries to predict defects and plan effective quality controls.

In the scientific literature, several models have been developed to identify defects in final products, and some of which rely on the close relationship between assembly complexity and defectiveness rate (Galetto, Verna, and Genta 2020; Verna et al. 2022b; Alkan et al. 2018). In this paper, the productivity and the generation of defects occurring in-process and offline in manual assembly and disassembly processes of different product variants are analysed and modelled by using designed experiments.

The topic of modelling assembly processes, and in particular the sequencing and reordering of assembly steps, has been explored in previous studies. For example, Bisgaard (1997) proposed a method for experimentally determining tolerance limits by reordering assembly steps for mating components of assembled products. Zhao et al. (2021) focused on incorporating "order of addition" into experimental design. Both studies demonstrate the importance of sequencing and the potential benefits of incorporating it into experimental design and assembly processes. In this paper, the authors apply a similar approach to monitoring assembly and disassembly processes using designed experiments by evaluating the effects of product variant complexity to develop some predictive models. These prediction models can forecast the productivity and quality of a manufacturing system based on the complexity of assembled/disassembled product variants. For instance, when introducing a new product variant with a certain complexity in the assembly line, these models can predict productivity and quality performance. Alternatively, the models can suggest the level of product variant complexity needed to achieve specific productivity and/or quality targets. These virtual models can be integrated into a DT for quality control management to continuously monitor the production process and improve final product quality (Zhu and Ji 2022; De Ketelaere et al. 2022). Thus, these models are a first step towards the construction of a DT in which real-time data from the production process may be collected to refine the virtual representation of the system. The DT can also simulate different production scenarios, identify process parameters or conditions leading to defects, and consequently send feedback to the physical system via the V2P connection. Additionally, retrospective analysis can speed up quality issue detection and improve system performance.

The rest of the paper is organized into five sections. Section 2 provides an overview of the DTs applied in quality engineering and management. Special attention is given to the integration of defect prediction models within the DTs framework. Section 3 presents the DT framework into which the proposed defect generation models can be integrated. Section 4 summarizes the main definitions of product complexity and the approaches used to assess complexity, and presents a case study where the proposed methodology is applied to manual assembly and disassembly. Section 5 illustrates the experimental results collected in the case study and discusses the relationships with product variant complexity. In addition, the obtained productivity and quality performance models are presented. Finally, Section 6 discusses the main results and concludes the paper.

2 Digital Twin in quality engineering and management

To remain competitive, manufacturing companies aim to quickly fulfil customer demand by ensuring the quality of their products (Galetto, Verna, and Genta 2021). Quality inspections are now standard practice in almost all manufacturing systems to prevent delivering defective products to end-users, which could harm a company's competitiveness (Genta, Galetto, and Franceschini 2018). Inspections cover semi-finished items, final goods, or production-related parts and may use human inspectors, automated sensing equipment, or a combination of both. Effective quality control management is critical for continuously monitoring production systems and improving market competitiveness (Montgomery, Runger, and Hubele 2010).

Product defects in a production system can be detected through in-process or offline quality controls. In-process inspections examine units during production, while offline inspections examine final products at the end of the production process (Genta, Galetto, and Franceschini 2018). Quality control and management of production processes are challenging due to the increasing complexity of manufacturing and the dynamic changes and unpredictability of manufacturing resources. The concept of Quality 4.0 has been introduced to address this, which involves using Industry 4.0 technologies for quality control and maintenance of products and services (Shivam and Gupta 2022). Quality 4.0 offers many advantages, such as increased enterprise efficiency, performance, innovation, and improved business models (Sony, Antony, and Douglas 2020).

Thanks to real-time data acquisition and data mapping on the shop floor, digital twins of production processes are continuously informed about the actual process and equipment condition. This information can be used to predict the final quality, making DTs an important alternative to traditional methods for quality control (see Fig. 1).

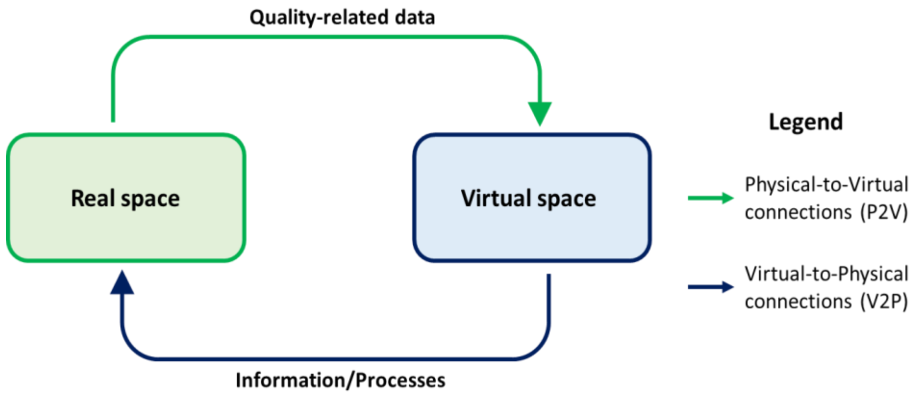

Figure 1. Digital Twin components for quality control.

Fig. 1 shows that the digital twin collects quality-related data, such as process and quality control data, through P2V connections. This data is then integrated with various experimental models, including geometric, mechanical, and material models (Guo and Lv 2022) and defect prediction models. By processing the data collected with the defect generation models, the virtual space sends information to the real world through V2P connections to prevent critical scenarios from occurring and predict the overall quality of the process. Real-time data from sensors embedded in the production process and simulations are tightly integrated, allowing for continuous improvement of both experimental models and real-space components. With fast sensing technology and powerful computing capabilities, DTs can move from offline monitoring to an entirely in-line solution, making them a useful tool for quality control in various applications (Ma et al. 2019).

A digital twin for quality control of a manufacturing process is a helpful approach in improving the quality of complex product variant assembly and disassembly due to its ability to predict unknown situations by delivering accurate simulation results and reducing the number of trials and errors. The DT control model should be created based on product properties, such as the types of components and connections, and process-related variables. Consequently, the first step toward developing the DT is defining the relationship between the assembly quality and process control parameters by implementing appropriate defect generation models (Ma et al. 2019). Integrating defect prediction models and DTs can support traditional quality control and optimize complex product assembly by combining physical and digital methods and moving from experience-driven to data-driven assembly. De Ketelaere et al. (De Ketelaere et al. 2022) propose integrating quality control and digital twin by using data from multiple sensors to create a simulation model for inspecting fruit quality during processing. The study emphasizes the need for field tests to verify the accuracy of defect generation models. By simulating scenarios in the virtual world first, only high-productivity and low-defect scenarios are tested in the real world. This approach reduces the need for massive testing in the real world. In another study, a DT of a rotating machine based on a dynamic model for gear fault diagnosis is presented, highlighting the monitoring, diagnosis, prognosis and analytical prescription capabilities of the DT (Kenett and Bortman 2022).

Thus, an integrated production management and control method is essential in a market with increasingly complex and customised products. DT systems provide real-time simulation, data-driven analysis, and dynamic feedback for optimizing the assembly of complex products (Ma et al. 2019).

3 Integrating defect generation models within Digital Twin for quality control

In the proposed study, the developed prediction models represent a virtual shadow of the physical process of assembly and disassembly process and can be integrated into a DT for in-process quality control. Fig. 2 shows an example of a structure that integrates the defect generation models within the digital twin architecture.

The real space consists of the assembly/disassembly system and all process and quality parameters related to the manufacturing process and quality control procedures, respectively. The process parameters are collected during normal production, while the quality parameters are used to monitor the quality of the assembly/disassembly process. Both process and quality data are sent to the virtual space, where they are processed by the different experimental models, that will be presented in Sections 4 and 5. In particular, these experimental models are, on the one hand, assembly/disassembly and quality control models based on product complexity and, on the other hand, quality and process performance models. These models are then used for predictive and optimization purposes in order to identify critical and optimal scenarios within production. At the end of the prediction and optimization process, the virtual space sends real-time feedback on the system status to the real space, modifying the process and quality parameters as necessary. Of course, in this case study, where the assembly is completely manual and there are no machines communicating directly with the digital twin, the operators could receive audible or visual feedback from the virtual space to improve the assembly. Furthermore, as new data is acquired from the physical system, continuous improvements can be made to the experimental models and, accordingly, to the real space components themselves. This integrated method can ensure an accurate, real-time prediction of assembly and disassembly performance, thus allowing the occurrence of defects to be prevented and/or corrective strategies to be adopted to minimize the occurrence of non-conformities. As a result, the cost and time of quality control can be drastically reduced, and the overall quality of the system improved.

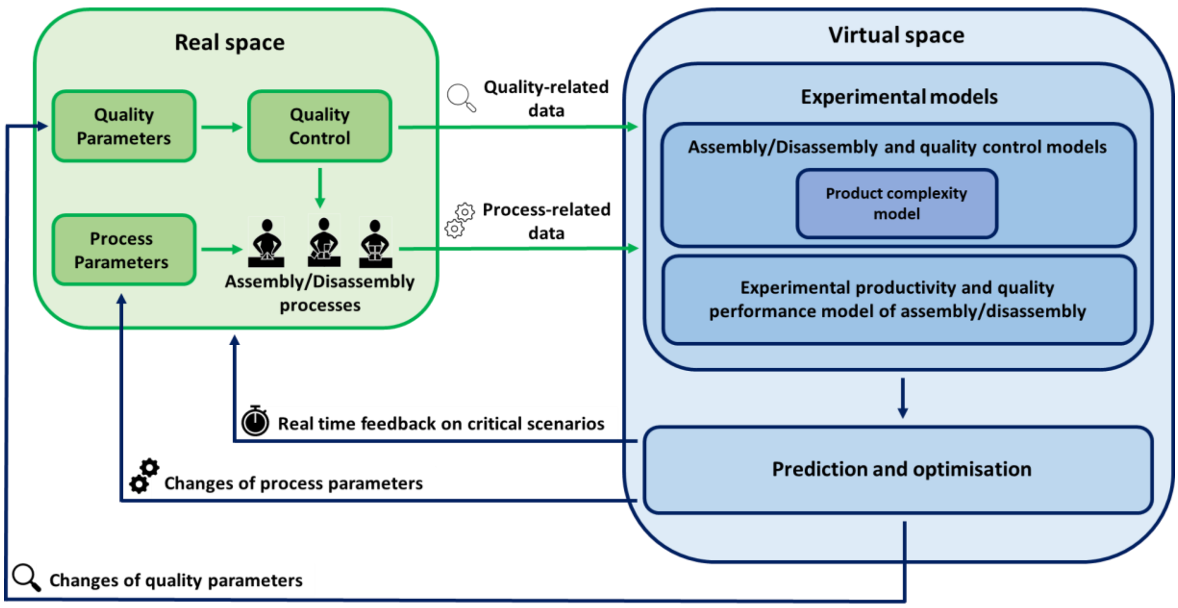

Figure 2. Structure of a digital twin for quality control in which the developed experimental models may be integrated.

However, the integration of defect generation models into the digital twin can be limited by inaccuracies in the data used to create the twin, which can lead to errors in the integration of the collected data. In addition, the digital twin may not fully capture the complexity and variability of real-world systems, making it difficult to calibrate and validate the predictive model. Accordingly, an initial calibration of the predictive models used is essential, in which data obtained from simulations are compared with data obtained in the real space (Yeratapally et al. 2020). In this case study, some initial tests could be performed to observe how assembly times and defects vary with the complexity of the products assembled in the real and virtual environment. The data obtained from the simulations can then be compared with the data obtained from the real assembly, and the parameters of the prediction model can be corrected in case of discrepancies.

4 Assembly and disassembly processes modelling

This section presents and discusses a modelling approach for assembly and disassembly processes. In detail, Section 4.1 describes the assembly and disassembly processes selected as a case study and related quality control procedures of several different product variants. Product variants are characterized by different levels of complexity, modelled according to the structural complexity paradigm defined in Section 4.2. The specific product variants exploited in this work are described in Section 4.3.

4.1 Case study: process and quality control description

In the proposed study, products used for assembly and disassembly are molecular structures made up of balls and sticks. Such structures are widely considered in the scientific community as reference products that may be used for research purposes to emulate real products (Sinha 2014; Alkan and Harrison 2019; Alkan et al. 2017; Alkan 2019; Verna et al. 2022a). Adopting such structures is dictated by the need to achieve general results regardless of product type and field of application. Furthermore, this enabled minimising the confounding effects typical of real productions (e.g., dynamic and organisational issues) and isolating and controlling the effects of product complexity, typical of product variants, on defect generation, as will be further detailed in the next Section 4.3. Twelve different variants of molecular structures are manually assembled and disassembled, characterized by different levels of complexity, as per Section 4.3. The structures are assembled from a molecular modelling kit (Orbit™ by 3B Scientific®) based on 2-D and 3-D work instructions. The molecular structures consisted of different atoms (i.e., balls) and bonds (i.e., sticks). As represented in Fig. 3, the different atom types are carbon (grey), hydrogen (white), nitrogen (blue), oxygen (red) and sulphur (yellow). Two types of chemical bonds make up the structure variants: single covalent bonds made with rigid connectors and double covalent bonds made with flexible connectors.

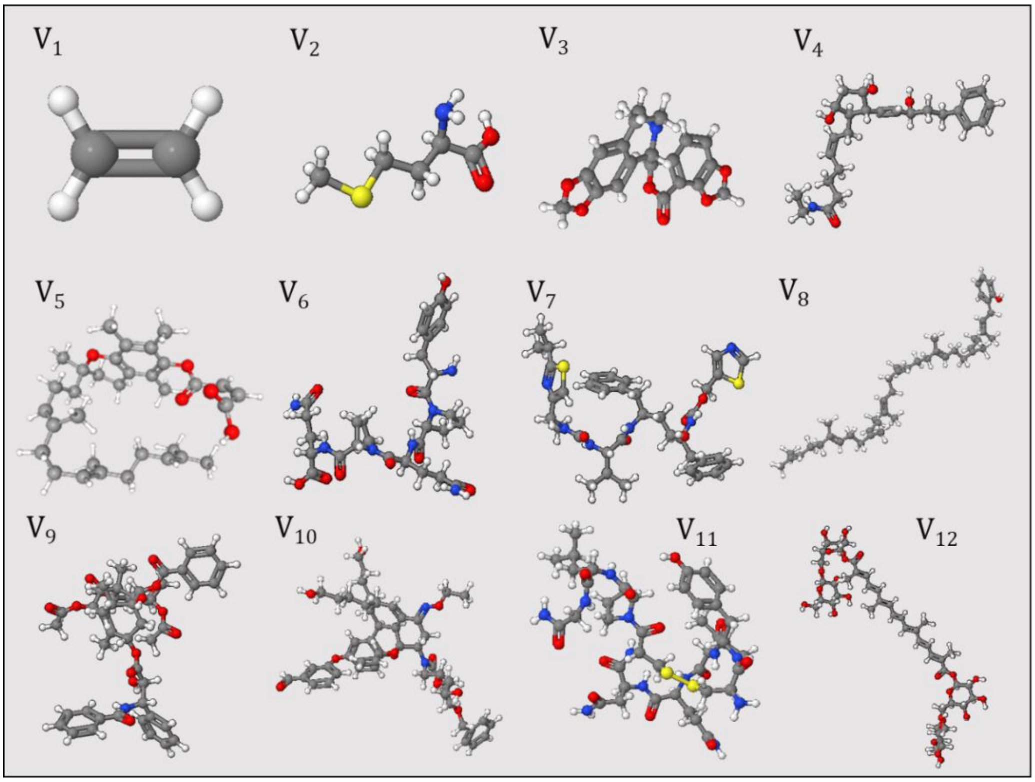

Figure 3. 3-D representation of the 12 variants of ball-and-sticks molecular structures used in the experiments ($V_1 - V_{12}$).

In order to develop prediction models of productivity and quality performance in both assembly and disassembly for the twelve different product variants, an experimental campaign was designed. The assembly and disassembly operators were students of the "Quality Engineering" course in the 2nd year of the Master's degree course in Management Engineering at the Politecnico di Torino (Italy). The experimental campaign lasted 8 working days, in which 52 assembly and 52 disassembly operators were involved, for a total of 104 operators. In detail, on each day, a maximum of 7 operators were involved in the product assembly, each supported by a quality controller who monitored the work. Each operator was responsible for assembling (or disassembling) all the 12 randomly assigned variants of molecular structures described in Section 4.3 (see Fig. 3 and Table 1), following a full factorial design. The operators were divided into pairs: while the first operator assembled a randomly assigned structure, the second operator performed the quality inspections and was responsible for the subsequent disassembly. Each structure was therefore assembled 52 times and disassembled an equal number of times. The assembly and disassembly of the variants did not follow a particular assembly sequence to minimize the effects of sequence complexity, whose effects on complexity are well known in the literature (Gulivindala et al. 2020; Raju Bahubalendruni et al. 2015).

The workstation of each assembler was equipped with several boxes containing one type of atom and bond, and a PC with assembly instructions and files to record the results of the experiments. For disassembly, empty boxes were set aside to store the atoms and bonds, divided by type.

During assembly, the quality control operator was instructed to measure each structure's total assembly time and count the number of in-process and offline defects. In-process defects are the errors that occur during assembly operations, e.g., missing, incorrect, or additional parts or connections, involving the disassembly of one or more parts/connections and the repetition of operations to correctly complete the structure. These defects, which the operator repairs during assembly, lengthen the assembly process and decrease the operator's productivity. On the other hand, offline defects are those found by the quality inspector on the finished product after assembly (such as additional, missing, or defective parts or connections). During disassembly, the quality inspector measured the total disassembly time and the number of defects, e.g., parts and connections placed in the incorrect box (misclassification errors), incorrectly disassembled or left on the assembly bench.

4.2 Modeling product complexity

Complexity has been widely adopted in the scientific literature to predict production performances, including production times and defects, as its increase leads to an increase in performances, often more than proportionally (ElMaraghy et al. 2012; Verna et al. 2022a; Genta, Galetto, and Franceschini 2018). The product structure and its complexity are also a crucial element of DT paradigm. Information about the individual components of a product and their complexities are essential for two main reasons. Firstly, product defects are often related to certain product components and their characteristics (Verna et al. 2022a). Secondly, the structure and complexity of the product itself may strongly affect production operations, e.g. assembly/disassembly order, and performances (Detzner and Eigner 2018). In the context of mass customisation and increasing product complexity, it is therefore crucial to assess product variant complexity to predict product quality. Accordingly, this paper analyses the possibility of integrating defect prediction models based on product complexity into a Digital Twin for quality control of product variants.

One of the most accredited models to assess product complexity, which has been used in the quality control and management field, is the one proposed for the first time by Sinha (Sinha et al. 2012) and then readapted by Alkan et al. (Alkan et al. 2017; Alkan 2019) and Verna et al. (Verna et al. 2022b, 2022a). The model relies on the analogism between molecular structures assembly and real cyber-physical systems assembly (Sinha et al. 2012), and it was successfully applied and validated for industrial applications, including pressure recording devices, printing systems and wrapping machines (Verna et al. 2021; Alkan and Harrison 2019; Sinha 2014; Alkan et al. 2017; Alkan 2019).

Product complexity is defined by this model as follows:

$$C = C_1 + C_2 \cdot C_3, \quad (1)$$

where C_1 is handling complexity of parts, C_2 is the complexity of connections and liaisons between parts, and C_3 is the topological, or architectural, complexity. In detail, C_1 is obtained by considering each constituent part's complexity, according to Eq. (2):

$$C_1 = \sum_{p=1}^{N} \gamma_p, \quad (2)$$

where:

- N is the total number of parts that compose a product;
- γ_p is the handling complexity of a generic part p. The parameter γ_p is related to the technical difficulty of handling and interacting with part p under isolated conditions.

C_2 is the complexity of connections and liaisons between parts and is calculated from the complexities of pairwise connections existing in the product structure, according to Eq. (3):

$$C_2 = \sum_{p=1}^{N-1} \sum_{r=p+1}^{N} \varphi_{pr} \cdot a_{pr}, \quad (3)$$

where:

- φ_{pr} is the complexity in achieving a connection between parts p and r.
- a_{pr} is the $(p, r)^{\text{th}}$ entry of the binary adjacency matrix ***AM*** of the product. ***AM*** is a symmetric matrix of size NxN where each element designates the existence of an assembly liaison or connection between two components. a_{pr} can assume two values: 1 if there is a connection between p and r and 0 otherwise.

C_3 is the topological complexity and represents the complexity related to the architectural pattern of the assembled product, as follows:

$$C_3 = \frac{E_{AM}}{N} = \frac{\sum_{q=1}^{N} \delta_q}{N}, \quad (4)$$

where:

- N is the total number of parts;
- E_{AM} is the matrix (or graph) energy of the adjacency matrix, which is designated by the sum of the corresponding singular values δ_q of ***AM***. The singular values are the absolute

eigenvalues of the matrix $\boldsymbol{AM}$, as it is a symmetric matrix of size NxN with diagonal elements equal to zero (Li, Shi, and Gutman 2012; Sinha 2014).

Three different regimes can be identified according to the value of E_{AM} (Li, Shi, and Gutman 2012): hyperenergetic regimes when $C_3 \geq 2(1-1/N)$, hypoenergetic regimes when $C_3 < 1$ and transition regimes when $1 \leq C_3 < 2(1-1/N)$. Note that for hyperenergetic regimes, C_3 can be approximated to 2 when N is sufficiently large. Thus, C_3 increases as the system topology shifts from centralised to more distributed architectures (Sinha 2014).

This approach to assessing product complexity was applied to different industrial applications in the electronics, electromechanical, and aerospace industries. Such a complexity model may be easily integrated into the virtual space of a DT for the development and implementation of assembly/disassembly and quality control models of product variants.

4.3 Assembled and disassembled product variants

As mentioned in Section 4.1, the twelve different variants of molecular structures manually assembled and disassembled are characterized by different ball and stick variety and quantities and different levels of product structural complexity, as specified in Table 1.

Table 1. Structural characteristics and complexity of the product variants used in the experiments.

Product variant label	Chemical formula	Atoms (balls)	Bonds (sticks)	Rigid sticks	Flexible sticks	E_{AM}	C_1	C_2	C_3	C
V_1	C_2H_4	6	5	4	1	6.00	1.72	4.67	1.00	6.40
V_2	$C_2H_{11}NO_2S$	20	19	18	1	20.95	5.75	17.52	1.05	24.10
V_3	$C_2H_{17}NO_6$	44	49	42	7	52.41	12.65	45.55	1.19	66.90
V_4	$C_{25}H_{37}NO_4$	67	68	62	6	75.85	19.26	62.90	1.13	90.47
V_5	$C_{33}H_{46}O_5$	84	85	76	9	89.79	24.14	78.75	1.07	108.32
V_6	$C_{29}H_{41}N_7O_9$	86	88	78	10	97.74	24.72	81.58	1.14	117.44
V_7	$C_{37}H_{48}N_6O_5S_2$	98	101	88	13	111.46	28.17	93.76	1.14	134.81
V_8	$C_{46}H_{70}O$	117	117	106	11	123.29	33.63	108.28	1.05	147.73
V_9	$C_{47}H_{51}NO_{14}$	113	119	103	16	130.25	32.48	110.53	1.15	159.88
V_{10}	$C_{50}H_{64}N_2O_{12}$	128	133	119	14	145.73	36.79	123.21	1.14	177.07
V_{11}	$C_{43}H_{66}N_{12}O_{12}S_2$	135	137	123	14	151.33	38.80	126.88	1.12	181.04
V_{12}	$C_{44}H_{64}O_{24}$	132	135	126	9	151.58	37.94	124.64	1.15	181.07

According to the complexity model described in Section 4.2, for obtaining complexity C_1 (see Eq. (2)), γ_p is estimated using the standard time for handling the parts, in the hypothesis that all the parts are identical in terms of handling complexity. In a preliminary experimentation, 94 operators performed three time measurements, for a total of 282 measures recorded, and an average handling time of 2.80 s was obtained, used as a reasonable estimation of the standard handling time. In detail, standard handling time (t_h) is calculated as follows (Alkan 2019):

$$t_h = t_{lp} + t_{rp} + t_m + t_p + t_w, \quad (5)$$

where:

- t_{lp} is the time to localize the relevant box of the part;
- t_{rp} is the time to select the proper response;
- t_m is the time to move arm to pick positioning;
- t_p is the picking time;
- t_w is the time to move to work position.

For obtaining complexity C_2 (see Eq. (3)), φ_{pr} is estimated considering the standard completion time of connecting the parts, separately for rigid and flexible connectors. In detail, the standard completion time of joining (t_j) is obtained as follows (Alkan 2019):

$$t_j = t_{lc} + t_{rc} + t_o + t_{pl} + t_a + t_c, \quad (6)$$

where:

- t_{lc} is the time to localize the relevant box of the connection holes;
- t_{rc} is the time to select the proper response;
- t_o is the time to orient and position the parts and connector;
- t_{pl} is the time for placing connectors to both parts;
- t_a is the time to adjust the connections;
- t_c is the time to final check.

Also in this case, 280 measures were performed by 94 operators and 8.95 s was obtained as the standard completion time for rigid connectors and 9.75 s for flexible connectors. Then, a normalization based on the maximum value (i.e., 9.75 s) was performed to derive the final values of complexities, which specifically are $\gamma_p = 0.287$, $\varphi_{pr}(rigid) = 0.918$, $\varphi_{pr}(flexible) = 1$.

It has to be noted that the abovementioned times were acquired in a series of short experiments conducted before the assembly/disassembly of the product variants by randomizing the trials to minimize the learning effects.

Then, by applying Eqs. (2) and (3), complexity values C_1 and C_2 reported in Table 1 are obtained.

In order to calculate complexity C_3, the adjacency matrix was built for each product variant, see e.g. Fig. 4 for structure V_3, and the corresponding singular values were derived according to the Singular Value Decomposition (SVD) (Sinha 2014).

	C1	O1	O2	C2	C3	C4	C5	C6	C7	C8	C9	N	C10	C11	C12	O3	C13	O4	C14	C15	C16	C17	C18	C19	O5	O6	C20	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12	H13	H14	H15	H16	H17
C1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O2	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C2	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C3	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C4	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C5	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
C6	0	0	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C7	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C8	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0
C9	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0
N	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C10	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0
C11	0	0	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
C12	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
O3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
C17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
C19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
H1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H3	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H4	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H5	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H6	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H7	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H8	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H9	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H10	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H11	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H17	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 4. Adjacency matrix of structure V_3 (layered architecture). Grey cells represent a connection between the part

5 Productivity and quality performance models in assembly/disassembly process

As mentioned above, assembly and disassembly times were recorded during the experiments to measure the productivity of the processes. Total assembly and disassembly times (T_t) are recorded in the experiments as follows:

$$T_t = T_p + T_m + T_e + T_c + T_r, \tag{7}$$

where T_p is the perception time, T_m is the mental decision time, T_e is the action execution time, T_c is the time for in-process quality control, T_r is the rework time.

In addition, the number of in-process and offline defects was recorded for assembly, and the total number of defects was counted for disassembly. The experimental data were collected and then statistically analysed to identify and delete any possible outliers. The exclusion rule used was the Modified Interquartile Range Method, which is widely recognized as a practical and effective method for identifying outliers taking into account sample size (Barbato et al. 2011). Data on times and defects (average values and standard deviations obtained after outliers' detection) are presented separately for assembly and disassembly in Tables 2 and 3, respectively.

Table 2. Productivity and quality data for assembly obtained from the experimental campaign.

	Assembly time (s)		Assembly in-process defects (-)		Assembly offline defects (-)		Assembly total defects (-)	
Product variant label	Average value	Std. dev.	Average value	Std. dev.	Average value	Std. dev.	Average value	Std. dev.
V_1	0.46	0.13	0.02	0.14	0.00	0.00	0.02	0.14
V_2	3.38	1.56	0.40	0.72	0.17	0.86	0.58	1.19
V_3	10.17	3.28	1.33	1.65	0.90	2.01	2.23	2.79
V_4	12.86	5.30	1.88	5.83	0.96	3.00	2.85	6.52
V_5	18.78	5.52	3.29	5.14	3.37	6.80	6.65	8.62
V_6	20.38	7.02	2.65	3.22	2.75	6.81	5.40	8.04
V_7	22.31	7.15	2.88	5.81	2.67	4.91	5.56	7.62
V_8	22.60	7.56	3.56	2.79	5.33	9.35	8.88	10.63
V_9	31.26	8.97	1.98	2.06	5.73	13.60	7.71	14.19
V_{10}	36.76	10.70	4.23	6.59	6.21	11.77	10.44	13.74
V_{11}	45.38	16.29	6.25	6.84	12.77	16.49	19.02	18.63
V_{12}	44.16	17.13	9.44	12.16	16.04	18.19	25.48	22.10

Table 3. Productivity and quality data for disassembly obtained from the experimental campaign.

	Disassembly time (s)		Disassembly total defects (-)	
Product variant label	Average value	Std. dev.	Average value	Std. dev.
V_1	0.30	0.11	0.04	0.28
V_2	1.63	0.50	0.04	0.28
V_3	3.28	0.87	0.27	0.66
V_4	4.36	0.96	0.56	0.98
V_5	5.34	1.30	0.60	0.96
V_6	6.02	1.50	0.67	0.81
V_7	6.80	1.69	0.69	2.17
V_8	7.00	1.48	0.81	1.07
V_9	7.87	2.32	1.08	1.23
V_{10}	8.49	2.21	1.58	2.40
V_{11}	8.75	1.79	1.65	1.94
V_{12}	9.23	2.24	1.67	2.17

The results show that the variability of times and defects is low for product variants with low complexity. Indeed, when assembling/disassembling less complex products, humans can more easily identify the optimal method, errors and rework are unlikely, and the time for cognitive processing is short. On the other hand, the time required for cognitive processing and subsequent reprocessing increases with increasing complexity, resulting in greater variability in the data (see Tables 2 and 3). Productivity and quality data of assembly (see Table 2) and disassembly (see Table 3) are related to the complexity of product variants (see Table 1) in order to model the functions relating them for prediction and optimization purposes. Average values of all operators are considered for the analysis since the factor operator was not significant at a confidence level of 95% by performing the Analysis of Variance (the null hypothesis is that the factor is not statistically significant, i.e., all means are equal). In detail, the *p*-value of "operator factor" when considering the response assembly times is 0.514, for assembly in-process defects, offline defects and total defects is, respectively, 0.761, 0.410

and 0.399; for disassembly time, the *p*-value is 0.094 and disassembly total defects is 0.911. Fig. 5(a) represents the two-term power curve fitting relating assembly time and product variant complexity. This is the best fitting model compared to various models defining the relationship between assembly time and product complexity, considering the goodness of fit statistics and residual analysis (Montgomery, Runger, and Hubele 2010). Fig. 5(b) illustrates the best fitting model for the relationship between disassembly time and the complexity of product variants. In this case, disassembly time increases linearly as product variant complexity increases, according to the function shown in Fig. 5(b). The 95% confidence and prediction intervals are represented on the plots, showing that the regression lines follow the curvature in the points closely and no systematic deviations from the fitted line appear (see Fig. 5(a) and 5(b)).

Confidence intervals of the prediction are calculated as follows:

$$\hat{y}_j \pm t_{1-\frac{\alpha}{2},N-P} SE(Fit), \tag{8}$$

where $\hat{y}_j$ is the predicted value of the *j*-th response, $t_{1-\frac{\alpha}{2},N-P}$ is the value of the Student's *t* distribution with α level of significance and *N*-*P* degrees of freedom (where *N* is the total number of observations and *P* the number of free parameters) and $SE(Fit)$ is the standard error of the fit.

Prediction intervals are calculated as follows:

$$\hat{y}_j \pm t_{1-\frac{\alpha}{2},N-P} \sqrt{[SE(Fit)]^2 + S^2}, \tag{9}$$

where *S* is the standard error of the estimate, also known as the standard error of the regression, derived from the sum of the squared residuals *RSS*, the number of observations *N* and the number of free parameters *P*, i.e. $S = \sqrt{\frac{RSS}{N-P}}$ (Bates and Watts 1988). Calculations were performed using the software Minitab®.

The statistical significance of the parameter estimate is then assessed by analysing the 95% confidence intervals for the parameters, calculated from the corresponding Standard Errors (SE) reported in Table 3. The parameter is statistically significant if the range excludes the null hypothesis value, *i.e.* zero in the case of parameters of power-law regression (Seber and Wild 1989; Bates and Watts 1988). The parameter estimates in both regression curves are verified to be statistically significant as their 95% confidence intervals - i.e. (0.0003, 0.0402) for *a* and (1.3212, 2.2817) for *b* in the assembly time model, and (0.0481, 0.0503) for *a* in the disassembly time model - do not include zero. Residual plots of the regression models for assembly and disassembly times are represented in Fig. 5 (c) and (d), respectively. The visual analysis of the residuals was accompanied by a statistical normality test, the Anderson-Darling test, where the null hypothesis is that residuals follow a normal distribution, and the alternative hypothesis is that the data do not follow a normal distribution. The Anderson-Darling goodness-of-fit statistic (AD) measures the area between the fitted line (based on the normal distribution) and the empirical distribution function (which is based on the data points). The Anderson-Darling statistic is a squared distance weighted more heavily in the tails of the distribution. For the residuals of the assembly time regression model shown in Fig. 5(a), the AD measure is 0.343 and the corresponding *p*-value is 0.425. Thus, since the *p*-value is greater than the 0.05 significance level, the decision is to fail to reject the null hypothesis of normality of residual distribution. Similarly, for the residuals of the disassembly time regression model shown in Fig. 5(b), the AD measure is 0.602 and the corresponding *p*-value is 0.272, leading to a non-rejection of the null hypothesis of normality of residual distribution. In Fig. 5(a) and (b), the *S* value is reported, i.e., the standard error of the regression, which is a measure of goodness of fit of the model and is used instead of R^2 for nonlinear models (Bates and Watts 1988). Overall, the residual analysis and the goodness of fit tests allowed the robustness and validity of the regression models presented in Fig. 5 (a) and (b) to be established.

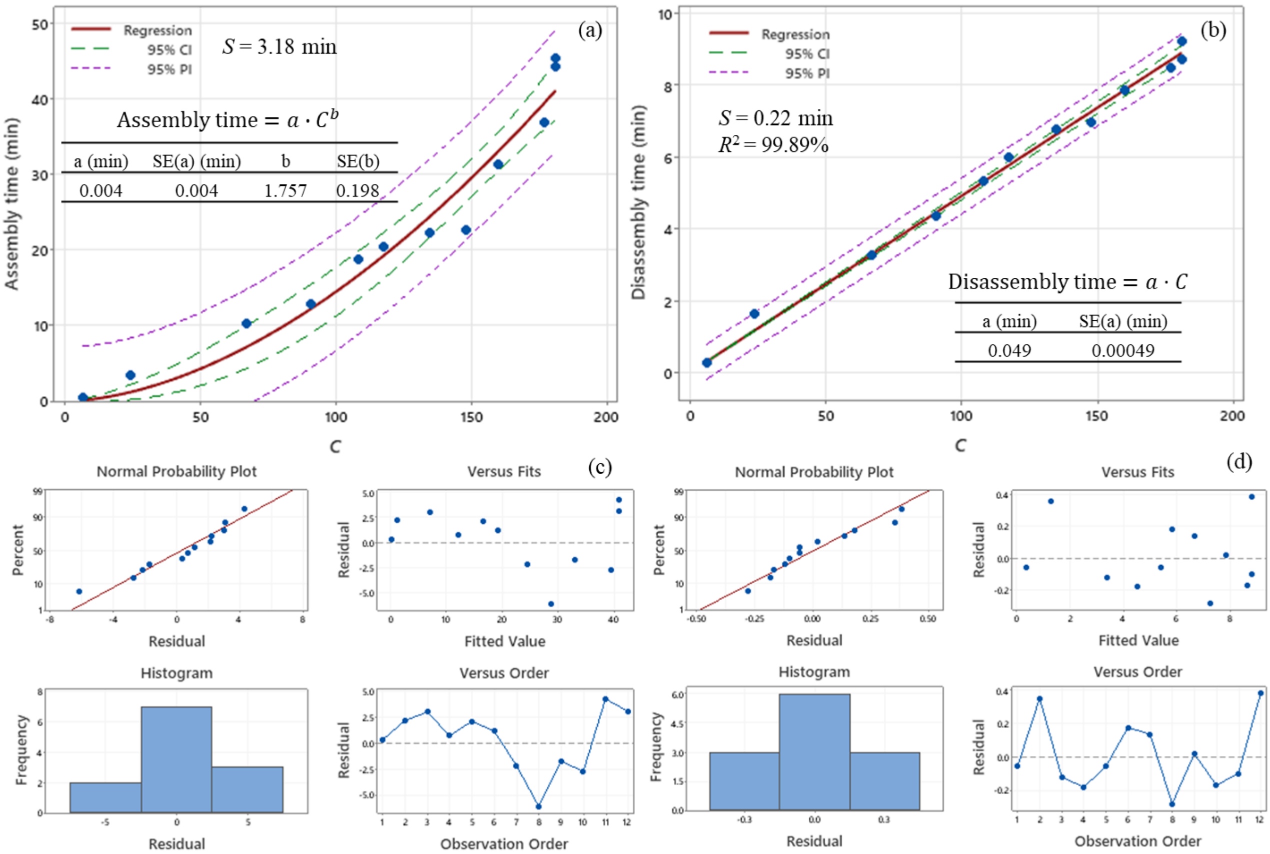

Figure 5. Assembly time vs product variants complexity: **(a)** experimental values, regression curve, 95% Confidence Intervals (CI) and Prediction Intervals (PI); and **(c)** residual plots. Disassembly time vs product variants complexity: **(b)** experimental values, regression curve, 95% Confidence Intervals (CI) and Prediction Intervals (PI); and **(d)** residual plots.

Fig. 5 shows a superlinear relationship between assembly time and product variant complexity. This implies that assembly operations require longer deliberation times, and thus more cognitive effort, as product complexity increases. On the other hand, disassembly times increase linearly with increasing product complexity, showing that cognitive effort remains constant regardless of the level of product complexity.

Regarding the relationship between defects and product variant complexity, Poisson regression models are adopted for the analysis since the responses are count data (Cameron and Trivedi 2013). Specifically, the link functions considered are the natural logarithm and the square root, and different models have been compared up to the third order of the predictor (i.e., the complexity). The selection of the most appropriate models is done according to Akaike's Corrected Information Criterion (AICc) and Bayesian Information Criterion (BIC), goodness-of-fit tests (Deviance and Pearson tests) and deviance residual plots (Cameron and Trivedi 2013). Deviance and Pearson tests assess whether the predicted number of events deviate from the observed number of events in a way that would not be predicted by the Poisson distribution. In detail, if the *p*-value is less than the significance level, the null hypothesis that the Poisson distribution provides a good fit can be rejected (Cameron and Trivedi 2013). According to the results, the most appropriate model to describe all the relationships between defects and complexity is the one using the square root link function, in the form $Y = (a \cdot C)^2$, where Y is the response, C is product complexity evaluated according to Eq. (1), and a is the regression coefficient. According to the results of the Poisson regression analysis reported in Table 4, the associations between responses and complexity are statistically significant, as the *p*-value of the coefficients is less than the significance level of 0.05 (the null hypothesis is that the coefficient is zero), the goodness-of-fit tests (Deviance and Pearson) indicate that the models fit the data well (i.e., the *p*-value is higher than the significance level of 0.05; the null hypothesis is that data follow a Poisson distribution), and very high values of the deviance R^2 are obtained.

Table 4. Main outputs from Poisson regression for assembly and disassembly defects vs product complexity. Models are in the form $Y = (a \cdot C)^2$.

	a	SE(a)	Coefficient *p*-value	Deviance R^2	Goodness-of-Fit Tests	
Assembly in-process defects vs Complexity	0.0137	0.0011	<0.0001	99.74%	Deviance Test *p*-value	0.900
					Pearson Test *p*-value	0.891
Assembly offline defects vs Complexity	0.0168	0.0011	<0.0001	99.71%	Deviance Test *p*-value	0.572

					Pearson Test *p*-value	0.538
Assembly total defects vs Complexity	0.0217	0.0011	<0.0001	99.79%	Deviance Test *p*-value	0.380
					Pearson Test *p*-value	0.331
Disassembly total defects vs Complexity	0.0069	0.0011	<0.0001	99.93%	Deviance Test *p*-value	0.368
					Pearson Test *p*-value	0.893

Fig. 6 shows the predicted curves of assembly and disassembly defects with the 95% confidence and prediction intervals. Confidence intervals of Poisson regression were calculated according to Eq. (8), while prediction intervals were derived empirically by bootstrapping due to the assumption of non-Gaussian distribution (Cameron and Trivedi 2013).

Finally, deviance residuals - used for Poisson regression (Cameron and Trivedi 2013) - are analysed, showing that the models are adequate and meet the assumptions of the analysis (see Fig. 7). In addition to the visual analysis of the deviance residual plots, the Anderson-Darling test was used to verify the normality of deviance residuals. The obtained *p*-values are 0.736 for deviance residuals of assembly in-process defects model, 0.066 for deviance residuals of assembly offline defects model, 0.464 for deviance residuals of assembly total defects model and 0.191 for deviance residuals of disassembly total defects model, respectively. Thus, the normality test led to the non-rejection of the hypothesis of normality of deviance residual distribution.

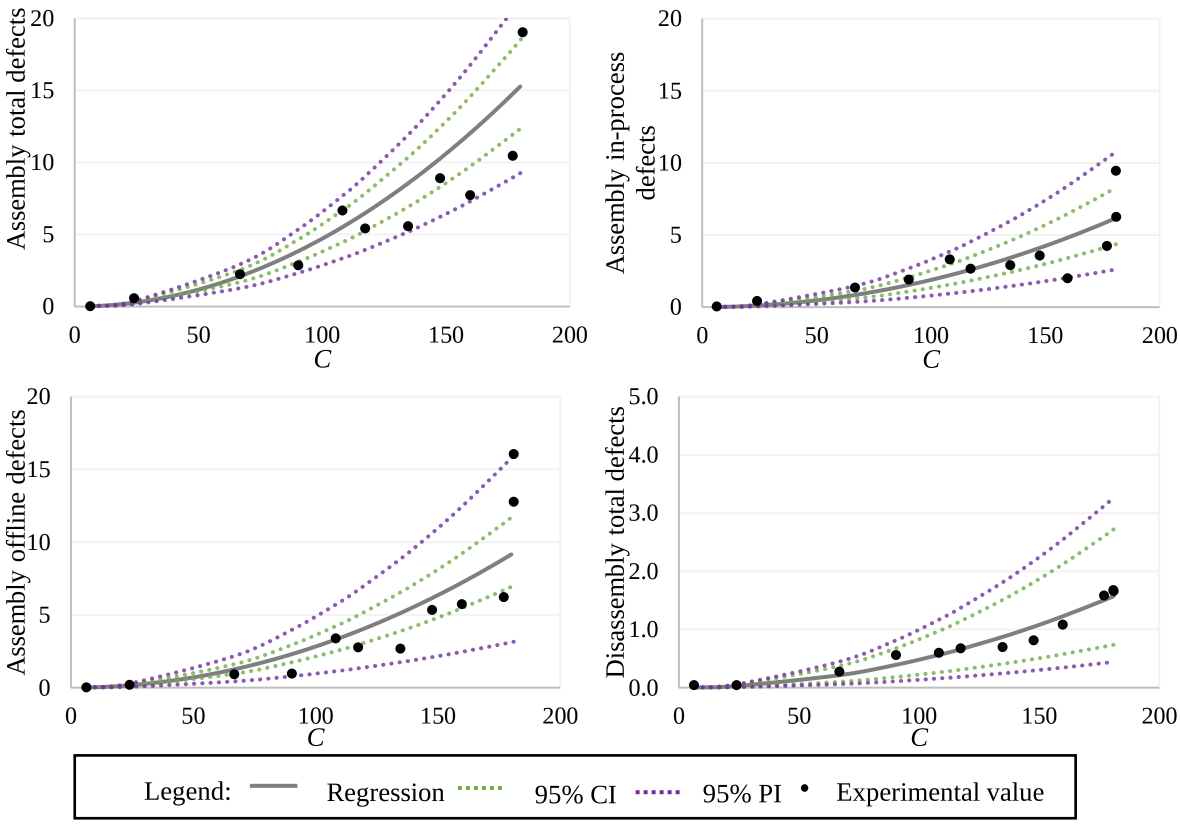

Figure 6. Regression models of assembly and disassembly defects vs product complexity with relevant 95% Confidence Intervals (CI) and Prediction Intervals (PI).

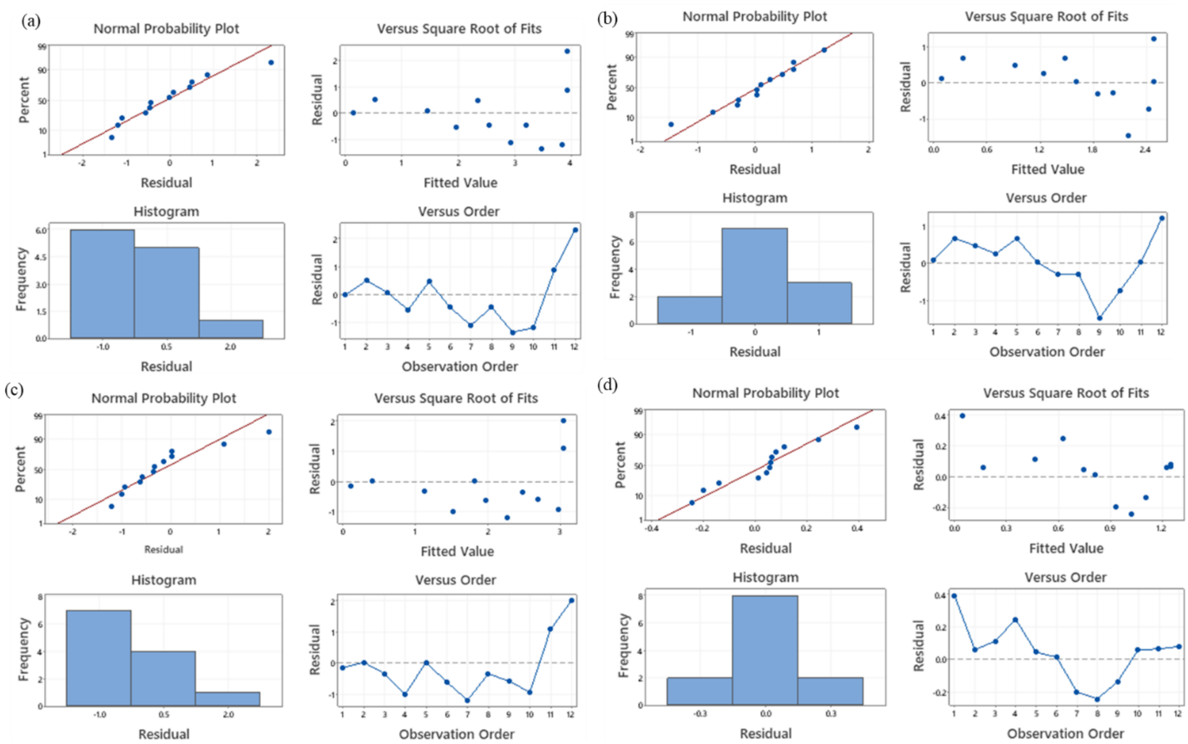

Figure 7. Deviance residual plots of regression models for **(a)** assembly total defects, **(b)** assembly in-process defects, **(c)** assembly offline defects, and **(d)** disassembly total defects vs product complexity.

Results obtained for defects occurring in assembly and disassembly processes show that the increase in product variant complexity leads to an increase in defects, following a nonlinear trend.

The Poisson regression models obtained (see Table 4 and Fig. 6) can be used by researchers and practitioners for prediction and optimization. For example, if a new product variant is introduced into the assembly/disassembly line, a defect prediction with a 95% confidence level can be obtained, as shown in Table 5 for a new product variant with complexity C=50.

Table 5. Assembly and disassembly defects predictions for a new product variant with complexity C=50.

Defect	**Fit**	**SE Fit**	**95% CI**
Assembly in-process defects	0.471	0.076	(0.333, 0.633)
Assembly offline defects	0.707	0.094	(0.535, 0.902)
Assembly total defects	1.178	0.121	(0.953, 1.427)
Disassembly total defects	0.120	0.039	(0.056, 0.207)

On the other hand, regression curves can be used for optimization purposes to achieve a certain level of product variant complexity that can lead to the achievement of a specific goal, e.g., minimizing defects or achieving a target value. Since the joint optimization must satisfy the requirements of all the selected responses, the more responses to be optimized, the more difficult it is to achieve high predictability due to the conflicting objectives (Galetto, Verna, and Genta 2021). In this study, the multi-response optimization is performed using a composite desirability function. The method uses an objective function called the desirability function and transforms an estimated response into a scale-free value (d_j) called desirability. In detail, for each j-th response, the index "d_j" represents the individual desirability and evaluates how the settings optimize the response. For instance, if the goal is to target the j-th response, the desirability d_j is calculated as:

(i) $d_j = ((\hat{y}_j - L_j)/(T_j - L_j))$ if $L_j \leq \hat{y}_j \leq T_j$

(ii) $d_j = ((U_j - \hat{y}_j)/(U_j - T_j))$ if $T_j \leq \hat{y}_j \leq U_j$

(iii) $d_j = 0$ if $\hat{y}_j < L_j$

(iv) $d_j = 0$ if $\hat{y}_j > U_j$

where $\hat{y}_j$ is the predicted value of the j-th response, U_j, L_j and T_j are the highest acceptable, the lowest acceptable and the target values for the j-th response, respectively. The desirability ranges from 0 to 1 (least to most desirable, respectively). A value of 1 represents the ideal case, while 0 indicates that one or more responses are outside their acceptable limits.

From the weighted geometric mean of the individual desirability of the responses, the composite desirability "D" is obtained, which assess how the settings optimize the total set of responses. The factor settings with the maximum total desirability are considered to be the optimal parameter conditions. In detail, the formula for the composite desirability, when the importance of each response is equal, is:

$$D = (d_1 \cdot d_2 \cdot d_3 \cdot \ldots \cdot d_n)^{1/n} \tag{10}$$

where n is the total number of responses optimized in the scenario and d_j represents the individual desirability for the j-th response (j=1, …, n).

For example, assuming that the assembly process should achieve 10 total defects and the disassembly process should achieve 1 disassembly defect, then the optimal complexity of the product variant should be around 144 (i.e., a product whose complexity is between V_7 and V_8), as shown in Table 6. In such a case, the composite desirability obtained according to Eq. (10) is D=98.59% since the desirability obtained for assembly total defects is $d_1 = 97.71\%$ and the desirability obtained for disassembly total defects is $d_2 = 99.48\%$.

Table 6. Optimal complexity to jointly achieve the target value of 10 assembly total defects and 1 disassembly total defect.

Optimal *C*	**Defects**	**Fit**	**SE Fit**	**95% CI**
144.01	Assembly total defects	9.77	1.00	(7.90, 11.84)
	Disassembly total defects	0.995	0.320	(0.466, 1.721)

Fig. 8 shows the correlation matrix of assembly/disassembly total defects and assembly/disassembly times, with indication of the Pearson correlation coefficient. All values are significant at 95% confidence level (p-value<0.002; the null hypothesis is that the correlation is not statistically significant). This analysis does not consider the distinction between in-process and offline defects as both are included in the total number of defects and therefore their trends reflect those of total defects.

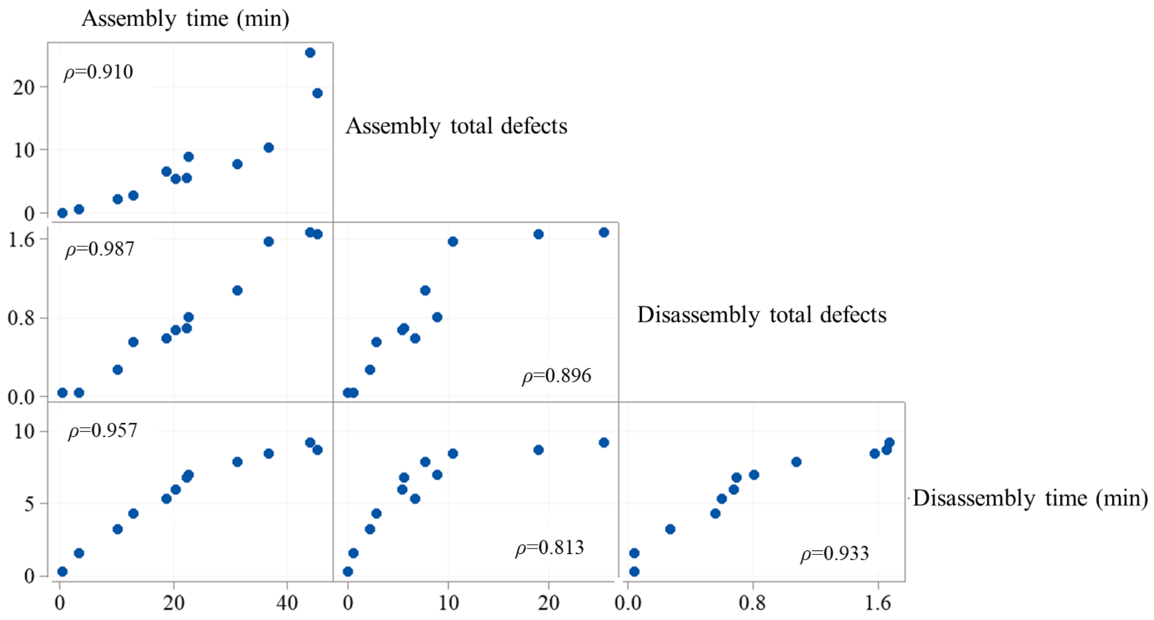

Figure 8. Correlation matrix containing pairwise Pearson correlation coefficients of assembly/disassembly times and total defects.

As shown in Fig. 8, the relationships between assembly/disassembly defects and times appear to follow a superlinear trend, as in the evident case of assembly total defects vs assembly time. Furthermore, when comparing assembly and disassembly processes, disassembly time (and disassembly total defects) appears to grow roughly linearly as the assembly time (and defects) increase, but after a certain threshold the growth is much slower or almost absent. This may be due to the fact that for low and medium-complex products, the cognitive effort required to assemble and disassemble the structure is comparable. However, for highly complex structures, the cognitive effort required for disassembly is much lower than that required for assembly. For instance, in the relationship between disassembly time and assembly time, from the V_8 product variant, the slope of the curve is much lower due to the lower cognitive effort required for disassembly compared to assembly, as shown in Fig. 9.

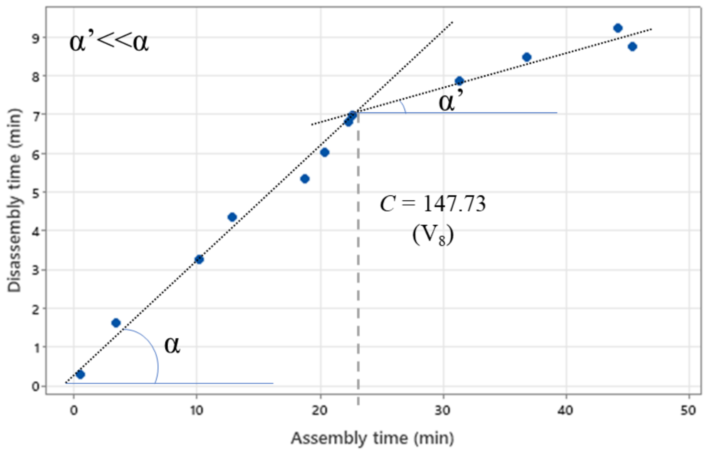

Figure 9. Disassembly time vs assembly time. The difference in cognitive effort influences the slope of the relationship at different levels of complexity.

From the correlation matrix (see Fig. 8), relationships between defects and times are derived using Poisson regression models, the results of which are shown in Table 7. For each of the models, the significance of the regression coefficients is checked (see coefficient *p*-value in Table 7), and goodness-of-fit tests (Deviance and Pearson) indicate that the models fit the data well (i.e., the *p*-value is higher than the 0.05 significance level), and very high values of deviance R^2 are obtained. The Deviance residuals of the models were analysed graphically, and the Anderson-Darling normality test was performed, which resulted in the non-rejection of the hypothesis of normality of deviance residuals distribution. Specifically, the *p*-values obtained from the Anderson-Darling test for the assembly total defects vs assembly times model are 0.116, 0.601 for the assembly total defects vs disassembly times model, 0.429 for the disassembly total defects vs assembly times model, and 0.854 for the disassembly total defects vs disassembly times model.

Table 7. Main outputs from Poisson regression for defects vs times in assembly and disassembly. Models are in the form $Y = (a \cdot C)^2$.

	a	SE(a)	Coefficient *p*-value	Deviance R^2	Goodness-of-Fit Tests	
Assembly total defects vs Assembly times	0.1063	0.0055	<0.0001	99.83%	Deviance Test *p*-value	0.596
					Pearson Test *p*-value	0.481
Assembly total defects vs disassembly times	0.4407	0.0226	<0.0001	99.81%	Deviance Test *p*-value	0.497
					Pearson Test *p*-value	0.477
Disassembly total defects vs Assembly times	0.0339	0.0055	<0.0001	99.69%	Deviance Test *p*-value	0.999
					Pearson Test *p*-value	0.760
Disassembly total defects vs Disassembly times	0.1406	0.0226	<0.0001	99.92%	Deviance Test *p*-value	1.000
					Pearson Test *p*-value	1.000

Furthermore, a nonlinear relationship between disassembly and assembly times is obtained, as shown in Fig. 10 (a). In Fig.10 (b), residual plots show that the power-law curve is an adequate model to describe the relationship between disassembly and assembly times (the *p*-value of the Anderson-Darling normality test on residuals is 0.922).

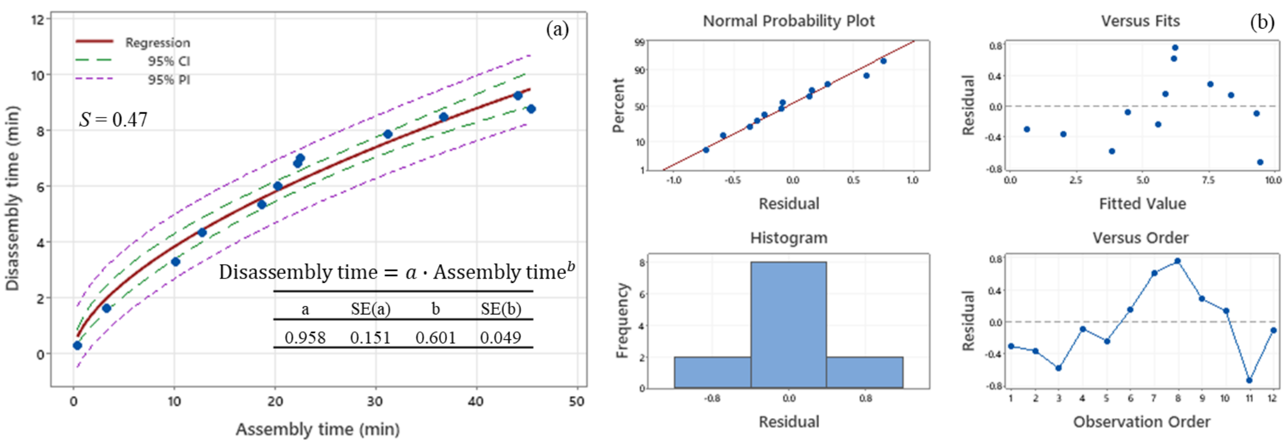

Figure 10. Disassembly time vs assembly time: **(a)** experimental values, regression curve, 95% confidence and prediction intervals, and **(b)** residual plots.

The developed experimental models of assembly/disassembly productivity and quality performance can be integrated into the virtual space of a DT, together with the assembly/disassembly and quality control models presented in Section 4. These models can be used as an infrastructure for constructing a twin model of the assembly/disassembly system to enhance predictive, diagnostic and monitoring capabilities. In particular, a DT tracking the process can be used to make predictions and optimizations according to the specific goal to be achieved, e.g., to identify optimal parameter settings for new product variants during the design phase and critical scenarios that may lead to undesirable quality and productivity performance.

6. Conclusions

Nowadays, companies are increasingly focusing on product customization, which involves changing production quickly to produce different product variants according to market demand. In this scenario, developing suitable approaches to perform quality control and monitoring of existing product variants, and to predict the quality of new variants is essential to increase business performance. To this end, this study investigates the integration of defect prediction models in Digital Twin for quality control and monitoring of product variants.

In order to develop predictive models of productivity and quality performance in both assembly and disassembly processes for different product variants, an extensive experimental campaign involving 104 operators was conducted. During the experiments, the operators were tasked with assembling and disassembling 12 different product variants, which were realized using a molecular kit to construct molecular structures typically considered in the scientific literature to effectively emulate the behaviour of real products (Sinha 2014). Data on assembly and disassembly times were recorded to measure the productivity of the processes. Additionally, the number of in-process and offline defects was collected for assembly, and the total number of defects was counted for the disassembly. The experimental results showed a superlinear relationship between assembly times and product variant complexity, while disassembly times increased linearly with increasing product complexity. Poisson regression models were used to model the relationships between defects and product variants. The regression models show that an increase in product variant complexity leads to an increase in defects, following a nonlinear trend. Moreover, the performance of assembly and disassembly processes is compared and related predictive models are derived. The results show a change in cognitive effort from assembly to disassembly when considering different product variants. Indeed, for simple product variants, the cognitive effort required to assemble and disassemble the structure is comparable, as shown by an almost linear trend. On the other hand, for complex product variants, the cognitive effort required for assembly is much higher than the cognitive effort required for disassembly, resulting in a change in the slope of the curve.

The proposed methodology for integrating predictive models into Digital Twin (DT) for quality control and monitoring of product variants has general validity and can be applied to different manufacturing contexts beyond the assembly and disassembly processes using molecular models, which were presented as an example in this paper. With the increasing focus on product customization, developing suitable approaches for quality control and monitoring of product variants is crucial for enhancing business performance. The developed prediction models, which are a virtual representation of the physical assembly and disassembly process, can be integrated into a DT for in-process quality control and to enhance the monitoring capabilities of the system. This integrated approach can ensure an accurate real-time prediction of assembly and disassembly performance, enabling defects to be avoided and/or preventive and corrective measures to be taken to minimise defect generation. Resource consumption (including cost, time, energy consumption) associated with quality control can therefore be significantly reduced, and the overall system quality may be improved. While this study is preliminary in nature, the authors believe that integrating defect generation models into a DT for quality control can lead to an overall increase in productivity and quality performance of the production system. Future work is planned to verify the effectiveness of the integration in the production environment by physically developing a digital twin for monitoring the assembly/disassembly of product variants.

Acknowledgements

The authors gratefully acknowledge M. Sc. Elena Martorana and M. Sc. Giammarco Arcieri for the fruitful collaboration on this project.

About the authors

Elisa Verna received the Master of Science Degree in Industrial Engineering and Management in 2016, and the PhD in Management, Production and Design in 2021 from Politecnico di Torino, Italy. She is currently Assistant Professor at the Department of Management and Production Engineering (DIGEP) of the Politecnico di Torino. She is Fellow of A.I.Te.M. (Associazione Italiana delle Tecnologie Manifatturiere) and E.N.B.I.S. (European Network for Business and Industrial Statistics). Her current research interests are Quality Engineering and Management, Statistical Process Control, and Innovative Production Systems. In particular, she is focusing on the study, implementation and planning of quality inspection procedures in manufacturing processes, and on the development of defect generation models in manual and human-robot collaborative assembly and disassembly processes.

Stefano Puttero is a second-year PhD Student in Management, Production and Design at Politecnico di Torino. He received the Master of Science Degree in "Engineering and Management" from Politecnico di Torino, Italy, in 2021. His main research interests are quality control and management in Human-Robot Collaboration and the development of innovative defect-generation models in the collaborative environment. His research also focuses on the central role of humans in the collaborative environment and the relationship between human well-being and the generation of defects.

Gianfranco Genta received the Master of Science degree in Mathematical Engineering from Politecnico di Torino, Italy, in 2005 and the PhD Degree in Metrology: Measuring Science and Technique from Politecnico di Torino in 2010. He is currently Associate Professor at the Department of Management and Production Engineering (DIGEP) of the Politecnico di Torino, where he teaches "Experimental Statistics and Mechanical Measurement" and "Industrial Quality Management". He is Research Affiliate of CIRP (The International Academy for Production Engineering) and Fellow of A.I.Te.M. (Associazione Italiana delle Tecnologie Manifatturiere). He is author and co-author of 3 books and more than 70 publications on national/international journals and conference proceedings. His current research focuses on Industrial Metrology, Quality Engineering and Experimental Data Analysis.

Maurizio Galetto received the Master of Science degree in Physics from University of Turin, Italy, in 1995 and the PhD Degree in Metrology: Measuring Science and Technique from Politecnico di Torino, Italy, in 2000. He is currently Head of Department and Full Professor at the Department of Management and Production Engineering (DIGEP) of the Politecnico di Torino, where he teaches "Quality Engineering" and "Experimental Statistics and Mechanical Measurement". He is Associate Member of CIRP (The International Academy for Production Engineering) and Fellow of A.I.Te.M. (Associazione Italiana delle Tecnologie Manifatturiere) and E.N.B.I.S. (European Network for Business and Industrial Statistics). He is Member of the Editorial Board of the scientific international journal Nanomanufacturing and Metrology and collaborates as referee for many international journals in the field of Industrial Engineering. He is author and co-author of 4 books and more than 100 published papers in scientific journals and international conference proceedings. His current research interests are in the areas of Quality Engineering, Statistical Process Control, Industrial Metrology and Production Systems. At present, he collaborates in some important research projects for public and private organisations.

References

Alkan, Bugra. 2019. "An Experimental Investigation on the Relationship between Perceived Assembly Complexity and Product Design Complexity." *International Journal on Interactive Design and Manufacturing (IJIDeM)* 13 (3): 1145–57. https://doi.org/10.1007/s12008-019-00556-9.

Alkan, Bugra, and Robert Harrison. 2019. "A Virtual Engineering Based Approach to Verify Structural Complexity of Component-Based Automation Systems in Early Design Phase." *Journal of Manufacturing Systems* 53: 18–31.

Alkan, Bugra, Daniel A Vera, Mussawar Ahmad, Bilal Ahmad, and Robert Harrison. 2018. "Complexity in Manufacturing Systems and Its Measures: A Literature Review." *European Journal of Industrial Engineering* 12 (1): 116–50.

Alkan, Bugra, Daniel Vera, Bilal Ahmad, and Robert Harrison. 2017. "A Method to Assess Assembly Complexity of Industrial Products in Early Design Phase." *IEEE Access* 6: 989–99.

Barbato, G., E. M. Barini, G. Genta, and R. Levi. 2011. "Features and Performance of Some Outlier Detection Methods." *Http://Dx.Doi.Org/10.1080/02664763.2010.545119* 38 (10): 2133–49. https://doi.org/10.1080/02664763.2010.545119.

Bates, Douglas M., and Donald G. Watts, eds. 1988. *Nonlinear Regression Analysis and Its Applications*. Hoboken, NJ, USA: John Wiley & Sons, Inc. https://doi.org/10.1002/9780470316757.

Bisgaard, Søren. 1997. "Designing Experiments for Tolerancing Assembled Products." *Technometrics* 39 (2): 142–52.

Cameron, A Colin, and Pravin K Trivedi. 2013. *Regression Analysis of Count Data*. Vol. 53. Cambridge university press.

Detzner, Alexander, and Martin Eigner. 2018. "A DIGITAL TWIN FOR ROOT CAUSE ANALYSIS AND PRODUCT QUALITY MONITORING." In , 1547–58. https://doi.org/10.21278/idc.2018.0418.

ElMaraghy, Waguih, Hoda ElMaraghy, Tetsuo Tomiyama, and Laszlo Monostori. 2012. "Complexity in Engineering Design and Manufacturing." *CIRP Annals* 61 (2): 793–814.

Franceschini, Fiorenzo, Maurizio Galetto, Gianfranco Genta, and Domenico A. Maisano. 2018. "Selection of Quality-Inspection Procedures for Short-Run Productions." *The International Journal of Advanced Manufacturing Technology* 99 (9–12): 2537–47. https://doi.org/10.1007/s00170-018-2648-8.

Galetto, Maurizio, Elisa Verna, and Gianfranco Genta. 2020. "Accurate Estimation of Prediction Models for Operator-Induced Defects in Assembly Manufacturing Processes." *Quality*

Engineering 32 (4): 595–613.

———. 2021. “Effect of Process Parameters on Parts Quality and Process Efficiency of Fused Deposition Modeling.” *Computers & Industrial Engineering* 156 (June): 107238. https://doi.org/10.1016/j.cie.2021.107238.

Genta, Gianfranco, Maurizio Galetto, and Fiorenzo Franceschini. 2018. “Product Complexity and Design of Inspection Strategies for Assembly Manufacturing Processes.” *International Journal of Production Research* 56 (11): 4056–66.

Grieves, Michael. 2014. “Digital Twin: Manufacturing Excellence through Virtual Factory Replication.” *White Paper* 1 (2014): 1–7.

Grieves, Michael, and John Vickers. 2017. “Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems.” In *Transdisciplinary Perspectives on Complex Systems*, 85–113. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-38756-7_4.

Gulivindala, Anil Kumar, M. V.A.Raju Bahubalendruni, S. S.Vara Prasad Varupala, and K. Sankaranarayanasamy. 2020. “A Heuristic Method with a Novel Stability Concept to Perform Parallel Assembly Sequence Planning by Subassembly Detection.” *Assembly Automation* 40 (5): 779–87. https://doi.org/10.1108/AA-01-2020-0017/FULL/PDF.

Guo, Jinkang, and Zhihan Lv. 2022. “Application of Digital Twins in Multiple Fields.” *Multimedia Tools and Applications* 81 (19): 26941–67. https://doi.org/10.1007/s11042-022-12536-5.

He, Zhen, Hao Hu, Min Zhang, and Xiaofei Yin. 2022. “Product Quality Prediction Based on Process Data: Taking the Gearbox of Company D as an Example.” *Quality Engineering* 34 (3): 409–22. https://doi.org/10.1080/08982112.2022.2063743.

Jones, David, Chris Snider, Aydin Nassehi, Jason Yon, and Ben Hicks. 2020. “Characterising the Digital Twin: A Systematic Literature Review.” *CIRP Journal of Manufacturing Science and Technology* 29 (May): 36–52. https://doi.org/10.1016/j.cirpj.2020.02.002.

Kaewunruen, Sakdirat, Panrawee Rungskunroch, and Joshua Welsh. 2018. “A Digital-Twin Evaluation of Net Zero Energy Building for Existing Buildings.” *Sustainability* 11 (1): 159. https://doi.org/10.3390/su11010159.

Kenett, Ron S, and Jacob Bortman. 2022. “The Digital Twin in Industry 4.0: A Wide-angle Perspective.” *Quality and Reliability Engineering International* 38 (3): 1357–66.

Ketelaere, Bart De, Bart Smeets, Pieter Verboven, Bart Nicolaï, and Wouter Saeys. 2022. “Digital Twins in Quality Engineering.” *Quality Engineering* 34 (3): 404–8. https://doi.org/10.1080/08982112.2022.2052731.

Li, Xueliang, Yongtang Shi, and Ivan Gutman. 2012. *Graph Energy*. Springer Science & Business

Media.

Liu, Mengnan, Shuiliang Fang, Huiyue Dong, and Cunzhi Xu. 2021. “Review of Digital Twin about Concepts, Technologies, and Industrial Applications.” *Journal of Manufacturing Systems* 58 (January): 346–61. https://doi.org/10.1016/j.jmsy.2020.06.017.

Ma, Yuanye, Hang Zhou, Honghong He, Guotao Jiao, and Sha Wei. 2019. “A Digital Twin-Based Approach for Quality Control and Optimization of Complex Product Assembly.” In *2019 International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM)*, 762–67. IEEE. https://doi.org/10.1109/AIAM48774.2019.00157.

Modoni, Gianfranco E., Benedetta Stampone, and Gianluca Trotta. 2022. “Application of the Digital Twin for in Process Monitoring of the Micro Injection Moulding Process Quality.” *Computers in Industry* 135 (February): 103568. https://doi.org/10.1016/j.compind.2021.103568.

Montgomery, DC, GC Runger, and NF Hubele. 2010. *Engineering Statistics*. New York: John Wiley & Sons Inc.

Raju Bahubalendruni, M. V.A., Bibhuti Bhusan Biswal, Manish Kumar, and Radharani Nayak. 2015. “Influence of Assembly Predicate Consideration on Optimal Assembly Sequence Generation.” *Assembly Automation* 35 (4): 309–16. https://doi.org/10.1108/AA-03-2015-022/FULL/PDF.

Seber, George A. F., and Chris J. Wild. 1989. *Nonlinear Regression*. New York: John Wiley & Sons.

Shivam, and Manish Gupta. 2022. “Quality Process Reengineering in Industry 4.0: A BPR Perspective.” *Quality Engineering*, July, 1–20. https://doi.org/10.1080/08982112.2022.2098044.

Sinha, Kaushik. 2014. “Structural Complexity and Its Implications for Design of Cyber-Physical Systems.” PhD dissertation, Engineering Systems Division, Massachusetts Institute of Technology.

Sinha, Kaushik, Olivier L de Weck, M Onishi, M Maurer, K Kirner, and U Lindemann. 2012. “Structural Complexity Metric for Engineered Complex Systems and Its Application.” In *Gain Competitive Advantage by Managing Complexity: Proceedings of the 14th International DSM Conference Kyoto, Japan*, 181–94.

Sony, Michael, Jiju Antony, and Jacqueline Ann Douglas. 2020. “Essential Ingredients for the Implementation of Quality 4.0: A Narrative Review of Literature and Future Directions for Research.” *The TQM Journal* 32 (4): 779–93.

VanDerHorn, Eric, and Sankaran Mahadevan. 2021. “Digital Twin: Generalization, Characterization and Implementation.” *Decision Support Systems* 145 (June): 113524. https://doi.org/10.1016/J.DSS.2021.113524.

Verna, Elisa, Gianfranco Genta, Maurizio Galetto, and Fiorenzo Franceschini. 2021. “Inspection

Planning by Defect Prediction Models and Inspection Strategy Maps.” *Production Engineering* 15 (6): 897–915. https://doi.org/10.1007%2Fs11740-021-01067-x.

———. 2022a. “Defect Prediction for Assembled Products: A Novel Model Based on the Structural Complexity Paradigm.” *International Journal of Advanced Manufacturing Technology* 120 (5–6): 3405–26.

———. 2022b. “Defects-per-Unit Control Chart for Assembled Products Based on Defect Prediction Models.” *International Journal of Advanced Manufacturing Technology* 119 (5–6): 2835–46. https://doi.org/10.1007/S00170-021-08157-1/FIGURES/8.

Yeratapally, Saikumar R., Patrick E. Leser, Jacob D. Hochhalter, William P. Leser, and Timothy J. Ruggles. 2020. “A Digital Twin Feasibility Study (Part I): Non-Deterministic Predictions of Fatigue Life in Aluminum Alloy 7075-T651 Using a Microstructure-Based Multi-Scale Model.” *Engineering Fracture Mechanics* 228 (April): 106888. https://doi.org/10.1016/j.engfracmech.2020.106888.

Zhao, Yuna, Dennis K J Lin, and Min-Qian Liu. 2021. “Designs for Order-of-Addition Experiments.” *Journal of Applied Statistics* 48 (8): 1475–95.

Zhu, Xiaoyang, and Yangjian Ji. 2022. “A Digital Twin-Based Multi-Objective Optimization Method for Technical Schemes in Process Industry.” *International Journal of Computer Integrated Manufacturing*, September, 1–26. https://doi.org/10.1080/0951192X.2022.2126013.