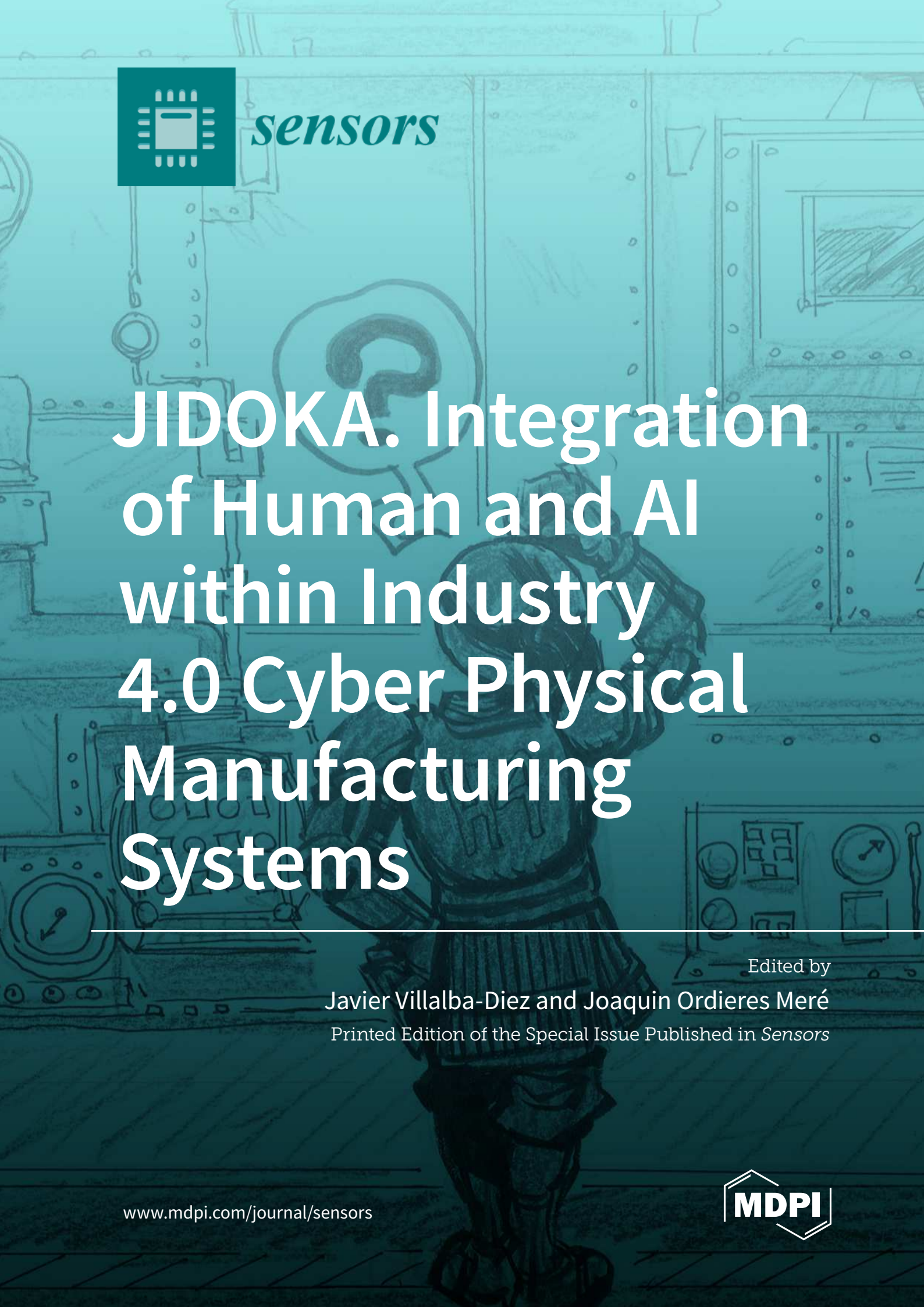


sensors

JIDOKA. Integration of Human and AI within Industry 4.0 Cyber Physical Manufacturing Systems

Edited by
Javier Villalba-Diez and Joaquin Ordieres Meré

Printed Edition of the Special Issue Published in *Sensors*

www.mdpi.com/journal/sensors

MDPI

JIDOKA. Integration of Human and AI within Industry 4.0 Cyber Physical Manufacturing Systems

JIDOKA. Integration of Human and AI within Industry 4.0 Cyber Physical Manufacturing Systems

Editors

Javier Villalba-Diez
Joaquin Ordieres Meré

MDPI • Basel • Beijing • Wuhan • Barcelona • Belgrade • Manchester • Tokyo • Cluj • Tianjin

Editors

Javier Villalba-Diez
Management und Vertrieb
Hochschule Heilbronn
Schwäbisch Hall
Germany

Joaquin Ordieres Meré
Industrial Engineering School
Universidad Politécnica de
Madrid
Madrid
Spain

Editorial Office
MDPI
St. Alban-Anlage 66
4052 Basel, Switzerland

This is a reprint of articles from the Special Issue published online in the open access journal *Sensors* (ISSN 1424-8220) (available at: www.mdpi.com/journal/sensors/special_issues/JIDOKA_Integration).

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

LastName, A.A.; LastName, B.B.; LastName, C.C. Article Title. *Journal Name* **Year**, *Volume Number*, Page Range.

ISBN 978-3-0365-3812-9 (Hbk)
ISBN 978-3-0365-3811-2 (PDF)

Cover image courtesy of Javier Villalba-Diez

Contents

About the Editors

Javier Villalba-Diez

Javier Villalba-Diez, PhD., aims to empower individuals and organizations to achieve their strategic goals while increasing trust. Dr. Villalba-Diez is a Mechanical Engineer with Technische Universita t Mu nchen, Germany and Industrial Engineer with Technical University Madrid, Spain (2003). He received his PhD in Engineering, Economics and Organizational Innovation with a focus on Strategic Organizational Design from the Universidad Polite cnica de Madrid in 2016. His PhD was awarded with the prize for the best doctoral thesis. In March 2022, he received a second PhD in Engineering from the Universidad Polite cnica de Madrid and Applied Physics with a focus on industrial applications of Quantum Computation. His current research interests include Quantum Computing, Deep Learning, Hoshin Kanri, Strategic Organizational Design, Business and Artificial Intelligence. He has 15 years'worth of experience as a lean consultant and production manager in a number of positions related to manufacturing operations in German, American and Japanese manufacturing facilities. His research and work, currently performed at Hochschule Heilbronn in Germany, has brought him to numerous companies and hundreds of factories, where he collaborates with people to test ideas and share lessons learned. He splits his time between Germany, USA, Japan, and Spain.

Joaquin Ordieres Meré

Prof Ordieres-Meré is a full professor at the Industrial Engineering School of the Universidad Politécnica de Madrid, Spain, and his research aims are to increase the understanding of Integrated Manufacturing processes and their identification and optimization with the help of Business Analytic tools, including artificial intelligence and quantum computing.

In particular, complex processes involving human and technological devices as well as complex socio-technical systems are targeted. The aim is to be able to translate the gained knowledge into advanced tools helping decision-making processes that managers need to carry out, by providing them with either support or guidance.

Different type of processes and industries have been explored, such as building physics, steel-making, and rubber industries. In addition, some interest was also paid to more scientific fields, such as pollution prediction or digital astrophysics, where similar tools helped to bring additional value and knowledge.

His research has been cited more than 8000 times, and he has published more than one hundred and fifty journal papers and a similar number of conference papers as well as fifteen patent applications.

Preface to "JIDOKA. Integration of Human and AI within Industry 4.0 Cyber Physical Manufacturing Systems"

This reprint is dedicated to Manuel Villalba, a boy with great talents. Remember that discipline is the root of all good qualities.

This book is about JIDOKA, a Japanese management technique coined by Toyota that consists of imbuing machines with human intelligence. The purpose of this compilation of research articles is to show industrial leaders innovative cases of digitization of value creation processes that have allowed them to improve their performance in a sustainable way. This book shows several applications of JIDOKA in the quest towards an integration of human and AI within Industry 4.0 Cyber Physical Manufacturing Systems. From the use of artificial intelligence to advanced mathematical models or quantum computing, all paths are valid to advance in the process of human–machine integration.

The editors would like to thank the different authors for their contributions. Special thanks are extended to the Escuela Técnica de Ingenieros Industriales de la Universidad Politécnica de Madrid, Spain, and the Hochschule Heilbronn, Germany.

Javier Villalba-Diez and Joaquin Ordieres Meré
Editors

sensors

MDPI

Article

Human–Machine Integration in Processes within Industry 4.0 Management

Javier Villalba-Diez [1,2] and Joaquín Ordieres-Meré [3,*]

[1] Hochschule Heilbronn, Fakultät Management und Vertrieb, Campus Schwäbisch Hall, 74523 Schwäbisch Hall, Germany; javier.villalba-diez@hs-heilbronn.de
[2] Department of Artificial Intelligence, Escuela Técnica Superior de Ingenieros Informáticos, Universidad Politécnica de Madrid, Boadilla del Monte, 28660 Madrid, Spain
[3] Escuela Técnica Superior de Ingenieros Industriales (ETSII), Universidad Politécnica de Madrid, 28006 Madrid, Spain
* Correspondence: j.ordieres@upm.es

Abstract: The aim of this work is to use IIoT technology and advanced data processing to promote integration strategies between these elements to achieve a better understanding of the processing of information and thus increase the integrability of the human–machine binomial, enabling appropriate management strategies. Therefore, the major objective of this paper is to evaluate how human–machine integration helps to explain the variability associated with value creation processes. It will be carried out through an action research methodology in two different case studies covering different sectors and having different complexity levels. By covering cases from different sectors and involving different value stream architectures, with different levels of human influence and organisational requirements, it will be possible to assess the transparency increases reached as well as the benefits of analysing processes with higher level of integration between them.

Keywords: Industry 4.0; Operator 4.0; process variability; JIDOKA; integration explaining variability

Citation: Villalba-Diez, J.; Ordieres-Meré, J. Human–Machine Integration in Processes within Industry 4.0 Management. *Sensors* **2021**, *21*, 5928. https://doi.org/10.3390/s21175928

Academic Editor: Leopoldo Angrisani

Received: 11 July 2021
Accepted: 30 August 2021
Published: 3 September 2021

1. Introduction

Value chains associated with Industry 4.0 (I4.0) are formed by complex cyber-physical networks in which humans and machines process information efficiently to supply a customer with the desired product [1–3]. I4.0 and industrial internet of things (IIoT) describe new paradigms for integrated human–machine interaction [4,5]. Both concepts are based on intelligent, interconnected cyber-physical production systems that are capable of controlling the process flow of industrial production. Since machines autonomously make many decisions and interact with production planning and manufacturing systems, the integration of human users requires new paradigms [6].

IIoT technology is significantly contributing to enlarge the data available for many manufacturing processes. In an I4.0 context, such as the IIoT [7,8], these data are produced by decentralized sources such as thousands of sensors in factories [9], i.e., the data are distributed over networks [10]. With the classical already collected dataset related to sensors located at the processing machines, now it becomes possible considering additional data coming from wearables of human operators [11]. The number of edge devices that are currently developed to support fitness and health monitoring is enormous [12]. Many of them aim at measuring body parameters to offer care related services [13]. At the same time, a lot of smart health applications are developed, often making decisions or offering feedback based on sensor data processing. Application developers often struggle to integrate and plug in novel sensor technologies, becoming available on the market at a fast pace [14]. These technologies enable describing processes in a more integrated way, including many more potential sources of variability [8]. Although the advantages are rather evident, still there are significant challenges to be better identified and faced when new useful solutions regarding knowledge and management are foreseen.

In the manufacturing IIoT I4.0 domain, the I4.0 vision has promoted smart manufacturing and smart factory concepts by augmenting all assets with sensor-based connectivity [15]. These intelligent sensors generate a large volume of industrial data helping to create digital twins (defined as a digital replication of both living and inanimate entities that enable seamless data transfer between the physical and virtual worlds) as support for a live mirror of physical processes [16,17]. Within this approach, the ambition is to capture the process variability, being able to process all relevant information by big data analysis on cloud computing so that manufacturers are able to find manufacturing processes' bottlenecks, identify the causes and impacts of problems in such a way that effective implementation of measures becomes useful either for product design or for manufacturing engineering including maintenance, repair and overhaul [18].

A critical aspect to be considered when the previous interest is addressed is the human influence on the processes. There is a gap between the information collected by the IIoT devices and their capability to capture the causes influencing process variations. This human–machine symbiosis presents great potential advantages, since on the one hand the human has a great cognitive flexibility that the machine lacks, while on the other hand the machine has a great computational capacity superior to the human [19]. However, there are also voices warning of the potential dangers of the *bionization* of human tasks [20,21]. Fundamental requirements for the future design of human–machine interactions in productive assembly systems are now being identified [22]. Expectations generated by Operator 4.0 (O4.0) in this context have implications for empowerment and management models [23]. The technical implications of realising a human–machine symbiosis have to enable the use of trustworthy and ethical artificial intelligence (ethical AI) [24].

The aim of this work is to use IIoT technology and advanced data processing to promote integration strategies between these elements to achieve a better understanding of the processing of information and thus increase the integrability of the human–machine binomial, providing appropriate management strategies for these configurations [25–27]. Thus, the major objective of this paper is to evaluate how human–machine integration helps to explain the variability associated with value creation processes. Therefore, the research question being addressed in this paper can be formulated as *RQ1: The I4.0 technology allows to increase the transparency to understand process variability when it is used to integrate different sources of uncertainty.*

In particular, we are interested in the case of natural intereffects between human workers and operating machinery. The approach selected is to implement an action research methodology through two different case studies covering different sectors and having different complexity levels, and the presentation is structured into four further major sections. First, in Section 2, we outline the main lines of research that deal with the human–machine integration in an I4.0 environment. Second, in Section 3 we present the results of two case studies that illustrate the usage of IIoT technology when integrating the human–machine binomial. Third, in Section 4 we discuss the possible implications for the management of creating processes. Finally, in Section 5 we present the conclusions, further steps, and possible limitations of this work.

2. State of the Art

Strategic organizational design is a scientific field [28,29] that studies the relationship between organizational entities and how its structure and functionality affects its performance [30]. Under the organizational network paradigm, modern organizations can be understood as a symbiotic socio–technical ecosystem of social networks [31] that interact with ever increasingly complex networked cyber-physical distributed interconnected sensors [32], whose readings are modelled as time-dependent signals on the vertices, human or cyber–physical, respectively.

Under this evolutionary information flow perspective [33], organisations can be adequately modelled [34]. One of these models is the Human–Cyber–Physical Systems (H–CPS) model, that integrates the operators into flexible and multi-purpose manufactur-

ing processes. The primary enabling factor of the resultant O4.0 paradigm is the integration of advanced sensor and actuator technologies and communications solutions. Although process automation reduces costs and improves productivity, human operators are still essential elements of manufacturing systems [35]. As discussed in [36], the degree of automation does not directly imply an enhanced operator performance, because handling human factors requires more complex dimensions related to human-to-machine interactions, including robotics. The integration of workers into an I4.0 system consisting of different skills, educational levels, and cultural backgrounds is a significant challenge. The new concept of O4.0 was created for the integrated analysis of these challenges [19].

The concept of O4.0 is based on the so-called H–CPSs designed to facilitate cooperation between humans and machines [36]. Although specific contributions regarding different dimensions have been proposed by different authors [23,37] still there is a significant room for improvement when an integrated perspective is required, because the available wearable devices lack of enough level of integration. In current industrial practice, most applications are developed in isolated circumstances aimed at addressing specific problems. Therefore, there is a gap in creating human-centred systems able to promote an operator learning context not only relying on single parameters but also providing a meaningful articulated set of relevant parameters both in the short and long term [37].

I4.0 envisions a future of networked production where interconnected machines and business processes running in the cloud will communicate with one another to optimise production and enable more efficient and sustainable individualised/mass manufacturing. Inside such a vision, there are different requirements to be considered, including cloud computing, data pseudo-anonymisation, as well as data micro-services. The shop-floor in virtual space is the reconstruction and digital mapping of the physical devices at shop-floor level. They exchange data/information/knowledge through by using a big data storage and management platform.

These shop-floor-management platforms construct a virtual shop-floor system that monitors the working progress and working status of assembly stations, products, and manufacturing resources in the physical shop-floor so that it can be dynamically, realistically, and accurately mapped in the virtual space through cloud services [38]. The main challenge to developing shop-floor in virtual spaces is addressed is the complexity of the IIoT solutions, as they suffer from poor scalability, extensibility, and maintainability [39,40]. In response to those challenges, microservice architecture has been introduced in the field of IIoT application, due to its flexibility [41], lightweight [42] and loose coupling [43].

The evolution of the human–machine integration that allows benefiting from the different information processing capabilities of both parties has been investigated in this context in a comprehensive manner [5]. Extensive and intensive research has shown that on the one hand, humans in shop-floor management environments in I4.0, have a holistic problem-solving capability where several brain areas are activated for problem solving [44–46], however humans have a limit to the cognitive load they can compute that affects their performance [47,48]. On the other hand, with the advent of artificial intelligence, machines are increasingly capable of performing a massive processing of information that can make up for human deficiencies: one approach is to use the machine, having greater computational capacity to reduce the cognitive load of humans [49–51], another approach has been to create a semantic framework that allows for machine recommendations for human problem-solving related to manufacturing tasks [52,53], while other scholars have proposed an open source web-based protocol to enhance inter-operability between human and machine assets [54]. The problem with all these proposals can be summarized in the fact that they try to adapt either the machine to man or vice versa, and as a natural consequence, they obviate a symbiosis between both forms of information computation.

Although previous research studies have addressed aspects related to human–machine integration, they were performed mainly from a dominant perspective, including process development design [55], but also analysing the relationship between management practices and Industry 4.0 as in [56]. However, a low number of contributions were focused on

how a more integrated view provided with the help of I4.0 allows to better understand causes of process variability, and this paper aims to contribute to reduce this gap.

3. Case Studies

As a first step to evaluate the effect of human–machine integration in I4.0 environments, two case studies are used. As argued by Byrd and Turner [57], a single case study can be seen as the only possible building block in the process of developing the validity and reliability of the proposed hypothesis. Following the recommendations of Eisenhardt [58], a clear case study road-map is followed for each one of them. This road-map has several phases:

1. Scope establishment;
2. Specification of population and sampling;
3. Data collection;
4. Standardisation procedure;
5. Data analysis.

3.1. Case Study 1. Reverse Logistics Process. Near Field Communication

In this case study, we focus on studying the variability experienced by a reverse logistics collection process of steel scrap, in which a human driver in a truck covers different routes and time periods. In particular, we concentrate on merging relevant and difficult to evaluate aspects such as the state of the drivers, their operational working conditions, and other health-related parameters with data related to the machine elements used to perform the logistics tasks.

3.1.1. Scope Establishment

The ambition in this case is to merge technologies covering different aspects as a way to better define the influence of different factors. We placed these devices on the human worker and we considered that they do not affect their work, which can be performed normally.

Integration of data flows is relevant to produce process-related information, making it possible to extract behavioural rules.

3.1.2. Specification of Population and Sampling

In this case study, we analysed the data of 5 users sharing 3 trucks, performing daily routes in 3 shifts. We monitored the data are monitored on a per second basis, but aggregated them by day to ensure a more consistent analysis.

3.1.3. Data Collection

The initial goal is to assess the technologies themselves in a real case implementation, which include, as shown in Figure 1:

- Health-related parameters are gathered through non-invasive Bluetooth Low Energy (BLE) devices. In this study, we consider as irrelevant the effect that the fact that their health is being measured could have on human behaviour.
- Trucks' condition monitoring is gathered through solid state based devices.
- Near-field communication (NFC) Technologies.

Specific Android based applications have been developed to enable data collection and sensor fusion [59], as well as the integration of the NFC tags with the process logic in a consistent way [60] through a mobile phone.

Specifically, from MongoDB [61], we have created a platform for storing vehicle driver wristband data. They are organized by date (day) and MAC identification of the wristband. We have also created a read user in the MariaDB manager cluster system, a community-developed fork of the MySQL relational database management system [62], which stores lower frequency data from different sensors and web services. Every few minutes they run processes that load data into their databases which can be accessed with a certain user and password.

Figure 1. Some devices implemented in the case study 1. Smartbands to collect stress levels on top subfigure and Single Board Computers with appropriate plugin sensors on bottom subfigure.

3.1.4. Standardization Procedure

The key performance indicators (KPIs) measured are standardized over the whole management system to ensure a comparable framework in which several processes can be benchmarked against each other. A list of the measured KPIs and its meaning is depicted in Table 1.

Table 1. Monitored KPIs Case Study 1.

Name	Meaning
date	date for the record.
shift	Number of shift. 1: 06:00–14:00; 2: 14:00–22:00; 3: 22:00–06:00+1
plate	Truck plate ID.
user	Anonymous user ID (pseudo-anonymity for the truck driver).
idcycle	Number of cycle in the working day.
kpi_unload	Duration in minutes to unload the truck at the headquarters.
kpi_trip1	Duration in minutes from headquarters to customer facilities for collecting the scrap.
kpi_customer	Duration in minutes inside the customer facilities.
kpi_input	From customer entrance to loading point.
kpi_output	From loading point to the exit.
kpi_load	Duration of scrap loading process.
kpi_trip2	Duration in minutes from customer facilities for collecting the scrap to headquarters.
kpi_total	Duration of the whole cycle without headquarters movements.
kpi_tot_cycle	Duration of the total cycle.
weight	Scrap weight.
t2	Absolute time for starting the cycle.

3.1.5. Data Analysis

To perform the analysis of the KPIs, a preliminary analysis of the coefficients of variation (CV), as the ratio between the standard deviation and the mean of the KPI. It becomes interesting to see how the CV is much more sensitive to the rather distant multimodal structure. To avoid such effects, the coefficient interquartile of dispersion (CQD) was introduced. Figure 2 shows that some KPIs have the largest CQD and these KPIs are good advisors for the variability of processes. Therefore, attention will be given to the reasons for such variability and to do this a paired view of the KPIs is presented in Figure 3.

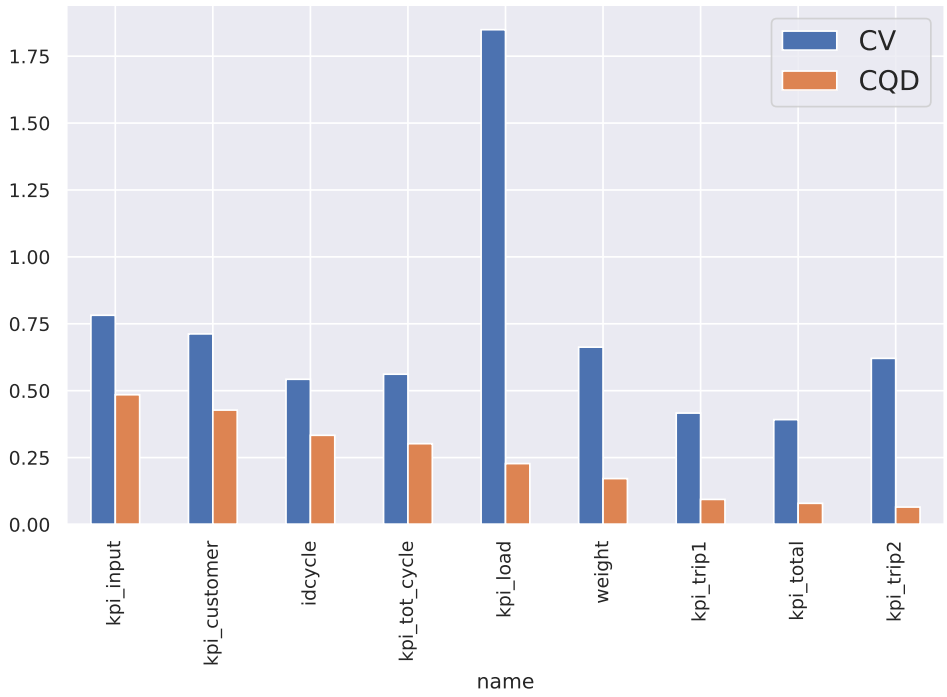


Figure 2. CV and CQD per variable in case study 1. Units for CQD are same that for original variables (see Table 1).

To assess the variability in more detail, it was decided to analyse records obtained because of the daily activity, by assigning to each variable its quartile class (see Table 2), additionally class-oriented variables such as Shift, TruckPlate and DriverID were included to represent all relevant elements potentially linked to process variability (see Table 3). In this way, we transform the records of KPIs related to the same operation cycle in terms of an orchestrated list of KPIs quartiles that can be then extended to all operations, creating a sort of item list. The goal is to apply data mining (DM) to obtain potentially useful, previously unknown, and ultimately understandable knowledge from the data. Association rule mining is one of the important portions of data mining and is used to find interesting associations or correlation relationships between item sets in mass data (item list) [63].

To apply the DM association rule technology, the selected algorithm was FP-Growth (frequent-pattern growth), which is an improved algorithm of the Apriori algorithm put forward by [64]. It compresses data sets to a FP-tree, scans the database twice, does not produce the candidate item sets in mining process, and greatly improves the mining efficiency [65].

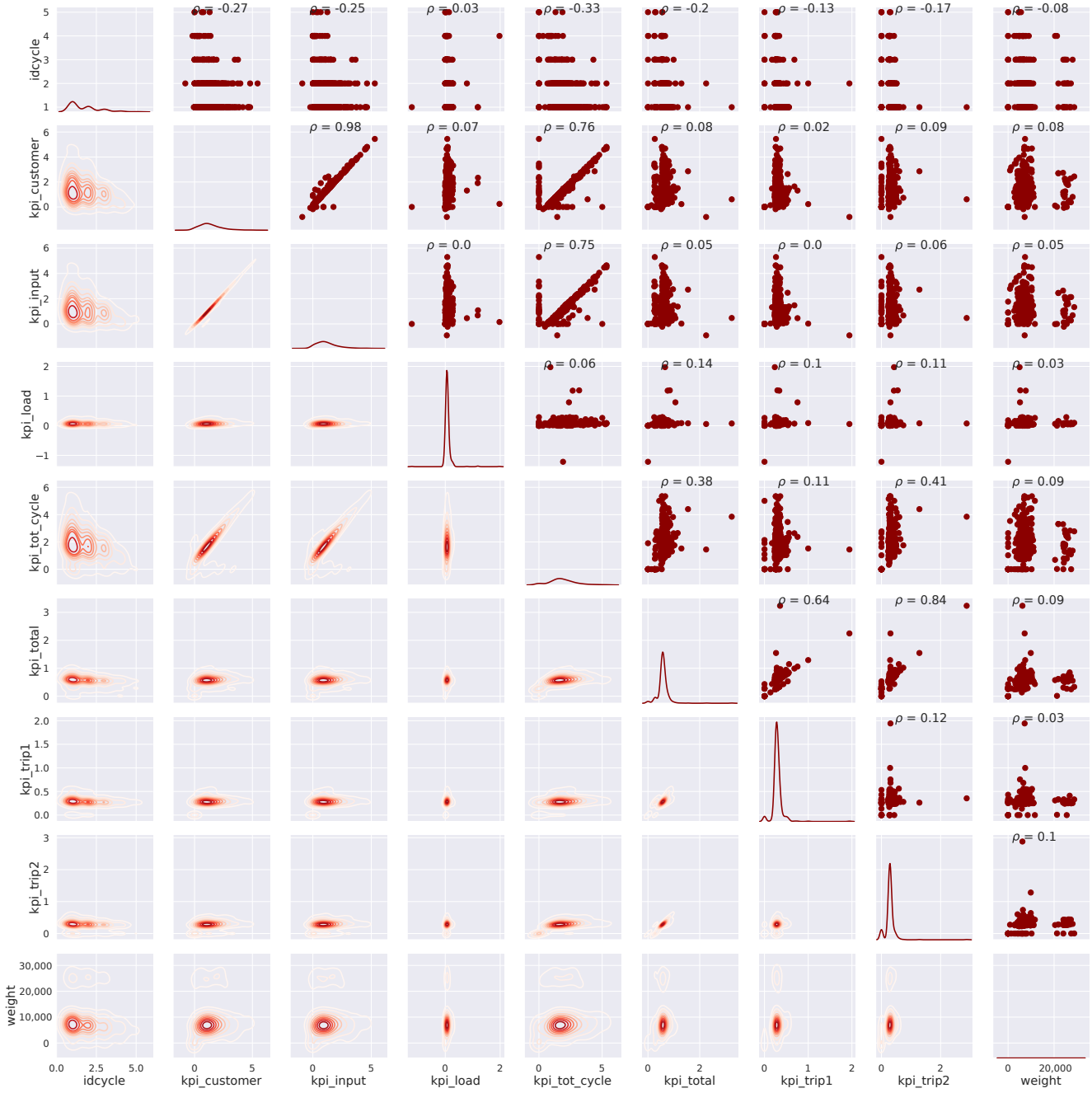


Figure 3. Correlation between KPIs in case study 1. Units for original variables were defined at Table 1.

Table 2. Quartile ranges for the interesting KPIs.

KPI (Unit)	Min	Q1-Init	Q2-Init	Q3-Init	Max	StDev
idcycle (h)	1.000000	1.000000	1.000000	2.000000	5.00000	0.945672
kpi_customer (h)	−0.801667	0.730278	1.187360	1.821110	5.44750	0.961633
kpi_input (h)	−0.905833	0.574653	1.036945	1.656458	5.29806	0.940865
kpi_load (h)	−1.217780	0.050208	0.063750	0.079792	1.97444	0.151037
kpi_tot_cycle (h)	0.000000	1.283052	1.745555	2.393260	5.34778	1.047631
kpi_total (h)	0.000000	0.534722	0.571111	0.626667	3.23611	0.228157
kpi_trip1 (h)	0.000000	0.259653	0.281805	0.313402	1.94278	0.123308
kpi_trip2 (h)	0.000000	0.267500	0.284722	0.304791	2.88056	0.177552
weight (Kg)	0.000000	5830.000000	6960.000000	8245.000000	28,780.00000	5271.877027

Table 3. Itemlist from the process records to be used for rule construction.

ItemListID	ItemList
1	(Plate_01, U_1, Shift_Q1, idcycle_Q0, kpi_customer_Q0, kpi_input_Q0, kpi_load_Q0, kpi_tot_cycle_Q0, kpi_total_Q0, kpi_trip1_Q1, kpi_trip2_Q0, weight_Q0)
2	(Plate_02, U_4, Shift_Q1, idcycle_Q0, kpi_customer_Q0, kpi_input_Q0, kpi_load_Q0, kpi_tot_cycle_Q0, kpi_total_Q3, kpi_trip1_Q3, kpi_trip2_Q3, weight_Q3)
...	...

After creating the item list, mining for rules having limited support but high confidence can start. Rules do not extract an individual's preference, rather find relationships between the set of elements of every distinct transaction. This is what makes them different from collaborative filtering. Normally, rules exhibiting a high level of support are the so-called 'well-known rules' the people involved in such activities are familiar with, but those having low support, although their confidence becomes even higher, are harder to learn and it is where DM can help to unveil those unknown behaviours.

In our application case, the threshold for support was established at 15% and the confidence threshold was established at over 95%. As a significant variability in CQD appeared regarding *kpi_tot_cycle*, which actually reflects all operations, it could be interesting to look for explanations with its highest range (kpi_tot_cycle_Q3) and the lowest one (kpi_tot_Cycle_Q0).

The relevant factors and combinations can be better understood, and filtering the right hand side (RHS) to contain *kpi_tot_cycle_Q3* it is possible to find what Table 4 presents. Similarly, lower values for the same KPI have been analysed with the same technology, where the findings are presented in Table 5.

Table 4. Rules explaining the Q3 for the *kpi_tot_cycle, where ^means logical and*.

	Antecedent_STR	Consequent_STR	Confidence
9	kpi_customer_Q3^kpi_total_Q3	kpi_tot_cycle_Q3	1.000000
10	kpi_input_Q3^kpi_total_Q3	kpi_tot_cycle_Q3	1.000000
15	kpi_input_Q3^kpi_total_Q2	idcycle_Q0^kpi_customer_Q3^kpi_tot_cycle_Q3	0.967742
16	kpi_customer_Q3^kpi_input_Q3^kpi_total_Q2	idcycle_Q0^kpi_tot_cycle_Q3	1.000000
18	idcycle_Q0^kpi_input_Q3^kpi_total_Q2	kpi_customer_Q3^kpi_tot_cycle_Q3	0.967742
22	kpi_customer_Q3^kpi_total_Q2	idcycle_Q0^kpi_tot_cycle_Q3	1.000000
24	idcycle_Q0^kpi_customer_Q3^kpi_total_Q2	kpi_tot_cycle_Q3	1.000000
58	kpi_customer_Q3^kpi_input_Q3^kpi_trip2_Q3	kpi_tot_cycle_Q3	1.000000
60	idcycle_Q0^kpi_customer_Q3^kpi_trip2_Q3	kpi_tot_cycle_Q3	1.000000
62	idcycle_Q0^kpi_input_Q3^kpi_trip2_Q3	kpi_tot_cycle_Q3	0.975000

Table 5. Rules explaining the lower quartile values for the *kpi_tot_cycle*.

	Antecedent_STR	Consequent_STR	Confidence
1	idcycle_Q1^kpi_customer_Q0	kpi_tot_cycle_Q0	0.972973
2	idcycle_Q1^kpi_input_Q0	kpi_customer_Q0^kpi_tot_cycle_Q0	0.972222
3	idcycle_Q1^kpi_customer_Q0^kpi_input_Q0	kpi_tot_cycle_Q0	0.972222

Based on the more than four hundred items in the item list, more than 250000 rules have been distilled when the minimum support is chosen to one percent. Analysis of the selected RHS rules presented in Tables 4 and 5 show several interesting aspects, such as that there are rules discovered that sometimes are meaningless from the practical business point of view, as in the case of Table 5, because the second rule assumes that the time spent by the customer during the collection of materials is a consequence, and it is part of the process as an antecedent and never a consequence. It also happens in Table 4 with rules 15, 16, 18, and 22.

Another relevant aspect is that the discovered rules establish a relationship between the total cycle KPI and the input time KPI as well as with the time used to collect materials at the customer's site. This is interesting because those two KPIs were the most sensitive to CQD in Figure 2.

3.2. Case Study 2. Integration between Robotics and Human Oriented Processes

In this case study, there are three different processes involved, each of them with different levels of automation and human operator engagement. In Figure 4, the sub-process design is presented, where the manual forming of different components are manufactured and assembled in six different configurations (Rework station). After assembling, robotic laser-based quality control stations (SCAN units) verify the geometric tolerances of each part and when they do not pass the quality checks, they are routed back for manual repairing, getting integrated again for checking afterwards. After successful robotic inspection, a set of three pairs of manual inspection stations are configured to finally assess the parts (CHECK units) and attach the individual report before packaging and delivering products to customers (Final Gate Storage).

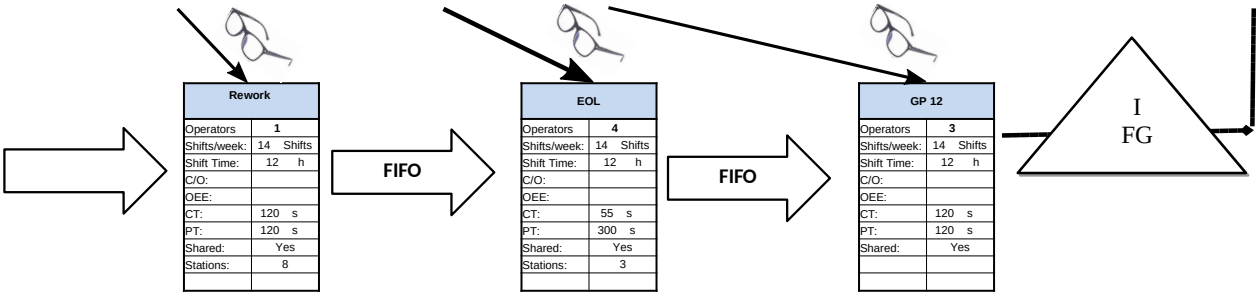


Figure 4. Value Stream Design (VSD).

Here there are different sources of variability as per part reference. First of all, since part manufacturing involves a relevant amount of manual work and component integration, including robotic welding stations, the number of goods per time unit has some uncertainty. The second source of variability is because of the robotised inspection, as the quality criteria are rather ambitious, because the success ratio is not constant and requires significant reworking and reinspection activities. The last and most visible impact is for CHECK units, where the shortage of parts after a geometrical check on SCAN units damages the whole performance. Shortage of parts to be processed at CHECK stations hinders the productivity of these workers, while a type of rigid planning is imposed because of labour regulations and the required union assessment before adoption. Therefore, since it is not possible to dynamically allocate workers to different working places, then the management reaction is to protect the CHECK capacity increasing the intermediate buffer, which is against the lean philosophy and it complicates the shop-floor layout.

3.2.1. Scope Establishment

To illustrate the significance of the issues captured by this study case Figure 5 presents the variability, where neither the finally delivered number of parts nor deviations from what it was initially planned are regular enough.

The strong variability found per day and shift looks interesting as there are robotic operations involved (SCAN units) (see Figure 4 which should add regularity because the more predictable cycle time values). Indeed, due to the labour regulations enforced in the country where such a facility operates, such variability (uncertainty) forces to allocate resources that sometimes are not able to perform as expected, which compromises the business dimension of the activity as a whole, either because of insufficient production or lack of productivity.

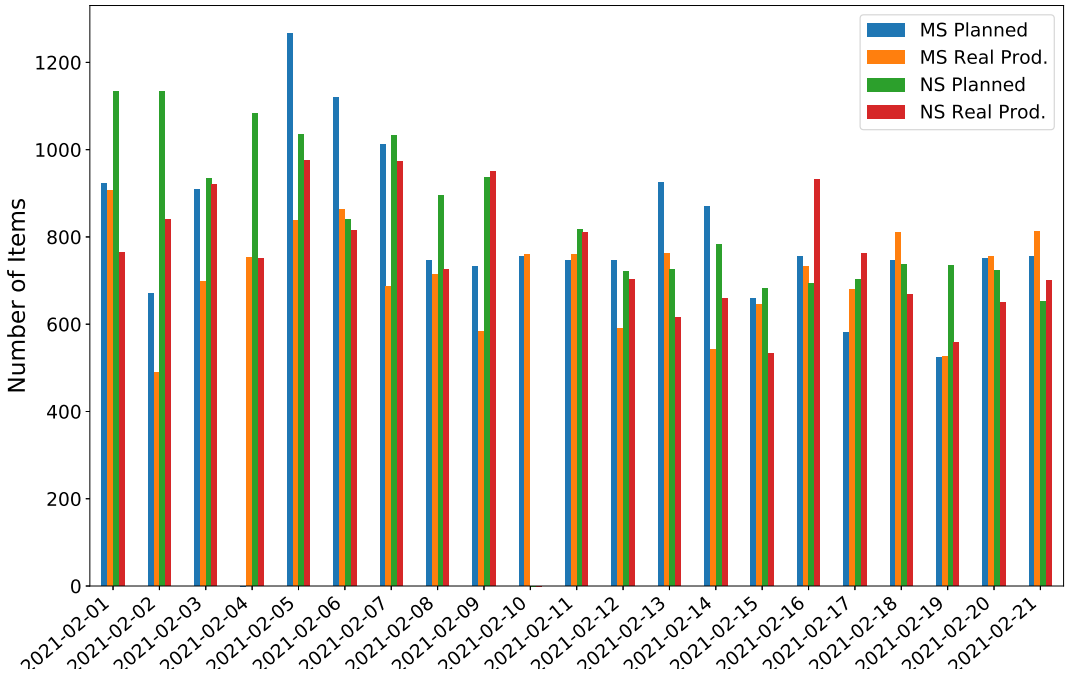


Figure 5. Variability between planned and delivered parts per shift: Morning (MS) and Night Shifts (NS) during February 2021.

3.2.2. Specification of Population and Sampling

To carry out a meaningful analysis, different datasets were collected inside individual processes daily based either on the automation system itself (SCAN units) or by slicing the time in a range of 30 min for manual operation of CHECK units. After several months of data, conclusions can be given in a clearer way. It was accepted that automation of the laser measuring system (SCAN units) continuously ingests products without delay, except those legal stops that are allowed for production, such as lunch time, which are well established. Therefore, more than 90,000 part components have been analysed, as they have been identified as the reference entity.

3.2.3. Data Collection

Two different data streams have been considered as an example for the process-oriented analysis involving both automatic and manual operations, which require more integration and better understanding from the managerial point of view.

The first data stream is related to robotised inspection workplaces, where based on the previous hypothesis it is possible to estimate the most frequent time duration for inspecting each reference successfully, and based on it, to estimate production losses at such stations, making it clear that earlier or later, such production losses will impact the final manual inspection units.

The second one is coming from manual processes and to perform a consistency analysis, fixed time slots where defined and production per slot was measured. Time ranges of 30 min where analysed, looking to identify where production losses occurred. Since every inspection is around three minutes including handling and shipping, that means nine items per every half hour. Therefore, such a ratio was considered the gold standard.

3.2.4. Standardisation Procedure

Standardisation plays an essential role here as it enables to define what the expectations for production are, and how big deviations appear.

By observing the delivery time of the individual parts, measured from the previous successful inspection, we can identify the performance losses occurring when the current part lasts for longer than expected as the most frequent value found for that part reference. This information is presented in Figure 6 just for a few references. Similar behaviour was identified for all different references regularly produced, and based on the findings, conclusions can be derived with implications at the managerial level.

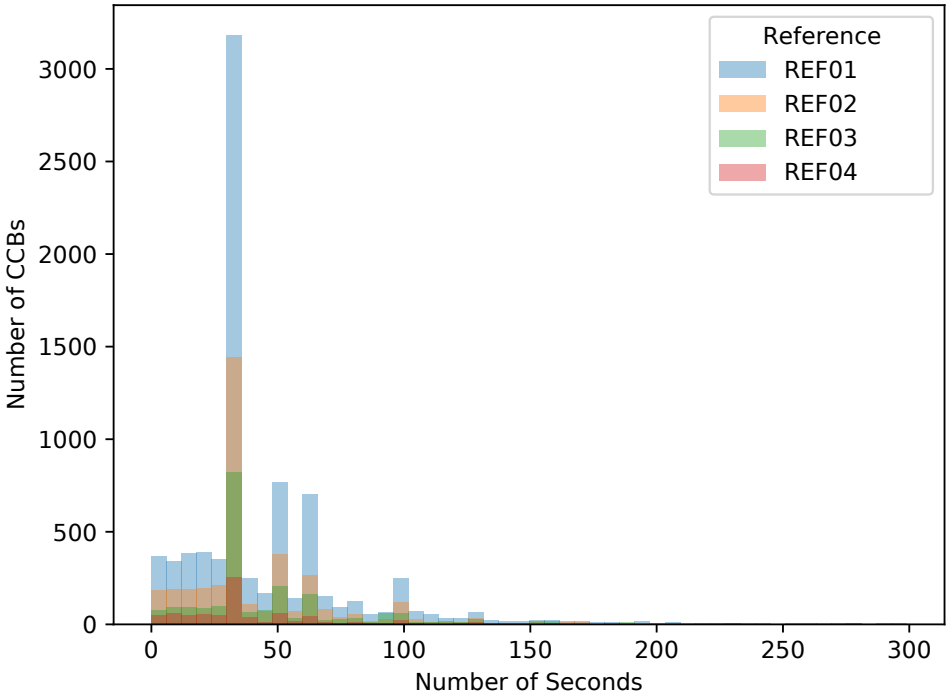


Figure 6. Standardisation for process duration at robotised inspection depending on the reference.

Indeed, when standardisation covers several processes at once, additional and not previously detected sources of errors appear, mainly because the analyses were carried out by individual process units, therefore, the derived impacts become hidden. It is related to references not matching the naming rules because of spelling issues, or different labelling rules for different working units. Such lack of integrity along the value chain does not help to provide a comprehensive perspective of the whole process and this is because standardisation is so relevant during the structured analysis.

3.2.5. Data Analysis

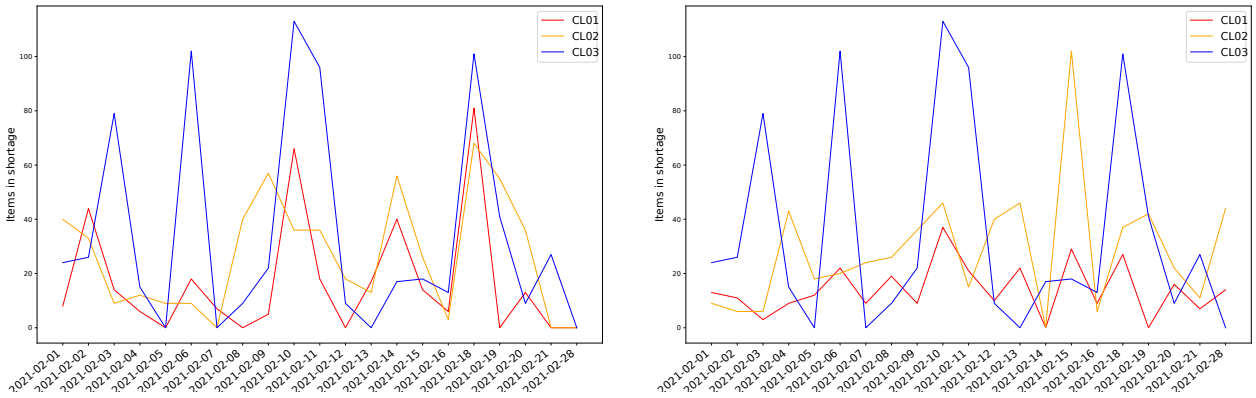
Real evidences show a high impact in the performance of the whole set of processes, where the nonproductive time in the last process exhibits significant variability depending on the process but also depending on the shift, as depicted in Figure 7.

For the robotised inspection units, losses can be also estimated, as depicted in Figure 8

A relevant aspect evidencing that the shortage issues observed at CHECK units are actually due to the previous production steps can be observed from Figure 6, and it can be observed equally for any of the more than fifteen different part references manufactured at the shop-floor. It is derived from the histogram used to identify the effective time required for a reference to be successfully processed.

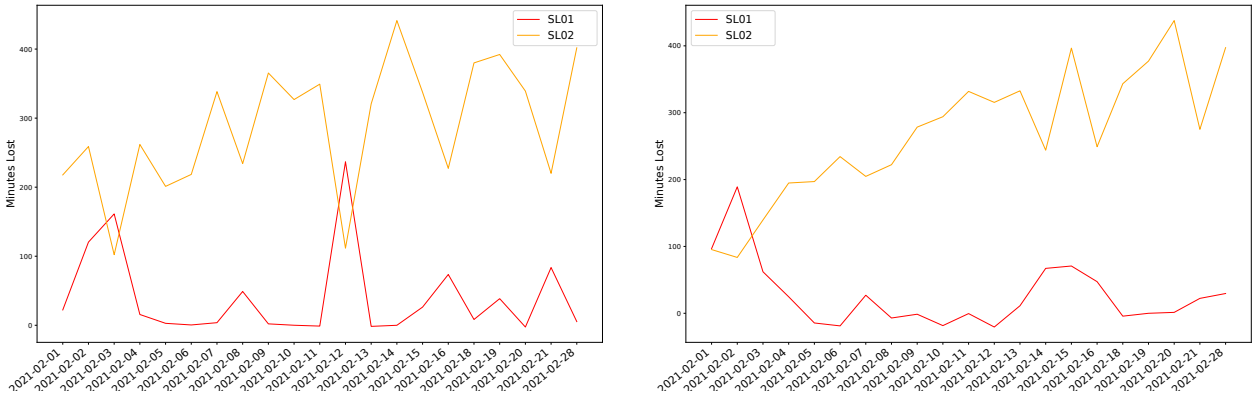
If the part reference REF01 is considered, it becomes clear that in most of the cases successful processing of this reference at the SCAN units lasts for 33 s. However, it is possible to realise that there are a relevant number of cases where it lasts for a shorter time, and in some cases it takes for longer. The main reason for shorter times is because when the robot decides the part fails, it does not continue to explore all positions and it rejects such a part, expending a shorter time than when the part is correct. On the opposite side, there are a significant number of parts lasting for 50 and 60 s, which means that after the

last part successfully processed, there were several others with the same reference failing in fulfilling the QC requirements, therefore, after a while another part was OK but it took much longer than 33 s. From this situation, it becomes evident that SCAN units are testing a very relevant number of parts requiring reworking and demanding inspection several times. The direct effect of loosing efficiency is that the situation hinders process stability for the next production unit.



(**a**) Morning Shift (**b**) Night Shift

Figure 7. Production Losses at last inspection station per production line and Shift.



(**a**) Morning Shift (**b**) Night Shift

Figure 8. Production Losses at SCAN inspection units per production line and Shift.

From the comparative analysis between SCAN and CHECK inspection units, it becomes clear how mitigation strategies such as enlarging part buffers between units have kept the situation bounded but with a high level of variability, but the effectiveness of SCAN continues to get degraded over time, mainly because managers are much more focused on the more evident problem of part shortage.

It is also worth to mention the different behaviour for the two robotic inspection lines presented in Figure 8, where SL01 is more stable and SL02 shows that it is not under control neither for the morning shift nor for the night one.

4. Discussion

When the first case study is considered, the integration of data from different sources, including routes, position, indoor and outdoor climate, NFC data provides accountability for different process steps. The integration was carried out because of the I4.0 principles and it allows to collect evidences capable of explaining process variability with increased confidence than ever before. Just to illustrate such reality, before deploying this project, managers in the company were explaining the variability in process duration because of weather conditions or because of driver attitudes, and in some cases because of route

congestion. Clearly, they based their analyses mainly on beliefs or personal experiences, but after the analysis carried out, it becomes obvious that the main sources of variability are not because of those reasons but because of bottlenecks in the input weighting system at the customer's site. With lower frequency but yet importantly, the time spent inside the customer site is also significant, but it was not the particular shift, the truck or the driver involved.

The deployed technology and analysis capability can be easily repeated on a regular basis, trying to inform managers about the effectiveness of the adopted measures, but also allowing to look for different objectives (RHS filtering), or even to predict the duration of specific process steps based on different parameters.

Regarding the second case study, it is needed to recognise that the processing logic is here rather complex because the components come from different manufacturers and at the shop-floor there are different stations involved in the whole Value Stream Mapping (VSM) producing different additional components, all of them welded and assembled at other production units, where the final quality control (QC) is done, both on SCAN units and then on CHECK ones. Because of such nonlinear circuit, since when the part is rejected due to QC reasons, it is extracted, reworked accordingly, and resubmitted, as well as because of the previously explained complex process with material flow through a large amount of steps and the parts becoming mixed between stations and part variants, it is very hard to track the different causes for failures. Therefore, the classical applications of Pareto techniques to identify main sources of failures do not work.

Sources of failure can be generally related to different root causes including workers in previous production stations, but also to the process and inspection systems, having more than one hundred of those identified causes, non-regularly distributed over time or part reference.

In this complex environment, only the I4.0 approach becomes effective, as the system collects information from workers' performance per working place, the period of time, and references produced, while additional postprocessing is needed because sometimes the number of nonconformity events per part is much larger than expected, for instance additional bending of the part can negatively impact on different distances and angles between points. It also collects information from the automatic scanning system and from workers' wearable elements, helping to complement the process perspective.

The advantage of the adopted approach as described in this case is that it increases the transparency of the processes as well as the effectiveness of the managerial decisions made. Instead of relaying on impressions from different people which, although very skilled, also have their own biases, the tools provide an agnostic perspective on the process and their impact.

Just as a small example, it was found that as opposite to the initial thoughts suggested by the internal experts, worker participation at CHECK units looked to moderate the impacts of uncertainties and part shortages. This is because different strategies can be promoted looking to minimise wasted extra capacity in specific periods, while automatic systems are much more rigid and require technical interventions lasting for a longer period of time. Therefore, human contribution provides significant flexibility to the processes, although the lack of production is still there.

The interest of the company now is to better integrate previous production steps, which in the beginning seemed rather independent, but new technologies are required to get them much more integrated into the analysis, continuing increasing the transparency.

5. Conclusions

This paper tries to highlight the significant contribution of the I4.0 framework to hybrid processes, where automated but also manual processes are required to cooperate and where most of the management strategies based on the 'split and win' approach fail to provide consistent evidences to improve the business.

Major contribution is related to increasing transparency from an agnostic perspective, avoiding bias because of beliefs or other reasons. Another significant contribution is the integrative perspective brought, as it enables rather easily to go through a very high number of items, being processed in different ways and different places and in different time periods, connecting all points much more consistently than even before. In addition, a hidden benefit is connected with the incremental characteristic of the analyses because of the different requests raised by the managerial aspects can be better addressed with the additional information provided by means of the I4.0 and the integration of process-oriented data streams. This is relevant because providing answers to a managerial question takes the interest towards the next one, which connects with scalability of the research.

A relevant aspect highlighted by this research is that a consistent, deep, and transparent analysis can be carried out, still protecting workers' identities by avoiding explicitly placing the focus on them, as required by the ethical AI principles.

It is needed to recognise that this paper is not claiming that the application of extended I4.0 technology to process-oriented integrated data flows will get the same level of benefits, because it will be case dependent (process and management). By exploring the presented two cases, the evident limitation related to the scarcity of the sample appears clearly. However, as they cover different business units from different sectors (logistics and production), with different complexity levels, and they brought abstract properties increasing the existing prior knowledge about the reasons for process variability, we believe they can be applied in other cases as well, with a similar increase in the existing knowledge. Based on these facts, this paper found enough support to positively answer the identified research question.

In terms of future activities, since the interests of companies are business driven, but at different speed. In the first case, they are interested in developing forecasting models and learning about their robustness able to better schedule the logistics activities, while in the second one they are just focused in increasing the understanding of the complex production process they are managing. In this particular case, they are aligned with requirements from the I4.0 paradigm, which is rather pervasive and to check influences from any source, looks to ingest significant behaviours from all relevant shop-floor units. Indeed, to complete the interaction requirements tracking, some wearable devices are under consideration, since they can help to check process variability and related effort from workers.

Finally, a common request is to bring a convenient way to present relevant information to managers in an automatic way and rule driven. To this end, the current implementation uses light clients with plotly and trello tools but probably they will come up with additional requirements in line with process evolution as well as the need for improvements based on the decisions made. Therefore, integrating forecasting capabilities as well as a more integrated way to describe the VSM are under further investigation. All of them look to contribute to the health of the VSM, as an extension of the well-known concept of assets' health.

Author Contributions: Conceptualization, J.O.-M.; methodology, J.V.-D.; software, J.O.-M. ans J.V.-D.; validation, J.V.-D.; formal analysis, J.V.-D.; investigation, J.O.-M.; resources, J.V.-D.; data curation, J.V.-D.; writing—original draft preparation, J.V.-D.; writing—review and editing, J.V.-D.; visualization, J.V.-D.; supervision, J.O.-M. and J.V.-D.; project administration, J.V.-D.; funding acquisition, J.O.-M. All authors have read and agreed to the published version of the manuscript.

Funding: The APC was funded by the Spanish Agencia Estatal de Investigacion through the project RTI2018-094614-B-I00 (Smashing).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to restrictions from the companies involved in the case studies.

Acknowledgments: The authors want to thank the WISEST consortium because of their support for the conducted research, as well as by the Spanish Agencia Estatal de Investigacion, through the research project with code RTI2018-094614-B-I00, which provided specific software tools.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
BLE	Bluetooth Low Energy
CQD	Coefficient Interquartile of Dispersion
CV	Coefficients of Variation
H–CPS	Human–Cyber–Physical Systems
I4.0	Industry 4.0
IIoT	Industrial Internet of Things
KPI	Key Performance Indicator
NFC	Near-field communication
O4.0	Operator 4.0
VSM	Value Stream Mapping
ethical AI	Ethical Artificial Intelligence
DM	Data Mining
FP-Growth	Frequent-Pattern Growth
R.H.S.	Right Hand Side of an Association Rule

References

1. Lee, J.; Bagheri, B.; Kao, H.A. A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems. *Manuf. Lett.* **2015**, *3*, 18–23. [CrossRef]
2. Song, H.; Rawat, D.; Jeschke, S.; Brecher, C. (Eds.) *Cyber-Physical Systems*; Academic Press: New York, NY, USA, 2017.
3. Schuh, G.; Zeller, V.; Stroh, M.F.; Harder, P. Finding the Right Way Towards a CPS—A Methodology for Individually Selecting Development Processes for Cyber-Physical Systems. In *Collaborative Networks and Digital Transformation*; Camarinha-Matos, L.M., Antonelli, D., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 81–90.
4. Lodgaard, E.; Dransfeld, S. Organizational aspects for successful integration of human–machine interaction in the industry 4.0 era. In Proceedings of the 13th CIRP Conference on Intelligent Computation in Manufacturing Engineering, Naples, Italy, 17–19 July 2019; Volume 88, pp. 218–222.
5. Nardo, M.; Forino, D.; Murino, T. The evolution of man–machine interaction: The role of human in Industry 4.0 paradigm. *Prod. Manuf. Res.* **2020**, *8*, 20–34. [CrossRef]
6. Krupitzer, C.; Müller, S.; Lesch, V.; Züfle, M.; Edinger, J.; Lemken, A.; Becker, C. A Survey on Human Machine Interaction in Industry 4.0. *arXiv* **2020**, arXiv:2002.01025.
7. Ordieres-Meré, J.; Villalba-Díez, J.; Zheng, X. Challenges and Opportunities for Publishing IIoT Data in Manufacturing as a Service Business. *Procedia Manuf.* **2019**, *39*, 185–193. [CrossRef]
8. Khan, W.Z.; Rehman, M.H.; Zangoti, H.M.; Afzal, M.K.; Armi, N.; Salah, K. Industrial internet of things: Recent advances, enabling technologies and open challenges. *Comput. Electr. Eng.* **2020**, *81*, 106522. [CrossRef]
9. Evjemo, L.D.; Gjerstad, T.; Grøtli, E.I.; Sziebig, G. Trends in Smart Manufacturing: Role of Humans and Industrial Robots in Smart Factories. *Curr. Robot. Rep.* **2020**, *1*, 35–41. [CrossRef]
10. Jardim-Goncalves, R.; Romero, D.; Grilo, A. Factories of the future: Challenges and leading innovations in intelligent manufacturing. *Int. J. Comput. Integr. Manuf.* **2017**, *30*, 4–14. [CrossRef]
11. Zheng, X.; Wang, M.; Ordieres-Meré, J. Comparison of Data Preprocessing Approaches for Applying Deep Learning to Human Activity Recognition in the Context of Industry 4.0. *Sensors* **2018**, *18*, 2146. [CrossRef]
12. Ghosh, A.; Edwards, D.J.; Hosseini, M.R.; Al-Ameri, R.; Abawajy, J.; Thwala, W.D. Real-time structural health monitoring for concrete beams: A cost-effective "Industry 4.0" solution using piezo sensors. *Int. J. Build. Pathol. Adapt.* **2021**, *39*, 283–311. [CrossRef]
13. Aceto, G.; Persico, V.; Pescapé, A. Industry 4.0 and Health: Internet of Things, Big Data, and Cloud Computing for Healthcare 4.0. *J. Ind. Inf. Integr.* **2020**, *18*, 100129. [CrossRef]
14. Bohé, I.; Willocx, M.; Naessens, V. An Extensible Approach for Integrating Health and Activity Wearables in Mobile IoT Apps. In Proceedings of the 2019 IEEE International Congress on Internet of Things (ICIOT), Milan, Italy, 8–13 July 2019; pp. 69–75.
15. Grangel-González, I.; Halilaj, L.; Coskun, G.; Auer, S.; Collarana, D.; Hoffmeister, M. Towards a Semantic Administrative Shell for Industry 4.0 Components. In Proceedings of the 2016 IEEE Tenth International Conference on Semantic Computing (ICSC), Laguna Hills, CA, USA, 4–6 February 2016; pp. 230–237. [CrossRef]

16. Batty, M. Digital twins. *Environ. Plan. Urban Anal. City Sci.* **2018**, *45*, 817–820. [CrossRef]
17. El Saddik, A. Digital Twins: The Convergence of Multimedia Technologies. *IEEE Multimed.* **2018**, *25*, 87–92. [CrossRef]
18. Qi, Q.; Tao, F. Digital Twin and Big Data Towards Smart Manufacturing and Industry 4.0: 360 Degree Comparison. *IEEE Access* **2018**, *6*, 3585–3593. [CrossRef]
19. Guerin, C.; Rauffet, P.; Chauvin, C.; Martin, E. Toward production operator 4.0: Modelling Human-Machine Cooperation in Industry 4.0 with Cognitive Work Analysis. In Proceedings of the 14th IFAC Symposium on Analysis, Design, and Evaluation of Human Machine Systems HMS, Tallinn, Estonia, 16–19 September 2019; Volume 52; pp. 73–78. [CrossRef]
20. Pacaux-Lemoine, M.P.; Trentesaux, D. Ethical risks of human-machine symbiosis in industry 4.0: Insights from the human-machine cooperation approach. In Proceedings of the 14th IFAC Symposium on Analysis, Design, and Evaluation of Human Machine Systems HMS, Tallinn, Estonia, 16–19 September 2019; Volume 52; pp. 19–24. [CrossRef]
21. Scafà, M.; Marconi, M.; Germani, M. A critical review of symbiosis approaches in the context of Industry 4.0. *J. Comput. Des. Eng.* **2020**, *7*, 269–278. [CrossRef]
22. Fletcher, S.R.; Johnson, T.; Adlon, T.; Larreina, J.; Casla, P.; Parigot, L.; del Mar Otero, M. Adaptive automation assembly: Identifying system requirements for technical efficiency and worker satisfaction. *Comput. Ind. Eng.* **2020**, *139*, 105772. [CrossRef]
23. Kaasinen, E.; Schmalfuß, F.; Özturk, C.; Aromaa, S.; Boubekeur, M.; Heilala, J.; Walter, T. Empowering and engaging industrial workers with Operator 4.0 solutions. *Comput. Ind. Eng.* **2020**, *139*, 105678. [CrossRef]
24. Zhou, J.; Zhou, Y.; Wang, B.; Zang, J. Human–Cyber–Physical Systems (HCPSs) in the Context of New-Generation Intelligent Manufacturing. *Engineering* **2019**, *5*, 624–636. [CrossRef]
25. Villalba-Diez, J.; Ordieres-Mere, J. Improving manufacturing operational performance by standardizing process management. *Trans. Eng. Manag.* **2015**, *62*, 351–360. [CrossRef]
26. Villalba-Diez, J. *The Hoshin Kanri Forest. Lean Strategic Organizational Design*, 1st ed.; CRC Press; Taylor and Francis Group LLC: Boca Raton, FL, USA, 2017.
27. Villalba-Diez, J. *The Lean Brain Theory. Complex Networked Lean Strategic Organizational Design*; CRC Press; Taylor and Francis Group LLC: Boca Raton, FL, USA, 2017.
28. Burton, R.M.; Obel, B. *Strategic Organizational Diagnosis and Design: The Dynamics of Fit*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 2004.
29. Alberts, D.S. Rethinking Organizational Design for Complex Endeavors. *J. Organ. Des.* **2012**, *1*, 14–17.
30. Burton, R.M.; Øbel, B.; Håkonsson, D.D. *Organizational Design: A Step-by-Step Approach*, 3rd ed.; Cambridge University Press: Cambridge, UK, 2015.
31. Cross, R.L.; Singer, J.; Colella, S.; Thomas, R.J.; Silverstone, Y. *The Organizational Network Fieldbook: Best Practices, Techniques and Exercises to Drive Organizational Innovation and Performance*, 1st ed.; Jossey-Bass: San Francisco, CA, USA, 2010.
32. Jabeur, N.; Sahli, N.; Zeadally, S. Enabling Cyber Physical Systems with Wireless Sensor Networking Technologies, Multiagent System Paradigm, and Natural Ecosystems. *Mob. Inf. Syst.* **2015**, *2015*, 15. [CrossRef]
33. Fujimoto, T. *Evolution of Manufacturing Systems at Toyota*; Productivity Press: Portland, OR, USA, 2001.
34. Durugbo, C.; Tiwari, A.; Alcock, J.R. Modelling information flow for organisations: A review of approaches and future challenges. *Int. J. Inf. Manag.* **2013**, *33*, 597–610. [CrossRef]
35. Powell, D.; Romero, D.; Gaiardelli, P.; Cimini, C.; Cavalieri, S. Towards Digital Lean Cyber-Physical Production Systems: Industry 4.0 Technologies as Enablers of Leaner Production. In *Advances in Production Management Systems. Smart Manufacturing for Industry 4.0*; Moon, I., Lee, G.M., Park, J., Kiritsis, D., von Cieminski, G., Eds.; Springer International Publishing: Cham, Switzerland, 2018; pp. 353–362.
36. Romero, D.; Wuest, T.; Stahre, J.; Gorecky, D. Social Factory Architecture: Social Networking Services and Production Scenarios Through the Social Internet of Things, Services and People for the Social Operator 4.0. In *Advances in Production Management Systems. The Path to Intelligent, Collaborative and Sustainable Manufacturing*; Lödding, H., Riedel, R., Thoben, K.-D., von Cieminski, G., Kiritsis, D., Eds.; Springer International Publishing: Cham, Switzerland, 2017; pp. 265–273.
37. Sun, S.; Zheng, X.; Villalba-Díez, J.; Ordieres-Meré, J. Indoor Air-Quality Data-Monitoring System: Long-Term Monitoring Benefits. *Sensors* **2019**, *19*, 4157. [CrossRef]
38. Hellebrandt, T.; Ruessmann, M.; Heine, I.; Schmitt, R.H. Conceptual Approach to Integrated Human-Centered Performance Management on the Shop Floor. In *Advances in Human Factors, Business Management and Society*; Kantola, J.I., Nazir, S., Barath, T., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 309–321.
39. Schilling, K.; Storms, S.; Herfs, W. Environment-Integrated Human Machine Interface Framework for Multimodal System Interaction on the Shopfloor. In *Advances in Human Factors and Systems Interaction*; Nunes, I.L., Ed.; Springer International Publishing: Cham, Switzerland, 2019; pp. 374–383.
40. Tao, F.; Qi, Q.; Wang, L.; Nee, A.Y.C. Digital Twins and Cyber–Physical Systems toward Smart Manufacturing and Industry 4.0: Correlation and Comparison. *Engineering* **2019**, *5*, 653–661. [CrossRef]
41. Al-Masri, E. Enhancing the Microservices Architecture for the Internet of Things. In Proceedings of the 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 10–13 December 2018; pp. 5119–5125. [CrossRef]
42. Ullah, I.; Ul Amin, N.; Zareei, M.; Zeb, A.; Khattak, H.; Khan, A.; Goudarzi, S. A Lightweight and Provable Secured Certificateless Signcryption Approach for Crowdsourced IIoT Applications. *Symmetry* **2019**, *11*, 1386. [CrossRef]

43. Sun, L.; Li, Y.; Memon, R.A. An open IoT framework based on microservices architecture. *China Commun.* **2017**, *14*, 154–162. [CrossRef]
44. Villalba-Diez, J.; Zheng, X.; Schmidt, D.; Molina, M. Characterization of Industry 4.0 Lean Management Problem-Solving Behavioral Patterns Using EEG Sensors and Deep Learning. *Sensors* **2019**, *19*, 2841. [CrossRef]
45. Schmidt, D.; Villalba Diez, J.; Ordieres-Meré, J.; Gevers, R.; Schwiep, J.; Molina, M. Industry 4.0 Lean Shopfloor Management Characterization Using EEG Sensors and Deep Learning. *Sensors* **2020**, *20*, 2860. [CrossRef]
46. Inshakova, A.O.; Frolova, E.E.; Rusakova, E.P.; Kovalev, S.I. The model of distribution of human and machine labor at intellectual production in industry 4.0. *J. Intellect. Cap.* **2020**, *21*, 601–622. [CrossRef]
47. López-Núñez, M.I.; Rubio-Valdehita, S.; Diaz-Ramiro, E.M.; Aparicio-García, M.E. Psychological Capital, Workload, and Burnout: What's New? The Impact of Personal Accomplishment to Promote Sustainable Working Conditions. *Sustainability* **2020**, *12*, 8124. [CrossRef]
48. Emami, Z.; Chau, T. The effects of visual distractors on cognitive load in a motor imagery brain–computer interface. *Behav. Brain Res.* **2020**, *378*, 112240. [CrossRef] [PubMed]
49. Carvalho, A.V.; Chouchene, A.; Lima, T.M.; Charrua-Santos, F. Cognitive Manufacturing in Industry 4.0 toward Cognitive Load Reduction: A Conceptual Framework. *Appl. Syst. Innov.* **2020**, *3*, 55. [CrossRef]
50. Chaijaroen, N.; Jackpeng, S.; Chaijaroen, S. The Development of Constructivist Web-Based Learning Environments to Enhance Learner's Information Processing and Reduce Cognitive Load. In *Innovative Technologies and Learning*; Huang, T.-C., Wu, T.-T., Barroso, J., Eika Sandnes, F.E., Martins, P., Huang, Y.-M., Eds.; Springer International Publishing: Cham, Switzerland, 2020; pp. 475–482.
51. Thees, M.; Kapp, S.; Strzys, M.P.; Beil, F.; Lukowicz, P.; Kuhn, J. Effects of augmented reality on learning and cognitive load in university physics laboratory courses. *Comput. Hum. Behav.* **2020**, *108*, 106316. [CrossRef]
52. Liu, H.; Wang, L. Remote human–robot collaboration: A cyber–physical system application for hazard manufacturing environment. *J. Manuf. Syst.* **2020**, *54*, 24–34. [CrossRef]
53. Li, X.; Wang, L.; Zhu, C.; Liu, Z. Framework for manufacturing-tasks semantic modelling and manufacturing-resource recommendation for digital twin shop-floor. *J. Manuf. Syst.* **2021**, *58*, 281–292. [CrossRef]
54. Caiza, G.; Nuñez, A.; Garcia, C.A.; Garcia, M.V. Human Machine Interfaces Based on Open Source Web-Platform and OPC UA. *Procedia Manuf.* **2020**, *42*, 307–314. [CrossRef]
55. Kinne, S.; Jost, J.; Terharen, A.; Feldmann, F.; Fiolka, M.; Kirks, T. Process Development for CPS Design and Integration in I4. 0 Systems with Humans. In *Digital Supply Chains and the Human Factor*; Springer: Cham, Switzerland, 2021; pp. 17–32.
56. Tortorella, G.; Sawhney, R.; Jurburg, D.; de Paula, I.C.; Tlapa, D.; Thurer, M. Towards the proposition of a lean automation framework: Integrating industry 4.0 into lean production. *J. Manuf. Technol. Manag.* **2020**, *32*, 593–620. [CrossRef]
57. Terry Anthony Byrd, D.E.T. Measuring the flexibility of information technology infrastructure: Exploratory analysis of a construct. *J. Manag. Inf. Syst.* **2000**, *17*, 167–208.
58. Eisenhardt, K.M. Building theories from case study research. *Acad. Manag. Rev.* **1989**, *14*, 532–550. [CrossRef]
59. Ordieres, J. *jbmere/HealthOperator4.0 v1.0*; Zenodo: Geneva, Switzerland, 2020.
60. Ordieres, J. *jbmere/GesOperNFC: First Release*; Zenodo: Geneva, Switzerland, 2021.
61. Bradshaw, S.; Brazil, E. *MongoDB: The Definitive Guide*, 3rd ed.; O'Reilly Media Inc.: Sebastopol, CA, USA, 2019.
62. Kenzler, E.; Razzoli, F. *MariaDB Essentials*; Packt Publishing: Birmingham, UK, 2015.
63. Fengyi, D.; Zhenyu, L. An ameliorating FP-growth algorithm based on patterns-matrix. *J. Xiamen Univ. (Nat. Sci.)* **2005**, *44*, 629–633.
64. Zhichun, L.; Fengxin, Y. An improved frequent pattern tree growth algorithm. *Appl. Sci. Technol.* **2008**, *35*, 47–51.
65. Jun, C.; Li, G. An improved FP-growth algorithm based on item head table node. *Inf. Technol.* **2013**, *12*, 34–35.

sensors

MDPI

Article

Quantum Strategic Organizational Design: Alignment in Industry 4.0 Complex-Networked Cyber-Physical Lean Management Systems

Javier Villalba-Diez [1,2,*] and Xiaochen Zheng [3]

[1] Fakultät Management und Vertrieb, Campus Schwäbisch Hall, Hochschule Heilbronn, 74523 Schwäbisch Hall, Germany
[2] Complex Systems Group, Escuela Técnica Superior de Ingenieros Agrónomos, Universidad Politécnica de Madrid, Av. Puerta de Hierro 2, 28040 Madrid, Spain
[3] ICT for Sustainable Manufacturing, SCI-STI-DK, École Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland; xiaochen.zheng@epfl.ch
* Correspondence: javier.villalba-diez@hs-heilbronn.de

Received: 24 September 2020; Accepted: 14 October 2020; Published: 16 October 2020

Abstract: The strategic design of organizations in an environment where complexity is constantly increasing, as in the cyber-physical systems typical of Industry 4.0, is a process full of uncertainties. Leaders are forced to make decisions that affect other organizational units without being sure that their decisions are the right ones. Previously to this work, genetic algorithms were able to calculate the state of alignment of industrial processes that were measured through certain key performance indicators (KPIs) to ensure that the leaders of the Industry 4.0 make decisions that are aligned with the strategic objectives of the organization. However, the computational cost of these algorithms increases exponentially with the number of KPIs. That is why this work makes use of the principles of quantum computing to present the strategic design of organizations from a novel point of view: Quantum Strategic Organizational Design (QSOD). The effectiveness of the application of these principles is shown with a real case study, in which the computing time is reduced from hundreds of hours to seconds. This has very powerful practical applications for industry leaders, since, with this new approach, they can potentially allow a better understanding of the complex processes underlying the strategic design of organizations and, above all, make decisions in real-time.

Keywords: quantum computing; strategic organizational design; Industry 4.0; complex networks; cyber-physical systems; lean management systems

1. Introduction

The most significant aspect of strategic planning in an organization is, according to Grant [1], the strategic process: "a dialog through which knowledge is shared and the consensus is achieved and commitment towards action and results is built". To unify common efforts and, hence, support the strategic organizational goals, it is during this dialogue, previously described as Nemawashi [2] or "catch-ball" [3] by scholars, that a sometimes delicate balance of forces is sought between the interests of different organizational agents [4]. Under a strategic organizational design paradigm [5,6], the interaction of these interdependent organizational agents shapes hierarchically nested complex networks [7] that support decision making towards, ideally, a coordinated effort to attain organizational strategic goal achievement, called organizational alignment. These alignment efforts can occur in different organizational settings, although, in this paper the authors focus on complex networked cyber-physical systems in an Industry 4.0 context.

For any given time t, complex cyber-physical networks have been formally described [8] as time-dependent graphs given by Equation (1):

$$\Omega(t) = [\Gamma(t); E(t)] \quad (1)$$

which can be understood as lists of $\Gamma(t)$ human and cyber-physical nodes and its standard communication $E(t) \subset (\Gamma(t)x\Gamma(t))$ edges [9]. The very emergence of complex networked organizational design configurations in the form of lean structural networks is only possible through a continuous improvement-oriented standardization of the organizational network edges, the business communication protocols, between the network elements [9]. These boundary conditions allow for representing the systems of Industry 4.0 as cyber-phsycial complex networks, allowing a systematic and quantitative analysis of the systems by means of lean management algorithms strategically oriented to the systematic reduction of the variability of the value creation processes. For this reason, this work is focused solely on lean management systems in an Industry 4.0 cyber-physical context. The standard repetition of *Check, Plan, Do* n-times—with a subsequent process standardization in *Act*, in short *(CPD)nA*, allows for a comprehensive management of such networks through HOSHIN KANRI FOREST [10] configurations, as shown in Figure 1.

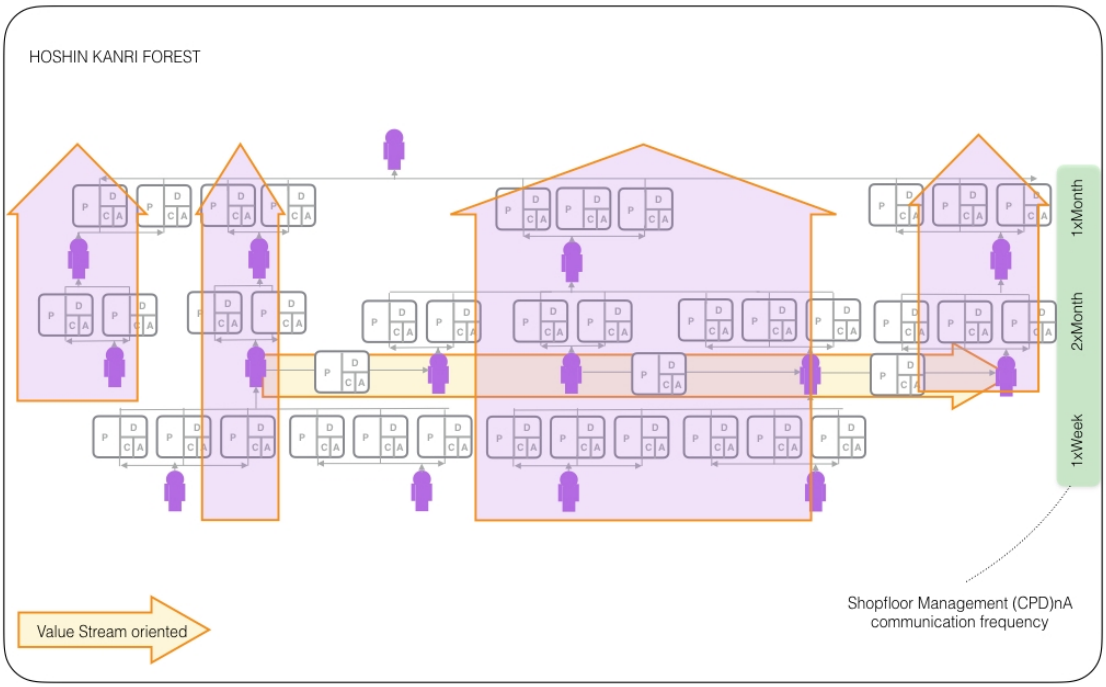


Figure 1. Hoshin Kanri Forest Structure [11].

Scholars have proposed approaches to qualitatively model organizational alignment [12–20]. Approaches that allow for a quantification of organizational alignment are less common [2], which shows the alignment state of each node is known at each discrete time interval. However, the NEMAWASHI approach, based on genetic algorithms, is computationally very costly and, therefore, difficult to implement in practice. While calculating the alignment state of the entire network is theoretically possible with this method, in practice, it is a challenge that leads to an exponential increase of calculation time with augmenting network size. For this reason, there is an urgent need to provide organizational leaders with a fast algorithm that allows for a calculation of the alignment status of the organization.

Quantum computing is a novel computation paradigm that might prove useful to this end [21]. Quantum computing examines the flow and processing of information as physical phenomena that follow the laws of quantum mechanics. This is possible, because quantum computing makes use of "superposition", which is the ability of quantum computers to be simultaneously in multiple different states [22]. By doing so, quantum computing has shown promising performance increases in solving certain unassailable problems for classic computing, such as Shor's algorithm [23] and Grover's algorithm [24]. The purpose of this work is to propose an efficient quantum computation algorithm that is capable of discerning the organizational state of alignment of such complex networks and, therefore, support the leaders of organizations in their decision-making process.

To do so, we start with an initial hypothesis formulation. The intrinsically interdependent nature of lean complex cyber-physical networks allows for us to reasonably establish the initial hypothesis H of this work: the decision-making associated with the alignment process lacks absolute certainty.

In other words, leading an organization toward the coordinated achievement of strategic objectives is a probabilistic process, in which decision-makers can never be certain that the choice being made is the correct one. Decision-makers are conditioned by other organizational agents' simultaneous decisions whose consequences cannot be completely foreseen a-priori. Consequently, these networks can be considered as decision networks or probabilistic directed acyclic graphical models [25] with known conditional probabilities of alignment.

The rest of the work hereinafter continues, as follows: first, Section 2 begins by defining some essential preliminary concepts for the precise use of terminology in the following sections of this work. Second, Section 3 presents the main contribution of this work by outlining the methodology and design principles proposed by this work to achieve a quantum formulation of strategic organizational design (QSOD). Third, Section 4 presents a case study to show the implementation of the theoretical design principles explained above. This is intended to allow for a better understanding and make the benefits that the application of this methodology may have for the leaders of the Industry 4.0 more explicit. Finally, Section 5 briefly summarizes the results that were obtained, as well as showing future lines of research and certain limitations to the work.

2. Background

This section starts by defining some preliminary concepts fundamental to the precise terminology use of the content presented in the following sections of this work:

- Industry 4.0. Industry 4.0 has gained a lot of attention since it was first released [26], claiming the necessity of a new paradigm shift in favour of a less centralized manufacture structure. It is regarded as the fourth industrial revolution, the first three being mechanization through the use of vapor energy, mass production through electricity generation, and ultimately the digital revolution through the integration of electronics and information technology. The industry 4.0 ought to allow for a larger independence of the manufacturing process, since technology is more interrelated and the machines can interact with each other creating a cyber-physical system [27–32].
- Cyber-Physical Systems. Cyber-physical system in the context of Industry 4.0 relates to the close bonding and alignment between computing and material resources. A new paradigm of technological systems that is based on embedded collaborative software is impacting the development of such systems [33–35].
- Lean Management. Lean management systems in a cyber-physical environment of Industry 4.0 are described as socio-technical structures that are designed to consistently reduce the variability of value creating processes and, therefore, increase their effectiveness and profitability [9,36–46].

  Variability in this context is understood as any deviation from the desired process state. In quantifiable terms, this work understands the variability of a process, as measured by the

systematic reduction of the standard deviation that is associated with the indicators measuring its performance [9,47].

- Complex Networked Organizational Design. According to the network organizational paradigm, modern cyber-physical systems that are oriented to the lean management of Industry 4.0 can be seen as a socio-technical symbiotic ecosystem of human networks [5] interacting with distributed physical sensors interconnected in an increasingly complex network interconnected sensors [48], which readings are modeled as time-dependent signals at the vertices, human, or cyber-physical, respectively. That means that in the structure of the network nodes you can find characteristics that represent them in the form of a certain time series that describes the key performance indicators (KPIs).

On the basis of the previous concepts, the lean management of complex cyber-physical systems networked in an Industry 4.0 context, may be defined as business systems that seek systematically to decrease the inherent variability of industrial value creation processes, considering them to be complex networks of interdependent computational and physical elements. Effective and efficient calculation of the information that flows through these elements is the key factor for achieving lasting and sustained business success.

- Alignment. This information is typically described by a series of KPIs. Such KPIs are interdependent and they describe certain trajectories in node-related orthonormal bases [2]. These scholars define a node to be in alignment at any given moment in time if the KPI's trajectory presents asymptotic stability at this point [49]. In other words, the condition for alignment at any given time interval t+Δt is given by Equation (2)

$$\forall \Delta t > 0 \quad D_{t,t+\Delta t} < D_{t-\Delta t,t} \tag{2}$$

where $D_{i,j}$ represents the euclidean distance between two points i and j in the KPI's trajectory. Consequently, the probability that the node is not in alignment is given by Equation (3)

$$\forall \Delta t > 0 \quad D_{t,t+\Delta t} \geq D_{t-\Delta t,t} \tag{3}$$

where $D_{i,j}$ represents the euclidean distance between two points i and j in the KPI's trajectory. Thus, alignment is a binary property of each node. Furthermore, since the trajectories are known $\forall t$, we can calculate the conditional probability that the nodes within the complex networked SOD are simultaneously in alignment or not, by the simple application of the well known Bayes theorems.

In fact, within this time interval Δt, the graph Ω that is described in Equation (1) converts into a decision network $\Omega' = [\Gamma', E']$ formed by a set of Γ' nodes and E' edges, where $\Gamma' = [\gamma_1, \gamma_2, \ldots, \gamma_N]$ represents the set of all the nodes being part of the network in Δt, and the edges are determined by the known probabilistic dependence of alignment occurrence in a node γ_j, depending on the alignment occurrence on another γ_i. The node γ_i is thus called *parent* and node γ_j the child. The root nodes are those that do not depend on any other. Subsequently, as described by [50], the joint probability on the nodes can be decomposed into the product of the marginal probabilities that are given by Equation (4):

$$P(\gamma_1, \gamma_2, \ldots, \gamma_N) = \prod_{i=1}^{N} P\left(\gamma_i | \prod \gamma_i\right) \tag{4}$$

where $\prod \gamma_i$ represents the set of parent nodes associated with γ_i. For the root nodes, $P(\gamma_i | \prod \gamma_i)$ becomes the marginal distribution $P(\gamma_i)$. This property shall be used later on for a proper representation of lean complex cyber-physical networks through quantum circuits.

The following paragraphs present several quantum computing fundamentals for the general reader.

- *Qubit*

  Information may be represented in many different ways. Quantum computing uses quantum discrete units of information, the qubit (quantum bit) [51]. Qubits represent elementary units of information exchange in quantum computing, similar to the "bits" of classical computing. A bit is always in two basic states, 0 and 1, while a qubit can be in both bases of these states simultaneously. The characteristic is also known as superposition. Quantum computing normally uses the Dirac notation that represents the two bases of computing of these states $|0\rangle$ and $|1\rangle$. The superposition of a $|\Psi\rangle$ qubit is merely a linear combination of the two basic states $|0\rangle$ and $|1\rangle$, expressed by the Equation (5):

$$|\Psi\rangle = c_1 |0\rangle + c_2 |1\rangle \tag{5}$$

  where c_1 and $c_2 \in \mathbb{C}$ such that the satisfy the Equation (6):

$$|c_1|^2 + |c_2|^2 = 1 \tag{6}$$

  in which $|c_1|^2$ and $|c_2|^2$ are, respectively, the probabilities of finding the *qubit* in $|0\rangle$ and $|1\rangle$ after a measurement in the $(|0\rangle, |1\rangle)$ basis.

- Bloch's sphere

  Bloch's sphere, as shown in Figure 2A, is commonly used to geometrically represent a *qubit* [52]. This is a useful and common geometric image of the quantum evolution of a single- or two-level system. On the Bloch sphere, of unitary radius, the Z-axis is the computational axis and its positive direction coincides with the state $|0\rangle$, and the negative with the state $|1\rangle$. A *qubit* can be represented as a point on the Bloch sphere with the help of two parameters (θ, ϕ), as expressed by Equation (7):

$$|\Psi\rangle = cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\phi} sin\left(\frac{\theta}{2}\right)|1\rangle \tag{7}$$

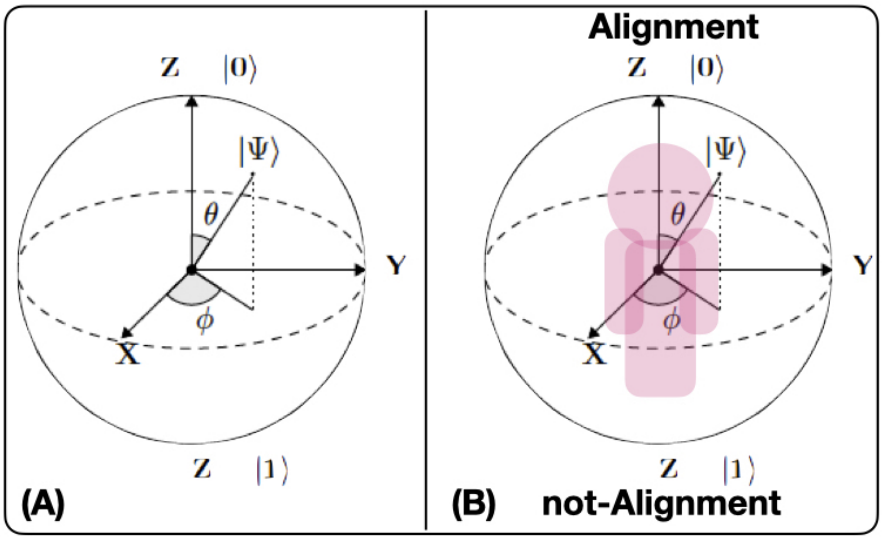


Figure 2. (**A**) Bloch sphere (**B**) Quantum Strategic Organizational Design (QSOD) Bloch sphere.

When several qubits are utilized, their aggregated state can be determined utilizing the tensorial product of the individual qubits. If the multiple qubit state can be expressed as a linear

combination of the $|0\rangle$ and $|1\rangle$ states, then the aggregated state can be represented, as in Equation (8):

$$|\Psi_1\rangle \otimes |\Psi_2\rangle = c_{11}c_{21}\,|00\rangle + c_{11}c_{22}\,|01\rangle + c_{12}c_{21}\,|10\rangle + c_{12}c_{22}\,|11\rangle \tag{8}$$

where $|\Psi_1\rangle = c_{11}\,|0\rangle + c_{12}\,|1\rangle$ and $|\Psi_2\rangle = c_{21}\,|0\rangle + c_{22}\,|1\rangle$. However, if the aggregate state cannot be expressed as the product of the individual states, in other words, if no qubit states $|a\rangle$ and $|b\rangle$ can be found, such that $|\Psi\rangle = |a\rangle\,|b\rangle$, this state is called entangled state, which is stronger than any other classical correlation [53].

The reduced purity κ_j of a qubit q_j in an $N - qubit$ state $|Psi\rangle$, as given by Equation (9), is a coefficient $\kappa_j \in [0.5, 1]$ that indicates the level of a qubit entanglement in the state [54]. A value of $\kappa_j = 1$ indicates that the qubit is not entangled with the other $N - 1$ qubits, and a value of $\kappa_j = 0.5$ indicates that the qubit is maximally entangled with the other qubits in the state.

$$\kappa_j = Tr[Tr_{i\in[0,N-1],i\neq j}\ |\psi\rangle\,\langle\psi|]^2 \tag{9}$$

- Quantum circuit

  A quantum circuit is a computational sequence that consists of performing a series of coherent quantum operations on qubits. By organizing the qubits into an orderly sequence of quantum gates, measurements, and resets, all of which can be conditioned and use data from the classical calculation in real-time, quantum computing can be simulated. These sequences typically follow a standardized pattern:

  1. Initialization and reset. First, we begin our quantum calculation with a specified quantum state for each qubit. This is achieved using the initialization operations, typically on the Z-computation axis, and reset. The resets can be done using a single-qubit gate combination that tracks whether we have succeeded in creating the desired state through measurements. Qubit initialization in a desired state $|\Psi\rangle$ can then continue to apply single-qubit gates.
  2. Quantum gates. Second, we implement a sequence of quantum gates that manipulate the qubits, as needed by the targeted algorithm following certain quantum circuit design principles.
  3. Measurement. Third, we measure the qubits. Classical computers translate the measurements of each qubits as classical results (0 and 1) and then store them in either one of the two classical bits. Measurement is understood to be projected into the Z-computational basis unless otherwise stated.

- Quantum gate

  A quantum gate consists of several mathematical operations applied to the qubits that change the amplitude of their probabilities and, thus, perform the intended computations [54]. The quantum computing basic elements are described in detail:

  – The $U_3(\theta, \phi, \lambda)$ gate is a single qubit gate that has three parameters θ, ϕ and λ which represent a sequence of rotations around the Bloch sphere's axes such that $[\phi - \pi/2]$ around the Z axis, $[\pi/2]$ around the X axis, $[\pi - \theta]$ around the Z axis, $[\pi/2]$ around the X axis, and a $[\lambda - \pi/2]$ around the Z axis. It can be used to obtain any single qubit gate. Equation (10) provides its mathematical representation,

$$U_3\,|\Psi\rangle = \begin{bmatrix} cos(\frac{\theta}{2}) & -e^{i\lambda}sin(\frac{\theta}{2}) \\ e^{i\phi}sin(\frac{\theta}{2}) & e^{i(\phi+\lambda)}cos(\frac{\theta}{2}) \end{bmatrix} |\Psi\rangle \tag{10}$$

and Equation (11) its quantum circuit equivalent:

$$|\Psi\rangle \quad \boxed{U3(\theta, \phi, \lambda)} \tag{11}$$

– The *CNOT* or conditional NOT gate is a two qubit computation gate with one qubit acting as control $|\Psi_1\rangle$ and the other as *target* $|\Psi_2\rangle$. The *CNOT* gate performs a selective negation of the target qubit. If the control qubit is in superposition, then *CNOT* creates entanglement. Equation (12) provides its mathematical representation,

$$CNOT\,|\Psi_1\rangle\,|\Psi_2\rangle = |\Psi_1\rangle\,|\Psi_1 \oplus \Psi 2\rangle \tag{12}$$

and Equation (13) its quantum circuit equivalent:

$$\begin{array}{l} |\Psi_1\rangle \\ |\Psi_2\rangle \end{array} \quad \begin{array}{c} \bullet \\ \oplus \end{array} \tag{13}$$

– The *ccX* or Toffoli gate is a three qubit computation gate with two qubits $|\Psi_1\rangle$ and $|\Psi_2\rangle$ acting as controls and one qubit $|\Psi_3\rangle$ acting as target. The *ccX* gate applies an X to the target qubit $|\Psi_3\rangle$ only when both controls $|\Psi_1\rangle$ and $|\Psi_2\rangle$ qubits are in state $|1\rangle$. Equation (14) provides its mathematical represetnation:

$$ccX\,|\Psi_1\rangle\,|\Psi_2\rangle\,|\Psi_3\rangle = |\Psi_1\rangle\,|\Psi_1 \oplus \Psi 2\rangle\,|\Psi_1\rangle\,|\Psi_1 \oplus \Psi 3\rangle \tag{14}$$

and Equation (15) its quantum circuit equivalent [54]:

$$\begin{array}{l} |\Psi_1\rangle \\ |\Psi_2\rangle \\ |\Psi_3\rangle \end{array} \quad \begin{array}{c} \bullet \\ \bullet \\ \oplus \end{array} \quad = \quad \begin{array}{l} |\Psi_1\rangle \\ |\Psi_2\rangle \\ |\Psi_3\rangle \end{array} \quad \begin{array}{c} \bullet \\ \oplus \\ \\ \end{array} \quad \begin{array}{c} \\ \bullet \\ \oplus \end{array} \tag{15}$$

– The *Z-measurement* of a quantum state—a self-adjoint operator on the Hilbert space—results in the measured object being in an eigenstate of the Z operator or computational basis, with the corresponding eigenvalue being the value measured. The measurement, also called observation, of a quantum state, is a stochastic non-reversible operation and, therefore, cannot be considered as a quantum gate, as it allocates a unique value to the variable observed. In mathematical terms, the probability p of a measurement result m occurring when the state $|\Psi\rangle$ is measured is given by Equation (16):

$$p(m) = \langle \Psi | M_m^\dagger M_m | \Psi \rangle \tag{16}$$

where $[M_m]$ represents a set of operators acting on the state space, such that $I = \sum_m p(m)$ and the state of the system after the measurement $|\Psi'\rangle$ is given by Equation (17):

$$|\Psi'\rangle = \frac{M_m\,|\Psi\rangle}{\sqrt{p(m)}} \tag{17}$$

Equation (18) shows its quantum circuit equivalent as a symbolic box:

$$|\Psi\rangle \quad \boxed{} \tag{18}$$

3. Quantum Strategic Organizational Design

This chapter outlines the main contribution of this work, presenting a model and design principles for enabling organizational leaders to perform quantum strategic organizational design. This model

consists of three steps: first, it introduces the definition of the QSOD qubit as a fractal unit of the decision Industry 4.0 complex-networked cyber-physical lean management systems. Second, it outlines the design principles that are to be applied in order to represent such systems as a quantum circuit. Finally, it provides guidelines for the interpretation of the results.

3.1. QSOD Qubit

To be able to create a quantum circuit that allows for contemplating decision Industry 4.0 complex-networked cyber-physical lean management systems it is necessary, to provide a standardized vision of the elements of this network. As mentioned above, scholars have previously defined a communication standard (CPD)nA that forms the edges of these complex networks [9]. Now, it is time to present a standard for the nodes.

As depicted in Figure 2B, we define the QSOD qubit, the node of a decision complex-networked cyber-physical lean management system, as a human or cyber-physical asset that is in the center of an imaginary Bloch sphere, in which the state of alignment and not-alignment references, respectively, with the QSOD qubit $|0\rangle$ and $|1\rangle$ computational states. This is possible because the QSOD qubit exchanges information with other QSOD qubits through standardized information flow behavioral (CPD)nA patterns that allow for constant monitoring of its performance and this implies, as shown in [2], that the state of individual alignment of all nodes is known at any point in time.

3.2. QSOD Design Principles

- Calculation of conditional probabilities

  The conditional probabilities that will give rise to the quantum circuit are derived from a preliminary analysis of the KPIs that are associated with each node of the complex network in question. This analysis, as indicated above, is based on the method based on genetic algorithms presented in [2]. Specifically, for each node, there are typically three related KPIs. The selected chromosome has subsequently 12 real numbers between 0 and 1, and we have used real value crossover and mutation with probabilities of 60% and 7%, respectively. The population was built over 8000 individuals and it ran over 1000 generations. Once the trajectories associated with each node have been calculated, by applying the Bayes theorem, it is trivial to calculate the relative probability of alignment or non-alignment at each node concerning those to which it is connected.
- Initialization and reset

  The initialization and reset of the qubits is typically standardized to the state $|0\rangle$ on the computational Z-axis.
- Rotation angle computation

  The conditional probabilities translate into qubit rotation angles depending on its decision network dependencies:

  - For a root node with no parents, the possible states are two $|0\rangle$ and $|1\rangle$. A trivial application of Equation (7), states that a qubit intialized to state $|0\rangle$ and rotated by a gate $U3(\theta, 0, 0)$, being $\phi = 0$, transforms it into $|\Psi\rangle = cos\left(\frac{\theta}{2}\right)|0\rangle + sin\left(\frac{\theta}{2}\right)|1\rangle$. Therefore, taking Equation (5) into account and the definition of the Bloch sphere's angles, the rotation angle θ that is required to calculate the probabilities of being in state $|0\rangle$ and $|1\rangle$ can be expressed by Equation (19):

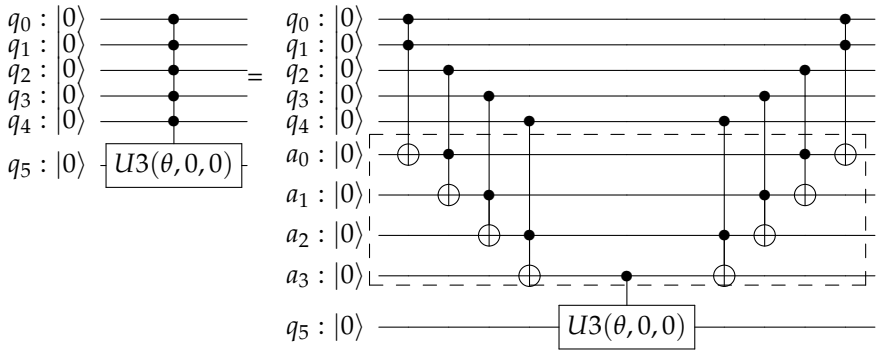
$$\theta = 2atan\left[tan\left(\frac{\theta}{2}\right)\right] = 2atan\sqrt{\frac{sin^2\left(\frac{\theta}{2}\right)}{cos^2\left(\frac{\theta}{2}\right)}} \overset{Equation\ (5)}{=} 2atan\sqrt{\frac{p(|1\rangle)}{p(|0\rangle)}} \tag{19}$$

– In general, for a child node γ_i with m parents, there are 2^m possible states $\prod \gamma_i^*$. Subsequently, taking Equations (4), (5), and (7) into account, as well as the definition of the Bloch sphere's angles, the rotation angle is given by Equation (20):

$$\theta_{\gamma_i, \prod \gamma_i^*} \overset{Equation\ (4)-Equation\ (5)}{=} 2atan\sqrt{\frac{p(|1\rangle \,|\, \prod \gamma_i = \prod \gamma_i^*)}{p(|0\rangle \,|\, \prod \gamma_i = \prod \gamma_i^*)}} \tag{20}$$

- Controlled rotations

  Controlled rotations are not elementary quantum gates and they need to be deconstructed into elementary operations. As described by Nielsen and Chuang [54], being m the maximum number of parent nodes a child has, the controlled rotation expressing the conditional probabilities needs of the addition of a_i "dummy" qubits $i = 1, \ldots, m-1$ in order to decompose the controlled rotation into $2(m-1)$ $CNOT$ gates and one $U3(\theta, 0, 0)$. This is exemplified in Equation (21) for $m = 5$ qubits, a_i "dummy" qubits $i = 1, \ldots, (m-1=4)$ and a total of $2(m-1) = 8$ $CNOT$ gates.



(21)

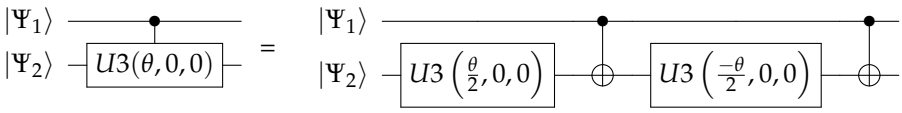
As first shown in [55], it is important to highlight that $U3(\theta, 0, 0)$ is best decomposed as by Equation (22):

$$U3(\theta,0,0)\,|\Psi_1\rangle\,|\Psi_2\rangle = U3\left(\frac{\theta}{2},0,0\right)|\Psi_2\rangle\,CNOT\,|\Psi_1\rangle\,|\Psi_2\rangle\,U3\left(\frac{-\theta}{2},0,0\right)|\Psi_2\rangle\,CNOT\,|\Psi_1\rangle\,|\Psi_2\rangle \tag{22}$$

Additionally, as a direct application of Equation (13), this equation converts into Equation (23):

$$U3(\theta,0,0)\,|\Psi_1\rangle\,|\Psi_2\rangle \overset{Equation\ (13)}{=} U3\left(\frac{\theta}{2},0,0\right)|\Psi_2\rangle\,|\Psi_1\rangle\,|\Psi_1 \oplus \Psi 2\rangle\,U3\left(\frac{-\theta}{2},0,0\right)|\Psi_2\rangle\,|\Psi_1\rangle\,|\Psi_1 \oplus \Psi 2\rangle \tag{23}$$

or in its quantum circuit equivalent that is shown in Equation (24):



(24)

This method works, because, if the control qubit $|\Psi_1\rangle$ is in state $|0\rangle$, all we have is $U3\left(\frac{\theta}{2}, 0, 0\right)$ followed by a $U3\left(\frac{-\theta}{2}, 0, 0\right)$ and the effect is trivial. If the control qubit $|\Psi_1\rangle$ is in state $|1\rangle$, the net effect is a controlled rotation $U3(\theta, 0, 0)$ on the $|\Psi_2\rangle$ qubit.

- Measurement of the QSOD qubits to obtain the probabilities of alignment states.

The measurements are mainly used in the end to extract computational results from the quantum states. This will allow for us to explore the quantum states of the qubits and make an interpretation that allows for improving the management system that is related to the industrial process.

4. Case Study. Qsod Circuit

We will propose a quantum circuit that allows calcullating the alignment states of the system that the state of the art based on genetic algorithms in order to illustrate the implementation of the QSOD method within a cyber-physical complex networked lean management system in an Industry 4.0 context. Following the recommendations of [56], we follow a clear case study roadmap to ensure the replicability and soundness of the results obtained. This roadmap has several phases:

- Section 4.1. Scope establishment.
- Section 4.2. Specification of population and sampling.
- Section 4.3. QSOD circuit design.
- Section 4.4. Analysis of results.

4.1. Scope Establishment

We aim to study the organizational alignment state of an Industry 4.0 factory that resembles a cyber-physical complex networked lean management system by modeling the strategic organizational design throughout a quantum circuit. The management networks that model such a system under study are based on the paradigm of formal standardized (CPD)nA communication among various leaders who follow the lean concept and represent value-creating processes that are driven by KPIs and their improvement of the system, as described in [39].

4.2. Specification of Population and Sampling

Data on four team members in a factory at three relevant hierarchical levels are taken over twelve weeks daily: at level 1, the factory leader, at level 2, a logistics leader, and a production leader both reporting to level 1 and at level 3 a production line leader reporting to the level 2 production leader. Each leaders' performance is measured through three KPIs, so there is a total of 12 KPIs being measured with the same frequency. The case study starts by creating a decision network from these data, as indicated in Section 3.2 and Figure 3. There are four nodes q_A, q_B, q_C, and q_D, which correspond to each process owner, and one "dummy" node q^*, because the maximum number of parents is two. This network resembles the conditional probabilities of alignment and the respective dependencies related to the complex networked system provided by the original Hoshin Kanri Forest [10] complex network. The details of these calculations are omitted for clarity, since they are explained in the reference literature [2,9,10].

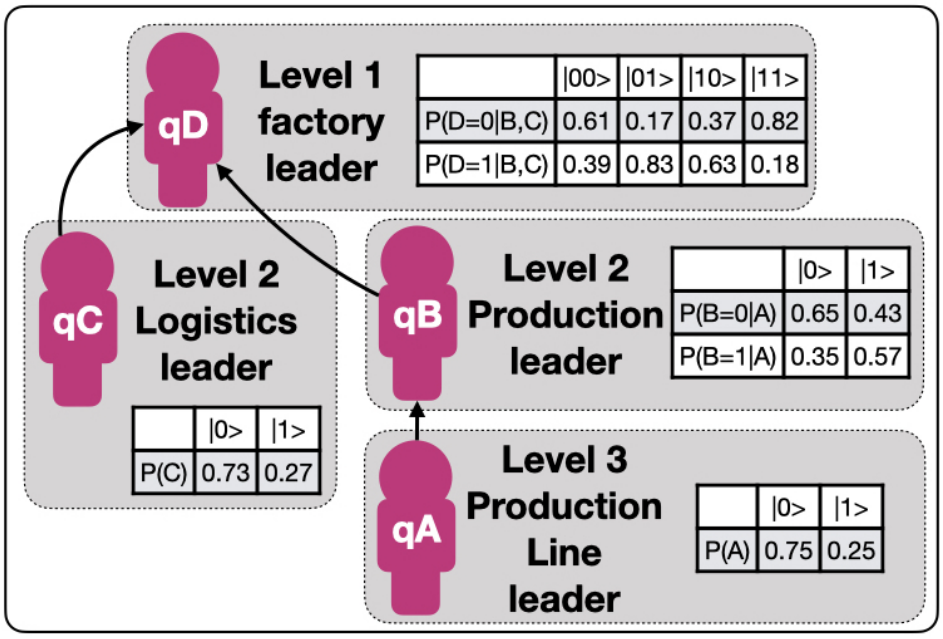


Figure 3. Case study.

4.3. QSOD Circuit Design

We will now proceed to implement each of the indicated steps in Section 3.2 in a systematic way in order to generate a QSOD circuit that represents the alignment probabilities of the whole system. Each one of the 38 steps of the circuit has been denoted with a number in order to allow for a better understanding of the elements of the quantum circuit and facilitate the visualization of the effect of each step on the whole system

- Calculation of conditional probabilities.

  For each node, there are typically three related KPIs. The selected chromosome has subsequently 12 real numbers between 0 and 1, and we have used real value crossover and mutation with probabilities of 60% and 7%, respectively. The population was built over 8000 individuals and it ran over 1000 generations. Once the trajectories associated with each node have been calculated, it is trivial to calculate the relative probability of alignment or non-alignment at each node concerning those to which it is connected. This process is executed for each node in parallel without the loss of performance.
- Initialization and reset.

  In Step 0, each one of the nodes is assigned to a qubit, q_A, q_B, q_C, q_D, and a "dummy" node q^* is created. The qubits are initialized and reset to $|0\rangle$ state. This allows for a controlled comparison of the probabilities through qubit rotations.
- Rotation angle computation.

  Following Equation (19) applied to each root qubit, we obtain the following results given by Equation (25):

$$\theta_A \overset{Equation\ (19)}{=} 2\arctan\sqrt{\tfrac{0.25327658}{0.74672341}} = 1.16479 \quad \theta_C \overset{Equation\ (19)}{=} 2\arctan\sqrt{\tfrac{0.27}{0.73}} = 1.0928 \tag{25}$$

Following Equation (20) applied to each child qubit, we obtain following results that are given by Equation (26):

$$\begin{aligned} \theta_{B,|0\rangle} &\overset{Equation\ (20)}{=} 2\arctan\sqrt{\tfrac{0.35}{0.65}} = 1.2661 & \theta_{B,|1\rangle} &\overset{Equation\ (20)}{=} 2\arctan\sqrt{\tfrac{0.57}{0.43}} = 1.7113 \\ \theta_{D,|00\rangle} &\overset{Equation\ (20)}{=} 2\arctan\sqrt{\tfrac{0.39}{0.61}} = 1.3489 & \theta_{D,|01\rangle} &\overset{Equation\ (20)}{=} 2\arctan\sqrt{\tfrac{0.83}{0.17}} = 2.29161 \\ \theta_{D,|10\rangle} &\overset{Equation\ (20)}{=} 2\arctan\sqrt{\tfrac{0.63}{0.37}} = 1.83382 & \theta_{D,|11\rangle} &\overset{Equation\ (20)}{=} 2\arctan\sqrt{\tfrac{0.18}{0.82}} = 0.87629 \end{aligned} \tag{26}$$

- Controlled rotations.

  By systematically applying Equation (21) to each qubit, we obtain the QSOD circuit that is shown in Figure 4.

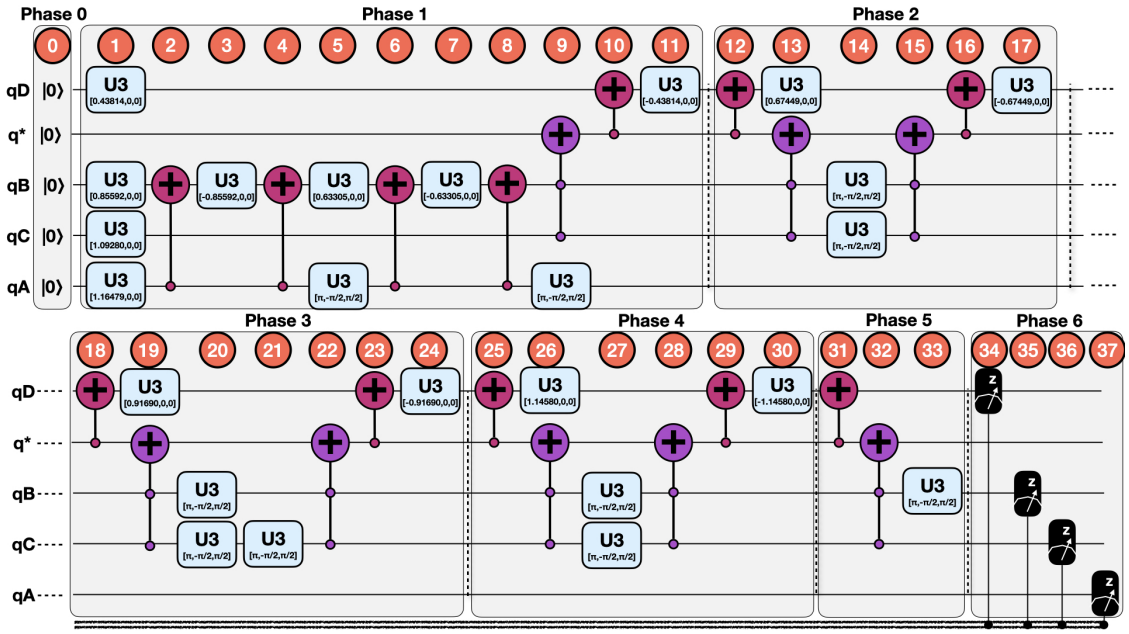


Figure 4. QSOD Circuit.

Figure 5A visualizes the quantum state of each step of the final state. The Bloch sphere provides a global view of a multi-qubit quantum state in the computational basis. Node size is proportional to state probabilities, and color reflects the phase of each basis state as shown in Figure 5B. The constant color blue that reflects the phase of each basis state does not change throughout the circuit. For the interested reader, in Appendix A, the Bloch sphere states of each step in the circuit are displayed in Figures A1–A4.

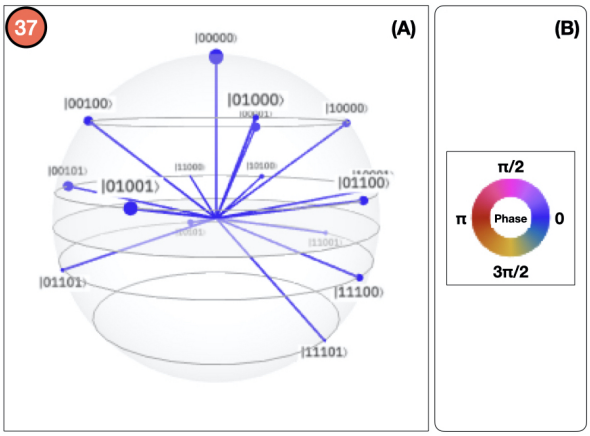


Figure 5. (**A**) QSOD Bloch sphere measurements of final state. (**B**) Phase state color code.

The 37 steps that make up the QSOD have been divided into six distinct phases in order to improve the clarity of the explanation. We will now comment on the logic behind each of them:

– Phase 0. Step 0.

  Phase 0 performs the initialization and reset, as previously explained in Section 4.3.

– Phase 1. Step 1–Step 11.

  This phase is subdivided in three conceptual parts:

  First. This phase starts in Step 1 by rotating q_A and q_C by $\theta_A = 1.16479$ and $\theta_C = 1.0928$ radians respectively as calculated in Equation (25). This is because both q_A and q_C are root qubits and Equation (19) applies.

  Second, two controlled rotations on qubit q_B in Steps 1–3 are performed, as calculated in Equation (26). As described in Equation (24), since qubit q_B has a parent *qubit* q_A, we need to perform two controlled rotations properly aligned by CNOT gates. Depending on the state of qubit q_A: $\frac{\theta_{B,|1\rangle}}{2} = 0.85592$ in Step 1 and $\frac{-\theta_{B,|1\rangle}}{2} = -0.85592$ in Step 3, in the case that q_A is in state $|1\rangle$, and $\frac{\theta_{B,|0\rangle}}{2} = 0.63305$ in Step 5 and $\frac{-\theta_{B,|0\rangle}}{2} = -0.63305$ in the case that q_A is in state $|0\rangle$. In this case, a $U_3(\pi, -\frac{\pi}{2}, \frac{\pi}{2})$ is performed, so as to generate proper alignment.

  Third, on qubit q_D, we need to perform a total of four controlled rotations throughout the circuit. This is because each of its parent qubits q_B and q_C can have two states, and we need to represent the rotations corresponding to the states $|00\rangle$, $|10\rangle$, $|01\rangle$, and $|11\rangle$. A controlled rotation $\theta_{D,|11\rangle}$ is performed on qubit q_D conditioned by the state $|11\rangle$ of its parent qubits q_B and q_C, as calculated in Equation (26). This is done with a controlled rotation $\frac{\theta_{D,|11\rangle}}{2} = 0.43814$ in Step 1 and a controlled rotation $\frac{-\theta_{D,|11\rangle}}{2} = -0.43814$ in Step 11, properly aligned through CNOT and ccX gates, as described in Equation (21) in Steps 9–10 and Steps 31–33.

– Phase 2. Step 12–Step 17.

  In phase 2, we perform the second controlled rotation on qubit q_D as related to the states $|00\rangle$ of qubit q_B and q_C. A controlled rotation $\theta_{D,|00\rangle}$ is performed on qubit q_D conditioned by the state $|00\rangle$ of its parent qubits q_B and q_C, as calculated in Equation (26). This is done with a controlled rotation $\frac{\theta_{D,|00\rangle}}{2} = 0.67449$ in Step 13 and a controlled rotation $\frac{-\theta_{D,|00\rangle}}{2} = -0.67449$ in Step 17, properly aligned through CNOT and ccX gates, as described in Equation (21) in

Steps 12 and Steps 14–16.

– Phase 3. Step 18–Step 24.

In phase 3, we perform the second controlled rotation on qubit q_D as related to the states $|10\rangle$ of qubit q_B and q_C. A controlled rotation $\theta_{D,|10\rangle}$ is performed on qubit q_D conditioned by the state $|10\rangle$ of its parent qubits q_B and q_C, as calculated in Equation (26). This is done with a controlled rotation $\frac{\theta_{D,|10\rangle}}{2} = 0.91690$ in Step 19 and a controlled rotation $\frac{-\theta_{D,|10\rangle}}{2} = -0.91690$ in Step 24, properly aligned through CNOT and ccX gates, as described in Equation (21) in Steps 18 and Steps 20–23.

– Phase 4. Step 25–Step 30.

In phase 4, we perform the second controlled rotation on qubit q_D as related to the states $|01\rangle$ of qubit q_B and q_C. A controlled rotation $\theta_{D,|10\rangle}$ is performed on qubit q_D conditioned by the state $|01\rangle$ of its parent qubits q_B and q_C, as calculated in Equation (26). This is done with a controlled rotation $\frac{\theta_{D,|01\rangle}}{2} = 1.1458$ in Step 26 and a controlled rotation $\frac{-\theta_{D,|01\rangle}}{2} = -1.1458$ in Step 30, properly aligned through CNOT and ccX gates, as described in Equation (21) in Steps 25 and Steps 27–29.

– Phase 5. Step 31–Step 33.

As mentioned earlier, in phase 5 the controlled rotation of qubit q_D as related to the states $|00\rangle$ of qubit q_B and q_C started in Phase 1, qubits 1 and 9–11, is completed.

– Phase 5. Step 34–Step 37.

Finally, in phase 6 each one of the qubits is measured, as expressed by Equation (17).

4.4. Analysis of Results

The circuit is simulated on qiskit tool, a Python-based [57] quantum computing platform developed by IBM [58]. A total number of 8192 runs were carried out on the simulation with a total runtime of 3.8 s. In contrast, a genetic algorithm that would solve a similar problem with 12 KPIs (three per process owner) would take hundreds of hours for determining 48 real numbers between 0 and 1 in the chromosome with a value crossover and mutation with probabilities of 60% and 7%, respectively, with a population built over 8000 individuals and would run over 1000 generations. When compared with that, the performance increase of QSOD is remarkable. This has very powerful practical applications for industry leaders, since with this new approach they can potentially allow a better understanding of the complex processes underlying the strategic design of organizations and above all make decisions in real-time.

The obtained results are summarized in Table 1, which shows the total probability of each process owner $P_j(|0\rangle)$ to be in alignment and the reduced purity κ_j of each *qubit* in the final state. We observe how the probability of alignment of the process owner D $P_D(|0\rangle) = 49\%$, which indicates that the management system, as configured does not give a probability of achieving the alignment better than chance. That is why the probability of alignment of the root node representing process owner C $P_C(|0\rangle) = 73\%$, the same as that presented in the decision network. The alignment probabilities representing process owners B $P_B(|0\rangle) = 58.335\%$ and A $P_A(|0\rangle) = 69.745\%$ are mixed probabilities. The nodes have a purity coefficient of over 90%, which indicates that there is almost no entanglement between them.

Table 1. Summary of the results.

Qubits	Probability of Alignment $P_j(\|0\rangle)$	Reduced Purity κ_j
q_D	49.135%	0.9178
q_B	58.335%	0.9026
q_C	73.000%	0.9178
q_A	69.745%	0.9794

Further detailed results on the measurement probabilities of the computational basis states are visualized in Figure 6 after measurement Step 37. As an example of how to interpret these results, it can be said that the probability of complete system alignment, as described by the state $|00000\rangle$, is 20.18%.

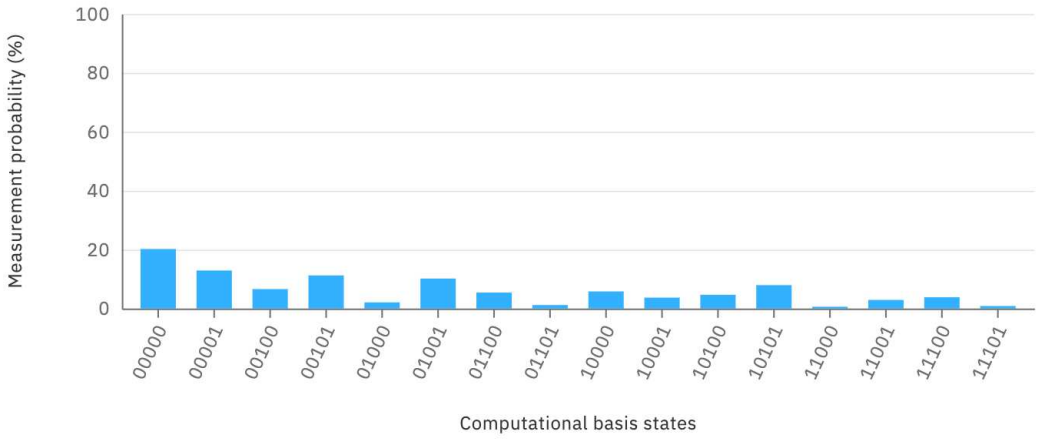


Figure 6. QSOD final Measurement Probabilities in step 37.

5. Conclusions, Further Research and Limitations

This study shows how the strategic design of an organization is quantified through adequate modeling of decision networks expressed as quantum circuits. Because of the speed that quantum computing offers, the obtained results are promising, because they are likely to allow in the future a real-time simulation of organizational designs that could not have been done until now. This will allow for the leaders of the Industry 4.0, as well as interested scholars, to perform simulations of possible organizational designs for a customized adaptation of the strategic configuration. More specifically, an industrial leader will be able to potentially implement the algorithms that are presented to perform in-vitro analysis of certain clusters of interest within their organization, in order to make decisions on how the optimal configuration of their strategic industrial design should be. All of this in real-time and at virtually no cost. Presently, this tool can be used for a maximum of 15 *qubits* in the IBM simulation environments mentioned.

From systems engineering point of view, the proposed QSOD approach has great potential for supporting complex system development and management. Modern industrial enterprises consist of multiple systems, subsystems and even system of systems that usually involve different stakeholders with heterogeneous requirements. The alignment of these requirements is critical for decision-makings. A more specific example is the design and integration of modern manufacturing systems that might contain many digital twin models across the entire life-cycle of a product, such as product design, simulation, manufacturing, and maintenance, etc. The alignment of these systems is very challenging, even if not impossible, with traditional approaches due to limited computing resources and time. The proposed quantum strategic approach provides a promising solution for this challenge.

In addition to the applications in the management field, the proposed quantum-based approach could also be inspiring to technological domains such as the Distributed Ledger Technology (DLT), which has been widely applied in recent years. For example, DLT-based platforms have been developed in order to facilitate industrial data sharing, supply chain management and process monitoring etc. One of the main concerns about the most popular distributed ledger architecture blockchain is the vulnerability against quantum computing attacks. The proposed approach makes possible creating an assessment mechanism that is based on quantum computing principles in order to evaluate the security and robustness of distributed ledger applications, especially in the industrial domain.

For this work, we have used a classic computational tool that ideally simulates quantum circuits. This brings with it a series of potential restrictions to the study: on the one hand, it is possible to foresee a certain distortion when performing the same simulations in real quantum circuits. On the other hand, the number of qubits that the simulation can support is limited, at this moment of the state of the art, and therefore the simulation of large organizational networks still presents a challenge for future studies.

Author Contributions: Conceptualization, J.V.-D. and X.Z.; Data curation, J.V.-D. and X.Z.; Formal analysis, J.V.-D.; Funding acquisition, J.V.-D.; Investigation, J.V.-D.; Methodology, J.V.-D.; Project administration, J.V.-D.; Resources, J.V.-D.; Software, J.V.-D.; Supervision, J.V.-D.; Validation, J.V.-D.; Visualization, J.V.-D.; Writing—original draft, J.V.-D.; Writing—review & editing, J.V.-D. All authors have read and agreed to the published version of the manuscript.

Funding: J.V.D. would also like to acknowledge the Spanish Agencia Estatal de Investigacion, through research project code RTI2018-094614-B-I00 into the "Programa Estatal de I+D+i Orientada a los Retos de la Sociedad".

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

QSOD	Quantum Strategic Organizational Design
KPI	Key Performance Indicator
DLT	Distributed Ledger Technology

Appendix A

This appendix shows, for the interested reader, the Bloch sphere states of each step in the circuit are displayed in Figures A1–A4. The states between steps 32 and 37 remain constant and are already represented in Figure 5A.

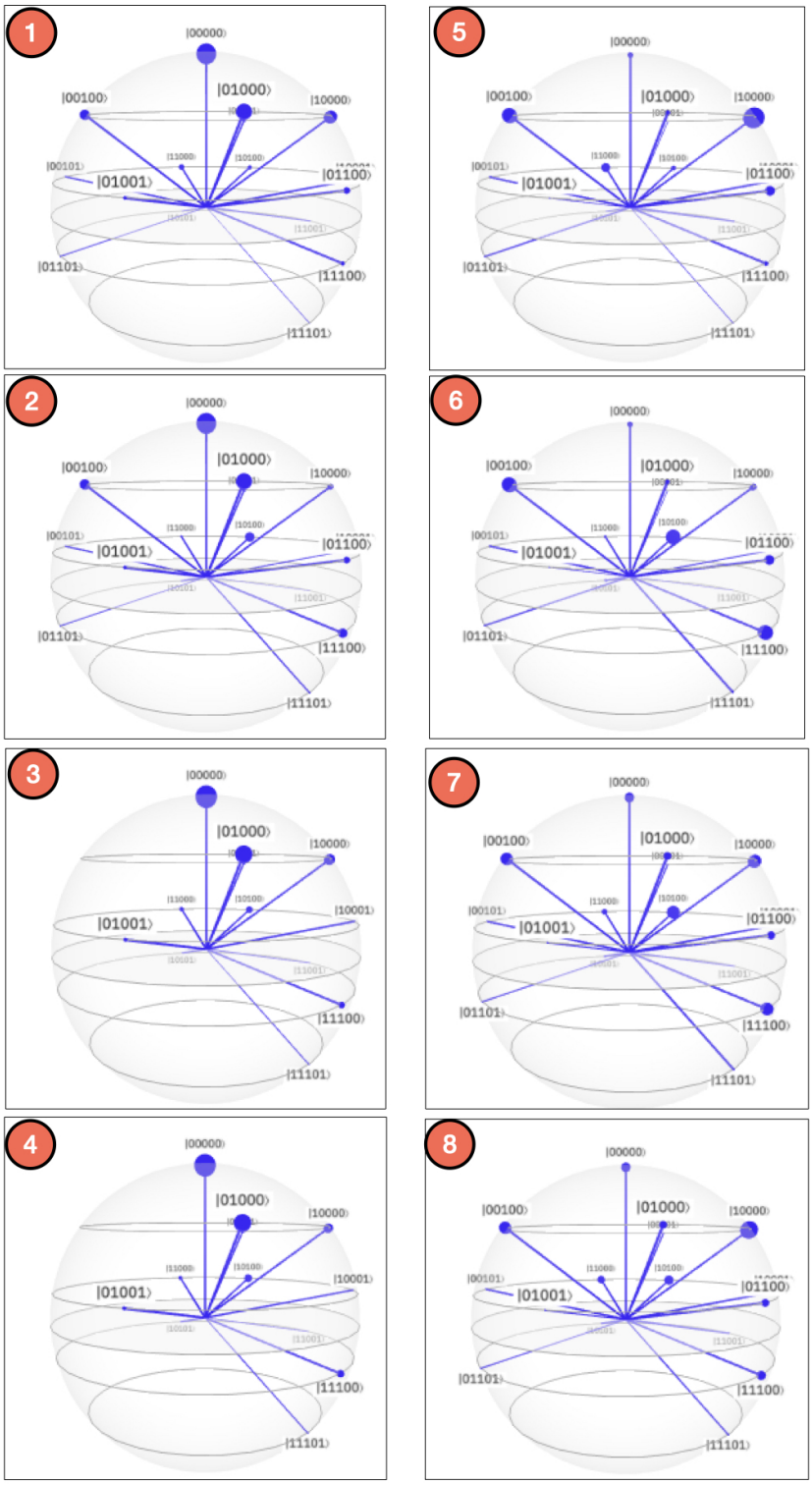


Figure A1. QSOD Bloch sphere measurements steps 1–8.

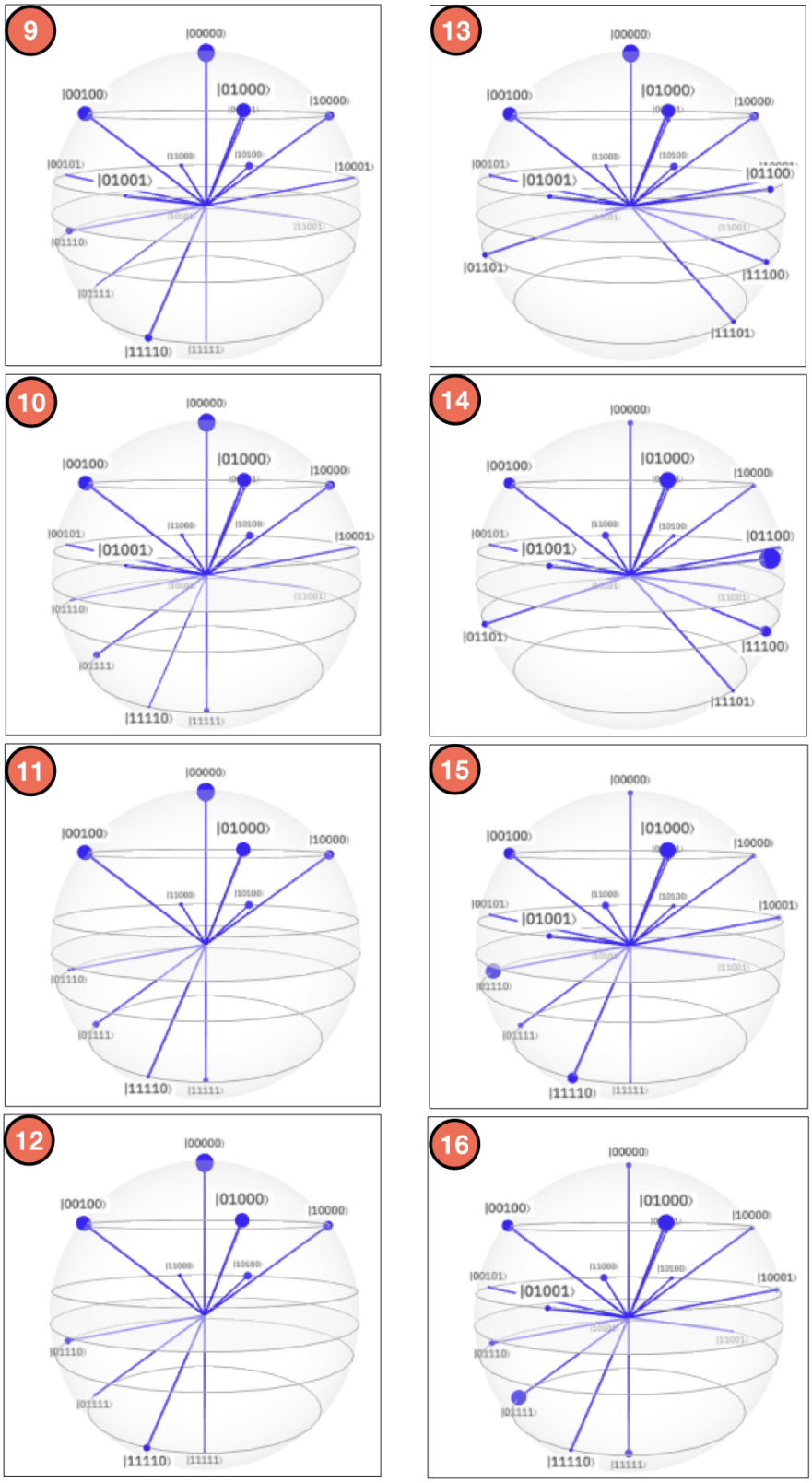


Figure A2. QSOD Bloch sphere measurements steps 9–16.

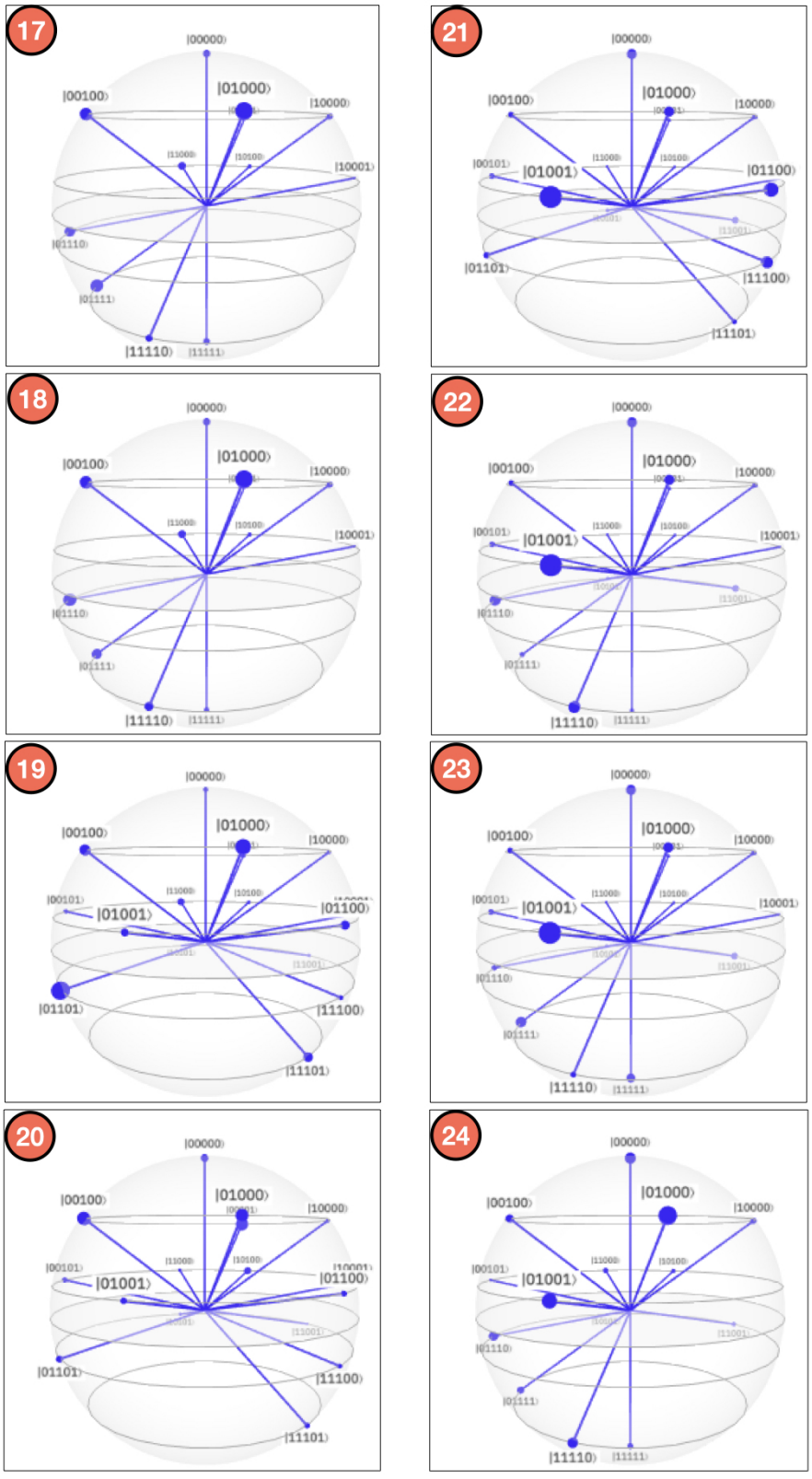


Figure A3. QSOD Bloch sphere measurements steps 17–24.

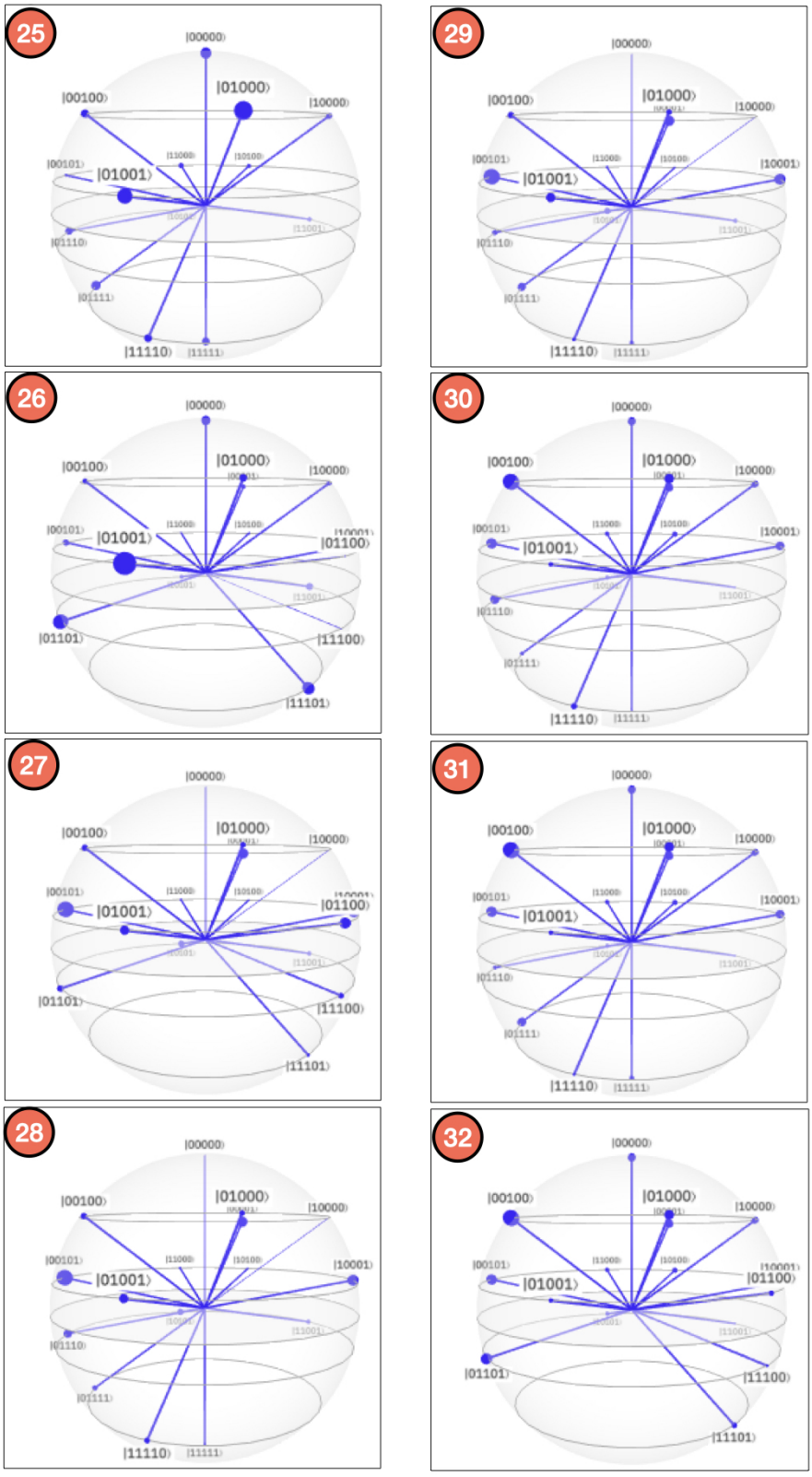


Figure A4. QSOD Bloch sphere measurements steps 25–32.

References

1. Grant, R.M. Organization Structure and Management Systems: The Fundamentals of Strategy Implementation. In *Contemporary Strategy Analysis*, 7th ed.; John Wiley & Sons: West Sussex, UK, 2010; pp. 174–206.
2. Villalba-Diez, J.; Ordieres-Mere, J.; Chudzick, H.; Lopez-Rojo, P. *NEMAWASHI: Attaining Value Stream alignment within Complex Organizational Networks*; Procedia CIRP; Elsevier: Cranfield, UK, 2015; Volume 37, pp. 134–139. [CrossRef]
3. Tennant, C.; Roberts, P. Hoshin Kanri: Implementing the Catchball Process. *Long Range Plan.* **2001**, *34*, 287–308. [CrossRef]
4. Cattani, G.; Ferriani, S.; Negro, G.; Perretti, F. The Structure of Consensus: Network Ties, Legitimation, and Exit Rates of U.S. Feature Film Producer Organizations. *Adm. Sci. Q.* **2008**, *53*, 145–182. [CrossRef]
5. Cross, R.L.; Singer, J.; Colella, S.; Thomas, R.J.; Silverstone, Y. (Eds.) *The Organizational Network Fieldbook: Best Practices, Techniques and Exercises to Drive Organizational Innovation and Performance*, 1st ed.; Jossey-Bass: San Francisco, CA, USA, 2010.
6. Burton, R.M.; Øbel, B.; Håkonsson, D.D. *Organizational Design: A Step-by-Step Approach*, 3rd ed.; Cambridge University Press: Cambridge, UK, 2015.
7. Barabási, A.-L. *Network Science*; Cambridge University Press: Cambridge, UK, 2016.
8. Villalba-Diez, J.; Molina, M.; Ordieres-Mere, J.; Sun, S.; Schmidt, D.; Wellbrock, W. Geometric Deep Lean Learning: Deep Learning in Industry 4.0 Cyber–Physical Complex Networks. *Sensors* **2020**, *20*, 763. [CrossRef]
9. Villalba-Diez, J.; Ordieres-Mere, J. Improving manufacturing operational performance by standardizing process management. *Trans. Eng. Manag.* **2015**, *62*, 351–360. [CrossRef]
10. Villalba-Diez, J. *The HOSHIN KANRI FOREST. Lean Strategic Organizational Design*, 1st ed.; CRC Press; Taylor and Francis Group LLC: Boca Raton, FL, USA, 2017.
11. Villalba-Diez, J. *The Lean Brain Theory. Complex Networked Lean Strategic Organizational Design*; CRC Press; Taylor and Francis Group LLC: Boca Raton, FL, USA, 2017.
12. Powell, T.C. Organizational alignment as competitive advantage. *Strateg. Manag. J.* **1992**, *13*, 119–134. [CrossRef]
13. Sender, S.W. Systematic agreement: A theory of organizational alignment. *Hum. Resour. Dev. Q.* **1997**, *8*, 23–40. [CrossRef]
14. Ravi, K.; Joshi, P.M.; Porth, J.S. Organizational alignment and performance: Past, present and future. *Manag. Decis.* **2007**, *45*, 503–517.
15. Bryan, A.; Arnold, M.; Lisa, E.; Barratt, A.M. Organizational alignment and supply chain governance structure: Introduction and construct validation. *Int. J. Logist. Manag.* **2009**, *20*, 169–186.
16. Quirós, I. Organizational alignment: A model to explain the relationships between organizational relevant variables. *Int. J. Organ. Anal.* **2009**, *17*, 285–305. [CrossRef]
17. Münch, J.; Fagerholm, F.; Kettunen, P.; Pagels, M.; Partanen, J. The Effects of GQM+Strategies on Organizational Alignment. *arXiv* **2013**, arXiv:1311.6221.
18. Sousa, H.P.; do Prado Leite, J.C.S. Modeling Organizational Alignment. In *Conceptual Modeling*; Yu, E., Dobbie, G., Jarke, M., Purao, S., Eds.; Springer International Publishing: Cham, Switzerland, 2014; pp. 407–414.
19. Carrillo, F.J.; Edvardsson, B.; Reynoso, J.; Maravillo, E. Alignment of resources, actors and contexts for value creation: Bringing knowledge management into service-dominant logic. *Int. J. Qual. Serv. Sci.* **2019**, *11*, 424–438. [CrossRef]
20. Yung-Chang, H.; Ming-Ho, W. How organizational structure and strategic alignment influence new product success. *Manag. Decis.* **2020**, *58*, 182–200. [CrossRef]
21. Piattini, M.; Peterssen, G.; Pérez-Castillo, R. Quantum Computing: A New Software Engineering Golden Age. *SIGSOFT Softw. Eng. Notes* **2020**, *45*, 12–14. [CrossRef]
22. Gyongyosi, L.; Imre, S. A Survey on quantum computing technology. *Comput. Sci. Rev.* **2019**, *31*, 51–71. [CrossRef]

23. Shor, P.W. Algorithms for quantum computation: Discrete logarithms and factoring. In Proceedings of the 35th Annual Symposium on Foundations of Computer Science, Santa Fe, NM, USA, 20–22 November 1994; pp. 124–134. [CrossRef]
24. Grover, L.K. A fast quantum mechanical algorithm for database search. In Proceedings of the STOC '96, Philadelphia, PA, USA, 22–24 May 1996.
25. Nielsen, T.D.; Jensen, F.V. *Bayesian Networks and Decision Graphs*, 2nd ed.; Springer: New York, NY, USA, 2009.
26. Hermann, M.; Pentek, T.; Otto, B. Design Principles for Industrie 4.0 Scenarios. In Proceedings of the 2016 49th Hawaii International Conference on System Sciences (HICSS), Koloa, HI, USA, 5–8 January 2016; pp. 3928–3937. [CrossRef]
27. Huh, J.H.; Seo, K. An Indoor Location-Based Control System Using Bluetooth Beacons for IoT Systems. *Sensors* **2017**, *17*, 2917. [CrossRef] [PubMed]
28. Lee, H.G.; Huh, J.H. A Cost-Effective Redundant Digital Excitation Control System and Test Bed Experiment for Safe Power Supply for Process Industry 4.0. *Processes* **2018**, *6*, 85. [CrossRef]
29. Park, S.; Huh, J.H. Effect of Cooperation on Manufacturing IT Project Development and Test Bed for Successful Industry 4.0 Project: Safety Management for Security. *Processes* **2018**, *6*, 88. [CrossRef]
30. Powell, D.; Romero, D.; Gaiardelli, P.; Cimini, C.; Cavalieri, S. Towards Digital Lean Cyber-Physical Production Systems: Industry 4.0 Technologies as Enablers of Leaner Production. In *Advances in Production Management Systems. Smart Manufacturing for Industry 4.0*; Moon, I., Lee, G.M., Park, J., Kiritsis, D., von Cieminski, G., Eds.; Springer International Publishing: Cham, Switzerland, 2018; pp. 353–362.
31. Sun, S.; Zheng, X.; Villalba-Diez, J.; Ordieres-Mere, J. Indoor Air-Quality Data-Monitoring System: Long-Term Monitoring Benefits. *Sensors* **2019**, *19*, 4157. [CrossRef]
32. Ordieres-Mere, J.; Villalba-Diez, J.; Zheng, X. Challenges and Opportunities for Publishing IIoT Data in Manufacturing as a Service Business. *Procedia Manuf.* **2019**, *39*, 185–193. [CrossRef]
33. Mosterman, P.J.; Zander, J. Industry 4.0 as a Cyber-Physical System study. *Softw. Syst. Model.* **2016**, *15*, 17–29. [CrossRef]
34. Jiang, P.; Ding, K.; Leng, J. Towards a cyber-physical-social-connected and service-oriented manufacturing paradigm: Social Manufacturing. *Manuf. Lett.* **2016**, *7*, 15–21. [CrossRef]
35. Sun, S.; Zheng, X.; Villalba-Díez, J.; Ordieres-Meré, J. Data Handling in Industry 4.0: Interoperability Based on Distributed Ledger Technology. *Sensors* **2020**, *20*, 3046. [CrossRef] [PubMed]
36. Shah, R.; Ward, P. Lean Manufacturing: Context, practice bundles and performance. *J. Oper. Manag.* **2003**, *21*, 129–149. [CrossRef]
37. Villalba-Diez, J.; Ordieres-Meré, J.; Nuber, G. The HOSHIN KANRI TREE. Cross-Plant Lean Shopfloor Management. *Procedia CIRP* **2015**, *32*, 150–155. [CrossRef]
38. Ma, J.; Wang, Q.; Zhao, Z. SLAE–CPS: Smart Lean Automation Engine Enabled by Cyber-Physical Systems Technologies. *Sensors* **2017**, *17*, 1500. [CrossRef] [PubMed]
39. Villalba-Diez, J.; Ordieres-Mere, J. Strategic Lean Organizational Design: Towards Lean World-Small World Configurations through Discrete Dynamic Organizational Motifs. *Math. Probl. Eng.* **2016**, *2016*, 1–10. [CrossRef]
40. Villalba-Diez, J.; Ordieres-Mere, J.; Rubio-Valdehita, S. Lean Learning Patterns. (CPD)nA vs. KATA. *Procedia CIRP* **2016**, *54*, 147–151. [CrossRef]
41. Jimenez, P.; Villalba-Diez, J.; Ordieres-Meré, J. HOSHIN KANRI Visualization with Neo4j. Empowering Leaders to Operationalize Lean Structural Networks. In *PROCEDIA CIRP*; Elsevier: Athens, Greece, 2016; Volume 55, pp. 284–289. [CrossRef]
42. Villalba-Diez, J.; DeSanctis, I.; Ordieres-Meré, J.; Ciarapica, F. Lean Structural Network Resilience. In *Complex Networks & Its Applications VI: Proceedings of Complex Networks 2017 (The Sixth International Conference on Complex Networks and Their Applications)*; Cherifi, C., Cherifi, H., Karsai, M., Musolesi, M., Eds.; Studies in Computational Intelligence; Springer International Publishing: Lyon, France, 2017; pp. 609–619.
43. Romero, D.; Gaiardelli, P.; Powell, D.; Wuest, T.; Thürer, M. Digital Lean Cyber-Physical Production Systems: The Emergence of Digital Lean Manufacturing and the Significance of Digital Waste. In *Advances in Production Management Systems. Production Management for Data-Driven, Intelligent, Collaborative, and Sustainable Manufacturing*; Moon, I., Lee, G.M., Park, J., Kiritsis, D., von Cieminski, G., Eds.; Springer International Publishing: Cham, Switzerland, 2018; pp. 11–20.

44. Villalba-Diez, J.; Schmidt, D.; Gevers, R.; Ordieres-Meré, J.; Buchwitz, M.; Wellbrock, W. Deep learning for industrial computer vision quality control in the printing industry 4.0. *Sensors* **2019**, *19*, 3987. [CrossRef]
45. Villalba-Diez, J.; Zheng, X.; Schmidt, D.; Molina, M. Characterization of Industry 4.0 Lean Management Problem-Solving Behavioral Patterns Using EEG Sensors and Deep Learning. *Sensors* **2019**, *19*, 2841. [CrossRef] [PubMed]
46. Schmidt, D.; Villalba Diez, J.; Ordieres-Meré, J.; Gevers, R.; Schwiep, J.; Molina, M. Industry 4.0 Lean Shopfloor Management Characterization Using EEG Sensors and Deep Learning. *Sensors* **2020**, *20*, 2860. [CrossRef]
47. Villalba-Diez, J.; Ordieres-Mere, J.; Molina, M.; Rossner, M.; Lay, M. Lean dendrochronology: Complexity reduction by representation of kpi dynamics looking at strategic organizational design. *Manag. Prod. Eng. Rev.* **2018**, *9*, 3–9. [CrossRef]
48. Jabeur, N.; Sahli, N.; Zeadally, S. Enabling Cyber Physical Systems with Wireless Sensor Networking Technologies, Multiagent System Paradigm, and Natural Ecosystems. *Mob. Inf. Syst.* **2015**, *2015*, 15. [CrossRef]
49. Linnea, C. Stable Coexistence of Three Species in Competition. Ph.D. Thesis, LinköPing Universitet, Linköping, Sweden, 2009.
50. Bossomaier, T.; Barnett, L.; Harré, M.; Lizier, J.T. Statistical Preliminaries. In *An Introduction to Transfer Entropy: Information Flow in Complex Systems*; Springer International Publishing: Cham, Switzerland, 2016; pp. 11–32.
51. Jaeger, G. (Ed.) *Quantum Information: An Overview*; Springer: New York, NY, USA, 2007; pp. 81–89.
52. Mosseri, R.; Dandoloff, R. Geometry of entangled states, Bloch spheres and Hopf fibrations. *J. Phys. A Math. Gen.* **2001**, *34*, 10243–10252. [CrossRef]
53. Doherty, A.C.; Parrilo, P.A.; Spedalieri, F.M. Distinguishing separable and entangled states. *Phys. Rev. Lett.* **2002**, *88*, 187904. [CrossRef]
54. Nielsen, M.A.; Chuang, I. *Quantum Computation and Quantum Information*; Cambridge University Press: Cambridge, UK, 2010.
55. Barenco, A.; Bennett, C.H.; Cleve, R.; DiVincenzo, D.P.; Margolus, N.; Shor, P.; Sleator, T.; Smolin, J.A.; Weinfurter, H. Elementary gates for quantum computation. *Phys. Rev. A* **1995**, *52*, 3457–3467. [CrossRef]
56. Eisenhardt, K. Building theories from case study research. *Acad. Manag. Rev.* **1989**, *14*, 532–550. [CrossRef]
57. van Rossum, G. *Python Tutorial, Technical Report CS-R9526*; Centrum voor Wiskunde en Informatica (CWI): Amsterdam, The Netherlands, 1995.
58. Wille, R.; Meter, R.V.; Naveh, Y. IBM's Qiskit Tool Chain: Working with and Developing for Real Quantum Computers. In Proceedings of the 2019 Design, Automation Test in Europe Conference Exhibition (DATE), Grenoble, France, 9–13 May 2019; pp. 1234–1240. [CrossRef]

sensors

MDPI

Letter

Industry 4.0 Quantum Strategic Organizational Design Configurations. The Case of Two Qubits: One Reports to One

Javier Villalba-Diez [1,2,*], Rosa María Benito [2] and Juan Carlos Losada [2]

[1] Hochschule Heilbronn, Fakultät Management und Vertrieb, Campus Schwäbisch Hall, 74523 Schwäbisch Hall, Germany
[2] Complex Systems Group, Escuela Técnica Superior de Ingenieros Agrónomos, Universidad Politécnica de Madrid, Av. Puerta de Hierro 2, 28040 Madrid, Spain; rosamaria.benito@upm.es (R.M.B.); juancarlos.losada@upm.es (J.C.L.)
* Correspondence: javier.villalba-diez@hs-heilbronn.de

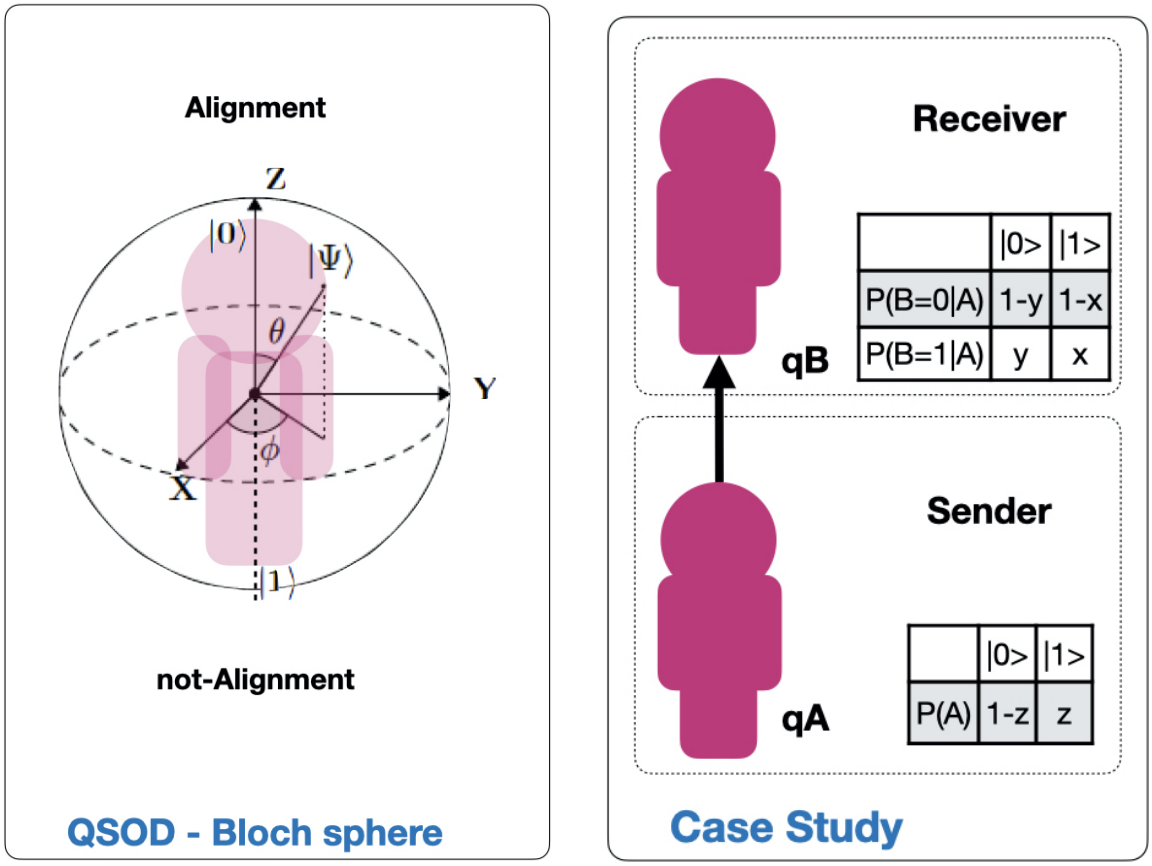
Received: November 2020; Accepted: 4 December 2020; Published: 6 December 2020

Abstract: In this paper we investigate how the relationship with a subordinate who reports to him influences the alignment of an Industry 4.0 leader. We do this through the implementation of quantum circuits that represent decision networks. In fact, through the quantum simulation of strategic organizational design configurations (QSOD) through five hundred simulations of quantum circuits, we conclude that there is an influence of the subordinate on the leader that resembles that of a harmonic under-damped oscillator around the value of 50% probability of alignment for the leader. Likewise, we have observed a fractal behavior in this type of relationship, which seems to conjecture that there is an exchange of energy between the two agents that oscillates with greater or lesser amplitude depending on certain parameters of interdependence. Fractality in this QSOD context allows for a quantification of these complex dynamics and its pervasive effect offers robustness and resilience to the two-*qubit* interaction.

Keywords: quantum strategic organizational design; industry 4.0; quantum circuits

1. Introduction

This work is designed as a brief disclosure of a significant new application of previous work on Quantum Strategic Organizational Design (QSOD) [1], which the interested reader should refer to as a framework. Within this framework, QSOD makes the real-time simulation of organizational alignment states of complex systems in Industry 4.0 possible. QSOD's simulation as decision networks and their equivalent quantum circuits undoubtedly opens a wide field of possibilities for the study of the design of complex networked strategic organizations. As represented schematically in Figure 1a, in the mentioned work we represented the individual process owner, the node of a complex network of Industry 4.0 represented in the form of a decision graph [2], as a quantum computing unit or *qubit* [3,4]. This *qubit* is allowed to have two fundamental states, one of alignment or asymptotic stability of the key performance indicators (KPIs) defining its performance [5–13], represented by the state $|0\rangle$ and another of non-alignment, absence of such stability, represented by the state $|1\rangle$. In this paper we intend to expand this study, taking the simplest case of two organizational agents, a subordinate A reporting to another agent B represented in Figure 1b, that will be simulated by means of a two-*qubit* quantum circuit.



(**a**) (**b**)

Figure 1. (**a**) Case study framework, (**b**) quantum simulation of strategic organizational design (QSOD)—Bloch sphere.

Bloch's sphere, shown in Figure 1a, is commonly used to geometrically represent a *qubit* [14]. This is a useful and common geometric image of the quantum evolution of a single- or two-level system. On the Bloch sphere, of unitary radius, the Z-axis is the computational axis and its positive direction coincides with the state $|0\rangle$, and the negative with the state $|1\rangle$. A *qubit* can be represented as a point on the Bloch sphere with the help of two parameters (θ, ϕ), as expressed by Equation (1):

$$|\Psi\rangle = cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\phi} sin\left(\frac{\theta}{2}\right)|1\rangle \quad (1)$$

Our goal is to determine the probability of alignment of agent B, $P(B = |0\rangle)$, as a function of the alignment probability of agent A, given by $P(A = |0\rangle) = z$, and the conditional alignment probabilities between agents A and B, given by $x = P(B = |1\rangle\,|A = |0\rangle) \in [0,1]$ and $y = P(B = |1\rangle\,|A = |1\rangle) \in [0,1]$. This is achieved through the simulation of more than 500 different quantum circuit configurations in which the relative alignment probabilities $x, y, z \in [0,1]$ vary and $P(B = |0\rangle) = f(x,y,z)$ is measured. In this work we present relevant conclusions about the alignment probabilities of the higher hierarchy agent depending on the alignment state of the leadership relationship.

The rest of the work hereinafter continues as follows: first, Section 2 begins with a description of the configuration of the quantum circuit computations necessary to simulate the outlined two-*qubit* organizational design configuration. Second, Section 3 presents the case study that will simulate numerous quantum circuits, varying the mentioned parameters in order to obtain an optimal configuration of them. Third, in Section 4 we discuss the results obtained and propose an interpretation in the perspective of previous studies and of the working hypotheses. Finally, in Section 5 we discuss

the findings and their implications in a broad context, and future research directions and limitations are highlighted.

2. QSOD Circuit—Two-Qubit Organizational Design Configuration

In this case, two *qubit*s are utilized, and therefore their aggregated state can be determined utilizing the tensorial product of the individual *qubit*s. The multiple *qubit* state can be expressed as a linear combination of the $|0\rangle$ and $|1\rangle$ states, then the aggregated state can be represented as in Equation (2):

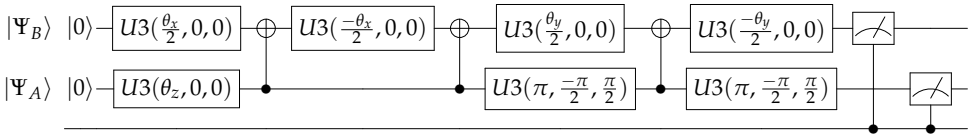
$$|\Psi_1\rangle \otimes |\Psi_2\rangle = c_{11}c_{21}\,|00\rangle + c_{11}c_{22}\,|01\rangle + c_{12}c_{21}\,|10\rangle + c_{12}c_{22}\,|11\rangle \tag{2}$$

where $c_{ij} \in \mathbb{C}^2$, $|\Psi_1\rangle = c_{11}\,|0\rangle + c_{12}\,|1\rangle$ and $|\Psi_2\rangle = c_{21}\,|0\rangle + c_{22}\,|1\rangle$.

Our initial hypothesis is that there is much intrinsic value for any organizational leader in Industry 4.0, to know their alignment status with the company's strategic objectives. Moreover, not only it is important for them to know, but the company has a great interest in having its leaders aligned with its strategic objectives, since this is expected to increase its overall organizational efficiency and effectiveness. We focus on finding answers to the question of how to maximize the probability of alignment of agent B, $P(B = |0\rangle)$, depending on agent A's individual no-alignment probability, $z = P(A = |1\rangle) \in [0,1]$, and the relative probability of alignment between the two agents, $x = P(B = |1\rangle\,|A = |0\rangle) \in [0,1]$ and $y = P(B = |1\rangle\,|A = |1\rangle) \in [0,1]$. Mathematically speaking, we intend to find the values of (x, y, z) that maximize the function $P(B = |0\rangle) = f(x, y, z)$. In other words, our challenge reduces to finding the values of $x, y, z \in [0,1]$ that maximize Equation (3):

$$P(B = |0\rangle) = (c_{11}c_{21})^2 + (c_{11}c_{22})^2 \tag{3}$$

Based on the principles of quantum circuit design that model decision networks presented in [1], the quantum circuit that models the interactions presented by Figure 1b translates into the quantum circuit Equation (4).



(4)

This circuit presents two *qubits* $|\Psi_A\rangle$, with rotation angle θ_z and initial state $|0\rangle$, and $|\Psi_B\rangle$ with rotation angles θ_x, θ_y and initial state $|0\rangle$. The respective interpretation of these rotations and the equations to calculate them are described in Table 1.

Table 1. Qubit angles of rotation.

Qubit	Interpretation	Equation
$\|\Psi_A\rangle$	The conditional probability $z = P(A = \|1\rangle)$ of *qubit* $\|\Psi_A\rangle$ to be in no–alignment translates into the rotation angle θ_z.	$\theta_z = 2\arctan\sqrt{\frac{z}{1-z}}$
$\|\Psi_B\rangle$	The conditional probability $x = P(B = \|1\rangle\,\|A = \|0\rangle)$ of *qubit* $\|\Psi_B\rangle$ to be in no–alignment depending on *qubit* $\|\Psi_A\rangle$ translates into rotation angle θ_x.	$\theta_x = 2\arctan\sqrt{\frac{x}{1-x}}$
	The conditional probability $y = P(B = \|1\rangle\,\|A = \|1\rangle)$ of *qubit* $\|\Psi_B\rangle$ to be in no–alignment depending on *qubit* $\|\Psi_A\rangle$ translates into rotation angle θ_y.	$\theta_y = 2\arctan\sqrt{\frac{y}{1-y}}$

3. Case Study

In the case study we proceed, as announced, to simulate a total of 500 configurations of the circuit shown in Equation (4). We intend to find the values of parameters $x, y, z \in [0,1]$ that maximize the probability of alignment of the agent B, $P(B = |0\rangle)$. To do this, the parameters $x, y \in [0,1]$ are varied in 10% incremental intervals in order to make a uniform mapping and create a proper

display of the results. However, not all $z \in [0,1]$ values are relevant. We are interested in values of $z \geq 0.5$, since they indicate that the alignment probability of the agent A, $P(A = |1\rangle)$, is greater than or equal to 50%. In other words it is equal to or better than a random process. We map the values $z \subset \{0.5, , 0.75, 0.9, 0.99, 0.9999\}$, thus generating 500 simulations, each with a run of 3.5 s, giving a total computation time of 1750 s. The circuits were simulated on *qiskit* tool, a Python-based [15] quantum computing platform developed by IBM [16], and the code and results can be accessed in this Open Access Repository. We summarize the obtained results in Figure 2 by representing $P(B = |0\rangle) = f(x, y, z)$ as a function of $x, y, z \in [0,1]$.

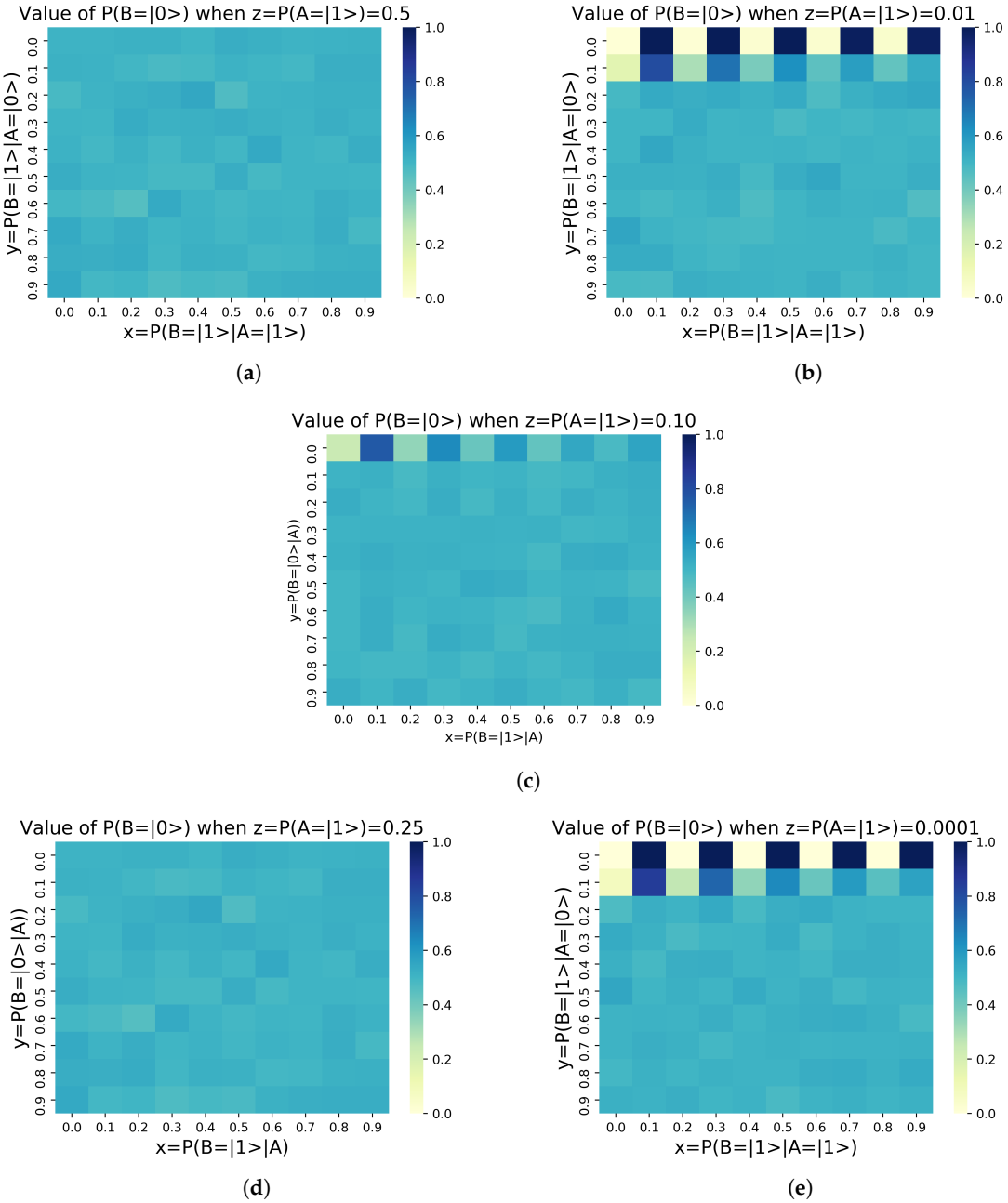


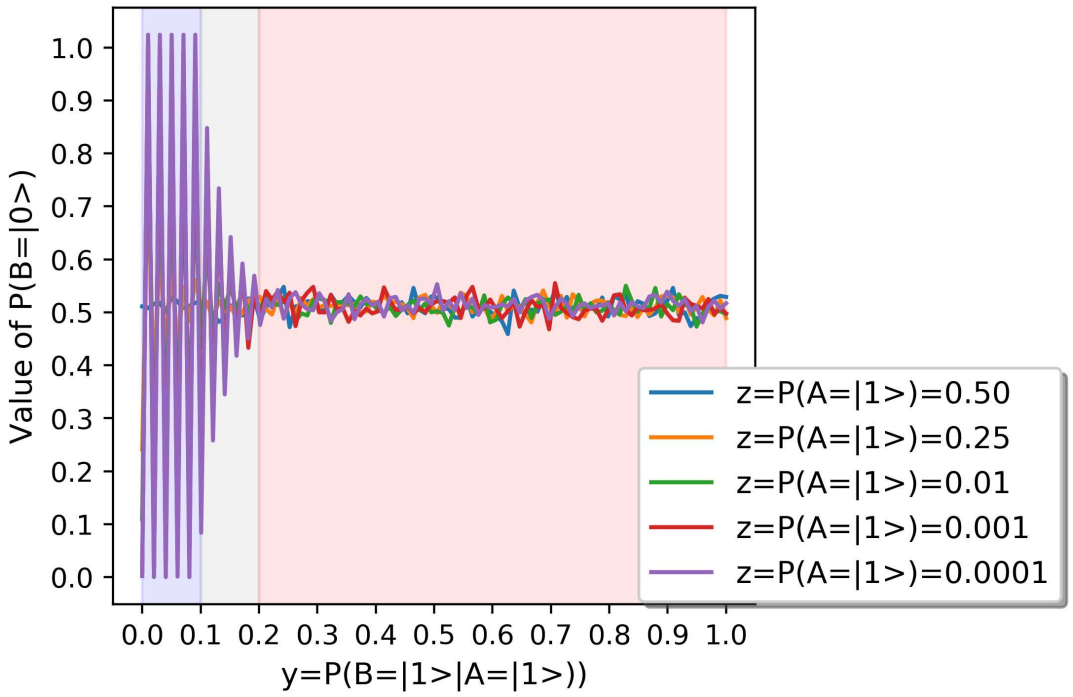
Figure 2. Results obtained for $P(B = |0\rangle)$ for different values of the no–alignment probability of agent A, $z = P(A = |1\rangle)$. (**a**) P(A = $|1\rangle$) = 0.50 (**b**) P(A = $|1\rangle$) = 0.25 (**c**) P(A = $|1\rangle$) = 0.1 (**d**) P(A = $|1\rangle$) = 0.01 (**e**) P(A = $|1\rangle$) = 0.0001.

These results can be interpreted as follows:

- Figure 2a describes the alignment state of agent B, $P(B = |0\rangle)$, for different values of conditioned alignment probability between agents A and B, $x = P(B = |1\rangle\,|A = |0\rangle), y = P(B = |1\rangle\,|A = |1\rangle) \in [0,1]$, being the alignment probability of agent A $P(A = |0\rangle) = 1 - P(A = |1\rangle) = 50\%$.
- Figure 2b describes the alignment state of agent B, $P(B = |0\rangle)$, for different values of conditioned alignment probability between agents A and B, $x = P(B = |1\rangle\,|A = |0\rangle), y = P(B = |1\rangle\,|A = |1\rangle) \in [0,1]$, being the alignment probability of agent A $P(A = |0\rangle) = 1 - P(A = |1\rangle) = 75\%$.
- Figure 2c describes the alignment state of agent B, $P(B = |0\rangle)$, for different values of conditioned alignment probability between agents A and B, $x = P(B = |1\rangle\,|A = |0\rangle), y = P(B = |1\rangle\,|A = |1\rangle) \in [0,1]$, being the alignment probability of agent A $P(A = |0\rangle) = 1 - P(A = |1\rangle) = 90\%$.
- Figure 2d describes the alignment state of agent B, $P(B = |0\rangle)$, for different values of conditioned alignment probability between agents A and B, $x = P(B = |1\rangle\,|A = |0\rangle), y = P(B = |1\rangle\,|A = |1\rangle) \in [0,1]$, being the alignment probability of agent A $P(A = |0\rangle) = 1 - P(A = |1\rangle) = 99\%$.
- Figure 2e describes the alignment state of agent B, $P(B = |0\rangle)$, for different values of conditioned alignment probability between agents A and B, $x = P(B = |1\rangle\,|A = |0\rangle), y = P(B = |1\rangle\,|A = |1\rangle) \in [0,1]$, being the alignment probability of agent A $P(A = |0\rangle)$=$1 - P(A = |1\rangle) = 99.99\%$.

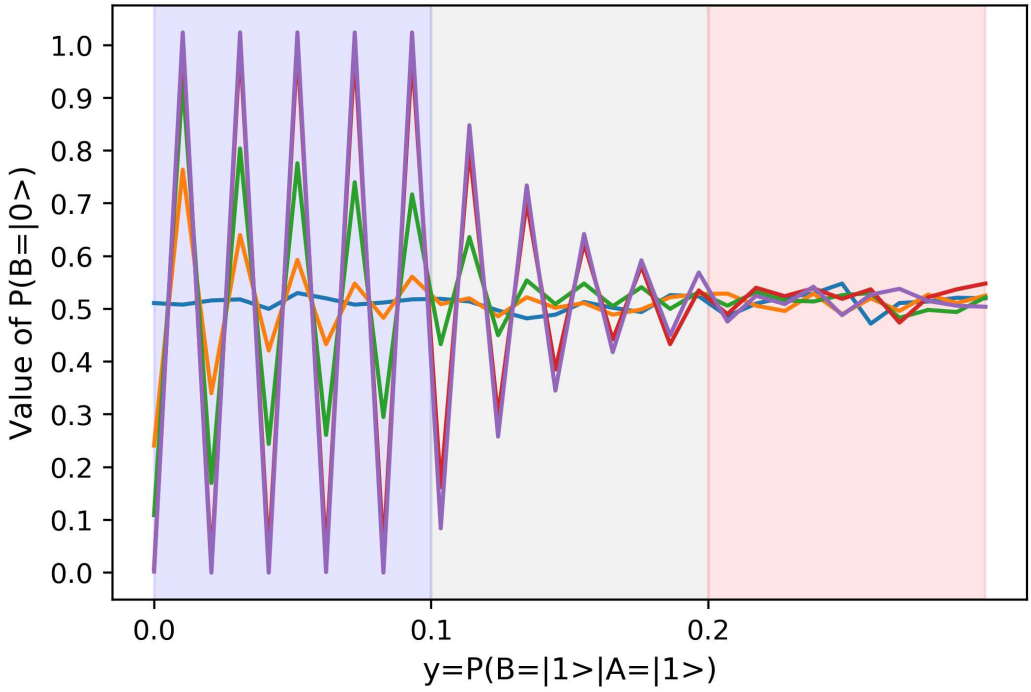
We can intuitively observe in Figure 2 that the first partial derivatives of the two dimensional functions $f(x,y) = P(B = |0\rangle)$ with changing $x = P(B = |1\rangle\,|A = |0\rangle)$ and with $y = P(B = |1\rangle\,|A = |1\rangle)$, given respectively by $\partial f(x,y)/\partial x$ and $\partial f(x,y)/\partial y$. We represent the values of $f(x,y) = P(B = |0\rangle)$ as a function of $y = P(B = |1\rangle\,|A = |1\rangle)$ in Figure 3a. As we indicate in Figure 3b, each [0.1] interval of interval of $y = P(B = |1\rangle\,|A = |1\rangle)$ contains ten values of $x \in [0,1]$.

If we increase the granularity of the mapping of the quantum circuits in search of a fractality within their behavior, and we make a mapping of the $y = P(B = |1\rangle\,|A = |1\rangle) \in [0, 0.1]$ for values of $z = P(A = |1\rangle) = 0.1$, then we obtain the results of Figure 4b. Similarly as in the previous diagrams, each [0.01] interval of $y = P(B = |1\rangle\,|A = |1\rangle)$ contains ten values of $x \in [0,1]$ for values of $z = P(A = |1\rangle) = 0.1$.



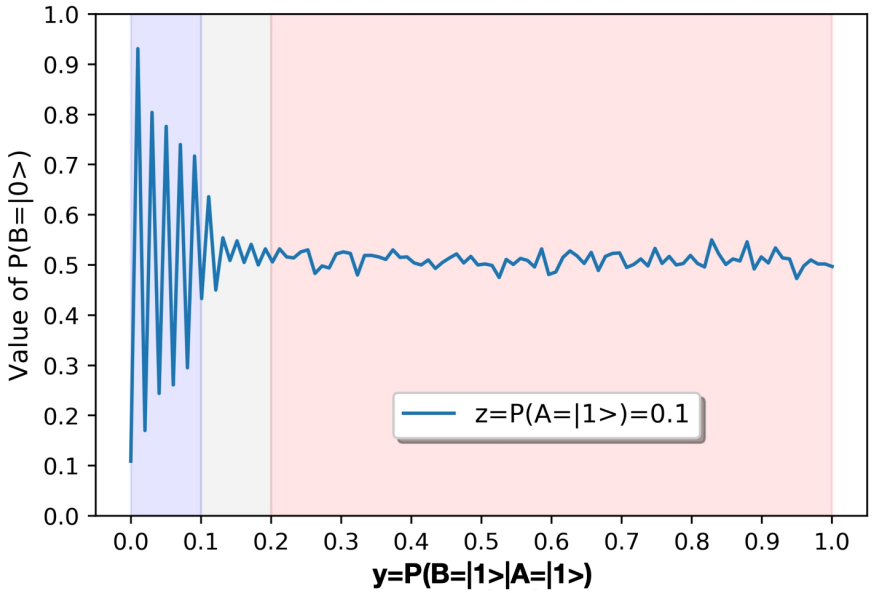
(**a**)

Figure 3. *Cont.*



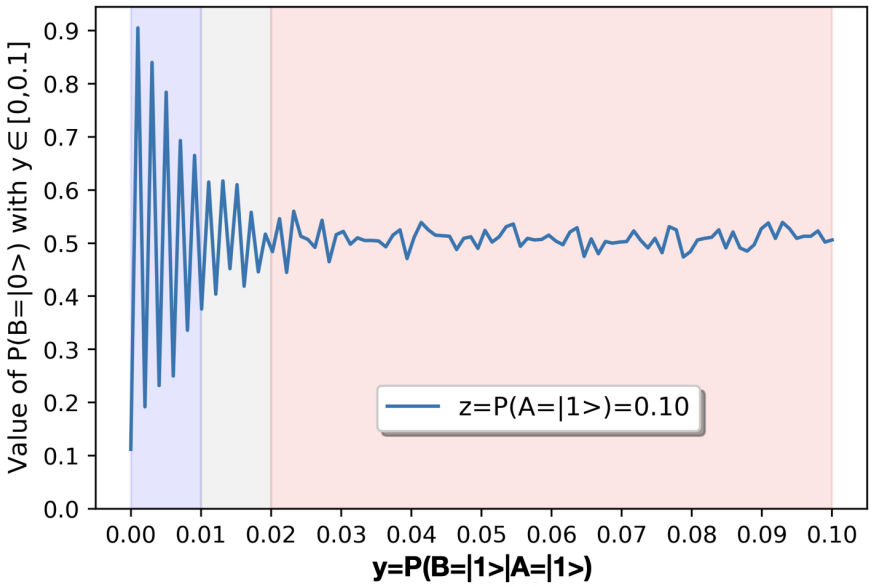
(**b**)

Figure 3. Summary of results of $P(B = |0\rangle)$. (**a**) Summary of results of $P(B = |0\rangle)$ as a function of $y = P(B = |1\rangle\,|A = |1\rangle)$. (**b**) Detail of results of $P(B = |0\rangle)$ with $y = P(B = |1\rangle\,|A = |1\rangle) \in [0, 0.3]$.



(**a**)

Figure 4. *Cont.*



(**b**)

Figure 4. Detail of results of $P(B = |0\rangle)$ for $P(A = |1\rangle = 0.1$. (**a**) Detail of results of $P(B = |0\rangle)$ with $y = P(B = |1\rangle\,|A = |1\rangle) \in [0, 1]$. (**b**) Detail of results of $P(B = |0\rangle)$ with $y = P(B = |1\rangle\,|A = |1\rangle) \in [0, 0.1]$.

In the following Section 4 we discuss these results in detail.

4. Discussion

We will now proceed to discuss the results **R** systematically. We will begin by discussing Figure 3, which describes the change in the alignment probability of the agent B described by the function $f(x, y) = P(B = |0\rangle)$, with increasing values of the relative alignment probability of B, depending on A, given by $y = P(B = |1\rangle\,|A = |1\rangle) \in [0, 1]$. Before we do so, a gentle reminder for the reader that taking into consideration Bayes's theorem, this probability can be expressed with Equation (5):

$$y = P(B = |1\rangle\,|A = |1\rangle) = \frac{P(B = |0\rangle \cap A = |1\rangle)}{P(A = |1\rangle)} \tag{5}$$

This means that growing values of the relative probability of agent B alignment, conditioned to agent A being in not-alignment, are caused by growing values of the intersection $P(B = |1\rangle \cap A = |1\rangle)$. In other words, increasing values of the counter $P(B = |1\rangle \cap A = |1\rangle)$ express that the probability of the intersection of $B = |0\rangle$ and $A = |1\rangle$ is high and therefore both present similar states.

Accordingly, the following results can be enumerated:

R1. In general, we can say that the probability of alignment of agent B, $f(x, y) = P(B = |0\rangle)$, oscillates consistently around the value 0.5 as a harmonic underdamped oscillator for different values of $z = P(A = |1\rangle)$, which is the equilibrium state of the system. This is plausible.

R2. At the scale represented in Figure 3a we observe that the angular frequency of this oscillator changes for different values of $y = P(B = |1\rangle\,|A = |1\rangle)$ and therefore we can separate the behavior of the function in three different regions, marked in Figure 3a, and depicted in Figure 3b in detail.

R3. For values of $y = P(B = |1\rangle\,|A = |1\rangle) \in (0.2, 1]$, the probability of alignment of agent B, $f(x, y) = P(B = |0\rangle)$, oscillates consistently around 0.5 with a minimal amplitude for all values of $z = P(A = |1\rangle)$.

R4. For values of $y = P(B = |1\rangle \,|A = |1\rangle) \in (0.1, 0.2]$, the probability of alignment of agent B, $f(x, y) = P(B = |0\rangle)$, oscillates consistently around 0.5 with an exponential decay that consistently increases with $1 - z = P(A = |0\rangle)$.

R5. For values of $y = P(B = |1\rangle \,|A = |1\rangle) \in [0, 0.1]$, the probability of alignment of agent B, $f(x, y) = P(B = |0\rangle)$, oscillates consistently around 0.5. The oscillation presents an exponential decay for $1 - z = P(A = |0\rangle) \in [0.5, 0.9)$, and presents no decay for values of $1 - z = P(A = |0\rangle) > 0.9$.

R6. However, the most striking observation of all is that if we increase the mapping of the circuits by a factor of 10, as shown in Figure 4b in which we make a mapping of the $y = P(B = |1\rangle \,|A = |1\rangle) \in [0, 0.1]$ with [0.01] intervals for $P(A = |1\rangle = 0.1$, we observe the same behavior as with the mapping of the $y = P(B = |1\rangle \,|A = |1\rangle) \in [0, 1]$ with [0.1] intervals for $P(A = |1\rangle = 0.1$ shown in Figure 4a. The results **R1–R5** are valid for both mapping intervals. The parameters (exponential decay, displacement, amplitude, and phase) of the signals represented in Figure 4 are very similar. Only the frequency is inversely proportional to the mapping interval, hence scaling the oscillation shape, and suggesting that the two graphics depict a similar process. This means that the behavior of this system is fractal. This has powerful implications, which we discuss in the conclusions.

In the following Section 5 we offer the conclusions derived from these results, we discuss the findings and their implications in a broad context, offer certain limitations of the study, and present possible next research paths to pursue.

5. Conclusions, Limitations and Further Steps

Quantum computing explores the processing and exchange of information as natural phenomena that respect the laws of quantum mechanics. The reason for this is that quantum computing makes use of "superposition", which is the ability of quantum computers to simultaneously be in multiple different states. Leading an organizational effort to achieve coordinated strategic goals is a probabilistic process in which the decision makers can rarely be sure that the choice made is the right one. Managers are conditioned by the simultaneous decisions of other agents in the organization whose consequences cannot be fully anticipated a priori. The purpose of this work has been to propose an efficient quantum computing algorithm that is capable of discerning the state of alignment of the organization in the interaction of one process owner informing another and, therefore, supporting the leaders of the organizations in their decision-making process.

We can summarize the main takeaway of this study with the following statement: the case of two qubits shown by QSOD, allows us to affirm that when the strategic objective of the organizational design is to increase the alignment of a process owner, increasing the network by adding a support agent is most likely to have no effect at all on the performance of the agent being reported to. We learn this from **R1**. In fact, the probability of alignment of the original agent always oscillates around the random state.

From result **R2, R3, R4**, and **R5** we also observe that there is an exchange of energy between the original agent and the added agent, so that the alignment probability of the original agent can be positively influenced for low levels of intersection between the alignment probability of the original agent and the added agent. This could be interpreted to mean that the original agent can benefit from the presence of its new partner as long as the new partner provides process information. In other words, if the added agent is able to explain some of the variability in the value creation process that could not previously be explained by the original agent, then the asymptotic stability probability of the original agent will increase. The immediate consequence of this reading is that to add hierarchy levels to strategic design models of organizations, it is necessary to ensure the asymptotic stability of the lower agents before implementing a stable aggregation. It is important to highlight at this point that, in the context of QSOD, a hierarchy does not only describe the rather classical hierarchical relationship between agents but rather a reporting relationship. This concept is more inclusive as it includes the relationship between an agent and his/her boss, but also the relationship between an agent and his/her customer, or the relationship between an agent and a supplier. Our work helps, therefore,

model interactions between organizational process owners and these interactions can potentially be scaled to any organizational context, including small and medium enterprises. This finding is very powerful for industry leaders, as well as for Strategic Organizational Design scholars because it imposes a severe constraint to ensure a sustainable and stable growth of Industry 4.0 organizations.

The implications of result **R6** are profound and reveal the essence of what some call the *fractal organization*. The interaction of two process owners reveals an energy interchange that oscillates with more or less amplitude depending on certain parameters—the conditional alignment probabilities. However, what is really striking is that, independently of the granularity in which these parameters are observed, the oscillation always follows the same pattern. Such pattern is expressed by the results **R1**–**R5** and represents the cornerstone of the bilateral interaction under study. Fractality in this QSOD context allows for a quantification of these complex dynamics and its pervasive effect offers robustness and resilience to the two-*qubit* interaction.

The main limitation of this study is that it only refers to two agents. Furthermore, the simulations of the quantum circuits have been made in a classic computer simulator. Although this undoubtedly reduces some statistical significance to the results, this fact is not relevant to our study at this time and can be neglected.

The results obtained studying the QSOD case of two *qubits* opens new interesting research questions. In order to continue offering a valuable contribution to Industry 4.0 leaders and the research community in general, the future steps we intend to take in this line of research will focus on studying the behavior of other more complex QSOD configurations.

Author Contributions: Conceptualization, J.V.-D.; methodology, J.V.-D.; software, J.V.-D.; validation, J.V.-D., R.M.B. and J.C.L.; formal analysis, J.V.-D.; investigation, J.V.-D.; resources, R.M.B. and J.C.L.; data curation, J.V.-D.; writing—original draft preparation, J.V.-D.; writing—review and editing, R.M.B. and J.C.L.; visualization, J.V.-D.; supervision, R.M.B. and J.C.L.; project administration, R.M.B. and J.C.L.; funding acquisition, J.V.-D. All authors have read and agreed to the published version of the manuscript.

Funding: J.V.D. would also like to acknowledge the Spanish Agencia Estatal de Investigacion, through research project code RTI2018-094614-B-I00 into the "Programa Estatal de I+D+i Orientada a los Retos de la Sociedad". This work was also partially supported by the Ministerio de Ciencia, Innovación y Universidades under Contract No. PGC2018-093854-B-I00.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

QSOD	Quantum Strategic Organizational Design
KPI	Key Performance Indicators

References

1. Villalba-Diez, J.; Zheng, X. Quantum Strategic Organizational Design: Alignment in Industry 4.0 Complex-Networked Cyber-Physical Lean Management Systems. *Sensors* **2020**, *20*, 5856. [CrossRef] [PubMed]
2. Nielsen, T.D.; Jensen, F.V. *Bayesian Networks and Decision Graphs*, 2nd ed.; Springer: New York, NY, USA, 2009.
3. Nielsen, M.A.; Chuang, I. *Quantum Computation and Quantum Information*; Cambridge University Press: Cambridge, UK, 2010.
4. Piattini, M.; Peterssen, G.; Pérez-Castillo, R. Quantum Computing: A New Software Engineering Golden Age. *SIGSOFT Softw. Eng. Notes* **2020**, *45*, 12–14. [CrossRef]
5. Villalba-Diez, J.; Ordieres-Mere, J. Improving manufacturing operational performance by standardizing process management. *Trans. Eng. Manag.* **2015**, *62*, 351–360. [CrossRef]
6. Villalba-Diez, J.; Ordieres-Meré, J.; Nuber, G. The HOSHIN KANRI TREE. Cross-Plant Lean Shopfloor Management. In Proceedings of the The 5th Conference on Learning Factories, Bochum, Germany, 7–8 July 2015; Elsevier: Bochum, Germany, 2015. [CrossRef]

7. Villalba-Diez, J.; Ordieres-Mere, J.; Chudzick, H.; Lopez-Rojo, P. NEMAWASHI: Attaining Value Stream alignment within Complex Organizational Networks. *Procedia CIRP* **2015**, *37*, 134–139. [CrossRef]
8. Villalba-Diez, J.; Ordieres-Mere, J. Strategic Lean Organizational Design: Towards Lean World-Small World Configurations through Discrete Dynamic Organizational Motifs. *Math. Probl. Eng.* **2016**, *2016*, 1–10. [CrossRef]
9. Jimenez, P.; Villalba-Diez, J.; Ordieres-Meré, J. HOSHIN KANRI Visualization with Neo4j. Empowering Leaders to Operationalize Lean Structural Networks. *Procedia CIRP* **2016**, *55*, 284–289. [CrossRef]
10. Villalba-Diez, J. *The HOSHIN KANRI FOREST. Lean Strategic Organizational Design*, 1st ed.; CRC Press, Taylor and Francis Group LLC: Boca Raton, FL, USA, 2017.
11. Villalba-Diez, J. *The Lean Brain Theory. Complex Networked Lean Strategic Organizational Design*; CRC Press. Taylor and Francis Group LLC: Boca Raton, FL, USA, 2017.
12. Villalba-Diez, J.; Zheng, X.; Schmidt, D.; Molina, M. Characterization of Industry 4.0 Lean Management Problem-Solving Behavioral Patterns Using EEG Sensors and Deep Learning. *Sensors* **2019**, *19*, 2841. [CrossRef] [PubMed]
13. Schmidt, D.; Villalba Diez, J.; Ordieres-Meré, J.; Gevers, R.; Schwiep, J.; Molina, M. Industry 4.0 Lean Shopfloor Management Characterization Using EEG Sensors and Deep Learning. *Sensors* **2020**, *20*, 2860. [CrossRef] [PubMed]
14. Mosseri, R.; Dandoloff, R. Geometry of entangled states, Bloch spheres and Hopf fibrations. *J. Phys. Math. Gen.* **2001**, *34*, 10243–10252. [CrossRef]
15. Van Rossum, G. *Python Tutorial, Technical Report CS-R9526*; Centrum voor Wiskunde en Informatica (CWI): Amsterdam, The Netherland, 1995.
16. Wille, R.; Meter, R.V.; Naveh, Y. IBM's Qiskit Tool Chain: Working with and Developing for Real Quantum Computers. In Proceedings of the 2019 Design, Automation Test in Europe Conference Exhibition (DATE), Florence, Italy, 25–29 March 2019; pp. 1234–1240, ISSN 1558-1101. [CrossRef]

sensors

MDPI

Communication

Quantum JIDOKA. Integration of Quantum Simulation on a CNC Machine for In–Process Control Visualization

Javier Villalba-Diez [1,2,*], Miguel Gutierrez [3], Mercedes Grijalvo Martín [3], Tomas Sterkenburgh [4], Juan Carlos Losada [2] and Rosa María Benito [2]

1 Hochschule Heilbronn, Fakultät Management und Vertrieb, Campus Schwäbisch Hall, 74523 Schwäbisch Hall, Germany
2 Complex Systems Group, Universidad Politécnica de Madrid, Av. Puerta de Hierro 2, 28040 Madrid, Spain; juancarlos.losada@upm.es (J.C.L.); rosamaria.benito@upm.es (R.M.B.)
3 Department of Organizational Engineering, Escuela Técnica Superior de Ingenieros Industriales, Universidad Politécnica de Madrid, 28006 Madrid, Spain; miguel.gutierrez@upm.es (M.G.); mercedes.grijalvo@upm.es (M.G.M.)
4 Independent Researcher, Niesmannshof 54, 46535 Dinslaken, Germany; sterkenburgh@t-online.de
* Correspondence: javier.villalba-diez@hs-heilbronn.de

Abstract: With the advent of the Industry 4.0 paradigm, the possibilities of controlling manufacturing processes through the information provided by a network of sensors connected to work centers have expanded. Real-time monitoring of each parameter makes it possible to determine whether the values yielded by the corresponding sensor are in their normal operating range. In the interplay of the multitude of parameters, deterministic analysis quickly becomes intractable and one enters the realm of "uncertain knowledge". Bayesian decision networks are a recognized tool to control the effects of conditional probabilities in such systems. However, determining whether a manufacturing process is out of range requires significant computation time for a decision network, thus delaying the triggering of a malfunction alarm. From its origins, JIDOKA was conceived as a means to provide mechanisms to facilitate real-time identification of malfunctions in any step of the process, so that the production line could be stopped, the cause of the disruption identified for resolution, and ultimately the number of defective parts minimized. Our hypothesis is that we can model the internal sensor network of a computer numerical control (CNC) machine with quantum simulations that show better performance than classical models based on decision networks. We show a successful test of our hypothesis by implementing a quantum digital twin that allows for the integration of quantum computing and Industry 4.0. This quantum digital twin simulates the intricate sensor network within a machine and permits, due to its high computational performance, to apply JIDOKA in real time within manufacturing processes.

Keywords: quantum simulation; JIDOKA; industry 4.0; shopfloor management

Citation: Villalba-Diez, J.; Gutierrez, M.; Grijalvo Martín, M.; Sterkenburgh, T.; Losada, J.C.; Benito, R.M. Quantum JIDOKA. Integration of Quantum Simulation on a CNC Machine for In–Process Control Visualization. *Sensors* **2021**, *21*, 5031. https://doi.org/10.3390/s21155031

Academic Editor: Leopoldo Angrisani

Received: 15 June 2021
Accepted: 21 July 2021
Published: 24 July 2021

1. Introduction

Driven by an unprecedented level of transparency based on the global availability of information, companies are facing extremely competitive global markets in which customers' expectations have risen to demand very high quality standards at a low price and ever-increasing speeds [1]. Adding to this, requests for customised products are growing as to become the pattern in certain industries [2]. Manufacturing industry is relying on technology to face this challenging environment, with Industry 4.0 emerging as a paradigm that can provide solutions to keep track of the markets [3,4].

Kagermann [5] places cyber-physical systems (CPS) as the key driver to trigger the Industry 4.0 paradigm, paralleling its role with the one played, respectively, by steam machines, mass production lines, and integrated circuits in the previous three industrial revolutions. CPS can be defined as "systems of collaborating computational entities

which are in intensive connection with the surrounding physical world and its on-going processes, providing and using, at the same time, data-accessing and data-processing services available on the Internet" [6]. As stated by Lee, Bagheri, and Kao [7] "a CPS consists of two main functional components: (1) the advanced connectivity that ensures real-time data acquisition from the physical world and information feedback from the cyber space; and (2) intelligent data management, analytic and computational capability that constructs the cyber space". Thus, CPS lead to a decentralised control system characteristic of the Industry 4.0, in which machines show great autonomy, share information with other machines, and handle large amounts of data.

For Industry 4.0 to be effective, interdisciplinary knowledge from engineering, computer science, business, and various other academic disciplines is crucial [8]. Lean Management stands out as a consolidated managerial paradigm that tries to tackle the previously mentioned challenges, with a large history since it was grounded in the 1940s by Toyota production managers [9–12]. The main goal of both Industry 4.0 and lean management is to find ways to deal with an ever increasing complexity in the manufacturing industry that nowadays partially stems from digitisation, as well as mass customisation. Although Industry 4.0 puts the focus on the technological elements, lean methods seek to find ways to reduce the complexity by designing clear and controllable processes that minimise non value-adding activities throughout the value chain [13]. To achieve this goal, a wide range of tools, initially conceived under the umbrella of the Toyota just-in-time (JIT) system, such as synchronised production, Kanban, single minute exchange of die (SMED), cross-functional work force [14], and others, that evolved to gain its own field of development as total productive maintenance (TPM) or total quality management (TQM) [15], have been developed and put into practice in manufacturing environments. These techniques have shown significant positive effects in different industries and even synergistic benefits through their combined implementation [16,17]. Combining the methodologies from Industry 4.0 and lean manufacturing has been an increasingly popular research topic, resulting in the so-called Lean 4.0 [18]. There are mixed opinions regarding whether lean management is needed to enable Industry 4.0 or Industry 4.0 advance lean management [19]. Yet, one of the key ideas about Industry 4.0 is the integration of varied technologies due to the limited effect of focusing on a single technology [20]. In fact, several of the many lean tools have been examined in the context of Industry 4.0 [21–23], leading to the conclusion that a leaner production is easier to integrate into an Industry 4.0 context and, in many cases, it is possible to combine the ideas and techniques from both frameworks [24–28]. Moreover, it seems to be a necessary evolutionary step for further raising the level of operational excellence (i.e., to coordinate actions, to optimise resource efficiency, to improve work safety, to decrease in cost) [18]. In this sense, one successful Lean 4.0 approach is the integration of JIDOKA(自働化) into the Industry 4.0 framework, which has shown great potential [29–31] and motivates this work.

JIDOKA(自働化) is currently considered a lean tool that "enables machines to work harmoniously with their human operators and features intelligent capabilities by automatically stopping a process by man or machine, in the event of an abnormally, a problem, such as equipment malfunction, quality issues, or late work" [30,32]. It was introduced by Ohno [14] as one of the two pillars of the Toyota production system, alongside with the JIT, needed to accomplish the elimination of waste. The Japanese term was translated from the coined term "autonomation" or "automation with human touch" and its origins traced back to the invention by Sakichi Toyoda (1867–1930) of a looming machine that would automatically stop as soon as a thread in the machine teared [14]. The core idea behind JIDOKA is to provide "intelligence" to machines with built-in automatic checking systems that would automatically stop to prevent any defective part passing to the next step in the value stream. In case of an abnormal situation, the intervention of an operator that can be in charge of monitoring several processes is necessary. Moreover, to eliminate any source of waste, JIDOKA aims at preventing this mistake from happening, again, reinforcing a culture of continuous learning and improvement. Thus, it is essential to identify the defective

part as soon as possible, trigger a signal to stop the work center, and even the production process if necessary, to determine what the root cause for the production of the defective item has been, but also achieve an effective human–machine interaction. For many years, JIDOKA principles were primarily built on mechanical tools and devices with electronic components playing an increasing role. Romero et al. [30] describe three prior generations of JIDOKA systems: mechanical gadgets that avoid mistakes (POKA–YOKE (ポカヨケ)), visual and audio alarms (ANDON (行灯)), sensor-based fault diagnosis (JIDOKA(自働化) rules). The immense possibilities brought up by Industry 4.0, through digitisation and wide availability of low-cost sensors, open the way to utilise these large amounts of data and be able to act on a variety of input variables and handle complex processes and lead to a new generation JIDOKA 4.0 [30]. Hence, it becomes an example of Industry 4.0 enabling the further integration of lean tools [31]. The impact of Industry 4.0 in the evolution of JIDOKA 4.0 is described in what follows.

As aforementioned, the main core of JIDOKA (自働化), which is the stopping of the machine, can only be one of the first steps as still no further production should take place until the operator has found the mistake. This has been one of the main reasons why many companies, outside Japan, were hesitant at first about implementing JIDOKA(自働化) [14]. Even the use of Andon-systems, which notify the correct person of the shutdown can only reduce but not eliminate those negative effects. The next step is to make the data available to the operator, in order to assist the search for the source of the defective machine. Industry 4.0 can act as a key enabler [33]. Ma, Wang, and Zhao [29] show a successful implementation of a CPS-based JIDOKA (自働化) system. In addition to its technological side, CPS offer multimodal interfaces for more effective human–machine collaboration [34]. It is part of the Industry 4.0 networking concept, production and people as entities, participating in value creation [35]. Another step further consists in directly pointing out the possible source of the problem. An ideal goal is the design of a self-regulating machine, which is able to adapt to different circumstances and prevent the need of stopping the machine. Although there are some limitations to identify the cause of defective parts in all cases, it should be possible to be able to at least better predict whether a part is likely to break. Deuse et al. [31] outline a generic proposal of a system aiming at this objective. This approach relates to maintenance systems where sensors are installed in different parts of the machine to monitor its operation in real time [36]. The data that these sensors send, consequently structured and appropriately processed, allows predicting the failure of the monitored spare parts of the machine. On the other hand, due to the availability of low-cost sensors and suitable mechanisms of sensor integration—e.g., by powerful mobile networks—the amount of available data is drastically increasing. With the benefits of availability comes the effort of analysis and handling of big data. In addition, the human interface needs an appropriate reduction in complexity to enable an effective JIDOKA (自働化) process [32]. Although the target range and probability distributions for the corresponding measured values can be specified for each sensor, the evaluation of sensor data due to networked conditional probabilities creates a high degree of complexity, which quickly overtaxes even experienced employees in the evaluation. Incorrect evaluations and time delays are often the result. In this respect, supporting a system that automatically performs the task of pre-assessing the overall situation for the human as a result of the interaction of the individual measured values can be of considerable benefit. Decision networks, which quickly and efficiently calculate the conditional probabilities of the interconnection of random variables, can be an effective tool to this aim [37]. However, even this approach faces limitations due to the computational power required and leads to delays in feedback to the worker [38]. Quantum simulations shows relevant advantages to accelerate this feedback significantly [39,40]. Hence the interest in studying the possibilities of quantum computing to implement the described JIDOKA (自働化) methodology more efficiently and responsively.

In sum, the growing possibilities of combining lean manufacturing and Industry 4.0 and their associated benefits have been posed by many studies [21–24,26,28], whereas the

rapid evolution of technological advancements opens new fields of applications. Previous works have shown that JIDOKA(自働化) benefits from some Industry 4.0 technologies such as CPS, connectivity, and operator wearable systems [26,29,33], and recently to improve process monitoring in order to predict quality defects [31]. An important aspect in the implementation of JIDOKA (自働化) is the operator, as the system only works through the cooperation of both [30]. Efficient Bayesian network computing algorithms can help firms to adopt JIDOKA (自働化) or "automation with human touch" approach, not only improving the efficiency of the manufacturing process but also making possible a human–machine cooperation system. In this context, we propose the integration of quantum computing and JIDOKA(自働化) to simulate the intricate monitoring sensor network within a machine. This quantum simulation of the machine sensor network is a quantum digital twin.

This quantum digital twin will allow—due to the speed that quantum computing offers—to apply real-time JIDOKA (自働化) in the manufacturing processes. Since this work is oriented to Industry 4.0 users and we want to give it an eminently experimental and practice-oriented character, we prefer to focus on the core of the problem and thus spare the reader from lengthy theoretical explanations on the design of quantum simulations. In references [41–44], interested readers can find papers in which the authors show both the theoretical and practical principles of how to transform decision networks with conditional probabilities into their counterpart quantum circuits.

The remainder of the paper is organised as follows: in Section 2 the case study chosen to test our hypothesis is introduced: we describe the structure and interaction of the overall system with our quantum digital twin and establish the scope in Section 2.1. In Section 2.2 the specific hardware and software are presented, followed by the description of the data collection components of the setup in Section 2.3. Results of our study are summarised in Section 3. Finally, further aspects are discussed in Section 4, where we also draw the main conclusions, limitations, and future applications of our work.

2. Case Study Quantum JIDOKA

Our initial working hypothesis is that we can model the internal sensor network of a machine with quantum simulations that show better performance than classical models based on decision networks. To quantitatively test this hypothesis, as a first step to evaluate the effect of the integration of quantum simulations in I4.0 environments, a case study is used. In this case study, we are going to generate a digital quantum twin to provide the 4.0 operator with a shopfloor management tool that allows him to visualize the status of the machine in real time, as shown schematically in Figure 1.

We integrated the proposed solution in a factory and tested it for 12 weeks before submitting the work. However, for privacy reasons imposed by the process owner, the data related to the process in question cannot be disclosed. For this reason a dummy dataset is presented to show the functionality of the solution, as well as to ensure reproducibility to interested scholars or industrialists. The machine in question consists of 5 sensors that measure the rotational speed of two motors that, in turn, drive two drills that make two holes, whose relative position determines the quality of the final product. As argued by Byrd and Turner [45], a single case study can be seen as the only possible building block in the process of developing the validity and reliability of the proposed hypothesis. Following the recommendations of Eisenhardt [46], a clear case study road-map is followed for each one of them. This road-map has several phases: (1) scope establishment, (2) specification of hardware and software, (3) data collection, and (4) quantum digital twin.

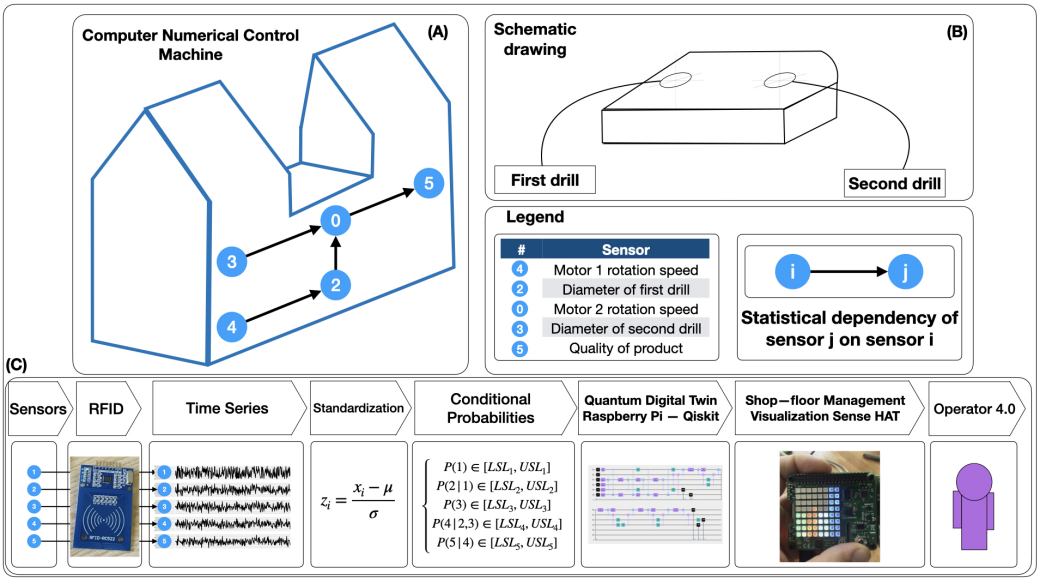


Figure 1. Scheme of our Quantum JIDOKA case study: the digital quantum twin to provide the 4.0 operator with a shopfloor management tool that allows him to visualize the status of the machine in real time. (**A**) CNN-machine and schematic sensor network. (**B**) Schematic product drawing. (**C**) Quantum（自働化) creation process.

2.1. Scope Establishment

The use of quantum computers in the implementation of larger decision networks and their use in a JIDOKA（自働化) environment should open up the possibility of performing high-performance analysis in larger sensor networks.

The objective of this case study is to generate a low-cost quantum digital twin that represents the statistical dependencies between five sensors positioned on a computer numerical control processing machine that measure parameters relevant to the quality of the manufactured product. For this purpose, we will install a quantum circuit in a low-cost component, which simulates the statistical dependencies derived from the value-creation process inside the machine. This component will receive data from the machine through a radio frequency identification (RFID) device and will compute the state of the machine in real time, generating a visualization that will allow the process owner to understand it. This interface will resemble a conditioning monitoring and will allow for a preventive operational lean shopfloor management.

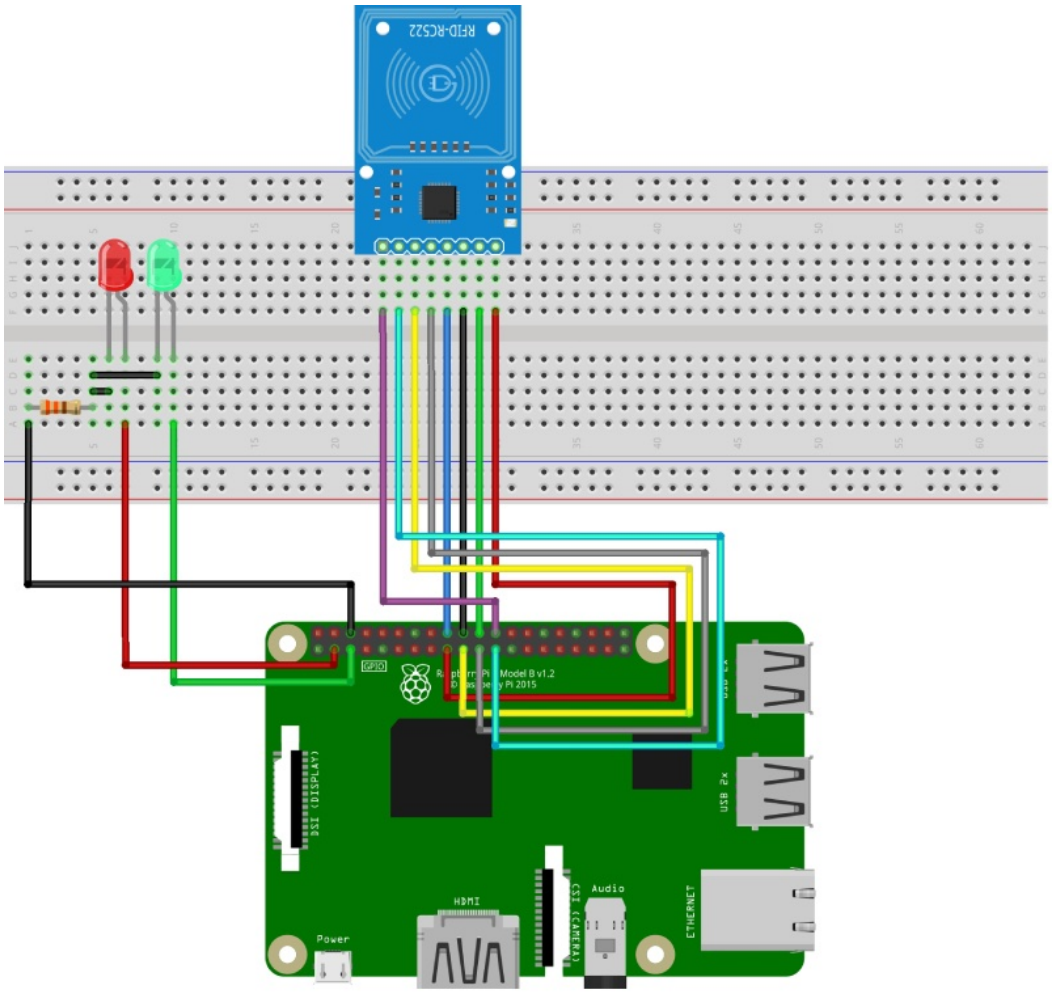
The implementation of a quantum computer simulation on a Turing machine is certainly accompanied by limitations, which are reflected in the performance and low number of manageable measured values. Nevertheless, the setup chosen here allows a feasibility study and can serve as a basis for later scaling when used in a real system environment and with the use of adequate quantum computers.

2.2. Specification of Hardware and Software

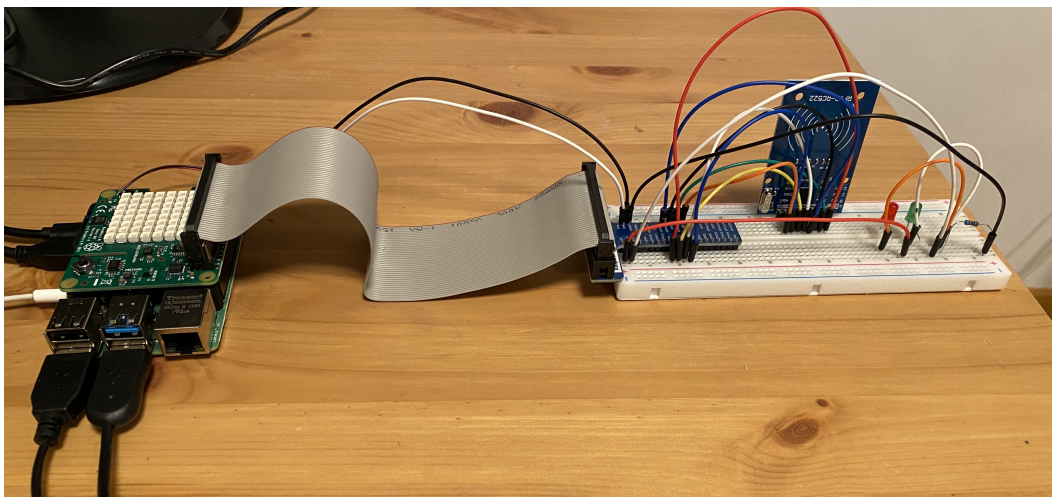
The proposed equipment has a low-cost standard configuration and the components can be easily available on the market. This design was consciously chosen for two reasons: on the one hand reliability is increased as the compounds are well proven, on the other hand an adequate stock of spare parts can be stored to ensure the continuous operation of the process at a very low cost.

2.2.1. Hardware

The necessary hardware needed to build the smart IoT sensor prototype is shown in Figure 2:



(**a**) Hardware circuit plan.



(**b**) Hardware circuit real.

Figure 2. RFID LEDs Raspberry Pi Visualization Hat-Hardware.

- Raspberry Pi 4.0. The Raspberry Pi 4.0 is a high-performance 64-bit quad-core processor, up to 8 GB of RAM, dual-band 2.4/5.0 GHz wireless local area network (LAN), Bluetooth 5.0, Gigabit Ethernet, USB 3.0, and a sense HAT add-on which is attached on top of the Raspberry Pi via the 40 general-purpose input or output pins (which

provide the data and power interface). It has several sensors and an 8 × 8 RGB (Red–Green–Blue) LED matrix display that can be used to visualize sensor states for multiple applications [47–50]. In the proposed design, the critical component, because of its value and relative complexity is the Raspberry Pi CPU. In the factory in question there are about two hundred CPUs of this type, and the annual failure rate, including human-caused failures, is 1%. This is acceptable and within standard maintenance parameters.

- RC522 RFID module. The RC522 RFID is a 13.56 MHz RFID module that is based on the MFRC522 controller from NXP semiconductors. Its operating voltage lies between 2.5 V to 3.3 V. It allows for serial peripheral interface (SPI), inter-integrated circuit (I^2C) and universal asynchronous receiver and transmitter (UART) communication protocols. Its maximum data rate is 10 Mbps with a read range of 5 cm and a current consumption of 13 to 26 mA. These characteristics are optimal for a number of industrial and educational applications [51,52].
- Set of cables, light-emitting diodes (LEDs), resistors, and test plates. An LED is a two-lead semiconductor light source, which emits light when activated. The LEDs used present a forward current of 30 mA and a forward voltage range between 1.8 V and 2.4 V. A resistor is a passive electronic component that offers a specific amount of electrical resistance to the flow of current when connected in a circuit. The resistors used in this project are standard 1 kΩ.

The Raspberry Pi is connected to the RFID card and the two red and green LEDs show the status of the connection. The electrical resistance allows for a proper functioning of the elements.

2.2.2. Software

The quantum digital twin circuits presented in Section 2.4 were simulated on *qiskit* tool, a Python-based [53] quantum computing platform developed by IBM [54], and the code and additional results can be accessed in this Open Access Repository: https://osf.io/24jrm/?view_only=a10d2e001e114807854b994616f8d4cf. We will analyze the data obtained in Section 2.3, evaluate these results in Section 3, and discuss the results obtained in Section 4.

2.3. Data Collection

The data input is realised by means of a radio frequency identification (RFID) module, connected to the serial peripheral interface bus (SPI) of the Raspberry Pi as presented in Figure 2b. This module acts as RFID-reader. To simulate a more realistic industrial process, we set up a data transfer by RFID consisting of the following components:

- An RFID-writer, connected to the sensors of the computer numerical control machine in Figure 1;
- The RFID-reader as described in Figure 2;
- A small machine that is physically rotating an RFID-card around a motor-driven axis, as shown in Figure 3, and by this transferring datasets from the writer to the reader.

One transfer process consists of one revolution of the RFID-card around the axis. Per revolution, one dataset is transferred. Each dataset consists of one measured value per sensor, as well as a time stamp and a quality assessment in the form of “ok” or “not ok” information. The process starts with the RFID-card in reach for the RFID-writer. Here, one dataset from the computer numerical control machine is written to the RFID-Tag. After that, the card is rotating around the motor axis. With a 180 degree rotation, the card comes into reach of the RFID reader. Here, the dataset is read from the RFID card, split into its components, entered into one list per sensor on the Raspberry Pi, and thereby made available on the digital twin.

Figure 3. RFID data collection.

2.4. Quantum Digital Twin

The quantum digital twin circuit that resembles the sensor network in the computer numerical control machine of Figure 1 is shown in Figure 4. Following the recommendations from [40,55], we build the quantum digital twin with a number of qubits equal to the number of sensors and one *ancilla* qubit that serves to perform the appropriate rotations, giving a total of 6 qubits in this case. In addition, as we have previously indicated in [41], after the proper qubit initialization, we perform a series of qubit rotation operations that allow us to simulate the conditional probabilities between the respective sensors. The interested reader can inspect several examples for several qubit configurations in these application examples from [42–44].

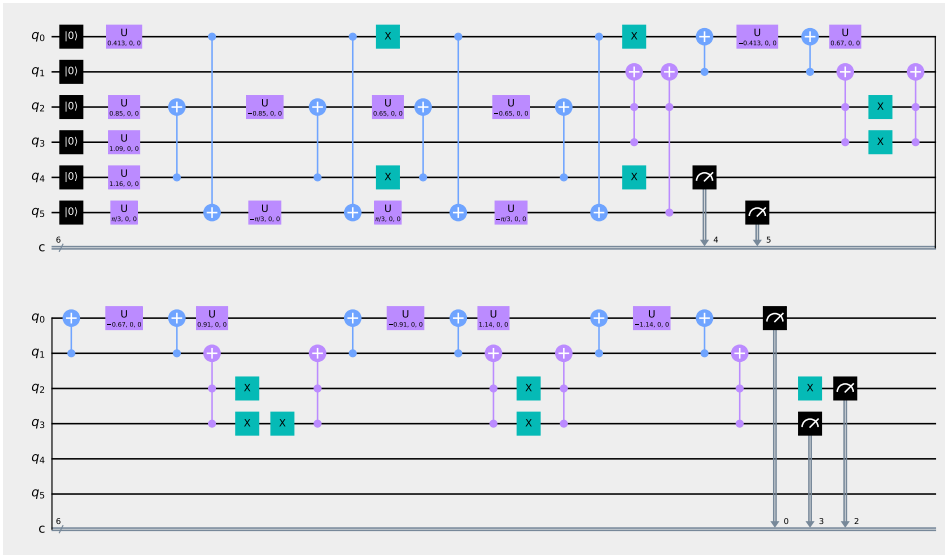


Figure 4. Quantum digital twin for the sensor network shown in Figure 1.

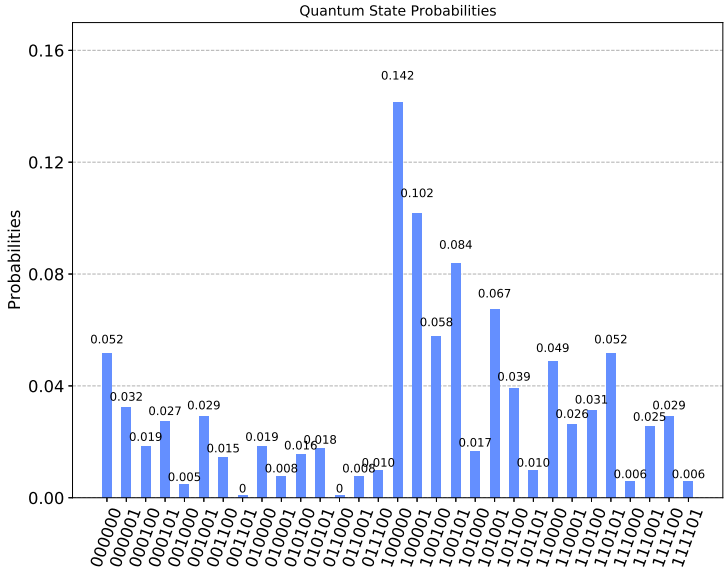
The factory in which this project was carried out has a workforce with a low level of education, which is why management has to look for intuitive solutions to visualize the state of the value creation processes so that workers can understand them and act on them. To this end, a visualization based on the logic of traffic lights, known to all workers, is used: green and red mean that the machine is operating within or outside specifications, respectively, while yellow shows a situation of caution as the machine is operating at the limit of specifications. The display of the state of the last qubit is performed

by means of a sense HAT that presents a linear colour degradation given by the expression $RGB255 \cdot [\|\langle 1|q_i\rangle\|^2, \|\langle 0|q_i\rangle\|^2), 0]$, $i = 0, \ldots, 5$. This yields naturally to a green colour $RGB[0, 255, 0]$ if the probability of the last qubit of the circuit $P(q_5 = |0\rangle) = \|\langle 0|q_5\rangle\|^2 = 1$ and a red colour $RGB[255, 0, 0]$ if the $P(q_5 = |1\rangle) = \|\langle 1|q_5\rangle\|^2 = 1$ which is the standard traffic-light colour code in the shopfloor: green, yellow, and red.

3. Results

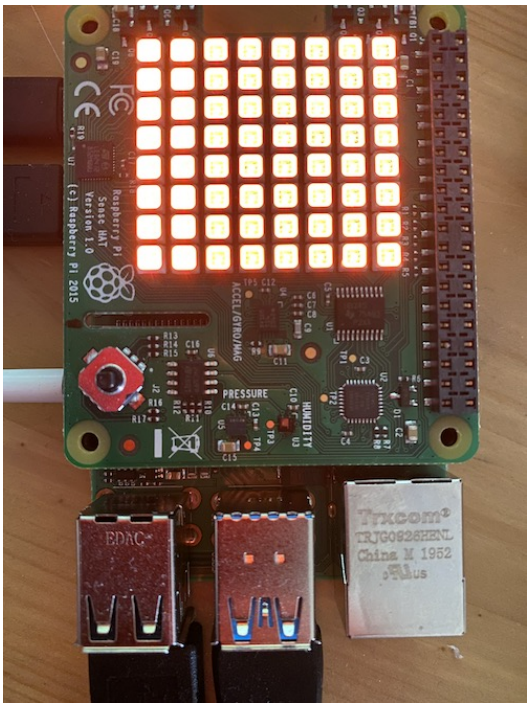
The results obtained by the simulation of the digital quantum twin can be represented in the form of a bar chart representing the quantum state probabilities of the quantum circuit, as shown in Figure 5a. However, this visualization is not intuitive and, therefore, has little chance of being interpreted satisfactorily by the process owner in an Industry 4.0 environment. Without this visual interpretation of the machine status, it is not possible to successfully perform a proper shopfloor management. For this reason, we have added an 8×8 RGB LED matrix display that allows a quick and intuitive visualization of the total state of the wave equation of the quantum circuit. This is exemplary shown in Figure 5b, which represent the sum of the 32 qubit combinations $P(q_5 = |0\rangle)$ given by $P(q_5 = |0\rangle) = \|\langle 0|q_5\rangle\|^2 = 0.25$, hence delivering a reddish visualization equivalent to $RGB[191, 64, 0]$.

To verify that the result is correct, we proceed to design the equivalent Bayesian network to the digital quantum twin [40] represented in Figure 6. The error percentage found when comparing both quantum digital twin and the classical Bayesian network is less than 2%, which is acceptable in the context of quantum simulations. The computation time of the equivalent Bayesian network doubles the quantum digital twin computational performance, which is a pre-requisite for the implementation to be carried out in real time.



(**a**) Quantum digital twin state probabilities

Figure 5. *Cont.*



(**b**) Quantum digital twin sense HAT shopfloor visualization $P(q_5 = |0\rangle) = \|\langle 0|q_5\rangle\|^2 = 0.25$

Figure 5. Quantum digital twin state probabilities and shopfloor visualization.

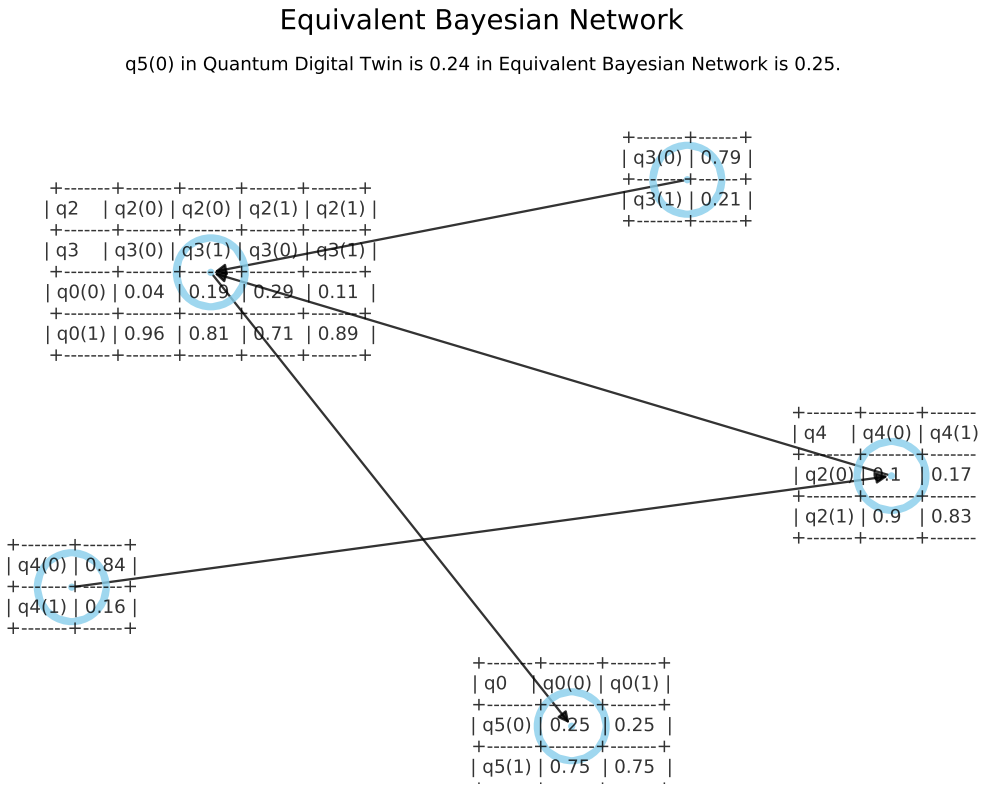


Figure 6. Equivalent Bayesian network to the quantum digital twin shown in Figure 4.

4. Discussion and Conclusions

In the initial situation, before starting the project, manual quality control was performed on 1% of the parts. The initial proposal of the process owner was to perform a condition monitoring of each of the sensors and aggregate this information by means of a

digital twin with machine learning methods, which would reduce the cost of personnel associated with quality control. However, it would be difficult to integrate this into the production process due to two reasons: on the one hand, the lack of knowledge of these methods by the operators, and, on the other hand, the computational cost of performing calculations by classical methods (Bayesian networks) that do not allow integration in the production. Our quantum digital twin allows to obtain the two advantages: to obtain the information on the product quality in the form of an intuitive visualization for the operator in real time and at low cost. The integration of the quantum digital twin has meant a reduction in the costs associated with quality control, as well as doubling the mean-time-between-failures associated with the CNC machine as the speed of reaction of the operator in case of error is increased.

In this work, we have successfully tested the integration of a digital quantum twin by means of quantum simulations on a conventional machine to enable a visualization of its systemic state in an Industry 4.0 environment. With the case study, we have achieved the much desired integration of quantum computing logic in industrial environments and opened a field of exploration that will allow, once emulated, managers of value creation processes to use these algorithms in a clean and simple way. Digital quantum twins is much more than just software for reaching the same goal in a slightly better way. Instead, it does not develop in isolation, but the socio-technical system enables the development, diffusion and use of technologies. Digital quantum twins necessarily change processes and the way in which people work. With this work, we have generated a bridge, so far non-existent in practice, between the world of quantum simulation and industrial environments. We have also confirmed the validity of the results by demonstrating that the quantum digital twin yields the same values as those obtained with traditional simulation methods, such as Bayesian networks.

We have integrated a quantum simulation that allows on the one hand to monitor the state of a sensor network inside the machine and on the other, through an intuitive traffic light visualization, shopfloor management system, to empower the process owner to benefit from the quantum digital twin results without any quantum knowledge. This means that our proposal has the potential to be widely applied in practice since it requires neither a high investment, nor a redesign of its components, nor a specific knowledge of quantum simulation principles. From an autonomation view, these characteristics of usability, selective provision of information, user acceptance, and profitability could result not only in better human-machine cooperation systems, but also may lead to changes in the range, depth, and content of tasks.

The objective of our work is the practical application of quantum simulation in a real environment. For this purpose, we present a real implementation of a prototype connected to a CNC machine in industrial use. Nevertheless, the proposal has some limitations. Quantum computers constitute a huge investment due to the required physical functioning conditions. Currently, quantum computing is available either using external free resources as IBM *qiskit*, the one used in the prototype that has a limit of 30-qubits, or through rental of proprietary equipment. Escalation to an extensive industrial deployment is neither expected to imply relevant barriers from the conceptual and technological point of view, nor it affects the design and logic of the prototype. Should the company hire a more powerful quantum computer, it would not affect the prototype functioning. The extra qubits would allow to compute Bayesian networks composed of a higher number of sensors. Computational benefits would then be significantly higher as the increase rate of quantum computation times with the number of network nodes is much slower than binary logic computers. However, challenges remain in structuring the condition monitoring offer, due to the different domains of application, the characteristics of the existing information, and the final goal of the monitoring activities.

Author Contributions: Conceptualization, J.V.-D.; methodology, J.V.-D.; software implementation, J.V.-D. and T.S.; validation, J.V.-D.; formal analysis, J.V.-D., T.S., and M.G.; writing—original draft preparation, J.V.-D., T.S., M.G., and M.G.M.; writing—review and editing, J.V.-D., M.G., T.S., J.C.L., and R.M.B.; visualization, J.V.-D. All authors have read and agreed to the published version of the manuscript.

Funding: J.V.-D. would like to acknowledge the Spanish Agencia Estatal de Investigacion, through research project code RTI2018-094614-B-I00 of the "Programa Estatal de I+D+i Orientada a los Retos de la Sociedad". J.V.-D. also wants to thank the EU RFCS research program for the contribution to this research through the grant with ID 793505 (WISEST project), mainly due to the industrial cases provided. This work was also partially supported by the Ministerio de Ciencia, Innovación y Universidades under Contract No. PGC2018-093854-B-I00.

Data Availability Statement: The code and additional results can be accessed in this Open Access Repository: https://osf.io/24jrm/?view_only=a10d2e001e114807854b994616f8d4cf.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CNC	Computer numerical control
CPS	Cyber-physical systems
I^2C	Inter-integrated circuit
JIT	Just-in-time
LAN	Local area network
QC	Quantum computing
LED	Light emitting diode
RFID	Radio frequency identification
RGB	Red–Green–Blue
SFM	Shopfloor management
SMED	Single minute exchange of die
SPI	Serial peripheral interface
TPM	Total productive maintenance
TQM	Total quality management
UART	Universal asynchronous receiver/transmitter

References

1. Wu, L.; Liu, H.; Su, K. Exploring the dual effect of effectuation on new product development speed and quality. *J. Bus. Res.* **2020**, *106*, 82–93. [CrossRef]
2. Wagner, R.M. (Ed.) *Industrie 4.0 für die Praxis*; Springer Gabler: Wiesbaden, Germany, 2018.
3. Kagermann, H.; Lukas, W.D.; Wahlster, W. Industrie 4.0: Mit dem Internet der Dinge auf dem Weg zur 4. industriellen Revolution. *VDI Nachrichten* **2011**, *13*, 1.
4. Lasi, H.; Fettke, P.; Kemper, H.G.; Feld, T.; Hoffmann, M. Industry 4.0. *Bus. Inf. Syst. Eng.* **2014**, *6*, 239–242. [CrossRef]
5. Kagermann, H. Change through digitization—Value creation in the age of industry 4.0. In *Management of Permanent Change*; Springer Science + Business Media: Berlin/Heidelberg, Germany, 2015; pp. 23–45. [CrossRef]
6. Monostori, L.; Kádár, B.; Bauernhansl, T.; Kondoh, S.; Kumara, S.; Reinhart, G.; Sauer, O.; Schuh, G.; Sihn, W.; Ueda, K. Cyber-physical systems in manufacturing. *CIRP Ann.* **2016**, *65*, 621–641. [CrossRef]
7. Lee, J.; Bagheri, B.; Kao, H.A. A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems. *Manuf. Lett.* **2015**, *3*, 18–23. [CrossRef]
8. Muhuri, P.K.; Shukla, A.K.; Abraham, A. Industry 4.0: A bibliometric analysis and detailed overview. *Eng. Appl. Artif. Intell.* **2019**, *78*, 218–235. [CrossRef]
9. Melton, T. The benefits of lean manufacturing: What lean thinking has to offer the process industries. *Chem. Eng. Res. Des.* **2005**, *83*, 662–673. [CrossRef]
10. Villalba-Diez, J.; Ordieres-Meré, J. Improving manufacturing performance by standardization of interprocess communication. *IEEE Trans. Eng. Manag.* **2015**, *62*, 351–360. [CrossRef]
11. Villalba-Diez, J. *The Hoshin Kanri Forest. Lean Strategic Organizational Design*, 1st ed.; CRC Press: Boca Raton, FL, USA; Taylor and Francis Group LLC: Boca Raton, FL, USA, 2017.
12. Villalba-Diez, J. *The Lean Brain Theory. Complex Networked Lean Strategic Organizational Design*; CRC Press: Boca Raton, FL, USA; Taylor and Francis Group LLC: Boca Raton, FL, USA, 2017.

13. Bick, W. Warum Industrie 4.0 und Lean zwingend zusammengehören. *VDI-Z* **2014**, *156*, 46–47.
14. Ohno, T. *Toyota Production System: Beyond Large-Scale Production*; Productivity Press: New York, NY, USA, 1988.
15. Cua, K.O.; McKone, K.E.; Schroeder, R.G. Relationships between implementation of TQM, JIT, and TPM and manufacturing performance. *J. Oper. Manag.* **2001**, *19*, 675–694. [CrossRef]
16. Shah, R.; Ward, P.T. Lean manufacturing: Context, practice bundles, and performance. *J. Oper. Manag.* **2003**, *21*, 129–149. [CrossRef]
17. Jolayemi, J.K. Hoshin kanri and hoshin process: A review and literature survey. *Total Qual. Manag.* **2008**, *19*, 295–320. [CrossRef]
18. Mayr, A.; Weigelt, M.; Kühl, A.; Grimm, S.; Erll, A.; Potzel, M.; Franke, J. Lean 4.0-A conceptual conjunction of lean management and Industry 4.0. *Procedia Cirp* **2018**, *72*, 622–628. [CrossRef]
19. Rüttimann, B.G.; Stöckli, M.T. Lean and industry 4.0—Twins, partners, or contenders? A due clarification regarding the supposed clash of two production systems. *J. Serv. Sci. Manag.* **2016**, *9*, 485–500. [CrossRef]
20. Stentoft, J.; Rajkumar, C. The relevance of Industry 4.0 and its relationship with moving manufacturing out, back and staying at home. *Int. J. Prod. Res.* **2020**, *58*, 2953–2973. [CrossRef]
21. Sanders, A.; Elangeswaran, C.; Wulfsberg, J.P. Industry 4.0 implies lean manufacturing: Research activities in industry 4.0 function as enablers for lean manufacturing. *J. Ind. Eng. Manag. (JIEM)* **2016**, *9*, 811–833. [CrossRef]
22. Mrugalska, B.; Wyrwicka, M.K. Towards lean production in industry 4.0. *Procedia Eng.* **2017**, *182*, 466–473. [CrossRef]
23. Wagner, T.; Herrmann, C.; Thiede, S. Industry 4.0 impacts on lean production systems. *Procedia CIRP* **2017**, *63*, 125–131. [CrossRef]
24. Kolberg, D.; Knobloch, J.; Zühlke, D. Towards a lean automation interface for workstations. *Int. J. Prod. Res.* **2017**, *55*, 2845–2856. [CrossRef]
25. Buer, S.V.; Strandhagen, J.O.; Chan, F.T. The link between industry 4.0 and lean manufacturing: Mapping current research and establishing a research agenda. *Int. J. Prod. Res.* **2018**, *56*, 2924–2940. [CrossRef]
26. Pereira, A.C.; Dinis-Carvalho, J.; Alves, A.C.; Arezes, P. How Industry 4.0 can enhance lean practices. *FME Trans.* **2019**, *47*, 810–822. [CrossRef]
27. Neto, J.M.; Salomon, V.A.; Petrillo, A.; Akabane, G.K. Impact of Industry 4.0 on the Lean Management concept: A systematic literature review. In Proceedings of the IJCIEOM 2020—International Joint Conference on Industrial Engineering and Operations Management, Rio de Janeiro, Brazil, 8–11 July 2020. [CrossRef]
28. Pagliosa, M.; Tortorella, G.; Ferreira, J.C.E. Industry 4.0 and Lean Manufacturing: A systematic literature review and future research directions. *J. Manuf. Technol. Manag.* **2021**, *32*, 543–569. [CrossRef]
29. Ma, J.; Wang, Q.; Zhao, Z. SLAE—CPS: Smart lean automation engine enabled by cyber-physical systems technologies. *Sensors* **2017**, *17*, 1500. [CrossRef]
30. Romero, D.; Gaiardelli, P.; Powell, D.; Wuest, T.; Thürer, M. Rethinking jidoka systems under automation & learning perspectives in the digital lean manufacturing world. *IFAC-PapersOnLine* **2019**, *52*, 899–903. [CrossRef]
31. Deuse, J.; Dombrowski, U.; Nöhring, F.; Mazarov, J.; Dix, Y. Systematic combination of Lean Management with digitalization to improve production systems on the example of JIDOKA 4.0. *Int. J. Eng. Bus. Manag.* **2020**, *12*. [CrossRef]
32. Baudin, M. *Working with Machines: The Nuts and Bolts of Lean Operations with JIDOKA*; Productivity Press: New York, NY, USA, 2007.
33. Kolberg, D.; Zühlke, D. Lean Automation enabled by Industry 4.0 Technologies. *IFAC-PapersOnLine* **2015**, *28*, 1870–1875. [CrossRef]
34. Broy, M.; Cengarle, M.V.; Geisberger, E. Cyber-physical systems: Imminent challenges. *Lect. Notes Comput. Sci.* **2012**, *7539 LNCS*, 1–28. [CrossRef]
35. Schneider, P. Managerial challenges of Industry 4.0: An empirically backed research agenda for a nascent field. *Rev. Manag. Sci.* **2018**, *12*, 803–848. [CrossRef]
36. Grijalvo Martín, M.; Pacios Álvarez, A.; Ordieres-Meré, J.; Villalba-Díez, J.; Morales-Alonso, G. New Business Models from Prescriptive Maintenance Strategies Aligned with Sustainable Development Goals. *Sustainability* **2021**, *13*, 216. [CrossRef]
37. Dey, S.; Stori, J.A. A Bayesian network approach to root cause diagnosis of process variations. *Int. J. Mach. Tools Manuf.* **2005**, *45*, 75–91. [CrossRef]
38. McCabe, B. Belief networks for engineering applications. *Int. J. Technol. Manag.* **2001**, *21*, 257–270. [CrossRef]
39. Low, G.H.; Yoder, T.J.; Chuang, I.L. Quantum inference on Bayesian networks. *Phys. Rev. A At. Mol. Opt. Phys.* **2014**, *89*, 062315. [CrossRef]
40. Borujeni, S.E.; Nannapaneni, S.; Nguyen, N.H.; Behrman, E.C.; Steck, J.E. Quantum circuit representation of Bayesian networks. *Expert Syst. Appl.* **2021**, *176*, 114768. [CrossRef]
41. Villalba-Diez, J.; Zheng, X. Quantum Strategic Organizational Design: Alignment in Industry 4.0 Complex-Networked Cyber-Physical Lean Management Systems. *Sensors* **2020**, *20*, 5856. [CrossRef]
42. Villalba-Diez, J.; Benito, R.M.; Losada, J.C. Industry 4.0 Quantum Strategic Organizational Design Configurations. The Case of Two Qubits: One Reports to One. *Sensors* **2020**, *20*, 6977. [CrossRef] [PubMed]
43. Villalba-Diez, J.; Losada, J.C.; Benito, R.M.; González-Marcos, A. Industry 4.0 Quantum Strategic Organizational Design Configurations. The Case of 3 Qubits: One Reports to Two. *Entropy* **2021**, *23*, 374. [CrossRef]
44. Villalba-Diez, J.; Losada, J.C.; Benito, R.M.; Schmidt, D. Industry 4.0 Quantum Strategic Organizational Design Configurations. The Case of 3 Qubits: Two Report to One. *Entropy* **2021**, *23*, 426. [CrossRef]

45. Byrd, T.; Turner, D. Measuring the flexibility of information technology infrastructure: Exploratory analysis of a construct. *J. Manag. Inf. Syst.* **2000**, *17*, 167–208.
46. Eisenhardt, K.M. Building theories from case study research. *Acad. Manag. Rev.* **1989**, *14*, 532–550. [CrossRef]
47. Johnston, S.J.; Cox, S.J. The Raspberry Pi: A Technology Disrupter, and the Enabler of Dreams. *Electronics* **2017**, *6*, 51. [CrossRef]
48. Dhillon, N.S.; Sutandi, A.; Vishwanath, M.; Lim, M.M.; Cao, H.; Si, D. A Raspberry Pi-Based Traumatic Brain Injury Detection System for Single-Channel Electroencephalogram. *Sensors* **2021**, *21*, 2779. [CrossRef] [PubMed]
49. Kim, W.; Jung, I. Simulator for Interactive and Effective Organization of Things in Edge Cluster Computing. *Sensors* **2021**, *21*, 2616. [CrossRef] [PubMed]
50. Chang, Y.H.; Sahoo, N.; Chen, J.Y.; Chuang, S.Y.; Lin, H.W. ROS-Based Smart Walker with Fuzzy Posture Judgement and Power Assistance. *Sensors* **2021**, *21*, 2371. [CrossRef] [PubMed]
51. Tan, P.; Wu, H.; Li, P.; Xu, H. Teaching Management System with Applications of RFID and IoT Technology. *Educ. Sci.* **2018**, *8*, 26. [CrossRef]
52. Laxmi, A.R.; Mishra, A. RFID based Logistic Management System using Internet of Things (IoT). In Proceedings of the 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 29–31 March 2018; pp. 556–559. [CrossRef]
53. Van Rossum, G. *Python Tutorial, Technical Report CS-R9526*; Centrum voor Wiskunde en Informatica (CWI): Amsterdam, The Netherlands, 1995.
54. Wille, R.; Meter, R.V.; Naveh, Y. IBM's Qiskit Tool Chain: Working with and Developing for Real Quantum Computers. In Proceedings of the 2019 Design, Automation Test in Europe Conference Exhibition (DATE), Florence, Italy, 25–29 March 2019; pp. 1234–1240. [CrossRef]
55. Nielsen, M.A.; Chuang, I. *Quantum Computation and Quantum Information*; Cambridge University Press: Cambridge, UK, 2010.

sensors

MDPI

Communication

Improvement of Quantum Approximate Optimization Algorithm for Max–Cut Problems

Javier Villalba-Diez [1,2], Ana González-Marcos [3] and Joaquín B. Ordieres-Meré [4,*]

[1] Hochschule Heilbronn, Fakultät Management und Vertrieb, Campus Schwäbisch Hall, 74523 Schwäbisch Hall, Germany; javier.villalba-diez@hs-heilbronn.de
[2] Complex Systems Group, Universidad Politécnica de Madrid, Av. Puerta de Hierro 2, 28040 Madrid, Spain
[3] Department of Mechanical Engineering, Universidad de La Rioja, San José de Calasanz 31, 26004 Logroño, Spain; ana.gonzalez@unirioja.es
[4] Escuela Técnica Superior de Ingenieros Industriales (ETSII), Universidad Politécnica de Madrid, José Gutiérrez Abascal 2, 28006 Madrid, Spain
* Correspondence: j.ordieres@upm.es

Abstract: The objective of this short letter is to study the optimal partitioning of value stream networks into two classes so that the number of connections between them is maximized. Such kind of problems are frequently found in the design of different systems such as communication network configuration, and industrial applications in which certain topological characteristics enhance value–stream network resilience. The main interest is to improve the Max–Cut algorithm proposed in the quantum approximate optimization approach (QAOA), looking to promote a more efficient implementation than those already published. A discussion regarding linked problems as well as further research questions are also reviewed.

Keywords: Industry 4.0; quantum approximate optimization algorithm; value–stream networks; optimization

Citation: Villalba-Diez, J.; González-Marcos, A.; Ordieres-Meré, J.B. Improvement of Quantum Approximate Optimization Algorithm for Max–Cut Problems. *Sensors* **2022**, *22*, 244. https://doi.org/10.3390/s22010244

Academic Editor: Juan M. Corchado

Received: 2 December 2021
Accepted: 24 December 2021
Published: 30 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

Value chains linked to Industry 4.0 (I4.0) involve complex cyber-physical networks in which information is processed efficiently by humans and machines to deliver the desired product to a customer [1–3]. I4.0 and the Industrial Internet of Things (IIoT) both describe further emerging landscapes for an integrated human–machine interaction [4,5]. Together, the two concepts are grounded in intelligent, interconnected cyber–physical manufacturing systems that are fully equipped and capable of controlling the process flow of industrial production. Given that many decisions are made independently by machines interoperating with production planning and fabrication systems, the integration of human users requires new paradigms [6].

In the realm of IIoT I4.0 manufacturing, I4.0 vision has advanced the notions of smart fabrication and smart factory by augmenting all assets with sensor–based connectivity [7]. These intelligent sensors generate a large amount of manufacturing data that helps to create digital twins as support for a live mirror of physical processes [8,9]. The ambition is to capture process variability within this approach, with the capability to process all relevant information by analyzing big data in cloud computation so that manufacturers are able to find bottlenecks in manufacturing processes, identify the causes and impacts of problems in such a way that the effective application of measures is useful for both product design and manufacturing engineering, including maintenance, repair and overhaul [10].

Quantum near–term simulations in classical computers have been recently used to solve different applications [6,11], including Industry 4.0 challenges such as the modelling of organizational decision networks as quantum circuits [12]. In this work, with the help of quantum simulations, a new solution for the combinatorial optimization problem is proposed, which can be applied to a wide range of applications including in the Industry

4.0 environment. It consists in finding the "optimal" partitioning of a value chain into two classes, such that the number of connections between them is maximized. Direct applications are linked to introduce flexibility in value chain models by enabling extra resilience to it, no matter whether the related processes are logistics or production related ones. Solving this industrial process design problem potentially allows maximizing the interaction between the elements of the value chain and thus maximizes its productivity [13,14]. Other applications can be connected to the Narrowband Internet of Things (NB–IoT) technology. NB–IoT is a cellular radio–based access protocol specified by 3GPP to tackle the quickly growing market for low–power wide–area connectivity significantly targeting mobile use cases. To realize the global outreach and broad adoption of NB–IoT services, mobile network operators (MNOs) need to guarantee end-to-end devices and services across several vendors connected to the deployed NB–IoT systems, and that the data transport capacity and connection modes are well understood. In this context, efficient dynamic partitions depending on the low power network available are relevant for providing a robust integrative configuration with limited transport overhead.

A solution to these problems can be formulated in terms of a combinatorial optimization approach, which involves finding an optimal object out of a finite set of objects. In this particular case, it involves finding "optimal" bitstrings composed of 0's and 1's among a finite set of bitstrings. Such bitstring represents a partition of nodes of a graph into two sets, such that the number of edges between the sets is maximum. Each of the sets represents the allocation of nodes in the value stream network or nodes in the IoT system to specific managerial structures giving a maximal flexibility by providing the highest degree of connectivity.

This optimization challenge is already known as the Max–Cut problem, and it is one of the most studied combinatorial optimization problems because of its wide range of applications and because of its connections with other fields of discrete mathematics [15]. Different solutions have been proposed for the Max–Cut type of problems, as it belongs to the so-called NP-hard complexity class problems, where no known algorithms are able to solve NP-hard problems in polynomial time and thus exact methods rapidly become intractable. Such solutions include search-based algorithms [16], Machine Learning alternatives [17], as well as Recurrent Neural Networks and Reinforcement Learning [18,19].

Quantum approaches were also proposed with a quantum approximate optimization algorithm (QAOA) by [20]. The QAOA belongs to the class of hybrid algorithms and requires, in addition to the execution of shallow quantum circuits, a classical optimization process to improve the quantum circuit itself. The QAOA is an algorithm that uses unitary transformations $U(\beta_i, \gamma_i)$, depends on two parameters β_i and γ_i, and is arranged in alternating blocks a number p of times ($i \in \{1, \ldots, p\}$) given by

$$|\psi(\vec{\beta}, \vec{\gamma})\rangle = \underbrace{U(\beta_i)U(\gamma_i)...U(\beta_i)U(\gamma_i)}_{p \text{ times}} |\psi_0\rangle \tag{1}$$

where $|\psi_0\rangle$ is a suitable state and parameters $\vec{\beta}, \vec{\gamma} \in \mathbb{R}^p$.

The goal of the algorithm is to find the combination of parameters that allows a quantum state $|\psi(\overrightarrow{\beta_{opt}}, \overrightarrow{\gamma_{opt}})\rangle$ to yield the optimal solution [21]. This heuristic algorithm produces then a bit string $x \in \{0,1\}^n$ that with high probability is expected to give a good approximation of the theoretical solution. The algorithm follows a classical optimization scheme: first prepares a parameterized quantum state $|\psi(\vec{\beta}, \vec{\gamma})\rangle$ (called the ansatz), then computes the parameters $(\overrightarrow{\beta_{opt}}, \overrightarrow{\gamma_{opt}})$ such that the expectation value of the quantum state is given by

$$F_p = \langle \psi(\vec{\beta}, \vec{\gamma}) | \, H_p \, | \psi(\vec{\beta}, \vec{\gamma}) \rangle \tag{2}$$

is maximized with respect to the problem Hamiltonian H_p, and finally performs a classical optimization until some convergence criterion is reached. An overview of this is shown schematically in Figure 1.

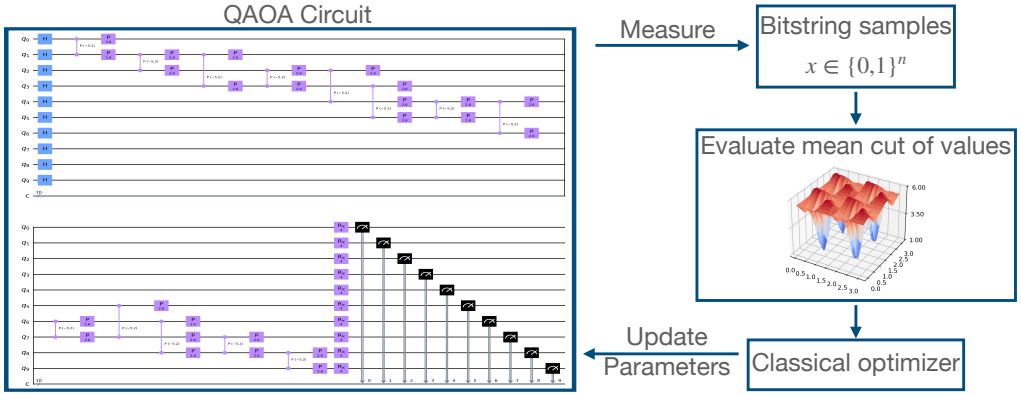


Figure 1. QAOA Overview.

The convergence criterion is in the Max–Cut cost function given by

$$C(x) = \sum_{i,j=1}^{n} x_i(1 - x_j), \tag{3}$$

which can be mapped to a Hamiltonian that is diagonal in the computational basis by

$$H = \sum_{x \in \{0,1\}^n} C(x) \left|x\right\rangle \left\langle x\right|, \tag{4}$$

in which $x \in \{0,1\}^n$ labels the computational basis states $\left|x\right\rangle \in \mathbb{C}^{2^n}$. The expansion of $Z_i = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ Pauli–Z operators can be obtained from the canonical expansion of $C(x)$ by substitution of every variable $x_i \in \{0,1\}$ by the matrix $\frac{1}{2}(1 - Z_i)$.

As indicated in the abstract, this paper aims to show that the already suggested approximate solution can be improved. The proposal for an alternative quantum algorithm configuration improves the existing solutions up to thirty nodes. The optimization algorithm proposed in Farhi et al. [20] promotes a specific sequence of unitary operators, which means an effective expression for the Hamiltonian. Finally, such a sequence of transformations will perform differently when the size of the circuit evolves. Our approach can be understood in the end as a proposal for a different sequence of unitary operators, providing a different configuration for the Hamiltonian. Then, what it is claimed is that our algorithm (our effective expression for the Hamiltonian) performs much better than the existing one.

The solution is implemented in a simulated quantum hardware environment, however there are already studies showing the time and noise effects over these algorithms when implemented in real hardware [22].

We structure the rest of the work hereinafter as follows: Section 2 outlines the modified architecture in a reasoned manner. Then, Section 3 presents the results of the algorithm as compared with the analytical solution, which for when $|\psi(\overrightarrow{\beta_{opt}}, \overrightarrow{\gamma_{opt}})\rangle$ is not too deep can be computed classically, and the results previously obtained by [20]. Finally, Section 4 briefly discuss the obtained results, outlines future lines of research, and presents limitations in the presented work.

2. Modified QAOA

In this section, we present the results of the algorithm applied to a value stream network of $n = 10$ nodes. The complete results for other configurations are available in open access in [23].

We start by representing in Figure 2 the value stream network as a graph $G = \{n, e\}$ of $n = 10$ unlabeled nodes and $e = 13$ edges.

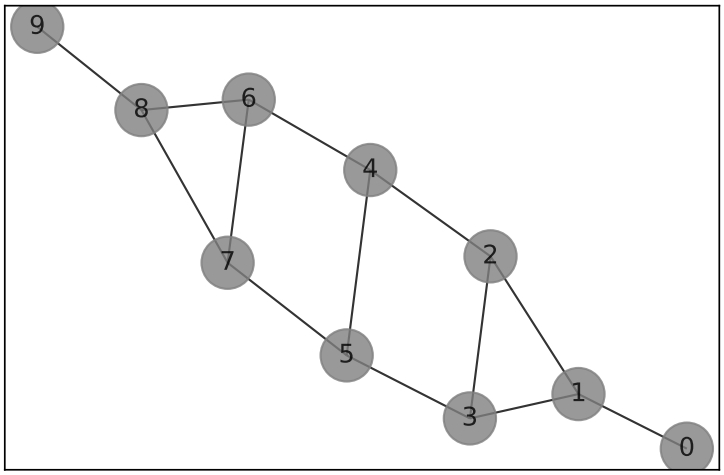


Figure 2. Value stream network with n = 10 nodes.

If the graph coincides with the connectivity of our logical network (either IoT topology or value stream network), the cost function $C(x)$ coincides with the hamiltonian H used to generate the state.

For a shallow approximation with p = 1, the analytical solution for the expectation value is given by

$$F_1(\beta, \gamma) = \langle \psi_1 | H | \psi_1 \rangle \tag{5}$$

Combining Equations (3) and (4), the Hamiltonian H makes use of the expectation value to measure the edges individually:

$$f_{i,k}(\gamma, \beta) = \frac{1}{2} \langle \psi_1(\gamma, \beta) | (1 - Z_i Z_k) | \psi_1(\gamma, \beta) \rangle \tag{6}$$

There are two types of edges: those that connect a node with degree one (A), and those that connect a node with degree three (B). For the A–class edges, an example of the encoding of the optimization function between nodes (0) and (1) is given by

$$2f_A = 1 - \langle +^1 | U_{01}(\gamma) U_{12}(\gamma) U_{13}(\gamma) X_0(\beta) X_1(\beta) Z_0 Z_1 X_1^\dagger(\beta) X_0^\dagger(\beta) U_{01}^\dagger(\gamma) U_{12}^\dagger(\gamma) U_{13}^\dagger(\gamma) | +^1 \rangle \tag{7}$$

and for the B–class edges, the encoding of the optimization function between nodes (1) and (2) is given by

$$2f_B = 1 - \langle +^3 | U_{21}(\gamma) U_{24}(\gamma) U_{23}(\gamma) X_1(\beta) X_2(\beta) Z_1 Z_2 X_1^\dagger(\beta) X_2^\dagger(\beta)(\gamma) U_{12}^\dagger(\gamma) U_{23}^\dagger(\gamma) U_{24}^\dagger(\gamma) | +^3 \rangle \tag{8}$$

in which $|+^n\rangle = \sum_{x \in \{0,1\}^n} \frac{1}{\sqrt{2^n}} |x\rangle$ prepares for an equal superposition state followed by a sequence of parametrized unitary operations. As shown in Equations (7) and (8) these unitary operations are a combination of parametrized Hamiltonian cost $U_C(\gamma) = e^{-i\gamma H_C}$ and mixer layers $U_M(\beta) = e^{-i\beta H_M}$. The subindexes in the unitary operations indicate the nodes on which the operators act upon.

In our case n = 10, there are two A–class edges and eleven B–class edges. This yields Equation (9), which is depicted in Figure 3 which shows the periodicity in both parameters and exhibits a highly nonlinear behaviour.Farhi et al. [20] proposed QAOA with the structure represented in Figure 4.

$$\begin{aligned} F_1(\beta, \gamma) = 2f_A(\beta, \gamma) + 11 f_B(\beta, \gamma) = \Big[\sin(4\gamma) \sin(4\beta) + \sin^2(2\beta) \sin^2(2\gamma) \Big] + \\ + \frac{11}{2} \Big[1 - \sin^2(2\beta) \sin^2(2\gamma) \cos^2(4\gamma) - \frac{1}{4} \sin(4\beta) \cos(4\gamma) \Big(1 + \cos^2(4\gamma) \Big) \Big] \end{aligned} \tag{9}$$

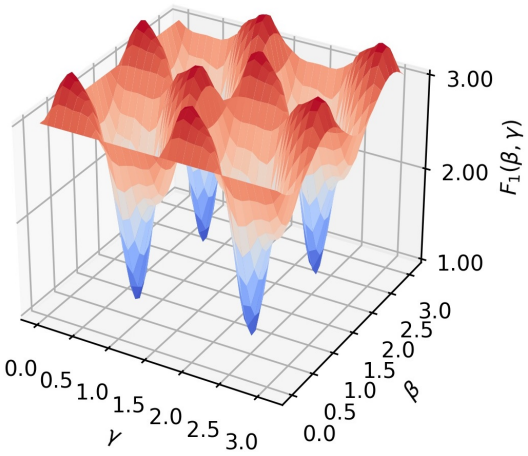


Figure 3. Analytic solution for $p = 1$ and value stream network configuration from Figure 2.

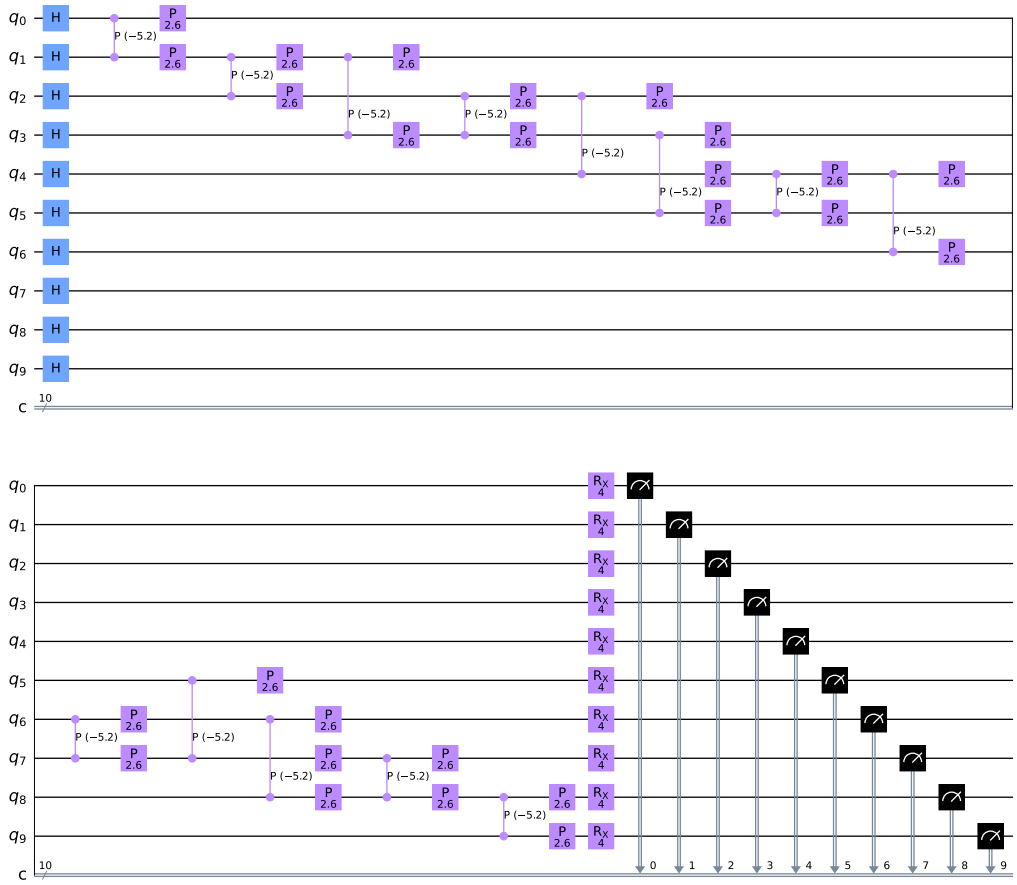


Figure 4. QAOA—Farhi et al. [20].

Such an algorithm starts by preparing the system in superposition with a Hadamard gate on all qubits. Next, a rotation of 2γ is applied to each of the edges if both are in state $|11\rangle$. This conditional rotation has the form given by

$$C_p(2\gamma) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & e^{-2i\gamma} \end{pmatrix}. \tag{10}$$

This allows the algorithm to be applied when both qubits are in state $|1\rangle$ simultaneously. A quantum phase correction of γ is then applied to each of the nodes joined by each edge. This rotation has the form given by

$$p(\gamma) = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\gamma} \end{pmatrix}. \tag{11}$$

Such configuration allows the previous rotation to be neutralized when the two qubits are in state $|11\rangle$. The result of these two operations allows a rotation of γ to be applied to all nodes as long as both communicating nodes are not simultaneously in state $|1\rangle$.

Finally, a rotation around the X-axis of 2β, perpendicular to the computing axis, is applied to all nodes. This rotation has the form given by

$$R_x(2\beta) = \begin{pmatrix} \cos\beta & -i\sin\beta \\ -i\sin(2\beta) & \cos(2\beta) \end{pmatrix}. \tag{12}$$

In summary, in [20] the QAOA algorithm applies, after a standard superposition, a quantum phase of γ to every node connected to each other, as long as both are not in state $|11\rangle$, and a rotation around the perpendicular to the computational axis of 2β to all the nodes.

On the other hand, this paper proposes a novel QAOA approach represented in Figure 5.

Analogous to the previous example, our algorithm starts by preparing the system in superposition with a Hadamard gate on all qubits. We then perform a conditional rotation of γ to each node connected to another if the second is in state $|1\rangle$ in both directions. This is done by concatenation of two $U3(\frac{\gamma}{2}, 0, 0)$ and $U3\left(\frac{-\gamma}{2}, 0, 0\right)$ gates given by

$$U3\left(\frac{\gamma}{2}, 0, 0\right) = \begin{pmatrix} \cos\left(\frac{\gamma}{4}\right) & -\sin\left(\frac{\gamma}{4}\right) \\ \sin\left(\frac{\gamma}{4}\right) & \cos\left(\frac{\gamma}{4}\right) \end{pmatrix}, \tag{13}$$

and a conditional CX rotation applied to one of the nodes q_0 taking the other q_1 as control given by

$$CX_{q_0,q_1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \tag{14}$$

This method works because when the control *qubit* $|\Psi_0\rangle$ is in state $|0\rangle$, all we have is $U3(\frac{\gamma}{2}, 0, 0)$ followed by a $U3\left(\frac{-\gamma}{2}, 0, 0\right)$ and the effect is trivial. On the other hand, when the control *qubit* $|\Psi_0\rangle$ is in state $|1\rangle$, the net effect is a controlled rotation $U3(\gamma, 0, 0)$ on the $|\Psi_1\rangle$ *qubit*. These rotations are taken in both directions because our network is not directed. This algorithm is expected to yield better results than the previous one because the transformations are differential as a function of node state.

Finally, as in the previous algorithm, a rotation around the X-axis of 2β, perpendicular to the computing axis, is applied to all nodes. This rotation has the form given by Equation (12).

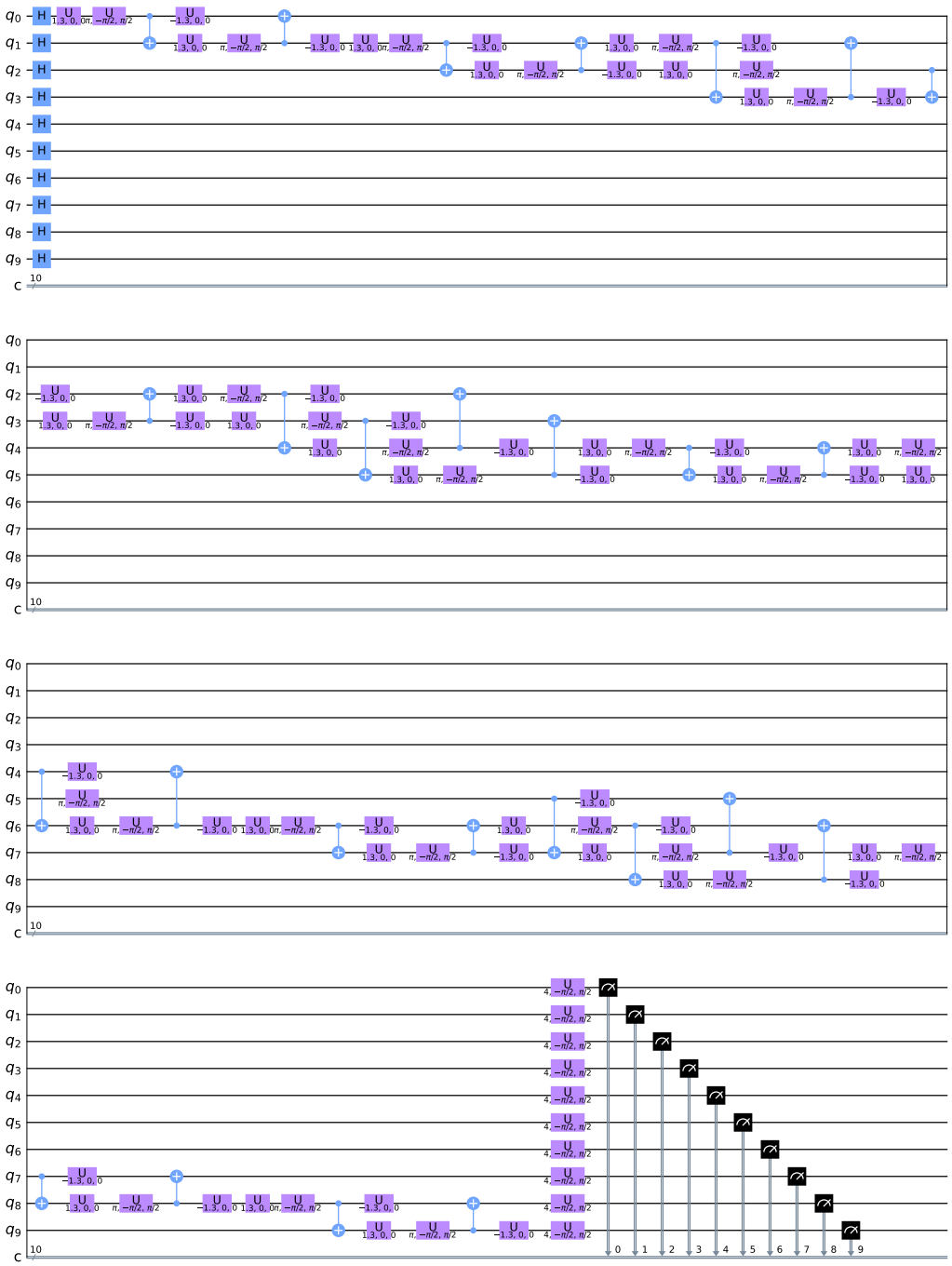


Figure 5. QAOA—Villalba–Diez et al.

3. Results

In this section, we present the results of the algorithm applied to a logical nondirected network of $n = 10$ nodes. The quantum simulations presented were simulated on *qiskit* tool, a Python–based quantum computing platform developed by IBM [24], and the code and additional results can be accessed in this Open Access Repository: [23].

The results confirm our expectations and our proposed QAOA algorithm predicts the analytical results better for a shallow quantum circuit with $p = 1$. A summary of the results for different numbers of nodes is shown in Figure 6. In Table 1 we represent the comparison of the analytical solution curve and the respective QAOA algorithms. Our solution shows better performance in all metrics.

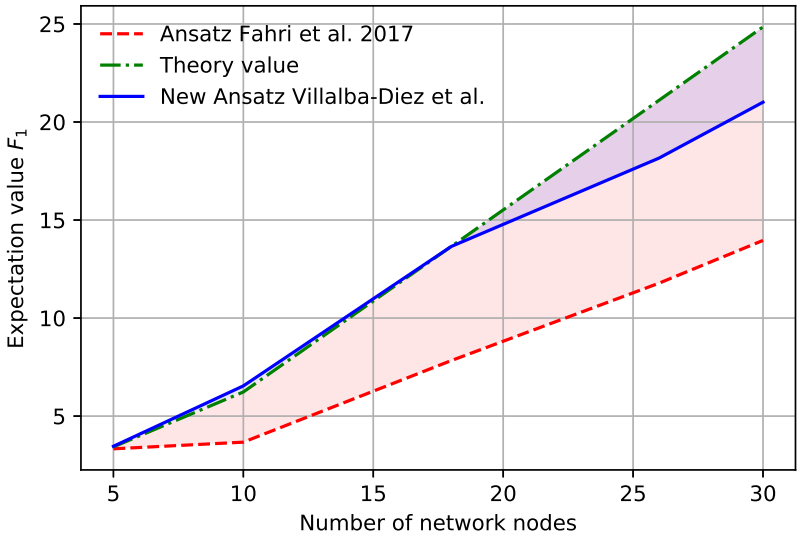


Figure 6. QAOA results comparison.

Table 1. Results comparison for different measures for identifying curve similarity [25].

	Analytic vs.	
	Farhi et al.	**Villalba et al.**
Directed Hausdorff distance	8.22	3.84
Discrete Fréchet distance	10.89	3.84
Dynamic Time Wrapping	28.70	7.13
Partial Curve Mapping	1.6893	0.3223
Area between two curves	1.2744	0.3642
Curve-Length distance metric	141.21	26.23

The bit string that delivers the optimal solution is $x = \{0110011010\}$, as shown in Figure 7. This graph clearly shows the configuration obtained by the QAOA algorithm presented with two types of nodes represented in two colors, green $\{0\}$ and blue $\{1\}$.

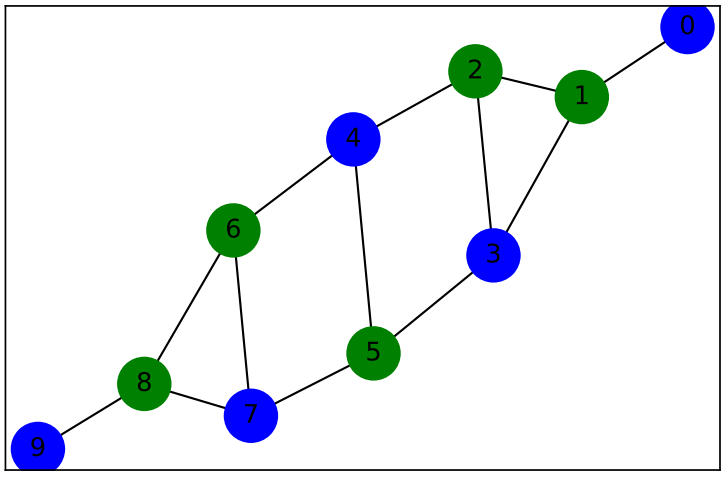


Figure 7. Value stream network clustering with QAOA.

4. Discussion, Future Lines of Research, and Limitations

The analysis of the solution presented in Table 1 shows to what end the quality of the solution found improves the previous one, which justifies the spent effort in considering smarter quantum circuits for the operation of the QAOA algorithm since there is not exist a universal strategy that works across a broad range of optimization problems. Based on

the proposal made, the benefit of the algorithm proposed is evident against other existing algorithms, at least in the case of $p = 1$. As a consequence, value–stream network design challenges can be better understood with the aid of this quantum optimization algorithm. More research is needed to analyze the evolution of potential benefits when the number of transformation blocks grows up. Indeed, resources and performance figures are also needed to get the whole perspective.

The network topology can be modified, however, if classes for the number of links per node are extended, then a new formulation for the cost function introduced in Equation (9) is required.

Future research lines will involve the implementation of this new proposal for quantum circuits in physical quantum computers, to analyze both the performance and the stability against noise, not only for $p = 1$ but also when the transformation blocks are increased. Moreover, since two-qubit gates (e.g., CNOT gates) are significantly more erroneous than single qubit gates, the proposal of smarter circuits with reduced number of two-qubit gates, such as the optimization proposed by [22], is an area of interest.

Because of the problem formulation, the network was defined and just the link of nodes with different managerial classes was the goal. However, it could be possible to reverse the problem and start from the node type distribution and look to connect those nodes with a number of edges optimizing the imputation rules between them.

Some limitations can be found regarding the applicability to real cases, because the existence of extra constraints applicable to nodes or edges. Therefore, additional aspects related to penalty terms when formulation of the $C(x)$ function could be a potential workaround. Following this line, another relevant research area is to extend the current formulation for the Max–Cut problem to the Max–k–Cut one, in line with the recent analysis provided by [26]. Although we have obtained satisfactory results with $p = 1$, we can expect a better approximation for a larger number of qubits if we increase the p parameter. This would, however, entail additional relative difficulties in factoring the Hamiltonian in the adiabatic hypothesis that may be problematic in practice.

Author Contributions: Conceptualization, J.V.-D.; methodology, J.V.-D.; software, J.V.-D.; validation, J.V.-D. and A.G.-M.; formal analysis, J.V.-D. and A.G.-M.; investigation, J.V.-D.; resources, J.B.O.-M.; data curation, J.V.-D.; writing—original draft preparation, J.V.-D.; writing—review and editing, J.V.-D.; visualization, J.V.-D.; supervision, J.B.O.-M. and A.G.-M.; project administration, J.V.-D.; funding acquisition, J.B.O.-M. All authors have read and agreed to the published version of the manuscript.

Funding: The authors thanks the Grant RTI2018-094614-B-I00 (SMASHING) into the "Programa Estatal de I+D+i Orientada a los Retos de la Sociedad" funded by MCIN/AEI/10.13039/501100011033.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The code can be found at https://shorturl.at/hnpC3 (accessed on 1 December 2021) [23].

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

I4.0	Industry 4.0
IIoT	Industrial Internet of Things
MNO	Mobile Network Operator
NB–IoT	Narrowband Internet of Things
QAOA	Quantum Approximate Optimization Approach

References

1. Lee, J.; Bagheri, B.; Kao, H.A. A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems. *Manuf. Lett.* **2015**, *3*, 18–23. [CrossRef]
2. Song, H.; Rawat, D.; Jeschke, S.; Brecher, C. (Eds.) *Cyber-Physical Systems*; Academic Press: Cambridge, MA, USA, 2017.
3. Schuh, G.; Zeller, V.; Stroh, M.F.; Harder, P. Finding the Right Way Towards a CPS—A Methodology for Individually Selecting Development Processes for Cyber-Physical Systems. In *Collaborative Networks and Digital Transformation*; Camarinha-Matos, L.M., Afsarmanesh, H., Antonelli, D., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 81–90.
4. Lodgaard, E.; Dransfeld, S. Organizational aspects for successful integration of human-machine interaction in the Industry 4.0 era. *Procedia Cirp* **2020**, *88*, 218–222. [CrossRef]
5. Nardo, M.; Forino, D.; Murino, T. The evolution of man–machine interaction: The role of human in Industry 4.0 paradigm. *Prod. Manuf. Res.* **2020**, *8*, 20–34. [CrossRef]
6. Ardanza, A.; Moreno, A.; Segura, Á.; de la Cruz, M.; Aguinaga, D. Sustainable and flexible industrial human machine interfaces to support adaptable applications in the Industry 4.0 paradigm. *Int. J. Prod. Res.* **2019**, *57*, 4045–4059. [CrossRef]
7. Grangel-González, I.; Halilaj, L.; Coskun, G.; Auer, S.; Collarana, D.; Hoffmeister, M. Towards a Semantic Administrative Shell for Industry 4.0 Components. In Proceedings of the 2016 IEEE Tenth International Conference on Semantic Computing (ICSC), Laguna Hills, CA, USA, 4–6 February 2016; pp. 230–237. [CrossRef]
8. Batty, M. Digital twins. *Environ. Plan. B Urban Anal. City Sci.* **2018**, *45*, 817–820. [CrossRef]
9. El Saddik, A. Digital Twins: The Convergence of Multimedia Technologies. *IEEE Multimedia* **2018**, *25*, 87–92. MMUL.2018.023121167. [CrossRef]
10. Qi, Q.; Tao, F. Digital Twin and Big Data Towards Smart Manufacturing and Industry 4.0: 360 Degree Comparison. *IEEE Access* **2018**, *6*, 3585–3593. [CrossRef]
11. Ma, H.; Govoni, M.; Galli, G. Quantum simulations of materials on near-term quantum computers. *Npj Comput. Mater.* **2020**, *6*, 1–8. [CrossRef]
12. Villalba-Diez, J.; Zheng, X. Quantum Strategic Organizational Design: Alignment in Industry 4.0 Complex-Networked Cyber-Physical Lean Management Systems. *Sensors* **2020**, *20*, 5856. [CrossRef] [PubMed]
13. Bortolini, M.; Ferrari, E.; Gamberi, M.; Pilati, F.; Faccio, M. Assembly system design in the Industry 4.0 era: A general framework. *IFAC-PapersOnLine* **2017**, *50*, 5700–5705. [CrossRef]
14. Cohen, Y.; Naseraldin, H.; Chaudhuri, A.; Pilati, F. Assembly systems in Industry 4.0 era: A road map to understand Assembly 4.0. *Int. J. Adv. Manuf. Technol.* **2019**, *105*, 4037–4054. [CrossRef]
15. Deza, M.; Laurent, M. *Geometry of Cuts and Metrics*, 1st ed.; Algorithms and Combinatorics; Springer: Berlin/Heidelberg, Germany, 1997; Volume 15.
16. Benlic, U.; Hao, J.K. Breakout Local Search for the Max-Cutproblem. *Eng. Appl. Artif. Intell.* **2013**, *26*, 1162–1173. [CrossRef]
17. Vesselinova, N.; Steinert, R.; Perez-Ramirez, D.F.; Boman, M. Learning combinatorial optimization on graphs: A survey with applications to networking. *IEEE Access* **2020**, *8*, 120388–120416. [CrossRef]
18. Beloborodov, D.; Ulanov, A.E.; Foerster, J.N.; Whiteson, S.; Lvovsky, A. Reinforcement learning enhanced quantum-inspired algorithm for combinatorial optimization. *Mach. Learn. Sci. Technol.* **2020**, *2*, 025009. [CrossRef]
19. Gu, S.; Yang, Y. A Deep Learning Algorithm for the Max-Cut Problem Based on Pointer Network Structure with Supervised Learning and Reinforcement Learning Strategies. *Mathematics* **2020**, *8*, 298. [CrossRef]
20. Farhi, E.; Goldstone, J.; Gutmann, S.; Neven, H. Quantum Algorithms for Fixed Qubit Architectures. *arXiv* **2017**, arXiv:1703.06199.
21. Farhi, E.; Goldstone, J.; Gutmann, S. A Quantum Approximate Optimization Algorithm. *arXiv* **2014**, arXiv:1411.4028.
22. Guerreschi, G.G.; Matsuura, A.Y. QAOA for Max-Cut requires hundreds of qubits for quantum speed-up. *Sci. Rep.* **2019**, *9*, 6903. [CrossRef] [PubMed]
23. Villalba-Diez, J. Value Stream Network Quantum Approximate Optimization Algorithm. 2021. Available online: https://shorturl.at/hnpC3 (accessed on 1 December 2021).
24. Wille, R.; Meter, R.; Naveh, Y. IBM's Qiskit Tool Chain: Working with and Developing for Real Quantum Computers. In Proceedings of the 2019 Design, Automation Test in Europe Conference Exhibition (DATE), Florence, Italy, 25–29 March 2019; pp. 1234–1240. [CrossRef]
25. Jekel, C.; Venter, G.; Venter, M.; Stander, N.; Haftka, R. Similarity measures for identifying material parameters from hysteresis loops using inverse analysis. *Int. J. Mater. Form.* **2019**, *12*, 355–378. [CrossRef]
26. Fuchs, F.G.; Kolden, H.Ø.; Aase, N.H.; Sartor, G. Efficient Encoding of the Weighted MAX-k-CUT on a Quantum Computer Using QAOA. *SN Comput. Sci.* **2021**, *2*, 1–14. [CrossRef]

sensors

MDPI

Article

Data Handling in Industry 4.0: Interoperability Based on Distributed Ledger Technology

Shengjing Sun [1], Xiaochen Zheng [1,2] and Javier Villalba-Díez [1,3,4] and Joaquín Ordieres-Meré [1,*]

1 Escuela Técnica Superior de Ingenieros Industriales (ETSII), Universidad Politécnica de Madrid, José Gutiérrez Abascal 2, 28006 Madrid, Spain; shengjing.sun@alumnos.upm.es (S.S.); xiaochen.zheng@epfl.ch (X.Z.); javier.villalba-diez@hs-heilbronn.de (J.V.-D.)

2 ICT for Sustainable Manufacturing, SCI-STI-DK, École polytechnique fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland

3 Fakultaet fuer Management und Vertrieb, Campus Schwäbisch-Hall, Hochschule Heilbronn, 74081 Heilbronn, Germany

4 Escuela Técnica Superior de Ingenieros Informáticos (ETSIInf), Universidad Politécnica de Madrid, Calle de los Ciruelos, Boadilla del Monte, 28660 Madrid, Spain

* Correspondence: j.ordieres@upm.es

Received: 23 April 2020; Accepted: 24 May 2020; Published: 27 May 2020

Abstract: Information-intensive transformation is vital to realize the Industry 4.0 paradigm, where processes, systems, and people are in a connected environment. Current factories must combine different sources of knowledge with different technological layers. Taking into account data interconnection and information transparency, it is necessary to enhance the existing frameworks. This paper proposes an extension to an existing framework, which enables access to knowledge about the different data sources available, including data from operators. To develop the interoperability principle, a specific proposal to provide a (public and encrypted) data management solution to ensure information transparency is presented, which enables semantic data treatment and provides an appropriate context to allow data fusion. This proposal is designed also considering the Privacy by Design option. As a proof of application case, an implementation was carried out regarding the logistics of the delivery of industrial components in the construction sector, where different stakeholders may benefit from shared knowledge under the proposed architecture.

Keywords: industry 4.0; reference architecture model; interoperability; digital twin; distributed ledger technology; GDPR; RAMI 4.0; LASFA

1. Introduction

The recent advances in Information Technology, Internet of Things (IoT) and Cyber-Physical Systems (CPS), among other fields, have enabled digitization and automation of production processes and led to the definition of the fourth industrial revolution, also known as Industry 4.0 (I4.0) [1–3]. In the manufacturing domain, the I4.0 vision has promoted smart manufacturing and smart factories concepts by augmenting all assets with sensor-based connectivity [4]. Intelligent sensors such as positioning tags, safety gloves [5], head-mounted displays (HMDs), and smart glasses [6] have been widely applied in industrial applications. These intelligent sensors generate a large volume of industrial data and it remains a challenging task to collect, store, analyze, and exploit this data in business, including in simulations, virtual reality, digital twins, and so on [7].

Human factors, such as fatigue indicators, have significant effects on product quality and factory productivity in manufacturing activities [8]. It is clear that increasing the integration of data (including wearable information) can help to increase understanding of the performance of a

business [9]. Indeed, when there is interest in increasing the existing information, a higher level of data integration considering both sub-systems (process and operator) is vital to achieve an effective and efficient human cyber–physical symbiosis [10]. Therefore, pervasive data interconnection and data-driven dynamic decision-making are the key points of Industrial Internet of Things (IIoT) infrastructures, in order to deliver a significant increase in business performance.

To ensure all participants involved in I4.0 share a common perspective and understanding, the *Reference Architecture Model for Industry 4.0* (RAMI 4.0) [11] was conceived as a standard architecture to describe the fundamental aspects of I4.0. It helps to leverage the process of transitioning classical manufacturing systems to I4.0. However, this model, as well as some existing frameworks [12,13], insufficient focus on those wearable devices in the integration layer, especially those related to human bio-sensing dimensions; in such a way, the value extended from data fusion from wearable devices is weak, from both functional and business points of view. Such imperfections mainly linked to aspects of privacy, data ownership, data silos and a lack of interoperability.

Recently, a new architecture model based on RAMI 4.0, named LASFA (LAsim Smart FActory), has been proposed [14]. It adopts the hierarchy of the layers from RAMI 4.0 and focuses on communication among the distributed systems in smart factories, which can help manufacturing companies to transform their manufacturing processes and systems towards the Industry 4.0 paradigm. Figure 1 presents an example of the lower level of such a local production system, as well as the data flows among different components. Compared with RAMI 4.0, the LASFA approach provides a simple configuration for smart factory-enabling agents and manufacturing units in different layers. Due to these advantages, the LASFA architecture model was selected for further development in this work.

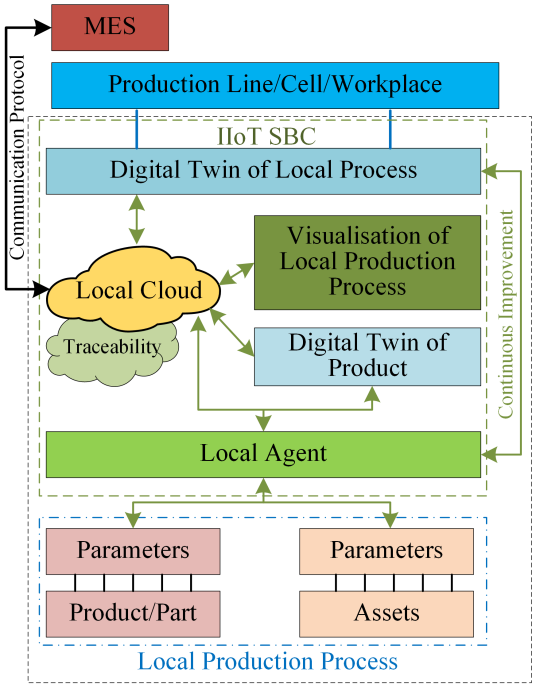


Figure 1. Lower level components of LASFA architecture model [14].

Integration of data requires data fusion capabilities, where data fusion from different systems demands interoperability capacities of such systems, as this accounts for the capability of a system to work with other products or systems; either for implementation or access without significant restrictions [15,16].

Interoperability is related to heterogeneity in information theory. Therefore, classical taxonomies identified Semantic Interoperability, Structural Interoperability, Syntactic Interoperability, and System Interoperability [17]. The focus of this work is to enable the syntactic and structural dimensions of interoperability by fostering terminology transparency and context-sensitive information processing.

This is due to the lack of common data standards (i.e., models, dictionaries, and conventions) which enable the interpretation of data across sub-systems inside an organization, hindering effective data re-usability under a standard format outside of the provided data environment. Therefore, addressing such a gap will help automatic information processes through CPS. The paradigm of cyber–physical equivalence (or digital twin) has emerged [18] to provide a comprehensive physical and functional description of a component, product, or system, in such a way that its information can be useful in any of its life-cycle phases [19].

When a value proposal requires the integration of different stakeholders with different Information Technology (IT) capabilities, such as customers or specific providers, granting access to a public database with immutable but encrypted data is a significant proposal. To this end, different cases can be figured out; on one hand, customers wish to have access not only to the product itself, but also to the production data of the product [20] for different reasons, such as knowledge of the environmental footprint, logistics, and so on.

On the other hand, providers of specific services (e.g., certification bodies) require certified production rates, environmental performance, or continuous improvement, which is another clear example of requiring access to such production-related data streams. Organizations such as regulatory agencies expect holistic and transparent information exchange with peer companies or factories. Furthermore, when asked, 88% of U.S. internet users and 87% of British internet users wish to control the data being collected through smart devices [21]. Encrypted but public databases enable data ownership of producers (especially those related to humans), which can ensure trust and business values resulting from data re-usability by other data consumers.

In modern interconnected societies, where different stakeholders aim to gather different views of different data flows and implement their workflows to make use of them, companies need to consider their value proposals in order to attract current and future stakeholders with relevant data, as well as being able to properly interpret them [22,23]. There is a need for clients, such as middleware platforms, end-users, services, and applications, which efficiently and effectively retrieve IIoT data resources, enabling the integrated and interoperable usage of their data streams.

Due to the different requirements that different stakeholders have in relationship to production-related data encoded into public databases, but also due to the different ways of processing such data, the convenience of integrating semantic meaning into the data itself has arisen. Thus, different stakeholders can easily integrate data processing into their own workflows, and the whole process becomes more resilient to changes in format, the creation of new attributes or entities, and so on. These aspects are linked to data transparency, understood (in the sense of [24–26]) as the degree of visibility and accessibility of information, which includes a component for providing an institutional infrastructure.

Of the secure and non-intensive technologies requiring third-party data access as previously described, the company itself may benefit from providing such standardized production process data flows; particularly when a lean management strategy is implemented. In the case of the usage of the $(CPD)_n A$ method [27–29], the expansion of Key Performance Indicator (KPI) at different levels through time can render traceability useful for the external certification of auditing processes. In addition, workflows can be activated based on their progress, as per the Value Stream Mapping method [30,31]. Therefore, the management dimension is also relevant for this proposal.

Under the described context, this paper aims to promote the principles of interoperability and data transparency in the I4.0 context by enabling the integration of data coming from different IIoT devices, including wearable devices, as well as enabling the integration from different data sources, with the aim of enhancing data fusion among processes from different stakeholders.

The contributions from this paper are the following:

- An enhanced framework providing flexibility to accommodate different data sources and digital twin, as well as an integrated mechanism to disseminate the desired content while limiting exposition of the related IT systems. This creates a common understanding of interoperability.

- Integration of ontologies with the data sets, enabling data fusion techniques and a higher degree of flexibility for data manipulation, enabling automation in machine learning applications.
- Integration of data dissemination in an encrypted way which is immutable and isolated from other IT areas of the company, through the usage of Distributed Ledger Technology (DLT) solutions. Such automation accounts for the transparency principle. Obviously, its contribution must be understood as not being attached to the any particular DLT ledgers, but as a convenient data management approach to show its capabilities.
- Validation of the proposal through Proof of Concept, as deployed in an industrial case involving a real manufacturing company.

Therefore, in Section 2, we review some background principles and contributions related to attributes connected with the value proposal. Section 3 introduces the proposed architecture, including the reference updated framework and its significant components. Section 4 presents the scenario for proof of concept, involving several devices with different semantic meanings in such a way that their data flows are integrated. In Section 5, a discussion of the business dimension is carried out. Finally, the last section is devoted to presenting our conclusions.

2. Literature review

There are several principles that need to be considered when a new or improved framework is presented. These principles are:

- *Data silos*. At present, data produced by IIoT deployment devices and wearable devices are controlled or owned by different device manufacturers. Most of the existing value proposals for the IIoT deployment or wearable devices include one platform to manage the physical devices in an integrated way, as well as collecting the data from them by placing it into a private cloud for processing [32,33]. Benefiting from Radio-frequency identification (RFID) and sensor network technologies, common physical objects can be connected and are able to be monitored and managed by a single system [34]. However, as increasing types of IIoT and wearable devices from different vendors become available, more data silos are generated. The existence of such data silos not only jeopardize potential data-driven services and applications; the limited and restricted features provided also dramatically hinder the learning and inference capability derived from the higher requirements of data integration. Most importantly, all related stakeholders in I4.0 are looking forward to a transparent and shared information platform including production and operator data, which is difficult to realize due to the security constraints when passing between various data silos.
- *Data ownership*. It is an additional challenge for data producers (normally referred to as operators) to reuse (e.g., monetize) their data outside the data provider's environments. The data producer cannot benefit from the data they generate, as the data is locked inside of the internal data environment of the enterprise. They lose their data ownership and the business opportunities stemming from those data are outside of the data owner's control. Under this perspective, when referring to data related to humans, the EU General Data Protection Regulation (GDPR) is as effective tool to add significant trust to this dimension, as users are now able to understand their rights and privacy.
- *Privacy*. When data are related to people, such as devices which are linked to apps by bluetooth, the apps are mostly designed to present the data to users and upload summaries to the cloud manufacturer [35]. However, data integration (especially when related to people) has limitations related to privacy and misuse [36,37]. Cyber security is vulnerable, as wearable device manufacturers have reduced their safety protocols and safety stack layers to enable cheap products, as they have been understood as only serving the end user in a local context. Therefore, there is a clear demand regarding the concept of building and embedding security and privacy controls into connected products, as well as the infrastructure itself. This is one of the implementations of the Privacy by Design (PbD) concept [38,39].

- *Interoperability*. IIoT devices, including wearable devices, are highly heterogeneous in terms of the underlying communication protocols, data formats, and technologies from different vendors. Such heterogeneous infrastructures, devices, and configurations have becomes a strong limitation for data integration and interoperability. By 2021, 25 billion sensor-enabled objects are expected to be connected to the IIoT, as reported by Gartner [40].

Traditional factory-floor control and interconnection data management solutions are mostly based on centralized systems. The German Federal Ministry of Economic Affairs and Energy has provided specifications for the exchange of information within the administration shell. The Open Platform Communications Unified Architecture (OPC UA)[41] is the core communication standard for I4.0-compliant communications [42]. Its adoption was a significant step towards the interconnection of devices, according to RAMI 4.0.

RAMI 4.0 is based on a three-dimensional co-ordinate system consisting of the Layers, Life-Cycle and Value Stream, and Hierarchy Levels dimensions, according to the IEC 62890 and IEC 62264/IEC 51512. It is derived from the Smart Grid Architecture Model (SGAM), which is a key outcome of the EU Mandate M/490 Reference Architecture group for the purposes of communication in networks of renewable energy sources.

In addition, several other proposals aimed at contributing to such common perspectives among stakeholders have been proposed by different institutions. Therefore, The National Institute of Standards and Technology (NIST) has promoted the Smart Manufacturing vision as collaborative manufacturing systems which are able to cope with the challenges of quality, efficiency, and personalization that manufacturing companies are presently facing. China has proposed the National Intelligent Manufacturing System Architecture (IMSA), looking to guide the upgrade of Chinese manufacturing towards intelligent manufacturing. However, wearable devices have been neither well-considered to be connected elements in the Integration Layer of RAMI4.0 [13], nor in the 5C architecture [12].

After RAMI 4.0, the Industrial Internet Reference Architecture (IIRA) for IIoT Systems has been introduced, promoting IIoT architects to use a standard-based architecture framework and adopt a common vocabulary based on ISO 42010. All of these frameworks have been discussed in [42], although such architectures did not specifically include the operator's contributions.

Compared with RAMI 4.0, the LASFA model can not only enable more reliable and simple modeling of smart factories, but also cover the functionality of Enterprise Resource Planning (ERP), Manufacturing Execution System (MES), and Product Lifecycle Management (PLM) sub-systems with new digital twins, which are integrated digital agents supported by artificial intelligence. Approaches such as cloud-based data storage methods have created the centralized approach, shifting the way in which franchisers interact with franchisees [32,33]. However, such systems (e.g., Programmable Logic Controllers (PLCs)) are expected to be too limited in their accessibility to keep up with the growing demand for higher level integration [43] and information transparency for all relevant stakeholders [44,45].

Distributed ledger technologies (DLTs), such as blockchain, provide a convenient replacement for the central administrator in guaranteeing the integrity of a database. Blockchain-based data sharing systems [46] were proposed to address data security and privacy issues and have gained trust from data owners [47,48]. However, blockchain-based protocols still have several drawbacks that hinder their usage in generic data IoT data sharing platforms [49]. The inherent transaction rate limit leads to low throughput, which brings about scalability issues. The concept of a transaction fee for any value transaction is difficult to get rid of in blockchain-based IIoT infrastructures and some concerns have been raised regarding the susceptibility of blockchains to quantum computers [49].

The lack of interconnected data and shared meaning is another limitation for most DLT-based IIoT data platforms [50,51]. Most sensor data sources, including wearable devices, are characterized by a high degree of heterogeneity, and the implementation of their provided interfaces is highly dependent on the underlying device hardware [52]. These data silos with heterogeneous data jeopardize potential data-driven services and applications. Delicato et al. [53] pointed out that discovering and retrieving

data from IoT is a challenge, which is worsened by a lack of standardization of formats in representing resources. However, efforts are still slow-paced to create interoperable platforms to bridge such huge data silos.

To ensure data privacy and to protect data ownership, the data need to be digitally signed and encrypted; even in a distributed platform supported by DLT. Besides blockchain-structured protocols, some recent DLTs have been developed which are specifically designed for the IoT industry. For example, the Directed Acyclic Graph (DAG)-structured IOTA surpassed conventional blockchains, in terms of its scalability, fee-less, and quantum-resistant features [54]. In the IOTA network (named the Tangle), a new transaction needs to approve two previous transactions to attach itself to the network; this transaction will then be validated by some subsequent transactions [49,54]. This mechanism theoretically eliminates throughput caps as the more transactions that are added, the faster the approval speed is.

IIoT requires interconnected data for predictive maintenance, industrial automation, operational efficiency, and better decision-making. Extending the ability to use sensor data integrated from a wide variety of heterogeneous sources facilitates the building of multiple applications and services, which is a vital step toward the success of the IIoT vision and I4.0. However, to effectively address the interconnection of data from heterogeneous entities ranging from simple sensing devices to complex robotic devices, service agents, or human actors in a consistent way, it is necessary to pay attention to the meanings of the individual data streams [55]. The functions of the information layer in LASFA+ also demand consistent integration of higher-quality data, with the provisioning of structured data [56].

Aimed at the data interconnection issue, semantic modeling (e.g., ontology-based techniques) proved successful in dealing with data interoperability and semantic compatibility [57,58] in the manufacturing domain. The Blockchain ontology with dynamic extensibility (BLONDiE) ontology [59] has been developed for Bitcoin and Ethereum but is still in its initial phase. Upper level ontologies, such as the Basic Formal Ontology (BFO) [60] and Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE) [61], work as hub-and-spokes strategies for bringing about interoperability throughout a domain. Ontologies, thus, are in development in various areas. BFO is currently being used in industry IT contexts to gain interoperability in digital manufacturing. Based on BFO, the aim of the Industry Ontology Foundry (IOF) [62] initiative is to create a suite of interoperable and high-quality ontologies covering the domain of industrial (especially manufacturing) engineering [63].

Integration not only happens among the data streams of wearable devices, but higher levels of knowledge can also be derived in industrial contexts when integration between production processes and operator data is considered. This allows more integral access to the infrastructures by means of digital twins [64].

To overcome the aforementioned concerns, this study improves upon an interoperable data handling architecture. The semantic modeling module encompasses heterogeneous resource handling by applying an ontology-based information model, where the ontologies take advantage of well-established and standardized concepts and relationships derived from existing ontologies, in order to semantically represent data resources in IIoT [65].

3. Proposed Architecture

To complement the RAMI 4.0 reference model, and taking advantage of the LASFA architectural model, we integrate all the key elements and their interconnections in a two-dimensional platform. It includes several production layers, as well as a business process. The production layers represent the manufacturing company in terms of production lines, production cells, warehouses, and manual workplaces, creating the local production systems (which are treated as distributed systems). The business layer of the LASFA model includes the company's strategy, business planning, and the monitoring and the delivery of orders [14].

As indicated in Section 1, our main ambition is to extend the proposal of the existing, modern, and valuable framework, LASFA. The general structure of the LASFA architectural model was

presented in [14], and its components and attributes were compared with those of RAMI 4.0 in Section 4 of the above reference.

It is evident that LASFA is much more specific and offers the end-user a simple visualization of the entire architecture of the smart factory, with the definitions of the exact locations and functions of the digital twins and agents, as well as the specific information and data flows between them. The model shows, very clearly, the distribution and autonomy of every single building block in the smart factory, from the product to the management.

In general terms, the LASFA architecture fits with the expectations raised in this work. However, it does have several limitations and, so, this paper adds some extensions to create the improved LASFA model, which we call LASFA+ (see Figure 2).

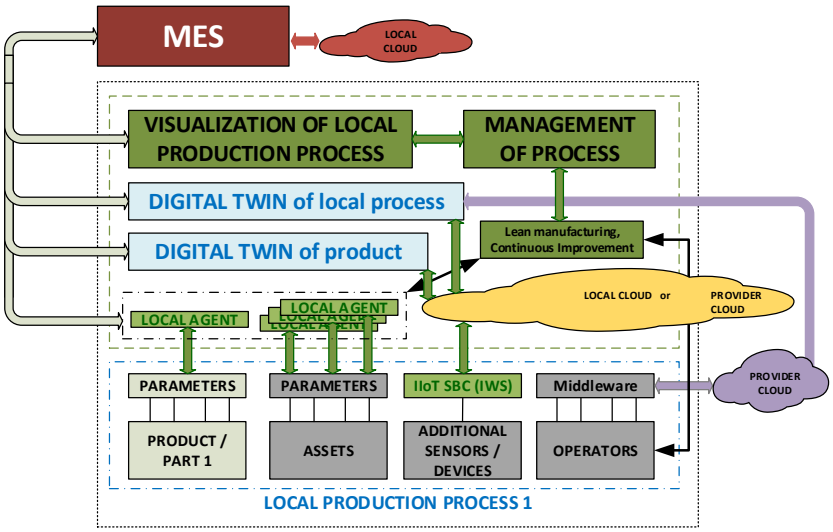


Figure 2. Improved lower layer for the LASFA+ architectural model. MES: Manufacturing Execution System.

The first difference is the wider perspective given to the elements participating in production, where not just traditional assets are considered but also extended ones, providing complementary information to the production process. Some of those services rely either on local or (more frequently) on provider-based clouds outside of the company boundaries. Therefore, the reference model needs to be consistent with such realistic conditions. Obviously, such data flows should be integrated with the traditional ones to build the digital twin.

Another difference is the role of the MES and ERP systems, where the MES used to be closer to production and as a support for visualization of the data of the production process.

Machines, devices, and people are equipped with sensors in order to connect and communicate with each other. This proposed architecture enables the connection of people by low-cost wearable devices which can be conveniently obtained in the market, such as personal health-tracking devices. This architecture makes it possible to not only collect physical features and events but also virtual ones, as a result of the model's output. This becomes useful, when integrated, to develop even more complex models.

The information from various sensors can be further processed in edge devices, such as local servers and single-board computers, where such processing includes data fusion, semantic modeling, and data encryption.

Data fusion is relevant to different aspects, and can include data summaries over certain periods of time from a certain type of sensor. To provide a higher view of data, raw data may need to be statistically summarized. Extracting data patterns or classification of raw data to high-level knowledge could also be included in data fusion. Machine learning (ML), deep learning (DL), or inference mechanisms can be applied for this purpose. A local server is preferable in this context, considering the high computation demands of ML or DL methods. The motion/positioning data collected from

sensors can be segmented and classified into different types of activity, such as running, walking, and so on.

To enable data interoperability, semantic information is annotated in each data message. Existing ontologies in any domain (besides industrial) are able to be reused under this architecture. Ontology-based structured data can be conveniently integrated with other data sets, in order to empower more complex applications (e.g., data-driven AI algorithms). Ontologies aligned with existing ones or newly proposed ontologies are also supportive in this semantic modeling module.

As shown in Figure 3, an application scenario was developed which enables interoperability among different data sources, further providing a secure, distributed database that empowers data access by different stakeholders. The main components of the prototype include a heterogeneous data source, ontology-based interoperability modeling, and incorporation within the DLT module for data dissemination.

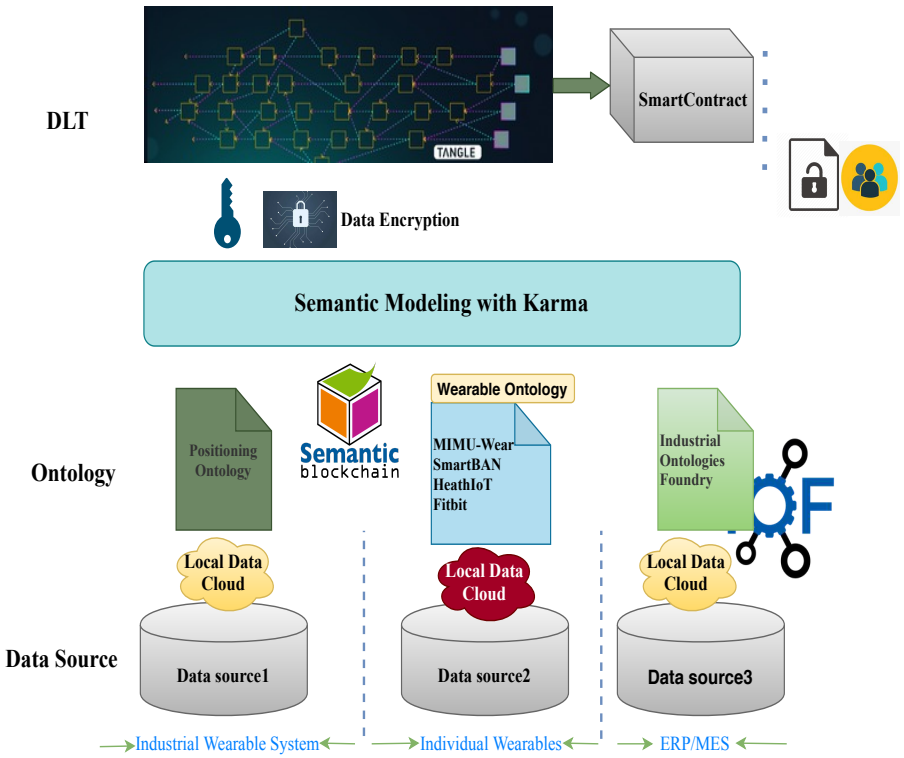


Figure 3. Prototype of data handling in an industrial context. DLT: Distributed Ledger Technology; ERP: Enterprise Resource Planning; MES: Manufacturing Execution System; MIMU: Magnetic and Inertial Measurement Units.

The corresponding ontology schema is functional as an information model for the automatic understanding of data during retrieval procedures. Data on the DLT, such as the IOTA Tangle, can be retrieved by data consumers. Various data transactions can be retrieved by having the address/tag/bundle or integrated based on defined ontology schema. In this way, the dominant paradigm of keeping data hidden from public knowledge by having protected access is broken and the data becomes part of the product being delivered for further usage. Data consumers may define their own ontology fitting their specific requirements, where this data will be seamlessly integrated.

Although data published to a DLT (e.g., IOTA Tangle), is protected in a secure infrastructure, the message within each transaction confronts a data security issue. If the message is not encrypted, once the transaction address is spread in the web, any party who has the tag, address of the receiver, or transaction ID is able to read the contents.

Here, different policies can be adopted, such as:

- *Public.* This is the already-described mechanism, where anyone knowing the address of the receiver has the right to freely access to the content.
- *Restricted.* This is the case where sensitive information needs to be shared among only a limited number of stakeholders. To prevent unauthorized entities from reading the data, the message content is obfuscated by encrypting it before uploading it to the Tangle. DLT access at message level, including the encryption/decryption process, is demonstrated in Figure 3.
- *Private.* In this case, due to the specific content of the data, to prevent unauthorized entities from reading the data and to respect the GDPR enforcement rules, the message content is obfuscated by encrypting it before uploading it to the Tangle by using a private certificate, making it Private by Design (PbD).

For messages which are determined to be *Private*, the framework proposed in this paper can be extended in such a way that data owners are able to share a concrete data set defined by matching some conditions (e.g., a period of time) with specific entities. The proposed solution is that the data owner installs a web-service which is able to handle such authorized data requests. The signature key of the data structure will be a public one, preserved under a smart contract and consumed by a web service which is in charge of providing the requested information. The designed data structure for the request can be provided in JSON format for the REST-type implemented protocol, which includes the following elements,

- The DLT address, representing the interested entity or device;
- The DLT tag, representing the interested data set;
- The public key, for the user requesting the access to the DLT stored data;
- Selection criteria for the required data, such as from [DateTime] to [DateTime]; and
- Proof of worth for accessing the data.

When either people or agents request the web service to access the specific data according to the above structure, when they have the right of access, the web-service will access the DLT repository under the specific criteria provided and verify the ownership of the requested information. This verification will be based on consuming the smart-contract to obtain access to the symmetric encryption key used by the node, thus being able to decrypt it and compare signatures. Then, such data can be aggregated (according to the requested period of time) and be sent to the requester by the web service after encrypting it using the provided public key of the user/agent. The data consumers will receive back the data and handle it by using its private key, in such a way that only they themselves are allowed to consume it.

The whole schema can be seen in Figure 4, where the role of different certificates is shown clearly. In this way, no third parties are aware of the encrypting/decrypting certificate, as the data is made available to them by re-encrypting it with their public asymmetric certificates, which only enables the third party to decrypt its content.

This type of solution can help companies from at least three different perspectives. They can redesign their adopted business models related to the IIoT and its integration into the management dimension of the organization [66,67]. However, it may also become useful to increase the knowledge of the existing processes in the existing business model, or may just help in the update process of such business models themselves.

In the case of public messages, there are no requirements for encryption. Under the restricted approach, the encryption certificate can be shared among the interested stakeholders, in such a way they can collect the information on their own.

Not only due to the transparency provided, but also because such data streams are delivered automatically by the different digital twins, the trust of shareholders and consumers will increase immediately. In addition, such knowledge will enable the management needs of owners, being more consistent with its potential value when shared.

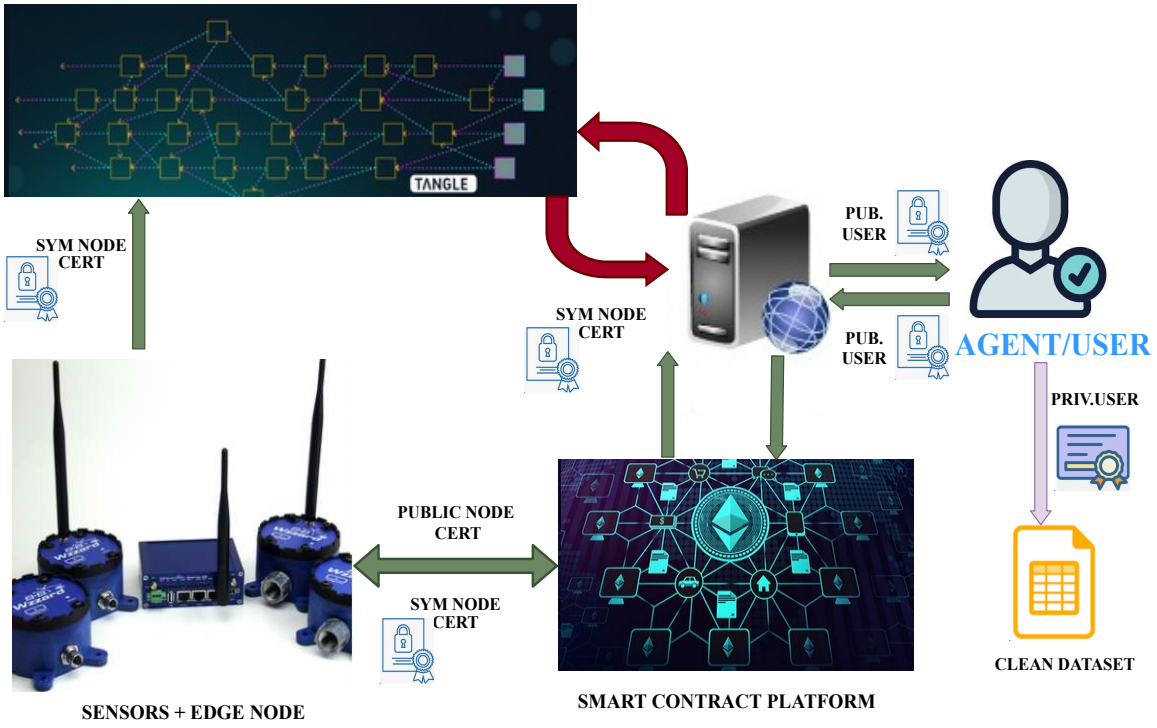


Figure 4. Global schema for Web Services operation including certificate usage.

4. Proof of Concept: An Industrial Scenario

To validate the proposed architecture, a data-handling system based on a DLT (IOTA), and domain ontologies was designed to improve data interoperability, with integrated privacy criteria through message encryption and data sharing under the Industry 4.0 paradigm.

The implemented system was adopted in a practical industrial scenario. The aim was to verify that our proposed architecture has practical potential in real-world applications, leveraging the implemented system to demonstrate its feasibility. The approach, which adopted a positivist view of research, relied on the literature and empirical data coming from the case itself, as well as on the insights of the researcher to build incrementally more powerful theories.

The application case was conducted in a Spanish factory manufacturing steel rebars to reinforce concrete in the construction sector. The general workflow is as shown in Figure 5. Its production process (operation) includes cutting, blending, and welding rebars into different size configurations, packing them into rebar bundles according to customer's order, and loading the rebar bundles into trucks for logistics. For each process, there are different data flows that may affect the production efficiency and effectiveness; the data flows came from data sources such as production line, IIoT divides such as industrial wearable systems (IWSs), and personal wearable devices.

Consider the truck loading process as an example. The requirement is that the right rebar bundles are loaded into the truck in a specific disposal sequence, which needs to be well managed as the bundles are delivered and distributed to different sites. As a lack of specific items could impact the delivery dates, an improper loading disposal sequence may negatively affect the scheduled work in the unloading process. In the present situation, the responsibility for such a decision relies on the crane operators, who take charge of loading the rebar bundles to the truck buffer using a crane machine.

To effectively assess this internal logistics, different data sources should be taken into consideration. Such an analysis requires knowledge about the order and manufacturing sequences, which are handled by the ERP system, MES, or PLM. The logistics information is required, such as where and when the items are loaded into the truck, in order to understand whether the items were loaded and well-placed in a proper storage buffer in the truck.

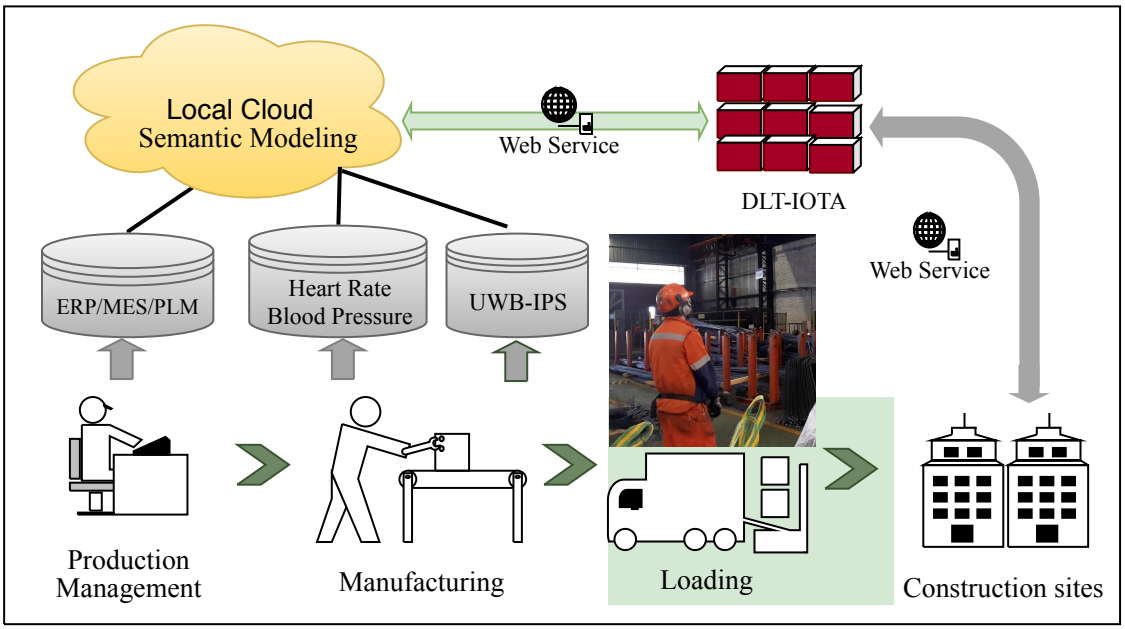


Figure 5. Application scenario of the proof of concept implementation. DLT: Distributed Ledger Technology; ERP: Enterprise Resource Planning; MES: Manufacturing Execution System; PLM: Product Lifecycle Management; UWB-IPS: Ultra Wide Band Indoor Positioning System.

Moreover, more details, such as the crane operator profile, their working conditions, and working status (e.g., health parameters and movement trajectory/speed) are also relevant. Body-related parameters need to be integrated, mainly as the stress of the crane operators may affect the loading process. The movement trajectory/speed need to be considered, as analysis can be made to understand whether the loading process requires excessive physical efforts.

When different stakeholders need to have access to specific process-related information, non-tampered data are required due to certification principles; therefore, an open system is a convenient tool, reducing the IT barriers and which is robust against facility ownership changes.

The system prototype will adapt the proposal presented in Section 3 by adopting specific mappings over the particular characteristics of the company and its processes.

Different devices, systems, and data sources composed the configured prototype, as indicated below:

- Production related information from ERP/MES/PLM system;
- Ultra wide band (UWB) indoor positioning system (IPS) to track crane position and crane operator movements, to better define the location for rebar bundles; and
- Smart band to monitor crane operator's heart rate and blood pressure.

The MES manages sequence planning and the bottom-up data flow on the shop-floor [68]. It provides the current state of the process, identifying the optimal resources allocation by taking consideration of the workstation capacity and organization-level requirements (e.g., manufacturing order). All data generated by manufacturing processes are stored in a single database in the company local cloud, based on a Microsoft SQL server database.

The UWB indoor positioning system (IPS) from the Tracktio™ company (https://tracktio.com/) was used to track the crane hook and crane operators. Rebar bundle locations in different buffer areas of production or positions in the truck for customer delivery were derived from the dynamic behavior of the process. The position of crane operator was tracked as a reference for understanding the movement trajectory and speed. The IPS had its own data repository; the data were acquired from a web service.

A smart band was worn by the crane operator, in order to monitor their stress factors, such as blood pressure and heart rate. Every minute, data was collected and uploaded to a smart phone (as an

example of edge computing); then, it was reprocessed and transmitted to a MongoDB database located in the local cloud.

For data interoperability, all data sources in this industrial scenario were semantically modeled, fostering their linkage to other domain knowledge. To this end, different existing domain ontologies were reused in this study. The data sources were in the three aforementioned sectors: production system, IPS, and individual wearable devices. The existing and shared/published ontologies for each sector were collected and one was selected to model the data source in this study, according to the mapping of data structure. The details of available ontologies for reusability are listed in Table 1.

Table 1. Ontology Reusability Selection.

Ontology List		
IPS	**Wearable**	**Production System**
Positioning Ontology [69,70]	MIMU-Wear Ontology [71]	MES ontology [72–74]
IndoorGML [75]	SmartBAN Ontology [76]	ERP ontology [77,78]
Navigation ontology [70,79,80]	HealthIoT Ontology [81]	PLM ontology [73]
Indoor space ontology [82,83]	Fitbit Ontology [84]	
	Vital Sign Ontology [85]	

ERP: Enterprise Resource Planning; IPS: Indoor Positioning System; MES: Manufacturing Execution System; MIMU: Magnetic and Inertial Measurement Units; PLM: Product Lifecycle Management.

To map the generated data sources in this industrial scenario, three ontologies were chosen for data modeling. The selection criteria were chosen on basis of the degree of matching between the ontological schema and data source structure. After deeper analysis of each ontology's schema, as listed in Table 1, the vital sign ontology [85] was used to model heart rate and blood pressure data from the Smart band; the rebar bundle and crane operator's position were modeled with the positioning ontology [69]; and the MES ontology [72] was applied for the model production system data source.

The mapping tool Karma (https://usc-isi-i2.github.io/karma/) provides a graphical user interface which automates the semantic modeling process. Karma learns to recognize the mapping of data to the chosen ontology classes and proposes a model that can generate JavaScript Object Notation for Linked Data (JSON-LD) for large data sets in a batch mode. JSON-LD is a lightweight Linked Data format, which was designed based on the concept of "context", linking object properties in JSON to concepts in an ontology. The JSON-LD data type was selected as it is lightweight and interoperate at Web-scale, but also provides embedded semantics. The data format is based on JSON and ontological schemas, maintaining a common space of understanding and supporting the evolution of schemes over time without requiring data consumers to change format. The data source, applied ontology schema, and the transformed JSON-LD format (containing semantic annotation) are listed in Table 2.

To distribute information about the process interaction, the IOTA network Tangle is required. For this, access to a node is needed and a public node (https://nodes.thetangle.org:443) was selected for data submission to the IOTA Tangle. The message sent to the IOTA Tangle is encrypted. A python script that implements the RSA (Rivest–Shamir–Adleman) public-key cryptosystem was used for encryption and validation of messages. PyOTA (https://github.com/iotaledger/iota.py), an IOTA python API library, was used to implement data sending and retrieval to and from the IOTA Tangle. The source code can be found at [86], which supports Python and NodeJS.

A data message describes relevant information combined from different data sources, in which the details about loading a rebar bundle are encrypted and delivered to the IOTA Tangle by the IIoT system, after integrating the meaningful information. In this application, every bundle loaded into a truck and delivered to a construction site was encoded as a message. As the unit is the truck, every truck has a specific encryption certificate, which could be shared with interested stakeholders (e.g., construction workers, transport agencies, insurance companies, or end consumers).

Table 2. Data semantic transformation based on ontology and Karma.

Semantic Modeling			
Data Source	**Data Storage**	**Ontology**	**JSON-LD**

MES (ISA-95)	Microsoft SQL database	MES ontology [72]	{ "@context": {
			"mes":"http://mesontology/schema",
			"gr": "http://purl.org/goodrelations/v1#",
			"pto": "http://www.productontology.org/id/",
			"xsd": "http://www.w3.org/2001/XMLSchema#",
			"gr:seriaNumber": {
			"@type": "xsd:int" },
			"gr:description": {
			"@type": "xsd: string" },
			"gr:amountOfThisGood": {
			"@type": "xsd:int" },
			"gr:eligibleRegions": {
			"@type": "xsd: string" },
			"gr:weight": {
			"@type": "xsd:float" }
			"mes:planificationId": "2270077239",
			"gr:seriaNumber": "AAAEEE3452XX",
			"gr:description": "bundle of rebars",
			"gr:amountOfThisGood": "20",
			"gr:eligibleRegions": "DE(Germany)",
			"gr:weight": "300",
			"gr:includes": {
			"@type": [
			"gr:Individual",
			"pto:Rebar"]} }

Smart band	MongoDB	Vital Sign Ontology [85]	{ "@context": {
			"vso": "https://bioportal.bioontology.org/ontologies/VSO",
			"xsd": "http://www.w3.org/2001/XMLSchema#",
			"vso: pulserate": {
			"@type": "xsd:int" },
			"vso: systolicbloodpressure": {
			"@type": "xsd:int" },
			"sumo": "http://www.adampease.org/OP/SUMO.owl",
			"sumo:timepoint": {
			"@type": "xsd:dateTime" },
			"vso:pulserate" : "79",
			"vso:systolicbloodpressure": "111",
			"sumo:timepoint": "2020-04-08T11:20:00Z" }

IPS (tracktio)	CSV file	Positioning Ontology [69]	{ "@context": {
			"positionpoint": "http://positioningontology/schema",
			"xsd": "http://www.w3.org/2001/XMLSchema#",
			"axis": "http://data.ign.fr/def/ignf#CoordinateSystemAxis",
			"axis: listofaxes": {
			"@type": "xsd:list" },
			"sumo": "http://www.adampease.org/OP/SUMO.owl",
			"sumo:timepoint": {
			"@type": "xsd:dateTime" },
			"positionpoint": {
			"axis: listofaxes": "7.299998 1.140002 1.499998" },
			"sumo:timepoint": "2020-04-09T10:00:00Z" } }

IPS: Indoor Positioning System; JSON-LD: JavaScript Object Notation for Linked Data; MES: Manufacturing Execution System.

As shown in Figure 6, the loading truck activity is labeled with the tag *[LOADTRUCKTESTCASE]* in IOTA Tangle. Related transactions could be looked up and fetched by tags or bundles. After data retrieval, a data message should be validated and decrypted by stakeholders, where the semantic meaning of the data are encoded in the message itself. This design fosters data reusability by learning processes and growth of models, as they can select the appropriate meaning of data sets or ask for automatic preprocessing before accomplishing further transformations.

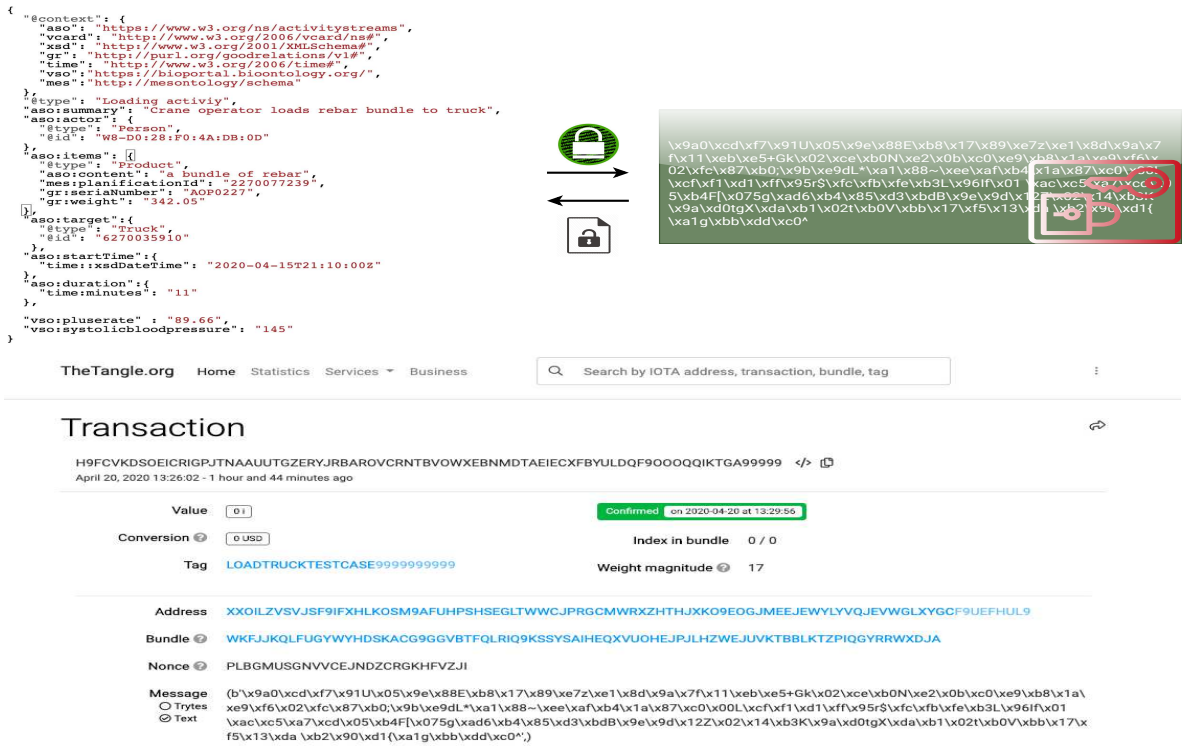


Figure 6. Data handling process supported by IOTA Tangle; loading activity as example.

5. Discussion

The needs of the near-future dimensions of I4.0 require an increasing level of integration of data from different sources, including data from not only workers, but also from different providers and sources. This enforces PbD requirements, due to both company security rules and regulatory requirements. A major aspect of the GDPR are the so-called legal grounds for lawfully processing personal data, one of which is consent. In several IoT applications where consent is used, it may even need to be explicit consent.

There still remain, after the GDPR, some problems with personal data, as follows:

- Data consistency and the use of different sources of information and/or different time periods or geographical positions;
- Data storage: although properly scrambled, masked, or blurred, the related persons probably are not aware such pieces of information do exist related to them; and
- Problems related to EU citizens when requesting products or services outside of the EU.

The harder problem is the second one, mainly due to the absence of accountability: as a user never knows who is gathering information about them, they cannot ask for access, erasure, or modification. Thus, there are two types of accountability:

- Accountability regarding users; but also,
- Accountability within the organization.

In our validation case, the delivery process to the customer has limited information from the process itself, as truck loading is mainly a human-driven activity. Therefore, it risks exposure to a significant time variability and existence of errors in missed or items wrongly included inside the truck. From a business perspective, adding transparency by maximization in the amount of automation involved is worthwhile for stakeholders. Furthermore, current private and centralized data management approaches limit the high-level integration of multi-modal data sources, as well as data ownership and re-usability (e.g., monetizing health data generated from data producers).

The deployed architecture, as proof of concept, was able to provide additional information, integrated in a convenient way in order to help to understand the type of items frequently creating problems, configurations of truck layout consistent with specific delivery sequences, the movements used to load particular item shapes, and so on.

There exist potential contributions in the direction of empowering operators of I4.0 [87], including virtual and augmented reality applications to facilitate the operator's activities and to reduce unwanted mistakes or damage by giving appropriate instructions for complex tasks or working environments. They are also useful for job role allocation and/or for determining training needs, especially in dynamic work environments with changing conditions [88]. Therefore, the impact of the physiological conditions of workers can be considered to be relevant, especially the influences of tiredness or stress.

Obviously, to have significance in such an analysis, large periods of data collection and close integration are required. Indeed, path dependency [89] does exist; however, such analysis exceeds the scope of the present research.

Another critical enabling technology for data interoperability adopted in this study is semantic modelling and ontology engineering. Ontologies provide standardized definitions for different data sources and make possible of reasoning and autonomous decision-making, by formalizing the structure of the knowledge. Existing ontologies, i.e., vital sign ontology, positioning ontology, and the MES ontology, are reused in this study to accelerate the development. To obtain a higher level of interoperability, such ontologies can be further exploited and adapted to keep align with widely recognized upper-level ontology such as BFO [60] and to be attached to industrial ontology initiatives such as IOF [62]. In this way, the developed ontologies will be able to interact with other domain ontologies thus to achieve wider range of interoperability in I4.0 context.

As one of the fundamental enabling technologies of I4.0, the digital twin concept is briefly discussed in this study but not explored with details as it is not the main focus of this paper. Nevertheless, the proposed data integration and handling approaches are important basis for implementation of digital twins. For example, in some existing digital twin studies [90,91], similar semantic modelling techniques have been adopted to improve data interoperability data heterogeneous issues. Aiming to identify the dynamics of virtual model evolution and enhancing the decision-making capabilities, the next generation of digital twin, i.e., cognitive twin concept, has been proposed in a recent study [92], which integrated augmented semantic capabilities with conventional digital twins. The work presented in this paper will have great potential in these digital twin and cognitive twin domains.

6. Conclusions

In this paper, our main contributions were to extend and complement the I4.0 reference model LASFA, with reference to some more general ones (e.g., RAMI4.0), by adding additional data flows to complete the production process description and including different provider add-ons for digital twin construction. The proposed update enables higher level integration and more efficient process description. By considering the PbD design for exchange of information over public DLT systems, it also accounts for information transparency and data ownership for related stakeholders.

The adopted configuration, which enables the semantic enrichment of data, also make it possible to implement rules enforcing interaction mechanisms and to derive new properties related to the

employed ontologies. Although such a direction was not implemented in the proof of concept carried out, it may be useful for stakeholders accessing the data.

The presented proposal benefits from the integration of semantically annotated data from different sources, including human health-related wearable device data, in an industrial context. When information related to people has been included, a signature layer is used to provide a PbD solution, in full accordance with the GDPR regulations and to enable positive-Sum [39].

The core idea of the proposed architecture is to enable the sharing of different obfuscated data streams over a public immutable network-oriented database with quick-answer capabilities and good scalability characteristics. The proposed architecture enables extensions, in order to facilitate higher levels of control and empowerment of data owners by integrating an explicit delegation of access to specific data sets and time windows through the use of web services.

Considering a simple application case, the benefits from a business perspective appear evident, with the integration of different dimensions not previously considered. There also exists additional value with the possibility of further performance analysis of procedures, when the proposed architecture is combined with artificial intelligence (AI) techniques. The theoretical implications of this research allow for a convenient framework providing a high level of control on the data produced by each agent, while also providing a platform to share data coherently with different business models. It also provides a flexible way to interconnect semantically powered data streams, which is a significant contribution for data driven ML/AI applications.

This paper also makes a point from a practical application perspective; in particular, by providing an interesting implementation of data handling strategies aligned with Industry 4.0 principles enabling the integration of data coming from workers. Such tools can have a significant impact on advanced lean manufacturing and lean management implementations, as they go a step beyond classical digitization approaches. This could be the case for the standardization of formal communications under Lean Management $(CPD)_n A$.

There are several aspects that provide us with further research paths, as we require long data collection approaches and the definition of consistent key performance indicators for individual applications. Furthermore, the definition of related business models for micro (smart contract-based) data monetization and traceability applications also provides avenues for further research.

A clear limitation of the existing research is the adoption of a single and specific DLT technology; namely IOTA. The proposal is not conceptually dependent on IOTA; it simply served as a driver to implement our architecture in the proof of concept. However, some other technologies, such as Obyte, may be considered to be potential substitutes.

Another limitation is the required digital infrastructure needed in the companies to successfully implement the digital twin concepts, which is capable of managing requirement of both internal and external stakeholders.

Future research steps will be to deeply analyze the meaning of the integrated data built, in terms of value for business, looking to identify behavioral patterns or to create forecasting models, and its ability to better adjust operation times and, most importantly, timely delivery.

Another interesting perspective to be considered is that wearable devices, such as smart watches [88] and smart wristbands, can be useful for estimating the operator's well-being (i.e., cognitive, psychological, and physical needs) in Human Cyber–Physical Systems (H-CPS) [93]. Wearable monitoring provides opportunities for health diagnostics and operator well-being in industrial environments, as was also identified in [94,95].

Author Contributions: S.S. was responsible for the development of the case study. S.S. and J.O.-M. proposed the methodology and wrote the original draft. J.O.-M. contributed with S.S. and J.V.-D. to provide the industrial dimension context for the application. The IOTA selection was the responsibility of J.O.-M. and X.Z. The review & editing was conducted by X.Z. and J.V.-D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by the European Commission through the RFCS program, grant number with ID 793505.

Acknowledgments: The authors thank China Scholarship Council (CSC) (No.201206730038) for the support they provided for this work. Also the authors want to thank the EU RFCS research program for the support of this research through the grant with ID 793505 (WISEST project). Indeed, JOM wants to acknowledge the Spanish Agencia Estatal de Investigacion, through the research project with code RTI2018-094614-B-I00 into the "Programa Estatal de I+D+i Orientada a los Retos de la Sociedad".

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
DAG	Directed Acyclic Graph
DL	Deep Learning
DLT	Distributed Ledger Technology
ERP	Enterprise Resource Planning
GDPR	General Data Protection Regulation
I4.0	Industry 4.0 paradigm
IIoT	Industrial Internet of Things
IoT	Internet of Things
IPS	Indoor Positioning System
IT	Information Technology
IWS	Industrial Wearable System
JSON	JavaScript Object Notation
JSON-LD	JavaScript Object Notation for Linked Data
KPI	Key Performance Indicator
LASFA	Lasim Smart Factory reference model
MES	Manufacturing Execution System
ML	Machine Learning Techniques
OPC UA	Open Platform Communications Unified Architecture
PbD	Privacy by Design
PLM	Product Lifecycle Management
UWB	Ultra Wide Band

References

1. Piccarozzi, M.; Aquilani, B.; Gatti, C. Industry 4.0 in management studies: A systematic literature review. *Sustainability* **2018**, *10*, 3821. [CrossRef]
2. Maresova, P.; Soukal, I.; Svobodova, L.; Hedvicakova, M.; Javanmardi, E.; Selamat, A.; Krejcar, O. Consequences of industry 4.0 in business and economics. *Economies* **2018**, *6*, 46. [CrossRef]
3. Villalba-Díez, J.; Molina, M.; Ordieres-Meré, J.; Sun, S.; Schmidt, D.; Wellbrock, W. Geometric Deep Lean Learning: Deep Learning in Industry 4.0 Cyber–Physical Complex Networks. *Sensors* **2020**, *20*. [CrossRef]
4. Grangel-González, I.; Halilaj, L.; Coskun, G.; Auer, S.; Collarana, D.; Hofmeister, M. Towards a Semantic Administrative Shell for Industry 4.0 Components. In Proceedings of the 2016 IEEE Tenth International Conference on Semantic Computing (ICSC), Laguna Hills, CA, USA, 4–6 February 2016; pp. 230–237.
5. Scheuermann, C.; Heinz, F.; Bruegge, B.; Verclas, S. Real-Time Support During a Logistic Process Using Smart Gloves. In Proceedings of the Smart SysTech 2017; European Conference on Smart Objects, Systems and Technologies, Munich, Germany, 20–21 June 2017; pp. 1–8.
6. Hao, Y.; Helo, P. The role of wearable devices in meeting the needs of cloud manufacturing: A case study. *Robot. Cim.-Int. Manuf.* **2017**, *45*, 168–179. [CrossRef]
7. Vrchota, J.; Pech, M. Readiness of Enterprises in Czech Republic to Implement Industry 4.0: Index of Industry 4.0. *Appl. Sci.* **2019**, *9*, 5405. [CrossRef]
8. Neumann, W.; Kolus, A.; Wells, R. Human Factors in Production System Design and Quality Performance - A Systematic Review. *IFAC-PapersOnLine* **2016**, *49*, 1721–1724. [CrossRef]

9. Sun, S.; Zheng, X.; Gong, B.; Paredes, J.G.; Ordieres-Meré, J. Healthy Operator 4.0: A Human Cyber–Physical System Architecture for Smart Workplaces. *Sensors* **2020**, *20*, 2011. [CrossRef]
10. Kong, X.T.R.; Yang, X.; Huang, G.Q.; Luo, H. The impact of industrial wearable system on industry 4.0. In Proceedings of the 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC), Zhuhai, China, 27–29 March 2018; pp. 1–6. [CrossRef]
11. Zvei, E. Plattform Industrie 4.0: The reference Architecture Model of Industrie 4.0 (2015). Available online: https://www.zvei.org/en/subjects/industrie-4-0/the-reference-architectural-model-rami-40-and-the-industrie-40-component/ (accessed on 25 April 2020).
12. Lee, J.; Bagheri, B.; Kao, H.A. A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems. *Manuf. Lett.* **2015**, *3*, 18–23. [CrossRef]
13. Alcácer, V.; Cruz-Machado, V. Scanning the Industry 4.0: A Literature Review on Technologies for Manufacturing Systems. *Eng. Sci. Technol. Int. J.* **2019**, *22*, 899–919. [CrossRef]
14. Resman, M.; Pipan, M.; Simic, M.; Herakovic, N. A new architecture model for smart manufacturing: A performance analysis and comparison with the RAMI 4.0 reference model. *Adv. Prod. Eng. Manag.* **2019**, *14*, 153–165. [CrossRef]
15. Wiemann, S.; Bernard, L. Spatial data fusion in spatial data infrastructures using linked data. *Int. J. Geogr. Inf. Sci* **2016**, *30*, 613–636. [CrossRef]
16. Gagnon, M. Ontology-based integration of data sources. In Proceedings of the 10th IEEE International Conference on Information Fusion, Quebec, QC, Canada, 9–12 July 2007; pp. 1–8.
17. Noura, M.; Atiquzzaman, M.; Gaedke, M. Interoperability in Internet of Things Infrastructure: Classification, Challenges, and Future Work. In *IoT as a Service*; Lin, Y.B., Deng, D.J., You, I., Lin, C.C., Eds.; Springer International Publishing: Cham, Switzerland, 2018; pp. 11–18.
18. Glaessgen, E.; Stargel, D. The digital twin paradigm for future NASA and US Air Force vehicles. In Proceedings of the 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 20th AIAA/ASME/AHS Adaptive Structures Conference 14th AIAA, Honolulu, HI, USA, 23–26 April 2012; p. 1818.
19. Boschert, S.; Rosen, R. Digital twin—The simulation aspect. In *Mechatronic Futures*; Hehenberger, P., Bradley, D.; Eds.; Springer: Cham, Switzerland, 2016; pp. 59–74.
20. Kagermann, H. Change through digitization—Value creation in the age of Industry 4.0. In *Management of Permanent Change*; Springer: Cham, Switzerland, 2015; pp. 23–45.
21. Groopman, J. Circles of Trust for Smart Devices for US citizens. Source: Privacy & The Internet of Things: The Importance of Transparency in Accounting for What We Can't See. Available online: https://www.trustarc.com/blog/tag/iot-summit/ (accessed on 25 April 2020).
22. Frysak, J.; Krenn, F.; Kaar, C.; Stary, C. Decision-Making Support for View-Oriented I4.0 System Architecture Design. In Proceedings of the 2018 IEEE 20th Conference on Business Informatics (CBI), Vienna, Austria, 11–14 July 2018; Volume 1, pp. 186–195.
23. Bakhshandeh, M.; Antunes, G.; Mayer, R.; Borbinha, J.; Caetano, A. A Modular Ontology for the Enterprise Architecture Domain. In Proceedings of the 17th IEEE International Enterprise Distributed Object Computing Conference Workshops, Vancouver, BC, Canada, 9–13 September 2013; pp. 5–12.
24. Zhu, K. Information transparency in electronic marketplaces: Why data transparency may hinder the adoption of B2B exchanges. *Electron. Mark.* **2002**, *12*, 92–99. [CrossRef]
25. Flatt, H.; Schriegel, S.; Jasperneite, J.; Trsek, H.; Adamczyk, H. Analysis of the Cyber-Security of industry 4.0 technologies based on RAMI 4.0 and identification of requirements. In Proceedings of the 2016 IEEE 21st International Conference on Emerging Technologies and Factory Automation (ETFA), Berlin, Germany, 6–9 September 2016; pp. 1–4.
26. Müller, J.M.; Kiel, D.; Voigt, K.I. What Drives the Implementation of Industry 4.0? The Role of Opportunities and Challenges in the Context of Sustainability. *Sustainability* **2018**, *10*. [CrossRef]
27. Villalba-Diez, J.; Ordieres-Meré, J. Improving manufacturing performance by standardization of interprocess communication. *IEEE Trans. Eng. Manag.* **2015**, *62*, 351–360. [CrossRef]
28. Villalba-Diez, J.; Ordieres-Meré, J.; Nuber, G. The HOSHIN KANRI TREE. Cross-Plant Lean Shopfloor Management. *Procedia CIRP* **2015**, *32*, 150–155. [CrossRef]
29. Villalba-Diez, J.; Ordieres-Mere, J.; Chudzick, H.; Lopez-Rojo, P. NEMAWASHI: Attaining Value Stream alignment within Complex Organizational Networks. *Procedia CIRP* **2015**, *37*, 134–139. [CrossRef]

30. Villalba-Diez, J. *The Lean Brain Theory. Complex Networked Lean Strategic Organizational Design*; CRC Press. Taylor and Francis Group LLC: Boca Raton, FL, USA, 2017.
31. Villalba-Diez, J. *The HOSHIN KANRI FOREST. Lean Strategic Organizational Design*, 1st ed.; CRC Press, Taylor and Francis Group LLC: Boca Raton, FL, USA, 2017.
32. Lee, J.; Ardakani, H.D.; Yang, S.; Bagheri, B. Industrial Big Data Analytics and Cyber-physical Systems for Future Maintenance & Service Innovation. *Procedia CIRP* **2015**, *38*, 3–7. [CrossRef]
33. Liu, Y.; Xu, X. Industry 4.0 and Cloud Manufacturing: A Comparative Analysis. *J. Manuf. Sci. Eng.* **2016**, *139*. [CrossRef]
34. Jiang, L.; Da Xu, L.; Cai, H.; Jiang, Z.; Bu, F.; Xu, B. An IoT-oriented data storage framework in cloud computing platform. *IEEE Trans. Ind. Inf.* **2014**, *10*, 1443–1451. [CrossRef]
35. Stergiou, C.; Psannis, K.E.; Kim, B.G.; Gupta, B. Secure integration of IoT and cloud computing. *Future Gener. Comp. Syst.* **2018**, *78*, 964–975. [CrossRef]
36. D'Antrassi, P.; Prenassi, M.; Rossi, L.; Ferrucci, R.; Barbieri, S.; Priori, A.; Marceglia, S. Personally Collected Health Data for Precision Medicine and Longitudinal Research. *Front. Med.* **2019**, *6*, 125. [CrossRef] [PubMed]
37. Antón, M.Á.; Ordieres-Meré, J.; Saralegui, U.; Sun, S. Non-Invasive Ambient Intelligence in Real Life: Dealing with Noisy Patterns to Help Older People. *Sensors* **2019**, *19*. [CrossRef]
38. Bednar, K.; Spiekermann, S.; Langheinrich, M. Engineering Privacy by Design: Are engineers ready to live up to the challenge? *Inf. Soc.* **2019**, *35*, 122–142. [CrossRef]
39. Cavoukian, A. Privacy by Design: The 7 Foundational Principles. Available online: https://www.ipc.on.ca/wp-content/uploads/resources/7foundationalprinciples.pdf (accessed on 25 April 2020).
40. Gartner. Gartner Identifies Top 10 Strategic IoT Technologies and Trends. Available online: https://gtnr.it/2JLiEbA (accessed on 06 February 2020).
41. OPC Foundation. OPC UA Specification: Part 1–Concepts, Version 1.04. OPC Foundation 2017. Available online: https://opcfoundation.org/developer-tools/specifications-unified-architecture/part-1-overview-and-concepts/ (accessed on 25 April 2020).
42. Fraile.; Sanchis.; Poler.; Ortiz. Reference Models for Digital Manufacturing Platforms. *Appl. Sci.* **2019**, *9*, 4433. [CrossRef]
43. Glebke, R.; Henze, M.; Wehrle, K.; Niemietz, P.; Trauth, D.; Mattfeld, M.; Bergs, T. A Case for Integrated Data Processing in Large-Scale Cyber-Physical Systems. In Proceedings of the 52nd Hawaii International Conference on System Sciences, Wailea, HI, USA, 8 August–1 November 2019
44. Al-Jabri, I.M.; Roztocki, N. Adoption of ERP systems: Does information transparency matter? *Telemat. Inf.* **2015**, *32*, 300–310. [CrossRef]
45. Ghobakhloo, M. The future of manufacturing industry: A strategic roadmap toward Industry 4.0. *J. Manuf. Technol. Manag.* **2018**, *29*, 910–936. [CrossRef]
46. Zhang, L.; Luo, M.; Li, J.; Au, M.H.; Choo, K.K.R.; Chen, T.; Tian, S. Blockchain based secure data sharing system for Internet of vehicles: A position paper. *Veh. Commun.* **2019**, *16*, 85–93. [CrossRef]
47. Wang, L.; Guo, S. Blockchain Based Data Trust Sharing Mechanism in the Supply Chain. In *Security with Intelligent Computing and Big-data Services*; Yang, C.N., Peng, S.L., Jain, L.C., Eds.; Springer International Publishing: Cham, Switzerland, 2020; pp. 43–53.
48. Shen, B.; Guo, J.; Yang, Y. MedChain: Efficient Healthcare Data Sharing via Blockchain. *Appl. Sci.* **2019**, *9*, 1207. [CrossRef]
49. Zheng, X.; Sun, S.; Mukkamala, R.R.; Vatrapu, R.; Ordieres-Meré, J. Accelerating Health Data Sharing: A Solution Based on the Internet of Things and Distributed Ledger Technologies. *J. Med. Internet Res.* **2019**, *21*, e13583. [CrossRef]
50. Ledwaba, L.P.I.; Hancke, G.P.; Isaac, S.J.; Venter, H.S. Developing a Secure, Smart Microgrid Energy Market using Distributed Ledger Technologies. In Proceedings of the 2019 IEEE 17th International Conference on Industrial Informatics (INDIN), Helsinki, Finland, 22–25 July 2019; Volumes 1, pp. 1725–1728. [CrossRef]
51. Nikander, P.; Autiosalo, J.; Paavolainen, S. Interledger for the Industrial Internet of Things. In Proceedings of the 2019 IEEE 17th International Conference on Industrial Informatics (INDIN), Helsinki, Finland, 22–25 July 2019; Volume 1, pp. 908–915. [CrossRef]
52. Gomes, P.; Cavalcante, E.; Batista, T.; Taconet, C.; Conan, D.; Chabridon, S.; Delicato, F.C.; Pires, P.F. A semantic-based discovery service for the Internet of Things. *J. Internet Serv. Appl.* **2019**, *10*, 10. [CrossRef]

53. Delicato, F.C.; Pires, P.F.; Batista, T. *Resource Management for the Internet of Things*; Springer International Publishing AG: Cham, Switzerland, 2017. [CrossRef]
54. Popov, S. The Tangle. Available online: https://assets.ctfassets.net/r1dr6vzfxhev/2t4uxvsIqk0EUau6g2sw0g/45eae33637ca92f85dd9f4a3a218e1ec/iota1_4_3.pdf (accessed on 25 April 2020).
55. Sahinel, D.; Akpolat, C.; Görür, O.C.; Sivrikaya, F. Integration of Human Actors in IoT and CPS Landscape. In Proceedings of the 2019 IEEE 5th World Forum on Internet of Things (WF-IoT), Limerick, Ireland, 15–18 April 2019; pp. 485–490.
56. Zezulka, F.; Marcon, P.; Vesely, I.; Sajdl, O. Industry 4.0 —An Introduction in the phenomenon. *IFAC-PapersOnLine* **2016**, *49*, 8–12. [CrossRef]
57. Giannetti, C.; Ransing, M.R.; Ransing, R.S.; Bould, D.C.; Gethin, D.T.; Sienz, J. Organisational Knowledge Management for Defect Reduction and Sustainable Development in Foundries. *Int. J. Knowl. Syst. Sci.* **2015**, *6*, 18–37. [CrossRef]
58. Cheng, H.; Zeng, P.; Xue, L.; Shi, Z.; Wang, P.; Yu, H. Manufacturing Ontology Development Based on Industry 4.0 Demonstration Production Line. In Proceedings of the 2016 Third International Conference on Trustworthy Systems and Their Applications (TSA), Wuhan, China, 18–22 September 2016; pp. 42–47. [CrossRef]
59. Blockchain Ontology with Dynamic Extensibility. Available online: https://github.com/hedugaro/Blondie (accessed on 6 February 2020).
60. Smith, B.; Grenon, P.; Goldberg, L. Biodynamic Ontology: Applying BFO in the Biomedical Domain. *Stud. Health Technol. Inf.* **2004**, *102*, 20–38.
61. Gangemi, A.; Guarino, N.; Masolo, C.; Oltramari, A.; Schneider, L. Sweetening Ontologies with DOLCE. In Proceedings of the EKAW: International Conference on Knowledge Engineering and Knowledge Management Sigüenza, Spain, 1–4 October 2002; pp. 166–181.
62. Industrial Ontologies Foundry (IOF). Available online: https://sites.google.com/view/industrialontologies/home (accessed on 25 April 2020).
63. Yun, J.; Ahn, I.Y.; Song, J.; Kim, J. Implementation of Sensing and Actuation Capabilities for IoT Devices Using oneM2M Platforms. *Sensors* **2019**, *19*, 4567. [CrossRef] [PubMed]
64. Cimino, C.; Negri, E.; Fumagalli, L. Review of digital twin applications in manufacturing. *Comput. Ind.* **2019**, *113*, 103130. [CrossRef]
65. Jha, S.B.; Babiceanu, R.F.; Seker, R. Formal modeling of cyber-physical resource scheduling in IIoT cloud environments. *J. Int. Manuf.* **2019**, pp. 1–16. [CrossRef]
66. Ikävalko, H.; Turkama, P.; Smedlund, A. Value creation in the internet of things: Mapping business models and ecosystem roles. *Technol. Innov. Manag. Rev.* **2018**, *8*. [CrossRef]
67. Zhang, Y.; Wen, J. The IoT electric business model: Using blockchain technology for the internet of things. *Peer Peer Netw. Appl.* **2017**, *10*, 983–994. [CrossRef]
68. D'Antonio, G.; Macheda, L.; Sauza Bedolla, J.; Chiabert, P. PLM-MES Integration to Support Industry 4.0. In Proceedings of the PLM: IFIP International Conference on Product Lifecycle Management, Seville, Spain, 10–12 July 2017; pp. 129–137.
69. Carbonera, J.L.; Fiorini, S.R.; Prestes, E.; Jorge, V.A.M.; Abel, M.; Madhavan, R.; Locoro, A.; Gonçalves, P.; Haidegger, T.; Barreto, M.E.; Schlenoff, C. Defining positioning in a core ontology for robotics. In Proceedings of the 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems, Tokyo, Japan, 3–7 November 2013; pp. 1867–1872.
70. Kun, D.P.; Varga, E.B.; Toth, Z. Ontology based navigation model of the ILONA system. In Proceedings of the 2017 IEEE 15th International Symposium on Applied Machine Intelligence and Informatics (SAMI), Herl'any, Slovakia, 26–28 January 2017; pp. 479–484. [CrossRef]
71. Villalonga, C.; Pomares, H.; Rojas, I.; Banos Legran, O. MIMU-Wear: Ontology-based sensor selection for real-world wearable activity recognition. *Neurocomputing* **2017**, 1–25. [CrossRef]
72. Long, W. Construct MES Ontology with OWL. In Proceedings of the 2008 ISECS International Colloquium on Computing, Communication, Control, and Management, Guangzhou, China, 3–4 August 2008; Volume 1, pp. 614–617.
73. Khedher, A.B.; Henry, S.; Bouras, A. Integration between MES and Product Lifecycle Management. In Proceedings of the ETFA2011, Toulouse, France, 5–9 September 2011. [CrossRef]

74. Fumagalli, L.; Pala, S.; Garetti, M.; Negri, E. Ontology-Based Modeling of Manufacturing and Logistics Systems for a New MES Architecture. In *Advances in Production Management Systems. Innovative and Knowledge-Based Production Management in a Global-Local World*; Grabot, B., Vallespir, B., Gomes, S., Bouras, A., Kiritsis, D., Eds.; Springer: Berlin/Heidelberg, Germany, 2014; pp. 192–200.
75. OGC. OGC® IndoorGML. Available online: https://www.ogc.org/standards/indoorgml (accessed on 6 February 2020).
76. ETSI. Smart Body Area Networks (SmartBAN) Unified data representation formats, semantic and open data model. Available online: https://www.etsi.org/deliver/etsi_ts/103300_103399/103378/01.01.01_60/ts_103378v010101p.pdf (accessed on 25 April 2020).
77. Nach, H.; Lejeune, A. Implementing ERP in SMEs: Towards an Ontology Supporting Managerial Decisions. In Proceedings of the 2008 International MCETECH Conference on e-Technologies (Mcetech 2008), Montreal, QC, Canada, 23–25 Janauary 2008; pp. 223–226.
78. Hepp, M.; Wechselberger, A. OntoNaviERP: Ontology-Supported Navigation in ERP Software Documentation. In *The Semantic Web-ISWC 2008*; Sheth, A., Staab, S., Dean, M., Paolucci, M., Maynard, D., Finin, T., Thirunarayan, K., Eds.; Springer: Berlin/Heidelberg, Germany, 2008; pp. 764–776.
79. Dudas, P.M.; Ghafourian, M.; Karimi, H.A. ONALIN: Ontology and Algorithm for Indoor Routing. In Proceedings of the 2009 Tenth International Conference on Mobile Data Management: Systems, Services and Middleware, Taipei, Taiwan, 18–20 May 2009; pp. 720–725.
80. Stevenson, G.; Ye, J.; Dobson, S.; Nixon, P. LOC8: A Location Model and Extensible Framework for Programming with Location. *IEEE Pervas. Comp.* **2010**, *9*, 28–37. [CrossRef]
81. Rhayem, A.; Mhiri, M.B.A.; Gargouri, F. HealthIoT Ontology for Data Semantic Representation and Interpretation Obtained from Medical Connected Objects. In Proceedings of the 2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA), Hammamet, Tunisia, 30 October–3 November 2017; pp. 1470–1477.
82. Maheshwari, N.; Srivastava, S.; Rajan, K.S. Development of an Indoor Space Semantic Model and Its Implementation as an IndoorGML Extension. *ISPRS Int. J. Geo Inf.* **2019**, *8*, 333. [CrossRef]
83. Khruahong, S.; Kong, X.; Sandrasegaran, K.; Liu, L. Develop An Indoor Space Ontology For Finding Lost Properties for Location-Based Service of Smart City. In Proceedings of the 2018 18th International Symposium on Communications and Information Technologies (ISCIT), Bangkok, Thailand, 26–29 September 2018; pp. 54–59.
84. universAAL IoT. Fitbit Ontology. Unified Data Representation Formats, Semantic and Open Data Model. Available online: https://github.com/universAAL/ontology/issues/479 (accessed on 25 April 2020).
85. Goldfain, A.; Smith, B.; Arabandi, S.; Brochhausen, M.; Hogan, W. Vital Sign Ontology. In Proceedings of the Workshop on Bio-Ontologies, Vienna, Austria, June 2011; pp. 71–74.
86. Sun, S. Interoperable Data Handling for Industry 4.0. Available online: https://gitlab.com/sunshengjing/interoperable-data-usage-for-industry-4.0 (accessed on 25 April 2020).
87. Longo, F.; Nicoletti, L.; Padovano, A. Smart operators in industry 4.0: A human-centered approach to enhance operators' capabilities and competencies within the new smart factory context. *Comp. Ind. Eng.* **2017**, *113*, 144–159. [CrossRef]
88. Aehnelt, M.; Urban, B. Follow-Me: Smartwatch Assistance on the Shop Floor. In *HCI in Business*; Nah, F.F.H., Ed.; Springer International Publishing: Cham, Switzerland, 2014; pp. 279–287.
89. Chapin, F.S.; Folke, C.; Kofinas, G.P. A framework for understanding change. In *Principles of Ecosystem Stewardship*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 3–28.
90. Kharlamov, E.; Martin-Recuerda, F.; Perry, B.; Cameron, D.; Fjellheim, R.; Waaler, A. Towards semantically enhanced digital twins. In Proceedings of the 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 10–13 December 2018; pp. 4189–4193.
91. Gómez-Berbís, J.M.; de Amescua-Seco, A. SEDIT: Semantic Digital Twin Based on Industrial IoT Data Management and Knowledge Graphs. In *International Conference on Technologies and Innovation*; Springer: Cham, Switzerland, 2019; pp. 178–188.
92. Lu, J.; Zheng, X.; Gharaei, A.; Kalaboukas, K.; Kiritsis, D. Cognitive Twins for Supporting Decision-Makings of Internet of Things Systems. *arXiv* **2019**, arXiv:1912.08547.
93. Zhou, J.; Zhou, Y.; Wang, B.; Zang, J. Human-Cyber-Physical Systems (HCPSs) in the Context of New-Generation Intelligent Manufacturing. *Engineering* **2019**, *5*, 624–636. [CrossRef]

94. Roda-Sanchez, L.; Garrido-Hidalgo, C.; Hortelano, D.; Olivares, T.; Ruiz, M.C. OperaBLE: An IoT-Based Wearable to Improve Efficiency and Smart Worker Care Services in Industry 4.0. *J. Sens.* **2018**, *2018*, 1–12. [CrossRef]
95. Awolusi, I.; Marks, E.; Hallowell, M. Wearable technology for personalized construction safety monitoring and trending: Review of applicable devices. *Automat. Constr.* **2018**, *85*, 96–106. [CrossRef]

sensors

MDPI

Article

Healthy Operator 4.0: A Human Cyber–Physical System Architecture for Smart Workplaces

Shengjing Sun [1,2], Xiaochen Zheng [1,3], Bing Gong [4], Jorge García Paredes [1] and Joaquín Ordieres-Meré [1,*]

1 Escuela Técnica Superior de Ingenieros Industriales (ETSII), Universidad Politécnica de Madrid, José Gutiérrez Abascal 2, 28006 Madrid, Spain; shengjing.sun@alumnos.upm.es (S.S.); xiaochen.zheng@epfl.ch (X.Z.); j.gparedes@alumnos.upm.es (J.G.P.)
2 Exposure, Epidemiology, and Risk Program, Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, MA 02115, USA
3 ICT for Sustainable Manufacturing, SCI-STI-DK, École polytechnique fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland
4 Jülich Supercomputing Center, Forschungszentrum Jülich GmbH, Wilhelm-Wohnen-Str, 52428 Jülich, Germany; b.gong@fz-juelich.de
* Correspondence: j.ordieres@upm.es

Received: 18 March 2020; Accepted: 1 April 2020; Published: 3 April 2020

Abstract: Recent advances in technology have empowered the widespread application of cyber–physical systems in manufacturing and fostered the Industry 4.0 paradigm. In the factories of the future, it is possible that all items, including operators, will be equipped with integrated communication and data processing capabilities. Operators can become part of the smart manufacturing systems, and this fosters a paradigm shift from independent automated and human activities to human–cyber–physical systems (HCPSs). In this context, a Healthy Operator 4.0 (HO4.0) concept was proposed, based on a systemic view of the Industrial Internet of Things (IIoT) and wearable technology. For the implementation of this relatively new concept, we constructed a unified architecture to support the integration of different enabling technologies. We designed an implementation model to facilitate the practical application of this concept in industry. The main enabling technologies of the model are introduced afterward. In addition, a prototype system was developed, and relevant experiments were conducted to demonstrate the feasibility of the proposed system architecture and the implementation framework, as well as some of the derived benefits.

Keywords: healthy operator 4.0; human–cyber–physical system; industrial internet of things; industry 4.0; smart workplaces

1. Introduction

The continuous technological innovations in the domains of information technology (IT), the Internet of Things (IoT), and artificial intelligence (AI), among others, have significantly changed production systems [1–3]. Recent advances of these technologies have enabled a systematical implementation of cyber–physical systems (CPS) in manufacturing, which has significantly improved the efficiency of production systems and made them perform more resiliently and collaboratively. These advanced technologies are transforming the manufacturing industry to the Industry 4.0 paradigm [4,5].

In the Industry 4.0 era, all items in a smart factory will be equipped with integrated communication and certain data processing capabilities, by the so called Industrial Internet of Things (IIoT). The IIoT scenario enables a close connection between the physical and digital worlds, and the paradigm of cyber–physical equivalence or digital twin emerges [6], where the vision of the digital twin

itself refers to a comprehensive physical and functional description of a component, product, or system, which includes more or less all information that could be useful in all lifecycle phases [7]. This represents a seamless integration between both worlds, meaning that the digital part can virtually replicate the behavior of the physical counterpart and, henceforth, be used to create new added value services in both directions.

The smart "artifacts", including people, will be connected to the CPS [8,9], which is different from computer integrated manufacturing (CIM), as the intention of Industry 4.0 paradigm is not creating unmanned production facilities. In contrast, it provides a great opportunity for operators to become part of the smart manufacturing systems in such a way where the individual skills and talents of the operators can be better realized [10–12].

The advanced technologies, such as IoT and CPS, in the Industry 4.0 paradigm provide new forms of interaction between operators and machines that will produce new intelligent workforces and will bring significant impacts on the nature of manufacturing [13,14]. Human-centricity is one of the most critical focuses in the transformation towards Industry 4.0. It allows for a paradigm shift from independent, automated, and human activities to human–cyber–physical systems (HCPSs) where machines are not designed to replace the skills and talents of humans, but rather to co-exist with, and to assist humans in being more efficient and effective [15,16]. A clear opportunity in such a path is to align the human central view in the decisions for process improvement, where improvement in the classical lean management, e.g., $(CPD)_nA$ [17], was performed through a strategy enabling taking advantage of the IoT devices, and looking to increase the knowledge analysis capabilities of the responsible worker. Therefore, the aims of HCPSs can be described [12,15] as:

- Empowering people to dynamically interact with machines in both the cyber and physical worlds to fit the operators' cognitive and physical needs supported by intelligent human–computer interaction techniques;
- Improving the physical, sensing, and cognitive capabilities of people taking advantages from diverse technologies, such as IoT and wearable technologies.

On the basis of HCPSs, the concept of Operator 4.0 was proposed, aiming at improving the cooperation between humans and machines [13,15]. Operator 4.0 represents "a new design and engineering philosophy for adaptive production systems where the focus is on treating automation as a further enhancement of the human's physical, sensorial, and cognitive capabilities" [15].

According to different enabling technologies and targeting aspects, Operator 4.0 can be empowered by different technologies [12,13], such as,

- Virtual operator, enabled by virtual reality (VR)/augmented reality (AR) [18],
- Super-strength operator, enabled by exoskeletons [19],
- Smarter operator, enabled by intelligent personal assistant (IPA)-based solutions [20],
- Healthy operator, enabled by wearable technologies combined with advanced data analytic techniques [21],
- etc.

As one of the core sub-types of Operator 4.0, the Healthy Operator 4.0 (HO4.0) aims to address the concerns regarding increasing workforce stress levels, the state of psycho-social health [22,23], and the new potential physical risks [24,25] in the cyber–physical production environments caused by the introduction of new Industry 4.0 technologies, including autonomous and collaborative robots, AR, and VR.

Although occupational health has been a popular topic for decades in both academic research and industrial applications, it has rarely had the advantage of real time quantitative measurements, mainly due to the limited capabilities for directly measuring such factors [21,26–31].

The recent advances in sensing technologies and IoT, especially wearable technologies, provide new solutions for real-time monitoring of an operator's activities, locations, vital signs, etc., as well as

the status of the surrounding workplace environment [30–32]. These technologies make it possible to develop health-related applications, such as alerting operators of possible exposure to hazardous environments, avoiding collisions with moving heavy equipment, and preventing anti-ergonomic body movements and postures [21,33]. However, in current industrial practice, most applications are developed in isolated circumstances aimed at addressing specific problems. Therefore, there is a gap in creating human-centered systems able to promote operators' learning context not only relying on single parameters but also providing a meaningful articulated set of relevant parameters both in the short and long term.

To cope with this challenge, a unified architecture for HO4.0 is required to support the integration of different enabling technologies, thus guiding the implementation of this concept in the context of occupational safety and health. The focus of the paper is established to:

- Formally define the HO4.0 concept, by carefully analyzing the contributions from previous research in relevant fields,
- Propose a HO4.0 application framework enabling relevant data integration in highly dynamic work environments, supported by the IIoT and wearable technology. According to the definitions, we presented a multi-layer architecture and an implementation framework to guide the their application in industry;
- Provide a prototype system, developed according to the proposed framework, where the assessment of its capabilities can be derived; and
- Conduct experiments to verify the capabilities of the prototype system to provide new knowledge on real applications.

From the methodological point of view, this study followed the critical action research method, in which the researcher and a field operator collaborated in the diagnosis of the problem and in the development of a solution based on the diagnosis [34]. Critical action research is based on the analysis, action, evaluation, and critical analysis of practices based on collected data, in order to introduce improvements in the relevant practices. This type of research is facilitated by the participation and collaboration of a number of individuals with a common purpose where the research focuses on specific situations and their context.

The rest of this paper is organized as follows: Section 2 introduces the formal definition for HO4.0, enabling us to build over the proposed framework, in Section 3. Section 4 introduces the application case supporting the prototype creation. The experiments and outcomes are presented in Section 5. Then Section 6 summarizes the findings and provides our interpretation in the context of the proposed HO4.0 framework. Finally, Section 7 provides a more strategic perspective of the knowledge gains, including the limitations and further research directions.

2. HO4.0 Concept

As the sub-type of Operator 4.0, the HO4.0 concept was first envisioned in [12], where it was roughly depicted as *Operator + Wearable Tracker = Healthy Operator (physical and cognitive interaction)*. In this stage, individual wearable devices, such as smartwatches, were used to track an operator's health-related metrics. However, an isolated and separate monitoring dimension can not fully represent comprehensive health aspects. Therefore, a unified health view of the operator is needed to enable holistic workforce health management and analytics. In this study, aligned with the Operator 4.0 definition [13,15], and extending the former concept, the HO4.0 is formally defined as a system focused on health and operator well-being, looking to facilitate the operators' empowerment by enabling relevant knowledge creation, including modeling their behaviors.

It is relevant that our adopted perspective differs from the one chosen in [12,21], where the the concept was closer to the operator. It was introduced as a specific type or vision for the Operator 4.0 concept. Instead, the ambition here was to consider the relevant information gathered, when appropriately fused, as a digital image of the operator's behavior, and due to this, the digital twin

approach comes in naturally. To facilitate the integration of concepts, the natural approach is to consider the HO4.0 as a system instead of a type of operator. Under this interpretation, the HO4.0 can be also presented as a virtual system: the HO4.0 digital twin, gathering all the health related information from operators and potentially enabling the learning of rules from the different behaviors. This does include benchmarks regarding the health impact of different operator routines, etc.

The HO4.0 digital twin is powered by IIoT networks, wearable technologies, ambient intelligence, and modeling technologies. It not only enables real-time health risk information, e.g., risk alert, but also make it possible to simulate the future behavior of operators; to infer and forecast the evolution from their behavior at mid-long term ranges, aiming to reduce the operator's cognitive and physical workload, and increase the operator's well-being such as Occupational Safety and Health (OSH), job satisfaction, work-related affect, and enhance the workforce productivity in the Industry 4.0 context.

3. HO4.0 Framework

The proposed framework has two main components, one is the architecture, and the other involves the required technologies.

3.1. Proposed Architecture

A traditional CPS structure usually includes five levels (5C structure), including Smart Connection, Data-to-information Conversion, Cyber, Cognition and Configuration [4]. Based on the Operator 4.0 definition, we adapted the 5C structure of CPS and designed a four-level architecture for HO4.0, compatible with the 5C CPS structure, as the proposed architecture shown in Figure 1 is composed of the Smart Connection layer, the Integration and Communication layer, the Modeling layer, and the Cognition layer.

- **Smart Connection layer:** Accurate and reliable data from operators, machines, the ambient environment, and other parts of the production system are the foundation of the HO4.0 system. The sensing layer collected these data utilizing wearable devices, ambient sensors, and the industrial sensors of manufacturing systems etc. Wearable devices, such as smart watches, smart bands, smart glasses, smart shoes, and smart helmets, contain a variety of sensors that can measure the operator's location, movements, heart rate, blood pressure, body temperature, concentration level (e.g., fatigue), foot pressure, etc. Stationary or portable ambient sensors, such as air quality, thermal (temperature and humidity) sensors, and acoustic and light sensors, are used to monitor the conditions of the workplace environment surroundings, such as the indoor air quality level, temperature, humidity, pollutant concentration, light intensity, and noise levels. The indoor positioning system (IPS) was adopted to detect the movement and location of the operators, vehicles, work pieces, machines, etc. in a workplace.
- **Integration and Communication layer:** This layer was mainly composed of edge computing devices or gateways, which could include a smartphone, tablet, router, or single-board computer (SBC) etc. This layer was required due to the limitations of the size, power supply, and computing capabilities of most wearable devices and ambient sensors from the sensing layer. They are not capable of performing operations such as data filtering and integration. This layer aimed to filter and integrate these diverse data, and if necessary, to convert and pre-process the raw data based on different data modeling approaches. The sensors used in the sensing layer are from many domains, e.g., monitoring human psychological behaviors and detecting environmental conditions. This layer also facilitated the data uploading to the server storage for further data analysis. The edge computing device in this layer can be applied to provide real-time health warnings that are critical for a human that interacts with machines and HCPSs.
- **Modeling layer:** The health related data was fused and modeled in this layer. The twin model of operator health was created and relevant simulations can be run to apply the derived high lever rules. This is typically conducted in a local or remote/cloud server with a higher computing

capability that is able to care about data streams from different communication devices. Semantic data description enables the interoperability and usage of semantic models, and knowledge graphs might also be employed to facilitate the knowledge extraction. Supported by advanced machine learning techniques and semantic-driven data fusion, useful knowledge, such as risk alerts, improved advice, and rules, can be extracted.

- **Cognition layer:** Based on the results from the model layer, the cognition layer can provide hints and insights from cyber space to physical space and acts as a monitoring system for the preventive decisions from operators, machines, or ambient environments. On this layer, analyzing the results can support the decision-making and can be presented with proper data visualization techniques. Certain user interfaces, applets, or web-based services can be also implemented. The alerts, advice, and orders can be delivered to different stakeholders to help them conduct corresponding actions. The target is to ensure work–life health, safety, and satisfaction.

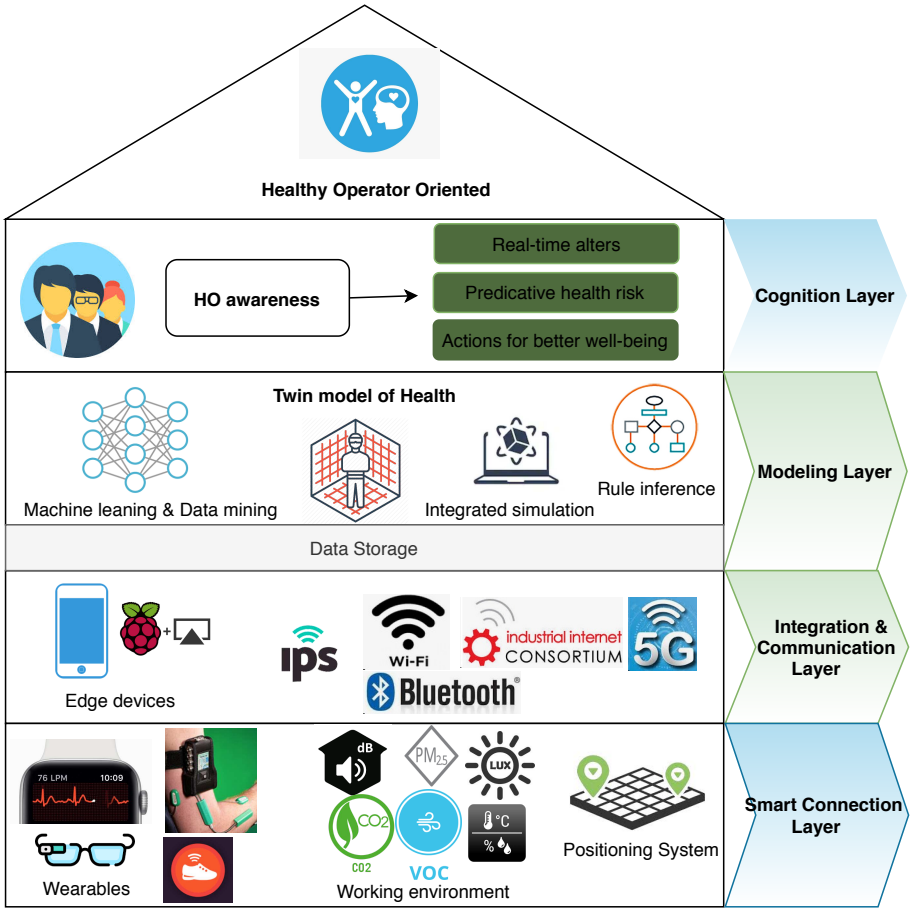


Figure 1. Healthy Operator 4.0 (HO4.0) architecture.

3.2. Enabling Technologies

The successful implementation of HO4.0 architecture requires the support of a series of advanced enabling technologies. Some of the ones that were mentioned in the model and used in the prototype are introduced as follows.

- **Wearable technology:** Wearable technology or wearable computing is the study or practice of inventing, designing, building, or using miniature body-worn computational and sensory devices [35]. With the rapid development of sensing technologies, various types of sensors can be embedded in wearable devices. According to their features, they can be divided into four major groups [36]:

1. Environmental sensors, such as light sensors, temperature sensors, sound sensors, humidity sensors, air quality sensors (e.g., CO_2 sensors, particulate matter (PM) sensors, and volatile organic components (VOCs) sensors), and barometric sensors.
2. Biosensors, including body temperature sensors, heart-rate-monitoring sensors, electrocardiogram (ECG), electroencephalography (EEG), electromyography (EMG) sensors, blood pressure sensors, galvanic skin response (GSR) sensors, eye tracking, weight insole, and glucose level sensors.
3. Location tracking sensors, such as GPS, altimeters, magnetometers, compasses, and accelerometers.
4. Other sensors, including camera sensors, communication sensors, ultrasonic sensors, infrared receiver (IR), sensors.

The application of wearable technologies in HCPSs provides rich information related to the operators and the surrounding environment.

- **Indoor Positioning System:** IPSs locate and track objects within a closed environment usually based on triangulation and multilateration methods using light, ultrasound, or radio signals to provide positional information [37]. The tracking of operators and production equipment can be significantly improved by the application of IPS [13]. Several wireless technologies might be used in IPS depending on the application situations by means of different positioning algorithms [38], such as Global Positioning System (GPS), Radio Frequency IDentification (RFID), Cellular networks, Ultra-wideband (UWB), wireless local area network (WLAN), Bluetooth etc. The adoption of IPS makes possible of a context-aware system for HO4.0 to help operators monitor the locations of machines, vehicles and workpieces to improve productivity and avoid potential collision risks.
- **Ambient environment monitoring:** The condition of the ambient environment plays a crucial role in a HO4.0 system. It may impact the operators' working performance or even harm their health. For example, indoor air pollution is one of the leading environmental risks, and indoor air quality (IAQ) is proven to have significant impacts on human comfort, health, and performance [29]. The advancement of low-cost IoT sensors, in recent years, has enabled the use of wireless communications and computing for interacting with the physical world. These sensors can measure indoor environmental parameters including IAQ, comfort, lighting, and acoustic conditions [39–41]. Many of these low-cost ambient sensors allow portable deployment and different sensors can be easily packed in one board to fit the requirements of different application scenarios. The data generated by such sensors can be collected by certain edge computing devices, such as single board computers (SBC). As an example, an ambient environment monitoring system based on a sensor island and a Raspberry Pi SBC was developed in a previous study [29]. The sensor island used in this system was composed of nine different ambient sensors to monitor different aspects of the ambient environment. More details of this system are explained in the prototyping section.
- **Knowledge engineering:** A semantic data model is a high-level representation of knowledge. It is designed to capture the meaning of data attributes and relationships, and obtain the meaning of each instance in an application context. The standard data model and knowledge engineering facilitates information exchange and data interpretation in a very heterogeneous data sources. Fusion and integration data based on rich semantic context gravitates toward applications, such as effective decision making processes, by fully making use of the various heterogeneous data. Ontologies, stemming from the Semantic Web, are used to define standard and common terms, vocabulary, and relationships for a particular area/domain. Aligning on common interpretations through ontologies, data/knowledge graphs can be generated to refer to the relationships among different data sources from different domains. Existing or adapted ontologies [42,43]

are recommended to be used for semantic modeling aiming to derive action rules toward the operator's health management.

- **Machine learning (ML):** The data gathered from the sensing layer can contain valuable information, which is the core of the HO4.0 system. Indeed, when scaled up appropriately, it can become a large volume of data; whereas, a global analysis can help in providing better interpretations for the impact of different behaviors. Therefore, advanced data modeling tools, such as machine learning, can be useful to create or emphasize such new knowledge. The wide applications of the IoT have promoted the concept of big data, which on the other hand, has fostered the explosive involvement of the advanced machine learning techniques represented by deep learning. Despite the great achievement of machine learning in many fields, its power has not been fully and systematically explored in the industry sector, particularly in the human-centric systems. Here the benefits from the adopted concept for HO4.0, such as the system proposed in this research (which enables the digital twin approach) facilitates modeling capabilities to detect patterns and to add value in the real world for human-centric applications. For example, deep learning can be used to monitor an operator's activities based on the motion data collected from wearable devices [9,44,45], but when combined with health related information and environmental conditions, this can provide a health index for such activities. In addition to the self learning process based on the real time data, the inter-comparison between workers would help to better understand the physiological and mental demands, but it can also help to find different patterns and to learn about the sustainability of them for the long term [46].

4. The Application Case and Experiments

To validate the proposed architecture, an application case was carried out, enabling the construction of a prototype helping with different practical implementations.

4.1. Case Selection

The case is connected with a logistic intra-facility, where the interest was in obtaining health related knowledge of crane operators in their working environment, who were in charge of loading trucks that moved manufactured steel rebar to build singular concrete structures.

Figure 2a shows how an operator moved around, looking for the pieces to be loaded, then attaching them and loading them into the trucks, where they went up and down a specific platform as presented in Figure 2b, which was designated to facilitate access to the truck itself and to maintain awareness on how the material was placed inside the truck. The preparation of the materials on trucks must be carefully planned as the unload process typically needs to visit different construction sectors to deliver specific sets of rebar. Therefore, a specific disposal sequence of items as per the layers in the truck must also be carefully planned and is prone to errors and stressful situations, where there are doubts about what was actually loaded into the truck. This is even more sensitive for sets involving small numbers of light rebar, where weight variations are not informative enough about errors in the loading process.

Outside of their specific job, the crane operator was working in a rather complex environment, with potential environmental quality concerns. This environment was a plant that manufactures steel rebar, where the production of specific items by cutting, blending, and welding may affect the air quality. Due to the working environment, as shown in Figure 2a, acoustic and illumination aspects were be other health threats.

Due to the different working conditions for crane operators and their complex mobility schema, their mental working environment, and physical workload, and because of the continuous over-pressure to quickly deliver products to trucks, we determined that these workers were suitable for inclusion in the validation for the HO 4.0 architecture.

(**a**) Crane operator at work.

(**b**) Operator unloading platform.

Figure 2. Crane operator working conditions.

4.2. Sensing Layer

In this application, based on the specific characteristics of the workers, the sensing layer was designed to include different dimensions, including the:

- **Environmental conditions**, including temperature, humidity, and noise, as they can impact the operators' health, as well as their performance.
- **Heart rate**, as related to their physical or mental effort.
- **Blood Pressure**, also as related to their physical or mental effort.
- **Arm angle**, referring to the horizontal line, as a reference for ergonomic working activities.
- **Position**, as a reference for understanding the movements and knowledge requirements in order to complete the work.
- **Steps**, as a reference for the physical effort.
- **Crane position**, in relation to the crane operator movements, allowing us to identify the effective working time.

Therefore, different sensing devices need to be attached to the crane operator to collect the information in relevant frequency patterns, although they must be as lightweight as possible, in order to be feasible for daily operations.

For the prototype, we decided to include three devices attached to the crane operators and one environmental station deployed on the shopfloor. Two of them were based on low cost devices, a MetaSensor (see Figure 3a) from mbientlab (https://mbientlab.com/store/metamotionr/), enabled us to monitor angles and acceleration and was used to track the arm angle through time. Another sensor we included was a smartband from HBand (see Figure 3b), able to track the heart rate, blood pressure, and steps.

For positioning the crane operator and the crane hook, we decided to use Ultra-wideband (UWB) technology with six stations able to track those positions [47]. The technology named $uRTLS^{TM}$, is a high precision real-time location solution (RTLS) based on UWB technology. The solution used several channels available at 3–7 GHz using the Decawave UWB chipset in compliance with IEEE 802.15.4.

The indoor environmental quality substantially affects a operator's health, comfort, working performance, and well-being. In this study, the working environment for operators was monitored with an IoT based environmental quality monitoring system developed by [29]. The measured indoor environment conditions included: the chemical environmental parameters: particulate matter (PM), formaldehyde (HCHO), Total Volatile Organic Compounds (TVOC), benzene (C_6H_6), carbon dioxide (CO_2), carbon monoxide (CO), ozone (O_3), nitrogen dioxide (NO_2), and the physical parameters: temperature, humidity, illumination, and noise.

Although these specific devices were selected for this implementation, without a lack of generalization, certain other devices, such as smart insoles, smart helmets, and skin conductivity sensors, could be also adequate, depending on the application case under consideration.

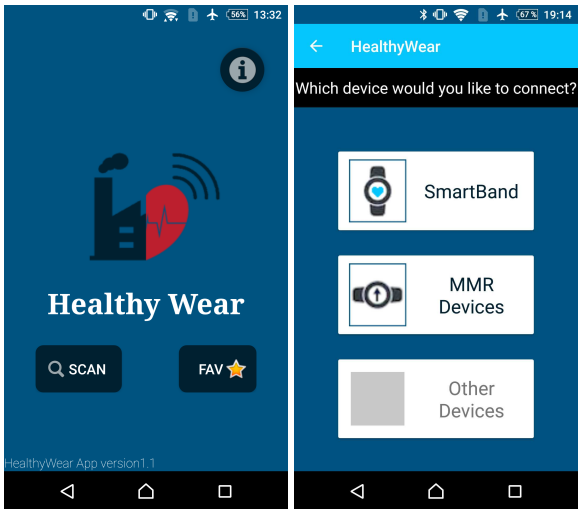
(**a**) Angle measurement device. (**b**) Smartband device.

Figure 3. Wearable devices used in this application.

4.3. Integration and Communication Layer

In this application, the usage of several technologies were required in order to properly handle the data streams. There were data collected from the factory that benefited from using the existing extranet to upload the datasets to the data repository. This is the case for the environmental monitoring station, which was based on a raspberry pi delivering the data throughout its internal network card. This was also the case for the positioning system, which was acquired from the tracktio company (https://tracktio.com/). This system had its own data repository, therefore an additional middleware was required in order to enable data integration.

There were other data streams that were collected only by bluetooth, as the devices only had this interface. To handle these streams, an Android based app was developed [48], making it convenient to integrate the features and enabling a flexible data collection architecture [49]. Figure 4 presents the main details for this application, which collected data according to a specific time sequence. For this application, a minute-based frequency was adopted for health and environmental parameters, and a second-based frequency for the position, which provided convenient granularity for the learning application. Indeed, it was relevant that the developed app showed real time values and informed the operator regarding significant values. Although it was possible to order smartband vibrations when specific conditions were met, this was not activated as it could distract the operators from their main task and possibly become an unacceptable safety risk.



(**a**) Main screen. (**b**) Device list.

Figure 4. The app developed for data collection.

The overall context for the data handling of the applied wearable in this study is presented in Figure 5.

Figure 5. Overview of the applied wearable data flow.

4.4. Modeling Layer

With all the data streams ingested, this layer facilitated the fusion and time alignment as per the device media access control (MAC) address and time. Therefore, the record-sets were derived providing context for the digital twin, according to the chosen application case. The data fusion could integrate data streams that in some cases belonged to the same individual or in other cases belonged to properties of the area, affecting several operators on the shop-floor. The collected data streams allowed us to consider semantic annotation to enable a machine based data handling strategy.

The existing Vcard ontology for people [50] was used to model each operator. The indoor environmental quality ontology: AIR_POLLUTION_Onto, proposed by [51], was used to model the operator working environmental exposure, risk, and control applications. The physiology factor of the operator was semantically represented with HUMAN STRESS ONTOLOGY [52], considering the data sources steaming from the physiological parameters, such as the heart rate and blood pressure. The indoor navigation ontology introduced in the ILONA system [53] was applied for the indoor model positioning of the operators. Therefore, the considered set of references, based on the integration of such ontologies was as shown in Table 1.

Table 1. Data streams with semantic annotation.

Entity	Message
crane operator	{ “Ontology”:”https://bit.ly/2OxeEkO”,”object”:”vcard”,”fn”:”José Luis Fernández”, ”nickname”:”Jose”,”hasEmail”:”mailto:perpalper@gmail.com”, “Gender”:”Male”,”bday”:”1988-06-23”,”adr”:”José Gutiérrez Abascal 2, 28006, Madrid” }
environmental situation	{“Ontology”: “AIR_POLLUTION_Onto”, “object”: “airPollutants”, “deviceId”: “Airmonitor1”, “PM”: “25 ug/m^3”, “CO2”: “0.04%”, “VOC”: “0.4 mg/m^3, “NOX”: “0”, “Timestramp”: “09/08/2019 09:10:00”}
stress factor	{“Ontology”: “HUMAN STRESS ONTOLOGY”, “object”: “Measurements”, “deviceId”: “Hband1”, “Timestramp”: “09/08/2019 09:10:00”, “stressPhysiology”: [{“heartRate”: “80”, “bloodPressureHigh”: “120”, “bloodPressureLow: “70”}]}
position	{“Ontology”: “ILONA”, “object”: “Position”, “deviceId”: “Crane1”, “Coordinate”: [{“x”:”40.342712”, “y”:”-3.123472”, “z”: “605.85”}], “Timestamp”: “09/12/2019 10:10:00”}

Based on the specific interest, the annotated data streams were combined to create the contextual datasets, where their structure was not fixed at all. Instead, it was problem related. Therefore, when the interest was to analyze the similarity of the operating conditions and operator behavior through time, the application for building projectors with general non-linear dimension reduction was needed.

Then, this layer will include different algorithms and tools, such as Principal component analysis (PCA), T-distributed Stochastic Neighbor Embedding (t-SNE), or Uniform Manifold Approximation and Projection (UMAP) dimension reduction techniques, which optimize a low-dimensional graph from twelve dimensions into two, in order to keep close records that are close in a higher dimensional space, and that preserved more of the global structure of the whole dataset.

When the interest is to look at specific behaviors that can be described by particular variations in variables, such variables were created and the ranges calculated. Based on those derived new variables, different Machine Learning techniques, such as association rules, can bring interesting rules explaining the causes for such behaviors.

When the case is to understand the behavior of operators through time, different regression algorithms were applied in order to forecast future trends, either by using the time series approach or the multidimensional one. Indeed, if the interest was to identify similar behavior between operators, different clustering algorithms were helpful, based on the medium term records of their digital twins.

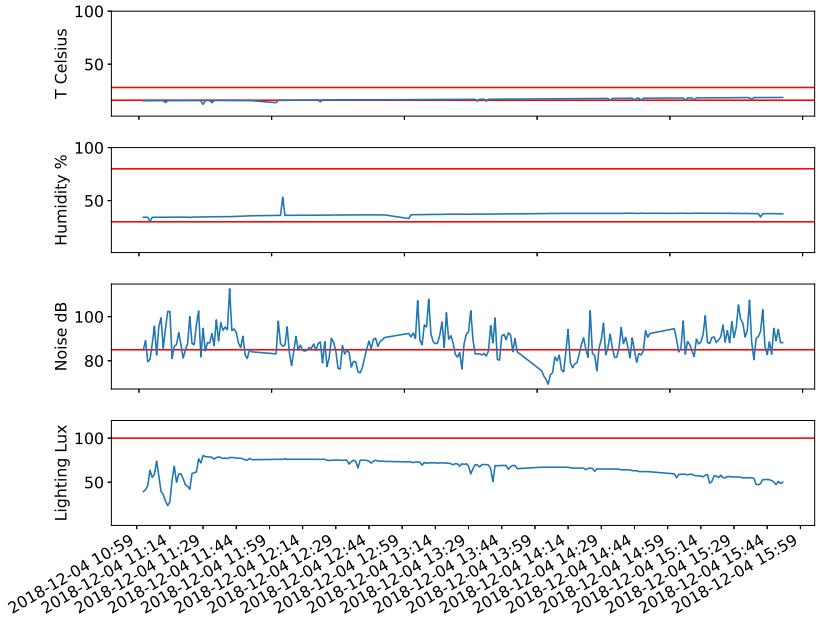
The previous examples illustrated the strong possibilities that this layer can provide when the interest is to derive new knowledge for operators at different level of aggregation and for different time scales.

5. Results

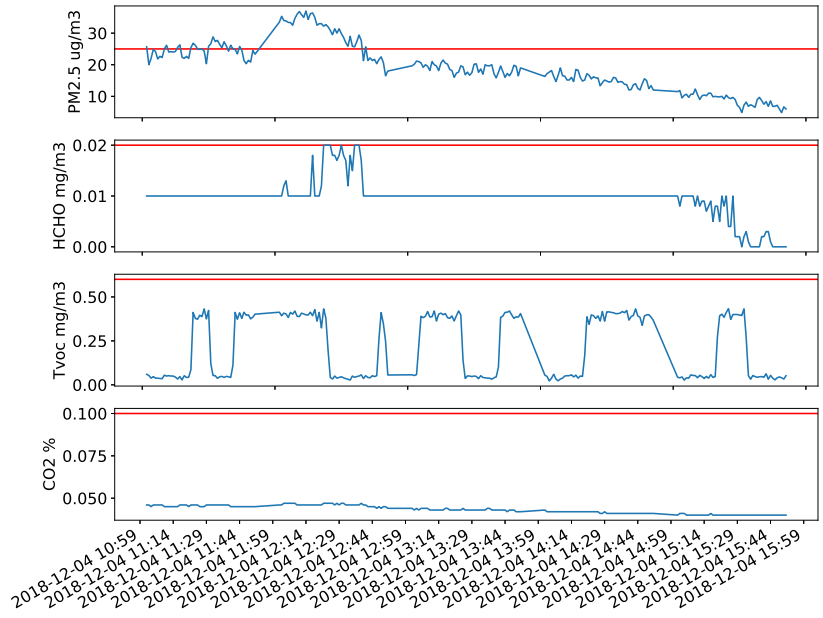
The working environment of operators for this particular case study is dissatisfying as shown in Figure 6, due to different aspects where the actual exposition levels exceeded the threshold limit value (TLV).

The TLV of a environmental substance (chemical or physical) is defined to be a level to which a worker can be exposed day after day for a working lifetime without adverse effects. The recommended comfort TLVs for temperature and humidity are 16–28 °C and 30–80%, respectively, according to World Health Organization (WHO) [54]. For work that requires the perception of details, such as offices, sheet metal work, or bookbinding, the minimal luminance TLV is 100 Lux, which is defined by the European Union (EU) standard. The noise levels defined by WHO are 85–90 dB(A) daily average [55], the smallest value (85 dB (A)) was taken as the TLV to ensure acoustic health to the maximal extent. The TLV for PM2.5 and HCHO are a 25 $\mu g/m^3$ average per day and a 0.016 ppm (0.02 mg/m^3) average per 8 hours as defined by WHO [56] and National Institute for Occupational Safety and Health (NIOSH) [57], respectively. The TLV of CO_2 (0.1%) and TVOC (0.60 mg/m^3 average per 8 h) are based on a daily average, which obtained references from Circulate App: EnvCon [58]. The Circulate company set TLVs by taking references from China's air-quality standards.

In this application case, the noise level was rather high consistently over time, as demonstrated in Figure 6a. In the vast majority of the working time, the noise level was higher than 85 dB(A) as defined by WHO [55], as the primary working activity is steel operations, which produces a huge amount of noise. Although operators are wear industrial headsets to protect their ears, constantly working in very noisy environments presents potential damage and hearing related problems, e.g., tinnitus. It also affects concentration at work, which was the case for crane operators. The environment was also rather dry and dim as depicted in Figure 6a (Humidity and Lighting). The chemical environmental conditions were inadequate as shown in Figure 6b. The PM2.5 and HCHO exceeded the contaminant level from time to time, and the TVOC is approached limits periodically. Therefore, these systems also help to identify dimensions where measures can be taken to improve the working environment, which would have a significant impact on the operators.



(**a**) Physical environmental conditions.

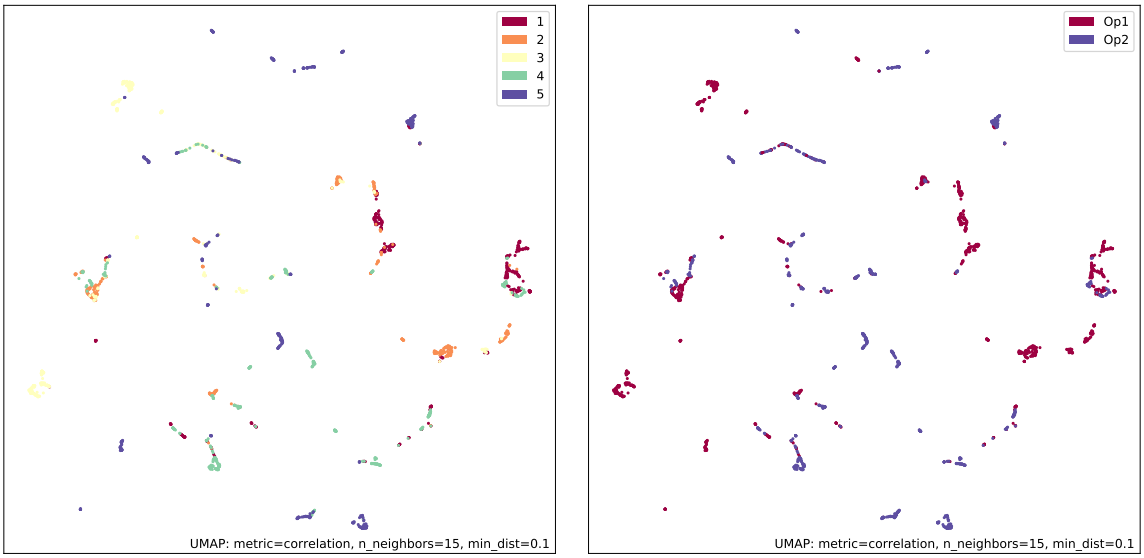


(**b**) Chemical environmental conditions.

Figure 6. Operator exposure to the environmental condition on site. The red line refers to the threshold limit value (TLV).

It is relevant to identify different clusters of behavior from the collected data as shown in Figure 7. In fact, this Figure presents a non-linear two dimensional projection from the originally 12-dimensional space of fused data from two different crane operators was embedded. It was decided to use the

UMAP projection technique to provide a clear vision regarding the sample distribution where the inter-distance between samples was maintained and good preservation of the data's global structure was granted [59].



(**a**) Segmented per period of time. (**b**) Segmented per Operator.

Figure 7. Projected patterns from the digital twin.

The projection result, shown in Figure 7, depicts the whole dataset's distributions and clusters phenomenon. Then, we decided to segment the whole dataset into five groups of the same length, to explore time dependency of clusters (see Figure 7a). The five groups were segmented via homogeneous distribution through time, and marked with different colors (labels 1–5). As far as the same color appears in different clusters, that means that they are not related to time mainly, but related to operator's behaviour. In Figure 7b, the behaviors from the two monitored crane operators are presented. The projection of the whole dataset was also segmented into two groups: operator 1 and operator 2 marked with two different colors. It becomes clear that there is not operator specificity as per cluster, therefore the existence of many of them is due to intrinsic reasons. On the other side, different operators behaved in different ways, as different colors appeared in different clusters, although similarities due to the common activities were also there. Indeed, there was room to identify each of the clusters and to enrich the collection with larger datasets and, thus, derive knowledge from such clusters. The evolution of operator health such as stress level can be identified via clusters and amount of time on each, which makes it possible to adapt individual healthy requirement from management point of view. The objective measures can be taken, e.g., to allocate specific operator to less demanding trucks in order to provide better wellness or convenient recovery process.

The interest in the prototype was also to evaluate its functionality to derive specific knowledge when considering digital twins. In particular, the interest was to analyze the stress levels and the impacts. We decided to focus on situations where large differences between high and low blood pressure happen. This is because no matter whether an operator can suffer hypertension, an unwanted situation related to high levels of stress will have locally larger differences between the high and the low pressures, and the risk for cardiovascular disease, diabetes, chronic kidney disease, high cholesterol levels, and aneurysms will be larger [60]. A segmentation in different discrete categories between Low to Intense for the original variables as well for the derived ones (mainly differences and ranges), was carried out, including ΔBloodPressure, where the threshold for a large difference was adopted to be 45 mmHg. After the data preparation an a priori algorithm for the association rules discovery was applied, which provided some rules like those presented in Table 2.

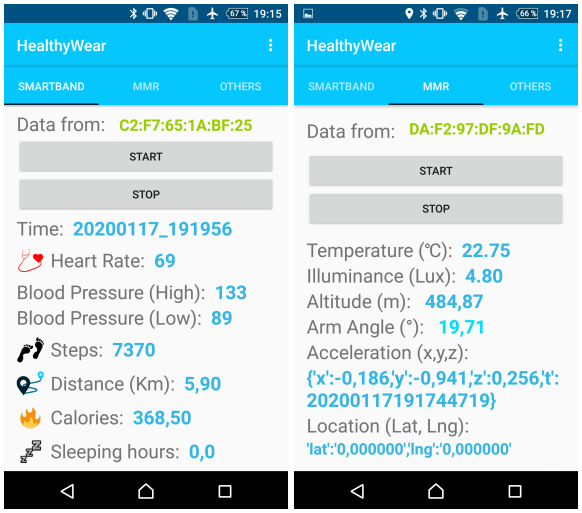
Table 2. The derived rules from the application of machine learning techniques looking at high differences in blood pressure and explaining the intense values.

Rule Antecedent	Confidence	Support	Lift
'Arm angle:UP FRONT' AND 'High Blood Pressure:Intense' AND 'Heart Rate:Moderate'	1.0	5.9%	3.86
'Sound Level:Noisy' AND steps:Low AND Arm angle:UP FRONT AND 'High Blood Pressure:Intense'	1.0	6.27%	3.86
'steps:Low' AND 'HCHO:Relative Moderate' AND 'High Blood Pressure:Intense'	0.84	10.0%	3.85

Only relevant rules were highlighted as the algorithm discovered many others related to normal conditions, including normal environmental conditions and normal body indicators, which imply a normal range for ΔBloodPressure, etc. All of them had strong support and moderately high confidence however the lift values were close to 1. Indeed, the biggest interest was to check the rules with high confidence values, high lift figures, and moderate or low support. Those rules became excellent candidates to create knowledge (maximum confidence); however, due to their relatively low support, they remain mostly unknown.

During the validation of this prototype there was an unexpected high interest from the operators to obtain the knowledge about their parameters (to learn from their digital twins), and as far as it was possible and it was understood it could also provide benefits to the operators as they can become more aware of the reasons for their parameters; therefore, a significant effort was conducted to address the increase of awareness.

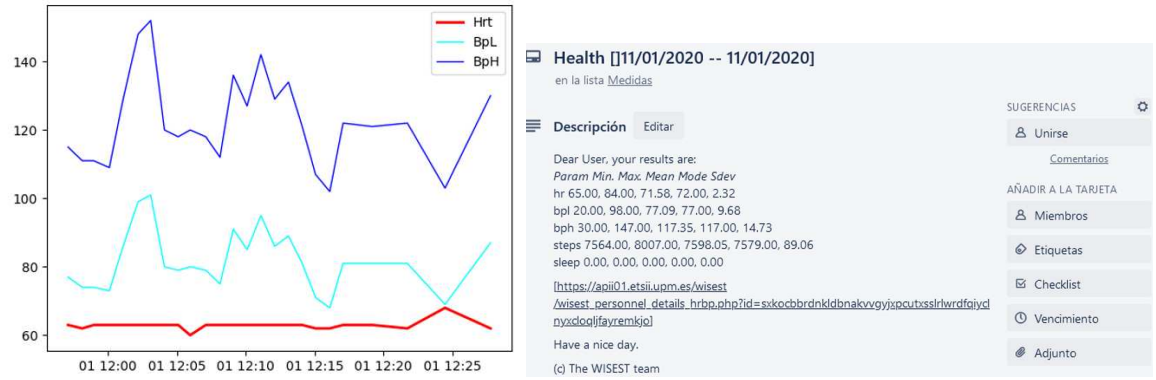
The operator were able to have the real-time information of their health parameters, as shown in Figure 8.



(**a**) Biometric parameters. (**b**) Ergonomic parameters.

Figure 8. Informative phone screens.

When lean management techniques are being used at the organization level, it can also help with the management of processes to communicate to the operators their figures on a daily basis, as described above for the $(CPD)_nA$ method [17]. Therefore, an application for the trello™system was created at the Cognition layer, where the detailed messages can be seen, as in Figure 9, bringing a way to deeply dig into the specific graphics Figure 10, making it possible to promote the operator self-learning process.



(**a**) Graphical view sent to user account. (**b**) Summary screen sent as message to trello.

Figure 9. Informative user phone screens where time granularity can be configured.

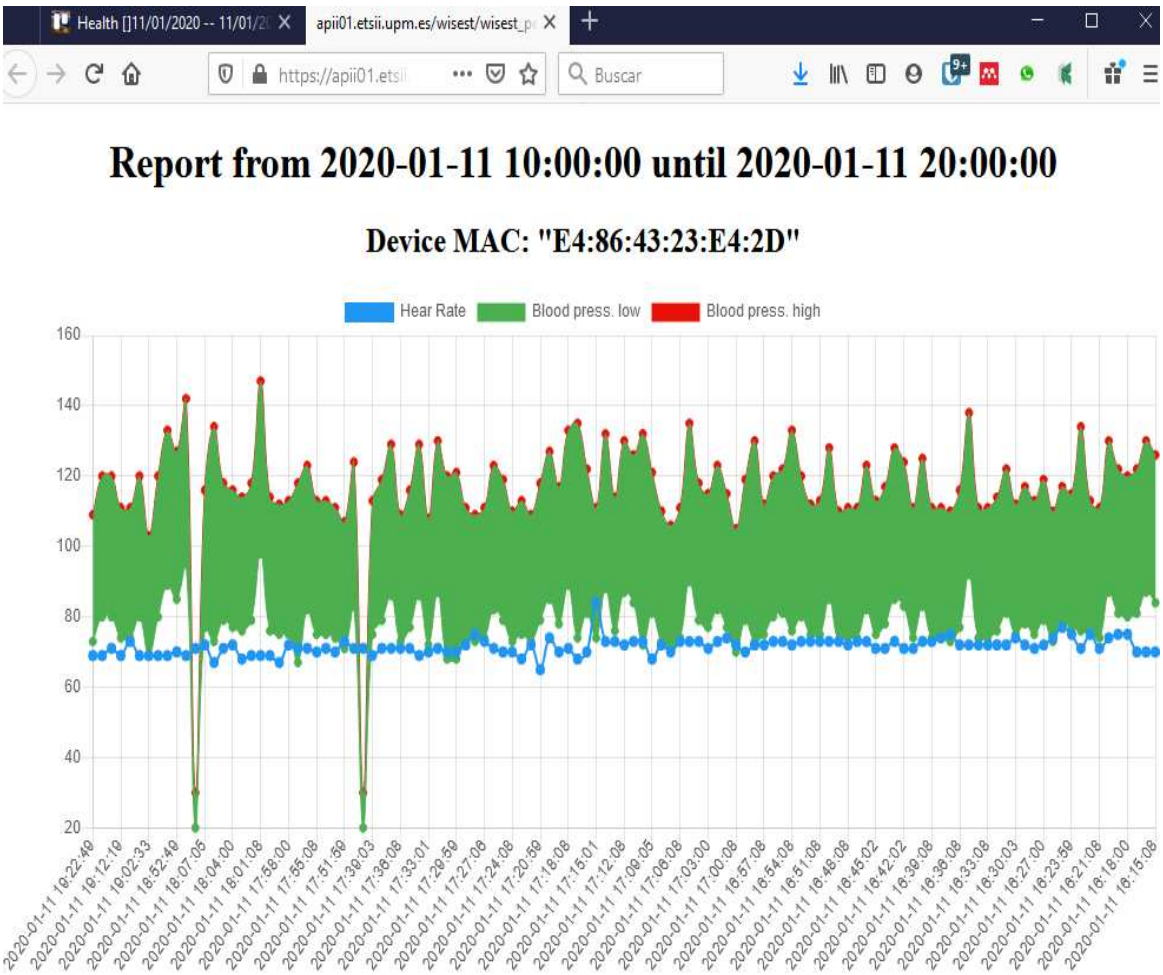


Figure 10. The graphical detailed view of the collected data at an individual level.

6. Discussion

The implementation of a prototype to validate the HO4.0 framework was performed successfully, where the layer structure helped in providing guidance to the required discussion in order to build the system. Data interconnection and interoperation further facilitated the making of health action decisions, aiming to improve the operator working conditions, and the strategies and actions were built on semantic-based data driver strategies.

The data from the sensing layer was collected via the data networking layer and was multimodal and heterogeneous when considering the different sensing instruments and data collection solutions. To make decisions, especially for the long-term policy in terms of health management, the data coming from different domains have to be interoperable and linked. We also needed to emphasize the meaning of each type of data streams, as well as the frequency of sampling, aligned with the relevant business model.

To further support decision making for health operator applications, different ontologies were applied in this study to make the data interoperable and semantically accessible while crossing domains. Although the already presented combination was retained for this prototype, there was no limitation other than to properly describe the entities and their inter-relationships to maintain a standard domain vocabulary and common understanding. Therefore, they can also be created on purpose. The metadata languages aligning with the ontology vocabulary convey that kind of shared information in a landscape of information management.

With the data inter-operation features and semantic annotation, a holistic view can be built up to enable the machine learning application to create new knowledge, as already presented. This can be operationalized and communicated looking to improve the operator health and safety.

Additionally, it is possible to combine the organizational analysis based on patterns with the additional value created from the operator level, by allowing them to be aware of their own values. This is another useful dimension of the HO4.0 concept, as awareness is a key element for empowerment [61].

7. Conclusions

Empowering people has significant implications for factories under the Industry 4.0 paradigm. Health, as one of main concerns affects operator's job satisfaction and well-being. However, it was rarely investigated in the context of HCPSs. Aiming to reduce this gap, holistic view of operator health dimension was proposed, where the main contributions of this paper are summarized as follows:

- We explored the HO4.0 concept.
- Based on the classical five-level CPS structure, we proposed a four-layer framework for the implementation of the HO4.0 concept.
- We investigated of the enabling technologies for HO4.0.
- We developed a prototype system to showcase the application process of the proposed framework.
- We applied advanced machine learning and preparation with semantic engineering technologies to manage and analyze heterogeneous data from different sources, bringing this support to the digital twin perspective.

The research scope is not limited to the outcome of conducted experiments, but to provide an integrated framework providing vertical and horizontal integration capabilities. The proposed architecture supports easy integration of different data sources as from Industrial Internet of Things, wearables such as smart helmet, smart insoles, smart glasses, to name a few. From a practical perspective, it also successfully addresses the need for data integration from different devices, as manufacturers aim to provide their own data lakes, which are not flexible enough to allow user defined sampling frequency or convenient restful based data access. In addition, enhancing the capability for monitoring different devices with the same middleware (phone, tablet, etc.) including reconnecting capabilities, makes the data collection process stronger, which is really valuable in industrial environments.

The gathered data sources with the defined data fusion techniques, including semantic interoperability, produce multidimensional datasets, which are suitable to be analyzed in real time with advanced modeling technologies (big analytics). By learning from operators' daily activities and parameters, inference about their evolution in long term perspective can be derived through their digital-twins. Such approach can have managerial value as adopted strategies can relay on transparent

measurement system and they can better preserve knowledge, experience, and well-being for longer, as well as company enlarges its social sustainability dimension.

In addition, it was also possible to realize the high degree of interest that such applications open to workers themselves, and how it becomes a way of discussion between them, as they try to identify patterns and explanations. We have found it very interesting for using it in a smarter way of management, taking advantage of such interest in order to increase the empowerment and engagement of workers into their processes.

Further than that, based on the adopted methodology, we concluded that the designed framework for HO4.0, as depicted above, based on the HO4.0 concept, was adequate and easy to follow in the integration of different sources of data at any time.

As a limitation, the prototype was applied in only one company with a wide range of operator profiles, but where the size and culture can also have an impact, and these dimensions were not analyzed. Another limitation is based on appropriate sensor availability, as there are safety regulations that can limit the application of specific sensors, and there are a lack of sensors for certain tasks, such as human concentration. However, the range of applications was very large as the integration between the operator's health and smart monitoring is just starting.

As a further research ambition, the prototype will be extended by integrating weight measurement through smart insoles, giving the information of the loads being carried by operators along the working time. This will be implemented by extending the 'Other devices' option from the Android app, following the created implementation rules. Indeed, further research is needed to better categorize the identified clusters, its distribution and stability against sample size. Another research line is connected to the information that could be derived from the digital-twins when more data populate them. It seems to be very promising. Finally, but yet importantly, to connect the HO4.0 with production processes can bring a truly integrated management perspective which demands further investigations.

Author Contributions: Data curation, S.S.; Investigation, S.S., X.Z. and J.O.-M.; Methodology, X.Z. and S.S.; Project administration, J.O.-M.; Resources, J.O.-M.; Software, J.G.P.; Validation, B.G.; Visualization, B.G.; Writing—original draft, X.Z., J.O.-M. and S.S.; Writing—review and editing, B.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by the European Commission through the RFCS program, grant number with ID 793505 (WISEST project), and by the Spanish Agencia Estatal de Investigacion, through the research project with code RTI2018-094614-B-I00.

Acknowledgments: The authors thank the EU RFCS research program for the support of this research through the grant with ID 793505 (WISEST project) as well as the industrial partners. Indeed, Joaquín Ordieres-Meré wants to acknowledge the Spanish Agencia Estatal de Investigacion, through the research project with code RTI2018-094614-B-I00 into the "Programa Estatal de I+D+i Orientada a los Retos de la Sociedad".

Conflicts of Interest: The authors declare no conflict of interest. The founders did not have any influence either in selecting the topic or in the experiment design.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence.
CIM	Computer Integrated Manufacturing.
CPS	Cyber–Physical Systems.
ECG	Electrocardiogram.
EEG	Electroencephalography.
EMG	Electromyography.
GPS	Positioning System.
HCPSs	Human-Cyber–Physical Systems.
HO4.0	Healthy Operator 4.0.
IAQ	Indoor Air Quality.
IIoT	Industrial Internet of Things.
IoT	Internet of Things.

IPA	Intelligent Personal Assistant.
IPS	Indoor Positioning System.
IR	Infrared Receiver.
IT	Information Technology.
MAC	Media Access Control
ML	Machine Learning
OSH	Occupational Safety and Health
RFID	Radio Frequency IDentification.
SBC	Single-Board Computer.
TLV	Threshold Limit Value
UWB	Ultra-wideband.
VR	Virtual Reality.
WLAN	wireless local area network.

References

1. Schmidt, R.; Möhring, M.; Härting, R.C.; Reichstein, C.; Neumaier, P.; Jozinović, P. Industry 4.0-potentials for creating smart products: empirical research results. In *International Conference on Business Information Systems*; Springer: London, UK, 2015; pp. 16–27.
2. Pereira, A.; Romero, F. A review of the meanings and the implications of the Industry 4.0 concept. *Procedia Manuf.* **2017**, *13*, 1206–1214. [CrossRef]
3. Zheng, X.; Sun, S.; Mukkamala, R.R.; Vatrapu, R.; Ordieres-Meré, J. Accelerating health data sharing: A solution based on the internet of things and distributed ledger technologies. *J. Med. Internet. Res.* **2019**, *21*, e13583. doi:10.2196/13583. [CrossRef] [PubMed]
4. Lee, J.; Bagheri, B.; Kao, H.A. A cyber-physical systems architecture for industry 4.0-based manufacturing systems. *Manuf. Lett.* **2015**, *3*, 18–23. [CrossRef]
5. Chen, S.L.; Chen, Y.Y.; Hsu, C. A new approach to integrate internet-of-things and software-as-a-service model for logistic systems: A case study. *Sensors* **2014**, *14*, 6144–6164. [CrossRef] [PubMed]
6. Glaessgen, E.; Stargel, D. The Digital Twin Paradigm for Future NASA and US Air Force Vehicles. In Proceedings of the 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 20th AIAA/ASME/AHS Adaptive Structures Conference 14th AIAA, Honolulu, HI, USA, 23–26 April 2012; p. 1818.
7. Boschert, S.; Rosen, R. Digital twin—The simulation aspect. In *Mechatronic Futures*; Springer: London, UK, 2016; pp. 59–74.
8. Hao, Y.; Helo, P. The role of wearable devices in meeting the needs of cloud manufacturing: A case study. *Robot. Comput.-Integr. Manuf.* **2017**, *45*, 168–179. [CrossRef]
9. Zheng, X.; Wang, M.; Ordieres-Meré, J. Comparison of data preprocessing approaches for applying deep learning to human activity recognition in the context of industry 4.0. *Sensors* **2018**, *18*, 2146. [CrossRef] [PubMed]
10. Putnik, G. Advanced manufacturing systems and enterprises: Cloud and ubiquitous manufacturing and an architecture. *J. Appl. Eng. Sci.* **2012**, *10*, 127–134.
11. Gorecky, D.; Schmitt, M.; Loskyll, M.; Zühlke, D. Human-Machine-Interaction in the Industry 4.0 Era. In Proceedings of the 2014 12th IEEE International Conference on Industrial Informatics (INDIN), Porto Alegre, RS, Brazil, 27–30 July 2014; pp. 289–294.
12. Romero, D.; Stahre, J.; Wuest, T.; Noran, O.; Bernus, P.; Fast-Berglund, Å.; Gorecky, D. Towards an Operator 4.0 Typology: A Human-Centric Perspective on the Fourth Industrial Revolution Technologies. In Proceedings of the International Conference on Computers and Industrial Engineering (CIE46) Proceedings, Tianjin, China, 29–31 October 2016.
13. Ruppert, T.; Jaskó, S.; Holczinger, T.; Abonyi, J. Enabling technologies for operator 4.0: A survey. *Appl. Sci.* **2018**, *8*, 1650. [CrossRef]
14. Lorenz, M.; Rüßmann, M.; Strack, R.; Lueth, K.L.; Bolle, M. *Man and Machine in Industry 4.0: How Will Technology Transform the Industrial Workforce through 2025*; Boston Consulting Group: Boston, MA, USA, 2015; Volume 2.

15. Romero, D.; Bernus, P.; Noran, O.; Stahre, J.; Fast-Berglund, Å. The operator 4.0: Human cyber-physical systems & adaptive automation towards human-automation symbiosis work systems. In *IFIP International Conference on Advances in Production Management Systems*; Springer: London, UK, 2016; pp. 677–686.
16. Tzafestas, S. Concerning human-automation symbiosis in the society and the nature. *Int. J. Fact. Autom. Robot. Soft Comput.* **2006**, *1*, 16–24.
17. Villalba-Díez, J.; Ordieres-Meré, J. Improving manufacturing operational performance by standardizing process management. *Trans. Eng. Manag.* **2015**, *62*, 351–360. [CrossRef]
18. Horváth, L.; Rudas, I.J. Role of Information Content in Multipurpose Virtual Engineering Space. In Proceedings of the 2017 IEEE 15th International Symposium on Applied Machine Intelligence and Informatics (SAMI), Her'lany, Slovakia, 26–28 January 2017; pp. 000099–000104.
19. Sylla, N.; Bonnet, V.; Colledani, F.; Fraisse, P. Ergonomic contribution of ABLE exoskeleton in automotive industry. *Int. J. Ind. Ergon.* **2014**, *44*, 475–481. [CrossRef]
20. Myers, K.; Berry, P.; Blythe, J.; Conley, K.; Gervasio, M.; McGuinness, D.L.; Morley, D.; Pfeffer, A.; Pollack, M.; Tambe, M. An intelligent personal assistant for task and time management. *AI Mag.* **2007**, *28*, 47.
21. Romero, D.; Mattsson, S.; Fast-Berglund, Å.; Wuest, T.; Gorecky, D.; Stahre, J. Digitalizing occupational health, safety and productivity for the operator 4.0. In *IFIP International Conference on Advances in Production Management Systems*; Springer: London, UK, 2018; pp. 473–481.
22. Buffet, M.; Gervais, R.; Liddle, M.; Eeckelaert, L. Well-being at work: Creating a positive work environment. In *Literature Review*; European Agency for Safety and Health at Work: Bilbao, Spain, 2013.
23. Salanova, M.; Del Líbano, M.; Llorens, S.; Schaufeli, W.B. Engaged, workaholic, burned-out or just 9-to-5? Toward a typology of employee well-being. *Stress Health* **2014**, *30*, 71–81. [CrossRef] [PubMed]
24. Vasic, M.; Billard, A. Safety Issues in Human-Robot Interactions. In Proceedings of the 2013 IEEE International Conference on Robotics and Automation, Karlsruhe, Germany, 6–10 May 2013; pp. 197–204.
25. Matthias, B.; Oberer-Treitz, S.; Staab, H.; Schuller, E.; Peldschus, S. Injury Risk Quantification for Industrial Robots in Collaborative Operation with Humans. In Proceedings of the ISR 2010 (41st International Symposium on Robotics) and ROBOTIK 2010 (6th German Conference on Robotics), Munich, Germany, 7–9 June 2010; pp. 1–6.
26. Page, K.M.; Vella-Brodrick, D.A. The 'what','why'and 'how'of employee well-being: A new model. *Soc. Indic. Res.* **2009**, *90*, 441–458. [CrossRef]
27. Podgorski, D.; Majchrzycka, K.; Dąbrowska, A.; Gralewicz, G.; Okrasa, M. Towards a conceptual framework of OSH risk management in smart working environments based on smart PPE, ambient intelligence and the Internet of Things technologies. *Int. J. Occup. Saf. Ergon.* **2017**, *23*, 1–20. [CrossRef]
28. Bernal, G.; Colombo, S.; Al Ai Baky, M.; Casalegno, F. Safety++: Designing IoT and Wearable Systems for Industrial Safety through a User Centered Design Approach. In Proceedings of the 10th International Conference on Pervasive Technologies Related to Assistive Environments, Rhodes, Greece, 21–23 June 2017; pp. 163–170.
29. Sun, S.; Zheng, X.; Villalba-Díez, J.; Ordieres-Meré, J. Indoor air-quality data-monitoring system: Long-term monitoring benefits. *Sensors* **2019**, *19*, 4157. [CrossRef]
30. Kassner, L.; Hirmer, P.; Wieland, M.; Steimle, F.; Königsberger, J.; Mitschang, B. The Social Factory: Connecting People, Machines and Data in Manufacturing for Context-Aware Exception Escalation. In Proceedings of the 50th Hawaii International Conference on System Sciences, Village, HI, USA, 4–7 January 2017.
31. Romero, D.; Wuest, T.; Stahre, J.; Gorecky, D. Social factory architecture: social networking services and production scenarios through the social internet of things, services and people for the social operator 4.0. In *IFIP International Conference on Advances in Production Management Systems*; Springer: London, UK, 2017; pp. 265–273.
32. Pavón, I.; Sigcha, L.; Arezes, P.; Costa, N.; de Arcas, G.; Lopez-Navarro, J. Wearable technology for occupational risk assessment: Potential avenues for applications. *Occup. Saf. Hyg. VI* **2018**, 447–452.
33. Perera, C.; Liu, C.H.; Jayawardena, S. The emerging internet of things marketplace from an industrial perspective: A survey. *IEEE Trans. Emerg. Top. Comput.* **2015**, *3*, 585–598. [CrossRef]
34. Bell, E.; Bryman, A.; Harley, B. *Business Research Methods*; Oxford University Press: Oxford, UK, 2018.
35. Mann, S. Wearable computing. In *The Encyclopedia of Human-Computer Interaction*, 2nd ed.; International Design Foundation: New York, NY, USA, 2013.

36. Mardonova, M.; Choi, Y. Review of wearable device technology and its applications to the mining industry. *Energies* **2018**, *11*, 547. [CrossRef]
37. Koyuncu, H.; Yang, S.H. A survey of indoor positioning and object locating systems. *IJCSNS Int. J. Comput. Sci. Netw. Secur.* **2010**, *10*, 121–128.
38. Liu, H.; Darabi, H.; Banerjee, P.; Liu, J. Survey of wireless indoor positioning techniques and systems. *IEEE Trans. Syst. Man Cybern. Part C (Appl. Rev.)* **2007**, *37*, 1067–1080. [CrossRef]
39. Zanni, S.; Lalli, F.; Foschi, E.; Bonoli, A.; Mantecchini, L. Indoor air quality real-time monitoring in airport terminal areas: An opportunity for sustainable management of micro-climatic parameters. *Sensors* **2018**, *18*, 3798. [CrossRef] [PubMed]
40. Zakaria, N.A.; Abidin, Z.Z.; Harum, N.; Hau, L.C.; Ali, N.S.; Jafar, F.A. Wireless Internet of Things-based air quality device for smart pollution monitoring. *Int. J. Adv. Comput. Sci. Appl.* **2018**, *9*, 65–69. [CrossRef]
41. Benammar, M.; Abdaoui, A.; Ahmad, S.; Touati, F.; Kadri, A. A modular IoT platform for real-time indoor air quality monitoring. *Sensors* **2018**, *18*, 581. [CrossRef]
42. Industrial Ontologies Foundry (IOF). Available online: https://sites.google.com/view/industrialontologies/home (accessed on 20 March 2020).
43. Sampath Kumar, V.R.; Khamis, A.; Fiorini, S.; Carbonera, J.; Olivares-Alarcos, A.; Habib, M.; Gonçalves, P.; Li, H.; Olszewska, J. Ontologies for Industry 4.0. *Knowl. Eng. Rev.* **2019**, *34*. doi:10.1017/S0269888919000109. [CrossRef]
44. Ward, J.A.; Lukowicz, P.; Troster, G.; Starner, T.E. Activity recognition of assembly tasks using body-worn microphones and accelerometers. *IEEE Trans. Pattern Anal. Mach. Intell.* **2006**, *28*, 1553–1567. [CrossRef]
45. Shoaib, M.; Bosch, S.; Incel, O.; Scholten, H.; Havinga, P. Complex human activity recognition using smartphone and wrist-worn motion sensors. *Sensors* **2016**, *16*, 426. [CrossRef]
46. Ordieres-Meré, J.; Prieto Remón, T.; Rubio, J. Digitalization: An opportunity for contributing to sustainability from knowledge creation. *Sustainability* **2020**, *12*, 1460. [CrossRef]
47. Macoir, N.; Bauwens, J.; Jooris, B.; Van Herbruggen, B.; Rossey, J.; Hoebeke, J.; De Poorter, E. Uwb localization with battery-powered wireless backbone for drone-based inventory management. *Sensors* **2019**, *19*, 467. [CrossRef]
48. Mobile Operating System Market Share Worldwide. Available online: https://gs.statcounter.com/os-market-share/mobile/worldwide (accessed on 20 March 2020).
49. García Paredes, J.; Ordieres-Meré, J. Healthy Operator 4.0. 2020. Available online: https://github.com/jbmere/HealthOperator4.0/releases (accessed on 20 March 2020).
50. vCard Ontology - for describing People and Organizations. Available online: https://www.w3.org/TR/vcard-rdf/ (accessed on 20 March 2020).
51. Oprea, M.M. AIR_POLLUTION_Onto: An ontology for air pollution analysis and control. In *Artificial Intelligence Applications and Innovations III*; Maglogiannis, I., Karpouzis, K., Eds.; Springer: London, UK, 2009; pp. 135–143.
52. Khoozani, E.N.; Hadzic, M. Designing the human stress ontology: A formal framework to capture and represent knowledge about human stress. *Aust. Psychol.* **2010**, *45*, 258–273. doi:10.1080/00050061003664811. [CrossRef]
53. Kun, D.P.; Varga, E.B.; Toth, Z. Ontology Based Navigation Model of the ILONA System. In Proceedings of the 2017 IEEE 15th International Symposium on Applied Machine Intelligence and Informatics (SAMI), Herl'any, Slovakia, 26–28 January 2017; pp. 479–484. doi:10.1109/SAMI.2017.7880357. [CrossRef]
54. Head, K.; Clarke, M.; Bailey, M.; Livinski, A.; Ludolph, R.; Singh, A. Report of the systematic review on the effect of indoor heat on health. In *WHO Housing and Health Guidelines*; World Health Organization: Geneva, Switzerland, 2018.
55. Noise Guidelines. Available online: http://www.euro.who.int/__data/assets/pdf_file/0008/383921/noise-guidelines-eng.pdf?ua=1 (accessed on 20 March 2020).
56. WHO Guidelines for Indoor Air Quality. Available online: http://www.euro.who.int/__data/assets/pdf_file/0009/128169/e94535.pdf (accessed on 20 March 2020).
57. NIOSH Pocket Guide to Chemical Hazards: Formaldehyde. Available online: https://www.cdc.gov/niosh/npg/npgd0293.html (accessed on 20 March 2020).
58. EnvCon App. Available online: https://itunes.apple.com/tr/app/envcon/id1045648840?mt=8 (accessed on 20 March 2020).

59. McInnes, L.; Healy, J.; Saul, N.; Grossberger, L. UMAP: Uniform Manifold Approximation and Projection. *J. Open Source Softw.* **2018**, *3*, 861. [CrossRef]
60. Soni, R.K.; Porter, A.C.; Lash, J.P.; Unruh, M.L. Health-related quality of life in hypertension, chronic kidney disease, and coexistent chronic health conditions. *Adv. Chronic Kidney Dis.* **2010**, *17*, e17–e26. [CrossRef] [PubMed]
61. Konijn, A.M.; Lay, A.M.; Boot, C.R.; Smith, P.M. The effect of active and passive occupational health and safety (OHS) training on OHS awareness and empowerment to participate in injury prevention among workers in Ontario and British Columbia (Canada). *Saf. Sci.* **2018**, *108*, 286–291. [CrossRef]

sensors

MDPI

Article

Industry 4.0 Lean Shopfloor Management Characterization Using EEG Sensors and Deep Learning

Daniel Schmidt [1,2,*,†], Javier Villalba Diez [1,3,4,*,†], Joaquín Ordieres-Meré [1], Roman Gevers [2], Joerg Schwiep [2] and Martin Molina [4]

1 Department of Business Intelligence, Escuela Técnica Superior de Ingenieros Industriales, Universidad Politécnica de Madrid, 28006 Madrid, Spain; j.ordieres@upm.es
2 Matthews International GmbH, Gutenbergstraße 1-3, 48691 Vreden, Germany; roman.gevers@saueressig.de (R.G.); joerg.schwiep@saueressig.de (J.S.)
3 Hochschule Heilbronn, Fakultät Management und Vertrieb, Campus Schwäbisch Hall, 74523 Schwäbisch Hall, Germany
4 Department of Artificial Intelligence, Escuela Técnica Superior de Ingenieros Informáticos, Universidad Politécnica de Madrid, 28660 Boadilla del Monte, Madrid, Spain; martin.molina@upm.es
* Correspondence: Daniel.Schmidt@saueressig.de (D.S.); javier.villalba-diez@hs-heilbronn.de (J.V.D.)
† These authors contributed equally to the work.

Received: 30 April 2020; Accepted: 12 May 2020; Published: 18 May 2020

Abstract: Achieving the shift towards Industry 4.0 is only feasible through the active integration of the shopfloor into the transformation process. Several shopfloor management (SM) systems can aid this conversion. They form two major factions. The first includes methodologies such as Balanced Scorecard (BSC). A defining feature is rigid structures to fixate on pre-defined goals. Other SM strategies instead concentrate on continuous improvement by giving directions. An example of this group is the "HOSHIN KANRI TREE" (HKT). One way of analyzing the dissimilarities, the advantages and disadvantages of these groups, is to examine the neurological patterns of workers as they are applying these. This paper aims to achieve this evaluation through non-invasive electroencephalography (EEG) sensors, which capture the electrical activity of the brain. A deep learning (DL) soft sensor is used to classify the recorded data with an accuracy of 96.5%. Through this result and an analysis using the correlations of the EEG signals, it has been possible to detect relevant characteristics and differences in the brain's activity. In conclusion, these findings are expected to help assess SM systems and give guidance to Industry 4.0 leaders.

Keywords: EEG sensors; manufacturing systems; shopfloor management; machine learning; deep learning

1. Introduction

The concept of Industry 4.0 tries to tackle many of the forthcoming challenges that get faced by business leaders in the 21st century [1,2]. Some may see it as focused on technological solutions for higher automation enabled by digitalization and the resulting possibilities [3]. Yet the full potential is only achievable through the holistic perspective of a sociotechnical system [4,5]. This stands in stark contrast to the extreme technical view of Computer Integrated Manufacturing (CIM) [6]. In the sociotechnical system, the technology and the workers are both viewed as interconnected parts. Because of this, placing a prime focus on the complex interaction between these two groups is necessary [7]. This presupposes that organizations will change old structures and adapt to the approaching demands [8]. The challenge presented here, which is closely linked to the research

question, is whether artificial intelligence is capable of providing future leaders with the tools necessary to understand the complex dynamics that occur in these Industry 4.0 environments.

The purpose of this work is to provide Industry 4.0 leaders with a better understanding of Lean Shopfloor Management (SM) methods through artificial intelligence techniques applied to information collected from portable devices that provide an electroencephalographic (EEG) signal. For this the initial hypothesis is that deep learning (DL) algorithms are capable of characterizing and discerning between different types of behavior, once the EEG signal has been properly treated. This is shown in the graphical abstract of the paper depicted in Figure 1. More specifically, the work aims to offer insights into the two major categories of SM systems. This is done through the study of the neurological activity of process owners and their leaders performing either method. Non-invasive EEG sensors capture the electrical activity of the brain. A DL soft sensor is used to categorize the data. If the hypothesis is correct, then this would confirm that distinct contrasts in the brain activity during the conduction of the different SM systems exist. Furthermore, the correlations of the sensor channels are compared. These correspond to brain regions and show existing differences.

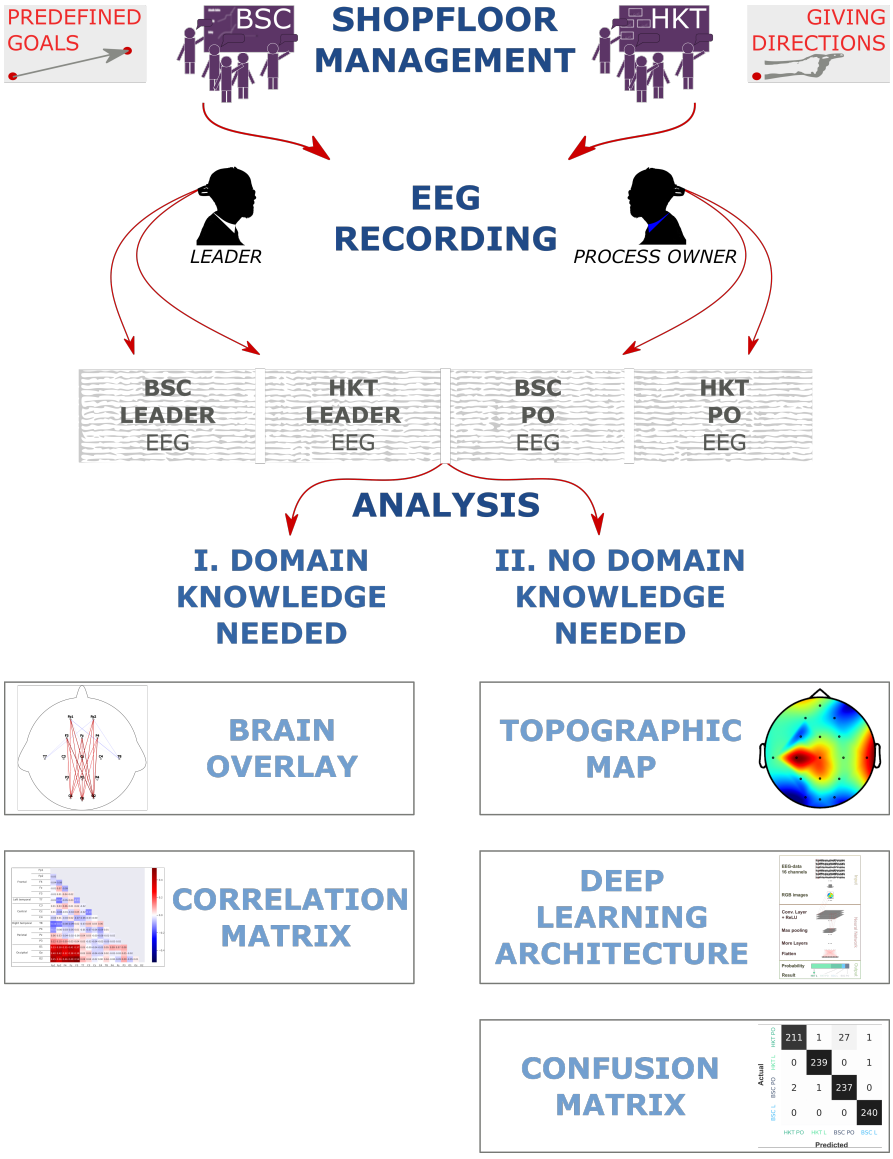


Figure 1. Graphical Abstract.

In the following the structure of the paper is described. Section 2 provides a background to the lean SM methodologies which presents fundamental preliminary concepts to the comprehensive understanding of the presented content in the following sections of this work. Section 3 provides the neurophysiological background in the context of SM, necessary to understand the EEG classification performed by the DL algorithms. Section 4 demonstrates the used materials and methods. The results attained are discussed in Section 5 and the implications in section 6 should be able to help judge disparate SM systems and give guidance to Industry 4.0 leaders.

2. Background

In order to achieve this goal, following points have received close attention. These and the rest of this section are, apart from the further description of the paper, taken from the conference paper The HOSHIN KANRI TREE. Cross-Plant Lean Shopfloor Management [9].

The points are:

- the need to sustainably empower the workforce (Learning Factory) (LF) as indicated by Narkhede et al. [10],
- the need to develop an autonomous and intelligent process management (PM) (Smart Factory) (SF) as presented by Lee et al. [11],
- the need to cope with increasing complexity of value-stream networks (VSN) as researched by Schuh et al. [12],
- the necessary paradigm shift towards strategic alignment as pointed out by Covey [13].

In the context of LF and SF, empowerment can be understood as a systematic way of learning that enables continuous improvement in an autonomous, intelligent, self-organized, and systematic manner. Coleman [14] defines empowerment as "the act of enhancing, supporting or not obstructing another's ability to bring about outcomes that he or she seeks." An "autonomous" management method ought to be able to function without a centralized controlling force by empowering all organizational elements. An "intelligent" management method ought to sustainably empower all organizational individuals to align and grow in the direction where value comes from for the organization.

A powerful paradigm to empower organizations focuses on value creation has flourished in the last two decades as lean management (LM). LM has been declared to be the industrial paradigm of the 21st century by Shah and Ward [15]. It is believed that the *problems* that LM endeavors to solve, the non-value adding activities, are embedded within processes, and therefore the response-able process owners (PO) that manage them are in charge of eliminating the non-value adding activities within them. Thus, the task set by LM is mainly a process management task and not a problem-solving one. Each individual of the organization is understood to be a PO, who is acting on his or her process on the shop floor.

This also enhances the need to act autonomously to make fast and flexible adjustments. Nonetheless, it is necessary to be aligned in the same direction (HOSHIN) by the strategic goals of the organizations. This leads to the fact that a balancing act between empowerment and alignment towards strategic goals is needed. According to Frow et al. [16] this makes multiple controls necessary. HOSHIN KANRI (HK) (management by giving direction) is a comprehensive management system that enables such alignment of complex systems as shown by Jolayemi [17].

Many of these points are in the general focus of SM. The term "shopfloor" has been used by western scholars, de Leeuw and van der Berg [18], to refer to processes close to production or distribution, excluding purposely strategic processes. In this sense, SM can be understood as a management system that can enhance shopfloor performance. The term "shopfloor" is used by Japanese scholars, Suzaki [19], in a broader sense, understanding "shopfloor" or "gemba" as the place, physical or virtual, where the value stream (VS) is performed. The definition of VS that Womack and Jones [20] gives is a "sequence of activities required to design, produce, and provide a specific good or service, and along which information, materials, and worth flows."

Because of the before mentioned challenges, efficient SM systems will be essential in the transformation process towards Industry 4.0 [21]. Thus, it is of chief benefit to rate these options. Earlier papers have focused on the implementation effects of case studies [22–25] or used theoretical considerations [26]. A big detriment is the enormous effort to measure the direct influence of the management strategies. For a comparison, similar starting positions are necessary. Only then are the outcomes contrastable. The many influencing factors make it challenging to get meaningful results from only a few comparisons. So, alternative concepts that produce more comparable findings are of immense value. Examining the brain activity during the practice of the SM methods opens up unique possibilities.

Earlier research demonstrates that significant neurological variations of a PO using different LM techniques such as KATA and (CPD)nA exist [27]. The aimed added value through this paper is manifold. Instead of only looking at the brain activity of the PO, this research further takes into account the interaction with his supervisor (in the following "leader") by recording the EEG data of both. It also concentrates on the differences of various *SM systems*. LM and SM have a large overlap, as SM uses many LM methods and tools. The primary distinction is that SM focuses on the aspects of leading and empowering people on the shopfloor [28], the place where the value creation occurs. Furthermore, an alternative way of pre-processing the EEG data for the DL soft sensor is implemented.

3. Literature Review

Much is still unknown about how the brain functions and above all the human brain. Despite this, many remarkable discoveries of the recent years promote the understanding [29,30]. These findings can help Industry 4.0 leaders in distinct ways. They make it easier to understand which factors are most important to foster progress on the shopfloor. This is attributed to the circumstance that the crucial component in the advancement of a company continues to be using the full potential of the employees [31].

The essential element for continuous improvement through the workers is the ability to form internal goals and to pursue these. In this context, the prefrontal cortex (PFC) has been identified as the most important region of the brain that contributes to this [32]. It is the center for cognitive control and makes it possible to act flexible to the outer world. For automatic "*hardwired*" behaviors, it doesn't play any significant role. This leads to the conclusion that substantial activity in the PFC should be expected and seems to be a requirement for POs practicing an SM system.

Another critical factor is the cooperation between workers to achieve improvements. Understanding the mental state of others is a prerequisite. In this context, there is sound evidence that not only the PFC plays a significant task. Likewise, the left and right temporoparietal junctions (TPJ) are indispensable. Though the roles they play seem to differ. While the left appears to be involved in strategic planning of choices concerning humans [33], the right plays a pivotal role in empathy, sympathy and perception [34,35].

Next to the PFC, the right TPJ additionally seems to play an important role for attention shifting [36]. This is a fundamental aspect in an Industry 4.0 setting, as every part of a manufacturing process is intertwined with other processes. These need to be put into consideration during any change process. Hence it is necessary to be capable of moving the mental focus.

To analyze the differences between SM systems, two distinct groups can be identified by focusing on the goal achievement that lays at the core of each management system.

1. Focus on pre-defined goals

   Through pre-defined goals, the focus is set on finding ways to achieve these. This is done through specific key figures, that in the best case give a balanced view on the different achievements or KPIs [37–39].
2. Continuous improvement by giving directions

This group of SM systems only provides a direction (HOSHIN) of improvement. A pre-defined goal is not set. Through this, improvements are approached in a more agile form and can be adapted along the road [40–43].

To narrow down the further analysis, one example was chosen for each of the two categories. The Balanced Scorecard (BSC) [44] as a representation of SM systems with pre-defined goals and the Hoshin Kanri TREE (HKT) representing the focus on continuous improvement by giving directions [45].

1. *Balanced Scorecard* is a SM system first described by Kaplan et al. [44] in 1992. The prime goal is to enable a *balanced* view on the driving measures of a business. This works by showing a handful of measurements that allow managers to interpret the complex interactions. Every measurement receives a specific target to motivate the employees to achieve this state. This is in contrast to more traditional approaches. A focus was only set on a few financial performance numbers. These only give a very short-sighted glance on the actual competitiveness of a company.

   One example of a balanced scorecard implementation is visible in Figure 2a showing the measures of **S**afety, **Q**uality, **D**elivery and **V**alue (*SQDV*). There are many variants in circulation such as **S**afety, **Q**uality, **D**elivery and **C**ustomer (*SQDC*) or **Q**uality, **D**elivery, **I**nventory, and **P**roductivity (*QDIP*) that can also show the priorities of a company by including or excluding specific categories such as **E**nvironment or **S**afety.

   Neely describes the standard way of using *BSC* [46] through the following steps:

   I **Check the current performance.** See how the development is progressing.
   II **Communicate performance.** Bring everyone to the same understanding of the current state.
   III **Confirm priorities.** Align the actions needed to improve the performance.
   IV **Compel progress.** Systematically achieve better performances.

   The prime aim is to measure and communicate the achievements towards **predetermined** goals [47]. Niven summarizes balanced scorecard as a conversation tool, a measurement system and a strategic management system [48].
2. *HKT* [45] in contrast is an example of a SM system that focuses on continuous improvement without pre-defined goals. A key feature is the standardization of communication between *PO*s in organizations using the *(CPD)nA* framework. Based on this, a feedback empowerment loop is implemented which makes it possible to build a sustainable process development.

   An implementation of the *HKT* can be seen in Figure 2b. The standard procedure encompasses the following steps as described by Villalba-Diez [49]:

   I **Evaluating progress.** Every PO checks if an improvement has been made to his or her *KPI* and places either a red or green magnet on his *(CPD)nA*.
   II **Reporting progress.** Only when the *KPI* receives a red magnet, does the progress need to be announced. The others can choose.
   III **(CPD)nA**. Every reporting PO ought to follow the *(CPD)nA* behavioral pattern.
   IV **Shopfloor visit.** One of the *(CPD)nA*s improvement is checked on site by the complete team.

As shown in Figure 2, a visual representation of both types of SM systems displays the main differences between them: BSC is basically depicting a set of key performance indicators as time series, and the HKT is representing a continuous improvement focused communication network between POs.

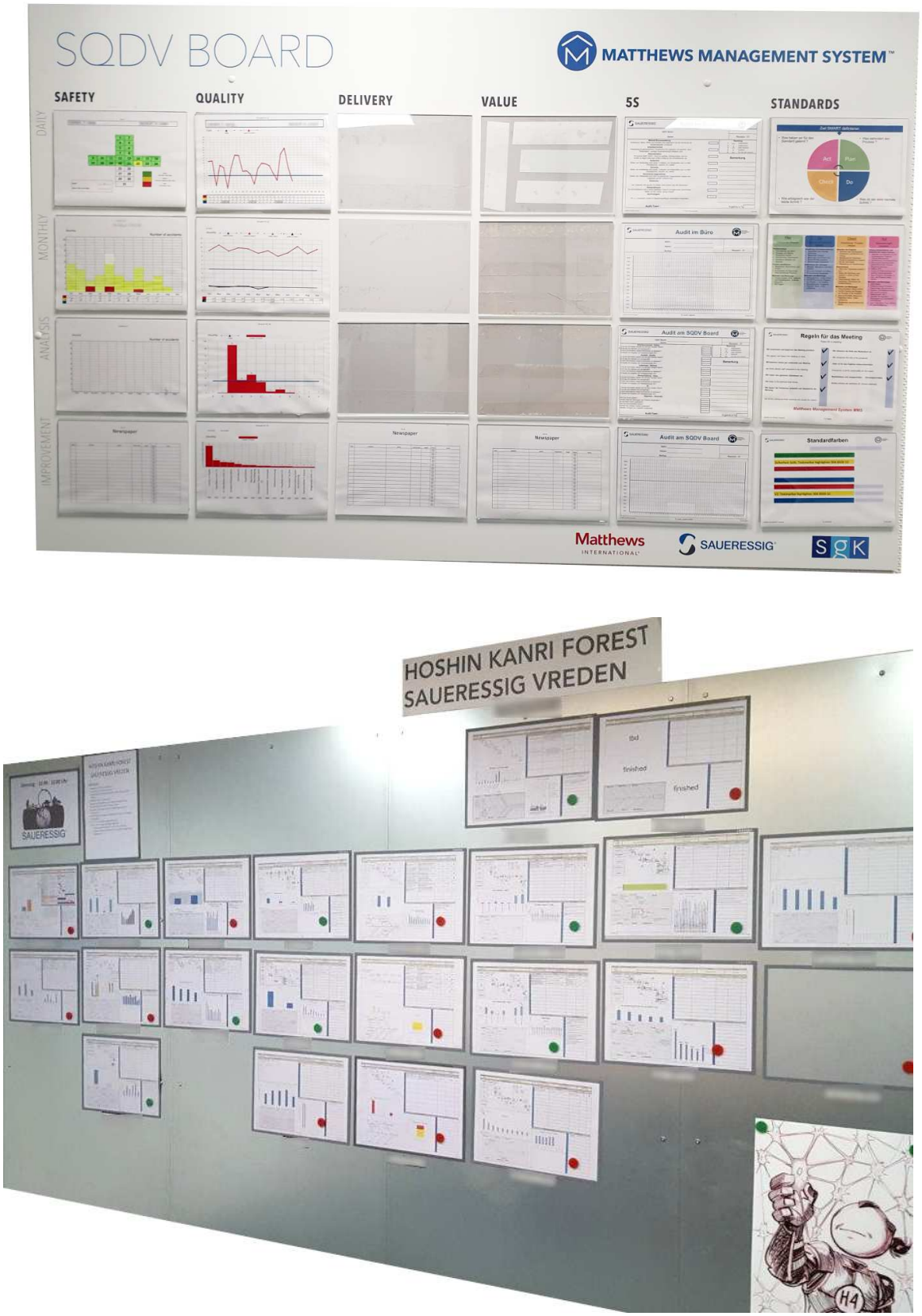


Figure 2. Shopfloor management systems. (**a**) SQDV (Safety, Quality, Delivery and Value) board at Matthews International GmbH, Vreden, Germany as an example of a balanced scorecard management system [44]. (**b**) Hoshin Kanri TREE [45] at Matthews International GmbH, Vreden, Germany. Process owner names are blurred out for privacy reasons.

To analyze these, an experimental setup using EEG recordings can be implemented. Depending on the results to be achieved with the EEG data, distinct methods of investigation are possible. A limited use case is the manual inspection for abnormalities in the EEG data, which can take a long time and presumes at least a basic level of domain knowledge [50]. The frequency analysis has been a more popular method that allows a wider range of applications [51–53].

Two distinct techniques are used in this research. These are the correlation function and a DL soft sensor.

1. Correlation Function

   The correlation function has found many use cases. It makes it feasible to classify EEG data [54], identify risk levels for developing schizophrenia [55], detect epileptic stages [56] or to classify EEG Motor imagery [57].

   With it, it is possible to determine the similarities between two signals and many application fields use it. In image processing it is for example used for template matching [58] or local image registration [59]. In geology for the location of earthquakes [60].

   The cross-correlation function works by sliding one signal along the other, calculating the product between the signals and finding the best fit [61]. Therefore, it is possible to work with time-shifted signals with the cross-correlation function.

2. Deep Learning

   In the last years DL has become a popular technique for analyzing EEG data and has been used to recognize emotions [62,63], detect Parkinson's [64] and Alzheimer's [65] disease, epileptic seizure prediction [66], the detection and diagnosis of seizures [67] or to decode and visualize the EEG data [68].

   An enormous advantage is that it can handle the complex EEG data with no prior domain knowledge, which allows a wider audience to work with this data. The neural network does this by *learning* the parameters to detect features from examples [69]. This has made it a popular choice for many other fields such as computer vision [70], audio processing [71] or bioinformatics [72]. The key challenge often hindering the further progress with neural networks is the limited data available. This is a prerequisite to represent a high range of input and parameters.

Our research aims to expand this approach on the characterization of complex LM shopfloor management associated behavioral patterns in an Industry 4.0 environment. In order to achieve this, this study outlines the following four research hypotheses (H) and their related LM interpretation shown in Table 1. Furthermore, as these hypotheses are based on neurophysiological expert knowledge, management needs to be provided with tools that allow a proper discernment of which behavior is followed, based only on the data. For this reason, a DL-based soft sensor is developed that is able to perform this task. The aim is to examine these with the mentioned methods.

Table 1. Hypotheses regarding the correlations of the recorded electroencephalography (EEG)-data.

#	Hypothesis	Interpretation
H1	The brain patterns of the leaders are expected to show strong correlations between the prefrontal-cortex and the occipital-cortex. This causes high correlations between the sensors *Fp1-Fp2-F3-Fz-F4* and *O1-Oz-O2*.	This result can be expected, as the leaders are listening for the majority of the time.
H2	In contrast to *H1*, the brain patterns of the process owners are expected to show significant correlations within the prefrontal-cortex and the occipital-cortex. This could be seen in strong correlations within the sensor groups *Fp1-Fp2-F3-Fz-F4* and *O1-Oz-O2*.	This result can be expected, as the process owner speaks for the majority of the time.
H3	Besides *H2*, all subjects are expected to show a high correlation of the prefrontal-cortex. This causes high correlations for the sensors *Fp1-Fp2-F3-Fz-F4*.	This could be understood in that way, that the conducted tasks are all executive behavioural patterns.
H4	The brain patterns of *HKT* practitioners are expected to show a strong correlation between the prefrontal-cortex and the TPJ. This could be seen by high correlations between the sensors *Fp1-Fp2-F3-Fz-F4* and *T7-T8* as well as *P3-P4*.	The interpretation is that *HKT* is a goal-oriented, context-independent *SM* problem-solving behavioral pattern.
H5	Compared to *HKT*, the brain patterns of *BSC* practitioners are expected to show a weak correlation between the prefrontal-cortex and the TPJ. This could be seen by low correlations between the sensors *Fp1-Fp2-F3-Fz-F4* and *T7-T8* and *P3-P4*.	This could be understood in that way, that *BSC* is a goal-oriented, context-dependent *SM* problem-solving behavioral pattern.

4. Materials and Methods

To test the hypotheses, a case study can provide meaningful first impressions if these are valid. Still it is necessary to note that a single study cannot give clear-cut proof. In the following the scope of the research is established Section 4.1, the population and sampling is specified Section 4.2, the data collection is further described Section 4.3, as well as the data pre-processing Section 4.4, the standardization procedure Section 4.5 and the data analysis Section 4.6.

4.1. Scope Establishment

For this study, the EEG data from a PO and his supervisor performing a *BSC* and a *HKT* implementation are recorded. All the recordings take place at one corporation. This is an automobile manufacturing facility based in Japan, where LM/SM methodologies were systematically implemented and accompanied by one author of this paper. The organization of the company can be seen through the HOSHIN KANRI FOREST STRUCTURE in Figure 3.

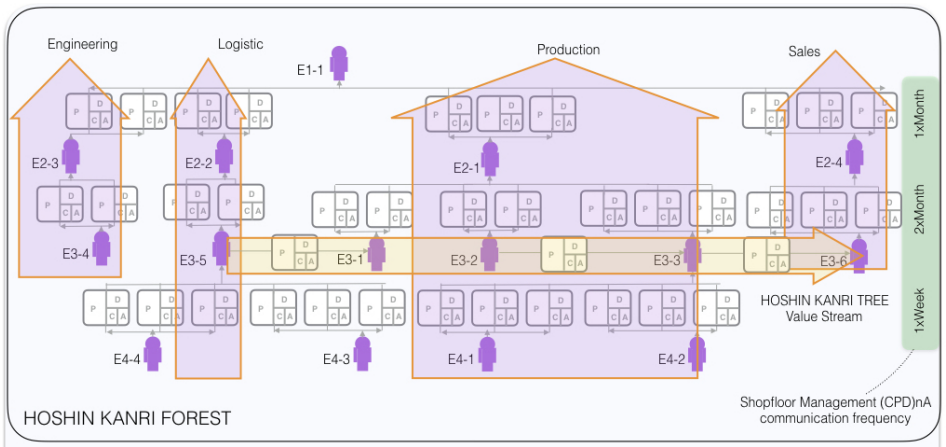


Figure 3. Part of the HOSHIN KANRI FOREST STRUCTURE [73].

4.2. Specifications of Population and Sampling

For this paper, data was collected from 14 subject pairs comprising a PO and his supervisor. The age range is between 20 and 60 years with a mean age of 40 years. As far as possible it was made sure that no significant neurological variations between the subjects should be expected. The subjects were male and had no history of any neurological or psychiatric disorder. Neither was any on chronic medication. All participants were left-hemisphere-dominant persons. Only between the PO and the leader differences could be possible that result from different levels of SM experience. Still, these are not expected to produce significant distinctions for the EEG recordings.

4.3. Data Collection

To record the EEG data, following 16 channels were chosen as standardized by the American Electroencephalographic Society [74]:

['Fp1', 'Fp2', 'F4', 'Fz', 'F3', 'T7', 'C3', 'Cz', 'C4', 'T8', 'P4', 'Pz', 'P3', 'O1', 'Oz', 'O2']

Through the regulated names, the respective locations on the head during the recording are defined. These can also be seen in Figure 4.

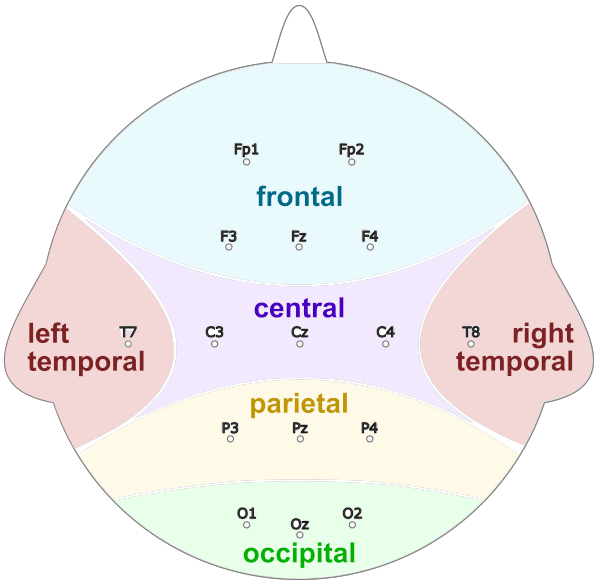


Figure 4. The 16 channels with standardized nomenclature that were recorded and the corresponding brain regions.

The used EEG sensor can be seen in Figure 5a,b and has these specifications [27]:

- Sampling method: Sequential sampling. Single ADC.
- Sampling rate: 128 samples per second (2048 Hz internal).
- Resolution: 14 bits 1 least significant beat = 0.51 μV (16 bit ADC, 2 bits instrumental noise floor discarded), or 16 bits.
- Bandwidth: 0.2–43 Hz, digital notch filters at 50 Hz.
- Filtering: Built in digital 5th order Sinc filter.
- Dynamic range (input referred): 8400 μV.
- Coupling mode: AC coupled.

60 s were recorded and the hair of all subjects was cut to <1 mm to receive the best possible quality of the data.

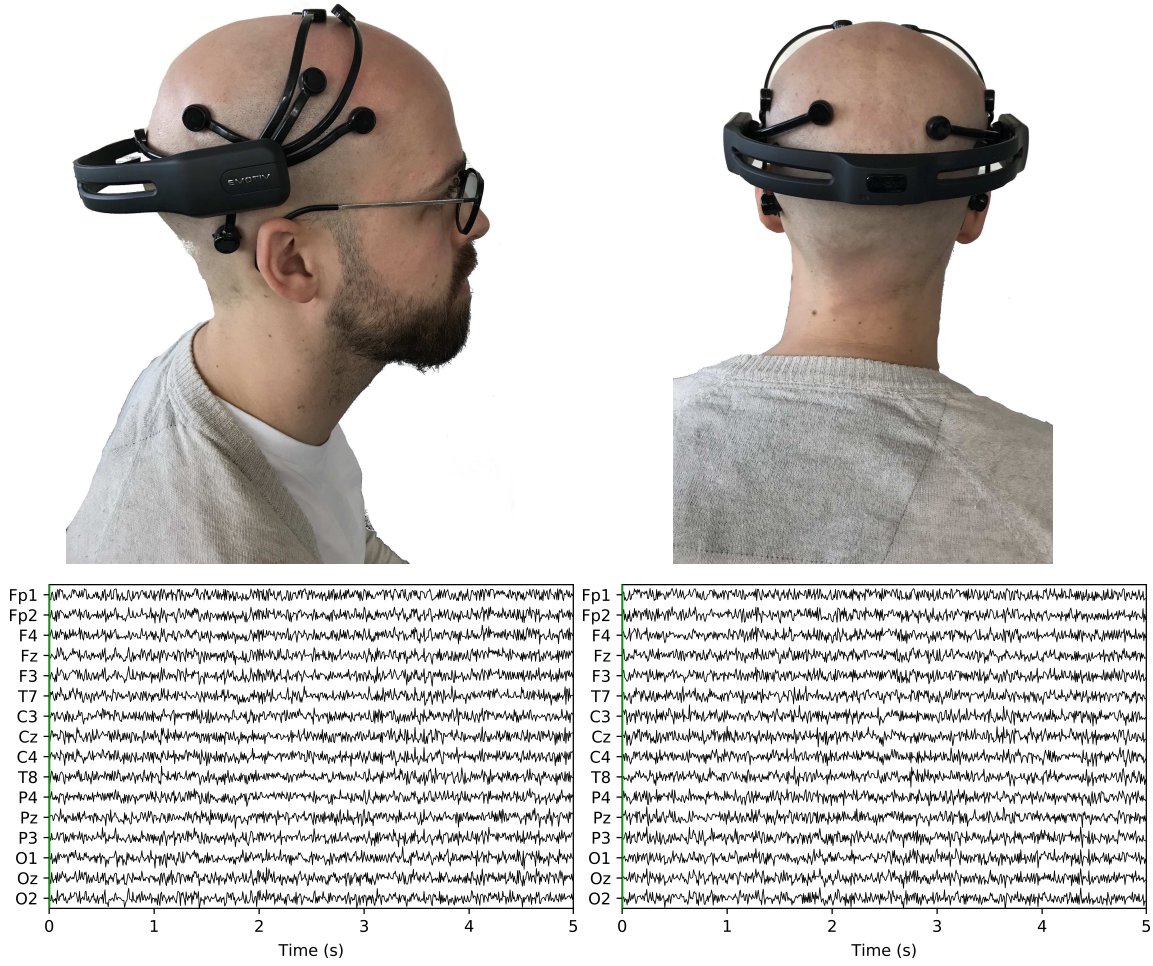


Figure 5. Data Collection. (**a**) EEG Low Cost Portable Sensor; (**b**) EEG Low Cost Portable Sensor; (**c**) 5 s of the recorded EEG-data for the *HKT* leader *Subject* 1; (**d**) 5 s of the recorded EEG-data for the *BSC* leader *Subject* 1.

4.4. Data Pre-Processing

After recording the EEG data, it needs to be pre-processed to improve the signal-to-noise ratio. This is done in multiple steps and has been achieved with ***Fieldtrip*** [75]. An open access toolbox for pre-processing and analyzing EEG data in ***MATLAB***. A high-pass filter is used to cut out the DC component of the signals, as large drifts were observed in the data. Also, a hardware embedded low-pass filter is used to remove frequencies above 50 Hz. Although it can be possible that gamma waves have an even higher frequency [76], those are usually the result of artifacts during the recording.

In the last step, the data is normalized to the range of -10 to 10, allowing the greatest level of anonymity for the probands. This doesn't remove any essential information for the further analysis, as relative differences between subjects are not relevant for this study.

4.5. Standardization Procedure

For every recording, a pair of process owner and his leader are examined. For the first 10 s of the recordings, the leader talks, and the PO listens. In the remaining 50 s, the process owner presents his results. In this time, the leader observes. This is in contrast to the earlier study [27], where only the process owner of the LM system was recorded without speaking.

During the recording, the pair of PO and leader sit in a room with 50 dBA artificially created noise. This noise level is akin to that of a fridge and makes it possible to have comparable background noises.

4.6. Data Analysis

In the following section of this paper the developed soft sensor is further described. This encompasses the experimental setup Section 4.6.1 of the used hardware and software as well as the DL based analysis Section 4.6.3.

4.6.1. Experimental Setup

The training and testing of the neural network and the pre-processing has been performed using following hardware:

- *CPU (Central Processing Unit):* Intel(R) Xeon(R) Gold 6154 CPU @ 3.00 GHz
- *GPU (Graphics Processing Unit):* NVIDIA Quadro P4000
- *RAM (Random-access Memory):* 192GB DDR4

The source code for the data pre-processing, the training, and testing of the neural network is available under Open Access Repository and was created with ***Jupyter Notebook*** Version 5.7.0.

4.6.2. Correlation Function

For this paper the Pearson correlation is calculated which returns a value between -1 for a high negative correlation, 0 for no interrelationship, and +1 for a strong positive correlation [77].

4.6.3. Deep Learning

After the recording and pre-processing of the EEG-data, further steps have to be taken to train the neural network.

1. Data Segmentation

   At first the pre-processed time-dependent EEG data is split into 0.5 s long segments. It is possible to work with shorter or longer segments such as 1 s [27,78], but 0.5 s was chosen because shorter lengths would decrease the amount of information of a data point and longer lengths would decrease the amount of data points that can be used for training.
2. Image generation

   Through the MNE library [79,80], these time segments are transformed into topographic maps, as shown in Figure 6. The topographic map displays the activity in the brain, using distinct color tones for the different strengths. To show the brain activity for the complete brain and not just the measured points, the points in between the sensors are interpolated, creating a topographic map for the complete brain. Using more sensors would increase the accuracy of the interpolated area.

   In the example Figure 6 four topographic maps are visible. These show the average brain activity in the first 0.5 s for the four different categories.

   The values of the 0.5 s segments stretch to the smallest and largest value, displaying the relative differences on the brain using the *jet* color map as shown in Figure 6e. This was chosen as the full range of colors is used and the color range does not need to be optimized for the human perception as this color range is not intuitive for a human [81]. The lowest values are shown as a dark blue, which turn to a green and then to a dark red with the highest value.

   The images have a size of $360 \times 360 \times 3$ pixel. A different size could have been chosen. Smaller images could decrease the accuracy and larger images would increase the training time.

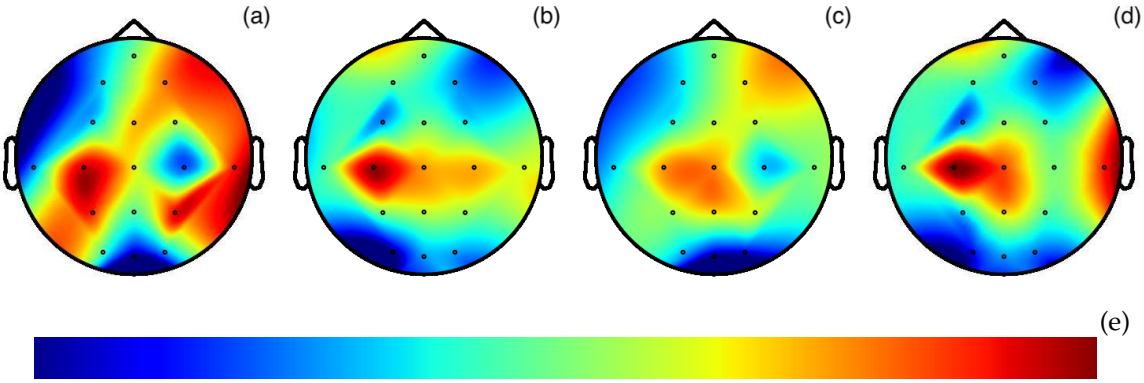


Figure 6. Topographic maps for the first 0.5 s of the four different categories. Each topographic map was generated from one subject. (**a**) *HKT* process owner. (**b**) *Hoshin Kanri TREE (HKT)* leader. (**c**) *Balanced Scorecard (BSC)* process owner. (**d**) *BSC* leader. (**e**) Colormap Jet used for the topographic maps.

3. Deep Learning Architecture

With these generated images, a neural network can be trained. In this paper a convolutional neural network [82] is used. It has shown astonishing results within the realm of image classification [83–85] and is by far the most adopted neural network for physiological signal data processing while also showing excellent results [86]. By using it, a manual feature engineering is unnecessary. In the training phase, the neural network finds the most important features. These become more complex in each convolutional layer.

Still, some specifications are set manually. These form the architecture. The coarse architecture of the used neural network can be seen in Figure 7, the exact composition and the code in the Open Access Repository. Parameters that are set include among others the number and the type of layers and the optimizer for the training of the neural network. Here we have chosen the *Adam* optimizer for the training of the network which stands for adaptive moment estimation [87]. The network architecture comprises four repeated layer groups consisting of a convolution, followed by a rectified linear unit (ReLU) activation and a max pooling layer. After this, the network is flattened and a regular, deeply connected neural network layer follows. To reduce the chance of over-fitting a dropout layer with a ratio of 0.2 is added. The network ends with four outputs that go through the Softmax function to display the probability of the four examined categories.

While statistical ways exist to help select some parameters [88,89] and few have become unofficial standards, no clear-cut way to know the best beforehand exist yet. Thus, it is an iterative process of finding the ideal specifications for the architecture by starting with a simple design and seeing which changes improve the results. The final optimal layout therefore depends on multiple factors. More complex features that have to be found in general increase the number of layers needed [90]. But these can also be limited because the training data wouldn't be sufficient for the increasing number of weights or a limit is reached solely because the hardware requirements can't be met.

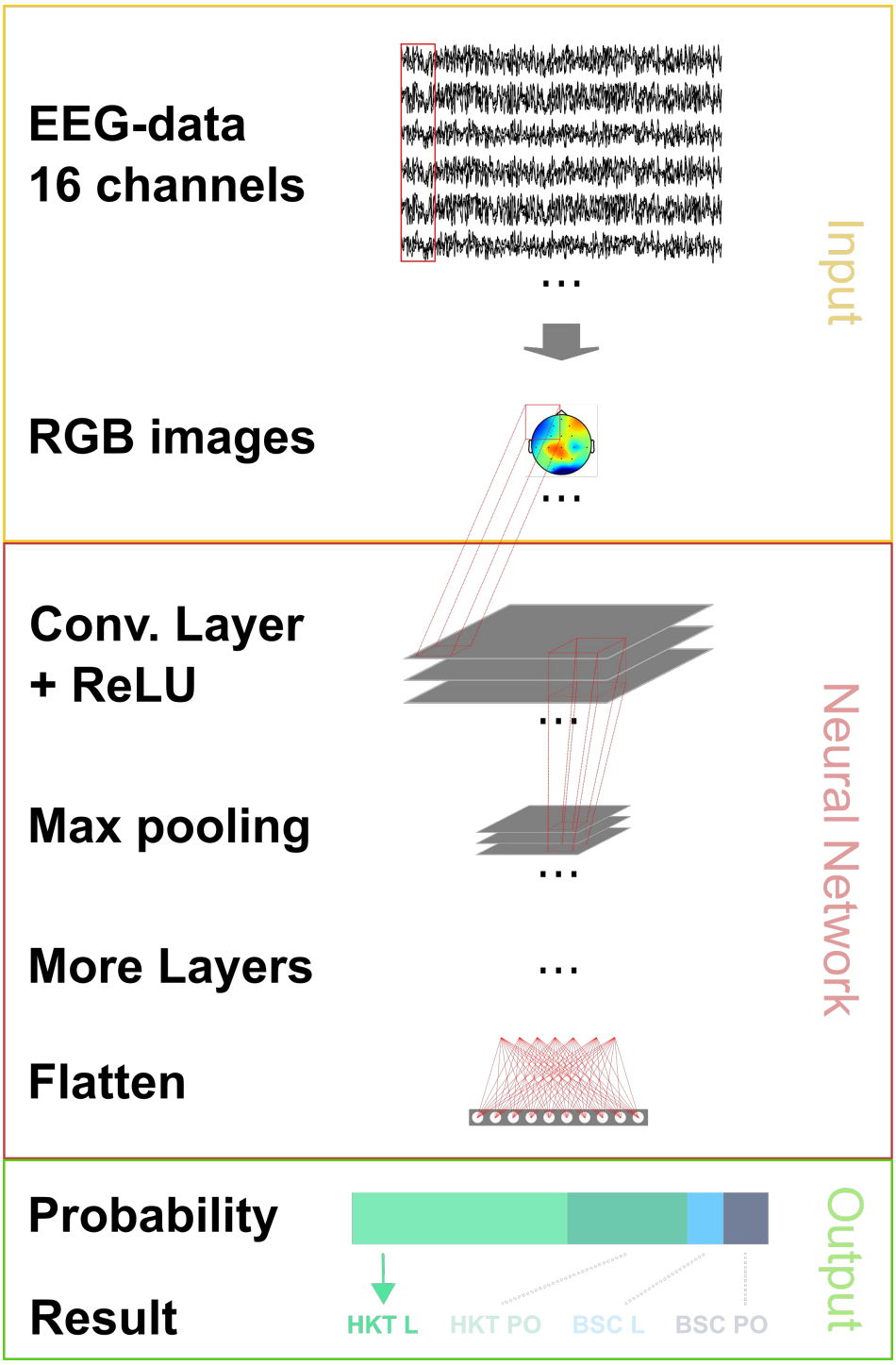


Figure 7. Deep Learning architecture for the classification of the EEG-data.

5. Results and Discussion

This section presents the obtained results using the recorded EEG-data. It comprises two major parts. Through the analysis of the EEG-data with the help of the correlation function, and through the analysis of the EEG-data with the help of the DL soft sensor. The real-world implications of these findings follow in Section 6.

5.1. Results and Discussion of the Correlation Function

The correlation function can verify the hypotheses presented in Table 1. In the following, the results for each hypothesis are examined.

1. Corresponding to *H1*

   A subject that is listening shows strong correlations between the prefrontal-cortex and the occipital cortex. For the recordings this would mean that a high correlation between the sensors *Fp1-F3-Fp2-F4-Fz* and *O1-Oz-O2* can be found.

   This can clearly be seen in the Figures 8 and 9. The leaders for both *HKT* and *BSC* show a strong correlation between the prefrontal-cortex and the occipital cortex with values ranging between 0.52 and 0.71 for the HKT leader and values between 0.32 and 0.61 for the BSC leader.
2. Corresponding to *H2*

   A subject that is listening shows significant correlations **within** the prefrontal-cortex and the occipital cortex. This would be seen in the recordings through high correlations within the sensor groups *Fp1-F3-Fp2-F4-Fz* and *O1-Oz-O2*.

   This is confirmed in Figures 10 and 11. Both *HKT* and *BSC* process owner show high correlations within the sensor groups but not between the sensor groups. The values for the frontal sensor group are in the range from 0.6 to 0.74 for the HKT process owner and 0.62 to 0.76 for the BSC process owner. For the occipital group, the values range between 0.55 and 0.72 for the HKT process owner and 0.58 to 0.73 for the BSC process owner. The values between the frontal and occipital sensor groups only reach up to 0.23 for both HKT PO and BSC PO.

   A clear difference between the leaders and process owners can be seen in Figures 12 and 13, where the biggest differences are the connections between the frontal and occipital group.
3. Corresponding to *H3*

   To confirm an executive behavioural pattern in the neurological activity, a strong correlation in the PFC should be found. For the recordings this would mean that a high correlation between the sensors *Fp1-F3-Fp2-F4-Fz* should be expected in all four examined categories.

   This is confirmed in Figures 9 and 10 showing very high correlations within the frontal sensor group. The minimum value for the weakest correlation of the four subject groups still reaching 0.6, which can already be considered a strong correlation.
4. Corresponding to *H4*

   To conclude hypothesis 4, it is anticipated that strong correlations can be found between the PFC and the TPJ for *HKT* practitioners.

   In the *EEG* recordings this would be confirmed, if strong correlations between the sensors *Fp1-F3-Fp2-F4-Fz* and *T7-T8*, as well as *P3-P4* can be found.

   This can be seen in Figures 8 and 10. Values range between 0.52 to 0.72 for the HKT process owner and 0.46 to 0.64 for the HKT leader.
5. Corresponding to *H5*

   To conclude hypothesis 5, it is anticipated that only weak correlations can be found between the prefrontal-cortex and the TPJ for *BSC* practitioners.

   In the *EEG* recordings this would be confirmed, if weak correlations between the sensors *Fp1-F3-Fp2-F4-Fz* and *T7-T8* as well as *P3-P4* can be found.

   The results for this hypothesis aren't as clear-cut as the previous. Although clearly lower correlations such as 0.28 for the correlation *Fp1_P3* of the BSC PO can be found, *Fp1_T8* for

the same category also shows a correlation of 0.65 in Figure 11. What can clearly be seen in the comparison of the BSC PO and the HKT PO in Figure 14, is that a handful of the expected correlations of the BSC PO show a significantly lower value compared to the HKT PO. These are the correlations *Fp1_P3*, *Fp2_T8*, *F4_T8*, *Fz_P3*, and *F3_P3*, that on average show a correlation with 0.2 less. This could mean that the right TPJ is less active for the BSC PO. Which in turn would signify that the BSC PO shows weaker activity in the brain region responsible for empathy or more general for the ability to switch between perspectives.

The difference between the HKT leader and BSC leader in Figure 15 shows similar results as between HKT PO and BSC PO. This means that the right TPJ is also more active for the HKT leader than for the BSC leader, which would again mean that also the leader of the BSC shows less activity in the brain region linked to the ability to switch between perspectives.

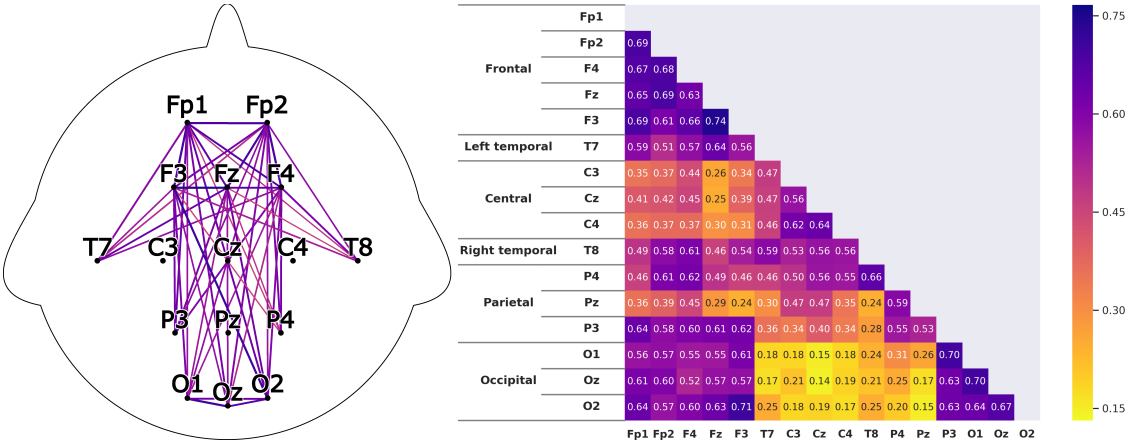


Figure 8. Average correlations of the EEG sensors for the *HKT* process owner displayed in two differing layouts. On the left through connecting lines. The strength is visible through the color and the thickness of the line. Only the correlations concerning the 5 hypotheses are shown to have a clearer view. On the right, the correlation matrix for all EEG sensors is displayed.

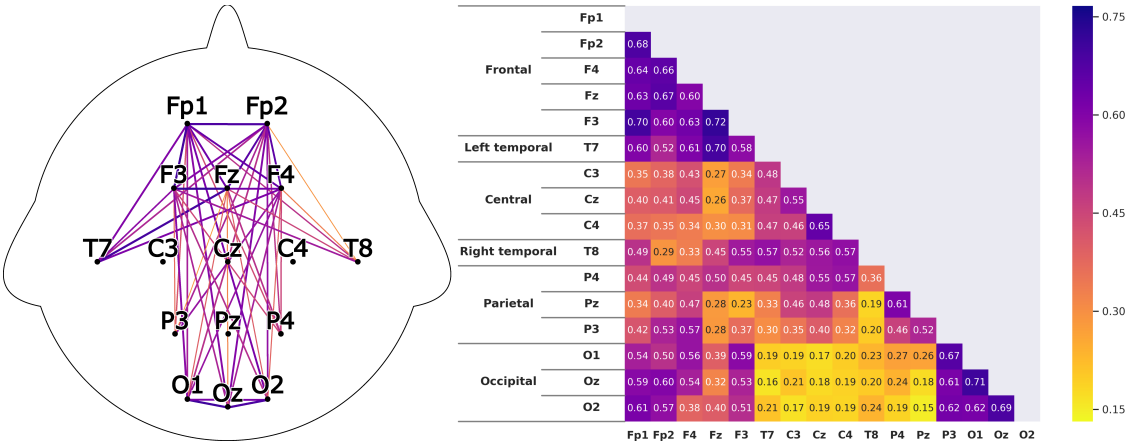


Figure 9. Correlations as described in Figure 8 for the *BSC* leader.

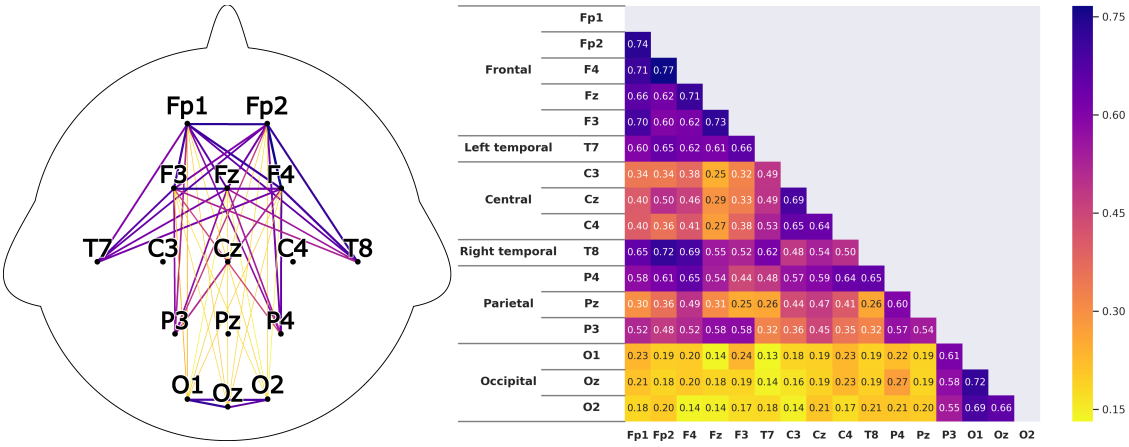


Figure 10. Correlations as described in Figure 8 for the *HKT* leader.

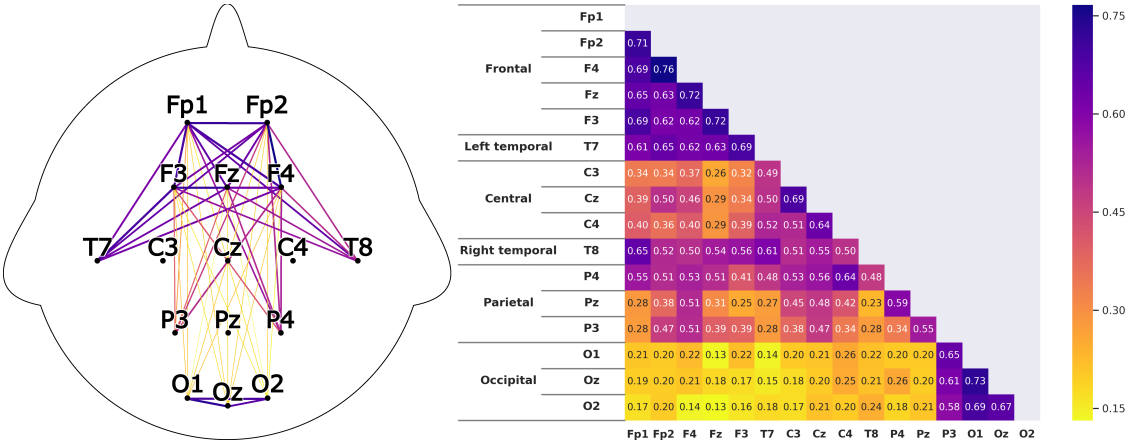


Figure 11. Correlations as described in Figure 8 for the *BSC* process owner.

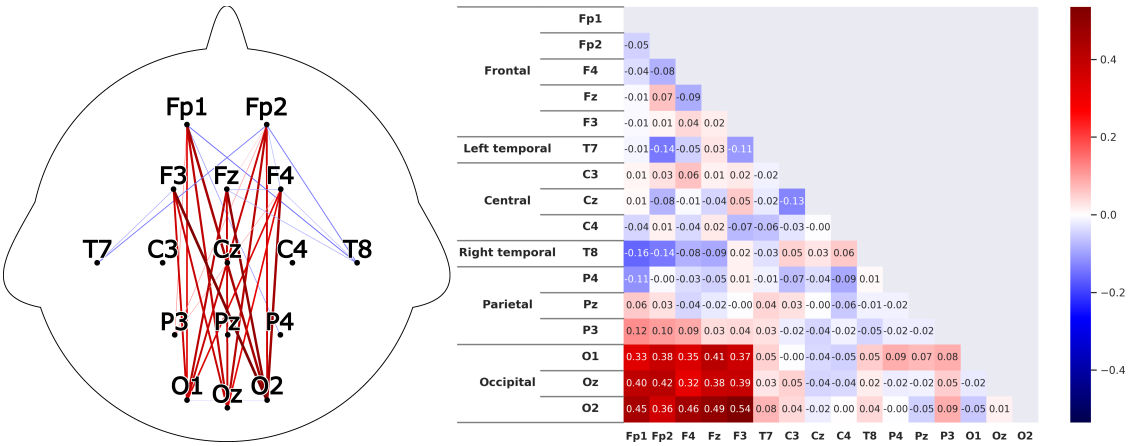


Figure 12. Difference of the results between the HKT leader and the HKT process owner as described in Figure 10.

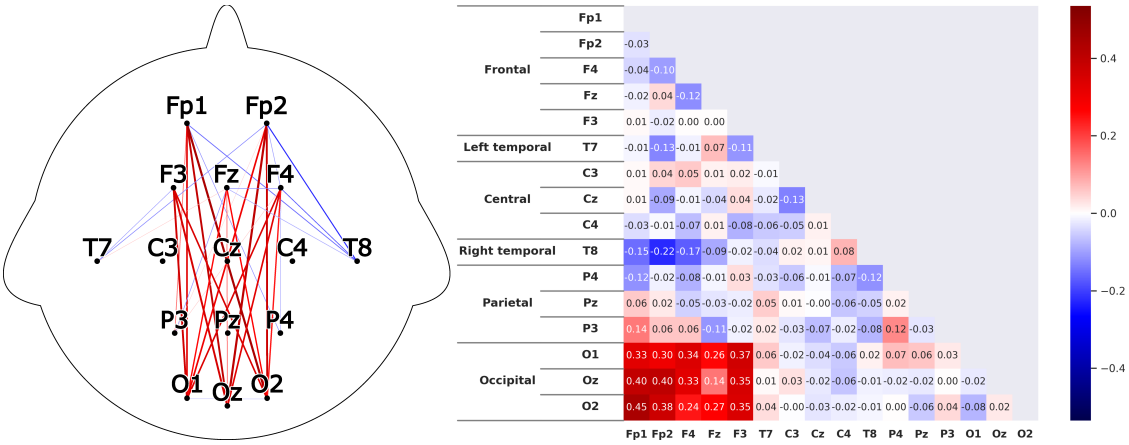


Figure 13. Difference of the results between the BSC leader and the BSC process owner as described in Figure 10.

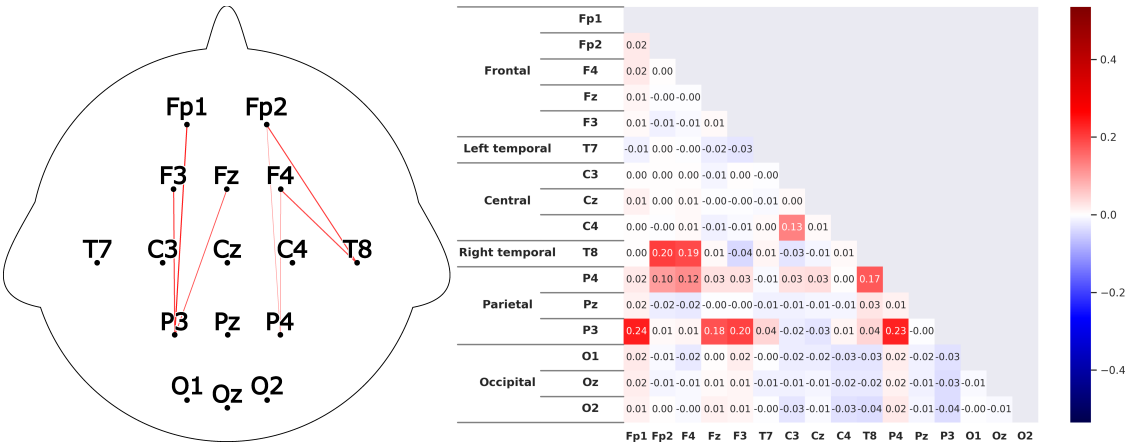


Figure 14. Difference of the results between the HKT process owner and the BSC process owner as described in Figure 10. Colorscale maximum and minimum are set to the +maximum and -maximum absolute value of the examined differences.

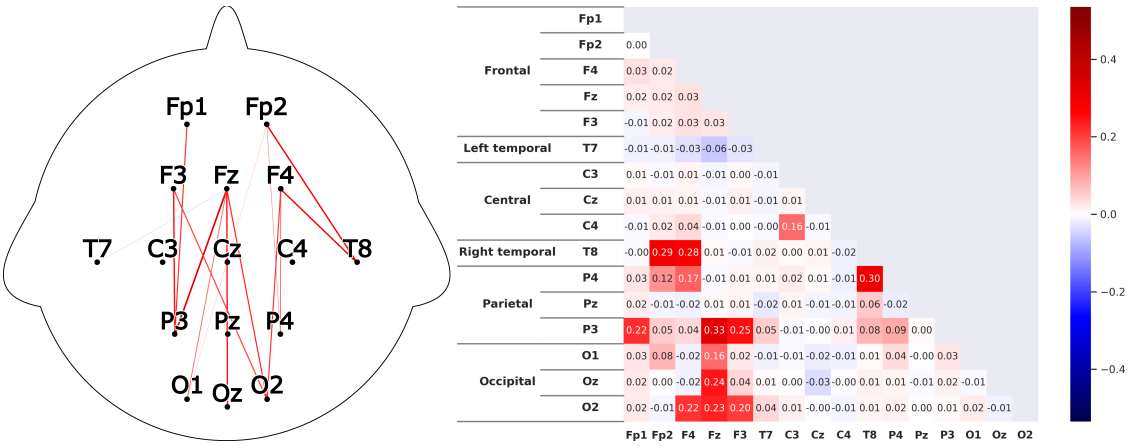


Figure 15. Difference of the results between the HKT leader and the BSC leader as described in Figure 10.

5.2. Results and Discussion of Deep Learning Soft Sensor

With the help of the DL soft sensor, a classification rate accuracy of 96.5% was able to be achieved. This proves that an automatic classification of the EEG data without the need of any domain knowledge is possible. Furthermore, it reveals that the differences of the neurological activity of the subjects practicing either SM system are significant enough to make this achievable.

To train the DNN, the recorded EEG data was split into three parts. The training, validation, and testing set. The set size ratio depends on the amount of data, as the test and validation data should be sufficient to check the validity. While big data cases with millions of examples allow for splits of 99.5% training data, 0.25% validation data and 0.25% test data, cases with examples in the range of 10,000 need a bigger split for the test and validation data [91]. For smaller data collections the standard is a 80%, 10% and 10% split. By keeping all of the time slices of a subject pair in one split, it can be assured that the DNN does not learn specifics from the particular pass. Therefore, a split of $10/14 \approx 72\%$, $2/14 \approx 14\%$ and $2/14 \approx 14\%$ was chosen. With 0.5 s slices this equates to 4800 images for the training set, 960 images for the validation set and 960 images for the test set. Though it would be possible to perform a k-fold cross-validation, this was not chosen, as the data set is big enough for a simple train/test/validation split and it would increase the training time significantly.

After training the DNN with the **training data** in multiple epochs (Figure 16), the **validation data** can be used to tune the hyperparameters of the DNN without exposing the **test data** to the DNN. Only at the very end, when no further changes to the DNN are being done, can the testing data be used to evaluate its effectiveness. A common method is to use the results of a confusion matrix [92]. These are usually separated into the categories True Negative, False Negative, True Positive and False Positive. As this is a multiclass categorization, it makes sense to have a True and False version for each of the four categories.

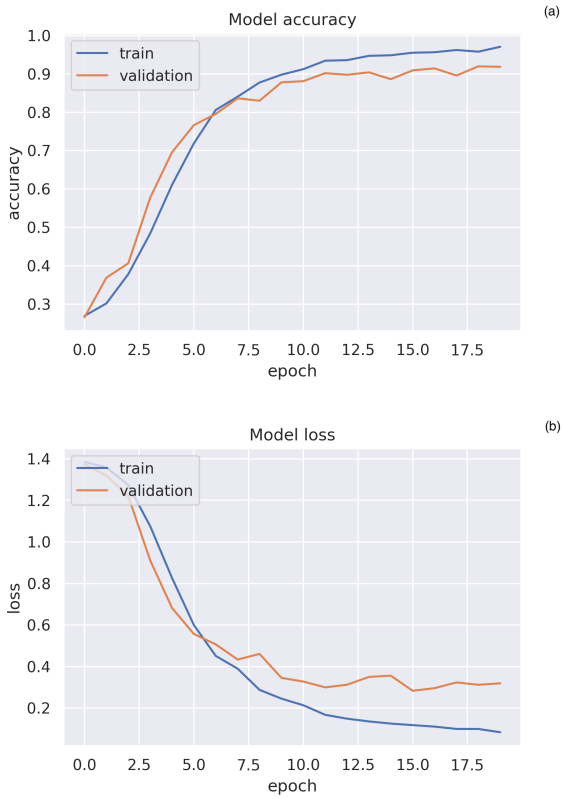


Figure 16. Results for the training of the deep learning (DL) network. (**a**) Training and validation loss of the trained DL network. (**b**) Training and validation accuracy of the trained DL network.

1. True *HKT PO* stands for the EEG data of POs practicing the *HKT* system, that has **correctly** been classified as a PO practicing *HKT*.
2. False *HKT PO* stands for all other EEG data, that has **falsely** been classified as a *HKT PO*.
3. True *HKT Leader* stands for the EEG data of Leaders practicing the *HKT* system, that has **correctly** been classified as a leader practicing *HKT*.
4. False *HKT Leader* stands for all other EEG data, that has **falsely** been classified as a *HKT Leader*.
5. True *BSC PO* stands for the EEG data of POs practicing the *BSC* system, that has **correctly** been classified as a PO practicing *BSC*.
6. False *BSC PO* stands for all other EEG data, that has **falsely** been classified as a *BSC PO*.
7. True *BSC Leader* stands for the EEG data of Leaders practicing the *BSC* system, that has **correctly** been classified as a leader practicing *BSC*.
8. False *BSC Leader* stands for all other EEG data, that has **falsely** been classified as a *BSC Leader*.

The results from the confusion matrix can be seen in Figure 17. With these values, further values for the evaluation of the model can be computed. The accuracy of the DNN can be calculated by dividing the sum of the True cases through the sum of all cases: $(THKTPO + THKTL + TBSCPO + TBSCL)/(THKTPO + THKTL + TBSCPO + TBSCL + FHKTPO + FHKTL + FBSCPO + FBSCL)$. In this study 96.5% has been achieved, which is extremely high, as EEG data usually has a poor signal-to-noise ratio and significant differences between subjects should be expected [93]. Other studies using EEG data and DL for the classification of brain data have also shown positive results. Though it is difficult to compare the applied classification algorithms, as there are few public data sets that are consistently used to evaluate the effectiveness of these. Next to CNN, Recurrent Neural Networks (RNN)/Long-short Term Memory (LSTM) implementations have been used extensively. Depending on the classification type, accuracies up to 100% could be achieved [86].

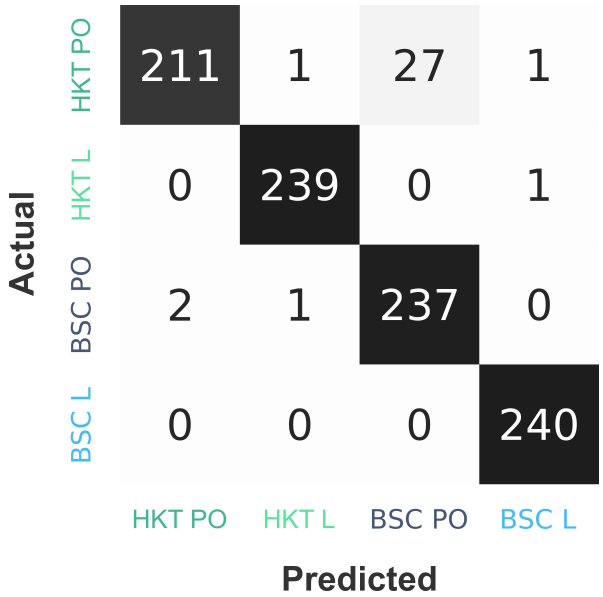


Figure 17. Multi-label confusion matrix (n = 960) without normalization for the four categories that have been used to train the neural network.

6. Management Conclusions and Future Steps

This paper uses a multidisciplinary approach to find ways that can help Industry 4.0 leaders guide the shopfloor. As it forms the center of the value-creation a specific focus needs to be laid on it. Through wide-ranging fields such as neuroscience, machine learning and management science, it has been possible to discover various interesting aspects that support interpreting the thinking processes during the practice of SM systems. It requires to be noted that these are just starting points that need

to be confirmed in further studies. The understanding of the brain regions functions is ongoing and the ability of the regions are not limited to one specific function. The interplay is even more complex and sometimes allows for a wide range of explanations.

Still, following interpretations can be drawn from the achieved results:

- The studied SM systems both cause large correlations within the PFC for the PO and the leader. This indicates that both systems and both PO and leader show brain activity that can be considered goal-oriented, which is the core requirement for any kind of improvement.
- Both BSC and HKT show a solid correlation between the PFC and the left TPJ. The left TPJ amongst other functions plays an important role in the strategic planning concerning other people. The difference between BSC and HKT is negligible, but the difference between PO and leader show interesting patterns. Both correlations *T7_Fp2* and *T7_F3* show noticeable stronger results for the PO. This arises from the different focuses of PO and leader.
- The HKT PO shows substantial correlations of the PFC to the right TPJ. This is especially noticeable in comparison to the results of the BSC participants. The direction-focused approach seems to enable a wider view, allowing diverse positions to be taken into perspective. This is a crucial characteristic for an Industry 4.0 setting where conflicting information needs to be taken into consideration to find optimal improvements that not only shift the focus but ensure sustainable improvements [94]. The HKT leader shows significantly weaker results in this aspect.
- The SM system BSC on the contrary demonstrates weak correlations for the right TPJ to the PFC for both PO and leader, while the leader also shows significantly weaker correlations. The reduced correlations of the PFC to the right TPJ indicate that the contemplated perspectives are restricted and the focus on pre-defined goals limits the aspects that are taken into consideration to receive a goal.

Although the results can be described as preliminary, due in part to the relatively small sample space, if confirmed, the conclusions of this study could have profound implications in the world of Western management. The fundamental reason for this assertion lies basically in the fact that practically the entire corporate culture of operational management is based on the achievement of objectives. In a business environment in which the complexity of processes is constantly increasing, management models that favour a holistic understanding of this reality. Studies such as the one presented, which is clearly multidisciplinary, present, with the help of artificial intelligence, the differences between types of management and the effects that management could be having on POs in a quantitative way.

The classification results of the EEG-data emphasize that significant and distinct differences in the brain activity of the examined leadership patterns exist and can be learned by a DNN. Furthermore, the implemented classifier allows for different usages. As the classification can be done in real time, direct feedback can be given. This could, for example, be used during the training of an SM system. An employee with an EEG-device connected to the trained DNN could receive valuable feedback to differentiate the systems and better understand the thinking patterns.

A comparison of the subjects would be of high interest in order to know how far these results are valid for all subjects and which differences can expected. In this way possible differences between experience, age and gender could be examined and it would be possible to see if alterations in the group size could possibly change the results. Due to compliance rules of the organization this was not possible.

Future research could test the influence of the cultural background and how experience in LM/SM changes the subjects brain patterns. It could also be of interest to examine the influence of the subject speaking on the results.

Author Contributions: Conceptualization, D.S. and J.V.D.; Data curation, D.S. and J.V.D.; Formal analysis, D.S. and J.V.D.; Funding acquisition, J.V.D.; Investigation, D.S.; Methodology, D.S. and J.V.D.; Project administration, J.O.-M., R.G. and J.S.; Resources, D.S.; Software, D.S.; Supervision, D.S. and J.V.D.; Validation, J.V.D. and M.M.; Visualization, D.S. and J.V.D.; Writing—original draft, D.S.; Writing—review & editing, J.V.D. and J.O.-M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: This research was partially supported by Hochschule Heilbronn, Fakultät Management und Vertrieb, Campus Schwäbisch Hall, 74523, Schwäbisch Hall, Germany. We thank Tomas Sterkenburgh from Matthews Europe GmbH, for his unconditional support. The second and third authors want to thank the Spanish Agency for Research for the support provided throughout the research project RTI2018-094614-B-I00 (SMASHING).

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

BSC	Balanced Scorecard
(CPD)nA	Check-Plan-Do-…-Act
dBA	Decibel A-weighting
DL	Deep Learning
DNN	Deep Neural Network
EEG	Electroencephalography
HK	HOSHIN KANRI
HKT	HOSHIN KANRI TREE
KPI	Key Performance Indicator
LM	Lean Management
PDCA	Plan-Do-Check-Act
PFC	Prefrontal Cortex
PO	Process Owner
SM	Shopfloor Management
TPJ	Temporoparietal junction
VS	Value Stream

References

1. Lasi, H.; Fettke, P.; Kemper, H.G.; Feld, T.; Hoffmann, M. Industry 4.0. *Bus. Inf. Syst. Eng.* **2014**, *6*, 239–242. [CrossRef]
2. Rüßmann, M.; Lorenz, M.; Gerbert, P.; Waldner, M.; Justus, J.; Engel, P.; Harnisch, M. Industry 4.0: The future of productivity and growth in manufacturing industries. *Boston Consult. Group* **2015**, *9*, 54–89.
3. Gilchrist, A. *Industry 4.0: The Industrial Internet of Things*; Apress: New York, NY, USA, 2016.
4. Imran, F.; Kantola, J. Review of industry 4.0 in the light of sociotechnical system theory and competence-based view: A future research agenda for the evolute approach. In *Proceedings of the International Conference on Applied Human Factors and Ergonomics*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 118–128.
5. Kopp, R.; Howaldt, J.; Schultze, J. Why Industry 4.0 needs Workplace Innovation: A critical look at the German debate on advanced manufacturing. *Eur. J. Workplace Innov.* **2016**, *2*. [CrossRef]
6. Vaidya, S.; Ambad, P.; Bhosle, S. Industry 4.0—A glimpse. *Procedia Manuf.* **2018**, *20*, 233–238. [CrossRef]
7. Baxter, G.; Sommerville, I. Socio-technical systems: From design methods to systems engineering. *Interact. Comput.* **2011**, *23*, 4–17. [CrossRef]
8. Kagermann, H. Change through digitization—Value creation in the age of Industry 4.0. In *Management of Permanent Change*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 23–45.
9. Diez, J.V.; Ordieres-Mere, J.; Nuber, G. The hoshin kanri tree. Cross-plant lean shopfloor management. *Procedia CIRP* **2015**, *32*, 150–155. [CrossRef]
10. Narkhede, B.; Nehete, R.; Mahajan, S. Exploring Linkages between Manufacturing Functions, Operations Priorities and Plant Performance in Manufacturing Smes in Mumbai. *Int. J. Qual. Res.* **2012**, *6*, *9–22*.
11. Lee, J.; Bagheri, B.; Kao, H.A. Recent advances and trends of cyber-physical systems and big data analytics in industrial informatics. In Proceedings of the International Conference on Industrial Informatics (INDIN), Porto Alegre, Brazil, 27–30 July 2014; pp. 1–6.
12. Schuh, G.; Potente, T.; Varandani, R.M.; Schmitz, T. Methodology for the assessment of structural complexity in global production networks. *Procedia CIRP* **2013**, *7*, 67–72. [CrossRef]
13. Covey, S.R. *The 8th Habit: From Effectiveness to Greatness*; Simon and Schuster: New York, NY, USA, 2013.

14. Coleman, P.T. Implicit Theories of Organizational Power and Priming Effects on Managerial Power-Sharing Decisions: An Experimental Study 1. *J. Appl. Soc. Psychol.* **2004**, *34*, 297–321. [CrossRef]
15. Shah, R.; Ward, P.T. Defining and developing measures of lean production. *J. Oper. Manag.* **2007**, *25*, 785–805. [CrossRef]
16. Frow, N.; Marginson, D.; Ogden, S. "Continuous" budgeting: Reconciling budget flexibility with budgetary control. *Account. Organ. Soc.* **2010**, *35*, 444–461. [CrossRef]
17. Jolayemi, J.K. Hoshin kanri and hoshin process: A review and literature survey. *Total. Qual. Manag.* **2008**, *19*, 295–320. [CrossRef]
18. De Leeuw, S.; van den Berg, J.P. Improving operational performance by influencing shopfloor behavior via performance management practices. *J. Oper. Manag.* **2011**, *29*, 224–235. [CrossRef]
19. Suzaki, K. *New Shop Floor Management: Empowering People for Continuous Improvement*; Simon and Schuster: New York, NY, USA, 1993.
20. Womack, J.P.; Jones, D.T. *Lean Thinking: Banish Waste and Create Wealth in Your Corporation*, 2nd ed.; Simon & Schuster: New York, NY, USA, 2003.
21. Hertle, C.; Siedelhofer, C.; Metternich, J.; Abele, E. The next generation shop floor management—How to continuously develop competencies in manufacturing environments. In Proceedings of the 23rd International Conference on Production Research, Manila, Philippines, 1–4 August 2015.
22. Davis, S.; Albright, T. An investigation of the effect of balanced scorecard implementation on financial performance. *Manag. Account. Res.* **2004**, *15*, 135–153. [CrossRef]
23. Speckbacher, G.; Bischof, J.; Pfeiffer, T. A descriptive analysis on the implementation of balanced scorecards in German-speaking countries. *Manag. Account. Res.* **2003**, *14*, 361–388. [CrossRef]
24. Doran, M.S.; Haddad, K.; Chow, C.W. Maximizing the success of balanced scorecard implementation in the hospitality industry. *Int. J. Hosp. Tour. Adm.* **2002**, *3*, 33–58. [CrossRef]
25. Schuster, B.; Herrmann, F. Analyse, Beurteilung und Entwicklung eines Umsetzungskonzeptes zur Optimierung des Warengruppenmanagements im Einkauf der Krones AG. *AKWI* **2017**, *5*, 14.
26. Mooraj, S.; Oyon, D.; Hostettler, D. The balanced scorecard: A necessary good or an unnecessary evil? *Eur. Manag. J.* **1999**, *17*, 481–491. [CrossRef]
27. Villalba-Diez, J.; Zheng, X.; Schmidt, D.; Molina, M. Characterization of Industry 4.0 Lean Management Problem-Solving Behavioral Patterns Using EEG Sensors and Deep Learning. *Sensors* **2019**, *19*, 2841. [CrossRef]
28. Kudernatsch, D. Eine Lean-Kultur im Unternehmen verankern. *wissensmanagement* **2013**, *3*, 48–49.
29. Fox, M.D.; Snyder, A.Z.; Vincent, J.L.; Corbetta, M.; Van Essen, D.C.; Raichle, M.E. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 9673–9678. [CrossRef] [PubMed]
30. Buckner, R.L.; Andrews-Hanna, J.R.; Schacter, D.L. The brain's default network: Anatomy, function, and relevance to disease. In *Annals of the New York Academy of Sciences: Volume 1124. The Year in Cognitive Neuroscience 2008*; Blackwell Publishing: Hoboken, NJ, USA, 2008.
31. Gabčanová, I. The employees—The most important asset in the organizations. *HUm. Resour. Manag. Ergon.* **2011**, *5*, 30–33.
32. Miller, E.K.; Cohen, J.D. An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* **2001**, *24*, 167–202. [CrossRef] [PubMed]
33. Samson, D.; Apperly, I.A.; Chiavarino, C.; Humphreys, G.W. Left temporoparietal junction is necessary for representing someone else's belief. *Nat. Neurosci.* **2004**, *7*, 499–500. [CrossRef] [PubMed]
34. Saxe, R.; Wexler, A. Making sense of another mind: The role of the right temporo-parietal junction. *Neuropsychologia* **2005**, *43*, 1391–1399. [CrossRef] [PubMed]
35. Ogawa, A.; Kameda, T. Dissociable roles of left and right temporoparietal junction in strategic competitive interaction. *Soc. Cogn. Affect. Neurosci.* **2019**, *14*, 1037–1048. [CrossRef]
36. Krall, S.C.; Rottschy, C.; Oberwelland, E.; Bzdok, D.; Fox, P.T.; Eickhoff, S.B.; Fink, G.R.; Konrad, K. The role of the right temporoparietal junction in attention and social interaction as revealed by ALE meta-analysis. *Brain Struct. Funct.* **2015**, *220*, 587–604. [CrossRef]
37. Kaplan, R.S.; Norton, D.P. Linking the balanced scorecard to strategy. *Calif. Manag. Rev.* **1996**, *39*, 53–79. [CrossRef]

38. Figge, F.; Hahn, T.; Schaltegger, S.; Wagner, M. The sustainability balanced scorecard–linking sustainability management to business strategy. *Bus. Strategy Environ.* **2002**, *11*, 269–284. [CrossRef]
39. Kaplan, R.S.; Norton, D.P. Transforming the balanced scorecard from performance measurement to strategic management: Part II. *Account. Horizons* **2001**, *15*, 147–160. [CrossRef]
40. Villalba-Diez, J.; Ordieres-Mere, J. Improving manufacturing operational performance by standardizing process management. *Trans. Eng. Manag.* **2015**, *62*, 351–360. [CrossRef]
41. Imai, M. *Gemba Kaizen: A Commonsense Approach to a Continuous Improvement Strategy*, 2nd ed.; McGraw-Hill Professional: New York, NY, USA, 2012.
42. Baba, F. Study on stable facility conservation activities based on PDCA cycle. *Yokohama Int. Soc. Sci. Res.* **2012**, *17*, 2–10.
43. Center, J.M.A.M. *PDCA Starting from C Works Faster*! Japan Management Association Management Center: Tokyo, Japan, 2013.
44. Kaplan, R.S.; Norton, D.P. The Balanced Scorecard: Measures that Drive Performance. *Harv. Bus. Rev.* **1992**, *70*, 71–79. [PubMed]
45. Villalba-Diez, J. *The Hoshin Kanri Forest: Lean Strategic Organizational Design*, 1st ed.; Productivity Press: New York, NY, USA, 2017, doi:10.1201/9781315155814. [CrossRef]
46. Neely, A.D. *Measuring Business Performance*; Profile Books: London, UK, 1998.
47. Hines, T. *Supply Chain Strategies: Customer Driven and Customer Focused*; Taylor & Francis: Abingdon, UK, 2004.
48. Niven, P.R. *Balanced Scorecard: Step-by-Step for Government and Nonprofit Agencies*; John Wiley & Sons: Hoboken, NJ, USA, 2008.
49. Díez, J.V. Hoshin Kanri Forest: Lean Strategic Organizational Design. Ph.D. Thesis, Universidad Politécnica de Madrid, Madrid, Spain, 2016.
50. Krumholz, A.; Wiebe, S.; Gronseth, G.; Shinnar, S.; Levisohn, P.; Ting, T.; Hopp, J.; Shafer, P.; Morris, H.; Seiden, L.; et al. Practice Parameter: Evaluating an apparent unprovoked first seizure in adults (an evidence-based review):[RETIRED]: Report of the Quality Standards Subcommittee of the American Academy of Neurology and the American Epilepsy Society. *Neurology* **2007**, *69*, 1996–2007. [CrossRef]
51. Nuwer, M.R. Quantitative EEG: II. Frequency analysis and topographic mapping in clinical settings. *J. Clin. Neurophysiol.* **1988**, *5*, 45–85. [CrossRef]
52. Neufeld, M.; Blumen, S.; Aitkin, I.; Parmet, Y.; Korczyn, A. EEG frequency analysis in demented and nondemented parkinsonian patients. *Dement. Geriatr. Cogn. Disord.* **1994**, *5*, 23–28. [CrossRef]
53. Coben, L.A.; Danziger, W.L.; Berg, L. Frequency analysis of the resting awake EEG in mild senile dementia of Alzheimer type. *Electroencephalogr. Clin. Neurophysiol.* **1983**, *55*, 372–380. [CrossRef]
54. Chandaka, S.; Chatterjee, A.; Munshi, S. Cross-correlation aided support vector machine classifier for classification of EEG signals. *Expert Syst. Appl.* **2009**, *36*, 1329–1336. [CrossRef]
55. Panischev, O.Y.; Demin, S.; Kaplan, A.Y.; Varaksina, N.Y. Use of cross-correlation analysis of EEG signals for detecting risk level for development of schizophrenia. *Biomed. Eng.* **2013**, *47*, 153–156. [CrossRef]
56. Poulos, M.; Georgiacodis, F.; Chrissikopoulos, V.; Evagelou, A. Diagnostic test for the discrimination between interictal epileptic and non-epileptic pathological EEG events using auto–cross-correlation methods. *Am. J. Electroneurodiagn. Technol.* **2003**, *43*, 228–240. [CrossRef]
57. Krishna, D.H.; Pasha, I.; Savithri, T.S. Classification of EEG motor imagery multi class signals based on cross correlation. *Procedia Comput. Sci.* **2016**, *85*, 490–495. [CrossRef]
58. Briechle, K.; Hanebeck, U.D. Template matching using fast normalized cross correlation. In Proceedings of the Optical Pattern Recognition XII. International Society for Optics and Photonics, Orlando, FL, USA, 16–20 April 2001; Volume 4387, pp. 95–102.
59. Avants, B.B.; Epstein, C.L.; Grossman, M.; Gee, J.C. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Med Image Anal.* **2008**, *12*, 26–41. [CrossRef] [PubMed]
60. Shearer, P.M. Improving local earthquake locations using the L1 norm and waveform cross correlation: Application to the Whittier Narrows, California, aftershock sequence. *J. Geophys. Res. Solid Earth* **1997**, *102*, 8269–8283. [CrossRef]
61. Bracewell, R. *Fourier Analysis and Imaging*; Springer: Boston, MA, USA, 2012.

62. Jirayucharoensak, S.; Pan-Ngum, S.; Israsena, P. EEG-based emotion recognition using deep learning network with principal component based covariate shift adaptation. *Sci. World J.* **2014**, *2014*. [CrossRef] [PubMed]
63. Zheng, W.L.; Zhu, J.Y.; Peng, Y.; Lu, B.L. EEG-based emotion classification using deep belief networks. In Proceedings of the 2014 IEEE International Conference on Multimedia and Expo (ICME), Chengdu, China, 14–18 July 2014; pp. 1–6.
64. Oh, S.L.; Hagiwara, Y.; Raghavendra, U.; Yuvaraj, R.; Arunkumar, N.; Murugappan, M.; Acharya, U.R. A deep learning approach for Parkinson's disease diagnosis from EEG signals. *Neural Comput. Appl.* **2018**, 1–7. [CrossRef]
65. Ortiz, A.; Munilla, J.; Gorriz, J.M.; Ramirez, J. Ensembles of deep learning architectures for the early diagnosis of the Alzheimer's disease. *Int. J. Neural Syst.* **2016**, *26*, 1650025. [CrossRef]
66. Hosseini, M.P.; Soltanian-Zadeh, H.; Elisevich, K.; Pompili, D. Cloud-based deep learning of big eeg data for epileptic seizure prediction. In Proceedings of the 2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP), Washington, DC, USA, 7–9 December 2016; pp. 1151–1155.
67. Acharya, U.R.; Oh, S.L.; Hagiwara, Y.; Tan, J.H.; Adeli, H. Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals. *Comput. Biol. Med.* **2018**, *100*, 270–278. [CrossRef]
68. Schirrmeister, R.T.; Springenberg, J.T.; Fiederer, L.D.J.; Glasstetter, M.; Eggensperger, K.; Tangermann, M.; Hutter, F.; Burgard, W.; Ball, T. Deep learning with convolutional neural networks for EEG decoding and visualization. *Hum. Brain Mapp.* **2017**, *38*, 5391–5420. [CrossRef]
69. Francois, C. *Deep Learning with Python*; Manning Publications: Shelter Island, NY, USA, 2017.
70. Voulodimos, A.; Doulamis, N.; Doulamis, A.; Protopapadakis, E. Deep learning for computer vision: A brief review. *Comput. Intell. Neurosci.* **2018**, *2018*, 7068349. [CrossRef]
71. Lee, H.; Pham, P.; Largman, Y.; Ng, A.Y. Unsupervised feature learning for audio classification using convolutional deep belief networks. In Proceedings of the Advances in Neural Information Processing Systems, Vancouver, BC, Canada, 7–10 December 2009; pp. 1096–1104.
72. Min, S.; Lee, B.; Yoon, S. Deep learning in bioinformatics. *Brief. Bioinform.* **2017**, *18*, 851–869. [CrossRef]
73. Villalba-Diez, J. *The Lean Brain Theory. Complex Networked Lean Strategic Organizational Design*; CRC Press, Taylor and Francis Group LLC: Boca Raton, FL, USA, 2017.
74. American Electroencephalographic Society Guidelines in Electroencephalography, Evoked Potentials, and Polysomnography. *J. Clin. Neurophysiol.* **1994**, *11*, 1–147.
75. Oostenveld, R.; Fries, P.; Maris, E.; Schoffelen, J.M. FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Comput. Intell. Neurosci.* **2011**, *2011*, 9, doi:10.1155/2011/156869. [CrossRef] [PubMed]
76. Henry, J.C. Electroencephalography: basic principles, clinical applications, and related fields. *Neurology* **2006**, *67*, 2092. [CrossRef]
77. Adler, J.; Parmryd, I. Quantifying colocalization by correlation: the Pearson correlation coefficient is superior to the Mander's overlap coefficient. *Cytom. Part A* **2010**, *77*, 733–742. [CrossRef]
78. Längkvist, M.; Karlsson, L.; Loutfi, A. Sleep stage classification using unsupervised feature learning. *Adv. Artif. Neural Syst.* **2012**, *2012*, 5. [CrossRef]
79. Gramfort, A.; Luessi, M.; Larson, E.; Engemann, D.A.; Strohmeier, D.; Brodbeck, C.; Goj, R.; Jas, M.; Brooks, T.; Parkkonen, L.; et al. MEG and EEG data analysis with MNE-Python. *Front. Neurosci.* **2013**, *7*, 267. [CrossRef]
80. Gramfort, A.; Luessi, M.; Larson, E.; Engemann, D.A.; Strohmeier, D.; Brodbeck, C.; Parkkonen, L.; Hämäläinen, M.S. MNE software for processing MEG and EEG data. *Neuroimage* **2014**, *86*, 446–460. [CrossRef]
81. Nuñez, J.R.; Anderton, C.R.; Renslow, R.S. Optimizing colormaps with consideration for color vision deficiency to enable accurate interpretation of scientific data. *PLoS ONE* **2018**, *13*, e0199239. [CrossRef]
82. LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**, *86*, 2278–2324. [CrossRef]
83. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. In Proceedings of the Advances in Neural Information Processing Systems, Lake Tahoe, NV, USA, 3–6 December 2012; pp. 1097–1105.

84. Zeiler, M.D.; Fergus, R. Visualizing and understanding convolutional networks. In *Proceedings of the European Conference on Computer Vision*; Springer: Berlin/Heidelberg, Germany, 2014; pp. 818–833.
85. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
86. Rim, B.; Sung, N.J.; Min, S.; Hong, M. Deep Learning in Physiological Signal Data: A Survey. *Sensors* **2020**, *20*, 969. [CrossRef]
87. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
88. Ripley, B.D. Statistical ideas for selecting network architectures. In *Neural Networks: Artificial Intelligence and Industrial Applications*; Springer: Berlin/Heidelberg, Germany, 1995; pp. 183–190.
89. Tishby, N.; Zaslavsky, N. Deep learning and the information bottleneck principle. In Proceedings of the 2015 IEEE Information Theory Workshop (ITW), Jerusalem, Israel, 26 April–1 May 2015; pp. 1–5.
90. Güçlü, U.; van Gerven, M.A. Deep neural networks reveal a gradient in the complexity of neural representations across the ventral stream. *J. Neurosci.* **2015**, *35*, 10005–10014. [CrossRef]
91. Ng, A. Train/Dev/Test sets. 2019. Available online: https://www.coursera.org/lecture/deep-neural-network/train-dev-test-sets-cxG1s (accessed on 8 November 2019).
92. Alpaydin, E. *Introduction to Machine Learning*; Adaptive Computation and Machine Learning; MIT Press: Cambridge, MA, USA, 2004.
93. Stober, S.; Sternin, A.; Owen, A.M.; Grahn, J.A. Deep feature learning for EEG recordings. *arXiv* **2015**, arXiv:1511.04306.
94. Villalba-Diez, J.; Ordieres-Meré, J.; Chudzick, H.; López-Rojo, P. NEMAWASHI: Attaining value stream alignment within complex organizational networks. *Procedia CIRP* **2015**, *37*, 134–139. [CrossRef]

sensors

MDPI

Communication

A Framework for the Optimization of Complex Cyber-Physical Systems via Directed Acyclic Graph

Manuel Castejón-Limas, Laura Fernández-Robles *, Héctor Alaiz-Moretón, Jaime Cifuentes-Rodriguez and Camino Fernández-Llamas

Department of Mechanical, Computer Science and Aerospace Engineering, Universidad de León, 24071 León, Spain; manuel.castejon@unileon.es (M.C.-L.); hector.moreton@unileon.es (H.A.-M.); jcifr@unileon.es (J.C.-R.); camino.fernandez@unileon.es (C.F.-L.)

* Correspondence: l.fernandez@unileon.es; Tel.: +34-987-29-3521

Abstract: Mathematical modeling and data-driven methodologies are frequently required to optimize industrial processes in the context of Cyber-Physical Systems (CPS). This paper introduces the `PipeGraph` software library, an open-source python toolbox for easing the creation of machine learning models by using Directed Acyclic Graph (DAG)-like implementations that can be used for CPS. `scikit-learn`'s `Pipeline` is a very useful tool to bind a sequence of transformers and a final estimator in a single unit capable of working itself as an estimator. It sequentially assembles several steps that can be cross-validated together while setting different parameters. Steps encapsulation secures the experiment from data leakage during the training phase. The scientific goal of `PipeGraph` is to extend the concept of `Pipeline` by using a graph structure that can handle `scikit-learn`'s objects in DAG layouts. It allows performing diverse operations, instead of only transformations, following the topological ordering of the steps in the graph; it provides access to all the data generated along the intermediate steps; and it is compatible with `GridSearchCV` function to tune the hyperparameters of the steps. It is also not limited to (X, y) entries. Moreover, it has been proposed as part of the `scikit-learn-contrib` supported project, and is fully compatible with `scikit-learn`. Documentation and unitary tests are publicly available together with the source code. Two case studies are analyzed in which `PipeGraph` proves to be essential in improving CPS modeling and optimization: the first is about the optimization of a heat exchange management system, and the second deals with the detection of anomalies in manufacturing processes.

Keywords: Cyber-Physical Systems; Lean Manufacturing; Directed Acyclic Graphs; scikit-learn; pipegraph; machine learning models

Citation: Castejón-Limas, M.; Fernández-Robles, L.; Alaiz-Moretón, H.; Cifuentes-Rodriguez, J.; Fernández-Llamas, C. A Framework for the Optimization of Complex Cyber-Physical Systems via Directed Acyclic Graph. *Sensors* **2022**, *22*, 1490. https://doi.org/10.3390/s22041490

Academic Editors: Javier Villalba-Diez and Alexios Mylonas

Received: 23 December 2021
Accepted: 11 February 2022
Published: 15 February 2022

1. Introduction

Continuous technological advancements in fields such as Information Technology (IT), Artificial Intelligence (AI), and the Internet of Things (IoT), among others, have drastically transformed manufacturing processes. Recent technological advancements have permitted a systematic deployment of Cyber–Physical Systems (CPS) in manufacturing, which allows intertwining physical and software components to control a mechanism by means of a computer system. CPS has considerably improved the efficiency of production processes while also making them more resilient and collaborative [1]. These cutting-edge technologies are advancing the manufacturing economic sector in the Industry 4.0 era [2].

In the Industry 4.0 paradigm, manufacturing industries must modify their management systems and look for new manufacturing strategies [3,4] to find solutions to tackle the issues faced nowadays. Lean Manufacturing (LM) has become one of the most generally accepted manufacturing methods and management styles used by organizations throughout the world to improve their business performance and competitiveness [5]. Since LM improves operational performance for manufacturing organizations in developing and

developed countries [6], it has spread all over the world [4]. Ref. [7] suggested that the future research methodologies for LM can be classified into meaningful themes, namely: the size of the research sample and its composition; several types of study (other than surveys); longitudinal studies; applying advanced statistical analysis and (mathematical) modeling techniques; objective, real and quantitative data; surveys; mixed/multiple research studies; reliability and validity analysis; using computer-aided technology for data collection, and processing and research collaborations.

This paper focuses on the application of mathematical modeling techniques and the use of computer-aided technology for data processing for LM CPS, at the core of Industry 4.0. LM deals with the optimization of the performance according to a specific set of principles [8]. In the context of CPS, mathematical modeling and data-driven techniques are usually needed to optimize industrial processes [9]. Ref. [10] presented a fully model-driven technique based on MontiArc models of the architecture of the CPS and UML/P class diagrams to construct the digital twin information system in CPS. Ref. [11] combined Random Forest (RF) with Bayesian optimization for large-scale dimension data quality prediction, selecting critical production aspects based on information gain, and then using sensitivity analysis to preserve product quality, which may provide management insights and operational guidance for predicting and controlling product quality in the real-world process industry. In [12], in the context of varying design uncertainty of CPS, the feasibility of appropriate evolutionary and Machine Learning (ML) techniques was examined. In [13], the Neural Network Verification (NNV), a software tool that offers a set of reachability methods for verifying and testing the safety (and robustness) of real-world Deep Neural Networks (DNNs) and learning-enabled CPS, was introduced.

ML is contributory in solving difficult problems in the domains of data-driven forecasting, classification and clustering for CPS. However, the literature of the ML field presents a high number of approaches and variations which makes difficult to establish a clear classification scheme for its algorithms [14]. Toolkits for ML aim at standardizing interfaces to ease the use of ML algorithms in different programming languages as R [15], Apache Spark [16], JAVA and C# [17], C++ [18,19], JAVA [20,21], PERL [22], JavaScript [23], command-line [24], among others. Python is one of the most popular and widely-used software systems for statistics, data mining, and ML, and `scikit-learn` [25] is the most widely used module for implementing a wide range of state-of-the-art ML algorithms for medium-scale supervised and unsupervised problems.

Data-driven CPS case studies usually need to split the data into training and test sets and to combine a set of processes to be applied separately to the training and test data. Some bad practices in data manipulation can end up in a misleading interpretation of the achieved results. The use of tools that allow the selection of the pertinent steps in an ad-hoc designed pipeline helps to reduce programming errors [26]. The `Pipeline` object of the `scikit-learn` module allows combining several transformers and an estimator to create a combined estimator [25]. This object behaves as a standard estimator, and `GridSearchCV` therefore can be used to tune the parameters of all steps. Nevertheless, `Pipeline` has some shortcomings such as: it is quite rigid since it strictly allows combining transformers and an estimator sequentially in such a way that the inputs of a step are the transformed outputs of the previous step; it requires (X, y)-like entries, the X is transformed by the transformers, and (X, y) is used by the estimator; `GridSearchCV` on a `Pipeline` is constrained to only split (X, y) and tune parameters that are variables of the functions, for example, a `fit_param` cannot be tuned.

In this work we present `PipeGraph`, a new toolbox that aims at overcoming some of the weaknesses of `Pipeline` while providing greater functionality to ML users. In `PipeGraph`, all kinds of steps, not only transformers and an estimator, can be combined in the form of a Directed Acyclic Graph (DAG), the entries of a step can come from the outputs of any previous step or the inputs to the graph, see Figure 1. For more information about the application of DAGs in ML see [27]. A `PipeGraph` accepts more variables than (X, y) as inputs that can be appended or split within the graph to make use of the standard

`scikit-learn` functions. This allows using hyperparameter tuning techniques such as `GridSearchCV` to easily manage data further than (X, y) inputs and tune other parameters apart from the variables of functions. The user can easily implement a new `PipeGraph` creating steps that make use of: (i) `scikit-learn` functions, (ii) provided custom blocks that implement basic functions, or (iii) their own elaborated custom blocks that implement custom functions. Moreover, in this paper we report two case studies where `PipeGraph` was essential to ease the modeling and optimization in the CPS area. The first case study deals with an optimization of heat exchange management system, whereas the second one deals with anomaly detection in manufacturing processes.

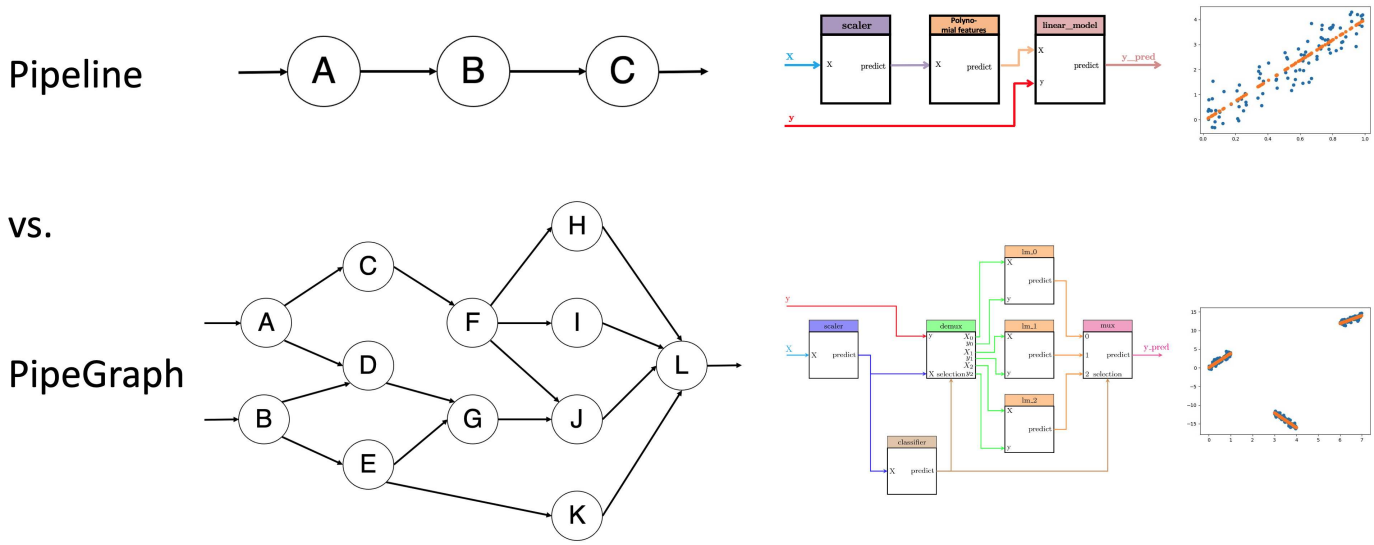


Figure 1. Graphical abstract. In the upper part, the `Pipeline` structure can be seen, which only allows sequential steps. At the bottom, the `PipeGraph` structure is shown. The combination of steps based on a directed acyclic graph makes a wide variety of operations feasible.

2. Data Leakage in ML Experiments

This section emphasizes the importance of avoiding data leakage in data driven numerical modeling. Data leakage in experimental learning is an important design defect that practitioners must consciously avoid. Best practices in ML projects establish that the information related to the case study must be split in a number of different data sets, i.e., training set, test set, and if necessary validation test [28]. This allows the practitioner to obtain a measure of the error expected on data unseen during the training process. It is crucial for the model to be evaluated on unseen data in order to confirm the generalization capability of the model. Moreover, cross–validation (CV) is one of the most popular strategies to obtain a representative value of the error measure. CV pursues to provide an error measure on unseen data by using the training set alone. To achieve that goal, it splits the training data set in a number k of subsets, also known as 'folds', and runs k training experiments by isolating one of those folds at a time, to later on measure the error on the specific set that has been isolated. Thus, in each of the k experiments the model is trained and tested on different sets. Finally, the error measures are condensed by using standard statistics like mean value and standard deviation.

One of the most important risks associated to CV is data leakage, by failing to correctly manage the different sets of data and their effect on the different stages of the training phase. Let us consider the following simple illustrative case: a linear model that is fitted using scaled data. Such an example requires the data to run through two processes: a scaler and then a predictive model. During the training phase the scaler annotates information related to the presented data, e.g., maximum and minimum, or mean value and standard deviation. This annotations allow the scaler process to eventually apply the same transformation to new and unseen data. A typical error that might occur is that the code for the CV presents

the whole training data set to the scaler and then loops different fit experiments using the *k*-fold strategy. By doing so, each of the *k* models fails to provide an error measure that is representative of the behavior of the model on unseen data, because indeed, what is meant to be considered as unseen data has effectively polluted the experiment by its potential impact on the parameters annotated on the scaler.

One of the main strategies to avoid the risk of allowing the user to mishandle the flow of information during CV is to embed the sequence of processes in an entity that is handled in the code as a single unit. A remarkable example of such solution is the `Pipeline` class provided by `scikit-learn`. According to `scikit-learn` documentation "*pipelines help avoid leaking statistics from your test data into the trained model in cross-validation blue (CV), by ensuring that the same samples are used to train the transformers and predictors*" [25].

In the former example, the scaler cannot be trained using the whole data set during the CV experiment because `Pipeline` is in charge of training the scaler as many times as the predictive model. This is a very successful strategy for migrating the responsibility of orchestrating the fit and predict phases from the user to the code of the framework, thus avoiding possible user errors while handling the data.

CV is one of the scenarios where data leakage can occur, other notable situations prone to this design defect can occur when augmenting the data or during the feature engineering stage, to name a few.

`PipeGraph` is another example of such encapsulation. As `Pipeline`, it provides the researcher with the safety that no data leakage will occur during the training phase. Moreover, it enhances the capabilities of the standard `Pipeline` provided by `scikit-learn` by allowing non–linear data flows, as we will show in the following section.

Thus the scientific goal of this paper is to present researchers from the ML community, in particular those in the CPS area, a novel framework aimed at providing expressive means to design complex models such as those typically present in CPS. `PipeGraph` thus combines the expressive power of DAGs with the intrinsic safety of encapsulation.

3. Library Design

3.1. Project Management

Quality assurance. In order to ensure code quality and consistency, unitary tests are provided. We granted a coverage of 89% for the release 0.0.15 of the `PipeGraph` toolbox. New contributions are automatically checked through a Continuous Integration (CI) system for the sake of determining metrics concerning code quality.

Continuous integration. In order to ensure CI when using and contributing to `PipeGraph` toolbox, Travis CI is used to integrate new code and provide back-compatibility. Circle CI is used to build new documentation along with examples containing calculations.

Community-based development. `PipeGraph` toolbox is fully developed under GitHub and gitter to facilitate collaborative programming, this is, issue tracking, code integration, and idea deliberations.

Documentation. We provide consistent Application Programming Interface (API) documentation and gallery examples (https://mcasl.github.io/PipeGraph/api.html, accessed on 20 December 2021) by means of sphinx and numpydoc. A user's guide together with a link to the API reference and examples are provided and centralized in GitHub (https://github.com/mcasl/PipeGraph, accessed on 20 December 2021).

Project relevance. At edition time of this paper `PipeGraph` toolbox has been proposed as part of the scikit-learn-contrib supported project.

3.2. Implementation Details

Here we describe the main issues that had to be solved for PipeGraph to work. First, `scikit-learn` step eligible classes provide a set of methods with different names for essentially the same purpose; providing the output corresponding to an input dataset. Depending on whether the class is a transformer or an estimator, this method can be called either `transform`, `predict`, or even `fit_predict`. This issue was originally solved by using

the wrapper design pattern, although it has been proposed by scikit-learn core developers to consider the usage of mixin classes to provide a similar functionality. The development branch of the package already implements this alternative approach.

The main difference from a users perspective between using a Pipeline object or a PipeGraph object is the need of defining a dictionary that establishes the connections. Again, a `scikit-learn` core developer suggested the implementation of an `inject` method for that purpose and that approach is also already available along with the optional dictionary for those cases in which the user finds it more convenient to use.

The proposed toolbox depends on `numpy`, `pandas`, `networkx`, `inspect`, `logging` and `scikit-learn` and is distributed under MIT license.

PipeGraph can be easily installed using `pip install pipegraph`.

3.3. Example

We describe here one example for illustrative purposes. The system displays a predictive model in which a classifier provides the information to a demultiplexer to separate the dataset samples according to their corresponding class. After that, a different regression model is fitted for each class. Thus, the system contains the following steps:

scaler: A `scikit-learn` MinMaxScaler data preprocessor in charge of scaling the dataset.

classifier: A `scikit-learn` GaussianMixture classifier in charge of performing the clustering of the dataset and the classification of any new sample.

demux: A custom Demultiplexer class in charge of splitting the input arrays accordingly to the selection input vector. This block is provided by PipeGraph.

lm_0, lm_1, lm_2: A set of `scikit-learn` LinearRegression objects

mux: A custom Multiplexer class in charge of combining different input arrays into a single one accordingly to the selection input vector. This block is provided by PipeGraph.

This PipeGraph model is shown in Figure 2. It can be clearly seen in the figure the non sequential nature of such a system that cannot be otherwise described as a standard `scikit-learn` Pipeline.

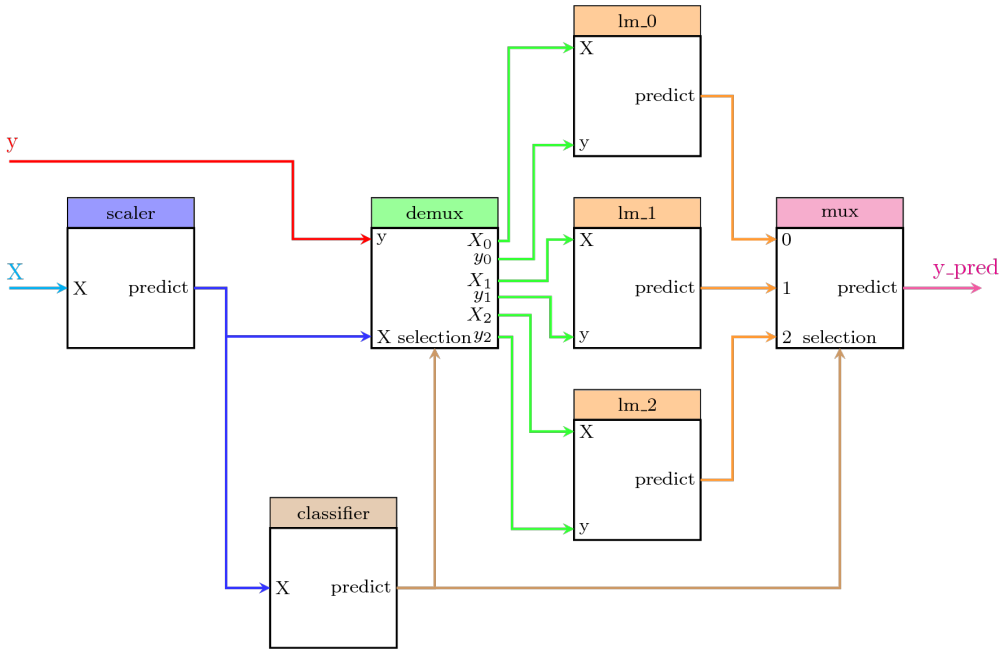


Figure 2. A PipeGraph system for fitting a different model per cluster.

The code for creating an artificial dataset and configuring the system is described in Listing A1 of Appendix A.

3.4. Implemented Methods

`PipeGraph` toolbox provides two interfaces, `PipeGraphRegressor` and `PipeGraphClassifier`, which are compatible with `GridSearchCV` and heavily based on `scikit-learn`'s `Pipeline` on purpose, as its aim is to offer an interface as similar to `Pipeline` as possible. By default, `PipeGraphRegressor` uses the regressor default score (the coefficient of determination R^2 of the prediction) and `PipeGraphClassifier` uses the classifier default score (the mean accuracy of the prediction on given test data with respect to labels). As for the rest, both interfaces are equivalent.

The following functions can be used by the user in both interfaces:

- `inject(sink, sink_var, source, source_var)` Defines a connection between two nodes of the graph declaring which variable (`source_var`) from the origin node (`source`) is passed to the destination node (`sink`) with new variable name `sink_name`).
- `decision_function(X)` Applies `PipeGraphClasifier`'s predict method and returns the decision_function output of the final estimator.
- `fit(X, y=None, **fit_params)` Fits the `PipeGraph` steps one after the other and following the topological order of the graph defined by the `connections` attribute.
- `fit_predict(X, y=None, **fit_params)` Applies `predict` of a `PipeGraph` to the data following the topological order of the graph, followed by the `fit_predict` method of the final step in the `PipeGraph`. Valid only if the final step implements `fit_predict`.
- `get_params(deep=True)` Gets parameters for an estimator.
- `predict(X)` Predicts the `PipeGraph` steps one after the other and following the topological order defined by the `alternative_connections` attribute, in case it is not *None*, or the `connections` attribute otherwise.
- `predict_log_proba(X)` Applies `PipeGraphRegressor`'s predict method and returns the `predict_log_proba` output of the final estimator.
- `predict_proba(X)` Applies `PipeGraphClassifier`'s predict method and returns the `predict_proba` output of the final estimator.
- `score(X, y=None, sample_weight=None)` Applies `PipeGraphRegressor`'s predict method and returns the score output of the final estimator.
- `set_params(**kwargs)` Sets the parameters of this estimator. Valid parameter keys can be listed with `get_params()`.

4. Case Studies

4.1. Anomaly Detection in Manufacturing Processes

The first case study deals with anomaly detection of machined workpieces using a computer vision system [29]. In that paper, a set of four classifiers were tested in order to choose the best model for identifying the presence of wear along the workpiece surface. Following a workflow consisting of a preprocessing phase followed by feature extraction and finally a classification step, the authors reported satisfactory results. In such scenario, the results of a deeper analysis could have been provided, where the quality of the classifier could have been improved by unleashing an additional parameter, the number of classes. Instead of assuming that two classes are present in the dataset, namely correct pieces and unacceptable pieces, according to the finishing quality, the result from considering more classes can be explored.

For that purpose, a pipegraph model as the one shown in Figure 3 allows for a two model workflow. In this case, a clustering algorithm partitions a dataset consisting of 10,000 artificially generated 5-dimensional samples. It is worth noting that the partition is performed considering a specific and predefined number of clusters. This experiment splits the dataset in 5 folds for cross–validation. For each configuration, the classifier is trained and the quality of its results is assessed according to some convenient metric.The classifier considered was the well–known *k*-means training for a maximum of 300 iterations and a stopping criterion of error tolerance smaller than 0.0001.

Figure 4 displays the quality results for such workflow in an illustrative case, described in more detail in the library documentation. It is worth noting that such a workflow cannot be constructed via standard `scikit-learn` tools as the two steps of the workflow are models and the standard Pipeline class only allows for a set of transformers and a single unique model in final position. We claim that for those purposes where different models are useful, even in a linear sequence, `PipeGraph` provides a viable and convenient approach.

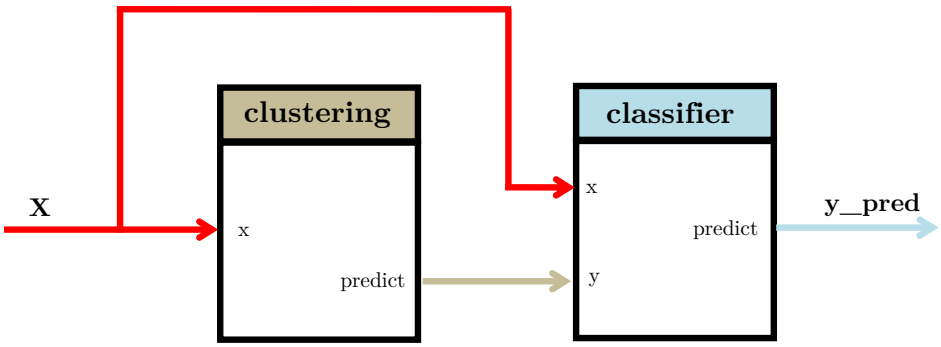


Figure 3. A pipegraph sporting two sequential model blocks.

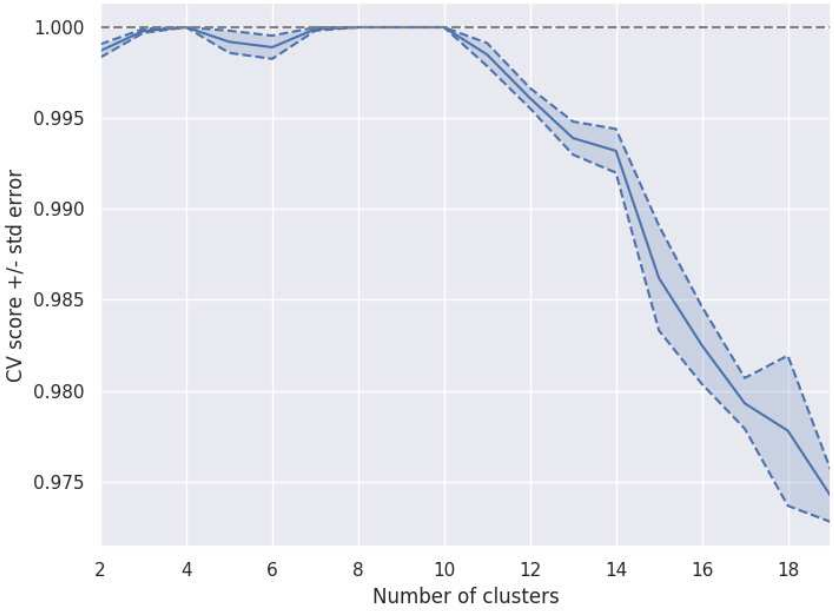


Figure 4. Number of clusters vs. a quality measure to assess the most appropriate number of clusters.

4.2. Heat Exchanger Modeling

The second case study deals with a problem that is common in manufacturing processes: the management of faulty sensors [30]. In that paper, 52,699 measurements from a sensor embedded in a heat exchanger system were compared to the predictions from a baseline model capturing the expected behavior of such CPS system. If the sensor measurements were significantly different from the predictions provided by the CPS model, an alarm was raised and specific actions performed according to the particular case study. In the paper, two classes were again considered, namely day and night, standing for the two particular periods in which the 24 h are split. The best model obtained from a Cross-Validation setup using 10 folds was an Extremely Randomized Tree whose training explored a range from 10 through 100 base estimators. A pipegraph similar to the one considered in Figure 2 was used in the experience reported in the paper for a two model for two classes case. For such scenario, an approach using more than two classes could have been considered to check if such an enhanced model can outperform the results reported. Figure 5 displays a unified workflow in which `PipeGraph` is capable of automatically wiring as many local

models as classes are defined in the dataset, thus relieving the user from the task of defining multiple configurations depending on how many classes were considered. It is again worth noting that such a workflow cannot be constructed via standard `scikit-learn` tools for the non linear workflow necessary for the purpose, and because of the automatic building procedure of the multiple prediction models put to play. We claim again that for those purposes where different models are used in a non linear sequence, `PipeGraph` provides a viable and convenient approach.

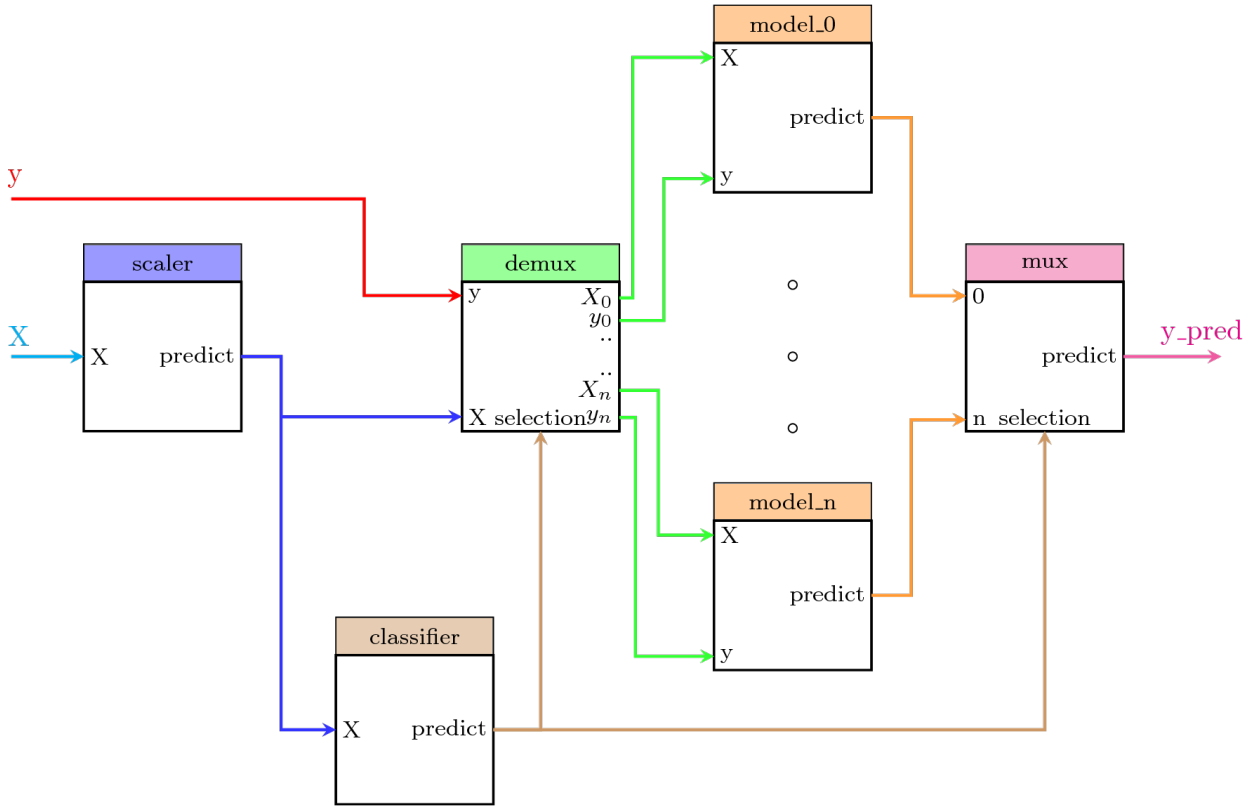


Figure 5. Parallel workflow with automatic wiring of the prediction models according to the labels contained in a dataset.

5. Conclusions

CPS and LM can greatly leverage on improvements in the tools and techniques available for system modeling. Data leakage, being one of the most common sources of unexpected behavior when the fitted models are stressed with actual demands, can largely be prevented by using encapsulation techniques such as the Pipeline provided by the `scikit-learn` library. For some specific complex problem appearing in the context of LM, and particularly in CPS modeling, the `PipeGraph` library provides a solution to building complex workflows, which is specially important for those cases that the standard `Pipeline` provided by `scikit-learn` cannot handle. Parallel blocks in non linear graph are available to unleash the creativity of the data scientist in the pursue of a simple and yet efficient model. In this paper we briefly introduced a novel `PipeGraph` toolbox for easily expressing ML models using DAG while being compatible with `scikit-learn`. We showed the potential of the toolbox and the underlying implementation details. We provided references to two case studies with hints on approaches for possible improvements by using `PipeGraph` and minor changes to the architecture proposed in two papers in the field of CPS and LM modeling. Future works on `PipeGraph` will include a Graphical User Interface (GUI) to use the API and new libraries which will encompass custom blocks related to specific areas connected to machine learning, such as computer vision, control systems, etc.

Author Contributions: Conceptualization, M.C.-L. and L.F.-R.; Formal analysis, M.C.-L., L.F.-R. and H.A.-M.; Investigation, M.C.-L. and L.F.-R.; Methodology, M.C.-L. and C.F.-L.; Software, M.C.-L., L.F.-R. and H.A.-M.; Validation, M.C.-L., L.F.-R., H.A.-M., J.C.-R. and C.F.-L.; Visualization, M.C.-L. and J.C.-R.; Writing—original draft, M.C.-L. and L.F.-R.; Writing—review & editing, M.C.-L., L.F.-R., H.A.-M., J.C.-R. and C.F.-L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Spanish Ministry of Economy, Industry and Competitiveness grant number DPI2016-79960-C3-2-P. The research described in this article has been partially funded by the grant RTI2018-100683-B-I00 funded by MCIN/AEI/10.13039/501100011033 and by "ERDF A way of making Europe".

Institutional Review Board Statement: Not applicable

Informed Consent Statement: Not applicable

Data Availability Statement: The source code is available online at https://github.com/mcasl/PipeGraph (accessed on 20 December 2021), and the API documentation is available at https://mcasl.github.io/PipeGraph/api.html (accessed on 20 December 2021).

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Listing A1. Example code for the PipeGraph shown in Figure 2.

```
import numpy as np
import pandas as pd
from sklearn.preprocessing import MinMaxScaler
from sklearn.linear_model import LinearRegression
from sklearn.model_selection import GridSearchCV
from pipegraph.base import PipeGraphRegressor
from pipegraph.base import Demultiplexer
from pipegraph.base import Multiplexer
import matplotlib.pyplot as plt
from sklearn.mixture import GaussianMixture

X_first = pd.Series(np.random.rand(100,))
y_first = pd.Series(4 * X_first + 0.5*np.random.randn(100,))
X_second = pd.Series(np.random.rand(100,) + 3)
y_second = pd.Series(-4 * X_second + 0.5*np.random.randn(100,))
X_third = pd.Series(np.random.rand(100,) + 6)
y_third = pd.Series(2 * X_third + 0.5*np.random.randn(100,))

X = pd.concat([X_first, X_second, X_third], axis=0).to_frame()
y = pd.concat([y_first, y_second, y_third], axis=0).to_frame()

scaler = MinMaxScaler()
gaussian_mixture = GaussianMixture(n_components=3)
demux = Demultiplexer()
lm_0 = LinearRegression()
lm_1 = LinearRegression()
lm_2 = LinearRegression()
mux = Multiplexer()


steps = [('scaler', scaler),
('classifier', gaussian_mixture),
('demux', demux),
```

```
('lm_0', lm_0),
('lm_1', lm_1),
('lm_2', lm_2),
('mux', mux), ]

connections = {
'scaler': {'X': 'X'},
'classifier': {'X': 'scaler'},
'demux': {'X': 'scaler', 'y': 'y',...
'selection': 'classifier'},
'lm_0': {'X': ('demux', 'X_0'), 'y': ('demux', 'y_0')},
'lm_1': {'X': ('demux', 'X_1'), 'y': ('demux', 'y_1')},
'lm_2': {'X': ('demux', 'X_2'), 'y': ('demux', 'y_2')},
'mux': {'0': 'lm_0', '1': 'lm_1', '2': 'lm_2',...
'selection': 'classifier'},
}

pgraph = PipeGraphRegressor(steps=steps,...
fit_connections=connections)
pgraph.fit(X, y)
y_pred = pgraph.predict(X)
plt.scatter(X, y)
plt.scatter(X, y_pred)
```

References

1. Lee, J.; Bagheri, B.; Kao, H.A. A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems. *Manuf. Lett.* **2015**, *3*, 18–23. [CrossRef]
2. Pereira, A.; Romero, F. A review of the meanings and the implications of the Industry 4.0 concept. *Procedia Manuf.* **2017**, *13*, 1206–1214. [CrossRef]
3. Nawanir, G.; Lim, K.T.; Othman, S.N.; Adeleke, A.Q. Developing and validating lean manufacturing constructs: An SEM approach. *Benchmark. Int. J.* **2018**, *25*, 1382–1405. [CrossRef]
4. Antony, J.; Psomas, E.; Garza-Reyes, J.A.; Hines, P. Practical implications and future research agenda of lean manufacturing: A systematic literature review. *Prod. Plan. Control* **2021**, *32*, 889–925. [CrossRef]
5. Ghobadian, A.; Talavera, I.; Bhattacharya, A.; Kumar, V.; Garza-Reyes, J.A.; O'Regan, N. Examining legitimatisation of additive manufacturing in the interplay between innovation, lean manufacturing and sustainability. *Int. J. Prod. Econ.* **2020**, *219*, 457–468. [CrossRef]
6. Maware, C.; Okwu, M.O.; Adetunji, O. A systematic literature review of lean manufacturing implementation in manufacturing-based sectors of the developing and developed countries. *Int. J. Lean Six Sigma* 2021, *ahead-of-print*. [CrossRef]
7. Psomas, E. Future research methodologies of lean manufacturing: A systematic literature review. *Int. J. Lean Six Sigma* **2021**, *12*, 1146–1183. [CrossRef]
8. Azadeh, A.; Yazdanparast, R.; Zadeh, S.A.; Zadeh, A.E. Performance optimization of integrated resilience engineering and lean production principles. *Expert Syst. Appl.* **2017**, *84*, 155–170. [CrossRef]
9. Kravets, A.G.; Bolshakov, A.A.; Shcherbakov, M.V. (Eds.) *Cyber-Physical Systems: Advances in Design & Modelling*; Springer International Publishing: Cham, Switzerland, 2020; Volume 259. [CrossRef]
10. Kirchhof, J.C.; Michael, J.; Rumpe, B.; Varga, S.; Wortmann, A. Model-Driven Digital Twin Construction: Synthesizing the Integration of Cyber-Physical Systems with Their Information Systems. In Proceedings of the 23rd ACM/IEEE International Conference on Model Driven Engineering Languages and Systems, MODELS '20, Virtual, 16–23 October 2020; Association for Computing Machinery: New York, NY, USA, 2020; pp. 90–101. [CrossRef]
11. Wang, T.; Wang, X.; Ma, R.; Li, X.; Hu, X.; Chan, F.T.S.; Ruan, J. Random Forest-Bayesian Optimization for Product Quality Prediction With Large-Scale Dimensions in Process Industrial Cyber–Physical Systems. *IEEE Internet Things J.* **2020**, *7*, 8641–8653. [CrossRef]
12. Banerjee, S.; Balas, V.E.; Pandey, A.; Bouzefrane, S. Towards Intelligent Optimization of Design Strategies of Cyber-Physical Systems: Measuring Efficacy through Evolutionary Computations. In *Computational Intelligence in Emerging Technologies for Engineering Applications*; Springer International Publishing: Cham, Switzerland, 2020; pp. 73–101. [CrossRef]
13. Tran, H.D.; Yang, X.; Manzanas Lopez, D.; Musau, P.; Nguyen, L.V.; Xiang, W.; Bak, S.; Johnson, T.T. NNV: The Neural Network Verification Tool for Deep Neural Networks and Learning-Enabled Cyber-Physical Systems. In *Computer Aided Verification*; Lahiri, S.K., Wang, C., Eds.; Springer International Publishing: Cham, Switzerland, 2020; pp. 3–17.

14. Portugal, I.; Alencar, P.; Cowan, D. The use of machine learning algorithms in recommender systems: A systematic review. *Expert Syst. Appl.* **2018**, *97*, 205–227. [CrossRef]
15. Bischl, B.; Lang, M.; Kotthoff, L.; Schiffner, J.; Richter, J.; Studerus, E.; Casalicchio, G.; Jones, Z.M. mlr: Machine Learning in R. *J. Mach. Learn. Res.* **2016**, *17*, 5938–5942.
16. Meng, X.; Bradley, J.; Yavuz, B.; Sparks, E.; Venkataraman, S.; Liu, D.; Freeman, J.; Tsai, D.; Amde, M.; Owen, S.; et al. MLlib: Machine Learning in Apache Spark. *J. Mach. Learn. Res.* **2016**, *17*, 1–7.
17. Heaton, J. Encog: Library of Interchangeable Machine Learning Models for Java and C#. *J. Mach. Learn. Res.* **2015**, *16*, 1243–1247.
18. Curtin, R.R.; Cline, J.R.; Slagle, N.P.; March, W.B.; Ram, P.; Mehta, N.A.; Gray, A.G. mlpack: A Scalable C++ Machine Learning Library. *J. Mach. Learn. Res.* **2013**, *14*, 801–805.
19. King, D.E. Dlib-ml: A Machine Learning Toolkit. *J. Mach. Learn. Res.* **2009**, *10*, 1755–1758.
20. Hall, M.; Frank, E.; Holmes, G.; Pfahringer, B.; Reutemann, P.; Witten, I.H. The WEKA Data Mining Software: An Update. *SIGKDD Explor. Newsl.* **2009**, *11*, 10–18. [CrossRef]
21. Abeel, T.; Van de Peer, Y.; Saeys, Y. Java-ML: A Machine Learning Library. *J. Mach. Learn. Res.* **2009**, *10*, 931–934.
22. Jing, R.; Sun, J.; Wang, Y.; Li, M.; Pu, X. PML: A parallel machine learning toolbox for data classification and regression. *Chemom. Intell. Lab. Syst.* **2014**, *138*, 1–6. [CrossRef]
23. Lauer, F. MLweb: A toolkit for machine learning on the web. *Neurocomputing* **2018**, *282*, 74–77. [CrossRef]
24. Gashler, M. Waffles: A Machine Learning Toolkit. *J. Mach. Learn. Res.* **2011**, *12*, 2383–2387.
25. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
26. Munappy, A.R.; Mattos, D.I.; Bosch, J.; Olsson, H.H.; Dakkak, A. From Ad-Hoc Data Analytics to DataOps. In Proceedings of the International Conference on Software and System Processes, ICSSP '20, Seoul, Korea, 26–28 June 2020; Association for Computing Machinery: New York, NY, USA, 2020; pp. 165–174. [CrossRef]
27. Ntalampiras, S.; Potamitis, I. A Concept Drift-Aware DAG-Based Classification Scheme for Acoustic Monitoring of Farms. *Int. J. Embed. Real Time Commun. Syst.* **2020**, *11*, 62–75. [CrossRef]
28. Géron, A. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*, 2nd ed; O'Reilly Media: Sebastopol, CA, USA, 2019.
29. Sánchez-González, L.; Riego, V.; Castejón-Limas, M.; Fernández-Robles, L. Local Binary Pattern Features to Detect Anomalies in Machined Workpiece. In *Hybrid Artificial Intelligent Systems*; de la Cal, E.A., Villar Flecha, J.R., Quintián, H., Corchado, E., Eds.; Springer International Publishing: Cham, Switzerland, 2020; pp. 665–673.
30. Aláiz-Moretón, H.; Castejón-Limas, M.; Casteleiro-Roca, J.L.; Jove, E.; Fernández Robles, L.; Calvo-Rolle, J.L. A Fault Detection System for a Geothermal Heat Exchanger Sensor Based on Intelligent Techniques. *Sensors* **2019**, *19*, 2740. [CrossRef]

sensors

MDPI

Article

Optimisation of Maintenance Policies Based on Right-Censored Failure Data Using a Semi-Markovian Approach

Antonio Sánchez-Herguedas [1], Angel Mena-Nieto [2], Francisco Rodrigo-Muñoz [3], Javier Villalba-Díez [4] and Joaquín Ordieres-Meré [5,*]

[1] Department of Industrial Management, School of Engineering, University of Seville, Camino de los Descubrimientos s/n, 41092 Seville, Spain; antoniosh@us.es
[2] Department of Electrical and Thermal Engineering, Design, and Projects, School of Engineering, Campus El Carmen, University of Huelva, 21819 Huelva, Spain; mena@uhu.es
[3] Department of Applied Mathematics II, School of Engineering, University of Seville, Camino de los Descubrimientos s/n, 41092 Seville, Spain; frodrigo@us.es
[4] Hochschule Heilbronn, Fakultät Management und Vertrieb, Campus Schwäbisch Hall, 74523 Schwäbisch Hall, Germany; javier.villalba-diez@hs-heilbronn.de
[5] Escuela Técnica Superior de Ingenieros Industriales (ETSII), Universidad Politécnica de Madrid, 28006 Madrid, Spain
* Correspondence: j.ordieres@upm.es; Tel.: +34-910677107

Abstract: This paper exposes the existing problems for optimal industrial preventive maintenance intervals when decisions are made with right-censored data obtained from a network of sensors or other sources. A methodology based on the use of the z transform and a semi-Markovian approach is presented to solve these problems and obtain a much more consistent mathematical solution. This methodology is applied to a real case study of the maintenance of large marine engines of vessels dedicated to coastal surveillance in Spain to illustrate its usefulness. It is shown that the use of right-censored failure data significantly decreases the value of the optimal preventive interval calculated by the model. In addition, that optimal preventive interval increases as we consider older failure data. In sum, applying the proposed methodology, the maintenance manager can modify the preventive maintenance interval, obtaining a noticeable economic improvement. The results obtained are relevant, regardless of the number of data considered, provided that data are available with a duration of at least 75% of the value of the preventive interval.

Keywords: maintenance interval; maintenance model; semi-Markov process; right-censored data; finite horizon; maintenance cost

Citation: Sánchez-Herguedas, A.; Mena-Nieto, A.; Rodrigo-Muñoz, F.; Villalba-Diez, J.; Ordieres-Meré, J. Optimisation of Maintenance Policies Based on Right-Censored Failure Data Using a Semi-Markovian Approach. *Sensors* **2022**, *22*, 1432. https://doi.org/10.3390/s22041432

Academic Editor: Anastasios Doulamis

Received: 29 December 2021
Accepted: 10 February 2022
Published: 13 February 2022

1. Introduction

The main goal of this work is to present a methodology that allows finding the optimal maintenance preventive interval when the failure data are right-censored. In particular, we are interested in finding out how the use of right-censored data affects the calculation of the optimal maintenance interval, considering the temporary maintenance schedule of maintenance interventions and the different types of costs incurred. This study is relevant, since these are the types of data that are generally available in the vast majority of companies, which the maintenance manager must use.

When the maintenance engineer is faced with the problem of managing equipment subject to predetermined preventive maintenance, one of the tasks is to determine the preventive maintenance interval that economically optimises (other factors could also be optimised) the series of interventions that can be carried out on a physical asset over a specified period [1]. For this predetermined maintenance, the interventions can be of two types: corrective interventions originated by a failure of some component of the equipment (failure mode) and preventive interventions carried out after a certain number of hours, kilometres, or units produced. The former is produced randomly, and their number and

location in time depend on the failure behaviour of the component with that particular failure mode. The second can be located in time, and the number of can be established along any considered time horizon. In those cases where the failure mode shows wear behaviour, these types of interventions are not independent of each other, since increasing the number of these interventions reduces the number of those due to component failures. However, the increase in the number of the second type of intervention implies an increase in the associated costs of preventive maintenance, but a reduction in the costs of corrective maintenance. Therefore, the optimal preventive maintenance interval that minimises maintenance costs (preventive and corrective) must be found.

To find the optimal interval, the starting point is the existing status and the available information, which must reflect the behaviour in the event of failure, the costs of maintenance interventions, and the income generated by the operation of the equipment [2]. The information on the cost of each intervention is generated during the operation and maintenance phase of the equipment and is usually collected in the Computer-Aided Maintenance Management Systems (CMMS). Once the learning period for maintenance tasks is over, this information is usually relatively stable over time. In the same way, the information on the income from the operation of the equipment is usually collected in the Enterprise Resource Planning systems (ERPs) [3]. Information on behaviour in the event of failure can be obtained from various sources, the equipment manufacturer, the history of failures collected in the plant itself, or any other external database that presents operating conditions similar to that of the equipment under study. The information referring to equipment failures is usually right-censored due to the performance of preventive maintenance. On rare occasions, the maintenance manager has failure data where there is no right-censored data. This information, if it exists, is usually in the hands of manufacturers who have tested their equipment up to its failure point. The maintenance manager cannot bear the test cost, so the adopted policy usually follows the preventive interval set by the manufacturer. Such settings are frequently conditioned by economic interests or brand image. In the research carried out, an extensive data set was available; therefore, the knowledge based impact of right censoring can be identified.

The literature on maintenance is vast, including Reliability Engineering [4–12], maintenance policies and modelling [13–20], and optimal preventive maintenance intervals [21–24]. However, fewer papers address the problem of uncertainty in lifetime distribution [25–28], either by analysing several time-based maintenance policies having uncertainty in the parameters of the lifetime distribution [29] or using right-censored data [30], and adopt the Markovian approach of transition between states (operational, preventive, and corrective maintenance intervention) [31–36].

The discrete-time and continuous-time Markov models satisfy the Markov property, i.e., the future only depends on the present, not on the past. Only the transitions between states and their respective probabilities are considered in the discrete-time model (Markov chain). The sojourn time in each state is irrelevant. In the continuous-time model, the sojourn time is a random variable with an exponential distribution due to the Markovian property. By contrast, in semi-Markov models, the sojourn time in each state does not follow an exponential distribution. It implies that semi-Markov models do not satisfy the Markovian property, which relaxes the constraints and improves the application value [37]. Therefore, successive transitions between states form a Markov chain, called an embedded Markov chain in the semi-Markov model. A wider review, which provides an interesting perspective, can be found in [7,38,39].

Additionally, other authors address the problem of using right-censored data by proposing alternative methods, for example, Li et al. [40] found the life distributions using a histogram-based technique to graphically obtain the maintenance interval. Mazzuchi and van Dorp [26] used the Newton–Raphson method and MLE to try to represent the remaining life of coupler knuckles in railway wagons. Taghipour and Banjevic [41] used the likelihood ratio test to check for trends in the failure data and the ME algorithm to find the parameters for trend analysis applied to pumps located in hospitals.

However, to the best of our knowledge, very few studies provide rules and guidelines to modify the default maintenance interval optimally, and this paper helps to fill that gap. When the maintenance manager has censored data due to the implementation of preventive tasks, the proposed methodology represents an advance with respect to previous works such as [21–24,42].

This article differs from previous studies in providing the maintenance manager with a prediction of the length of the optimal preventive interval as well as a decision rule to improve the interval being used, based on the analysis of the results of various scenarios, and by using the methodology able to deal with right-censored data. This rule guides the manager in the direction of modification of the preventive interval to be used (increasing or decreasing it) and the intensity of the modification. In this way, the uncertainty induced by using censored data can be overcome. Therefore, to fill the identified gaps, two research questions are formulated:

RQ1: How far is the obtained result from the optimal solution when the failure data are censored?

RQ2: What strategy would help to move closer to the optimum based on censored data?

The developed methodology guarantees that the solution found is optimal from a mathematical point of view, incorporating the time right-censored data, together with different types of costs and income. In addition, it applies a Semi–Markovian approach of transition between states [43–45], plus the z-transform, to find the optimal maintenance interval, which, as far as we know, is the first time it has been applied with right-censored data. It is also relevant that the developed methodology enables the assessment and comparison of alternative maintenance policies.

The rest of the paper is organised as follows. Section 2 provides the real data for this study, the research questions formulated, and the methodology applied to develop the mathematical optimisation model based on semi-Markov processes and z-transform. Section 3 presents the results for the different cases studied. In Section 4, those results are highlighted, discussed, and compared. Finally, Section 5 provides the conclusions drawn from the study.

2. Materials and Methods

To adapt the preventive interval to the particular equipment conditions usage, the manager responsible for maintenance should use the information generated by the equipment itself. The failure information generated is right-censored and it is a challenge to determine the behaviour before the failure of the equipment. The method followed to find the optimal interval when the available failure data is right-censored is explained throughout the article. This process is graphically summarised in Figure 1.

2.1. Real Case—Data Selection and Information Processing

This work analyses the behaviour of the O-rings located in the cooled crossover of a 12V diesel engine with 2 litres per cylinder. The crossover is a jacketed tube through which exhaust gases flow. The engine coolant cools this tube, and two O-rings are installed at each end of the tube to prevent coolant leakage. High temperatures affect the O-rings and are responsible for their degradation. In this experiment, failure data have been collected for several years, following the preventive replacement policy of O-rings every 4000 h (scenario A). During this time, it was observed that none of the O-rings had reached this limit, remaining at values very far from this value. This fact allows us to consider that the values of the observed failures can be considered as "not" censored. These failures are listed in Table 1. In addition, in some cases, the O-rings were replaced preventively because the last failure occurred close to the time of the preventive replacement. These are the censored values that are collected in Table 2.

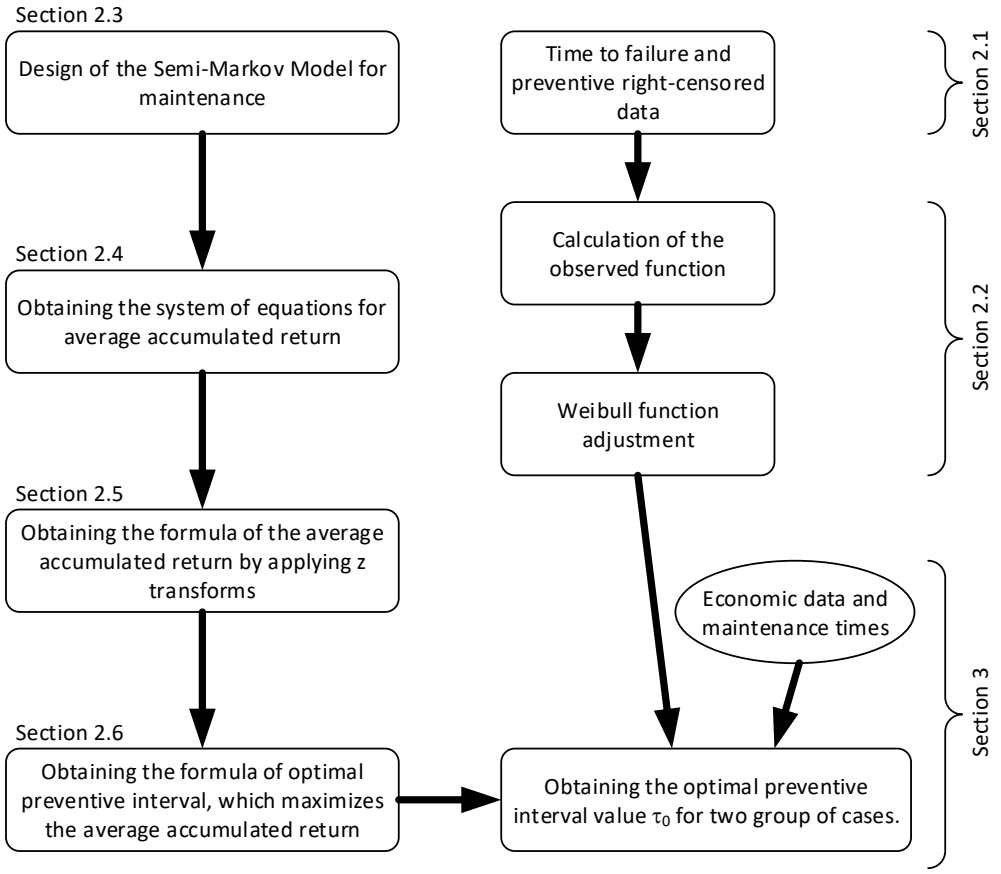


Figure 1. Graphical overview of the process.

Table 1. Hours of operation of the O-rings until failure in Scenario A.

Operation Hours to Failure											
190	276	296	409	429	430	437	454	481	492	498	499
543	552	552	552	552	577	603	604	604	612	619	658
675	683	696	702	742	754	773	797	812	836	881	889
912	913	942	974	994	994	1014	1015	1024	1025	1041	1105
1183	1203	1211	1236	1238	1240	1249	1274	1295	1304	1312	1343
1345	1407	1413	1421	1442	1447	1486	1492	1542	1581	1601	1621
1630	1675	1735	1892	1926	1960	2006	2101	2242	2437	2450	

Table 2. Hours of operation of the O-rings with censorship due to preventive maintenance policy in scenario A.

Operation Hours to the Censorship											
45	84	84	103	121	176	199	217	259	324	371	387
428	508	514	519	638	655	661	735	760	764	766	769
790	867	986	1006	1011	1111	1167	1188	1188	1395	1396	1505
1752	2016										

Subsequently, the value of the preventive interval was modified, settling at 1000 h. In this case, the failure data were considerably reduced as many observations were right-censored by the predetermined preventive maintenance performed at 1000 h (scenario B). This censorship radically changed the way of analysing the problem, since the results will depend less on the type of technique or method used and more on the type of censorship [46] and the extent of the censorship. Scenario B, corresponding to the preventive interval of 1000 h, is the one that is usually presented to the maintenance manager.

Failure data and data censored by preventive corresponding to scenario B are, respectively, represented in Tables 3 and 4. In Table 3, only seven of the nine failures that occurred during the period studied appear, since two of them, those corresponding to 171 h, did not correspond to the failure mode studied.

Table 3. Hours of operation of the O-rings until failure in Scenario B.

Operation Hours to Failure						
242	480	522	633	663	839	845

Table 4 contains 25 censored values and 87 values corresponding to a 1000-h preventive cycle completed.

Table 4. Hours of operation of the O-rings with censorship due to preventive maintenance policy in scenario B.

Operation Hours to the Censorship											
93	93	128	128	128	128	155	161	220	220	240	240
337	367	380	380	467	467	478	485	485	520	758	829
829	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
1000	1000	1000	1000	1000	1000	1000	1000	1000	1000		

2.2. Determination of the Failure Distribution Function

Data from maintenance interventions are collected at the machine or through an automated procedure using a network of sensors. Its mathematical use for optimisation requires treating these data until obtaining information on the appropriate protocol that can be understood by the model proposed for the simulation. This treatment often supposes a loss of veracity. In other cases, it may happen that the information on which the treatment is started is not adequate. Achieving the most reliable information to the original exposed in the desired protocol is one of the objectives to be achieved in any optimisation process.

In our case, the original starting data could be classified into two groups, those generated during the maintenance activities and those generated during the economic management of those maintenance activities. The first group included failure data and preventive activity data. These data are the hours of operation of the equipment in which preventive and corrective interventions are carried out and the duration of these interventions. In the first group, they were counted from the last repair or preventive change until the failure appeared or until a preventive task was carried out, and the equipment was inactive due to the intervention. The second group of data included the costs of each of the corrective and preventive interventions that were carried out and the income obtained from the use of the equipment.

For the first data group, it was necessary to establish the method to determine the method of mathematically obtaining the failure distribution. Many examples can be found in the literature [47,48]. In our case, a Weibull probability distribution function was used, although other authors analyse different methods to estimate the parameters of the Weibull distribution [49,50]. Other authors have previously used other distributions. However, the Weibull distribution is the one that best adapts to the process of failure appearance in industrial assets. Furthermore, the Weibull distribution includes the exponential distribution and has the advantage of using two or three parameters instead of a single parameter of the

exponential function. The exponential survival function has the hazard rate constant, while the Weibull survival distribution extends the exponential distribution to allow constant, increasing, or decreasing hazard rates.

The estimation of the model parameters becomes a difficult task when censorship radically changes the available information and when the opinion of experts does not help to understand the system's behaviour [51].

Many authors have estimated the parameters of the Weibull distribution for censored data through the Maximum Likelihood Estimation method (MLE), using algorithms for their numerical resolution such as the Maximised Expectation (EM) algorithm, see [52–54]. Other authors use and compare various methods such as the Synthetic Minority Oversampling TEchnique (SMOTE) [55], MLE, Least Square Estimation (LSE) [56], Weibull [57] probability plot, regression of the range of medians with the Benard approximation, and even combine methods, see [58,59]. Bayesian estimators [60] and other types of linear estimators [61] have also been used.

At the same time, the repair and preventive replacement average times and the cost and income average of each type of intervention (corrective or preventive) are also calculated. This last type of data generally poses more difficulties to obtain than to calculate it. However, despite not being as obvious as the calculation of the previous means, the calculation of the failure distribution function can lead to significant management errors, mainly due to the lack of coherence of the starting data. When these failure data are right–censored, due to the performance of preventive maintenance, poor optimisation of the preventive interval will be achieved. This article focuses on the influence of the maintenance management data (first group) to determine the distribution function to perform the preventive interval optimisation calculations.

In this research, a reduced number of data will be used, with which we will obtain an exact solution for the optimal preventive interval. If we had used a large number of data, as is the case with information collected online through sensors, the result would have been the same (exact solution). Nevertheless, in both cases, as will be shown, the results will be far from the optimal solution.

The equipment was failing and was repaired, with few cases of preventive interventions (when the scheduled time according to predetermined maintenance was reached), which also lasted for a few hours of operation. This case represents a situation where the censorship was practically nonexistent, and the observed distribution function was very close to the real one. On the other hand, Tables 3 and 4 represent the failure and preventive data when a predetermined preventive threshold was established at 1000 h of operation. In this second case, the behaviour of the O-rings would not be known beyond 1000 h. In both cases, it does make sense to estimate the behaviour beyond the predetermined threshold or to analyse the differences between the optimum intervals in both cases. Therefore, a procedure to obtain estimations from the second case closer to those obtained for the first one can be distilled. Looking to gather more information, the following procedure was adopted:

- From the failure data and preventive maintenance data collected in Tables 3 and 4, the observed function was calculated applying the range of medians method and using the Benard approximation.
- From the observed function, the theoretical Weibull function that fit the best was determined. Then, the least-squares method was used to obtain the theoretical function's two or three parameters.
- Subsequently, the theoretical function was used to optimise the preventive interval economically.

In Figure 2, the procedures for the data in Tables 3 and 4 are summarised. The Weibull failure distribution function that best fit the observed function had two parameters, the shape parameter with value 1.88 and a scale parameter (characteristic life) with value 3603. These data were very different from the data obtained for the case of uncensored data and would give rise to a very different preventive interval. The detailed process for

the data in Tables 1 and 2 can be verified in Sánchez-Herguedas et al. [42]. The Weibull functions adjusted to the uncensored data were, in that case, W(2.36, 1317, 0) and W(1.95, 1202, 117), where the first value corresponded to the shape parameter, the second to the scale parameter, and the third to the guaranteed life parameter.

Hours to failure, t	Observed Distribution, F(t)	ln t	ln ln (1/(1-F(t)))
242	0.00679	5.488938	-4.988423
480	0.01675	6.173786	-4.080898
522	0.02702	6.257668	-3.597559
633	0.03729	6.450470	-3.270167
663	0.04756	6.496775	-3.021595
839	0.05817	6.732211	-2.814639
845	0.06878	6.739337	-2.641473

α =	1.88
β =	3,603.25
γ =	0.00

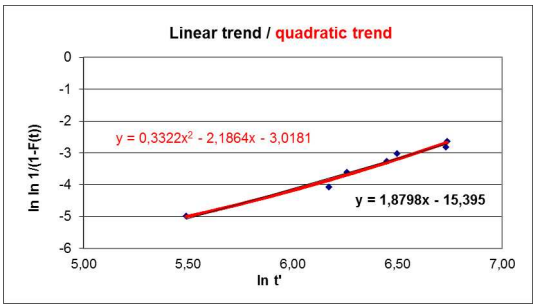


Figure 2. Weibull distribution function obtained from the observed function, which has been determined by Median Rank Regression (MRR) procedure and the Benard approximation.

2.3. Semi-Markov Maintenance Model with Returns for a Finite Period

To study the behaviour of O-rings concerning failure, we need a mathematical model that reflects this behaviour, such as the semi–Markovian model of three states, which is used with the data that appear in Tables 1 and 2. This model applies to the predetermined preventive maintenance of much industrial equipment, since it contains the two most representative maintenance states, the state of the equipment when corrective maintenance is performed after a failure (State S2) and the state of the equipment when preventive maintenance is performed after a predetermined number of operating hours (State S3). In these two cases analysed, the data correspond to preventive intervals of 4000 and 1000 h of operation, respectively. State S1 corresponds to the period in which the equipment is performing its required function. In this model, the equipment and others with the same characteristics evolve over time, changing the state according to the law of probability. When the system is in the operational state, the probability of equipment failure is distributed as a Weibull distribution. If the equipment fails, the system goes to the corrective state (S2) and develops a corrective intervention. Suppose the equipment does not fail after a time τ, the system transitions to the preventive state, developing a preventive intervention. Once these maintenance interventions (corrective and preventive) have been carried out, the system returns to its operational state. Both types of interventions have associated costs.

On the other hand, during the operational state, the equipment generates income. Over time, these costs and income are accumulated in a variable called the average accumulated return, $V(m)$. The optimisation of this variable will allow the calculation of the optimal preventive interval, τ_0. These transitions between states and the accumulation of returns are graphically expressed in Figure 3.

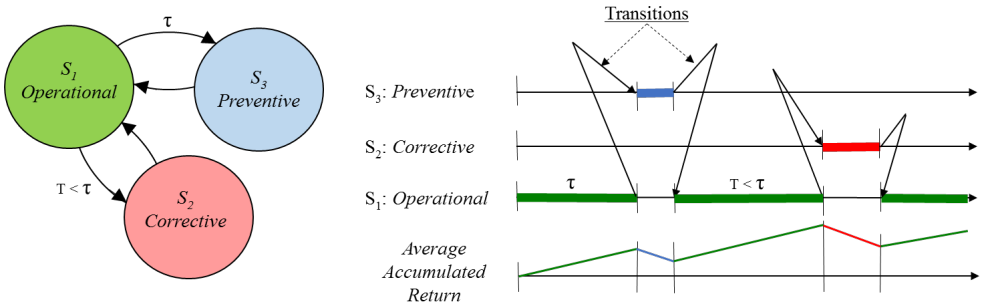


Figure 3. Transition process between states and accumulation of returns.

2.4. Formulating a Difference Equation System for Average Accumulated Return

The model that reflects the behaviour of the system is a semi–Markovian model. The variable to optimise is the average accumulated return in the successive transitions between states. In m transitions, starting from state i, the system accumulates m returns, which added to their respective signs (+ for income and − for costs) will constitute the accumulated return $Q_i(m)$ in m transitions from state i. For each value of i and m, the accumulated return is a random variable because once the initial state is determined, the next state is unpredictable, as well as the following states. Thus, in m transitions, the system can evolve in many ways. For this reason, it is interesting to know the average accumulated return instead of the accumulated return itself.

Let $v_i(m) = E(Q_i(m))$ be the average accumulated return in m transitions when the system starts from the initial state i. To calculate $v_i(m)$, a difference equation is constructed, separating the m transitions into two stages. The first stage is the one constituted by the transition from the initial state i to the next state j. As the second state can be any of the states of the system, the return $v_i(1)$ in a single transition is a random variable that can reach the values $r_{i1}(1), r_{i2}(1), \cdots, r_{in}(1)$, with the respective probabilities $p_{i1}, p_{i2}, \cdots, p_{in}$. The average return in that transition can be formulated as follows:

$$v_i(1) = \sum_{j=1}^{n} r_{ij}(1) \cdot p_{ij}. \tag{1}$$

Now, the process continues with the remaining $m-1$ transitions. Once the first transition has been made, the system is placed in a state called j, where j takes one of the values $1, 2, 3, \cdots, n$. The average accumulated return $v_j(m-1)$ in the following transitions is a random variable that can reach the values $v_1(m-1), v_2(m-1), \cdots, v_n(m-1)$, with probabilities of $p_{j1}, p_{j2}, \cdots, p_{jn}$, respectively, which remain constant throughout the m transitions, since the process is homogeneous. Therefore, the expected value of the return of the remaining $m-1$ transitions can be formulated as: $v_j(m-1)) = \sum_{q=1}^{n} v_j(m-1) \cdot p_{jq}$.

We conclude that the expected average return of the system in m transitions is calculated according to Equation (2):

$$v_i(m) = v_i(1) + v_j(m-1) \cdot p_j = v_i(1) + \sum_{q=1}^{n} v_j(m-1) \cdot p_{jq} \tag{2}$$

The vector $V(m)$ is defined as $V(m) = (v_1(m), v_2(m), \cdots, v_n(m))^t$, where the superscript t indicates the transpose in the matrix sense. This last equality can be written as a difference equation:

$$V(m) = V(1) + P \cdot V(m-1) \tag{3}$$

where P is the matrix of transition probabilities. The Equation (3) gives the average accumulated return in m transitions from any possible starting state i.

2.5. Solving the System of Difference Equations by Applying Transforms Z

Solving this difference equation requires the use of z transform and Laurent's series. To reach its resolution, it is necessary to include the data and functions involved in the calculation: distribution functions of the time spent in each state, the returns for each state, and the matrices that describe the process.

Time-related probabilistic functions for the tree states are as follows:

- $F(t)$ is the Cumulative Distribution Function (CDF) of equipment failures, and $f(t)$ is the Probability Density Function (PDF). Based on the discussion in Section 2.3, we will use the three-parametric Weibull distribution.
- $G(t_c)$ is the distribution function of the time the equipment remains under corrective maintenance and $g(t_c)$ is its probability density function.
- $H(t_p)$ is the distribution function of the time that the equipment remains under preventive maintenance and $h(t_p)$ is its probability density function.

Returns (costs for corrective and preventive states and income from operational state) are as follows:

- R_1, income per time unit that the system remains in State 1 (S_1: Operational), 6 €/h.
- R_{12}, the cost of transition from State 1 to State 2, −4320 €.
- R_{13}, the cost of transition from State 1 to State 3, −1 €.
- R_2, the cost per time unit that the system remains in State 2 (S_2: Corrective), −95 €/h.
- R_{21}, cost of transition from State 2 to State 1, −620 EUR.
- R_3, the cost per time unit that the system remains in State 3 (S_3: Preventive), −82 €/h.
- R_{31}, the cost of transition from State 3 to State 1, −620 €.

Matrices to describe the process are as follows:

- P, the transition probability matrix between states (p_{ij} is the probability of going from State i to j).
- F, stay time matrix (F_{ij} is the average time in State i before the system goes to State j).
- R, returns matrix, where $r_{ij} = R_i + R_{ij}$.

With these data, making the z-transform of $V(m+1) = V(1) + P \cdot V(m)$, during the calculation, it is required to determine the matrix $I - z^{-1}P$ to reach Equation (4).

$$Z[V(m)] = \frac{1}{z-1}V(1) + (I - z^{-1}P)^{-1} \cdot V(1) \tag{4}$$

Developing the matrix $(I - z^{-1}P)^{-1}$, the value of $Z[V(m)]$ is reached and decomposed into simple fractions for each of its terms $Z[v_1(m)]$, $Z[v_2(m)]$, and $Z[v_3(m)]$. Subsequently, using the appropriate Laurent expansion, the inverse z-transform of each one is calculated, obtaining the equations that describe the average accumulated returns in each transition for each of the starting states of the process, $v_1(m)$, $v_2(m)$, and $v_3(m)$. The full development can be followed in Sánchez Herguedas et al. [62]. The average accumulated return in m transitions starting in the operating state is described by Equation (5).

$$\begin{aligned} v_1(m) = \frac{1}{4}\Big[&(2m+1+(-1)^{m-1}) \cdot \Big(R_1 \cdot \int_0^\tau t \cdot f(t)dt + R_{12} \cdot F(\tau) + (R_1\tau + R_{13}) \cdot (1 - F(\tau))\Big) + \\ &(2m-1-(-1)^{m-1}) \cdot \Big(\Big(R_2 \cdot \int_0^\infty t_c \cdot g(t_c)dt_c + R_{21}\Big) \cdot F(\tau) + \\ &\Big(R_3 \cdot \int_0^\infty t_p \cdot h(t_p)dt_p + R_{31}\Big) \cdot (1 - F(\tau))\Big)\Big] \end{aligned} \tag{5}$$

The average accumulated return in m transitions starting in the corrective state is described by Equation (6). To deduce it, the same reasoning is followed as in the previous case.

$$\begin{aligned} v_2(m) = \frac{1}{4}\Big[&(2m-1-(-1)^{m-1}) \cdot \Big(R_1 \cdot \int_0^\tau t \cdot f(t)dt + R_{12} \cdot F(\tau) + (R_1 \cdot \tau + R_{13}) \cdot (1 - F(\tau))\Big) + \\ &(2m+1+(-1)^{m-1}) \cdot \Big(R_2 \cdot \int_0^\infty t_c \cdot g(t_c)dt_c + R_{21}\Big) + (2m-3+(-1)^{m-1}) \cdot \\ &\Big(\Big(R_3 \cdot \int_0^\infty t_p \cdot h(t_p)dt_p + R_{31}\Big) - \Big(R_2 \cdot \int_0^\infty t_c \cdot g(t_c)dt_c + R_{21}\Big)\Big) \cdot (1 - F(\tau))\Big] \end{aligned} \tag{6}$$

The average accumulated return in m transitions starting in the preventive state is described by Equation (7).

$$\begin{aligned} v_3(m) = \frac{1}{4}\Big[&(2m-1-(-1)^{m-1}) \cdot \Big(R_1 \cdot \int_0^\tau t \cdot f(t)dt + R_{12} \cdot F(\tau) + (R_1 \cdot \tau + R_{13}) \cdot (1 - F(\tau))\Big) + \\ &(2m+1+(-1)^{m-1}) \cdot \Big(R_3 \cdot \int_0^\infty t_p \cdot h(t_p)dt_p + R_{31}\Big) + \\ &(2m-3+(-1)^{m-1}) \cdot \Big(\Big(R_2 \cdot \int_0^\infty t_c \cdot g(t_c)dt_c + R_{21}\Big) - \Big(R_3 \cdot \int_0^\infty t_p \cdot h(t_p)dt_p + R_{31}\Big)\Big) \cdot F(\tau)\Big] \end{aligned} \tag{7}$$

2.6. Optimisation of the Averaged Accumulated Return When Starting from the Operational State

Deriving and equalising to zero is possible to achieve the mathematical formula of the preventive interval that economically optimises the average accumulated return for

each transition, starting from each of the states $\frac{dv_1(m)}{d\tau} = 0$, $\frac{dv_2(m)}{d\tau} = 0$, and $\frac{dv_3(m)}{d\tau} = 0$. The results obtained for each state are:

For the case of starting in the operational state, Equation (8).

$$\tau_0^{\alpha-1} = \frac{\beta^\alpha}{\alpha} \cdot \frac{-R_1}{R_{12} - R_{13} + \frac{(2m-1-(-1)^{m-1})}{(2m+1+(-1)^{m-1})} \cdot (R_2 \cdot B + R_{21} - R_3 \cdot C - R_{31})} \quad (8)$$

For cases of starting in the corrective and preventive states, Equation (9).

$$\tau_0^{\alpha-1} = \frac{\beta^\alpha}{\alpha} \cdot \frac{-R_1}{R_{12} - R_{13} + \frac{(2m-3+(-1)^{m-1})}{(2m-1-(-1)^{m-1})} \cdot (R_2 \cdot B + R_{21} - R_3 \cdot C - R_{31})} \quad (9)$$

3. Results

Applying the procedure followed in Section 2.2 to obtain the Weibull distribution function and Equation (8) to obtain the optimal preventive interval, the results expressed in Table 5 are reached. The average time of corrective and preventive tasks is eight and seven hours, respectively.

Table 5. Results of the cases (cases A, B, B-1, B-2, B-3, B-4, B-5, B-6) analysed. Parameters of the Weibull distribution function and the optimal preventive interval.

	Weibull			**Optimal**
	Shape Parameter α	**Scale Parameter β**	**Guaranteed Life γ**	**Preventive Interval τ_0**
Case A (uncensored)	2.36	1317	0	1059
Case B	1.88	3603	0	10,456
Case B-1	2.42	695	0	353
Case B-2	1.85	3729	0	11,914
Case B-3	1.90	3541	0	9805
Case B-4	2.65	1939	0	1908
Case B-5	2.40	2265	0	2663
Case B-6	2.48	2150	0	2369

Case A corresponded to the uncensored data presented in Tables 1 and 2. Case B corresponded to the data with censorship, when the maintenance policy implies a preventive change of the O-rings when reaching 1000 h of operation without failure or when the equipment reaches a total of hours that is a multiple of 1000.

In this case, the Weibull function was altered. The failures were spread over a more extended period, and 38% of them occurred after 3600 h. Obviously, this is not what happened when the data were uncensored, and 62% of the failures appeared before reaching 1300 h. The value of the optimal preventive interval given by the model was very high, 10456. This is due to the flattening and lengthening that the failure distribution curve suffered, as a consequence of the right-censored data. In cases B-2 and B-3, this situation also occurred. Case B-2 considered a failure and the first preventive maintenance at 1000 h. It affected the list of times considered in the Median Rank Regression method. Case B-3 was similar to the previous one; it considered a failure of a preventive maintenance of 1000 h ordered in half of the 87 preventives analysed in Table 4. If the calculations are made only taking into account the failure data and ignoring the censored data at 1000 h, case B-1 is the situation where there is a tendency to group the failures. However, this grouping occurred very early in the life of the O-rings and led to optimising the preventive interval at 353 h. Another group of cases was made up of cases B-4, B-5, and B-6. The distribution of failures was more concentrated, as in case A. However, it was concentrated towards a greater number of hours. For this reason, the optimal preventive interval was established

around 2000 h, well above 1059 in case A. All the cases were far from optimal, failing to find a strategy that was close to optimal.

Of course, using this censored data, the optimal preventive interval suitable for a correct economic management of maintenance would not be found. Finding a reason that explains this behaviour is not easy, but it could be thought that the data from A or B cases are not adequate, which could be expected, for example, when two or more failure modes are being mixed.

A further step can be taken in verification. The study can be carried out using the data from case A, Tables 1 and 2. We can take the data from Table 1 and artificially censor the failures greater than 1000 h. This is how Table 6 is generated. We are in a new scenario C, where we are going to show three different cases. In the case of C-3, only failure data less than 1000 h (42 failures) were considered. The case of C-2 employed failure data up to 1000 h, and censorship corresponding to the cases that failed after more than 1000 h. A total of 41 censored data appeared due to the 41 failures produced over 1000 h. This was intended to generate cases from case A. The data used in the C-3 case were those of the first part of Table 6 (darkened part), and the data used in case C-2 were all the data of Table 6.

Table 6. Failure hours of the O-rings in the case A, adapted to a policy of preventive replacement every 1000 h.

Time to Failure (Hours) and Censored Data (1000 h)					
190	276	296	409	429	430
437	454	481	492	498	499
543	552	552	552	552	577
603	604	604	612	619	658
675	683	696	702	742	754
773	797	812	836	881	889
912	913	942	974	994	994
1000	1000	1000	1000	1000	1000
1000	1000	1000	1000	1000	1000
1000	1000	1000	1000	1000	1000
1000	1000	1000	1000	1000	1000
1000	1000	1000	1000	1000	1000
1000	1000	1000	1000	1000	1000
1000	1000	1000	1000	1000	

Using these data for cases C-2 and C-3, the results shown in Table 7 were obtained. Case C-1 used data from Table 6 and, in addition, from Table 2.

Case C-3 involved an intense concentration of failures around 600 h of operation. The optimal preventive interval was considerably reduced, settling at 418 h. However, case A-2 showed more expected behaviour. There was also a concentration of failures but around 1000 h. The optimal preventive interval increased considerably to 705 h. If cases A-3, A-2, and A-1 are analysed as a whole, they showed a logical progression of the results.

Table 7. Results of the cases (cases A, C-1, C-2, C-3, and D-2) analysed. Parameters of the Weibull distribution function and the optimal preventive interval.

	Weibull			Optimal
	Shape Parameter α	**Scale Parameter β**	**Guaranteed Life γ**	**Preventive Interval τ_0**
Case A (uncensored)	2.36	1317	0	1059
Case C-1 (1000 + censored)	2.79	1149	0	822
Case C-2 (1000 + censored)	3.34	715	0	418
Case C-3 (1000 + censored)	2.76	1042	0	705
Case D-2 (900 + censored)	2.94	991	0	656

Case D-2 corresponded to a new scenario D, where the data from modified Table 1 were used censoring the data in the range of 400 to 2500 h of operation. Case D-2 corresponded to artificial censorship at 900 h, in the same way that case C-2 corresponded to artificial censorship at 1000 h.

The calculations were carried out for two transitions, starting from the operational state $v_1(2)$. These cases were compared to obtain the economic differences using Equation (5), which corresponded to the average accumulated return. The results are expressed in Table 8. The fifth column represents the return per hour of operation for each case. It would be the same in all even transitions.

Table 8. Economic comparison of cases A, C-1, C-2, C-3, and D-2.

Cases	**Average Return of Two Transitions $v_1(2)$ (€)**	**Average Transition (h)**	**Number of Transitions**	**Average Return (€/h)**
Case A (uncensored)	2174.36	899.5	2	1.209
Case C-1 (1000 + censored)	1815.24	745.6	2	1.217
Case C-2 (1000 + censored)	1388.05	647.0	2	1.073
Case C-3 (1000 + censored)	530.29	402.7	2	0.658
Case D-2 (900 + censored)	1308.91	610.4	2	1.072

4. Discussion

The results in Table 7 show that, when calculating the optimal preventive interval using the formula of our model, its value was always less than the predetermined preventive interval used (censorship). It is observed that, for a preventive interval adopted of 1000 h, case C-1, the derived optimal preventive interval reached the value of 705 h, while, when the preventive interval used was 900 h, case D-1, the optimal preventive interval reached the value of 656 h.

If with the same data used for cases C-1 and D-1, we modify the value of the predetermined preventive interval used between the range of values 500–2500 h, the behaviour of the optimal preventive interval shows a tendency to increase until reaching the value 899 when the preventive interval used was 2500 h, which was the case for which all fault data were used (see Figure 4). Why would the value 1059 not be reached? This value was not reached because, in these calculations, only the values in Table 2 were considered. This value would have been reached in the calculations if all censored data were considered

(Tables 2 and 6). The increase in the optimal preventive interval between the D-1 case of censorship to 900 h (656) and the case of 2500 h (899) was 27%. This value is comparable to the 22.4% value obtained when verifying the increase in the optimal interval between case A (1059) and case A-0 (822) using the data from Tables 2 and 6.

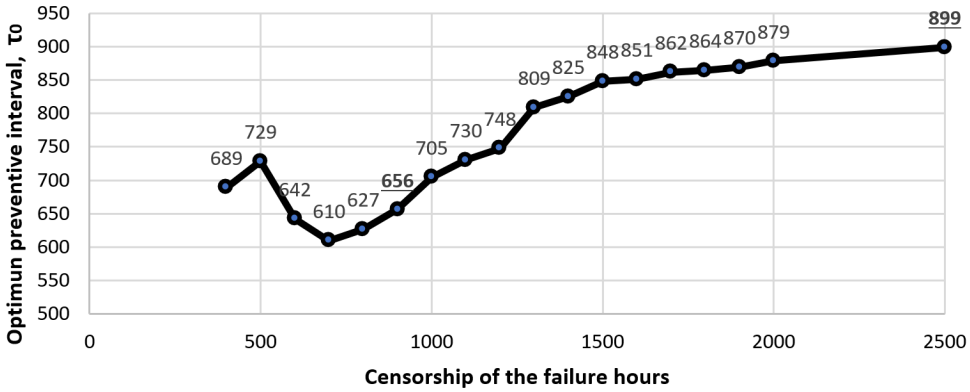


Figure 4. Optimal preventive interval values for case A failures when data are artificially censored.

In both cases and in other cases with different data, values close to 25% were obtained. This is because the preventive intervals used were 1000 and 900 h, values very close to the optimal 1059 and 899.

This value close to 25% is of particular interest, because it gives a reference value for the possible variation of the preventive interval when the starting data are censored. We have investigated and obtained a preventive interval of 1059 or 899 h, respectively. However, the maintenance manager does not know if their preventive interval used is close to its optimal value (because they do not let their equipment work until failure). If the value of the optimal interval that the maintenance manager obtains with the model is 30% less than the interval used, the maintenance manager would have a preventive maintenance policy with a high interval, which should be reduced. On the contrary, if the value obtained from the model is 20% less than the interval used, the preventive maintenance would have a low interval, which should be increased. This reduction or increase will be greater or lesser depending on how far the value of the optimal interval obtained from the model was from 25% lower than the interval used, Figure 5. This provides the maintenance manager with a forecast on the length of the optimal preventive interval and a decision rule to improve the interval.

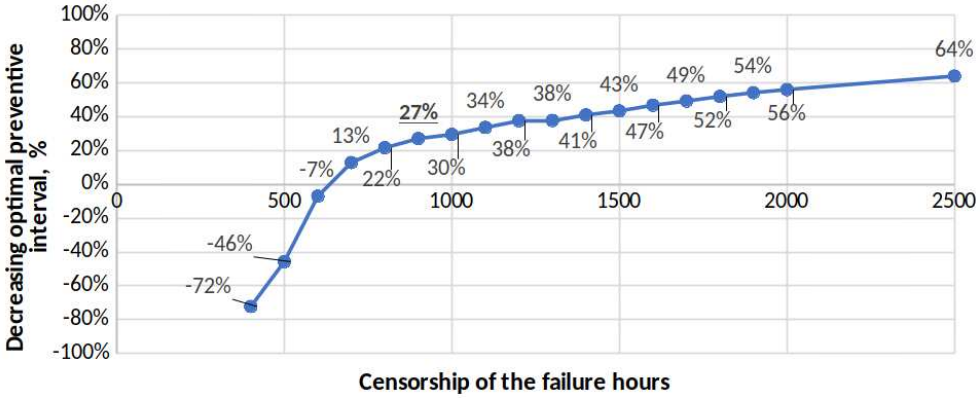


Figure 5. Values of the optimal preventive interval for the failures of case A when the data are artificially censored.

In Figure 4, an abrupt change in the direction of the curve is observed. It can be seen that for values greater than 700 h (point 700 on the abscissa), the curve experiences a change in trend that is maintained until the end. This indicates that the behaviour of the fault

distribution function should be studied when the number of failures is very small compared to a large number of censored data due to preventive maintenance. In this case, from values above 75% of the right-censored value (value of the adopted preventive interval), we can find values with which to obtain good to good results.

5. Conclusions

To carry out a study where the preventive interval is calculated, it is necessary to have adequate failure data, eliminating from the study those failure data due to other factors (for example, other failure modes). It is usually necessary to have uncensored failure data to optimise the preventive interval, but these are rarely accessible. Most of the data collected by maintenance managers are right-censored due to the performance of predetermined preventive maintenance tasks. The inclusion of censored data in the study introduces uncertainty in calculating the observed function, the theoretical function, the average accumulated return, and the preventive interval. These censored data disturb the calculation of the preventive interval, since they introduce uncertainty. The proposed model allows maintenance managers to use the censored data to guide the modification of the maintenance policy, by optimising the preventive interval, which will have an economic improvement.

The following conclusions can be drawn from using our model's formula to calculate the preventive interval for right-censored data:

- If the optimal preventive interval obtained using the model is more than 30% lower than the preventive interval used, the maintenance manager can decrease the interval used.
- If the optimal preventive interval obtained is less than 20% less than the preventive interval used, the maintenance manager can increase the interval used.
- This increase or decrease in the value of the optimal preventive interval will be more significant or less depending on how far the value obtained from the model is from 25% lower than the interval used.

This decision rule is of particular importance for the maintenance manager, since it allows them to modify the preventive interval in the correct direction. If censored data are used in the study, the optimal preventive interval increases as we also consider failure data of longer duration (adopt longer preventive intervals). These results are equally relevant regardless of the number of data considered when the number of failures is sufficient to define a failure distribution function according to the behaviour of the physical asset.

This research will continue in the future along two different lines, the first one by using other families of estimators and comparing the results, but the second one intends to extend the model by considering a fourth state for the system. The intention is to split the operational status between regular operational status and degraded operational status, where maintenance would be required by either legal or business constraints forcing consideration of its operation and whether in the new state the probabilistic function for failure becomes different. In this way, it would be possible to enable a stronger alignment between production and maintenance policies, where a more integrated perspective for the business as a whole is envisaged.

Author Contributions: Conceptualisation, A.S.-H.; methodology, A.S.-H., A.M.-N. and F.R.-M.; software, J.O.-M. and J.V.-D.; validation, A.S.-H., A.M.-N. and F.R.-M.; formal analysis, A.S.-H. and F.R.-M.; investigation, A.S.-H., A.M.-N. and F.R.-M.; resources, J.O.-M. and J.V.-D.; data curation, A.S.-H.; writing—original draft preparation, A.S.-H. and A.M.-N.; writing—review and editing, A.S.-H., A.M.-N. and F.R.-M.; visualisation, A.S.-H. and A.M.-N.; supervision, A.S.-H., A.M.-N. and J.O.-M.; project administration, J.V.-D.; funding acquisition, J.O.-M. All authors have read and agreed to the published version of the manuscript.

Funding: APC was funded by the Grant RTI2018-094614-B-I00 (SMASHING), and the "Programa Estatal de I+D+i Orientada a los Retos de la Sociedad" through the MCIN/AEI/ https://doi.org/10.13039/501100011033.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sample can be obtained from the authors upon request.

Acknowledgments: The authors want to thank the Spanish Agencia Estatal de Investigación for the support provided through the program "Retos para la Sociedad edition 2018".

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CDF	Cumulative Distribution Function
CMMS	Computer-Aided Maintenance Management Systems
ERPs	Enterprise Resource Planning systems
LSE	Least Square Estimation
ME	Maximised Expectation algorithm
MLE	Maximum Likelihood Estimation
MRR	Median Rank Regression
PDF	Probability Density Function
SMOTE	Synthetic Minority Oversampling Technique

References

1. Yang, L.; Zhao, Y.; Peng, R.; Ma, X. Hybrid preventive maintenance of competing failures under random environment. *Reliab. Eng. Syst. Saf.* **2018**, *174*, 130–140. [CrossRef]
2. Zhong, S.; Pantelous, A.A.; Goh, M.; Zhou, J. A reliability-and-cost-based fuzzy approach to optimize preventive maintenance scheduling for offshore wind farms. *Mech. Syst. Signal Process.* **2019**, *124*, 643–663. [CrossRef]
3. Damant, L.; Forsyth, A.; Farcas, R.; Voigtländer, M.; Singh, S.; Fan, I.S.; Shehab, E. Exploring the Transition from Preventive Maintenance to Predictive Maintenance Within ERP Systems by Utilising Digital Twins. In *Transdisciplinary Engineering for Resilience: Responding to System Disruptions*; IOS Press: Amsterdam, The Netherlands, 2021; pp. 171–180.
4. Zio, E. Reliability engineering: Old problems and new challenges. *Reliab. Eng. Syst. Saf.* **2009**, *94*, 125–141. [CrossRef]
5. Petritoli, E.; Leccese, F.; Ciani, L. Reliability and Maintenance Analysis of Unmanned Aerial Vehicles. *Sensors* **2018**, *18*, 3171. [CrossRef]
6. Sembiring, N.; Panjaitan, N.; Angelita, S. Design of preventive maintenance system using the reliability engineering and maintenance value stream mapping methods in PT. XYZ. In Proceedings of the IOP Conference Series: Materials Science and Engineering, TALENTA–Conference on Engineering, Science and Technology 2017 (TALENTA-CEST 2017), Sumatera Utara, Indonesia, 7–8 September 2017; Volume 309, p. 012128. [CrossRef]
7. Mizutani, S.; Zhao, X.; Nakagawa, T. WIB (Which-Is-Better) Problems in Maintenance Reliability Policies. In *Handbook of Advanced Performability Engineering*; Springer International Publishing: Cham, Switzerland, 2021; pp. 523–547. [CrossRef]
8. Chiodo, E.; Mazzanti, G. The Decreasing Hazard Rate Phenomenon: A Review of Different Models, with a Discussion of the Rationale behind Their Choice. *Electronics* **2021**, *10*, 2553. [CrossRef]
9. Parra, C.; Viveros, P.; Kristjanpoller, F.; Crespo, A.; Gonzalez-Prida, V. Audit and diagnosis in asset management and maintenance applied in the electrical industry. *DYNA* **2021**, *96*, 238. [CrossRef]
10. Sanchez-Robles, J.; de Leon Hijes, F.C.G.; Alarcon-Garcia, M.; Martinez-Garcia, F.M.; Belen-Rivera, E. Predictive maintenance for the efficient use of industrial agitators in distillers and reactors. *DYNA* **2021**, *96*, 17–21. [CrossRef]
11. Chiodo, E.; De Falco, P.; Di Noia, L.P. Challenges and New Trends in Power Electronic Devices Reliability. *Electronics* **2021**, *10*, 925. [CrossRef]
12. Birolini, A. *Reliability Engineering: Theory and Practice*; Springer Science & Business Media: Cham, Switzerland, 2013.
13. Melchor-Hernández, C.L.; Rivas-Dávalos, F.; Maximov, S.; Coria, V.; Guardado, J. A model for optimizing maintenance policy for power equipment. *Int. J. Electr. Power Energy Syst.* **2015**, *68*, 304–312. [CrossRef]
14. Legát, V.; Mošna, F.; Aleš, Z.; Jurča, V. Preventive maintenance models–Higher operational reliability. *Eksploat. Niezawodn./Maint. Reliab.* **2016**, *19*, 134–141. [CrossRef]
15. Kitagawa, T.; Yuge, T.; Yanagi, S. Maintenance modeling for a system equipped on ship. *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.* **2017**, *100*, 629–638. [CrossRef]
16. De Rivas, B.L.; Vivancos, J.L.; Ordieres-Meré, J.; Capuz-Rizo, S.F. Determination of the total acid number (TAN) of used mineral oils in aviation engines by FTIR using regression models. *Chemom. Intell. Lab. Syst.* **2017**, *160*, 32–39. [CrossRef]
17. Junliang, L.; Yueliang, C.; Yong Zhang, Z.Z.; Weijie, F. Availability modelling for periodically inspected systems under mixed maintenance policies. *J. Syst. Eng. Electron.* **2021**, *32*, 722–730. [CrossRef]

18. Martín, M.G.; Álvarez, A.P.; Ordieres-Meré, J.; Villalba-Díez, J.; Morales-Alonso, G. New Business Models from Prescriptive Maintenance Strategies Aligned with Sustainable Development Goals. *Sustainability* **2020**, *13*, 216. [CrossRef]
19. Ansari, F.; Glawar, R.; Nemeth, T. PriMa: A prescriptive maintenance model for cyber-physical production systems. *Int. J. Comput. Integr. Manuf.* **2019**, *32*, 482–503. [CrossRef]
20. Młynarski, S.; Pilch, R.; Smolnik, M.; Szybka, J.; Wiązania, G. A model of an adaptive strategy of preventive maintenance of complex technical objects. *Eksploat. Niezawodn./Maint. Reliab.* **2019**, *22*, 35–41. [CrossRef]
21. White, J.R.; Widup, R. Factors to Consider When Determining Maintenance Intervals. *IEEE Trans. Ind. Appl.* **2014**, *50*, 188–194. [CrossRef]
22. Coria, V.; Maximov, S.; Rivas-Dávalos, F.; Melchor, C.; Guardado, J. Analytical method for optimization of maintenance policy based on available system failure data. *Reliab. Eng. Syst. Saf.* **2015**, *135*, 55–63. [CrossRef]
23. Zhao, X.; Cai, J.; Mizutani, S.; Nakagawa, T. Preventive replacement policies with time of operations, mission durations, minimal repairs and maintenance triggering approaches. *J. Manuf. Syst.* **2020**, *61*, 819–829. [CrossRef] [PubMed]
24. Zhao, X.; Chen, M.; Nakagawa, T. Periodic replacement policies with shortage and excess costs. *Ann. Oper. Res.* **2020**. [CrossRef]
25. Bunea, C.; Bedford, T. The effect of model uncertainty on maintenance optimization. *IEEE Trans. Reliabil.* **2002**, *51*, 486–493. [CrossRef]
26. Mazzuchi, T.A.; van Dorp, J.R. A Bayesian expert judgement model to determine lifetime distributions for maintenance optimisation. *Struct. Infrastruct. Eng.* **2012**, *8*, 307–315. [CrossRef]
27. Zhu, J.; Wang, L.; Spiryagin, M. Control and decision strategy for a class of Markovian jump systems in failure prone manufacturing process. *IET Control. Theory Appl.* **2012**, *6*, 1803–1811. [CrossRef]
28. van Dorp, J.R.; Mazzuchi, T.A. Three-point lifetime distribution elicitation for maintenance optimization in a bayesian context. In *International Series in Operations Research & Management Science*; Springer International Publishing: Cham, Switzerland, 2021; pp. 147–177. [CrossRef]
29. Fouladirad, M.; Paroissin, C.; Grall, A. Sensitivity of optimal replacement policies to lifetime parameter estimates. *Eur. J. Oper. Res.* **2018**, *266*, 963–975. [CrossRef]
30. Wang, C.; Guo, J.; Shen, A. Sensitivity analysis of censoring data from component failure analysis and reliability evaluation for the aviation internet of things. *Comput. Commun.* **2020**, *157*, 28–37. [CrossRef]
31. Hu, Q.; Yue, W. Optimal replacement of a system according toa semi-Markov decision process in a semi-Markov environment. *Optim. Methods Softw.* **2003**, *18*, 181–196. [CrossRef]
32. Gualeni, P.; Perrera, F.; Raimondo, M.; Vairo, T. Accessibility for maintenance in the engine room: Development and application of a prediction tool for operational costs estimation. *Ship Technol. Res.* **2022**. [CrossRef]
33. Wu, B.; Cui, L. Reliability analysis of periodically inspected systems with competing risks under Markovian environments. *Comput. Ind. Eng.* **2021**, *158*, 107415. [CrossRef]
34. Tran, H.; Setunge, S.; Shi, L. Markov Chain–Based Inspection and Maintenance Model for Stormwater Pipes. *J. Water Resour. Plan. Manag.* **2021**, *147*, 04021077. [CrossRef]
35. Wu, B.; Maya, B.I.G.; Limnios, N. Using semi-Markov chains to solve semi-Markov processes. *Methodol. Comput. Appl. Probab.* **2021**, *23*, 1419–1431. [CrossRef]
36. Hu, J.; Shen, J.; Shen, L. Periodic preventive maintenance planning for systems working under a Markovian operating condition. *Comput. Ind. Eng.* **2020**, *142*, 106291. [CrossRef]
37. Nie, R.; He, S.; Liu, F.; Luan, X. Sliding mode controller design for conic-type nonlinear semi-Markovian jumping systems of time-delayed Chua's circuit. *IEEE Trans. Syst. Man Cybern. Syst.* **2019**, *51*, 2467–2475. [CrossRef]
38. De Jonge, B.; Scarf, P.A. A review on maintenance optimization. *Eur. J. Oper. Res.* **2020**, *285*, 805–824. [CrossRef]
39. Ruschel, E.; Santos, E.A.P.; Loures, E.d.F.R. Industrial maintenance decision-making: A systematic literature review. *J. Manuf. Syst.* **2017**, *45*, 180–194. [CrossRef]
40. Li, D.; Li, Q.; Mingcheng, E.; Jiang, Z.; Ma, J. Failure analysis of coupler knuckle considering truncated and censored lifetime data. In Proceedings of the 2019 Prognostics and System Health Management Conference (PHM-Qingdao), Qingdao, China, 25–27 October 2019; IEEE: Piscataway, NJ, USA , 2019; pp. 1–5.
41. Taghipour, S.; Banjevic, D. Trend analysis of the power law process using Expectation–Maximization algorithm for data censored by inspection intervals. *Reliab. Eng. Syst. Saf.* **2011**, *96*, 1340–1348. [CrossRef]
42. Sánchez-Herguedas, A.; Mena-Nieto, A.; Rodrigo-Muñoz, F. A new analytical method to optimise the preventive maintenance interval by using a semi-Markov process and z-transform with an application to marine diesel engines. *Reliab. Eng. Syst. Saf.* **2021**, *207*, 107394. [CrossRef]
43. Wen, Y.; Rahman, M.F.; Xu, H.; Tseng, T.L.B. Recent Advances and Trends of Predictive Maintenance from Data-driven Machine Prognostics Perspective. *Measurement* **2021**, *187*, 110276. [CrossRef]
44. Wu, B.; Cui, L.; Fang, C. Reliability analysis of semi-Markov systems with restriction on transition times. *Reliab. Eng. Syst. Saf.* **2019**, *190*, 106516. [CrossRef]
45. Yi, H.; Cui, L.; Shen, J.; Li, Y. Stochastic properties and reliability measures of discrete-time semi-Markovian systems. *Reliab. Eng. Syst. Saf.* **2018**, *176*, 162–173. [CrossRef]
46. Klein, J.P.; Moeschberger, M.L. *Survival Analysis: Techniques for Censored and Truncated Data*; Springer: Cham, Switzerland, 2003; Volume 1230.

47. Shen, A.; Guo, J.; Wang, Z.; Jia, W. A novel reliability evaluation method on censored data. *J. Mech. Sci. Technol.* **2017**, *31*, 1105–1117. [CrossRef]
48. Ahmed, E.A. Estimation of some lifetime parameters of generalized Gompertz distribution under progressively type-II censored data. *Appl. Math. Model.* **2015**, *39*, 5567–5578. [CrossRef]
49. Teimouri, M.; Hoseini, S.M.; Nadarajah, S. Comparison of estimation methods for the Weibull distribution. *Statistics* **2013**, *47*, 93–109. [CrossRef]
50. Genschel, U.; Meeker, W.Q. A Comparison of Maximum Likelihood and Median-Rank Regression for Weibull Estimation. *Qual. Eng.* **2010**, *22*, 236–255. [CrossRef]
51. Ranjan, R.; Sen, R.; Upadhyay, S.K. Bayes analysis of some important lifetime models using MCMC based approaches when the observations are left truncated and right censored. *Reliab. Eng. Syst. Saf.* **2021**, *214*, 107747. [CrossRef]
52. Balakrishnan, N.; Mitra, D. Left truncated and right censored Weibull data and likelihood inference with an illustration. *Comput. Stat. Data Anal.* **2012**, *56*, 4011–4025. [CrossRef]
53. Ferreira, L.A.; Silva, J.L. Parameter estimation for Weibull distribution with right censored data using EM algorithm. *Eksploat. Niezawodn./Maint. Reliab.* **2017**, *19*, 310–315. [CrossRef]
54. Joarder, A.; Krishna, H.; Kundu, D. Inferences on Weibull parameters with conventional type-I censoring. *Comput. Stat. Data Anal.* **2011**, *55*, 1–11. [CrossRef]
55. Starling, J.K.; Mastrangelo, C.; Choe, Y. Improving Weibull distribution estimation for generalized Type I censored data using modified SMOTE. *Reliab. Eng. Syst. Saf.* **2021**, *211*, 107505. [CrossRef]
56. Elmahdy, E.E. A new approach for Weibull modeling for reliability life data analysis. *Appl. Math. Comput.* **2015**, *250*, 708–720. [CrossRef]
57. Zhang, L.; Xie, M.; Tang, L. Bias correction for the least squares estimator of Weibull shape parameter with complete and censored data. *Reliab. Eng. Syst. Saf.* **2006**, *91*, 930–939. [CrossRef]
58. Musleh, R.M.; Helu, A. Estimation of the inverse Weibull distribution based on progressively censored data: Comparative study. *Reliab. Eng. Syst. Saf.* **2014**, *131*, 216–227. [CrossRef]
59. Jia, X.; Wang, D.; Jiang, P.; Guo, B. Inference on the reliability of Weibull distribution with multiply Type-I censored data. *Reliab. Eng. Syst. Saf.* **2016**, *150*, 171–181. [CrossRef]
60. Ducros, F.; Pamphile, P. Bayesian estimation of Weibull mixture in heavily censored data setting. *Reliab. Eng. Syst. Saf.* **2018**, *180*, 453–462. [CrossRef]
61. Jabeen, R.; Ahmad, A.; Feroze, N.; Gilani, G.M. Estimation of location and scale parameters of Weibull distribution using generalized order statistics under type II singly and doubly censored data. *Int. J. Adv. Sci. Technol.* **2013**, *55*, 67–80.
62. Sánchez Herguedas, A.; Crespo Márquez, A.; Rodrigo Muñoz, F. Optimizing preventive maintenance over a finite planning horizon in a semi-Markov framework. *IMA J. Manag. Math.* **2020**, *33*, 75–99. [CrossRef]

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland
Tel. +41 61 683 77 34
Fax +41 61 302 89 18
www.mdpi.com

Sensors Editorial Office
E-mail: sensors@mdpi.com
www.mdpi.com/journal/sensors

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland

Tel: +41 61 683 77 34
Fax: +41 61 302 89 18

www.mdpi.com

ISBN 978-3-0365-3811-2