ABSTRACT

LATIF, HASAN HABIBUL. A Methodology for Developing Digital Twins in the Manufacturing Plant. (Under the direction of Dr. Dr. Binil Starly).

Manufacturing is transforming into extremely complex systems because of more diversified and customized product lines. A lack of visibility from a manufacturing operations management perspective and real-time control of the process leads to delayed decision-making. There is a need for digital "visibility" tools to integrate different enterprise software systems and provide feedback on real-time manufacturing process changes. The visibility tool will enable better data utilization and decision making for the manufacturing industries.

Digital twins have become a key enabling technology to provide the visibility tool with real-time data across different platforms. It ensures seamless communication between different data sources in the manufacturing which can enable rapid analysis without altering actual setup. In the manufacturing industry, digital twins are still being deployed mostly at the machine level. However, digital twins can also be applied at the entire manufacturing plant level. This study fills the gap in the existing literature and proposes a plant level application by making multi-fold contributions. 1) This study demonstrates digital twin applications from the machine level to the manufacturing plant level. 2) It proposes a methodology to develop the digital twin. The methodology provides a systematic guideline to build digital twins for manufacturing plants. 3) Two cases are instantiated to validate the proposed methodology. One case study covers discrete and another covers continuous manufacturing.

The first case study integrates a dynamic simulator and an advanced process controller in a manufacturing system using the OPC-UA messaging protocol. The OPC-UA middleware communication protocol acts as a common interface between server and client. The case study is a proof of concept of the OPC-UA standard implementation to support interoperability for different domains following the methodology. The second case study focuses on building a
digital twin of a human involved manufacturing process following the proposed concept and methodology. It combines simulation with data from the physical world and uses reinforcement learning to improve decision making by creating a recommendation tool. In both cases, a digital twin is successfully implemented. The Digital twin implementation across these two case studies demonstrates better insights and efficiency from a manufacturing operation management perspective.
A Methodology for Developing Digital Twins in the Manufacturing Plant

by
Hasan Habibul Latif

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Industrial Engineering

Raleigh, North Carolina
2020

APPROVED BY:

_________________________  ____________________________
Dr. Guodong Shao          Dr. Yuan-Shin Lee
External Member

_________________________  ____________________________
Dr. Paul Cohen             Dr. Michael Kay

_________________________
Dr. Binil Starly
Chair of Advisory Committee
DEDICATION

My mother, Umme Habiba Sultana

and

My wife, Sriparna Ghosh

without these two rocks, couldn’t have done it.
Hasan Latif is a Sr. Manufacturing Engineer of Missile Systems at Raytheon. Raytheon Company (NYSE: RTN), with 2019 sales of $29 billion and 70,000 employees, is a technology and innovation leader specializing in defense, civil government, and cybersecurity solutions. He has been working for Raytheon for about three years and already contributed to great capacities. Being responsible for improving supplier performance with more effective cross-functional engagement, he is currently supporting 50+ suppliers across a portfolio of 20 programs. He has been awarded two achievement awards in 2019. He volunteers for Raytheon’s Southern Arizona MESA (Mathematics Engineering Science Achievement) Adopt-A-School initiative by teaching at a local high school. In total, Hasan has 5 years of working experience, having worked for National Institute of Standards and Technologies (NIST). NIST is a US national lab under the Department of Commerce.

Additionally, Hasan is serving US Army Reserve force as a Lieutenant. Last five years, he has actively engaged and led critical missions for the US army. He holds Master’s degree in Industrial Engineering from West Virginia University and Bachelor’s degree in Industrial Engineering from Bangladesh University of Engineering & Technology (BUET), Bangladesh, with 11 peer-reviewed publications and 120+ citations till date. He is an ASQ certified six sigma black belt holder.
ACKNOWLEDGEMENTS

I would like to thank my advisor Dr Binil Starly for his guidance and help. Dr. Guodong Shao has played a significant role in my research as well. Without their guidance and supervision, this will not be possible. Dr. Kevin Lynch, Dr. Yuan-Shin Lee, Dr. Paul Cohen, and Dr. Michael Kay have provided their valuable comments to improve the overall quality. I am grateful to all of my committee members to inspire me and continue this long journey.
# TABLE OF CONTENTS

LIST OF TABLES ................................................................. vii

LIST OF FIGURES ............................................................... viii

Chapter 1 Introduction ..................................................... 1
  1.1 Motivation ............................................................... 2
  1.2 Problem Statement ................................................... 4
  1.3 Research Aim & Objective ......................................... 5
  1.4 Proposed Work ....................................................... 6
  1.5 Contribution .......................................................... 6
    1.5.1 Contributions to the Manufacturing Domain ................. 6
    1.5.2 Novel Content of the Research ............................. 7
  1.6 Dissertation Outline ............................................... 8

Chapter 2 Literature Review ............................................ 10
  2.1 Introduction .......................................................... 10
  2.2 Current Methods and Research Gap .............................. 11
  2.3 Digital Twin Concept .............................................. 14
    2.3.1 Digital Twin Definition for Manufacturing plant ............ 15
    2.3.2 Digital Twin History ......................................... 15
  2.4 ISA-95 .................................................................. 17
  2.5 Conclusion ............................................................. 18

Chapter 3 Proposed Methodology ..................................... 19
  3.1 Introduction ........................................................... 19
  3.2 Concept of Digital Twins for the Manufacturing Plants .......... 19
  3.3 Methodology to Develop Digital Twins ......................... 23
  3.4 Proposed vs Other Methodologies ................................ 26
  3.5 Conclusion ............................................................. 27

Chapter 4 Digital Twin Implementation by Integrating Different Platforms ............. 28
  4.1 Introduction ........................................................... 29
  4.2 Background ............................................................. 29
    4.2.1 OPC-UA .......................................................... 30
    4.2.2 Tennessee Eastman Problem ................................. 31
  4.3 Case Study .............................................................. 31
    4.3.1 Simplified Tennessee Eastman Problem .................... 32
    4.3.2 Optimization Problem Formulation ......................... 33
    4.3.3 OPC-UA Client Development: Advanced Process Control .... 35
    4.3.4 OPC-UA Server Development: Simulation ................ 38
4.3.5 Communication Protocol: OPC-UA ........................................ 39
4.4 Lessons Learned ............................................................... 40
4.5 Discussion and Conclusion ................................................. 43

Chapter 5  Digital Twin Implementation Integrating different manufacturing processes Involving Human Interactions ........................................ 45
5.1 Introduction ................................................................. 46
5.2 Background ............................................................... 47
   5.2.1 Digital Manufacturing .............................................. 49
   5.2.2 Digital Twin in Manufacturing ................................... 50
5.3 Case Study ................................................................. 53
   5.3.1 Physical Elements ................................................... 55
   5.3.2 Digital Elements .................................................... 56
   5.3.3 Integration ............................................................ 75
5.4 Lessons Learned ............................................................ 77
5.5 Discussion and Conclusion ............................................... 78

Chapter 6  Conclusions & Future Work .................................... 80
6.1 Overall Summary & Research Contribution ......................... 80
6.2 Future Work .............................................................. 83
   6.2.1 Hybrid Approach .................................................... 84
   6.2.2 Product-Oriented Effort .......................................... 84
6.3 Conclusion ................................................................. 84

Bibliography ................................................................. 86
LIST OF TABLES

Table 2.1 Key Issues for Effective Resource Management with respect to Current Strategies .................................................. 14
Table 3.1 Comparison between Different Methodology in Manufacturing Domain .............................................................. 27
Table 4.1 Summary of variables and nominal operating conditions ....................................................................................... 36
Table 5.1 Structure of Preprocessed Data .......................................................................................................................... 60
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Concept of a Digital Twin for the Manufacturing Plant</td>
<td>20</td>
</tr>
<tr>
<td>3.2</td>
<td>Methodology of Developing a Digital Twin</td>
<td>24</td>
</tr>
<tr>
<td>4.1</td>
<td>Information Flow between modules of the case study</td>
<td>32</td>
</tr>
<tr>
<td>4.2</td>
<td>Process Schematic of the Simplified TE Problem</td>
<td>33</td>
</tr>
<tr>
<td>4.3</td>
<td>Transfer Function Graph of a Model Predictive Controller (MPC)</td>
<td>38</td>
</tr>
<tr>
<td>4.4</td>
<td>Test and Deployment of Node Address</td>
<td>40</td>
</tr>
<tr>
<td>4.5</td>
<td>Methodology Application for the Case Study</td>
<td>41</td>
</tr>
<tr>
<td>5.1</td>
<td>Data Flow of the Case Study</td>
<td>55</td>
</tr>
<tr>
<td>5.2</td>
<td>Partial Schematic Diagram Block 1/3</td>
<td>57</td>
</tr>
<tr>
<td>5.3</td>
<td>Partial Schematic Diagram Block 2/3</td>
<td>58</td>
</tr>
<tr>
<td>5.4</td>
<td>Partial Schematic Diagram Block 3/3</td>
<td>59</td>
</tr>
<tr>
<td>5.5</td>
<td>Adaptive Simulation Logic</td>
<td>63</td>
</tr>
<tr>
<td>5.6</td>
<td>Reinforcement Learning Principal</td>
<td>65</td>
</tr>
<tr>
<td>5.7</td>
<td>Reinforcement Learning Architecture for Recommendation List</td>
<td>65</td>
</tr>
<tr>
<td>5.8</td>
<td>Simulation Flow Part 1</td>
<td>67</td>
</tr>
<tr>
<td>5.9</td>
<td>Simulation Flow 2</td>
<td>68</td>
</tr>
<tr>
<td>5.10</td>
<td>Starting Screen of the Software</td>
<td>69</td>
</tr>
<tr>
<td>5.11</td>
<td>Simulation Moving Forward</td>
<td>69</td>
</tr>
<tr>
<td>5.12</td>
<td>Worst 5 Operations</td>
<td>70</td>
</tr>
<tr>
<td>5.13</td>
<td>Applying Recommendation</td>
<td>70</td>
</tr>
<tr>
<td>5.14</td>
<td>Default Recommendation List</td>
<td>71</td>
</tr>
<tr>
<td>5.15</td>
<td>Simulation Moving after Applying Recommendation</td>
<td>71</td>
</tr>
<tr>
<td>5.16</td>
<td>Re-ranked Recommendation List</td>
<td>73</td>
</tr>
<tr>
<td>5.17</td>
<td>Output Estimation</td>
<td>74</td>
</tr>
<tr>
<td>5.18</td>
<td>Error Comparison between Traditional and Proposed Approach</td>
<td>75</td>
</tr>
<tr>
<td>5.19</td>
<td>Overall Concept</td>
<td>77</td>
</tr>
<tr>
<td>5.20</td>
<td>Implementation Steps for the Case Study</td>
<td>78</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

A McKinsey Global Institute report states that U.S. manufacturing can maintain its competitiveness by applying an optimized, autonomous factory approach in a digitized and integrated value chain [Ram17]. Achieving such digitization and integration within a manufacturing factory is heavily dependent on factory data, communication techniques, and common semantics. A free communication flow among different software and hardware systems, in turn, typically relies on the proper application of proven methods [Lu16]. Exploitation and deployment of advanced technologies into production is one of the major challenges facing the manufacturing plants [Cwi17]. Therefore, the manufacturing plants are pushing for transformation in their approach. They are trying to adapt a range of planning and control decisions at all levels; and hence, smart manufacturing systems are introduced. According to Zheng et al. smart manufacturing systems are comprised of smart design, smart machining, smart monitoring, smart control, smart scheduling, and industrial applications [Zhe18]. The ’smart’ refers to the ability of performing the task automatically. Machine learning, artificial intelligence, recommendation-based system, deep learning, sensor-based automation, internet of things, big data analytic - anything can be integrated in any of the decision-making hierarchy level to make it productive and smart. The widespread applications of smart manufacturing marks the advent of the fourth state of
industry production - Industry 4.0. The fourth era, namely, Industry 4.0, referred to as the “Fourth Industrial Revolution,” also known as “smart manufacturing,” “industrial internet,” or “integrated industry,” is on-going, with the characteristics of cyber-physical system production based on heterogeneous data and knowledge integration. Industry 4.0 encompasses numerous technologies and associated paradigms [Lu16]. The Industry 4.0, enablers and features in digitalized manufacturing, could be described as the key elements that provide the typical quality for the digital connectivity and communication of the physical and digital elements. Industry 4.0 is dominating the intelligent products, 3D printing, and autonomous vehicles market [Iva19]. However, industry 4.0 does not have any systematic framework as well that could be clearly identified. It varies widely from industry to industry. This is in a work in progress stage. However, all the tools for smart manufacturing are shaping the path for industry 4.0. Digital Twin is one of the similar concepts that is helping the industry 4.0.

1.1 Motivation

The aerospace, defense, and space industry are experiencing immense growth in recent days. The importance of having a successful production plan is also becoming more significant. The manufacturing can get very complicated when it comes to a low volume high value product. For example, in the military defense industry, there is usually 500-15000 parts involved with different arrival times for each part. Add uncertainty and uncontrollable breakdowns, it becomes a frightening task for production/factory manager to optimize or strategize the floor. Without optimization, the factory performs poorly. Certain trends start to pick up in the manufacturing: illogical material flow, poor layout, inefficient automation, poor organization, skewed utilization of resources, and much more.

A successful manufacturing plant largely depends on an effective resource management plan. A resource management plan is a branch of production planning which includes a set
of combinatorial optimizations with seamless communication of data. Job shop scheduling problem, optimizing raw material arriving time, floor plan layout, material flow, inventory management, data integration, simulation - these are the common techniques to achieve an effective resource management plan for the manufacturing plant. Generally, it requires a lot of time from a production manager's standpoint to collect all the data, analyzing it, and act accordingly. Although plenty of research has been done on this specific topic, none of them have been embraced by the plant. First, these algorithm/mathematical tools are overly simplified, not taking into account practical constraints and the uncertainties involved on the floor. Then, the techniques are not easily replicable or customized. The solution is not easy to visualize, and comparative analysis between a current state and future state is not always easy to conduct. Additionally, there is no convenient what-if situation analyzing power of these tools. "Simulation" method can be an answer to these problems and it is being used in the manufacturing floor a lot. However, the plants accepts simulation method as a quick and dirty way to problem solving but are rarely used to make important decisions with. Simulation methods are also very simple and traditional. Most of the cases, it does not provide what-if analysis, doesn’t take into account uncertainty, and does not represent complex scenarios. It also does not focus on the real issues like integration of different platforms.

Therefore, the plant has an intent not only to gather data to run the simulation, but also using the data to drive specific business outcomes. The simulation needs to represent more realistic outcome where stakeholders can visualize the outcomes. Digital twins provide much more than visibility into how remotely-located equipment is functioning. The data collected from digital twins, and the value that digital twins provide, can and should be used to evolve how the business functions. At the most basic level, digital twins give organizations visibility and insight into the operation of their machines in the context of the environment or IoT ecosystem they are a part of. The data gathered from sensors empowers businesses to have the parts or staff on hand before problems become problems. Companies can use the data collected to be
predictive, adapting business processes to be smarter in the future. In effect, digital twins are a management model that can be used to improve enterprise decisions.

1.2 Problem Statement

When a manufacturing plant is operating poorly, the entire enterprise suffers. This is why teams who are setting priorities for digital initiatives should place a high priority on improving manufacturing processes first and foremost. Enterprise interoperability has become a thriving key factor for the collaboration of a manufacturing plant. It determines to what extent the companies can utilize their resource’s unique capability and produce added value through synergistic effects. In a nutshell, manufacturing requires visualization and seamless communication to increase factory efficiency, robustness, and profitability. A digital twin can become the visibility tool with the integration capability. It enables production systems to become more agile and interoperable.

However, manufacturing industries are not embracing digital twins at a rapid rate. Deloitte performed a survey on the manufacturing industries and found that the main challenges industries face in creating digital twins include 1) a lack of clear standards to implement digital twins, 2) train people to use digital twins, 3) difficult to integrate available data, and 4) an organized method to implement digital twins [Del17]. A digital twin development can be challenging for the manufacturing processes with a complex layout. A systematic approach is needed for the developers to develop digital twins for the manufacturing plants. Moreover, a lack of real-world examples with complex manufacturing layout hurts the usage of digital twins. These challenges are immensely important to address because enabling digital twins will increase growth in manufacturing plants.
1.3 Research Aim & Objective

This research aims to develop a methodology to implement digital twins for the existing manufacturing plants. Any plant manager or significant stakeholder can use this research work as a guideline to transform an existing manufacturing plant into a digital twin by solving various issues specific to the respective problem scenarios. This research addresses the problems of integrating digital twins into the established manufacturing execution and enterprise resource management systems. The core focus areas of this research are as follows:

- Develop a methodology that can be used as a guideline to execute the digital twin for the manufacturing plants.

Guided by an established literature review methodology and grounded on an in-depth analysis, the key obstacles are identified in the digital twin domain. A proposed step-by-step guideline is provided as a methodology to implement digital twin in manufacturing plants. The proposed methodology encompasses all the critical elements so that it can overcome the identified obstacles.

- Validate the proposed digital twin methodology through real-world manufacturing case studies and demonstrate the impact of the digital twin applications in manufacturing plants.

To validate the proposed methodology, two real-life problems have been carefully chosen. The first one is a well known Tennessee Eastman problem and the second one is a complex aerospace product with real data (anonymized) from an actual manufacturing plant. These cases represent all the identified key obstacles and provide ideal platforms to implement digital twins. The case studies pose actual challenges and difficulties faced by manufacturing plants. These two problems demonstrate a systematic approach to the
stakeholders. Besides, the overall improvement is shown to reinstate the benefits of digital
twin applications.

1.4 Proposed Work

There is a gap in the current manufacturing industry for the digital twin implementation. The
stakeholders in the manufacturing industry are not yet confident about digital twins, while other
industries are successfully embracing digital twins. The integration, inter-operability issues, real-
time simulation - all pose significant obstacles to develop digital twins for manufacturing plants.
This research provides a methodology that is adaptive and easy to execute. This methodology
can be used as a beginner’s guide to the manufacturing plant managers. Two case studies are
provided to recognize and tackle the aforementioned critical obstacles in a structured way. A
structured approach towards digital twin development will surely bring more positive outcomes.
However, manufacturing industries are not fully aware of the structure or implementation steps.
Therefore, this study is intended to help prepare a development strategy for a digital twin of the
existing products in manufacturing plants.

1.5 Contribution

This section will outline the expected outcomes and potential impacts of the research. The
outcomes are described in the context of contributions to the manufacturing domain. The
impacts are described in the context of the novel content of the research.

1.5.1 Contributions to the Manufacturing Domain

From the problem statement, it is evident that a systematic approach or guideline is required for
developing digital twins in the manufacturing plants. The main contribution the research is a
verified and validated methodology for the digital twin in the manufacturing plant level. The problem statement has been addressed by contributing the following ways.

1) The literature supports the need of a formalized digital twin methodology for a manufacturing plant level. Existing research focuses heavily on the machine level (e.g., digital twin of a grinding wheel, digital twin of a CNC machine). However, this research demonstrates digital twin applications from the machine level to the manufacturing plant level.

2) Much of the literature is focused on the digital twin – the positive contributions of surrogate models in the cyber-space to replicate the physical product. However, the digital twin methodology, which enables the digital twin in manufacturing context, has received minimal attention in the literature. Industry and solution providers are searching for a common technology approach to enable the digital twin. Therefore, a digital twin development methodology has been provided to address the issue.

3) Real-world problems bolster the case for digital twins. In this research, two case studies are performed to validate the methodology. These two real industrial problems will fill the research gap to some extent.

1.5.2 Novel Content of the Research

Aside from the digital twin methodology, several novel concepts are included in this research proposal. While many of the technologies (e.g., OPC UA, Advanced process control, Adaptive simulation, Decision support, reinforcement learning) discussed in this proposal or required to complete the research are not novel, the integration of those technologies is unique. There is no evidence in the literature to support that prior work tried or was successful in completely integrating technologies across the manufacturing plant. Industry has long sought to link together data from different sources in the manufacturing plant. However, incompatible information models and systems remain a barrier to achieving industry’s need. The digital twin methodology
and case study from Chapter 4 provide integration example that could support closing the gap and enable plant to curate, discover, and retrieve data more effectively and efficiently across the manufacturing plant. The case study from Chapter 5 provides an example of integrating existing manufacturing processes in a manufacturing plant. It also enable a recommendation based system which can proactively identify potential bottlenecks to improve efficient usage of resources. A python based software is also created to build recommendation based system. The software is stored at following github repository.

https://github.com/bpi06/digital_twin/blob/master/Mfg%20Process

1.6 Dissertation Outline

The study is elaborated into six sections. The first chapter introduces the problem along with existing literature and defines the digital twin of a manufacturing plant. The second chapter analyzes the existing digital twin concept and research gap. Chapter three proposes a step-by-step development methodology for digital twins in the manufacturing plants. This strategy should work as a primary guideline and replicate in various types of industrial applications. The fourth chapter provides a case study with real-time data integration with the efficient design of information flow. The fifth chapter provides another case study integrating different manufacturing operations within the existing system. Finally, the last chapter discusses the findings and future work scope.

The proposed methodology is validated using two real-world problems. These case studies are chosen to address complex and pragmatic issues. However, every product is different in its way. The manufacturing process, technical specification, quality - everything is highly customized. More case studies are required to perform. Open access storage where all the digital twin implementation case studies can be stored. Later, people can refer to that storage and improve their understanding of the digital twin. Henceforth, industry 4.0 can take over
the manufacturing industries and the world can achieve another revolution in the history of mankind.
Chapter 2

Literature Review

2.1 Introduction

A myriad of research have been performed to improve the manufacturing industry. However, the manufacturing plants are still in the rudimentary level. Outdated shop floor machinery and software solutions simply cannot provide the agility and visibility to meet today’s pressing demands. The manufacturing plant is where profits are made or lost, where customers are pleased or disgruntled, where streamlined activities with precision or where they fall apart into chaotic spurts of rush jobs, unexpected delays, and stock-outs. This is why teams who are setting priorities for digital initiatives should place a high priority on improving manufacturing plants first and foremost. No functional area can be seen as an isolated silo, though. The overall strategic priorities for the entire company must be kept in perspective. That is where digital twin can play an important role. Digital twin eases to identify potential bottlenecks, helps improving the overall efficiency and functionality of the job as well as the enterprise. Within the profound digitalization in manufacturing industries, the concept of a virtual, digital equivalent to a physical product has gained increasingly attention.

In this chapter, the current available literature is studied from manufacturing plant standpoint.
Current technologies in the industry, the challenges of previous technologies, future directions - all are studied. This chapter helps to identify the research gap and how digital twin fits to close the gap.

### 2.2 Current Methods and Research Gap

Even a couple decades ago, manufacturing plant was focused on one specific task or one machine at a time. In contrast, today’s plant deals an increasing range of tasks, demands and complexities. The development of the manufacturing is, of course, a mirror of the changing conditions faced by the production plants in general. In the last century, the manufacturing plants were in need of firm supervision and a rigid structure, in order to reach high productivity [Wit16]. In 1948, the concept Method Time Management (MTM) emerged and during the following years it was introduced to industry. In the 1980’s, the Toyota Production System (TPS), a production and management concept, began impacting the manufacturing industry. Along with another management philosophy, Total Quality Management(TQM), the focus was set on the expectations and requests of the customer, as well as the production variables, quality and flexibility. The production philosophies emerging from TPS, TQM and Lean have strongly influenced most of the world’s production companies on both management and floor levels, and still do so today [Lin13]. In this study, they are categorized as traditional methods. Even though these methods are effective and simple, they do not provide in-detail guidelines. These approaches are quite simplistic and holistic in nature. The methods can support overall policy making but it does not facilitate visualization or machine level improvisation.

Having a pro-active approach is an important strategy that facilitates greater flexibility and a possible reduction of the total lead time at plant’s assembly stations. Such approaches can counteract possible uncertainty caused by stochastic events (errors, breakdowns, etc.) during production [Hol18]. In a recent study, stochastic block model (SBM) concept is introduced in
the manufacturing plant. The SBM is a generalization of the block model settled in the social networks theory decades ago with a stochastic element. It is capable of handling a wide range of structures like core-periphery, hierarchies and many more. These capabilities are important for situations, where the type of resulting structure is not known before like in manufacturing plant [Fun18].

There are other algorithm based approach where it mostly includes optimization-based problem formulation. Most of the literature in early decades used aggregated single objective algorithms. These algorithms contain a set of constraints which are created from factory resources. Advanced algorithm includes all the latest hybrid methods and heuristic such as Gradient Based Search Methods, Stochastic Optimization, Response Surface Methodology, genetic algorithms, simulated annealing, evolutionary strategies, etc. Furthermore, in the past decade, many research projects have been implemented using a new algorithm called Pareto based multi objective evolutionary algorithms (MOEAs) [Deb01]. Genetic algorithm (GA) proved to be one of the best evolutionary algorithms over random search and particle swarm optimization [Chi13][Gao07]. Then, non-dominated sorting genetic algorithm (NSGA) provides better techniques and results in solving job shop scheduling problem [Sri94]. Even NSGA-II version came out to solve large multi-objective problems [Deb01]. Another modified version of NSGA-II was proposed integrating random search, which was called non-dominated ranking genetic algorithm [AJ08]. In another paper, uncertainty was integrated in the manufacturing plant and tested among the discussed genetic algorithm techniques [Vel00]. Apart from genetic algorithm, researchers tried several other techniques as well. Zhang et. al proposed a hybrid algorithm which combined particle swarm optimization and neighborhood function [Zha09]. Xiong worked on a robust scheduling with machine breakdowns. He used surrogated model for approaching the multi objective problem [Xio13]. Yuan and Zu created memetic algorithm which incorporates local search algorithm into NSGA-II [Yua15]. He and Sun proposed an approach in which right shift scheduling and route changing scheduling are used [He13]. Overall,
it is obvious that algorithms has not been considered adaptability extensively to cover a broad area of manufacturing industry. Over the time, researchers have proposed and created new algorithm to address different scenario. However, all these algorithms are not proven in the pragmatic case studies. Most of the case studies are overly simplified. It is difficult to customize and adapt to one manufacturing plant to other. There is robustness, but it is very specific to the field / manufacturing industry. Visualization is missing and too much computational effort is required to perform these algorithms. Therefore, all these algorithm-based techniques failed to make an impression in the real plant where the situation changes, uncertainty looms, and intricacy increases often.

There is another approach prevails in the manufacturing plant, which is simulation. The Simulation Based Design (SBD) concept refers to the use of simulations during the whole product life cycle. The objective is to validate the behaviour of the product during its life cycle. It can be implemented by the integration of new digital technologies and new methodologies. Other work aims to facilitate the integration of design and manufacturing modeling at the concept design stage, including cost [Cur07]. In particular, this methodology allows easiest exploitation of digital manufacturing simulation capabilities. These various approaches emphasise the simulation of product to be developed by taking into account some simple constraints of the product life cycle. This methodology does not integrate the product development phase (analysis of task, conceptual and embodiment design etc.) in the simulation one.

Simulation can be a simple solution to the current manufacturing plant problems. But simulation method is overly simple. It only considers the ideal situation where it generates the result without much consideration. It always assumes an ideal scenario and skips to the finish line without adapting inter-process uncertainties. It does not have the critical portion where every decision requires a lot of effort and change. Hence, a simulation with machine learning capabilities aka predictive analyzing capabilities can be a game changer. That is where, the new concept digital twin can play a big role.
Digital twin can bridge the gap between current industry needs and available technologies. Digital twin has a broad scope with a lot of integration opportunities. It is repeatable, adaptive, and convenient. Further analysis and newer technologies are easier to incorporate. Industries have been using it in a small scale without knowing much of it. Table 2.1 has provided the key issues and difference between various domains in manufacturing plant. The ‘-’ sign indicates ‘not covered’ and the ‘+’ indicates ‘covered’.

<table>
<thead>
<tr>
<th>Issues for the effective resource management plan</th>
<th>Traditional Methods</th>
<th>Advanced Algorithms</th>
<th>Simulation</th>
<th>Digital Twin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Practical Problem</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Replicable</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Adaptive</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Easy Visualization</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>What if Analysis</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Uncertainty Integration</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Required Time to perform</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Fast Feedback</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Industry Acceptance</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>N/A</td>
</tr>
</tbody>
</table>

2.3 Digital Twin Concept

The advancement of new generation technologies, such as internet of things (IOT), big data, cloud computing, artificial intelligence (AI), and their wide applications are driving manufacturing plant into an exciting era. New concepts are coming along to improve the manufacturing issues. The performance can be enhanced by enabling all the smart manufacturing or industry 4.0 techniques that have been discussed already. Digital twin is one such concept which can improve product performance, reliability, and maintenance. A digital twin can be defined as
a digital representation of the historical and present outline of a physical entity or process to optimize performance. Digital twin is the virtual representation of physical process. It is a high-fidelity virtual model for a physical system in digital way to simulate their behaviors. It is getting more and more attention. Gartner classified digital twin as a one of the top ten strategic technology trends for 2017 and 2018 [Pan16]. For a manufacturing plant, the digital representation aka the simulation of the manufacturing plant includes all the material path, machines, layout, part arrival rate, machine breakdown rate, etc.

2.3.1 Digital Twin Definition for Manufacturing plant

Industry and academia have defined digital twin in various ways. However, neither of the groups places emphasis on process aspects of digital twin. Specially in the manufacturing plants, the digital twin concept is still in an exploring zone. A digital twin in manufacturing plant is an evolving profile of the process which provides insights based on real-time, real-world data from different business/process levels. For instance, a simulation of the manufacturing plant with real time data can provide the virtual representation for digital twin. AI based recommendation, integration of different software platforms to exchange data, and optimization of the process will make it a true digital twin of a manufacturing plant.

2.3.2 Digital Twin History

The concept of the Digital Twin dates to a University of Michigan presentation to plant in 2002 for the formation of a Product Lifecycle Management (PLM) center. The University of Michigan used this conceptual model back in 2002 in their first executive PLM courses. It was referred as Mirrored Space Model and referenced in a journal article. In the seminal PLM book, Product Life cycle Management: Driving the Next Generation of Lean Thinking, the conceptual model was referred to as the Information Mirroring Model [Gri05]. The concept was greatly
expanded in Virtually Perfect: Driving Innovative and Lean Products through Product Lifecycle Management [Gri11]. The term, Digital Twin is used to describe the model. By this time, digital twin gets enough attention by the industry and academic domain.

Boschert and Rosen regarded digital twin concept as an amendment of existing approaches in system modeling and simulation by real product usage data and information [Bos16]. In the early stages of digital twin, it embraced the whole product life cycle "from cradle to cradle". Thereby, it represented a holistic model-based description of a product for current and future life cycle stages. Later, digital twin concept pursues to optimize product usage and service and maintenance activities for manufacturing plant. During the product service and maintenance, condition monitoring and operations/service support considering all available historical data and models (e.g., as-manufactured and as-maintained state) lies in the center of interest.

Digital twin has been used in different industries on a limited scope. Nikolakis et al. proposed a study that proposes a digital twin approach as part of a wider cyber-physical system (CPS) to enable the optimization of the planning and commissioning of human-based production processes using simulation-based approaches [Nik19]. Schroeder et al. proposed the use of AutomationML to model attributes related to DT and proposed that this model was very useful for data exchange between different systems connected with DT [Sch16]. Moreno et al. presented the process to construct a DT for a sheet metal punching machine to support the interactive design of optimal NC machining programs [Mor17]. Luo et al. presented a modelling method where he used strategy of DT for CNC machine. He provided a demonstration of DT concept in CNC machine tool era [Luo19]. On another paper, Botkina et al. demonstrated a study on a digital twin of a cutting tool. DT is represented as a digital replica of a physical tool, its data format and structure, information flows and data management, as well as possibilities for further applications and analysis of productivity [Bot18]. Liu et al. presented a detailed implementation process of the proposed DT method for the key parts of the marine diesel engine. They also mentioned the research gap in fulfilling digital twin-based smart process planning for complex
products [Liu19]. DT has contributed in other industries as well. For instance Tao et al. put forward the concept of Digital Twin Workshop and Digital Twin driven prognostics and health management for complex equipment [Tao17]. Many leading companies such as PTC, ANSYS, GE, and Siemens also explored more applications under the guidance of DT concept. In recent years, digital twin focuses on the general and specific applications which do no target a specific life cycle stage. Rather, digital twin focuses on integration, virtual representation at instance, and replication of physical objects. In sum, digital twin concepts are likely to go through more revolution and become more relevant over the next decade. For this development, the paper at hand offers initial insights and acts as guidance for prospective work.

2.4 ISA-95

ISA-95 is an international standard from the International Society of Automation for developing an automated interface between enterprise and control systems. This standard is developed for global manufacturers. It is applied in all industries, and in all sorts of processes, like batch processes, continuous and repetitive processes. The objectives of ISA-95 are to provide consistent terminology across supplier and manufacturer communications, information models, and operations models which is a foundation for clarifying application functionality and information usability. ISA-95 incorporates the layers model of technology and business process for manufacturing enterprises as levels for the standard. These levels are:

- Level 0: Defines the actual physical processes.
- Level 1: Defines the activities involved in sensing and manipulating the physical processes.
- Level 2: Defines the activities of monitoring and controlling the physical processes.
- Level 3: Defines the activities of workflow to produce the desired end products.
• Level 4: Defines the business-related activities needed to manage a manufacturing operation.

Chapter three will discuss the proposed methodology. In chapter four and chapter five, two case studies will be implemented from two different perspective to validate the proposed methodology. First case study integrates different platforms between level 2 and level 3 of ISA-95 and picked from continuous manufacturing. Other case study integrates different manufacturing operations of a discrete manufacturing and operates at level 3 of ISA-95. ISA-95 is widely accepted among industries. Therefore, manufacturing plants can relate to the case studies and apply at their levels accordingly.

2.5 Conclusion

The digital twin creates concrete value, generate revenue streams, and help them taking the key strategic decisions. With the aid of newer technologies, manufacturing industries can create a digital twin of their production plant in a short amount of time. Digital twin has many applications across its life cycle. It has been easier to answer the critical what-if questions in real-time more accurately. To keep pace with the progressing manufacturing world, it is highly recommended to leverage the usage of digital twin for manufacturing plant. However, existing literature lacks concrete examples for manufacturing industries from business perspective.
Chapter 3

Proposed Methodology

3.1 Introduction

From the literature review, it is evident that digital twins have the ability to address the current manufacturing problem. However, there are a couple of roadblocks as well to establish a digital twin in the manufacturing plant. The major challenges of digital twin are 1) how to effectively convert data into useful information across different enterprise control systems including the manufacturing plant. 2) how to effectively integrate a variety of data about the process, machine, operators, and environment from diversified sources using different methods 3) how to rapidly respond to a current event based on real-time data to make a better decision?

In this chapter, a methodology is proposed to answer all three questions of implementing digital twins in the manufacturing plant.

3.2 Concept of Digital Twins for the Manufacturing Plants

Creating a digital twin for the manufacturing plant starts with establishing new pipelines of manufacturing data sources. The system can automate the data collection, for example, of
materials and design data. When integrated with historical operations performance data, the system requires real-time data to support the creation of a digital twin. According to the literature, a digital twin must have the following: a physical element, a digital element, and integration across the two. In Figure 3.1, the concept of digital twins of a manufacturing plant is illustrated.

![Figure 3.1 Concept of a Digital Twin for the Manufacturing Plant](image)

From the manufacturing plant, data is being collected and transferred to the digital twin. The digital twin can be a simulation model, an algorithm, a mathematical model, etc. A historical dataset can be provided to the digital twin model as the initial dataset to train the model. Based on technologies such as machine learning, deep learning, fuzzy algorithms, and optimization, the digital twin generates recommendations or feedback. The feedback includes actionable decision guidance for the user. This decision aid should be intuitive and help tweak or adjust the manufacturing plant operational parameters. This process repeats throughout the manufacturing process until the best-case scenario or production target is achieved. The authors have identified 3 important elements to implement digital twins for a manufacturing plant.
• Physical Element: In the manufacturing world, physical element refers to the manufacturing processes in the manufacturing plant. First, a digital twin contains data collected from and about its physical counterpart, which is a manufacturing plant in this case. The data includes various parameters that are crucial for modeling. It includes product design specifications, process and engineering data, as-manufactured data (production equipment, material, method, quality data, and operators), and as-maintained data (real-time and historical configuration and operation state data, and maintenance records) of the real-world counterpart. The data may include the time required to complete an operation or number of workers in the system. Secondly, a digital twin may contain a variety of computational or analytic models pertaining to its real-world counterpart, ranging from first-principle-oriented (natural laws), data-oriented (statistical, machine learning/artificial intelligence) and geometrical or visualization-oriented (3D modeling and augmented reality). Lastly, a digital twin provides service interfaces for software applications to access its data and invoke its models. The connection between a digital twin and its real-world counterpart is dynamic, possibly real-time, and bi-directional. Therefore, With a digital twin, it enables decision makers to simulate and predict the state and behavior of its real-world counterpart based on analytics on historical and real-time processed data. It optimally responds to changing conditions of the real-world counterpart. Creating the digital twin of a manufacturing plant will require a good understanding of the processes.

• Digital Element: The digital element of a manufacturing plant comprises the digitalized manufacturing processes with an integration capability. From the process map of the manufacturing plant to the digital representation of the stations, all of them together form part of the digital twin. It can be modeling where constraints and objective functions are added, or a simulation where simulation logic and reasoning mechanisms are designed
to capture the physical model. The digital element of a digital twin virtually represents
the physical elements of the factory, where the quantitative and visual analysis factory
live-action, process parameters, product information, production line layout, production
equipment, logistics equipment, work-pieces, material, tooling/fixture, and equipment
operation, process parameters can be utilized. The difference between simulation and the
actual physics experiment is that the simulation is not based on the actual environment,
but the model in the virtual environment mapped from the actual environment. The corre-
sponding design process simulation, operations, logistics characterization methods are
built, such as digital logic configuration, event set, constraint conditions, and optimization
strategy. The different settings of process parameters and process among simulation can
be more real and fully describe the operation and evolution of the physical element. The
‘possible’ scenarios can be completed in simulation, which can improve and optimize the
physical space according to the optimal results in virtual space, so as to more effectively
improve the quality and efficiency, reduce cost and production cycle.

• Integration: The connection between a digital twin and its real-world counterpart should
be dynamic, possibly real-time and bi-directional. The communication between the
digital and physical elements allows potential problems to be identified early. That is the
benefit and goal of digital twins. Data collected from and about its physical counterpart
includes various parameters that are crucial for modeling the manufacturing plant. It may
include product design specifications, process and engineering data, as-manufactured data
(production equipment, material, method, quality, and operators), and as-maintained data
(real-time and historical configuration and operation states, and maintenance records).
The data may also include the time and resources required to complete an operation.

Integration is the most important part of the digital twin. A digital twin’s efficacy entirely
depends on it. A feedback loop (even better if automated) based on real-time data and
adjusting capability can strengthen the digital twin immensely. Integration is the key to the feedback loop. A database or repository is important for the digital twin of a manufacturing plant. Structured Query Language (SQL) database or No-SQL database both can be used as long as the parameters of interest are defined. Any suitable programming language can be used to extract the processed data from the database automatically if available. The data collection is manual because it is a manual assembly, statistical tools can be deployed to nullify the variances and outliers.

One important consideration is the whole concept should be adjustable and scalable for the newer and better technologies. This can allow the proposed methodology to be adaptive.

### 3.3 Methodology to Develop Digital Twins

Digital twin technology is still shaping the manufacturing industry. A proper guideline or methodology can help the effort. The Digital Twin methodology ensures the usage of factory data and produces a digital representation that can be deployed. A user needs to collect data about the phenomenon or event or physical entity. If the user has operational data like sensor data and event data like maintenance records, the analytic can correlate these to build a model that can predict them in the future.

This chapter intends to set the tone by proposing a methodology. Users can apply this methodology to implement digital twins in their manufacturing plant. Figure 3.2 shows the proposed methodology that is generic and that which can be applied to a variety of manufacturing plants.

From a digital twin methodology standpoint, users provide input to formulate the digital twin problem. The digital twin problem includes identifying the objective, constraints, and key factors. Here, the digital twin starts acquiring knowledge. A Digital Twin included both static and dynamic knowledge. Static knowledge includes geometrical dimensions, bill of materials,
processes, etc. The dynamic information is the one that changes with time along the product life cycle. Once the knowledge is gathered, a conceptual model is designed. The conceptual model virtually represents the physical processes in the manufacturing plant. It will enable the exchange of information of the problem attributes in the virtual world. A parallel activity takes place after problem formulation, which is specifying the data requirement. The data requirement decides which data should enter into the digital twin. Data requirement step takes physical properties with sensors that measure critical inputs. It is decided based on the operation standpoint and external factors. The physical plant is equipped with a capability, by means of various actuators, to adjust its function, behavior, and structure in the physical world. Sensors and actuators are the two technological backbones of a digital twin. The former plays a role in sensing the external world, whereas the latter plays the role in executing the desirable adjustments. In practice, the
commonly used actuators that are suitable for consumer products include, for example, hydraulic, pneumatic, electric, and mechanical actuators. The specification of data requirements can be used as a foundation for digital twin. After the data requirement specification, the communication protocol is determined. Data should be communicated by exchanging different functions using integration protocols and data repository. Pub/Sub, OPC-UA, TCP/IP, ERP systems are few examples through which data transfers in an efficient way [Beň19]. The connections are enabled using a number of technologies, such as network communication, cloud computing, and network security. The feasible networking technologies for consumer products include, for example, Bluetooth, QR code, barcode, Wi-Fi, Z-Wave, etc. Since product data are directly and indirectly concerning user-product interactions, it is critical to guarantee the security of connections. In light of the Internet of Things, much effort has been devoted to connecting the physical and virtual products, which can be adapted for digital twin research. The data aggregation and processing can be done on the local hardware or in the cloud. The technology is progressing rapidly, and scalable architectures are coming out to support as well. Whenever a data transfer and digitalization process of a plant exists, security protocols also becomes an issue. The security aspects should be considered and planned carefully. Now, data gets collected and processed. The heterogeneous data coming from different sources need to be accumulated in a database. Historical data is already stored in a database. Real-time data from manufacturing plant also comes into the database and processes altogether. Available data, data requirement, and strategy - coherently act together to organize the data in a useful format in the database. The data quality is also a point of concern. Product-data quality (PDQ) must be a crucial focus to ensure successful digital twin implementation. Manufacturing plant data consists of product-related specifications, manufacturing plant parameters, machining sequences, etc., and all are typically stored using a database. There are two uses of product data: (1) lateral direction and (2) vertical direction. Lateral direction means using product data within a phase of the product life cycle. Vertical direction means reusing product data in subsequent product life cycle phases [Kik10]. PDQ is
important to both uses.

After collecting & processing data and conceptual design development, the most critical phase of the methodology occurs - building the digital twin. A digital model of the physical plant will be created by modeling input and output factors, such as material flow, time, yield rate, product output, etc. The data should come from the real plant. There should be a model structure – designed and integrated among various platforms. The model structure should facilitate data transfer seamlessly between the platforms in real-time. In the developing method logic step, data is visualized, monitored, and analyzed. The advanced analytic platform and technologies can be used to develop insights and recommendations. The optimization of the manufacturing plant can be done by establishing model logic successfully.

After building the digital twin, a user executes the digital twin model. The output of the digital twin provides feedback to the user, so that users can make an informed decision. The outcome of the recommendation step is visualized and highlighted. The performance can be evaluated, and different scenarios can be tested. The digital twin will facilitate and leverage quick changes and long-term impacts on the manufacturing plant. From this developed knowledge, stakeholders can update the back-end system to achieve a more favorable outcome.

3.4 Proposed vs Other Methodologies

Throughout literature and different industrial use cases, a multitude of methodologies had been proposed in order to create a methodology for digital twins development and implementation. The analysis of the different proposed methodologies indicates the necessity of a generic one for the manufacturing plant [Ghi20]. The analysis includes digital twin methodology for big data, augmented reality, internet of things (IoT), complex systems engineering, cloud service, and advanced simulation techniques. However, none of the methodology directly addresses digital twin for manufacturing plants, especially when two or more domains combine. For instance, if
a manufacturing plant requires to combine cloud and IoT, the given methodology will not work. The proposed methodology in this study supports quick diagnostics, human involved manual process, existing system integration, and combined domains. In the following Table 3.1, shows Digital Twin Methodology (DTM) in different domains and how proposed methodology fares against other existing DTM in the literature in the manufacturing domain.

<table>
<thead>
<tr>
<th>Domains of Digital Twin Methodology (DTM)</th>
<th>Identify Bottlenecks</th>
<th>Support Manual Process</th>
<th>Integrate with Existing System</th>
</tr>
</thead>
<tbody>
<tr>
<td>Big Data DTM</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Augmented Reality DTM</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>IoT DTM</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Complex System DTM</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Cloud Service DTM</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Advanced Simulation DTM</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Proposed DTM</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

3.5 Conclusion

In the next two chapters, two case studies are presented to instantiate the proposed methodology. The first case study integrates different platforms between level 2 and level 3 of ISA-95. Another case study integrates different manufacturing operations and operates at level 3 of ISA-95. Both the case studies are real industry problems and focus on improving without discarding the existing system.
Chapter 4

Digital Twin Implementation by Integrating Different Platforms

This chapter’s work has been partially published in a peer reviewed conference and in a journal. The publication list is given below.


This chapter presents research results and experience for the establishment of a Digital Twin. The digital twin consists of integrating a dynamic simulator and an advanced process controller in a manufacturing system using OPC-UA. The OPC-UA communication protocol, which is middleware, acts as a common interface between these systems. The case study works between ISA-95 level 2 and 3. The case study is a proof of concept of the proposed digital twin development methodology.
4.1 Introduction

In recent years, there has emerged a significant shift in the interconnection of physical components on the manufacturing floor where transmitted information is used for control purposes [Bra15]. Many quintessential requirements have been identified: ubiquitous connectivity, local intelligence, safety, self-organization, flexibility, massive data monitoring, and efficiency, to name a few [Gon17]. The most common and crucial identified factor is efficient and reliable communication when dealing with heterogeneity and interoperability of various entities on a manufacturing plant. Advanced communication and information technologies can help achieve reliable, smooth, and robust integration between manufacturing levels through various physical media and protocols [Wag03]. This section reports a case study of digital twin that integrates a simulator and controller via communication protocol: OPC UA following proposed methodology. OPC UA is a sophisticated, scalable and flexible mechanism for establishing secure connections between clients and servers. The case study uses Tennessee Eastman problem to formulate the base problem. [Dow93].

This study also provides a baseline for digital twin methodology through system integration, automation, and control. It establishes a standard-based communication between two different platforms on a manufacturing plant. The flexible and scalable communication approach can be used for similar manufacturing problems.

4.2 Background

A common infrastructure model and communications methodology can improve interoperability, enable more secure and efficient data transmission, and facilitate smart data usage. The research community has made significant efforts to introduce digital communications in control and field networks [EB17]. In this study, we use the OPC-UA standard to enable communication and
integration between the controller and the simulation of the Tennessee Eastman process. The reasons behind selecting OPC-UA over publish-subscribe or TCP/IP technologies are multi-folds. It is platform independent, scalable, user friendly, and secure. OPC-UA is a completely new paradigm for systematic communication.

4.2.1 OPC-UA

Increasing demand for data exchange in a manufacturing plant requires better efficiency in communication networks [Mej17]. As a result, newer advanced automation and control domains continue to emerge [Alc13]. These new domains face a continuously increasing requirement of integration and interoperation. Therefore, standardized communication protocols are crucial for integrating manufacturing systems [Her12] [Frü15]. OPC emerged as an automation standard primarily driven by automation vendors in process industry [Roh13]. OPC defines a standard set of objects, interfaces, and methods to facilitate interoperability between control devices and systems. OPC’s connectivity layer helps improve system interoperability.

In the 1990s, Microsoft introduced the Component Object Model (COM) and the Distributed COM (DCOM) interface standards. In 1995, Rockwell, Opto22, Intellution, and Fisher Rosemount developed a data-access standard based on COM and DCOM, and called OPC. Classical OPC include DA (Data Access), AE (Alarm & Events), HDA (Historical Data Access), and DX (Data Exchange). Each of these interfaces has a unique read and write command structure that impacts only one interface at the time. OPC-UA can be implemented on multiple platforms and no longer relies on COM/DCOM technologies [Toi16].

The objective of the OPC-UA is to fulfill all the requirements for platform-independent system interfaces with versatile modeling capabilities that satisfy the needs of even complex systems. Independence of platform and scalability are necessary to facilitate the integration of OPC interfaces directly into a system that runs on various platforms. Access control and security
are also crucial requirements because communication should be allowed through firewalls. The basic premise of OPC-UA is that the client can access small pieces of data without having to understand the entire complex model.

Therefore, it is widely accepted as an enabling technology for digital manufacturing [Pau16]. So far, OPC-UA has been implemented by almost 20 different industrial sectors including tobacco, pharmaceuticals, and automation industries. These users have documented limitations including insufficient semantics, data models, dependence on COM/DCOM technology, inadequate security, and lack of implementation Application Programming Interfaces (API).

### 4.2.2 Tennessee Eastman Problem

Tennessee Eastman (TE) is a well-known industry problem [Dow93]. The original TE problem has a complex structure. It includes a three-unit operation: an exothermic, two-phase reactor; a flash separator; and a stripper. The TE problem contains 41 measured output variables and 12 manipulated variables. The TE problem has been solved with efficient algorithms using different modeling languages and tools. Downs & Vogel (1993) provided FORTRAN code of the model but did not publish the model equations. As a replacement, they provided a flow sheet, a steady-state material balance, and a qualitative description of the critical process characteristics. So, researchers who adopt the case need to make some assumptions to fulfill the missing information. In this case study, a simplified version of the TE problem has been adopted [Ric93]. The simplified TE problem is in the steady state with a relatively modest structure. The details of the simplified TE problem will be discussed in the next Section.

### 4.3 Case Study

According to the proposed methodology, first the TE problem is formulated. After that, value stream map or physical process map is required. To create the value stream map, this case
study has adopted the simplified version of the TE process as the case study and performed the simulation and control modeling. The model includes a controller and the TE problem simulation using two different applications between which information must constantly flow. OPC-UA acts as middleware between the applications. The scenario is pragmatic and can be reused for other similar real-world cases. Figure 4.1 illustrates the overview of information flow of the simplified TE process. First, we have derived an optimization problem from the simplified TE case. Then, we identify the optimal parameters for controllers. These optimal parameters, as control set points, are then sent to the process simulator via OPC-UA. OPC-UA plays the communication medium in this case study. Therefore, a lot of focus is given to the OPC-UA implementation which is critical for digital twin development. Then comes the recommendation which can be done through OPC-UA as well. Simulator sends the feedback to the controller at regular intervals. With this feedback, controller adjusts the new optimal parameters and send them back to process simulator. Therefore, a continuous and effective communication takes place at regular interval.

![Figure 4.1 Information Flow between modules of the case study](image)

### 4.3.1 Simplified Tennessee Eastman Problem

As shown in Figure 4.2, the simplified TE problem includes a combined reactor and separator vessel. The model has two input flows (Feed 1 and Feed 2) and two output flows (Feed 3 and
Feed 4). Feed 1 admits gas compounds A and C into the reactor while pure A is used to control the ratio between A and C through Feed 2. Product D, a liquid, exits through Feed 4, while the purge vapor flows out through Feed 3. In summary, the inputs of the system are A and C, and the outputs are D and the vapor purge as seen in Equation 4.1.

\[ A + C = D + \text{Purge} \]  

\[(4.1)\]  

**Figure 4.2** Process Schematic of the Simplified TE Problem

### 4.3.2 Optimization Problem Formulation

The optimization problem is used to derive the optimal parameter values to serve as control set points. The optimization problem has been formulated based on the model developed by Ricker (1993), which assumes that the plant is at the steady state. The optimization objective of the
problem is to minimize the instantaneous cost of producing a given amount of product D per hour, which depends on three user-provided input parameters: the product flow rate in kmol per hour, the cost per kmol of A, and the cost per kmol of C. The optimization result includes optimal values for six parameters that allow users to enact the most cost-effective setup. The parameters manipulated to achieve minimum cost are the valve positions (as a percentage open) of Feeds 1 to 3 as well as the total pressure of the system. From these values, the valve position of Feed 4 can also be calculated. These five variables, as well as the instantaneous cost, are returned after the optimization execution. The mathematical model is described below:

\[
\text{Minimize } C = \frac{1}{F_4} \left[ C_A (y_{A1} \chi_1 F_{1\text{max}} + \chi_2 F_{2\text{max}} - F_4) 
+ C_C (y_{C1} \chi_1 F_{1\text{max}} - F_4) \right] \quad (4.2)
\]

such that

\[
k_0 \left( \frac{P}{\chi_3 C_{v3} \sqrt{P - 100}} \right)^{1.6} \left( y_{A1} \chi_1 F_{1\text{max}} + \chi_2 F_{2\text{max}} - F_4 \right)^{1.2} 
\ast (y_{C1} \chi_1 F_{1\text{max}} - F_4)^{0.4} - F_4 \leq 0 \quad (4.3)
\]

and

\[
y_{C1} \chi_1 F_{1\text{max}} \geq 0.8 (y_{A1} \chi_1 F_{1\text{max}} + \chi_2 F_{2\text{max}})
\]

where,
\[
\chi_1 = \text{Feed 1 valve position (\%, expressed as decimal)}
\]
\[
\chi_2 = \text{Feed 2 valve position (\%, expressed as decimal)}
\]
\[
\chi_3 = \text{Purge valve position (\%, expressed as decimal)}
\]
\[
\chi_4 = \frac{P}{\chi_3 C_{v3} \sqrt{P - 100}} = \text{Product valve position (\%, expressed as decimal)}
\]
\[
P = \text{Total pressure of system (kPa)}
\]
\[
F_4 = \text{Product flow (kmol/h)}
\]
\( C_A = \) Cost of A ($/kmol)  
\( C_C = \) Cost of C ($/kmol)  
\( y_{A1} = \) Concentration of A in Feed 1 (\%), expressed as decimal  
\( y_{C1} = \) Concentration of C in Feed 1 (\%), expressed as decimal  
\( F_{1_{\text{max}}} = \) Maximum flowrate of Feed 1 (kmol/h)  
\( F_{2_{\text{max}}} = \) Maximum flowrate of Feed 2 (kmol/h)  
\( k_0 = \) Constant value associated with reaction

Equation 4.2 represents the relationship between the reaction rate of the system and the product flow rate based on the time-based equations from the model. Since the problem was assumed to be in steady state, equation 4.2 was derived by setting Ricker's state equations (1) through (4) equal to zero [Ric93]. The cost equation naturally favors A, so equation 4.3 ensures that an ideal ratio between A and C is maintained. Table 4.1 lists the variables and their descriptions. The variables are sorted into three categories: output variables, input parameters, and nominal values. An optimal value is assigned the output variables by the optimization solver, input parameters, as constants, are provided by the user, nominal values are taken from Table 3.1 [Ric93].

After executing the optimization, the optimal and target values (feed valves 1, 2, 3, and 4) are derived. These target values are used by the controller as set points values.

4.3.3 OPC-UA Client Development: Advanced Process Control

Process control plays a vital role ensuring conformity to process rules and protecting the process environment. Real-time optimization (RTO) can be deployed in a controller to determine the optimum control set points for the current operating conditions and constraints. The operating constraints for a plant are identified as part of the process design. During plant operations,
Table 4.1 Summary of variables and nominal operating conditions

<table>
<thead>
<tr>
<th>Output Variable</th>
<th>Set Value</th>
<th>Description</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\chi_1$</td>
<td>0.609533</td>
<td>Feed 1 valve position</td>
<td>(%)</td>
</tr>
<tr>
<td>$\chi_2$</td>
<td>0.250223</td>
<td>Feed 2 valve position</td>
<td>(%)</td>
</tr>
<tr>
<td>$\chi_3$</td>
<td>0.392578</td>
<td>Feed 3 valve position</td>
<td>(%)</td>
</tr>
<tr>
<td>$\chi_4$</td>
<td>0.470302</td>
<td>Feed 4 valve position</td>
<td>(%)</td>
</tr>
<tr>
<td>$P$</td>
<td>2700</td>
<td>Total system pressure</td>
<td>kPa</td>
</tr>
<tr>
<td>$C$</td>
<td>0.2415</td>
<td>Instantaneous cost</td>
<td>$/kmol</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Manipulated Variable</th>
<th>Set Value</th>
<th>Description</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_4$</td>
<td>100</td>
<td>Product flowrate</td>
<td>Kmol/hour</td>
</tr>
<tr>
<td>$C_A$</td>
<td>2.206</td>
<td>Cost of A</td>
<td>$/kmol</td>
</tr>
<tr>
<td>$C_C$</td>
<td>6.177</td>
<td>Cost of C</td>
<td>$/kmol</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Constants Variable</th>
<th>Set Value</th>
<th>Description</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y_{A1}$</td>
<td>0.485</td>
<td>Concentration of A in Feed 1</td>
<td>(%)</td>
</tr>
<tr>
<td>$y_{C1}$</td>
<td>0.510</td>
<td>Concentration of C in Feed 1</td>
<td>(%)</td>
</tr>
<tr>
<td>$F_{1max}$</td>
<td>330.46</td>
<td>Max flow rate of Feed 1</td>
<td>Kmol/hour</td>
</tr>
<tr>
<td>$F_{2max}$</td>
<td>22.46</td>
<td>Max flow rate of Feed 2</td>
<td>Kmol/hour</td>
</tr>
<tr>
<td>$k_0$</td>
<td>0.00117</td>
<td>Constant for assumed isothermic reaction</td>
<td>–</td>
</tr>
</tbody>
</table>

The optimum operating conditions can change regularly owing to product throughput, process disturbances, by-product as wastes, and economic evaluations. Therefore, it is profitable to recalculate the optimum operating conditions on a regular basis.

In this study, a model predictive control (MPC) is designed to control the TE process simulation. A predictive model controller is part of a multi-level control hierarchy in modern processing plants [Wag03]. We use Aspen DMC3 to develop the MPC controller. Three different types of variables are used: manipulated (MV), controlled (CV), and disturbance variables (DV). The three manipulated variables are three valve positions: $U_1$, $U_2$, and $U_3$ respectively. The three controlled variables are product flowrate $F_4$, pressure ($P$), and product A in the by-product $Y_{A3}$. 

36
The relationship between controlled variables and manipulated variables are adapted from Ricker (1993) [Ric93]. The connections are derived in transfer function format from a state space model of the TE problem.

The constraints of the model are given as below.

- Pressure (P) has an upper bound (3000 Kpa).
- $y_{A3}$ in the purge has a range ($0.429 < y_{A3} < 0.886$).
- Product flow $F_4$ has a set point (100 Kmol/hr).
- All three manipulated variables are unconstrained.

Using the transfer functions, library models are created in Aspen DMC3. Different types of state space models can be stored in a library. These library models can be reused to establish the relation between manipulated and controlled variables. For instance, the first order transfer function’s library model formula is $\frac{K}{T_{ss} + 1} e^{-Ds}$. where $T =$ Time Constant, $D =$ Delay, $K =$ Gain.

In this problem, transferring $g_{32}$ to model library provides $T = 10$ mins; $K = 1.5$; $D = 6$ sec.
Aspen DMC3 provides a visual representation of the library model as well. The graph of transfer function $g_{32}$ is given in Figure 4.3.

![Transfer Function Graph of a Model Predictive Controller (MPC)](image)

**Figure 4.3** Transfer Function Graph of a Model Predictive Controller (MPC)

After storing all the transfer functions in the library model, a master model is prepared. After being simulated offline, the controller is ready to deploy.

### 4.3.4 OPC-UA Server Development: Simulation

A Modelica model has been developed [MV18] to simulate the dynamic behavior of the simplified TE process, and it is based on the mathematical description provided by Ricker (1993). This Modelica model library includes the model of the open loop plant. It consists of a model named Reactor, a connector designated pCon, models for setting the boundary conditions, and two models describing the input and output source valves. The Reactor model represents
the processing unit that combines the behavior of the reactor and the separator. These models have been used to compose the ReactorOpenLoop model, a model describing the behavior of the open loop plant. The model library is written in Modelica 3.3 and has been tested using Dymola 2018 and OpenModelica 1.11.0 64 bits under Windows 2010. The model has been used for the ISO 15746 standard implementation.

4.3.5 Communication Protocol: OPC-UA

The Modelica simulation, discussed in previous subsection, is acting as an OPC server. The controller, designed in Aspen DMC3, is acting as an OPC client. To make a successful OPC-UA connection, OPC client and server need to communicate via nodes. The OPC server needs to identify the nodes and read data successfully. To setup an OPC client using Aspen DMC3, the Cim-IO interface manager first needs to be started. The CIM-IO interface is a communication interface that provides a communication standard for interfacing with various AspenTech products like InfoPlus.21 and third-party software such as Modbus, OPC servers. Through DMC3’s CIM-IO interface manager, the OPC-UA interface gets active and ready to communicate with the server. Next, the OPC-UA client requires connecting to the OPC-UA server via nodes. The nodes’ addresses are provided in the modeling. Then the OPC-UA connection via nodes is tested and deployed. Figure 4.4 is a screen capture that shows the variable names and types, as well as the node address assignment.

After a successful OPC-UA server/client connection, the controller "MPC1" is deployed and starts running. In the Aspen Web Interface module, the feedback from controller and simulation is observable. The history, data exchange information, and controller application can be seen and changed from this module according to the user need.

The optimization execution result provides the controller with target set values for feed
Figure 4.4 Test and Deployment of Node Address

valve positions. The controller uses the constraints to acquire real-time feed valve positions and communicates with the simulation through the OPC-UA messaging protocol. The simulation also provides feedback to the controller via OPC-UA and the controller acts as a check and balance element in the simulation by providing the next set of real-time feed valve positions. Overall, this case study is an implementation of a standard-based communication protocol in the manufacturing domain. The approach can be applied to similar problems in the plant to enable real-time communication between different enterprise levels. Automotive, medical device, consumer electronics, aerospace & defense industry can adapt the technology and march towards Digital Twin.

4.4 Lessons Learned

Even though in this case study, the OPC-UA has been successfully implemented between a controller and a simulator of the simplified TE problem, OPC-UA implementation requires some complicated procedures. Multiple challenges regarding digital twin need to be addressed in a proficient manner to have a smooth OPC-UA implementation. We have identified the following
such major challenges based on our experience. The following figure 4.5 maps the major steps according to the implementation strategy.

Figure 4.5 Methodology Application for the Case Study

- To develop a digital twin, it is critical to select the right applications. For instance, this study has the challenge of selecting OPC-UA server/client enabled applications and specification of a well-defined architectures. Not all the control and simulation applications are OPC-UA enabled, so effort is needed to select an application that is not only capable of modeling the problem (e.g., control) but also establishing a server or a client. Also, OPC-UA has a large set of specifications. It is difficult to assess and estimate the project effort and development time in the beginning of the project. The existing physical system has a limited capacity that can sometimes hinder additional functionality. For instance,
the existing system has only a limited amount of RAM space or processor clockwork speed available for additional OPC-UA accommodation. As OPC-UA memory utilization increases, it poses a threat to the existing infrastructure to crash.

- Because digital twin connects a multitude of applications across firewalls and networks, server security becomes a concern. For instance, like many other message protocol systems, OPC-UA uses authentication, authorization, and encryption via an address space concept. OPC-UA address space provides a standard way for servers to represent objects to clients. It defines objects in terms of variables and methods. The elements of a model are represented in the address space as nodes that are assigned to a node class, e.g., objects, variables, and methods. On the other hand, the software components have different levels of maturity for creating the address space model. Some of these software components provide a graphical user interface (GUI) to model address space and to add nodes and references, while others do not. The GUI generates code to establish OPC-UA connection between servers and clients. With a GUI, therefore, it will be easier for building up the server/client connection.

- Sometimes security protocol is pre-defined in digital twin. It requires huge infrastructure as well as investment to overcome pre-defined digital twin application. For instance, OPC-UA claims all the required security features are built in to minimize the efforts from the developers. In this case study, certificate authority is used as a security measure to ensure data protection. But, using an entire different security measure would be very complicated to establish.

- There is no way to measure the performance of the digital twin currently. At any point of digital twin establishment, it is hard to understand the level of reliability and quality of connection. Moreover, digital twin cannot determine data quality which impacts the performance as well. Since application of digital twin varies in a broad spectrum, it is
difficult to use bench-marking techniques for digital twin.

- Finally, to accomplish a complete semantics interoperability, digital twin alone is not sufficient because it only enables syntactic interoperability between clients and servers. There is a great need for semantics to support analytics and scalability across various application from different vendors. Other pillars of smart manufacturing like Internet of Things (IoT), Smart Manufacturing, etc. should be explored and deployed as well.

Overall, these key challenges contribute to the cost of digital twin implementation and create additional uncertainties.

### 4.5 Discussion and Conclusion

Interoperability is a very critical issue that manufacturers have to deal with. By integrating with other manufacturing interoperability standards with semantics Digital Twin could become an important piece of in semantic interoperability for industrial applications. This case study presents an approach for implementing Digital Twin as an architectural template between different manufacturing systems. The integration of process simulation and advanced process control from two different application environments has not been done through digital twin before. The case study identified the development requirements for the applied problem, standards, and technologies. The feasibility of scenario is also verified in the case study. Valuable lessons learned have been discussed.

However, in this work, the digital twin technology was tested with only a small stream of data in a laboratory environment. In a real-work application, enormous amounts of data have to be transferred and exchanged, which may complicate the implementation. The performance of the data exchange needs to be studied more closely. There are several ways that this work could be continued in the future. A manufacturing case study in which more vendors’ products are
involved for integration; a real process that generates more complex data and more realistic amounts of data could be used to replace the process simulator. In addition, the methodology and scenario of the implementation could be enhanced. More functionalities could be implemented within the system.
Chapter 5

Digital Twin Implementation Integrating different manufacturing processes Involving Human Interactions

This chapter’s work has been partially published in peer reviewed conferences. The publication list is given below.


An adaptive simulation-based, digital twin is developed for a real manufacturing case. The digital twin demonstrates the improvement in predicting overall production output and solutions to existing problems. The case study works at ISA-95 level 3. The case study is another proof of concept of the proposed Digital Twin development methodology and published in two peer
reviewed conferences.

5.1 Introduction

Digital representations develop a major challenge in improving the accuracy of existing and future simulation tools. On top of that, to leverage the full potential of a simulation, latest techniques should be incorporated. Hence, a digital twin, may bridge the gap between the physical and virtual experimentation by improving the added-value of plant’s simulation tools. In this case study, a manufacturing process is presented and studied as a part of an overall cyber-physical system (CPS); while its application to manufacturing is analysed and evaluated with respect to digital twin. Uncertainty on the plant area is the main the reason to deviate from the ideal simulation result. There are various reasons behind that uncertainty. Uncertainty on the shopping floor can constitutes from part unavailability, machine breakdown, human variance, etc. The case study consists a complex aerospace product with a complicated parts list. The data is masked to protect the intellectual property of the company.

The enrichment of the digital twin simulations, with sensor data, is proposed for enhancing the realism of digital simulations and improving the accuracy of their results. Thus, uncertainty and production aspects (e.g. cycle time) of shop floor operations can be improved in a cost effective, rapid and efficient way. A software prototype has been implemented to verify the proposed approach with respect to digital twin implementation steps. The main contributions of this case study are: (1) Applying the digital twin methodology for modeling a complex product in a manufacturing plant. (2) The utilization of critical parameters for the generation of realistic simulations. (3) Providing implementation step by step comparison for the proposed methodology.
5.2 Background

Current algorithms, computations, and solutions that predict how humans will engage in smart manufacturing are insufficient for real-time activities. Improving shop-floor performance is critical to the future success of smart manufacturing. In recent years, major manufacturing countries have made strategical responses that will help make those improvements. For example, there are “Industrial Internet of Things” and in the USA, “Industry 4.0” in Germany, “Made in China 2025” in China, and “Industrial Value Chain Initiative” in Japan. If successful, these collective responses will enable the transition from today’s automated manufacturing plant to tomorrow’s “smart” manufacturing plant. Digital Twin is considered as a set of technologies which enables efficient production and addresses emerging demands for customised production. Such a smart shop floor will require smart production systems, smart manufacturing resources, smart manufactured products, smart raw materials, and smart human operators. Smart production systems, for example, will include order planning, production planning, job scheduling, quality control, on-time delivery, and automated fabrication. All production-system functions have two major objectives: minimize energy and minimize cost.

In a smart shop-floor, the manufacturing resources should be easily reconfigured to respond to the changing market demands and changing shop conditions [Wan17]. The former includes customer orders and raw materials; the latter includes the real-time status of the operators, processes, equipment, and environment. That status is used as inputs to real-time, data-analytics tools whose goal is to make optimal decisions [Zha19]. Despite the significant progress, generating that status and making those decisions is still difficult given the typical challenges and problems on the shop-floor. One of those challenges involves the large number of sources of uncertainty, including the humans who work there. According to Tao and Zhang (2017), within a smart shop-floor, there exists numerous interactions between the physical space and the virtual space. However, those interactions are not deterministic due to the complexity and uncertainty
of suitable technologies that manage those interactions [Tao17]. With the current advancement of digital manufacturing and simulation technologies, managing production processes can be carried out in real time. The related technologies include Digital Factory, Internet of Things (IoT), Cloud Computing, and Service oriented Manufacturing System.

The aerospace, defense, and space industry are some of the top users of these technologies. Nevertheless, since their parts are normally very sophisticated and in low volume, humans are still an essential part of their production processes. There are usually about 5000 to 15 000 parts in production with different arrival times. The manufacturing process can get very complicated when it comes to a low-volume high-value products. This will add uncertainty and unpredictable breakdowns, which makes it difficult for production/factory managers to optimize or strategize the manufacturing operations. The authors have interviewed several production engineers from a recognized defense company and found out that most of the manufacturing plant inefficiencies can be overcome with better decision making capabilities related to the production plans. The use of digital twins will provide the production managers with timely meaningful insights to improve demand management, forecasting, schedule planning, production control, inventory management, and procurement.

There is a common misconception in the small-to medium-sized enterprises (SMEs) that digitalization is difficult or even impossible, especially in the manual process-based manufacturing industries. Therefore, the SMEs are concerned about making any significant changes in their manufacturing floor. This paper addresses the concerns by developing (1) A digital twin of a manufacturing process that has intensive human involvement and (2) a real case study to validate the adaptive simulation-based, digital twin.
5.2.1 Digital Manufacturing

Digital manufacturing is an integrated approach to manufacturing that is supported by information technologies. The integration of products, processes, and resources helps manufacturers make better decisions. It not only enables data-driven, decision-support tools but also stimulates the development of new production forms such as smart manufacturing and Industry 4.0. Digital manufacturing will be essential to the creation of new solutions for manufacturing industries [Zho11]. With the advancement of information technologies, many scholars conduct research on manufacturing digitalization from the system-optimization perspective. Typical, AI-based optimization methods these days include machine learning, artificial neural networks, and artificial intelligent algorithms. Current literature in the digital manufacturing domain usually concentrate on building static models where digitalization is completely separated from the actual production floor. For increasingly complex production needs, the digitalized model lacks adaptability because of its stagnant model premise. Digitalization provides manufacturers with more data for their products, production, and systems. Meanwhile, computing has emerged as the cheapest, most abundant resource that we can deploy to analyze that data. Stochastic simulation has been used to generate future “what if” scenarios; and, manufacturers use those scenarios to improve cost, quality, time-to-market, and throughput. In general, by combining the capabilities of data analytics, simulation, optimization, and real-time synchronization [Sha19]; a digital twin for manufacturing problems can be created. Of course, a specific digital twin implementation will totally depend on the scope, objective, and technologies selected for the manufacturing problem. In this paper, we focus on two of those solutions: modeling and optimization. Conventional studies of digitization for manufacturing usually focus on system-modelling methods. These methods mainly include the definitions of functional requirements, design parameters, and process flow of a manufacturing system. These definitions provide a solid foundation for generating and evaluating design solutions of manufacturing processes. However, little attention
has been devoted to the coupling relationships among different elements. Hence, no single model is suitable for addressing all types of manufacturing problems.

Digital Twin is considered as a set of technologies which enables efficient production and addresses emerging demands for customised production [Chr09]. A survey of modeling techniques in production planning is provided [Jeo16]. Simulation technologies, allowing cost-effective experimentation and validation of manufacturing and product solutions, constitute a key enabling technology in the field of digital manufacturing [Mou15]. The virtual representation and experimentation, enabling early product and process verification, provide promising potential savings in time and cost [Cur07][Law15][Nik19]. With these circumstances, it is evident that modern automation solutions are pivotal for production [Ped16]. However, manual operations are still dominant in low volume and complex manufacturing processes [Alk16]. Uncertainty still plays a major role in the performance of production systems. There are numerous studies, referring to digital human modelling and simulation of human-centred operations. In another research, digital human representation is analysed, and an implementation of a digital human environment is presented [Yan07]. In conclusion, in an industrial environment, human simulation can facilitate the assembly process design, accelerate, and improve product development in a cost-effective approach.

5.2.2 Digital Twin in Manufacturing

An on-going ISO standard effort defines a digital twin in the manufacturing context as “fit for purpose digital representation of an observable manufacturing element with a means to enable convergence between the element and its digital representation at an appropriate rate of synchronization (ISO 2020; Shao and Helu 2020).” The concept of a digital twin was first adopted in spacecraft design by NASA [Bre17] [Bos16][Fer17] [Gri14] [Gri17], viewed a digital twin as a combination of modeling-based methods and optimization-based methods
on the other hand, as a real-time simulation with the capability of transferring information from adjacent, product, life-cycle phases. Currently, although literature on digital twins in manufacturing exist, a number of issues with both the development and the testing of digital twins remain [Kri18]. One of those issues involves the bidirectional data flow between the physical space and the digital space. Răileanu et al. (2020) developed an architecture to address this issue [Răi19]. The architecture consists of four layers: the first layer is dedicated to the physical space where the data are collected and processed. The second layer enables the communication with the third layer. Layers three and four reside in the cloud, the layer three is responsible for data update and aggregation, and the layer four performs analysis and decision making. This architecture was demonstrated by using a shop-floor conveyor, where Radio Frequency Identification (RFID) technology was used to identify and locate the pallets on the conveyor. Open Platform Communications Unified Architecture (OPC UA), as the communication protocol, resides on the second layer of the architecture. Schleich et al. (2017) introduced a conceptual framework for building digital twins for specific applications [Sch17a]. Their application was managing the geometrical variation of the product throughout its lifecycle. That framework ensured certain model properties such as scalability, interoperability, expansibility, and fidelity. Latif et al. (2019) discussed an industrial, information-integration method using OPC-UA between the real process control and a simulation of that same process [Lat19]. Redelinghuys et al. (2019) propose a six-layer architecture that comprises physical twin as levels 1 and 2, local data repositories as level 3, IoT gateway as level 4, cloud-based repositories as level 5, and emulation and simulation as level 6 [Red19]. ISO is also developing a digital twin manufacturing framework standard. It is to provide a generic guideline and a reference architecture for case-specific digital twin implementations (ISO 2020). Currently, many case studies of digital twin implementations are within a laboratory environment. For example, an Automated Guided Vehicle (AGV) or Cyber Guided Vehicle (CGV) with self-adapting behavior was developed for solving a material handling problem (Bottani et al. 2017). There are
also a few industrial cases of the digital twin implementation and evaluation. For instance, Liu et al. (2018) introduce a digital twin for an automated flow-shop [Liu19], Zhang et al. (2019) show a digital twin driven cyber physical production system [Zha19], and Lin et al. (2019) describe a digital twin case study for the steel industry [Li17]. However, the literature about digital twin in manufacturing does not cover the manual assembly-based situation yet. SMEs are still in the rudimentary level where workers are receiving raw materials, processing parts, assembling component parts, and inspecting the final product. The digitalization of this type of process will be challenging, but once implemented, it will help eliminate non-value-added production time and other inefficiencies. The digital twin of a human-involved operation will guide the operators interactively and provide production managers with actionable information that enables them to make their decisions more effectively. This paper provides a specific prototype of a digital twin implementation for a manufacturing process with intensive human involvement, where most of the operations and related data collection are manual or semi-automatic, to demonstrate the benefits and value of a digital twin.

The inclusion of human real behaviour, part unavailability, and machine breakdown can make a modern production floor difficult to simulate and assess the impact. Introduction of digital twin can facilitate data acquisition, accurate simulation, and better analysis. The innovations that have been achieved in the last few years, have provided rich data sources. Low-cost sensors and embedded processors, along with advanced storage capabilities, enable the virtual representations of realistic models of the. An approach towards improving the accuracy of simulation results of virtual tools through the integration of sensor data is presented [Cai17]. The benefits of integrating real-world heterogeneous data into digital models, investigating their behaviour through advanced simulation tools, and changing the initial physical system based on the results, constitute the extensive research during the past years. The companies can either transform their processes gradually or they can undergo radical changes by exchanging entire processes. The problem is a substantial amount of time and equity is required to make
the transition towards digital twin. But, a proper guideline along with a proven framework can setup a gradual roll out of specific measures. Some researchers have already provided some frameworks. For instance, Paulus-Rohmer et al. presented a generalised implementation road map for manufacturing companies and their strategic positioning that comprises four steps [PR16]. The first basic step concerns the identification of the present position and the actual state of the enterprise in its current environment from an internal as well as an external perspective. The following step focuses on the analysis and definition of the desired target state, including potentials and feasibility studies. The third phase refers to the realisation of the strategic position by implementing prototypes and designing business models. Finally, the overall roll out and change management processes are required for adjusting the whole organisation. In another study, Schuh et al. formulated a general procedure that comprises six stages [Sch17b]. It focuses on the benefits of digital competencies based on the accessible information. Stoldt et al. showed the overall planning for digitization using tools of the digital factory [Sto18]. They proposed four steps management plan to transform traditional factory floor into a digital one. However, none of these studies provided a step by step implementation idea. They have proposed the framework and described it through a simple holistic approach. These frameworks are adaptive and insightful, but from a factory floor manager’s perspective, it does not provide guidance to implement the framework. Therefore, a detailed guideline with respect to a complex product can help the stakeholders to take appropriate measures. In this study, the proposed framework is described along with step by step analysis. At the same time, further analysis is performed to show how the digital twin helps to improve machine utilization and productivity.

5.3 Case Study

One of the greatest challenges in enabling realistic simulations is the acquisition and combination of real-world heterogeneous data into the digital models, thus creating digital representations
of real-world situations with high accuracy. In this study, we focus on a defense product that requires a manual, assembly process and a receiving, staging area. The objective of the study is to find the best sequence of assembly operations, given the uncertainties in part-arrival times, machine-breakdown times, data-communication times, integration, and part-obsolescence times. The physical system is approximated as a linear, discrete-event, time-invariant system. The users of the digital twin are the production managers. All data are collected monthly from the production floor.

A high-level instantiation of the digital twin concept is shown in Figure 5.1. The top half depicts the digital elements and the bottom half represents the physical elements of the manufacturing process. From a database, historical data is fed into the digital twin, which is an adaptive simulation, as initial inputs. Initial data comes into two varieties: first pass yield (FPY) and hour per unit (HPU). The digital twin model provides process recommendations as the feedback loop. Then, the production managers can select and apply the recommendations to the manufacturing process. Real-time data is fed to the digital twin from the physical manufacturing process.

Creating the digital twin of a manufacturing plant requires a good understanding of the structure, activities, and processes of the target factory or even the business. As per the proposed framework, a value stream map is required. In that case, it is important to list all the incoming materials and operations. A bill of materials and work instructions can help to prepare the process map. The enrichment of the simulations, with process map, is proposed to enhance their realism and improve the accuracy of their results. After creating the process map, the data points needs to be defined. Rather than collecting random data points, a detailed planning helps to strategize the data collection. The physical system is approximated in this study as a discrete event linear and time-invariant system. It consists of a production station/area, which receives an input denoting the product to be assembled with expected production metrics, and generates an output, which denotes the evaluated outcome of the assembly. A schematic block diagram of
the physical system is presented in Figure 5.2, 5.3, and 5.4.

5.3.1 Physical Elements

The manufacturing process produces one product, called Z, which needs 44 raw materials - at different production stages - from various suppliers. Those stages involve 79 operational workstations including testing, assembling, soldering, torqueing, and inspecting. Each workstation is denoted as "opxx" where xx is the operation number. Operations along with FPY and HPU are collected for a two-year time frame starting from April 2017 to April 2019. Important assumptions about the process, for the purpose of data processing, are given below.

- Each day is 8 h and each month is 30 d
- Only one operation at a time can be processed at an individual operation.
- There are no interruptions during the processing of an individual operation, meaning that
the work on an operation cannot be stopped in the middle and then continued later.

- All workstations are at their own location and the material is transported from one operation to another where the subsequent operation is performed. Due to missing information about transport, the time needed to transport material from one to another operation is set to zero.

- Breaks at work, failures, troubleshooting, etc. are included in the processing times.

- Maintenance work is not considered.

- The work orders are dependent. first block (Z1) and second block (Z2) finishes before the final block starts.

5.3.2 Digital Elements

5.3.2.1 Physical Process Map

Creating a digital twin of a manufacturing process requires a good understanding of the process. First, a process map needs to be created with information about all the raw material and operations. Since the core module of the digital twin is a simulation model, the process map provides the requirements for the development and execution of the model. The key requirements include (1) the definitions of the assembly scenario to be carried out, (2) the operational data captured and analyzed for identifying the key parameters, (3) the critical process parameters and operational constraints to determine the behavior of the physical elements, and (4) the simulation of the assembly scenarios and optimization according to a set of constraints.

The schematic block diagram is divided into three parts for the ease of representation. The diagram represents one single part. There are 44 raw materials coming at different stages from various suppliers. It is going through approximately 85 stations such as testing, assembling,
soldering, torquing, inspecting, etc. Finally, it brings one output, $Z$. Each operation stations are denoted as "opxx" where xx is the operation number.

Figure 5.2 Partial Schematic Diagram Block 1/3
The key requirements for the implementation of the presented approach are the (1) definition of the assembly scenario to be carried out, (2) operational data capture and analysis towards
the identification of key parameters for the manual assembly process, (3) creation of digital twin models with the integration of key motion parameters and operational constraints to them, modelling the behaviour of the physical assets, and (4) simulation of the assembly process and its optimisation according to a set of optimisation constraints.

5.3.2.2 Data Processing

The developed framework needs to explicitly represent all the essential information of all the manufacturing resources in a factory and the developed digital twins can be easily inferred by adaptive simulation and will support self-organizing decision making at a later stage. The digital twin serves as a virtual controller to the physical system. The integration of these two
parts, physical and cyber, where the physical system is controlled by the virtual one through the digital twin. To create the seamless transformation, data needs to be collected through sensors, manual operations, routers, etc. Moreover, it is important to identify the parameters of collecting the data. In this study, operations along with hour per unit (HPU), first pass yield (FPY), and failure data are collected for one year time frame. The manufacturing assets talk to each other in the cyberspace and the system itself constantly optimizes production activities using streamed machine status and capability information.

For this project, The data covers a period of 24 months, starting from April 2017. This experimental data set includes operation numbers along with HPU and FPY, incoming raw materials and their estimated numbers, and assembly operations. There are several real-world issues that need to be considered before further analysis of the data, e.g., missing records of operations, missing operation-end timestamps, unintentional and deliberate errors, security issues related to privacy and nondisclosure of business secrets, etc. Raw data was pre-processed and three relational tables without missing, meaningless, and extreme values were created. The structure of these relational tables is given in Table 5.1. After processing the data, 12 excel files are generated each one representing one month time period from April 2017. The simulation result is compared with April 2018 data.

<table>
<thead>
<tr>
<th>Table 5.1 Structure of Preprocessed Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operation</td>
</tr>
<tr>
<td>-----------</td>
</tr>
<tr>
<td>op10</td>
</tr>
<tr>
<td>AA</td>
</tr>
</tbody>
</table>

An actual start timestamp of a work order is determined as the actual start timestamp of the first corresponding operation. The sequence of operations for an individual work order is
determined by the sequence of the operations actual start timestamps. The block diagrams are showing the operations sequence. Important assumptions in data are given as below.

Despite incompleteness of information, other issues in the data under consideration, assumptions, and simplifications are defined. A simulation tool and further analyses on simulation results are valuable to obtain the big picture, new insights and ideas, and to identify further issues and challenges.

5.3.2.3 Simulation Logic

the core of the digital twin is an adaptive simulation model. Traditionally, the off-line simulation has minimal feedback loop, runs for the entire time period at once, and rarely provides assistance to the user for the next cycle. The adaptive simulation-based digital twin provides the decision-making assistance, allows real-time data input, and enables adjustment capability. The simulation runs for a specific time period (i.e., 1 month) with an initial/historical dataset (HPU and FPY for each operations). When a real-time dataset is collected, the simulation replaces that month’s data with the real-time data. However, if it does not find the real-time dataset, the simulation continues with the existing dataset. The real-time dataset comes from the manufacturing process, and the data is collected, cleaned, processed, and stored in the CSV format. The simulation ends once the entire time period completes (i.e., 12 months). Then a recommendation list is generated by the digital twin; the list gets re-ranked constantly based on the score of each item on the list.

The proposed digital twin approach takes each of the real time excel file and simulate for each month. After accumulation of twelve months, the final outcome is compared with April 2018 data to validate. Again, April 2017 data is simulated for 12 months at a time and the final outcome is compared with April 2018 data. It is found that, digital twin of the manufacturing process provides better results to the real values compared to the basic simulation. Additionally, the digital twin helps to identify critical areas and accounts for incoming material uncertainties.

Figure 5.5 shows the simulation logic in a nutshell. When the adaptive simulation moves into
the “wait” mode, it asks the users if they want to review the process for the worst performing operations. If the user wants to review the recommendation list, based on the latest information, the digital twin shows the five worst performing operations and the user selects one of them to review. Once the user selects and applies the recommendations from the list, the digital twin generates a dataset with improved parameters for that operation. Since the operations are picked from future time periods, the generated dataset waits for the actual operation to happen. Once the actual operation happens and the real-time data for that operation updates the dataset, the generated dataset is compared with the updated dataset. If the difference between the datasets stays below a pre-defined threshold value (30% of the real-time dataset), it signifies that the recommendation has improved the process output. On the contrary, if the difference stays above the threshold value, it shows the selected recommendation has not worked better. Either ways, the feedback goes to the recommendation list and re-ranks the list for the next cycle of usage.

5.3.2.4 Simulation Execution

For simulation execution, it has multiple stages. First, the time period needs to be defined and initial data needs to be provided. The simulation runs and stops after each time period and waits for “new data input”. So, after each simulation run, the simulation stops and asks two questions: 1) is there new data input? 2) does user want to review the worst performed operations? Based on the answers, the simulation algorithm branched out to 4 different ways. The 4 ways are described below.

a) “NO” new data input, and user does “NOT” want to review the performance: the simulation moves to the next time period using the previous data input. The performance review will not be executed either. The algorithm will continue afterwards.

b) New data input “exists”, and user does “NOT” want to review the performance: the simulation moves to the next time period using the new data input. The performance review will not be executed. The algorithm will continue afterwards.
Figure 5.5 Adaptive Simulation Logic
c) “NO” new data input, and user “wants” to review the performance: the simulation moves to the next time period using the previous data input. For performance review, the worst 5 operations will pop up based on user selected performance criteria. Then, a knowledge-based recommendation list will pop as well with respect to each of the worst performing operations. User will pick one or two from the recommendations list and implement in the actual manufacturing floor. At this point, the simulation will be divided into two sub-parts. First, it will assume a certain percentage of improvement will occur based on the recommended selection. The performance criteria parameter will be changed/improved and generate a new data input termed as “Estimated Data Input.” Using the estimated data input, the simulation will go all the way to the end of the total time period without waiting for any more user selection. For the second sub-part, the simulation will continue stopping after each time period and will wait for a “new data input.” Then, it will continue as per the 4 branch ways simulation algorithm.

d) New data input “exists” and user “wants” to review the performance: the simulation moves to the next time period using the new data input. The performance review segment will be executed as per aforementioned (c) segment.

Here, the recommendation list follows reinforcement learning (RL). RL is widely used in control systems to optimize the decisions. It is a “trial-and-error” approach that the learning agents learn optimal decisions by interacting with the environment. The “trial-and-error” rule means RL agents make a trade-off between known decision exploitation and new decision exploration to achieve optimal policy. Figure 5.6 shows the RL principal in a nutshell.

In Figure 5.7, the RL Architecture is instantiated for the case study 2. The recommendation list is the environment. In different state, which is the different simulation time period, new data input enters into the equation and threshold value comparison with generated input. Based on the agent, which is the threshold value in this case study determines the action. Here, the action is re-ranking the recommendation list.

Here,
Figure 5.6 Reinforcement Learning Principal

Figure 5.7 Reinforcement Learning Architecture for Recommendation List

T: The set of iteration time. T includes a sequence of discrete refresh after each new data feed. At each time period t, the agent completes an iteration with the environment. The environment is the recommendation list.

S: The finite set S includes the possible states of the environment. Here, the possible outcome from the recommendation list are the finite set.

A: The finite set $a_t$ is the set of actions available in state $s_t$. 
\[ p_t(s_{t+1}|s_t, a_t) \] denotes the state transition probability that environment transfer from \( s_t \) to \( s_{t+1} \) under action \( a_t \in A(s_t) \).

\[ \pi \] is the policy that map from S and action A. \( \pi(s,a) \) denotes the probability of acting a under state s. Here the probability of the an action under a state is binary, which is either 1 or 0.

In the \( t^{th} \) iteration, the agent observes the current environment state \( s_t \), and chooses an action \( a_t \). After that, the environment transfers from the state \( s_t \) to \( s_{t+1} \) following the probability \( p_t(s_{t+1}|s_t, a_t) \) and returns a reward \( r_t(s_t, a_t) \) according to the performance of \( a_t \).

Figure 5.8 and Figure 5.9 describe the whole simulation flow for an instantiated case for the simplicity. Figure 5.9 is a continuation of Figure 5.8. It is representing simulation flow c – subpart 2, where it mentions "Continues to next page".

The adaptive simulation model was developed using Python 3.6. When a specific operation (e.g., op560) is executed, assuming it is from the worst 5 operations, the user can compare the real-time dataset (FPY and HPU) with the generated dataset (10% improved FPY and 10% decreased time). After applying the recommendations, the real-time data should demonstrate significant improvement from the generated data. If the parameters (HPU and FPY) of the real-time dataset are more than 30% of the parameters of the generated dataset, the applied recommendations will be regarded as a failure, will get a score of 0, and will get pushed down on the recommendation list. If the parameters are similar (real-time data is within 30% of generated data), it indicates that the recommendations work well, it will get a score of 1, and will get push up in the recommendation list.

In either cases, the recommendation list will be updated and stored in the database for use in the next cycle as the initial dataset. With this rule, the recommendation list for the operations always gets updated; this reflects the reinforcement learning. Eventually, all items in the recommendation list will truly represent the recommendations that have been tested and verified with a proven track record. In figure 5, the default initial recommendation list and re-ranked recommendation list for Op560 in the next cycle are shown. , the re-ranked
recommendation list is showing “Find An Alternative Material” at number 1 because in the
previous run, the user has selected and successfully found that the recommendation is helpful.

5.3.2.5 Software Demonstration

In this section, the software is demonstrated step by step. The iteration starts by loading month
1 data. Based on the required time and first pass yield, the operation moves from operation 1
to assembly 1. After that, the iteration stops and provide the 2 options: 1) review; 2) Continue
Simulation without Review. Figure 5.10 is the starting screen of the software.
After choosing option 2, the following 5.11 comes up. The simulation picks month 2 data and keep moving forward.

At this point of the simulation, lets choose option 1 – Review. The review option provides worst 5 operations based on the FPY and time required to completion. Each operation is assigned with a score and based on the score; worst 5 operations are selected.

Now, it asks if the user wants to apply recommendation to any of the worst 5 operations. If the user wants to apply recommendation to any of the worst 5 recommendations, the program looks
for an existing list for that specific operation. If the program finds an existing recommendation list, it will be shown to the user. Otherwise, a new default recommendation list will be presented.

In this case, the program does not find any recommendation list for op560. So, it generates the default recommendation list and shows to the user. Here, user can select any recommendation from the list or provide own recommendation with option ‘X’. Once user selects one recommendation, the program asks if user wants add further recommendation. If user is content with the applied recommendations, the program moves forward to the next stage.

At this point, the simulation is divided into two parts. First, a generated data input is produced where FPY improves by 10% and time decreases by 10% for the specific operation.
The recommendation list gets updated with JSON format and stores in a separate folder. The simulation uses the generated data input and finish the simulation all the way to the end. Secondly, the simulation keeps moving forward asking the next existing data input and follow the original simulation flow. This is the part where the simulation flow mentioned "Continues to next page."
Figure 5.14 Default Recommendation List

Figure 5.15 Simulation Moving after Applying Recommendation
When that specific operation (op560 in that case) will be executed, the user get to compare the existing parameter (FPY and time) with generated data input’s parameter (10% improved FPY and 10% decreased time). Since the recommendations have been implemented, the existing data should demonstrate significant improvement or similarity with generated data input. If the difference between existing data and generated data goes more than 30%, the applied recommendations will get deleted. If the parameters are similar (less than 30%), therefore, the recommendations have been worked. So, the saved recommendation list will be updated. Next time, when a use wants to see the same operation’s performance, the previously selected recommendations will be shown at first. With this setup and added recommendation capabilities, the recommendation lists for the operations will always get updated. Eventually, all the recommendation lists will truly represent recommendations that have been tested and verified.

In the following figure 5.16, on a different run, the user wants to review op560. This time, the recommendation list is showing “Find An Alternative Material” at number 1 because on the previous run, user selected and successfully found the recommendation useful.

### 5.3.2.6 Model Validation

To validate the model, 24 months of actual data is collected. First 12 months data is used as an initial data. The average of the 12 months data is counted for the parameters. The rest of the 12 months data is used to validate the model. Each month data is entered into the simulation following simulation algorithm. The simulation model is predicting the output after each month. The output is compared with the actual real estimated number in the factory. In figure 5.17, 3 data points are taken into account. The basic simulation runs in the first month and estimated all the way to the month 12. Thus, it is showing a flat line of 38 units all through the 12 months. The revised data feeds to the adaptive simulation after each month. Therefore, the estimation varies with each month while capturing the real scenario. For basic simulation output numbers, only the
Figure 5.16 Re-ranked Recommendation List

<table>
<thead>
<tr>
<th>ID</th>
<th>Op Name</th>
<th>TIME</th>
<th>FPY</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>op560</td>
<td>0.75</td>
<td>1.0</td>
</tr>
<tr>
<td>2</td>
<td>op740</td>
<td>0.75</td>
<td>1.0</td>
</tr>
<tr>
<td>3</td>
<td>Assy5</td>
<td>0.5</td>
<td>1.0</td>
</tr>
<tr>
<td>4</td>
<td>Assy6</td>
<td>0.5</td>
<td>1.0</td>
</tr>
<tr>
<td>5</td>
<td>Assy7</td>
<td>0.5</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Do you want to keep applying recommendations?
1. Yes, 2. No

1

Worst 5 Ops

<table>
<thead>
<tr>
<th>ID</th>
<th>Op Name</th>
<th>TIME</th>
<th>FPY</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>op560</td>
<td>0.75</td>
<td>1.0</td>
</tr>
<tr>
<td>2</td>
<td>op740</td>
<td>0.75</td>
<td>1.0</td>
</tr>
<tr>
<td>3</td>
<td>Assy5</td>
<td>0.5</td>
<td>1.0</td>
</tr>
<tr>
<td>4</td>
<td>Assy6</td>
<td>0.5</td>
<td>1.0</td>
</tr>
<tr>
<td>5</td>
<td>Assy7</td>
<td>0.5</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Choose one (by ID number): 1
You have selected 1 ('op560', [0.75, 1.0])

eco/op560.json

1. Find alternative material
2. Add an additional machine
3. Add more human resources
4. Improve process yield
5. Investigate further
6. Add fixture & tools
7. Improve training
8. Collect more data to analyze
9. Standardize the work procedure
10. Strengthen supplier base
11. Collect feedback from worker
12. Automate the process
13. Supervise the process
14. Perform preventive maintenance

X. I Will Enter Recommendation Manually (Type X to choose this option)
Choose one recommendation: (by ID) 3
You have selected recommendation 3 Add more human resources

Finished applying reco for op op560
Do you want to keep applying recommendations?
1. Yes, 2. No
initial data is used to predict. The model remains same, and next 12 months data does not require into the basic simulation model. The actual value changes because of the modified situation over the months. The adaptive simulation shows better tracking and estimation compared to the basic simulation. Therefore, the figure validates the model as well as shows the benefits of implementing a digital twin.

Figure 5.17 Output Estimation

Figure 5.18 depicts the error percentage between two simulation methods. It is clearly evident that adaptive simulation picks up the trend better and shows less error percentage. With application of machine learning in the future, the digital twin will be able to use the adaptive simulation in a more effective way.
5.3.3 Integration

The digital twin includes two-way communications: (1) the real-time data is collected from the manufacturing process, processed, and updated in the digital twin and (2) the recommendation list is generated by the digital twin and applied to the manufacturing process. In this case study, most of the data are collected manually, processed in a separate application. A python-based parser was developed to interpret the CSV files, a text file provides the operation sequence. The recommendation list is saved in JSON. Each time the user manually enters a recommendation or updates the recommendation list, a new JSON file replaces the older one for that specific operation. For instance, if a user makes an update on OP560, the simulation first looks for a JSON file in the recommendation list. If it does not find the JSON file, it shows the default recommendation list. Based on user’s selection, the recommendation list gets updated and saved in the OP560 JSON file. Next time, whenever the user wants to see the recommendation list for OP560, the OP560 JSON file is shown for the user. The python parser does all the data transfer except the initial CSV file generation. The data is transferred once in every month, so the overall

![Error Comparison](image)

**Figure 5.18** Error Comparison between Traditional and Proposed Approach
effort to run the digital twin is minimal. Even though the data transfers monthly, the data comes from the real-manufacturing floor and therefore, it is considered as real-time data. In this case study, the initial dataset comes from the historical database. It is the data from April 2017 to March 2018. The simulation will ask for the real-time data for each month, e.g., April 2018, May 2018, or June 2018, once the real-time data is available, simulation replaces the existing dataset with the real-time data, and recalculate the output at the end of the simulation. If the user wants to improve the efficiency or reduce the processing time of the operation, the simulation provides the worst, five, performing operations that have the highest potential to be improved and a recommendation list for the improvement of each operation. The equation 5.1 used to find the worst five operations is given below.

\[ x = HPU \times FPY^2 \]  

(5.1)

For instance, OP560 has two parameters: FPY as 1.0 and HPU as 0.75. OP560’s ‘x’ value calculates as 0.75. With a similar calculation, the top, five, worst operations are derived based on the lowest ‘x’ value. Therefore, operations with a lower yield rate and higher HPU will show up as a worse performing operation.

Figure 5.19 captures the overall concept of the case study. The collected data is fed into the digital twin to replace the previous dataset. Therefore, the digital twin can periodically adjust its prediction output and point out the areas to improve. The recommendation list brings useful insight to the worst operations. Based on the application and processed data feedback, the recommendation list constantly gets evolved. The successful recommendations come at the top of the list and the failed ones go to the bottom of the list. Therefore, every operation shows a unique list that has been tested and verified over the period of time. For instance, after multiple runs of the digital twin, a recommendation list can show the recommendations that have been successful for the similar situation multiple times. Therefore, the user can make a
better decision.

![Figure 5.19 Overall Concept](image)

### 5.4 Lessons Learned

Once the simulation is completed and the result is shown the better output, the benefits of digital twin is realized. However, it is far more complicated than it seems on paper. We have identified the following such major challenges based on our experience. The following figure 5.20 maps the major steps according to the implementation strategy.

- The real life data collection is a real challenge. Even after having the full access to the data, it takes a lot more time to accumulate and clean the data file.

- It is important to understand the critical parameters to collect the data. Physical process map, previous experience, and crisp communication with all shareholders can resolve the issue. A detailed planning is required to capture these critical parameters. In this study,
HPU and FPY are the two critical parameters which is used for further analysis.

- The simulation is performed in python scripting language. The problem statement and scope determination pose another challenge. Often, the critical parameters e.g. incoming raw materials, HPU, or FPY changes drastically. So, the scripting should have enough adaptability to cover any unforeseen events.

5.5 Discussion and Conclusion

The method presented in this study describes an approach for transferring real-world information of a complex manufacturing system to its cyber counterpart, where data can be further processed
and evaluated in a cost-effective approach. Hence, the digital twin concept demonstrates the potential for being used as a digital test bed for empowering production-wise applications. A test-driven resource virtualization process is proposed as the recommendation for the industry to adopt to create digital twins for their smart factory solutions. The proposed process draws inspiration from past resource virtualization outcomes with special attention paid to the usability of the proposed solution. It provides a straightforward process for companies to create digital twins especially with the assist of an already developed simulation tool with a different approach. Equally important, the research highlighted that the data to be virtualized in the cyberspace consists of three integral components: technical properties, functional capabilities, and real-time status. The most notable contribution of the resource virtualization methodology is its flexibility in the virtualization process, making it easy for plant to adopt. The methodology can be extended to any type of resource virtualization in developing a cyber-physical system. Future work will focus on the extension of the data points, improving the accuracy of the data processing approaches for recognising operations. Moreover, the evaluation of alternative approaches for parameter extraction with the support of additional sensor systems will be investigated, towards increasing the accuracy of the digital twin.
Chapter 6

Conclusions & Future Work

6.1 Overall Summary & Research Contribution

Companies nowadays are required to transform the current practice of product development management into a smart factory solution that enables rapid production configuration to achieve the fast production of individualized products at changing scales. The digital twin creates concrete value by helping production personnel to make key strategic decisions. This directly impacts generating revenue streams. The digital twin has many applications across a product life cycle. It is easier to answer the critical what-if questions in real-time more accurately. It is expected that this kind of analysis can contribute to better management of operations and thus, to the more efficient execution of operations, better utilization of resources, shorter lead times, and higher due-date reliability. With the advancement of the manufacturing domain, the digital twin will become more crucial. As discussed in Chapter 2, the study focuses on existing research and finds the literature gap. A digital twin for the manufacturing plant lacks a holistic approach and real-life use cases. From the literature review, it is obvious that industries are not embracing the digital twin concept. Over time, researchers have proposed and created new algorithms to address different scenarios. However, all these algorithms are not proven in
the pragmatic case studies. Most of the case studies are overly simplified or too hypothetical. It is often difficult to customize and adapt to other manufacturing problem contexts. There is robustness in the solution, but it is very specific to the field. Visualization is missing and significant computational effort is required to perform these algorithms. Therefore, all these algorithm-based techniques failed to make an impression in the real plant with a changing situation and looming uncertainties. This study proposes a digital twin methodology to leverage recent advantages in computing and networking, known as Digital Twin, in supporting the industry 4.0. The proposed methodology is generic and provides an easy-to-follow guideline for the industries. Two case studies are provided to validate the proposed methodology: 1) a continuous product from an established industry problem; 2) a discrete product from a reputed manufacturing industry. In chapter 3, the study focuses on the integration of different platforms and how the digital twin methodology can help to achieve integration. A long-time industry problem – Tennessee Eastman Problem is modified. A controller and a simulator are created in two different platforms. The real-time messaging protocol system, OPC-UA is applied to get feedback from the simulation and use a controller to change the values of the controllable parameters. The digital twin proposed methodology guided step by step through the process of establishment. In chapter 4, a complex product’s manufacturing process is taken. The objective is to create a digital twin of the manufacturing process so that it can support the decision making at operations management level. Unlike a usual simulation, an adaptive simulation-based approach helps to predict and adjust the uncertainties along the production line better. Feedback is a very important part of the proposed methodology. In this problem, the digital twin aka adaptive simulation is based on the continuous monthly feedback. A recommendation list is created which evolves continuously for each operation. The recommendation list leverages reinforcement learning after each simulation. It helps to understand the bottleneck of the process and articulates suggestions based on an agent-based reward system, which is also known as reinforcement learning. After sharing the proposed methodology with the aerospace and
defense industry’s employees, it is well appreciated and recognized as a value added tool to the organization. According to the employees, the proposed methodology can significantly advance existing processes, products and services, and often lead to new business opportunities. They can significantly cut operating costs, leading to real bottom-line improvements. The technology can be widely adopted across the current setup. One of the ancillary benefits of digital twin is the capability of assimilating cross-functional teams to drive digital twin project success is the opportunity to link organizational tools, skills, and knowledge bases. However, there are a couple of improvement suggestions. The implementation strategy is case-specific and overall implementation strategy is not applicable for a more complex layout. The proposed methodology is quite simple to follow; however, factory managers may find it hard to follow when cloud computing, sensor data, and updated function appear. There is no guideline on how to cover the missing step/piece in the methodology. The proposed methodology is heavily dependent on the user input. There is no fall back method or plan to ensure user integrity. the methodology is heavily reliant on humans following through on recommendations made by the digital twin.

Taking all of the limitations into account, some of them are out of the scope of the research, and the rest of them can be addressed by the future work. For instance, maintaining user integrity is out of scope of the research. However, cloud computing, updated functions, and new sensor data can be included in the future work scope. More case studies can be performed within the same proposed methodology to capture more problems in the domain.

Overall, the contributions of this study include:

• A review of research publications dedicated to the digital twin area in the manufacturing domain.

• Identification of literature gap in the digital twin and propose a digital twin development methodology for manufacturing.

• Two real-life case studies to validate the proposed methodology. The results are demon-
strated to establish the potential benefits of a digital twin.

• The case studies can be used as a step-by-step guideline for the industries to follow. Both the case-studies are picked to cover most bases of the manufacturing domain: continuous and discrete product, automatic and human involved process, real-life industrial product, and established industry problem set, between level 2 & level 3 and at the level 3 of ISA 95, etc. Therefore, industries can pick and choose any of the guidelines as per their need.

• A python-based software is created which provides a recommendation list based on adaptive simulation. The structure of the software can aid the manufacturing industries to build their own.

According to Wagner et al. (2019), Current digital twin literature lacks 1) Integration of different domains; 2) Interfaces of information exchange; 3) Efficient design of information flow [Wag19]. Both case studies bridge the literature gap and the proposed methodology shows improvement over the existing systems.

6.2 Future Work

Digital twin as any other technology has many limitations and challenges, such as the high cost related to construction and acquiring the right expertise, which is rare at the same time, data availability, and so forth. Using digital twins during the life cycle of any project would be accompanied by many threats and opportunities. This relates to many factors such as the type of the manufacturing process, the phase, the circumstances, etc. Although this study addresses some of the common problems of the digital twin, the cost reduction and relevant expertise require more time to develop. Future work can include the following areas.
6.2.1 Hybrid Approach

Future work should involve the digital twin application with a more hybrid approach. For instance, some proven algorithms (e.g., advanced fuzzy theorem, neural network, genetic algorithm, etc.) can be applied to improve adaptive simulation output. The fidelity of the model can be increased by inserting more data and variables. The tests should also be made exploring more data transmission technologies and protocols (e.g., MTConnect, serial communication standard RS232, RS422, ModBUS, etc.) used by the existing IoT platforms to define the ways of development for better adoption of the proposed approach.

6.2.2 Product-Oriented Effort

The current study uses real-life industrial problems from the aerospace industry. Similar case studies can be performed in a different field to help a more product-oriented digital twin effort. Instead of reinforcement learning, other machine learning or fuzzy applications can be applied and compared to find the best alternate plan. Both studies have not considered uncertainty in the calculation. Future studies can integrate uncertainties and take a more realistic approach to the implementation of the digital twin.

6.3 Conclusion

In recent years, digital twin achieved great success by showing an important role among different industries and domains, as well as by embracing new technologies. There has been an increased focus on it as there is a strong belief that digital twins could provide a proper tool for balancing the conflicts inherent in exploring opportunities in the digital and physical – both the worlds. Different standards and methodologies have been developed to identify, assess, and manage digital twins. So, with the rapid development in technology, it is necessary to get the advantage
of these new technologies to enhance digital twin. Looking at the current situation and what is on the horizon, it is very clear that digital twins will be a major part of the risk management future.

Thus, the use of well-known and verified methods for the physical world with the capabilities of current information technologies in the digital world opens wide opportunities in the areas of Digital Twin. The current work has considered some of the possible applications, but many other of them should be considered and developed in the nearest future.
BIBLIOGRAPHY


[Chr09] Chryssoulakis, G., Mavrikios, D., Papakostas, N., Mourtzis, D., Michalos, G. & Georgoulias, K. “Digital manufacturing: history, perspectives, and outlook”. Pro-


