Digital twin-enabled automated anomaly detection and bottleneck identification in complex manufacturing systems using a multi-agent approach

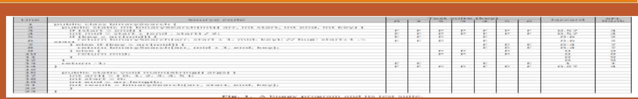
PRESENTED BY

ABIR EL

SEOUL NATIONAL UNIVERSITY OF SCIENCE AND TECHNOLOGY

DEPT. OF COMPUTER SCIENCE AND ENGINEERING

2023.05.15



4.1.1. Physical layer

- Physical layer refers to the operators working at the cryostorage company cryocarts and trolleys for the transportation of the cryomaterials on the shop floor , and cryomaterials and liquid nitrogen (LN2) for the preservation of cryomaterials.

- RFID tags are embedded to the containers and packaging items.

- In this study, three types of RFID readers have been implemented to the shop floor of the warehouse, including: (i) ‘shipping readers’ for reading dry shipper tags attached to cryostorage containers; (ii) ‘proximity readers’ for close up reads of bags and racks; and (iii) ‘cold 10 × 10 readers’ for reading cryoboxes containing up to 99 vials.

4.1.1. Physical layer

- Each vial slot in the cold readers has a unique antenna, enabling individual readings. All reader types can automatically update the location of stored items without requiring human intervention. Moreover, the physical processes carried out on the shop floor are captured, as demonstrated in Fig. 5.

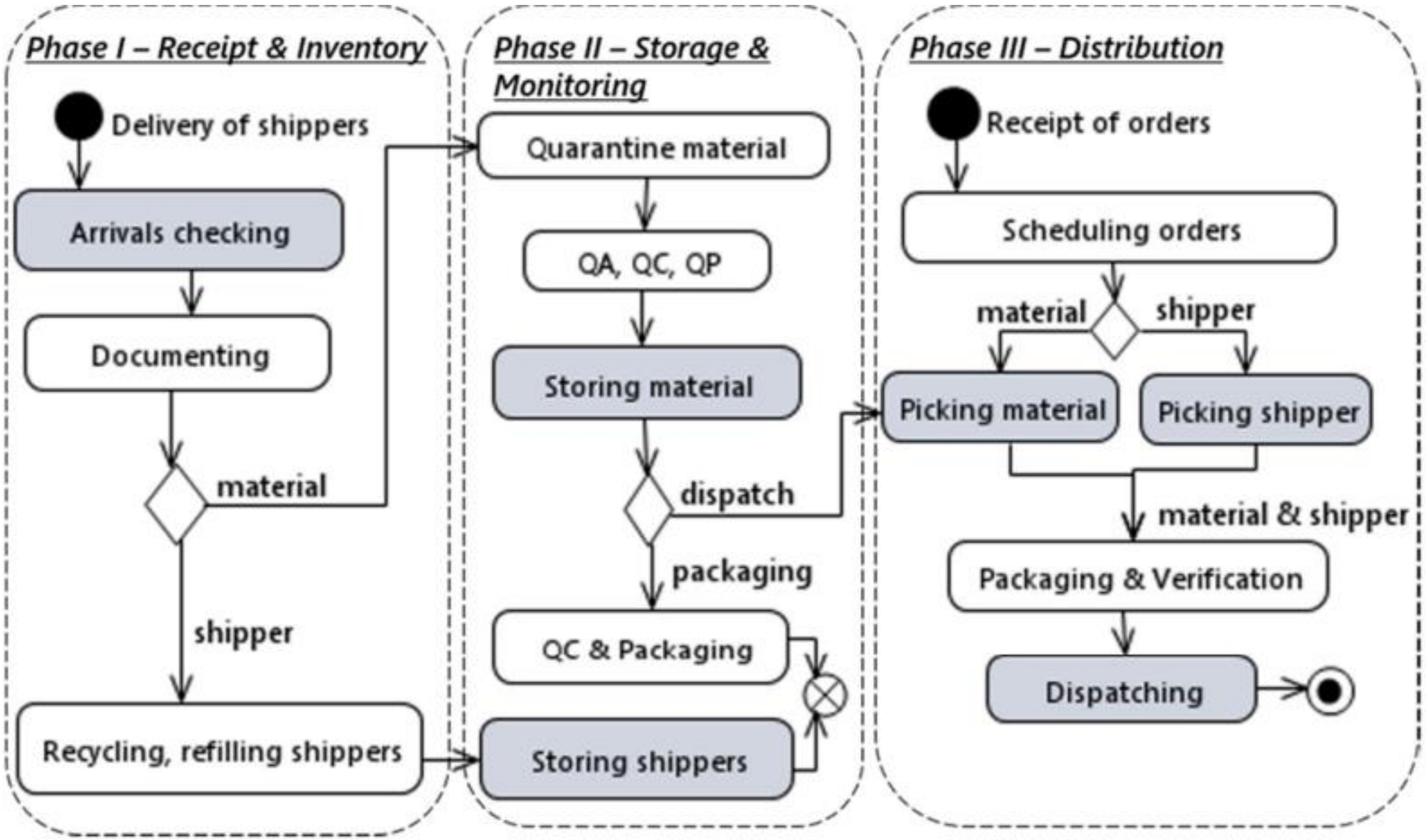


Fig. 5. Case study: UML activity diagram for the CGT cryogenic warehouse; RFID implementation (in blue).

4.1.2. Smart connection layer

- Smart connection layer acquires sensor and other input data associated with the operations occur at the shop floor of the cryogenic warehouse.
- The innovative cryogenic RFID system, integrated at the warehouse, has the ability to read, interpret and process RFID signals.
- The RFID system, used for the automated data capture, is deployed on MS Azure, a cloud computing service.
- The system driver is a component that reads RFID signals from a tag and produces an open standard file format and data interchange, JSON, which is then consumed by the RFID software. In this work, the Internet, the world's largest WAN, has been selected to transmit data between physical system and digital twin due to its wide availability and applicability. In terms of the software, the data captured by the RFID system is stored to MS SQL Server 2019 that is hosted on Azure Virtual Machines and transmitted using the Internet.

4.1.2. Smart connection layer

- To facilitate data communication among different software systems, a web server based on FLASK micro web framework, written in Python, was created.

- FLASK has an extension called Flask-RESTful that provides support to quickly building REST APIs for Create, Read, Update, and Delete (CRUD) endpoints. Hence, to retrieve the data to AnyLogic, MS SQL Server is connected to Google Sheets to read the RFID data from the database, by whitelisting the IP, creating 'Apps Script' project, creating a connection to MS SQL Server database, reading data from MS SQL Server database and writing data to Google Sheets.

- Google Drive API is also used to allow leverage Google Drive storage.

- AnyLogic software has Cloud APIs in Python that enable compatibility with other programs and processing JSON files.

4.1.2. Smart connection layer

- Other input data obtained from the enterprise system employed in the studied system is considered. The company's working hours is between 8:30 am and 16:30. The number of operators, working at the cryogenic facility.

- All data acquired in the smart connection layer is modelled at the micro-level agent. For the implementation of micro-level agents, a population of agents of the same type living in the same environment is created in AnyLogic software.

- These micro-agents have dynamic properties including movement speed (metre/second), location (X, Y, Z, rotation Z coordinates of Java type double), shape (2D/3D animation sketch) and recurrence update time. A function that returns the colour type value, using getFillColor () function, has been considered to get the fill colour of the (human, equipment and material) resources when simulation is animated in 2D.

4.1.3. Conversion layer

- Conversion layer transforms sensor data stored in the database to meaningful information for the health status of RFID tagged products.

- Anomalies identified in the database can be related to an increased number of deliveries and/or orders than it is expected on a daily basis, or delivery of wrong quantity of products, or increased time required for receiving, processing or dispatching cryomaterials compared to the nominal time required to complete these tasks.

- Anomalies are detected by the 'monitoring agent' at the exo-level agent that can then predict dynamically unplanned emergent bottlenecks related to human and equipment resources availability and inventory levels and storage space availability. Moreover, the root cause of the bottleneck is identified in terms of unstored products, queues of products waiting to be processed and any increase identified in the TH, LTs, WIP, HRU or ERU compared to average numbers observed during the normal daily operations of the cryogenic warehouse.

4.1.4. Cyber layer

- In this layer, information base, modelled at the micro-level agent, is deployed as a built-in fully integrated database for reading input data from the database table and writing simulation output. As discussed in the smart connection layer, the raw data collected from the RFID system is uploaded to AnyLogic software. In the cyber layer, this data is then processed using SQL queries to create timestamps and calculate the time required for performing various activities within the cryogenic warehouse. In AnyLogic, information base tables and views are created.

- Information base table is a collection of related data held in a structured format consisting of fields (i.e. columns) and rows. Each table has a column storing unique IDs of the table rows.

- Additional fields are the activities in which the RFID system has been implemented (e.g. arrivals checking, storing material, etc.), date and time stamps recorded after each tag is scanned, cycle time required for each activity to be carried out and users' UID.

4.1.4. Cyber layer

- Information base views, are relational tables representing a subset of data contained in the information base table. A view is computed dynamically from data in the table when access to that view is requested. In this work, one view has been created for each RFID activity. SQL queries for creating the views are developed in the ‘View definition’ field using the SELECT statement.

- The multi-agent simulation method discussed in Section 3.3 is employed to develop the cyber-twin model of the cryogenic warehouse. The multi-agent architecture of macro, exo, meso and micro level agents is developed.

- The ‘monitoring agent’ is modelled as a single agent type at the exo-level agent in the global manufacturing system, macro-level agent, of the cryogenic warehouse. The three manufacturing phases, including Phase I – Receipt & Inventory, Phase II – Storage & Monitoring, and Phase III – Distribution, are implemented as single type agents $\Phi_W = \{\Phi_1, \Phi_2, \Phi_3\}$ at the exolevel agent.

4.1.4. Cyber layer

- the stochastic phase space of Φ_W is Γ_W , the probabilities are p1 = p2 = p3 = 1/3 and can be obtained as:

$$\Gamma_W = p_1\Gamma_1 + p_2\Gamma_2 + p_3\Gamma_3$$

- It is noted that in this case study there are no repeated manufacturing modules, and hence the meso-level agent is not considered. Human and equipment resources are implemented as sub-sub-agents, X, at the micro-level agent with dynamic properties including population, working hours, breaks and shifts, movement speed, home location and 2D/3D animation shape and capacity.

- The collection of micro-level agents is described in a finite set X = {x1, x2, …, x9} modelling operators for receiving deliveries, general activities, shippers filling, QA, QC, QP, trolleys, cryocarts and cryotanks. Each sub-sub-agent contains parameters to capture the dynamic parametric operation of the corresponding agent and entire system.

4.1.4. Cyber layer

- With regard to the development of the multi-layer network the logical network (i.e. transition(s) from one activity to another) of the manufacturing system at the cryogenic warehouse.

- The parallel interactive activities in the three manufacturing phases that include highly interactive and manual handling processes are initiated once a delivery arrives at the company in Phase I and/or an order to dispatch cryoproducts to healthcare institutions is received in Phase III. After the shippers are delivered at the cryogenic warehouse, in Phase I, they are verified and documented (arrivals checking).

- The shippers are recycled and refilled, before they are stored in pallet racks in Phase II. The cryomaterials are initially stored in quarantine storages, checked in terms of policies and regulations and once approved, they are stored in the cryotanks. Phase III initiates once an order is received followed by the shipment planning and scheduling. After picking the right material and shipper from storage, and assign the material into the shipper, secondary packaging (if necessary), verification and dispatch are carried out. For the network of agents, let two sets of nodes x21 and x22 that represent micro-level agents, e.g., operators for general tasks,

4.1.4. Cyber layer

- interacting with each other. The state for each node set is represented by a canonical vector $\varepsilon 2_1$ and $\varepsilon 2_2$ as there are 20 operators for performing general tasks, the interactions between the agents can be expressed as:

$$\Phi 2^{\widetilde{\alpha\alpha}}_{\underset{\sim}{\beta\beta}} = \sum_{\underset{\sim}{i,j=1}}^{20} \sum_{i,j=1}^{20} w2_{12}\left(\widetilde{12}\right) \varepsilon 2^{\widetilde{\alpha\alpha}}_{\underset{\sim}{\beta\beta}}\left(12\widetilde{12}\right)$$

where $w2_{12}(\widetilde{12})$ is the intensity of the relationship between nodes $x2_1$ and $x2_2$; $\varepsilon 2^{\alpha}(1)$ and $\varepsilon 2^{\beta}(2)$ are the *αth* and βth components of the 1st and 2nd contravariant canonical vectors $\varepsilon 2_1$ and $\varepsilon 2_2$ in $\mathbb{R}^{20}$, respectively; and $\varepsilon 2^{\widetilde{\alpha\alpha}}_{\underset{\sim}{\beta\beta}}\left(12\widetilde{12}\right)$ is the fourth-order canonical basis in space $\mathbb{R}^{20\times 20\times 20\times 20}$.

4.1.5. Cognition layer

- Cognition layer is employed to transfer knowledge to the users to make appropriate decisions for maintaining or improving the performance and productivity of the cryogenic warehouse.

- Such knowledge obtained from the computational results includes the system TH, modelled at macro-level agent, time required for performing the several activities within the three manufacturing phases modelled at the exolevel agent, and WIP, HRU, ERU and space utilisation rates, levels of inventory size and availability of storage space modelled at the microlevel agent.

4.1.5. Cognition layer

- Additionally, the daily numbers of shippers delivered and dispatched and cryomaterials stored in the cryotanks, the space availability of cryotanks, and stock size of cryomaterials and consumables are considered.

- For utilisation rates of human and equipment resources, the billable hours over the eight working hours of the company are obtained from the simulation model.

- The performance and productivity of the cryogenic warehouse is continuously monitored and measured, while knowledge is obtained by running the simulation experiment with animation displayed.

- In the case of a disruptive event is diagnosed, at the exo-level agent, the user is warned about the abnormality in the system's performance through an alert that appears in the screen with the associated bottleneck to be highlighted.

4.1.6. Configuration layer

- The configuration layer is employed to automatically optimise the performance of the cryogenic warehouse by providing feedback to the smart connection, conversion and cognition layers.

- Self-optimisation is implemented at the macro-level agent of the cryogenic warehouse system.

- In the studied system, decision strategies including sourcing and procurement, risk mitigation and management, environment and sustainability strategies, are considered according to managers’ knowledge and experience, but also with the help of computational models where various simulation and optimisation scenarios can be executed.

- By realising the automated knowledge feedback from the cognition layer, actionable insights can be derived for improvement in the control and decision making.

4.1.6. Configuration layer

- Thus, relevant data analytics can be performed to make informed decisions considering dispatch planning, queue management to reduce WIP limits and queue sizes, resource planning, space layout planning and inventory control.

- In this study, the selected decision strategy proposes reallocation of operators within existing groups in order to handle bottlenecks identified in the conversion layer and increase system's flexibility, while maximising the number of deliveries completed, while minimising the WIP and excessive use of human resources.

- The self-optimisation is deployed in AnyLogic employing OptQuest® search engine that uses the metaheuristic algorithms of Scatter Search, Tabu Search and Neural Networks, combining them into a single search heuristic.

4.1.6. Configuration layer

- The results obtained from the optimisation experiment are used to modify the properties of the asset, (e.g. resource planning) and its environment (e.g. shop floor planning and control).

- In this study, the optimisation results, i.e. optimal values of operators required within each group, are automatically updated to the human resources parameters at the micro-level in the smart connection layer.

- Monitoring agent in the conversion layer is then updated based on the new reallocation of human resources and checks if the bottlenecks previously identified have been removed.

- Based on the updated parameters, the cyber-twin model in the cyber layer is then simulated and provides updated results visualised in the cognition layer.

- In the case of the bottlenecks remain or new are identified, a new self-optimisation for handling bottlenecks should be carried out.

4.1.6. Configuration layer

- Additional decision strategies can be explored finding the optimal number of equipment resources to improve the performance and productivity of the digital twin and, by extension, physical system.
- It can be also explored the optimal TH, initial stock size and storage space to improve the warehouse capacity and shipping speed, minimise the LTs of manufacturing phases and avoid the occurrence of queues, and emergence of bottlenecks.

4.2. Digital twin – cyber physical system model validation

- Validation of the simulation model and DT-CPS architecture is accomplished using real data obtained from the studied cryogenic warehouse.

- The validation of the model has been carried out at three stages. At Stage 1, the simulation model developed for the shop floor warehouse is validated against real data for the state without the DT-CPS and RFID system implementation (‘without DT-CPS’ scenario). This is important for later testing the validity of Stage 2 that builds on and extends the model of Stage 1.

- At Stage 2, the DT-CPS architecture is validated using real-time data collected from the RFID system for the normal operation of the warehouse (‘DT-CPS without anomaly’ scenario).

4.2. Digital twin – cyber physical system model validation

- This is also necessary for validating Stage 3 that builds on and extends the model of Stage 2. At Stage 3, anomalous values are captured in the RFID data due to disruptions that occur on the shop floor of the warehouse. The 'monitoring agent' is validated in terms of its capability to automatically detect these anomalies and capture their impact on the system's performance ('DT-CPS with anomaly' scenario).

- At Stage 1 ('without DT-CPS' scenario), the total number of daily orders (No D) and dispatches (Nd D) over time are obtained from the simulation model as illustrated in Fig. 6.

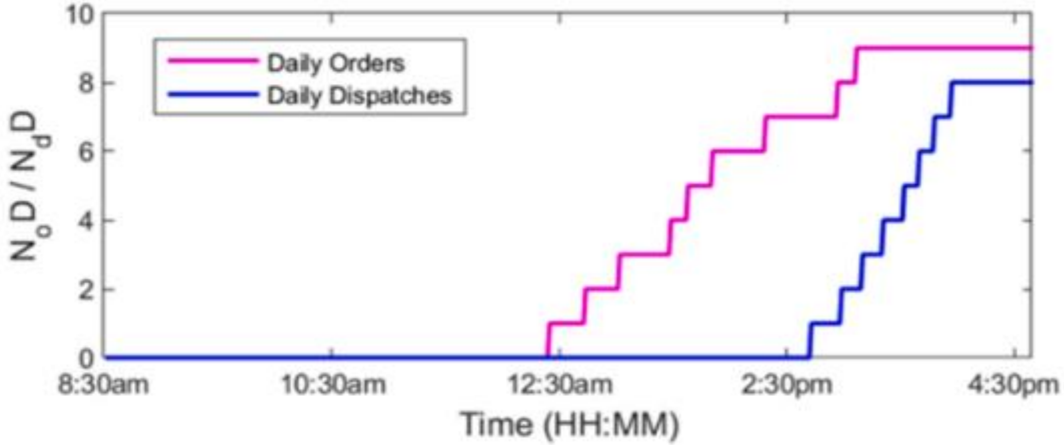


Fig. 6. Case study: simulation results for the total number of daily orders (N_o^D) and dispatches (N_d^D) for the 'without DT-CPS' scenario.

4.2. Digital twin – cyber physical system model validation

- The graph shows that the warehouse receives orders during the daily working hours between 8:30 am and 16:30. However, orders are dispatched between 14:30 and 16:30 when the trucks are available at the company. Moreover, real data for the cumulative total of monthly dispatches, provided by the studied company for the validation, is compared to the results obtained from the simulation model for an eight-month period, between January and August, as viewed in Fig. 7

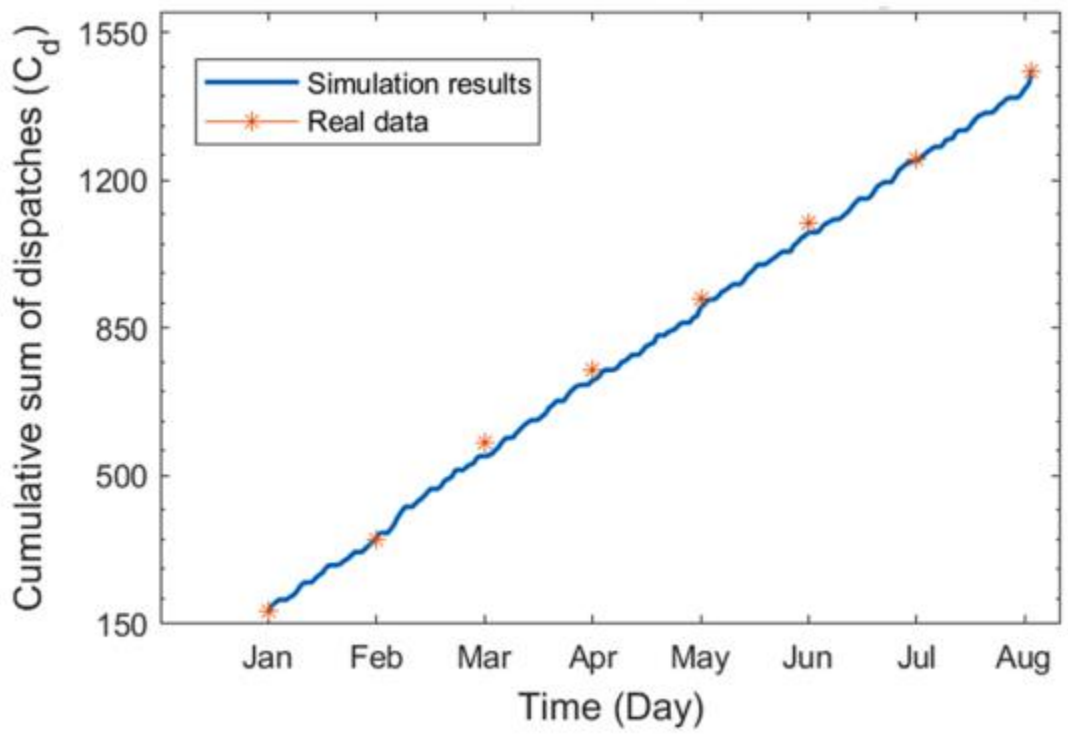


Fig. 7. Case study: simulation results against real data for the cumulative total of dispatches (C_d) for an eight-month period for the 'without DT-CPS' scenario.

4.2. Digital twin – cyber physical system model validation

- The simulation time has been set accordingly. The number of monthly dispatches obtained from the simulation results fall into the monthly ranges provided by the company, which are 150 – 185. The graph for the cumulative monthly dispatches shows the accuracy of measurements with a highly representative comparison between the simulation model and the real data, having an average percentage error of 0.81% in terms of the company performance.

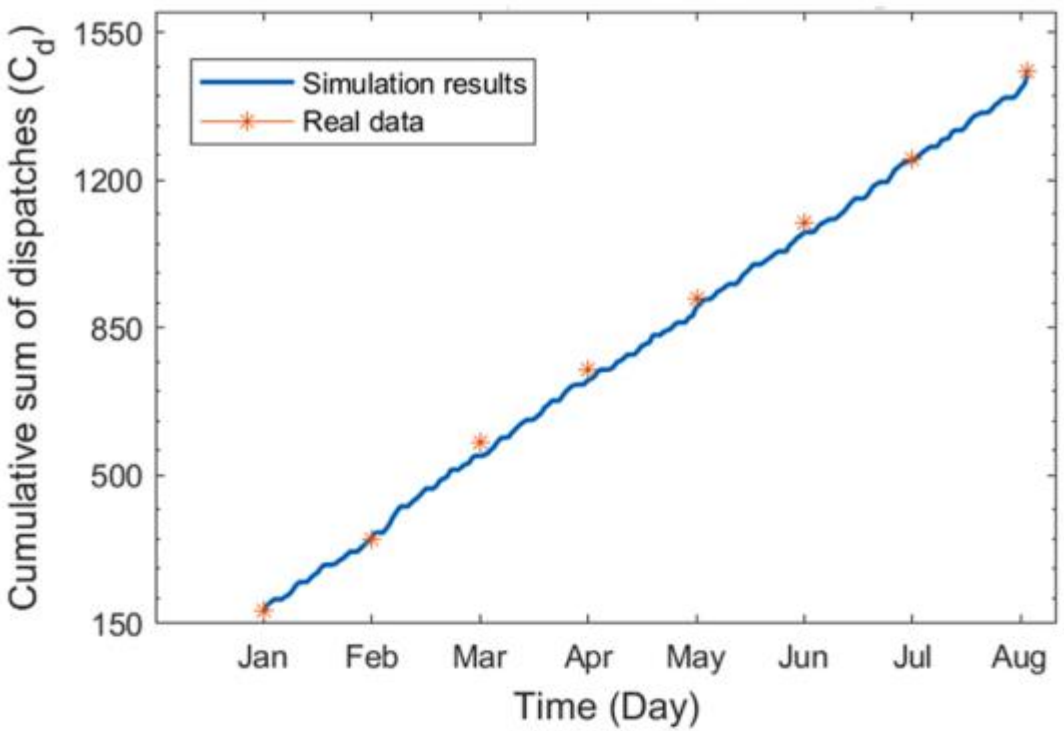


Fig. 7. Case study: simulation results against real data for the cumulative total of dispatches (C_d) for an eight-month period for the 'without DT-CPS' scenario.

4.2. Digital twin – cyber physical system model validation

- At Stage 2 ('DT-CPS without anomaly' scenario), after validating successfully the simulation model for the 'without DT-CPS' state, the proposed DT-CPS architecture is validated using RFID data collected under the normal operation of the cryogenic warehouse. Real data on the RFID cycle times for a six-week period has been collected from the shop floor of the company. The average cycle times taken for each test procedure carried out within a trial are summarised in Table 1.

Table 1
Case study: real-time RFID input data – cycle times for the 'DT-CPS without anomaly' scenario.

Activity with RFID	Test Procedures							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Arrivals checking (seconds)	9	9	9	15	15	15	15	15
Storing material (minutes)	2	2	2	4	4	4	4	4
Picking material (minutes)	1	1	3	3	3	1	3	1
Dispatching (seconds)	9	9	15	15	15	9	15	9

4.2. Digital twin – cyber physical system model validation

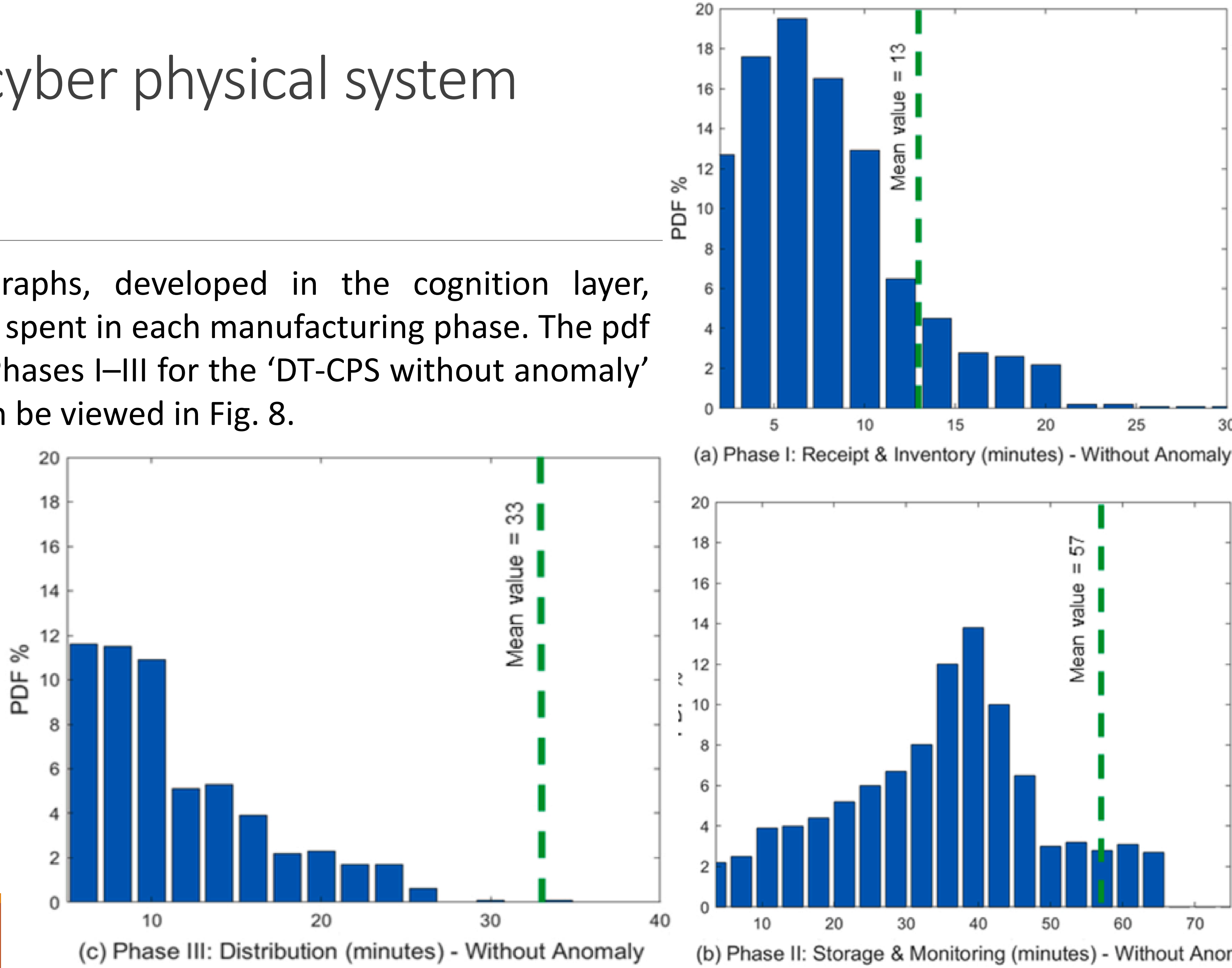
- According to this data, obtained from the database, the histogram graphs have been developed to calculate the Probability Density Function (pdf) for the arrivals checking, storing material, picking material and dispatching cycle times. The mean values of the corresponding pdf graphs are λArrivals checking = 12.81sec, λStoring material = 3.3min, λPicking material = 2.05min and λDispatching = 12.3sec.

Table 1
Case study: real-time RFID input data – cycle times for the 'DT-CPS without anomaly' scenario.

Activity with RFID	Test Procedures							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Arrivals checking (seconds)	9	9	9	15	15	15	15	15
Storing material (minutes)	2	2	2	4	4	4	4	4
Picking material (minutes)	1	1	3	3	3	1	3	1
Dispatching (seconds)	9	9	15	15	15	9	15	9

4.2. Digital twin – cyber physical system model validation

- Moreover, the histogram graphs, developed in the cognition layer, calculate the pdf for the time spent in each manufacturing phase. The pdf graphs for the lead times in Phases I–III for the ‘DT-CPS without anomaly’ scenario and mean values can be viewed in Fig. 8.



4.2. Digital twin – cyber physical system model validation

- According to the pdf graphs, it is observed that 95.5% of deliveries are being received and documented in less than 20 min, see Fig. 8(a). The corresponding pdf graph has a Poisson distribution with λPhase I− Without Anomaly = 13min. Additionally, the pdf for the storage and monitoring lead time in the 'DT-CPS without anomaly' state shows that 89% of the products are being stored in about 50 min, and only 11% of the products in between 50 and 70 min, see Fig. 8(b). The corresponding pdf graph has a Poisson distribution with λPhase II− Without Anomaly = 57min.

- Finally, in Phase III, the pdf graph for the distribution lead time has an exponential distribution, as illustrated in Fig. 8(c). The average time needed to complete a product dispatch is much less as 97% of the orders are being dispatched in less than 30 min. The corresponding pdf graph has an exponential distribution with λPhase III− Without Anomaly = 0.1158min. The pdf graphs and analysis are included as they will be used for the validation of the next stage for the 'DT-CPS with anomaly' scenario.

4.2. Digital twin – cyber physical system model validation

- At Stage 3 ('DT-CPS with anomaly' scenario), an anomaly detection scenario with ten test procedures has been considered to validate that the proposed 'monitoring agent' can detect anomalous values in input RFID data and realise their impacts to the system performance. In these trials, the cycle times for picking materials from storage and assigning them to shippers for dispatch have been deliberately increased compared to the normal operations of the system. The average cycle times taken for each test procedure carried out within a trial are summarised in Table 2.

Table 2

Case study: RFID input data of disruption scenario – cycle times for the 'DT-CPS with anomaly' scenario.

Activity with RFID	Test Procedures									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Picking Material (minutes)	8	10	10	12	25	27	15	17	20	22

4.2. Digital twin – cyber physical system model validation

- The cycle time for 'Picking material' under normal operating conditions is between 1 and 3 min, as seen in Table A3. The cryogenic warehouse carried out these scenarios and collected the data using the RFID system. The cycle time distributions are implemented to the simulation model at micro-level agent and the mean value of the corresponding pdf graph isλPicking material = 16.8min. Comparing the cycle times in Table 2 and these obtained from the pdf graph for 'Picking material, excellent agreement is found, with an average error of 1.21%.

Table 2
Case study: RFID input data of disruption scenario – cycle times for the 'DT-CPS with anomaly' scenario.

Activity with RFID	Test Procedures									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Picking Material (minutes)	8	10	10	12	25	27	15	17	20	22

4.2. Digital twin – cyber physical system model validation

- The ‘Anomaly Detected’ state in AnyLogic statechart is activated, validating the ability of the ‘monitoring agent’ to detect anomalous values in input sensor data. Additionally, average lead times for the three manufacturing phases, for the DT-CPS with anomaly’ scenario are obtained. The results are compared with the corresponding computational data obtained for the ‘DT-CPS without anomaly’ scenario (Stage 2) and the average lead times are summarised in Table 3.

Table 3

Case study: average lead times for the ‘DT-CPS without anomaly’ and ‘DT-CPS with anomaly’ scenarios.

Lead time for Phases I–III	‘Without anomaly’ (minutes)	‘With anomaly’ (minutes)	Increase (%)
Receipt & Inventory	13	14	7.7
Storage & Monitoring	57	65	14
Distribution	33	47	42.4

4.2. Digital twin – cyber physical system model validation

- Similarly, the average human resource utilisation rates for the three manufacturing phases for the 'without anomaly' and 'with anomaly' scenarios are obtained, as seen in Table 4. From the computational results in Tables 3 and 4, obtained in the cognition layer, it is seen that the model can capture the impacts of these anomalies to the operation of the manufacturing phase (i.e. Phase III - Distribution) and to entire system in terms of lead times and human resource utilisation rates.

Table 4
Case study: average human resource utilisation rates for the 'DT-CPS without anomaly' and DT-CPS with anomaly' scenarios.

Human resource utilisation for Phases I–III	'Without anomaly' (minutes)	'With anomaly' (minutes)	Increase (%)
Receipt & Inventory	40.2	40.8	1.5
Storage & Monitoring	50.7	53.6	5.7
Distribution	45.6	77.4	69.7

4.2. Digital twin – cyber physical system model validation

- According to the results, the lead time of Phase III – Distribution has increased by 42.4% (Table 3), while the utilisation rates of human resources by 69.7% (Table 4) compared to the normal operations of the cryostorage warehouse. Although the anomaly occurs in Phase III, an increase in the lead times and resource utilisation in the other two phase (Phases I and II) has been observed due to parallel dynamic interactions within the three manufacturing phases.

Table 4

Case study: average human resource utilisation rates for the 'DT-CPS without anomaly' and DT-CPS with anomaly' scenarios.

Human resource utilisation for Phases I–III	'Without anomaly' (minutes)	'With anomaly' (minutes)	Increase (%)
Receipt & Inventory	40.2	40.8	1.5
Storage & Monitoring	50.7	53.6	5.7
Distribution	45.6	77.4	69.7

4.3. ‘Monitoring agent’ validation

- The architecture enables real-time communication between the RFID system and DT-CPS, and the computational model represents the actual behaviour of the interactive system.

- In this section, a ‘Disruption’ scenario is studied to demonstrate that after the ‘monitoring agent’ at the exo-level captures anomalous values in input real-time data, collected by the RFID system, can analyse the impact of anomalies (i.e. bottlenecks identification) to the system at the macro, exo and micro level agents.

- Self-optimisation is then employed to automatically update the micro-level agents and remove the identified bottlenecks. To demonstrate the impact of the ‘monitoring agent’ and self-optimisation, key performance indicators (KPIs) are tested and compared for the ‘Disruption’ scenario for two cases: ‘without feedback’ and ‘with feedback’, obtained from the configuration layer.

4.3. ‘Monitoring agent’ validation

- The simulation experiments for the ‘Disruption’ scenario have been performed for a three-day period. For this experiment, the daily number of orders and deliveries have been increased by 100% and the time required for ‘Picking material’ (Phase III) by about 500% times on average compared to the normal operation of the facility.

- After obtaining the RFID data to the database in the smart connection layer, the ‘monitoring agent’ at exo-level agent in the connection layer informs the user that an anomaly has been detected in the input data in terms of large numbers of deliveries and orders, and increased time for materials picking.

- Analysing the simulation results, the bottlenecks, identified during the daily practices for this scenario in the cognition layer, are:

4.3. 'Monitoring agent' validation

- Shortage of human resources in the refilling and recycling zones at the warehouse between 9:00 am – 14:30. Queues of cryomaterials waiting to be stored are identified, with average waiting time 21 min. The utilisation levels of the operators trained in the shipper filling and verification tasks are 81% and 74%, respectively. Considering that only LN2 cryogenic products have been studied in this work, these utilisation levels are greater than the maximum allowable limit (50%) set by the company. This ultimately may result in shortage of operators to perform the tasks on the shop floor.

- Shortage of human resources in the storage zone between 11:30 am - 16:30, due to interactive actions between the three manufacturing phases.

- Shortage of human resources in the dispatching zone between 14:30–16:30, due to queues in the quality check completion, due to interactive actions between the receipt (Phase I) and distribution (Phase III) zones.

4.3. ‘Monitoring agent’ validation

- • Shortage of validated shippers (≤ 5). This bottleneck, identified by the ‘monitoring agent’, makes the simulation model to stop, as there is no available shipper to assign the cryomaterials for dispatch. This bottleneck informs users that based on the demand and supply there is insufficient initial inventory and the company may miss out on sales opportunities.

- The root causes of the bottlenecks are further explored. Thus, the computational model captures that the cryogenic warehouse accepts 18 orders daily from which only the 10 are completed and dispatched. The total number of daily orders (No D) and dispatches (Nd D) over time are obtained from the simulation model as illustrated in Fig. 9(a).

4.3. ‘Monitoring agent’ validation

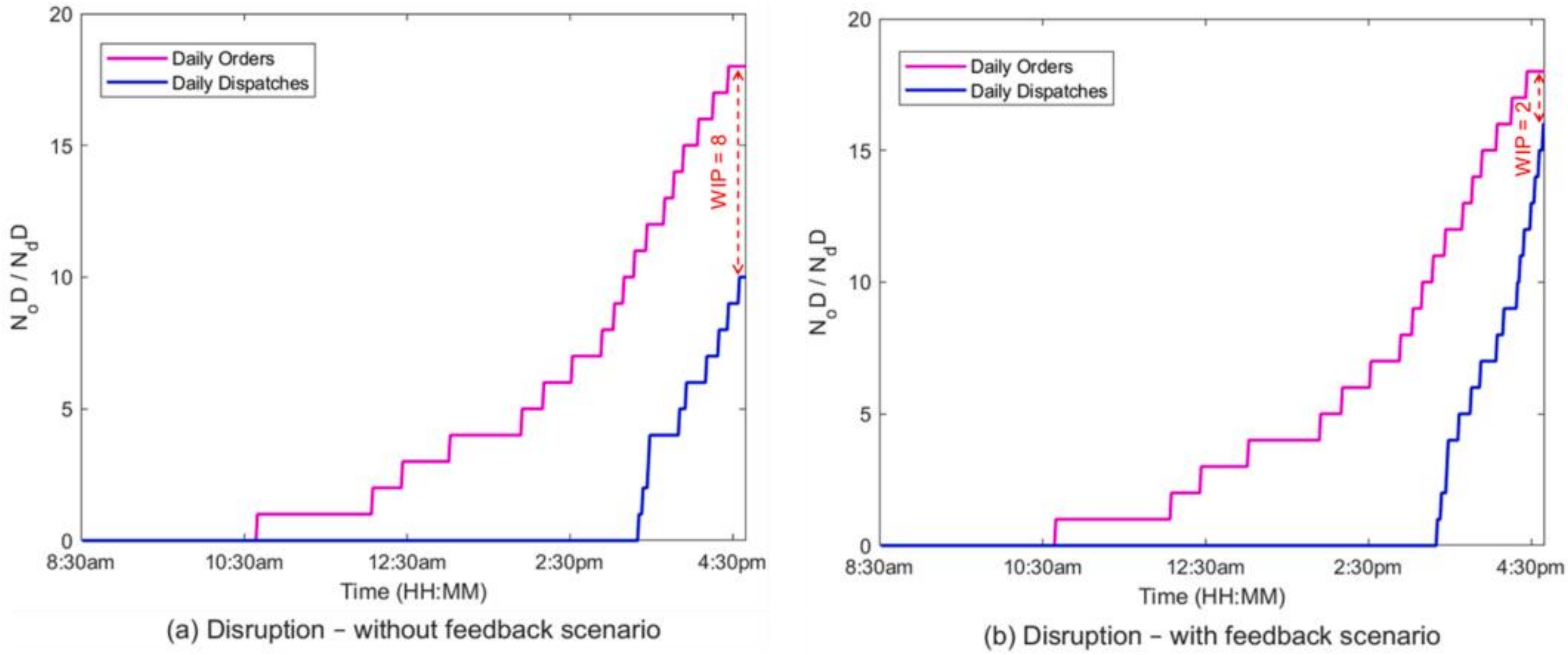


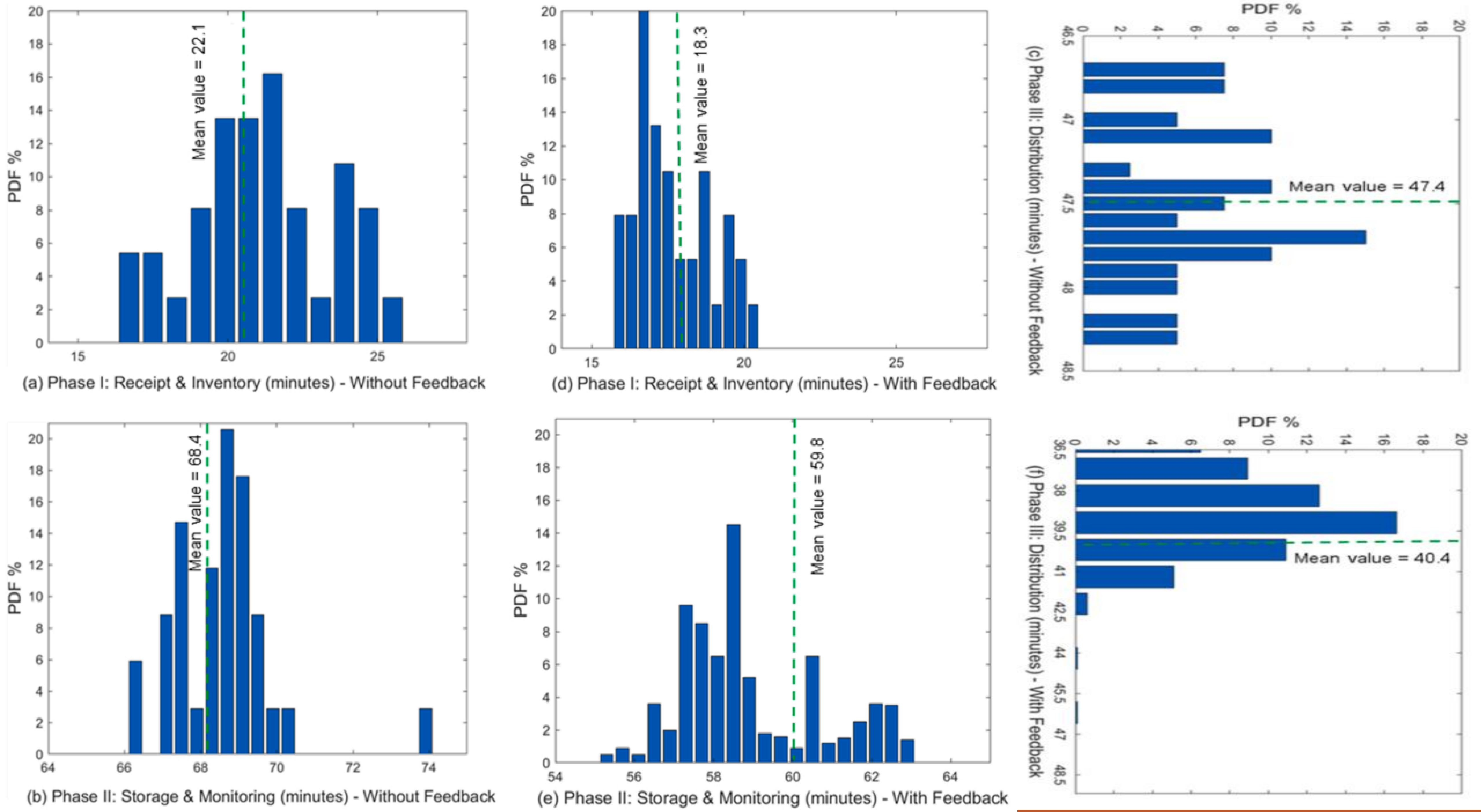
Fig. 9. Case study: Simulation results for the total number of the daily orders (N_o^D) and dispatches (N_d^D) for the: (a) ‘Disruption – without feedback’ and (b) ‘Disruption – with feedback’ scenarios.

4.3. ‘Monitoring agent’ validation

- The graph shows that the average WIP per day (at micro-level agent) is 8 orders, while only 56% of the orders accepted daily can be dispatched.

- The bottleneck root causes are further investigated through a stochastic data analysis, quantifying the uncertainty in lead times as visualised in the cognition layer.

- The results are compared against these from the ‘DT-CPS without anomaly’ scenario to show the capability of the ‘monitoring agent’ to capture the impact of the anomalies on KPIs. The histogram graphs have been developed to calculate the pdf for the time spent in each manufacturing phase at the exo-level agent. The pdf graphs for the lead (i.e. processing) times in Phases I–III for the ‘Disruption – without feedback’ scenario can be viewed in Fig. 10 (a–c).



(a) Phase I: Receipt & Inventory (minutes) - Without Feedback

(d) Phase I: Receipt & Inventory (minutes) - With Feedback

(c) Phase III: Distribution (minutes) - Without Feedback

(b) Phase II: Storage & Monitoring (minutes) - Without Feedback

(e) Phase II: Storage & Monitoring (minutes) - With Feedback

(f) Phase III: Distribution (minutes) - With Feedback

4.3. ‘Monitoring agent’ validation

- The pdf graphs have a Poisson distribution with λPhase I = 22.1min, λPhase II = 68.4min and λPhase III = 47.4min respectively. It is observed that 23% of deliveries are being received and documented within 16 – 18 min and for about 61% the process takes more than 20 min see Fig. 10 (a).

- On the contrary, in the ‘Without anomaly’ scenario, 95.5% of deliveries are being received and documented in less than 20 min, see Fig. 8(a). Additionally, the pdf for the storage and monitoring processing time in the ‘Disruption’ scenario shows that 35% of the products are being stored in about 68 min, 62% of the products in between 68 and 73 min, and for about 3% the process takes more than 73 min, see Fig. 10 (b). Moreover, the pdf for the storage and monitoring processing time in the ‘Without anomaly’ scenario, shows that 95.5% of the products are being stored in about 60 min, and only 4.5% of the products in between 60 and 70 min see Fig. 8(b). Finally, in Phase III, the pdf graphs for the distribution processing time for the ‘Disruption’ scenario, 87.5% of orders are being dispatched in less than 48 min see Fig. 10 (c). In terms of the ‘Without anomaly’ scenario, the average time needed to complete a product dispatch is much less as 97% of the orders are being dispatched in less than 30 min, see Fig. 8(c).

4.3. 'Monitoring agent' validation

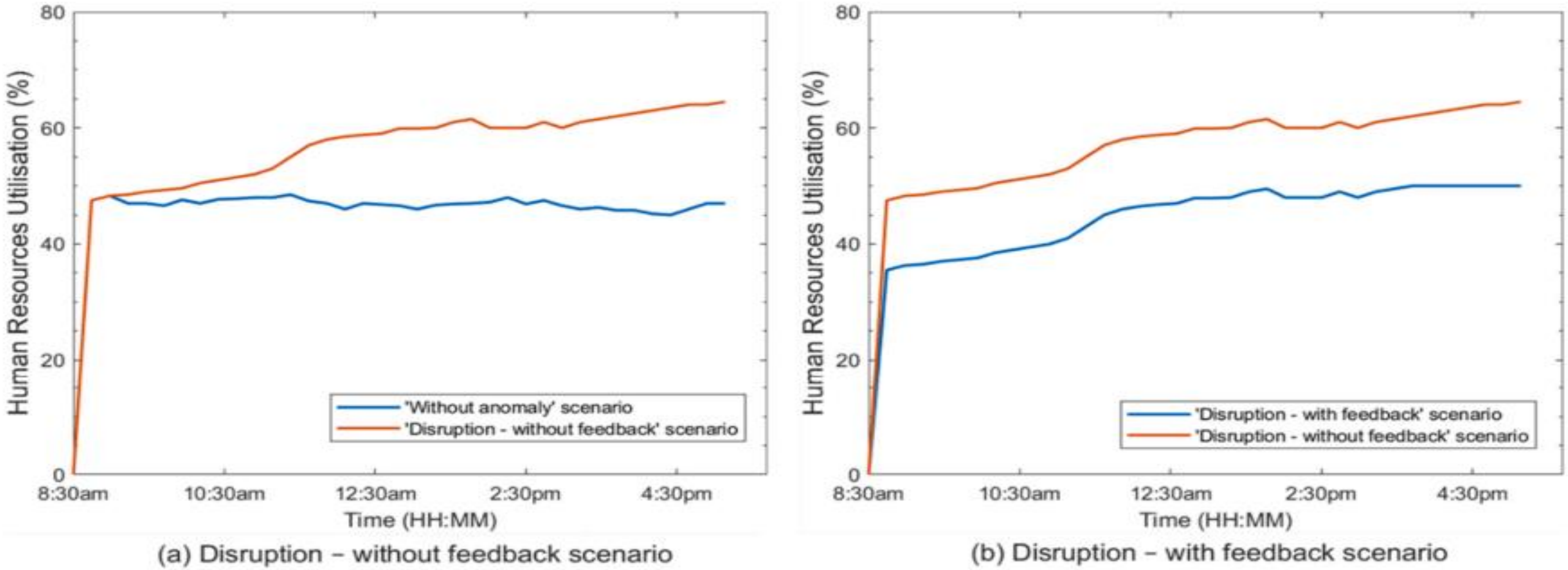


Fig. 11. Case study: Daily human resources utilisation rates for the: (a) 'Disruption – without feedback' and (b) 'Disruption – with feedback' scenarios.

4.3. ‘Monitoring agent’ validation

- Additionally, the average daily utilisation rates of human resources for the ‘Without anomaly’ and ‘Disruption – without feedback’ scenarios are 47% and 57%, respectively. The daily utilisation rates of the human resources for the two scenarios are illustrated in Fig. 11 (a).

- According to the results, the ‘monitoring agent’ can capture the impact of the studied anomalies on the utilisation rates of the operators at the warehouse.

- In the ‘Disruption – without feedback’ scenario, an increase in the utilisation rate is observed from 9:00 am, exceeding the corresponding rate of the ‘Without anomaly’ scenario during the daily operations (Fig. 11 (a)). This rise is explained due to the unexpected increase in TH and LTs for picking materials for dispatch at the macro and exo level agents, respectively.

4.3. 'Monitoring agent' validation

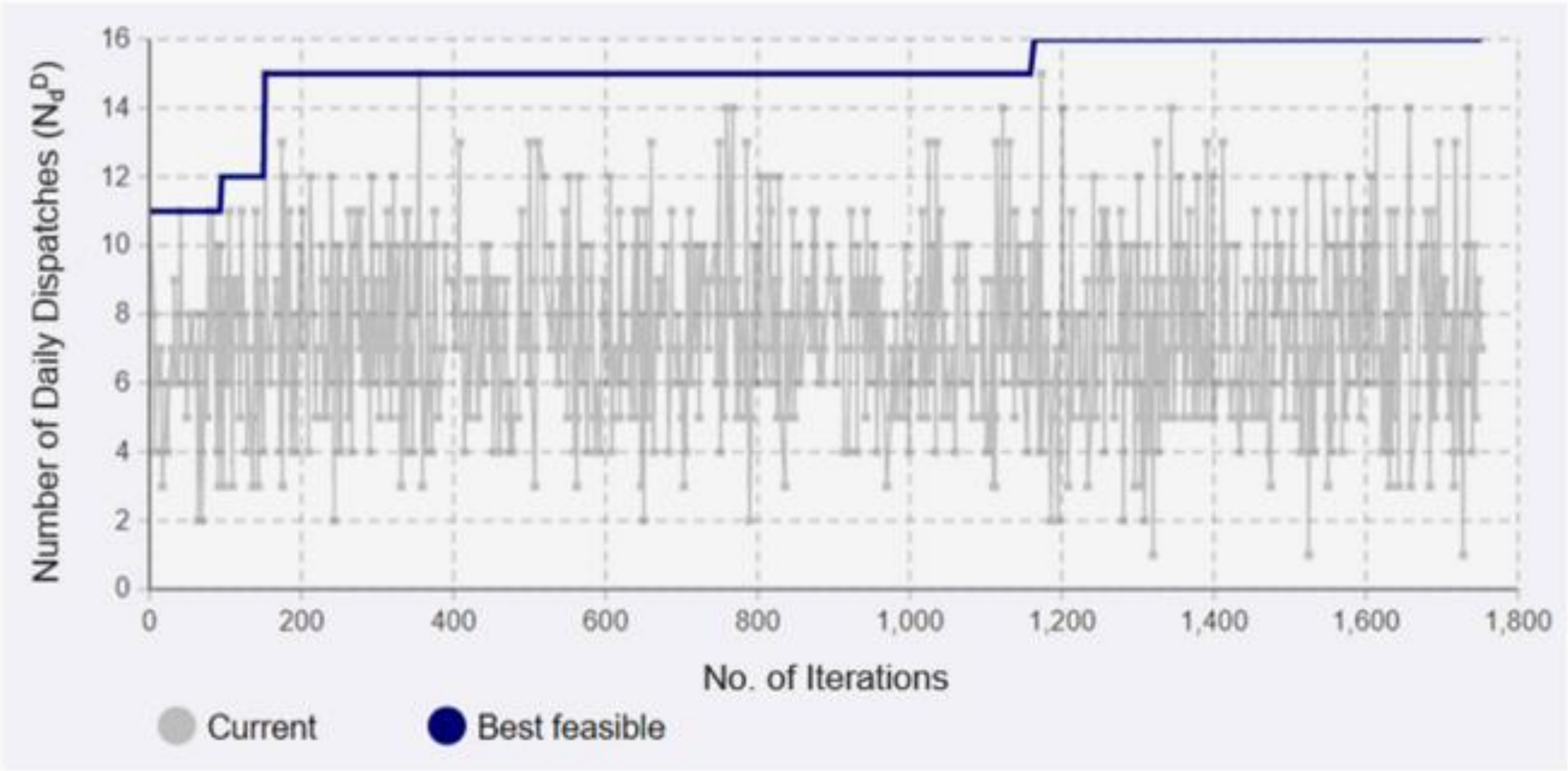


Fig. 12. Case study: optimisation results for the total number of the daily dispatches (N_d^D) for the 'Disruption' scenario.

4.3. ‘Monitoring agent’ validation

- The optimisation results show that if the proposed allocation of human resources is adopted by the cryogenic company, the number of daily dispatches can increase up to 16, i.e., 60% more orders can complete compared to the current figures (see Fig. 9(a)).

- Once the optimisation experiment is conducted, the proposed allocation of human resources is automatically embedded as feedback into the corresponding human resources parameters, modelled at the microlevel agent, in the smart connection layer using the getBestParamValue () method for the best iteration. After the human resources-related parameters are updated, the ‘monitoring agent’ detects the anomalous values in RFID data, but no bottlenecks related to the human resources utilisation rates are identified. The simulation model runs and new results are visualised in the cognition layer as illustrated in Figs. 9(b). Thus, the total number of daily orders (No D) and dispatches (Nd D) for the ‘Disruption – with feedback’ scenario are obtained from the simulation model as presented in Fig. 9(b). The computational results show that the cryogenic warehouse accepts 18 orders daily from which the 16 can be completed and dispatched.

4.3. ‘Monitoring agent’ validation

- It is also found that with the new allocation of the operators, the average queue waiting time in the refilling and recycling zones at the warehouse between 9:00 am – 14:30 has been reduced from 21 to 15 min.

- The utilisation levels of the operators trained in the shipper filling and verification tasks have been reduced from 81% and 74%, to 50% and 47%, respectively.

- Minor decrease of 4 min is also observed in the queue waiting time for quality check in the dispatching zone between 14:30 am - 16:30. The pdf graphs for the lead times in Phases I–III for the ‘Disruption – with feedback’ scenario can be viewed in Fig. 10 (d – f).

4.3. ‘Monitoring agent’ validation

- The pdf graphs have a Poisson distribution with λPhase I = 18.3min, λPhase II = 59.8min and λPhase III = 40.4min respectively. The average lead times for the three manufacturing phases, for the ‘without feedback’ and ‘with feedback’ scenarios are summarised in Table 5.

Table 5
Case study: average lead times for the ‘the Disruption – without feedback’ and ‘Disruption – with feedback’ scenarios.

Lead time for Phases I–III	‘Without feedback’ (minutes)	‘With feedback’ (minutes)	Reduction (%)
Receipt & Inventory	22	18	18.2%
Storage & Monitoring	68	60	11.8%
Distribution	47	40	14.9%

4.3. ‘Monitoring agent’ validation

- According to the results, the reallocation of human resources as proposed by the optimisation experiment can reduce the lead time of Phase III – Distribution by 15% compared to the ‘without feedback’ scenario. From the simulation results, it can be seen that the proposed reallocation of human resources can maximise the daily dispatches, complete all the deliveries, eliminate the WIP and prevent excessive use of human resources, while satisfying the constraints defined during the optimisation experiment in terms of WIP and HRU rates.

- Therefore, it has been demonstrated that the optimal values for the reallocation of human resources can be effectively applied for eliminating the bottlenecks emerged in the cryogenic warehouse from the occurrence of anomalies in the sensor data.

5. Discussion

- In manufacturing, systems are built by increasingly dynamic complexity at different levels of an agent-based model [7,9,20]. These systems, typically, consist of multiple manufacturing phases where various manufacturing activities operate simultaneously.

- Dynamic complexity, arising from manual activities with interactive behaviour, can create parallel dynamic interactions (i.e. collaborative interdependencies) in the system, affecting its productivity and performance. Advanced computational modelling such as bottom-up ABM approaches help represent such interdependencies and obtain a formal and flexible description of the system.

- This work contributes to the literature of complex manufacturing systems by proposing a generic, yet novel approach using the ABM technique for developing a DT-based multi-agent CPS model. The scope of the DT is to improve the operation of complex manufacturing systems, while the purpose of the CPS is to support the implementation of DT by automatically enabling anomaly detection and emergent bottlenecks identification (exo-level) through communicating with other agents in macro, exo, meso and micro levels dynamically.

6. Concluding remarks

- This paper has presented a DT-CPS approach, composed of multiple agents, for automated anomaly detection, and bottlenecks identification and removal for complex manufacturing systems with dynamic parallel interactions, using the bottom-up ABM technique. Anomalous values in model input data, captured from RFID sensors, are detected at the microlevel agent and bottlenecks that deteriorate the system’s performance are identified at the micro, meso, exo and macro level agents. The theoretical aspects and the mathematical formulation of the DT-based multi-agent CPS method have been introduced as an extension to the hybrid simulation method.

- The DT-based multi-agent CPS architecture, model, mathematical method, and simulation model can be used as an automated monitoring tool of anomalies detection, and bottlenecks identification and removal for more informed decision making and control in manufacturing sectors with a highly regulated and complex nature. The proposed architecture and method differ from the existing models as anomalies in input data are detected and unplanned bottlenecks are identified and eliminated automatically over time using real-time data.

6. Concluding remarks

- The bottom-up approach of the model using the multi agent-based technique for DTs can enhance the flexibility, interactivity and modularity of DT-CPS design. The bidirectional communication between the physical and twin spaces is also considered. Additionally, the method and simulation model follow a stochastic bottom-up approach for DT-CPS, to detect anomalies and identify bottlenecks in complex manufacturing systems using the ABM, DES and pdf techniques.

- Further to this work, the applicability of the DT-CPS approach can also be explored in other manufacturing or production systems and supply chains. Further research can also be conducted to quantify the impact of the ‘monitoring agent’ in terms of sustainability and evaluate the cost of goods and energy consumption. In this regard, cost information could be added to the simulation model to calculate the cost and profit for different scenarios considering unexpected and emergency events and their financial and environmental impacts.

Opinion

- Although the paper discusses an interesting topic of botnet detection and prevention in IIoT systems. The proposed system is not related to digital twins in my opinion.

- The deployed methods and simulation background is a classical CPS model.

- The author also mentioned that the paper provides solutions to prevent bottleneck, however, the paper only discussed detection methods using simple analytic techniques.

- The description of the steps in the paper could be summarized in a simple workflow as well, the authors used multiple repetition that made the paper difficult to read.