Automated Planning and Scheduling Method for Modular Construction Manufacturing

Angat Pal Singh Bhatia

A Thesis

In the Department of

Building, Civil and Environmental Engineering

Presented in Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy (Civil Engineering) at

Concordia University

Montreal, Quebec, Canada

December 2024

CONCORDIA UNIVERSITY

SCHOOL OF GRADUATESTUDIES

This is to certify that the thesis prepared

By: Angat Pal Singh Bhatia

Entitled: Automated Planning and Scheduling Method for Modular Construction Manufacturing

and submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy (Civil Engineering)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

______________________________________ Chair
Dr. Christian Moreau

______________________________________ External Examiner
Dr. Farook Hamzeh

______________________________________ Arms-Length Examiner
Dr. Ming Yuan Chen

______________________________________ Examiner
Dr. Mazdak Nik-Bakht

______________________________________ Examiner
Dr. Ashutosh Bagchi

______________________________________ Thesis Supervisor
Dr. Osama Moselhi

______________________________________ Thesis Co-Supervisor
Dr. Sang Hyeok Han

Approved by ______________________________________
Dr. Po-Han Chen, Graduate Program Director

12/2/2024

____
Dr. Mourad Debbabi, Dean, Gina Cody School of Engineering and Computer Science

ABSTRACT

Automated Planning and Scheduling Method for Modular Construction Manufacturing

Angat Pal Singh Bhatia, Ph.D.

Concordia University, 2024

Modular construction is a promising alternative to conventional construction; offering improved productivity, high-quality end products, and reduced labour requirements. To realize these benefits, sequencing module components during prefabrication process in a manner that ensures efficient allocation and utilization of labor resources at workstations is essential. However, one of the significant challenges in modular construction manufacturing (MCM) is that it follows a make-to-order process, resulting in customized module components. This customization leads to variations in the design specifications of module components, causing different processing times at each workstation. These imbalances in production lines result in increasing the waiting time for module components between the workstations, ultimately extending the makespan. This poses a challenge for production line managers, requiring frequent adjustments to plans and schedules related to the sequencing of module components at workstations using conventional methods.

To address these challenges, this thesis introduces a framework composed of three modules: (i) a simulation-based statistical method for planning in modular construction; (ii) a deep neural network (DNN)-based method for predicting production process times; and (iii) a hybrid optimization technique for scheduling in modular construction. In the first module, a simulation based statistical method is developed to plan the sequencing of module fabrication and the allocation of workers at workstations. The method encompasses data collection process to obtain historical/near real-time data and identification of significant impact factors affecting process times at workstations along the production line. In the second module, a newly developed method for predicting processing time at each workstation is introduced utilizing Deep Neural Network (DNN), Artificial Neural Network (ANN), and Multiple Linear Regression (MLR) for predicting production process time spent at each workstation in a manufacturing plant. The third module focuses on planning and scheduling method that ensures optimal sequencing of module components at workstations using Genetic Algorithm (GA), Simulated Annealing (SA), and Hybrid Genetic Algorithm Simulated Annealing (HGASA).

Two case studies were analyzed to demonstrate the use of the developed methods and test their performance. The first case is of a light gauge steel (LGS) wall panel production line operated by a modular fabricator in Edmonton, Canada, and the second is of a wood-based semi-automated wall panel production line also in Edmonton, Canada. These cases involve the production of 200 wall panels in the first case and 39703 wall panels in the second at various workstations along the production line. The simulation-based statistical method developed in the first module yielded 89.39% accuracy in prediction of process time and indicate a 44.42 hr duration to produce 309 wall panels with regards to first case. The results of the second case showing process time predictive method developed in the second module for most workstations had a mean absolute error (MAE) of under 2.50 minutes, with symmetric mean absolute percentage error (SMAPE) ranging between 22 % - 28%, respectively. The developed scheduling method of the third module provided an optimal sequence of wall panels for prefabrication, minimizing makespan. As a result, the hybrid optimization reduces makespan to 105.63 hr from those generated by GA (138.08 hr) and SA (108.06hr).

ACKNOWLEDGEMENTS

First and foremost, I wish to express my deepest gratitude to my supervisors, Dr. SangHyeok Han and Dr. Osama Moselhi for their continuous encouragement, help and valuable advice throughout all stages of this research. I would also like to extend my gratitude to Dr. Sadiq Altaf and Dr. Mohamed Al-Hussein for their help in my research by sharing the details of the production line of the case study company. I am grateful to my examining committee members —Dr. Farook Hamzeh, Dr. Ming Yuan Chen, Dr. Mazdak Nik-Bakht and Dr. Ashutosh Bagchi—for their comments and feedback. During the early stages of my PhD, I was fortunate enough to get the opportunity from Dr. SangHyeok Han and Dr. Osama Moselhi to visit the Fortis LGS Structures Inc. which really shaped my understanding about the modular construction production lines. In this respect, I would like to thank the management team at Fortis LGS Structures Inc. for their help during this process. Also, special thanks to Dr. Zhen Lei for helping and giving time during my visit to the Fortis LGS Structures. Additionally, I want to thank all my friends and colleagues at Construction Automation Lab and Concordia University for their friendship and support over the years, especially- Hassan Bardareh, Shrouk Gharib, Ahmed Zaalouk, Mizanoor Rahman, Islam Hamdan, Abdelhady Omar, Moazz Elkabalawy, Haindhavi Reedy, Aida Azarian, Ehab Elhosary.

I am especially grateful to Dr. Sungkon Moon, for introducing me to Dr. SangHyeok Han and recommending me for this PhD program.

With love and respect, I would like to express my heartfelt gratitude to my mother and father, who are always there to give and add to my life.

TABLE OF CONTENTS

List of Figures

List of Tables

List of Acronyms

MBI	Modular Building Institute
MCM	Modular Construction Manufacturing
SOP	Standard Operating Procedure
CPM	Critical Path Method
AON	Activity on Node
LSM	Linear Scheduling Method
SIF	Significant impact factors
DNN	Deep neural network
ANN	Artificial Neural Network
MLR	Multiple Linear Regression
DES	Discrete Event Simulation
GA	Genetic Algorithm
SA	Simulated Annealing
HGASA	Hybrid Genetic Algorithm Simulated Annealing
RFID	Radio-Frequency Identification
LGS	Light Gauge Steel
MAPE	Mean Absolute Percentage Error
LPS	Last Planner System
PBS	Production Line Breakdown Structure
VSM	Value Stream Mapping
CCPM	Critical Chain Project Management
JIT	Just-in-Time

BIM	Building Information Modelling
NP	Non-Deterministic Polynomial-Time
PSO	Particle Swarm Optimization
ACS	Ant Colony System
MCMSP	Modular Construction Manufacturing Scheduling Problem
CTF	Cycle Time Formula
PCA	Principal Component Analysis
PCC	Pearson Correlation Coefficient
RMSE	Root-Mean-Square Error
MAE	Mean Absolute Error
SQL	Structured Query Language
WPS	Work Planning Structure
Int	Interior
Ext	Exterior
K–S	Kolmogorov–Smirnov
CNC	Computer Numerical Control
SD	Standard Deviation
KPI	Key Performance Indicator
UT	Updated Temperature
IRT	Initial Read Time
PT	Process Time
SVD	Singular Value Decomposition

List of Notations

V'	Normalization
A	Independent Variables
μ	Population Mean
σ	Standard Deviation
$\bar{x}$	Mean of The Sample Data
Q	Eigenvectors
Δ	Diagonal Matrix
W	Workstation
W+1	Next Workstation
M	Module Component
A_i	Actual Process Time
P_i	Predicted Process Time
Y	Dependent Variable
β_0	Bias Value
β_k	Weighting Coefficients for Independent Variables
M-1	Previous Module Component
$S_{m, w}$	Start Time of Module Component at Workstation
X	Workstation Capacity
$C_{m, w}$	Completion Time of Module Component at Workstation
S_n	Neighbour Solution
S	Current Solution
T	Temperature Value
X_k	Independent Variables

$PT_{m,w}$	Process time of module component at workstation
D	Decoding
P_w	Parallel Workstation
P_{LB}	Production load balancing
WP_{mt}	Makespan of a work package

CHAPTER 1: INTRODUCTION

1.1 Background

Canada requires 4.3 million affordable homes for individuals with low incomes and more facilities like hospitals and long-term care homes to support a growing population, according to the CSA Public Policy Centre (Dragicevic and Riaz 2024). In the meantime, it is projected that 700,000 skilled workers will retire between 2019-2028, increasing labor shortage (ESDC 2022). Decision-makers are considering modular construction as a means to address such challenges. Modular construction, also known as panelized construction, off-site construction, volumetric manufacturing, and industrialized construction, is noted for its efficiency. According to the modular building institute (MBI 2023), modular construction can be utilized to construct condominium, dormitory, and duplex homes in about half the time required in conventional construction. Figure 1.1 shows that multifamily projects held the largest market share in the modular industry at 32% in 2022. The United Kingdom's housing sector brought significant attention to modular construction in 2004 by aiming to build 25% of new houses using these techniques to meet housing demands (Boyd et al. 2013). Modular construction has also been utilized in other countries such as the United States, Canada, China, and Australia (Steinhardt and Manley 2016). In this type of construction, module components such as wall, floor, and roof panels are prefabricated in mass production under a controlled factory environment and transported on-site for assembly of built facilities. The principal idea underlying this construction process is to reduce the on-site work and instead perform the majority of work, off-site, at factories under a controlled work environment.

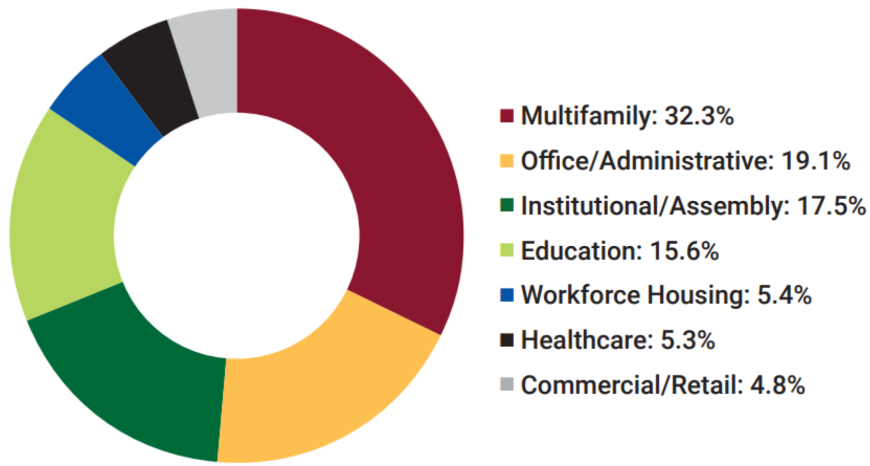
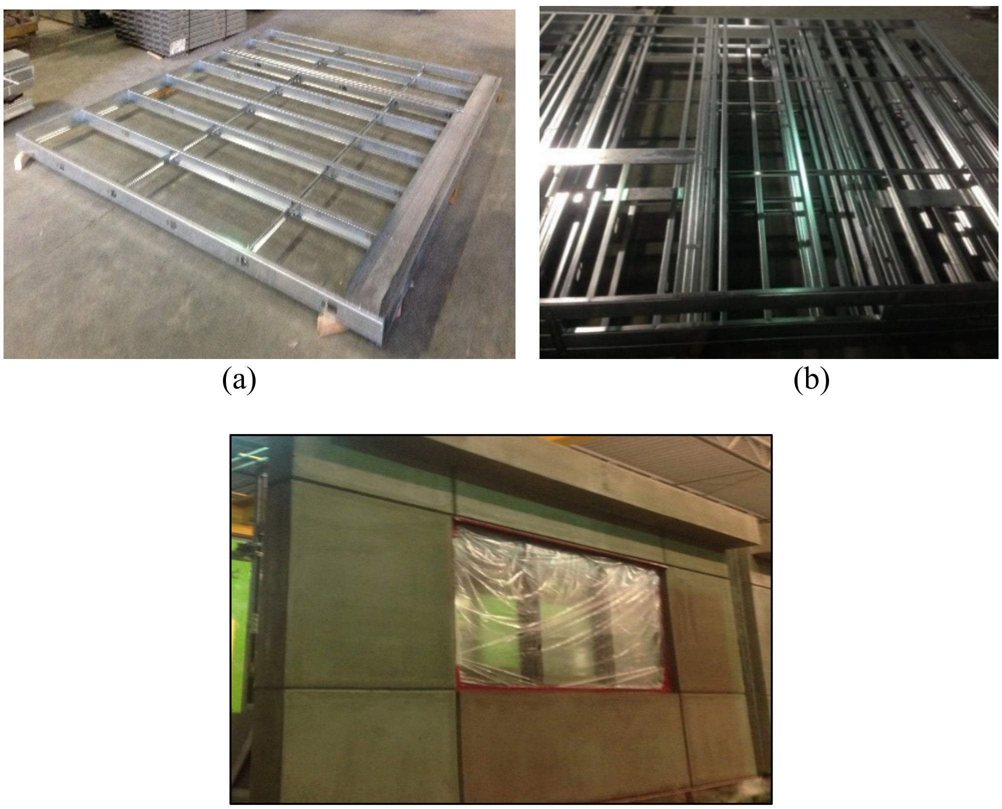


Figure 1.1: Market share of various projects for the modular industry (MBI 2023)

Previous studies (Nick et al. 2019; Zhang 2020) indicate that modular construction manufacturing (MCM) adheres to standardized operating procedures (SOPs) at each workstation in order to prefabricate the module components. These SOPs assists production managers in standardizing the tasks and efficiently implementing innovative technologies such as automation and robotics, thereby reducing the time and cost of produced modules. Modular construction, also, offers opportunities to: *(i)* improve productivity by allowing tasks to move towards the workers at the workstations via conveyor belt rather than the traditional method of workers moving towards tasks, the common practice in conventional construction; *(ii)* minimizing material waste and actively supporting the initiatives of circular economy; and *(iii)* achieve faster return on investment, as the modular construction process enables developers to open building doors sooner for the residents, thus starting to generate profits earlier (Garusinghe et al. 2023; Nik-Bakht et al. 2021). Due to these features, the MBI reports that gross revenues increase from $ 3.3 billion in 2016 to $ 3.97 billion in 2017 (MBI 2018), and the modular construction industry generated $12 billion in North America in 2022, accounting for 6.03 percent of all new construction projects. However, a certain level of customization (i.e., different design specifications) is required to build these homes in accordance with owners demands. Customized module components when prefabricated at the same production line lead to varying production rates, imbalanced production lines and increased makespan (i.e., total completion time).

1.2 Research Motivation and Problem Statement

Despite the advantages highlighted above, design customization of module components causes significant challenges in modular construction manufacturing (MCM). Such customization calls for variation in the design specifications of module components (e.g., different number of studs, cripples, doors and windows). Figure 1.2 illustrates the different types of module components such as interior and exterior wall panels, which vary in sizes and design specifications. Due to these variations, it is difficult for production line managers to accurately forecast the process times (Mohsen et al. 2022; Altaf et al. 2018). Process time (i.e., time taken to complete one module component at each workstation), being a critical component for measuring performance in MCM, requires accurate forecasting in order to: *(i)* controlling hourly/daily production line operations by effectively managing the workloads; *(ii)* gaining insights into underlying patterns of the production line; and *(iii)* making data-driven decisions with respect to resource planning (e.g., labors) and scheduling sequences of module components in the production line. In practice, production managers often rely on the average process times and linear fixed rate (sq. ft. per minute) to estimate the process time of module components at workstations. This leads to inaccurate prediction of process times, given that process time at workstations largely depends on the type of module components (interior and exterior wall panels) and its design specifications, for example, number of doors, windows and studs (Altaf et al. 2018). Inaccurate forecasting of process times combines with the dynamic nature of MCM that requires frequent hourly/daily changes, makes it challenging to develop a robust method for planning, and scheduling the sequences of module components and resource allocation in MCM. Additionally, it is challenging to optimize the sequences of module component and balance the production line with reasonable model runtime, considering this sequencing problem as a non-deterministic polynomial-time hard (NP-hard) problem. This results in an imbalanced production line, causing module components to experience waiting between workstations and workers to face idle time at workstations, ultimately extending the total project completion time.



(a) (b) (c)

Figure 1.2: Different types of wall panels: (a) Interior wall panel; (b) Interior wall panel with door; and (c) Exterior wall panel with window

To reduce subjectivity, various researchers have used probabilistic and statistical methods to predict duration and productivity in construction management (Altaf et al. 2018; AbouRizk and Halpin 1992; Lu and AbouRizk. 2000, Lee 2005; Barkokebas et al. 2018). However, probability distributions can lead to misleading results as they do not account for variations arising from clients' customization requirements when predicting production durations. Effective planning and scheduling are crucial for completing projects on time and within budget (Salama et al. 2021). Traditional methods, such as the critical path method (CPM), commonly used for project planning and scheduling has been proven to be ineffective due to: *(i)* time-consuming and error prone when rescheduling and re-optimization of sequences of panels at workstations due to dynamic factors such as change orders, weather conditions, and on-site demands (Lee and Hyun 2019); *(ii)* lack of crew work continuity, which helps to prevent idle time during repetitive tasks by planning the advancing from one task to another in accordance with the task demands and not considering crew work continuity leads to reduced productivity (Hegazy and Kamarah 2022); *(iii)* inability to make the real-time adjustments in the production planning and scheduling, which is essential for a dynamic production line environment; and *(iv)* being challenged to effectively model uncertainties arising from variations in process times due to customization requirements from clients and minimize makespan when scheduling production line operations (Wei et al. 2024). While the linear scheduling method (LSM) can adequately address these limitations in the scheduling of repetitive tasks (Salama et al. 2018), it does not: *(i)* account for decision variables such as module dimensions and wall openings in their sequencing and labour allocation, which significantly affects the cycle time of the production line; and *(ii)* consider multiple scenarios of resources and sequences of modules which is critical for planning effectively and improving production performance of MCM lines. To address the challenges associated with LSM and CPM, various studies have developed planning methods using job sequencing rules (Shafai 2012). However, the later method accounts only for a limited number of job sequencing rules for prefabricating module components, thereby falling short of providing an optimal sequence of the module components to be prefabricated in the production line. In terms of optimization algorithms, no prior research has employed a hybrid optimization method and compared it with other algorithms to determine the most suitable algorithm for scheduling modular construction projects.

1.3 Research Scope and Objectives

This research focuses on the planning and scheduling method for modular construction manufacturing (i.e., off-site) as illustrated in Figure 1.3. To tackle the previously stated challenges, this research aims to:

Minimize the total completion time (i.e., makespan) of prefabricating in the production line based on the sequences of module components and allocation of workers at workstations.

This is achieved through the following objectives:

1) Development of predictive method for forecasting process times of module components at each workstation.

2) Investigate a simulation-based method facilitating the generation of multiple scenarios based on sequences of module components and allocation of workers at workstations.

3) Optimize the sequences of the module components at workstations by experimenting with the utilization of Genetic Algorithm (GA), Simulated Annealing (SA) and hybrid Genetic Algorithm Simulated Annealing (HGASA) in order to develop an optimal production line plan and schedule.

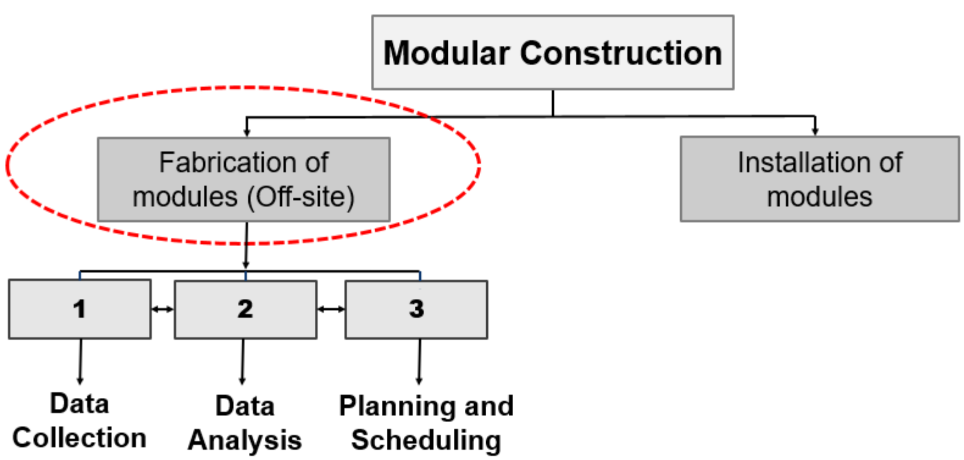


Figure 1.3: Research Scope

1.4 Thesis Organization

This thesis is organized in five chapters as shown in Figure 1.4. Chapter 2 provides an overview of the literature, focusing on: *(i)* data analytics including statistical methods, multiple linear regression and artificial neural network; *(ii)* planning in modular construction using lean manufacturing; *(iii)* discrete event simulation for planning in construction industry and *(iv)* production line scheduling. Chapter 3 is the main chapter and it describes the framework of the developed method, which is composed of three modules: *(i)* a simulation-based statistical method for planning in modular construction production lines; *(ii)* a deep neural network (DNN)-based method for predicting process times of module components at each workstation; and *(iii)* an optimization method for scheduling in modular construction. Chapter 4 demonstrates the use of the developed methods in two case studies, along with a discussion on their performance. Chapter 5 concludes with a summary of the thesis, the academic and industrial contributions of the work. It also highlights the research limitations and opportunities for future works.

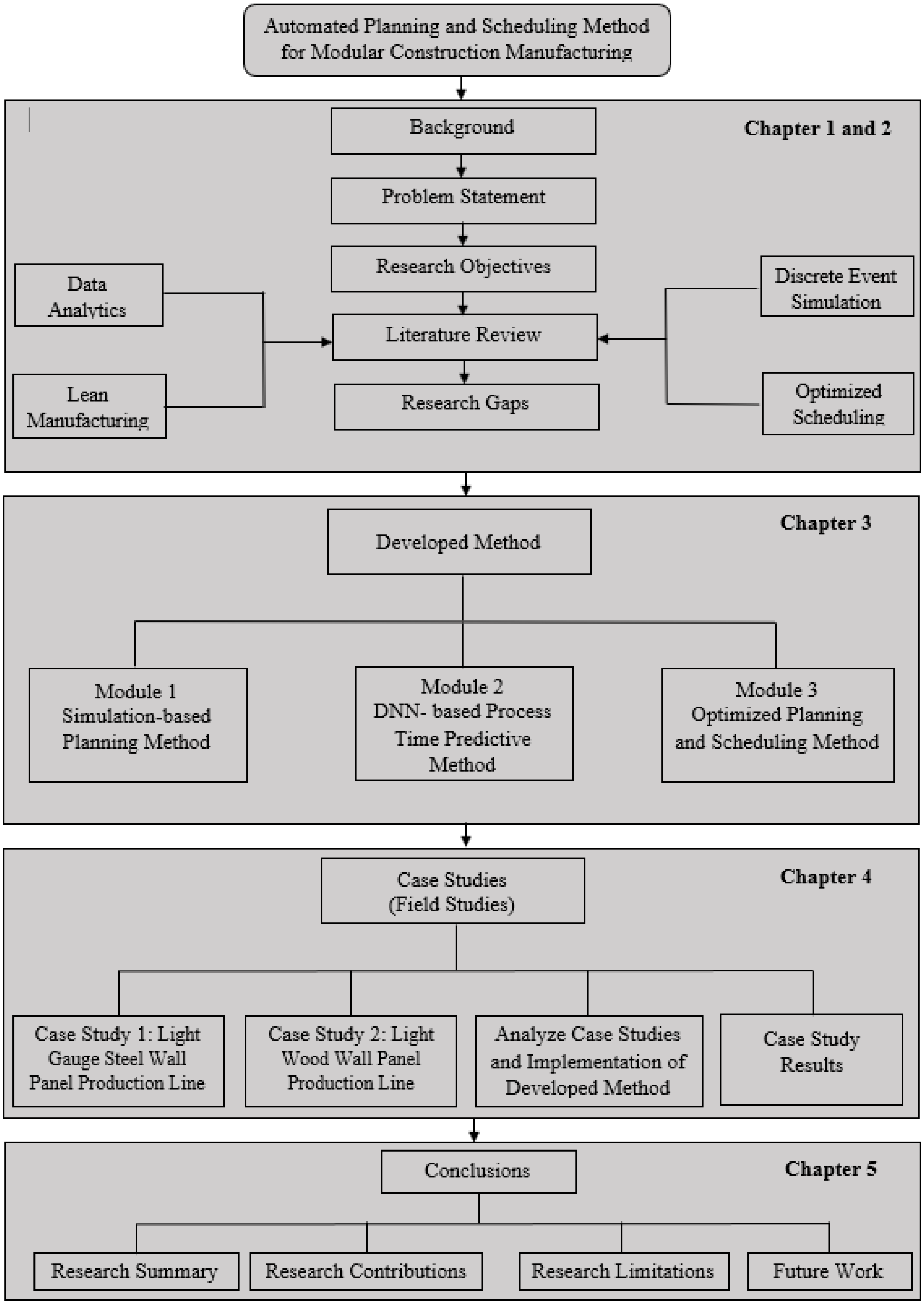


Figure 1.4: General overview of research framework

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

Literature review is conducted to support the understanding of the current practices, existing studies and research opportunities in the field of planning and scheduling for modular construction manufacturing (MCM). Figure 2.1 presents the overview of the literature review chapter that focuses on the following research areas: (i) data analytics; (ii) planning in modular construction using lean manufacturing; (iii) application of simulation in construction; and (iv) production line scheduling.

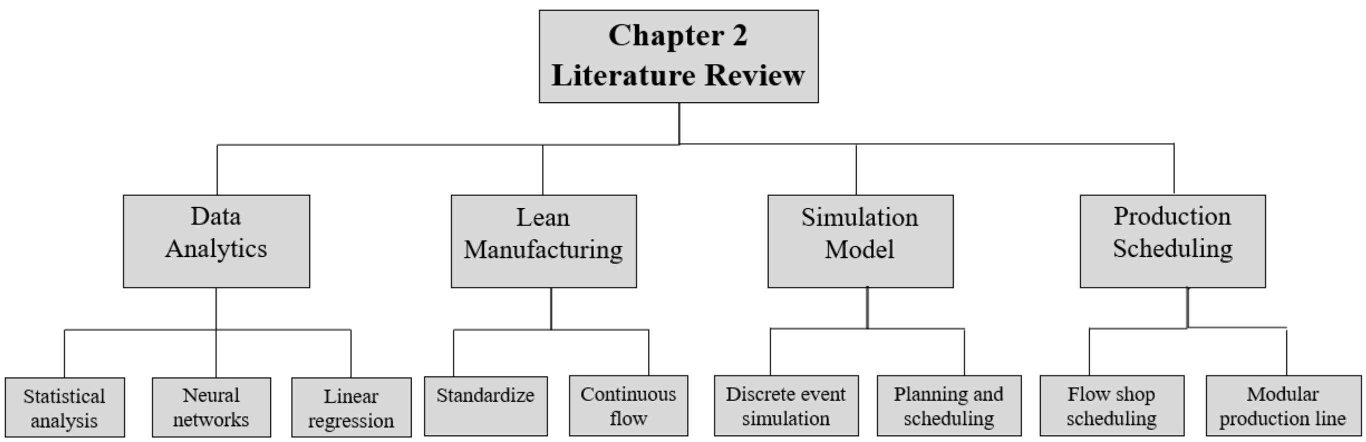


Figure 2.1: Literature review overview

2.2 Application of Data Analytics

The modular construction production line must run at its full capacity in order to elevate productivity to its maximum limit and meet the on-site demands successfully. However, since the production line receives orders of multiple projects, which are highly customized, this results in an uneven distribution of historical data related to the process time of module components across the workstations and creates uncertainty in predicting the process times (Mohsen et al. 2022). As a highly customized product, which increases the complexity of predicting process times, an accurate prediction model is essential and becomes a primary concern for the production line managers in MCM. In today's manufacturing environment, it is important for the production line managers to provide accurate prediction of processing times by methods that has less computational time in order to make real-time changes in the production (Taiwo et al. 2022). Accurate prediction of module component process time at workstations will assist production managers with resource allocation and reduction in makespan. According to a research report by McKinsey Global Institute (Manyika et al. 2024), big data and data analytics can help in managing manufacturing production lines by developing robust quantitative decision support solutions. In this respect, knowledge discovery (predictive modelling), which is defined as a process of identifying novel and potentially useful patterns from the data set in order to make it more understandable, has been applied by various researchers (Cheng et al. 2018). Moreover, in modular construction production lines, the amount of data plays a significant role in order to efficiently train the prediction model. For instance, if the collected data only includes the working period of the experienced worker, the prediction models are likely to perform differently when the work at that workstation will be performed by the new worker. Additionally, if the collected data period is for a project where design factors are not complicated for the module components as compared to a project with complex design factors, then also the prediction model is likely to perform differently (Mohsen et al. 2023). It should be noted that there are number of design factors (e.g., number of studs, number of doors, number of windows and net area) that affects the process times of workstations at the production line according to the type of module component (i.e., interior or exterior wall panel) as shown in Figure 2.2. In this respect, accurate prediction of workstation process times based on a module's design specifications is essential for production planning in MCM. Therefore, identifying the key features affecting the process times and developing predictive method that use historical data is vital to support intelligent decision making in production planning that promote productivity improvement. Feature selection methods in data analysis are used to select important features from available subset of variables in order to develop efficient and accurate predictive models (Guyon and Elisseeff, 2003). According to Mohsenijam and Lu (2019), a predictive model with key input features significantly reduces the collinearity between input variables and overfitting issues. Various feature selection methods, such as correlation matrix, principal component analysis, and t-test, have been used to identify key features (e.g., work durations, worker's skill, profit margin, and module type) in order to improve understanding of the underlying processes and overall performance of projects (Xu et al., 2016; Xie et al., 2018). For instance, Chanmeka et al. (2012) carried out a correlation analysis and statistical test of significance to determine critical factors related to the performance of oil & gas projects in Alberta, Canada. Mohsenijam et al. (2017) used stepwise regression to identify key design features such as rebar, bolts, nuts, etc. The project's labor hours were predicted using a regression model created using these smaller subsets of design features. Despite the appropriate use of feature selection methods and regression-type predictive models in the aforementioned studies, though, they cannot be effectively applied to MCM because of its process-oriented nature. In the MCM approach, the entire production process is divided into sequences of smaller repetitive processes, such that the productivity of the manufacturing production line can be improved. As such, in the case of MCM, an analytical framework is required in order to identify significant module design factors and establish a method to accurately predict the workstation process times. In consideration of these characteristics, the research presented in this thesis seeks to identify which feature selection methods are most effective for production planning in MCM.

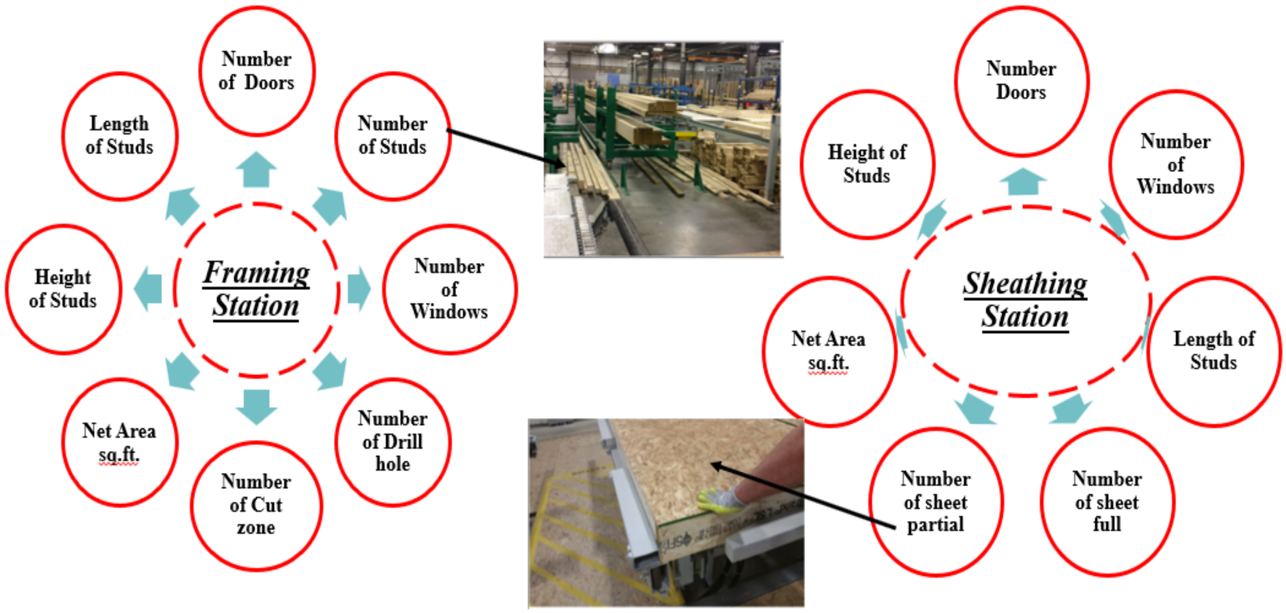


Figure 2.2: List of design factors (pictures provided by Dr. Sadiq Altaf)

Multiple linear regression (MLR) have been implemented to develop data-driven predictions metrices in various construction and fabrication projects. Multiple linear regression analysis is a statistical method, and this approach relies on input variables to predict the output variable as a way of gaining insights into underlying patterns. In construction management, MLR has been applied from predicting duration for earthmoving productivity (Smith 1999) to forecasting cycle time for one-span installation in precast bridges (Mohsenijam et al. 2017). Additionally, there are various examples of applying multiple linear regression in production lines. For instance, Akpinar et al. (2020) applied an MLR in order to describe the relationship between operations' cycle time and productivity in the case of a vehicle production line. Bhatia et. al. (2022) applied MLR to predict the processing time of wall panels at workstations of the production line. The predicted model achieves an R square of more than 70% with a dataset of 200 data points. However, the multiple linear regression works with the underlying assumption that there is a linear relationship between the independent variables and the dependent variable (e.g., process time) and does not perform well when the relationship is non-linear (Mohsenijam et al. 2017). Additionally, with the introduction of more complex manufacturing systems (e.g., modular construction manufacturing), the limitations of linear regression came to attention. For instance, in the case of aircraft assembly lines, which is a non-linear assembly system, regression has not provided a satisfactory result for predicting productivity (Mattsson et al. 2017).

Therefore, the application of artificial neural networks (ANN) to enhance the accuracy of predicting models has been used by various researchers. In particular, ANN are utilized as a method in the field of engineering operations and management in order to develop predictive models and help managers to enhance their decision-making process (Moon et al. 2023). Moreover, ANN: (i) provides superior performance for highly uncertain, nonlinear, and complicated problems; and (ii) frequently used in prediction for nonlinear problems in manufacturing and supply chain problems (Moon et al. 2023; Ambrogio and Gagliardi 2013). According to Moselhi et al. (1991) ANN is the most common supervised learning method to analyze the relation between input and output. The model encompasses a collection of processing elements, usually organized into layers (i.e., input, hidden, and output layers) and the target is to predict one/more dependent variable(s) from independent variables. The input layer accepts the data (i.e., independent variables), which is used by the hidden layers to represent the relationship, and the output layer produces the network response (i.e., dependent variable). In this study, ANN is defined as a neural network that has one hidden layer, whereas DNN (deep neural network) is characterized by having two or more hidden layers (Aggarwal 2018). In this thesis, 'deep' refers to the presence of multiple hidden layers, enabling the network to learn complex representations from the input data. ANN has been extensively applied in the construction and manufacturing industry. For example, Ambrogio and Gagliardi (2013) applied ANN to predict the performance of the production line process in manufacturing and improved prediction accuracy. Basma and Moselhi

(2022) implemented ANN in order to predict the duration and cost of highway projects, with Mean Absolute Percentage Error (MAPE) of 7.4% and 4.5% for the duration and cost. El-Sawah and Moselhi (2014) applied back propagation neural network, probabilistic neural network, generalized regression network, and regression analysis in order to estimate the cost at the pre-design stage for structural steel building and short-span timber bridges. The results of the different neural network models were compared, and it showed that the probabilistic neural network outperformed the regression method with a mean absolute percentage error of 1.91 %. Moselhi and Siqueira (1998) designed an ANN in order to estimate the direct costs for low-rise structural steel buildings. The model was developed using NeuraShell 2, a commercial software, and its performance outperformed that of regression method. Moon et al. (2023) successfully implemented a multilayer perceptron ANN to predict production and latency days for manufacturing production facilities. However, some studies, for example, Mohsen et al. (2022) and Alsakka et al. (2023) have advanced the development of forecasting process times at workstations along the production line in modular construction by applying various machine learning models, including ANN. However, the prediction model was not developed for all of the workstations, manual tuning and GridSearchCV were used to find the optimal parameters for the neural networks. This approach leads to subjectivity in the process of tuning the parameters and does not ensure optimal parameters (Callens 2020).

Various researchers have utilized GA to optimize hyperparameters for machine learning algorithms in the fields of construction and infrastructure engineering. For instance, Assad and Bouferguene (2022) used the GA in order to improve the accuracy of predicting the water mains condition by finding the optimal hyperparameters of various data mining techniques (e.g., deep neural network). However, this optimization approach omitted key hyperparameters such as learning rate and momentum, which inclusion could enhance convergence efficiency and reduce overall training time. Koc et al. (2021) applied tree based (e.g., Random Forest) machine learning models to predict the post-accident disability status of workers in the construction industry. The machine learning parameters were tuned using GA in order to improve the prediction accuracy. Considering the successful application of GA in tuning the parameters of machine learning algorithms this thesis implements GA to select the optimal neural network hyperparameters for predicting the process times of module components in the production line.

In summary, as shown in Table 2.1, the existing methods for predicting process times of module components in MCM involves following limitations: (i) manual tuning was employed to determine the optimal hyperparameters for the neural networks, which introduces subjectivity into the hyperparameter tuning process and does not guarantee optimal hyperparameters; (ii) predictive method have not been created for all workstations along the production line; (iii) the dataset used to train the model was small (i.e., 200), which may not yield better prediction performance for projects with complex designs and larger datasets; and (iv) probability distributions can produce misleading results as they do not account for design variations stemming from clients' customization requirements, which are unique to module components when predicting process time.

Table 2.1: Research gaps for data analytics

Number	Author and Year	Method	Gaps
1	Alsakka et al. (2023)	Used computer vision-based data to develop cycle time prediction framework for one of the workstations of modular construction production line. Various machine learning models, including ANN were trained to estimate the cycle time.	Prediction method was not created for all workstations along the production line. Manual tuning was employed to determine the optimal parameters for the neural networks. This method introduces subjectivity into the parameter tuning process

			and does not guarantee optimal neural network configuration. The dataset used to train the prediction models covers only a shorter period of operation. This may not yield better prediction performance for projects with complex designs and larger datasets.
2	Mohsen et al. (2022)	Predicted the cycle time of wall panels in a production line using ML models and utilizing RFID based historical data.	Predictive method was not developed for every workstation along the production line. Manual tuning was used to identify the parameters for the neural networks. This approach introduces subjectivity into the tuning process and does not ensure optimal parameters.
3	Bhatia et al. (2022)	Applied MLR models to predict the processing time of wall panels at each workstation of the production line	The dataset used to train the model was small (i.e., 200), which may not yield better prediction performance for projects with complex designs and larger datasets.
4	Altaf et al. (2018)	Predicted wall panel process time using probability distribution functions (e.g., beta, triangular and gamma distributions)	Probability distributions can yield misleading results as they do not consider design variations arising from clients' customization requirements, which are unique to module components when predicting production durations.
5	Taiwo et al. (2022)	Developed multiple linear regression model to predict the productivity of installing modules in modular integrated construction projects	The regression results were not compared with other machine learning algorithms. The dataset used to train the prediction models covers only a shorter period of operation. This may not yield better prediction performance for projects with

			complex designs and larger datasets.
6	Rashid et al. (2020)	Predicted duration at each workstation using triangular distribution with 15% upper and lower bound.	Probability distributions can yield misleading results as they do not consider design variations arising from clients' customization requirements, which are unique to module components when predicting production durations.
7	Moon et al. (2023)	Developed multilayer perceptron artificial neural network to predict the production days for the cable manufacturing production line.	Manual tuning was employed to determine the optimal parameters for the neural networks. This method introduces subjectivity into the parameter tuning process of neural networks and does not guarantee optimal parameters.
8	Barkokebas et. (2018)	Predicted task durations at workstations using the coefficient of variation (i.e., ratio of standard deviation to mean).	Relying on average times can yield to misleading results as they do not account for variations arising from clients' customization requirements (i.e., different design specifications).
9	Aghajamali et al. (2022)	Applied linear regression and artificial neural networks to predict cycle time of fitting and welding workstations at steel fabrication production line.	Manual tuning was employed to determine the optimal hyperparameters for the neural networks. This method introduces subjectivity into the parameter tuning process of neural networks and does not guarantee optimal hyperparameters. The dataset used to train the model was small, which may not yield better prediction performance for projects with complex designs and larger datasets.
10	Mohsen et al. (2023)	Developed a data-driven machine-learning model using Random forest and gradient	Manual tuning was employed to determine the optimal hyperparameters for the neural

		boosted decision trees to predict the total duration of prefabricating pipe spool	networks. This method introduces subjectivity into the parameter tuning process of neural networks and does not guarantee optimal hyperparameters.
11	Rashid and Louis (2020)	Developed a framework to identify different manual activities (e.g., nailing and hammering) performed in modular construction production line using audio signals and machine learning (e.g., SVM).	The method did not consider predicting the cycle time required to complete the module components at the workstations of the production line. The method did not identify which manual activities (e.g., nailing and hammering) have the most significant impact on the cycle time completion of module components.
12	Backus et al. (2006)	Implemented regression trees in order to predict factory cycle time.	The dataset used to train the model was small. Prediction algorithms like neural networks, which perform well with non-linear data, were not tested and compared.

2.3 Planning in Modular Construction Using Lean Manufacturing

The construction industry is moving towards industrialized construction (i.e., modular and off-site construction) by increasingly embracing the principles of mass production and standardization (Lee and Hyun 2019). Modular construction closely resembles the philosophy of the manufacturing industry, but encompassing both mass production and customization to meet specific customer requirements. Moreover, the objective of integrating manufacturing into construction is to introduce the concept of continuous flow, emphasizing the movement of module components through the workstations of the production line with waiting time minimized or eliminated (Gann 1996). In keeping with this concept, MCM production facilities often feature the simultaneous prefabrication of the module components for two to three projects in their production lines. In this regard, ensuring synchronization between the sequences of module components at each workstation and their on-site installation is critical. However, due to customization requirements from clients, the use of conventional construction methods for planning and scheduling (i.e., sequencing of jobs) in MCM poses challenges for production managers (as discussed above). Moreover, since the process times and release dates of the jobs in the production lines are uncertain at the planning stage, scheduling in MCM resembles a stochastic scheduling problem. In this respect, researchers have proposed methods based on lean manufacturing principles to enhance the process of planning in MCM (Zaalouk et al. 2023). The objective of applying lean manufacturing principles in construction is to reduce waste, achieve continuous improvement and increase value for the customer. While various lean principles and tools from manufacturing industry are applicable in the construction, but a number of lean principles and tools are developed specifically for construction industry. Koskela (1992) was the first to implement the concepts of lean production in the construction industry. The study introduced a new production philosophy, which served as a guideline to create continuous process flow in the production line. Later, last planner system (LPS) was introduced by Ballard (2000), which served as a system for project planning and control the process of production line. LPS supports the pull system and continuous improvement, which helps in creating smooth production line process. Given that, the implementation of lean principles has substantially enhanced the efforts in process improvement by identifying and eliminating non-value-added activities, it was implemented in the production line to improve planning in MCM (Yu et al. 2011). For example. Zhang et al. (2020) integrated the production line breakdown structure (PBS) with value stream mapping (VSM) to assess the status of the production line, identify the current issues and propose the solutions for future implementation. Salama et al. (2021) integrated BIM, LSM, and critical chain project management

(CCPM) in a novel scheduling method for modular construction projects, incorporating the concepts of Takt time and just-in-time (JIT) to increase productivity. Moghadam et al. (2012) integrated BIM and lean manufacturing principles in order to identify and eliminate bottlenecks (e.g., idle time of workers) in the production line and develop the production line schedules. Yu et al. (2009) used a value stream mapping (VSM) tool to determine steady production flow for productivity improvement by analyzing the production process and controlling fluctuation of sources. Yu et al. (2013) implemented lean production principles and techniques such as 5s and standardized work in the real case study. Future state maps were developed in order to implement these principles and improvement in terms of production line productivity was observed. However, based on these techniques, designing plans without validating and understating their effects on the production line can be costly and time-consuming. Also, lean based plans are required to be adjusted at regular basis in order to keep production line balanced. Considering the high amount of variation in design specification of modules and its effect on process times of workstations, manual adjustment is difficult and time consuming. Therefore, computer simulation can be employed as a validation tool for future planning by imitating production line processes as a way of assessing the effect of proposed solutions prior to incurring the cost and disruption of actual implementation (Han et al. 2012).

2.4 Application of Simulation in Modular Construction

The modular construction industry has complex processes, which consist of multiple module components and various production lines. In order to deal with these complex processes, a reliable technique and/or tool is required, which is capable of: (i) performing 'what if' analysis by generating multiple scenarios and assessing the effectiveness of these scenarios in terms of productivity improvement; (ii) generating multiple performance statistics; and (iii) efficiently model the interaction between processes and resources (Abourizk 2010). Due to advancements in the field of computer software, the above stated requirements are often achieved through simulation modelling. Abourizk et al. (2016) defined simulation as "*the use of computer software (e.g., Simphony) to represent the dynamic responses of a construction system by the behaviour of a model made to represent it. A simulation uses mathematical descriptions, graphical constructs, computer algorithms (as well as other means) that are generally encapsulated in a simulation software model to represent the real system*". The application of simulation is applied for decision making in the construction and manufacturing industry as it helps in understanding the process of complex systems. In this respect, previous researchers (Altaf et al. 2018; Azimi et al. 2011) have used the simulation technique in: (i) planning and scheduling the sequences of the modules produced in the production line based on the 'what if analysis'; and (ii) allocating resources at workstations. There are various types of simulation techniques depending on the nature of the dependent variables. As one of simulation techniques, discrete event simulation (DES), which is defined as the change of the dependent variables taking place at the specific event points. On the other hand, continuous simulation takes place when dependent variable changes continuously during the process being modeled (Marzouk 2003). To support the simulation, considerable efforts are made to introduce simulation applications, such as Cyclone, Stroboscope, Simphony.NET, AnyLogic and Arena. Cyclone as a foundation of various construction simulation applications (Halpin 1997) is deployed to model and analyse the construction process. CYCLONE, which stands for CYCLic Operations Network, is introduced by Halpin (1977) as a construction simulation tool. It follows the concept that construction operations can be considered in the form of cyclic networks of modelling elements that characterize the change of construction resources between an active state (productive) and an idle state (waiting). Later, special-purpose simulation is proposed by Hajjar and Abourizk (1999) for the experts in the area of construction simulation. Abourizk and Mohamed (2000) introduced a powerful and user-friendly simulation tool called Simphony.NET as a discrete event simulation (DES) system that evaluates various scenarios of the production line operation before implementing in the real production. While developing models using Simphony. NET, users have access to a domain-specific set of modelling elements library with names that users can relate. Figure 2.4 shows the various modelling elements of Simphony. NET that allows the users to mimic the system using process interaction concepts. For example, icon for ‘create’ element represents the entities (e.g., wall panels), icon for ‘task’ element represents an activity (e.g., assembly workstation) in the model, icon for ‘resource’ element allows representation of various real-world resources (e.g., crane) and icon for ‘condition’ element routes entities by evaluating true or false condition (e.g., if wall panel is exterior, then condition is true and the wall panel will move to framing workstation, if false then if will go to sheathing workstation).

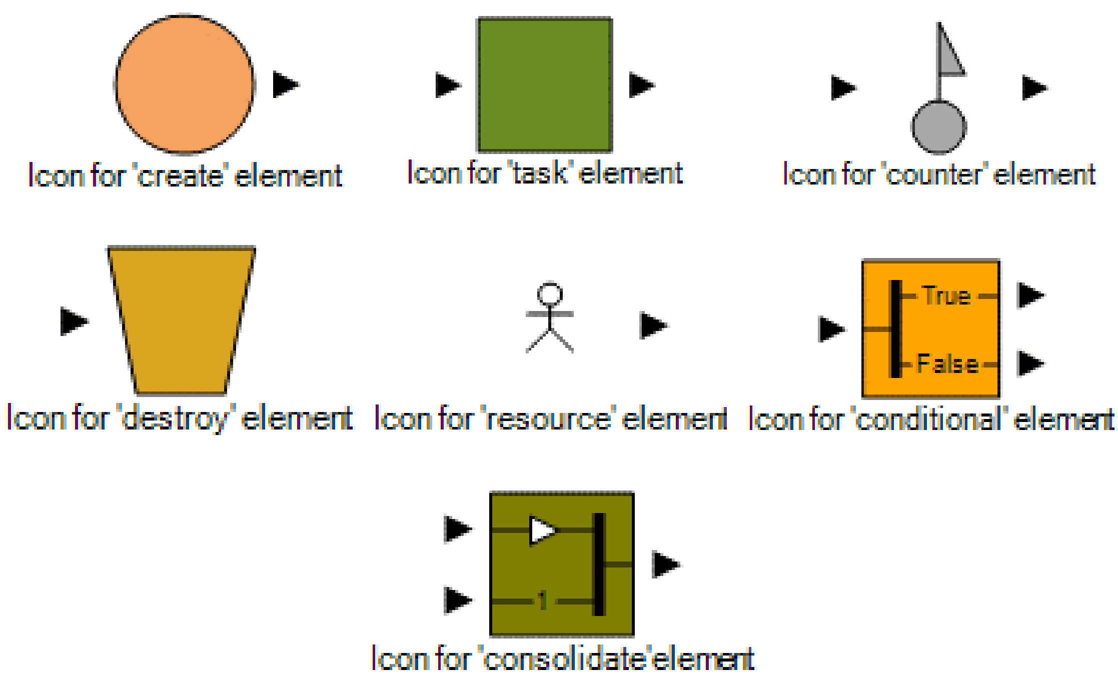


Figure 2.3: Modeling elements in Simphony. NET (Abourizk et al. 2016)

The outputs of the simulation model include statistical averages, utilization of resources and time graph. All these simulation systems in general focus on process-oriented concepts to reflect the characteristics of construction process and its working environment. Previous research (Hong et al. 2011) have contributed in advancing the production line planning in MCM using simulation platforms. For instance, Senghore et al. (2004) developed a simulation model to improve the production process of manufactured housing plants. Altaf et al. (2018) utilized RFID technology, data analysis and simulation-based optimization for planning and controlling the production line. Barkokebas et al. (2021) applied BIM with simulation and lean principles for planning in off-site construction and suggested production line process improvements. Bhatia et al. (2022) and Mohsen et al. (2008) developed a simulation model using Simphony.NET to predict the productivity of the production line. Azimi et al. (2011) integrated data acquisition and simulation for developing a decision support tool for production managers to take corrective actions ahead of time. Lee and Kim (2017) used BIM based 4D simulation method in order to improve the process, material, and quality management for manufacturing modules in the factory. Wei et. al. (2024) developed a pull-based hybrid simulation planning method for the modular and off-site construction supply chain. In general, simulation have been applied individually or integrated with lean and BIM to develop plans and schedules, thereby improving the performance of MCM. However, there are the following issues: *(i)* a limited number of scenarios were tested; therefore, not giving optimal sequences of module components (e.g., wall panels) to be prefabricated in the production line; *(ii)* these methods do not provide a systematic way to identify significant impact factors affecting process times of workstations, which can improve the accuracy of predictive results in a simulation model; and *(iii)* these methods do not simultaneously schedule the sequences of modules and allocation of resources for further productivity improvement. The simulation models are developed based on data collection which is generally implemented by manual tasks leading to error-prone and time-consuming. To address these limitations, some researches introduced radio-frequency identification (RFID) based data collection (Altaf et al. 2018; Azimi et al. 2011). Figure 2.4 illustrates the typical RFID system consists of an RFID tags (passive), stationary readers and database in order to store the data.

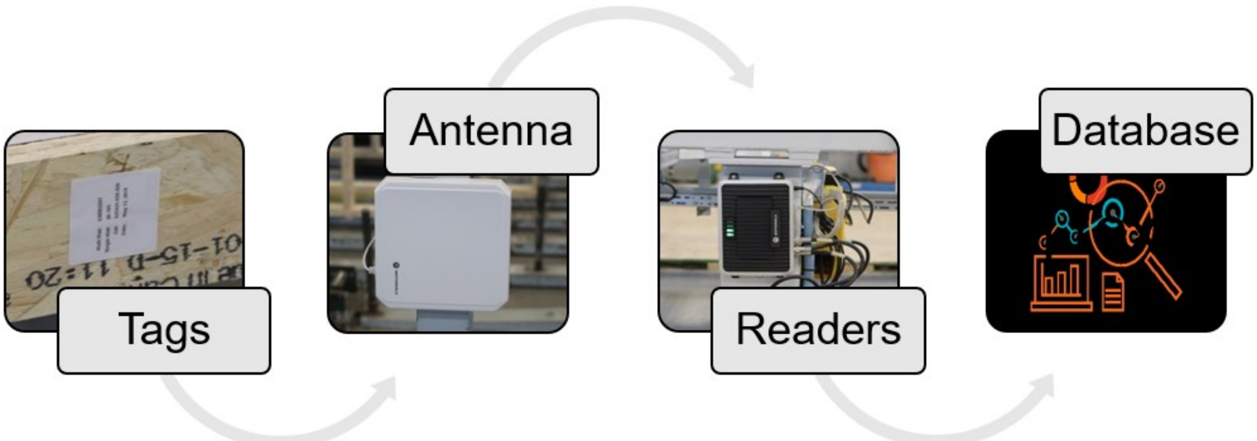


Figure 2.4: RFID system description (pictures provided by Dr. Sadiq Altaf)

2.5 Production Line Scheduling

Production line scheduling involves organizing a given number of predefined operations (i.e., jobs) in a specific sequence and time frame allocated to particular machines (i.e., workstations) within the production line (Varela et al. 2022). The most common production line scheduling problem is flow-shop scheduling (permutation flow shop model), where the machines (i.e., workstations) are arranged in series, and each job (i.e., module component) must pass through all the workstations in the predefined sequence (Framinan et al. 2014). In this respect, a modular construction production line, resembling the flow shop model and necessitating optimal sequencing of module components at workstations, is defined as a production line scheduling problem. This scheduling problem belongs to the class of non-deterministic polynomial-time hard (NP-hard) problem. In such problems, there are multiple sequences (n^m) in which the modules can be scheduled to pass through the different workstations, and, with an increase in the number of combinations, the complexity of the problem (i.e., the search space) increases, making it increasingly difficult for existing techniques to effectively find the optimal sequence within a reasonable model runtime (Baker and Trietsch 2013). For example, Alsakka et al. (2024) integrated computer vision, machine learning-based prediction models, and simulation in order to estimate cycle times, simulate operations, generates schedules for the wall panel production line. Barkokebas et al. (2023) integrated statistical and machine learning with the digitalization techniques in order to balance the production line and hence make process improvements. However, these method does not apply optimization algorithms in order to find optimal production sequences which can minimize the production duration.

As such, in order to find the optimal solution within a reasonable computational time, researchers in industrial engineering have applied metaheuristic optimization techniques—e.g., GA, particle swarm optimization (PSO), and SA—to address scheduling problems (e.g., flow-shop scheduling). For example, Chen et al. (2019) proposed a GA-based method to generate schedules for a hybrid flow shop, dynamically considering order arrivals. Their study demonstrated the effectiveness of GA in reducing job waiting time and meeting order deadlines. Meanwhile, An et al. (2014) illustrated the application of GA in minimizing the production time and cost for a metal-cutting production process. Ji et al. (2009) applied SA to find the optimal sequencing of jobs in terms of minimizing total duration in the flow shop. Previous studies have shown that optimization is a promising approach in developing schedules and minimizing the duration in the manufacturing industry's flow shop production lines. Modular construction production line problems, which resemble the flow-shop scheduling problem, can be solved using optimization techniques. In this respect, as shown in Table 2.2 various researchers have developed optimized schedules for modular and off-site construction. For example, Altaf et al. (2018) developed a simulation-based optimization model using PSO/SA to minimize the production line's cycle time and determine the optimal sequences of wall panels in the production line. Rahman and Han (2024) integrated linear scheduling method and multi objective optimization model to balance the off-site construction production line and minimizes project completion time, work in progress and workstation idle time. Hyun et al. (2021) developed a multiobjective optimization model for modular unit production line using a genetic algorithm. Various performance indicators such as minimization of cost, makespan, resource consumption and idle time, are used in multiobjective optimization. Lee and Hyun (2019) developed a multi-objective optimization model based on GA to reduce both duration and cost in modular construction production lines. Rashid and Louis (2020) integrated GA and discrete-event simulation to minimize the makespan (i.e., completion time) for a modular construction production line by optimizing the allocation of workers to the various workstations. Liu et al. (2015) developed optimized schedules for panelized construction projects by integrating BIM, discrete-event simulation, and PSO. However, these methods have the following limitations: (i) they do not account for the sequence of on-site installation of module components; (ii) the optimization does not account for multiple projects being carried out simultaneously; and (iii) they assume that the durations of workstations on a production line follow a triangular distribution, meaning that they do not consider unique design factors in predicting duration. In this respect, there is a need for scheduling methods that utilize historical production data to develop predictive models for scheduling the sequences of module components in the production line.

Hybrid optimization, which can solve complex optimization problems, has been employed in a number of research studies to obtain optimal and efficient solutions. The most prevalent approach has been to combine the strengths of two metaheuristic algorithms (e.g., GA and SA) and thereby overcome the shortcomings of each. For example, GAs converge prematurely and slowly, lacking local search capability, whereas SA excels at avoiding local optima by searching around the initial solution (Hassani et al., 2021). As such, combining these algorithms results in a more efficient hybrid approach, especially for solving large-scale scheduling problems (Mehdi 2011). Various researchers have applied this hybrid optimization approach to solve job/flow-shop scheduling problems, to reduce production costs for turning operations in metal, and to solve aggregate production planning problems. For instance, Hassani et al. (2021) implemented a hybrid SA/GA approach to minimize cost in a flow-shop context. Rameriz et al. (2019) developed a novel hybrid algorithm combining the Ant Colony System (ACS), SA, and GA to minimize the makespan in a flow-shop scheduling environment. Ganesh and Punniyamoorthy (2005) combined GA and SA to solve the production planning problem. In their study, GA was implemented to find a global solution while allowing SA to optimize each solution locally. Uslu et al. (2022) applied a hybrid algorithm combining GA and Ant colony optimization to minimize makespan in the context of flow-shop scheduling. Their results showed that hybrid optimization performed better than either GA and SA applied individually for small-scale problems.

The application of hybrid optimization (i.e., GA + SA) in sequencing module components (e.g., wall, floor, and roof panels) has yet to be explored. As illustrated through various studies, hybrid optimization algorithms have been effectively implemented in the manufacturing industry, yielding better results compared to those of individual metaheuristic algorithms. However, modular construction presents distinct challenges and differs from traditional manufacturing in terms of: (i) the high degree of customization of module components and (ii) the dynamic nature of production lines, necessitating frequent adjustments to production schedules to meet on-site demands. It is worth noting that, to the best of the authors' knowledge, no study has applied a hybrid optimization approach in MCM for production scheduling. Therefore, in the present study, a hybrid optimization technique is utilized to address the modular construction manufacturing scheduling problem (MCMSP). Although the effectiveness of optimization algorithms varies depending on the scheduling problem and objective, the positive reviews of GA and SA algorithms, known for providing effective solutions and good computing capabilities in solving production line scheduling problems, justifies their selection to solve MCMSP in this study.

Priority rules are a common approach for addressing the multi-project scheduling problem and are used in the manufacturing industry to determine the sequence of jobs on machines (ElFiky et al., 2020). These rules can be applied as constraints when developing production schedules for multiple projects to minimize total duration and tardiness (Kruger and Scholl, 2010). While numerous priority rules are available for multi-project scheduling in the literature, the challenge lies in selecting the rule that performs best under specific conditions. According to Kolisch (1994) and ElFiky et al. (2020), there are 17 priority rules categorized into activity-based, project-based, and resource-based groups: (i) the earliest due date rule, where projects with earlier due dates are given higher priority and processed first; (ii) the modified due date rule, where the project with the least modified due date is prioritized and processed next; and (iii) the shortest processing time rule, used to break ties among projects with the same due date, to minimize total tardiness. In this context, various studies have effectively implemented these priority rules for project scheduling. For example, Chen et al. (2019) applied 20 priority rules to address the multi-project scheduling problem, incorporating new project arrivals into their existing schedule to enhance practicality. Similarly, Pickardt et al. (2010) integrated genetic algorithms with discrete event simulation to generate dispatching rules, reducing tardiness for multiple projects in the semiconductor industry. Mizrak and Bayhan (2006) developed various dispatching rules to ensure projects were completed within their due dates, aiming to minimize assembly line lead time. Existing scheduling methods in modular construction manufacturing generally focus on single projects, which is unrealistic for actual factory settings where modular construction companies frequently undertake multiple projects at the same time. Therefore, it is crucial to develop a method that incorporates various priority rules along with optimization algorithms (i.e., the sequence in which wall panels of different projects should be prefabricated) to minimize project duration, and meet contractual obligations.

Table 2.2: Research gaps for production line scheduling

Number	**Author and Year**	**Method**	**Gaps**
1	Altaf et al. (2018)	Developed a simulation-based optimization model using PSO/SA to minimize the production line cycle time	• The optimization does not account for multiple projects being carried out simultaneously. • Not utilized a hybrid optimization approach and compared it with other algorithms
2	Rahman and Han (2024)	Integrated linear scheduling method and multi objective optimization model to balance the off-site construction production line and minimizes the completion time of wall panels	• Do not account for the sequence of on-site installation of module components • Not considered all the workstations of the production line
3	Alsakka et al. (2024)	Integrated computer vision, machine learning and simulation to estimate cycle times and generates schedules for the wall panel production line	• Does not apply optimization algorithms in order to find optimal production sequences • Not considered all the workstations of the production line
5	Hyun et al. (2021)	Developed a multiobjective optimization model for minimizing cost, makespan, and resource consumption in modular unit production line using a genetic algorithm	• Do not account for the sequence of on-site installation of module components • Not utilized a hybrid optimization approach and compared it with GA

6	Lee and Hyun (2019)	Developed a multi-objective optimization model based on GA to reduce both duration and cost in modular construction production lines	• The optimization does not account for multiple projects being carried out simultaneously • Does not consider practical case of parallel workstations aimed at reducing makespan in the optimization algorithms
7	Rashid and Louis (2020)	Integrated GA and discrete-event simulation to minimize the makespan for a modular construction production line by optimizing the allocation of workers	• Assume that the durations of workstations on a production line follow a triangular distribution, meaning that they do not consider unique design factors in predicting duration • Does not apply optimization algorithms in order to find optimal production sequences that highly affect the makespan

2.6 Summary of Research Gaps

In summary, the following limitations are identified:

1. The developed simulation lacks a systematic way to identify the SIFs affecting process times at workstations, which provides critical information that can improve the accuracy of the simulation results.
2. The analysis of the allocation of workers at workstations for the purpose of productivity improvement in MCM is not considered in the previous simulation-based planning methods.
3. Manual tuning was used to find the optimal parameters for the neural networks. This approach leads to subjectivity in the process of tuning the parameters and does not ensure optimal hyperparameters.
4. The predictive method to forecast the process time of module components were not developed for all the workstations in the production line.

5. Assumed the durations of workstations on a production line as only triangular distribution, which can lead to misleading results because it does not consider design factors in predicting duration, which are unique for module components.
6. Lean based plans are required to be adjusted at regular basis in order to keep production line balanced. Considering the high amount of variation in design specification of modules and its effect on process times of workstations, manual adjustment is difficult and time consuming.
7. There is a lack of practical methods for optimized production line scheduling that consider the installation sequences of module components on-site (e.g., interior wall panels should be prefabricated, delivered and installed before exterior wall panels)
8. The existing methods for planning and scheduling do not adequately consider the status (i.e., workload) of production lines, where multiple module components for multiple projects are managed and prefabricated simultaneously.
9. Previous study has not utilized a hybrid optimization approach (e.g., GA and SA) and compared it with other algorithms to identify the optimal algorithm that is best fitted for planning and scheduling of MCM.
10. No previous study has considered practical case of parallel workstations (i.e., a workstation capable of prefabricating two wall panels at the same time), aimed at reducing makespan in the optimization algorithms.

CHAPTER 3: DEVELOPED METHOD

3.1 General

The aim of this chapter is to introduce a framework composed of three modules regarding the developed method. The three modules include: (i) simulation-based statistical method for planning; (ii) deep neural network-based method for predicting the process times of module components at workstations; and (iii) optimized scheduling method based on metaheuristic algorithms (e.g., GA, SA and Hybrid GASA). Figure 3.1 depicts the details for automated planning and scheduling method for modular construction manufacturing (MCM) production line using historical and near real-time data. The process of simulation- based planning method in the first module, starts with work and time study in order to understand the standard operating procedures and collect near real time data of module components at each workstation in the production line. Based on this historical data, probability distribution functions and cycle time formula (CTF) using statistical techniques are developed to predict the process times, which are later input into the simulation method to plan the sequences of module components and allocation of workers at workstations. It should be noted that the simulation method helps to balance the production line by assigning the sequences of module components and allocation of workers. In the second module, a Deep Neural Network (DNN), based predictive method is developed. The predictive method is trained on large dataset (i.e., 416950 timestamps), which is collected using RFID based system to predict the process times of module components at each workstation of the production line. Artificial Neural Network (ANN) and multiple linear regression (MLR) are also utilized for predicting production process time spent at each workstation in a manufacturing plant. Also, genetic algorithm (GA) based optimization is used to optimize the architecture of the DNN and, as such, finds a near optimum number of hidden layers and nodes in each layer. In the third module, the production line schedules (i.e., optimal sequences of module components) are generated by utilizing genetic algorithm (GA), simulated annealing (SA) and Hybrid Genetic Algorithm Simulated Annealing (HGASA). These developed methods collectively provide the following outputs: (i) predictive method for MCM; (ii) lists the sequence of module components; (iii) allocation of workers at workstations; and (iv) minimum makespan. The details description of each developed method is provided in the following sections:

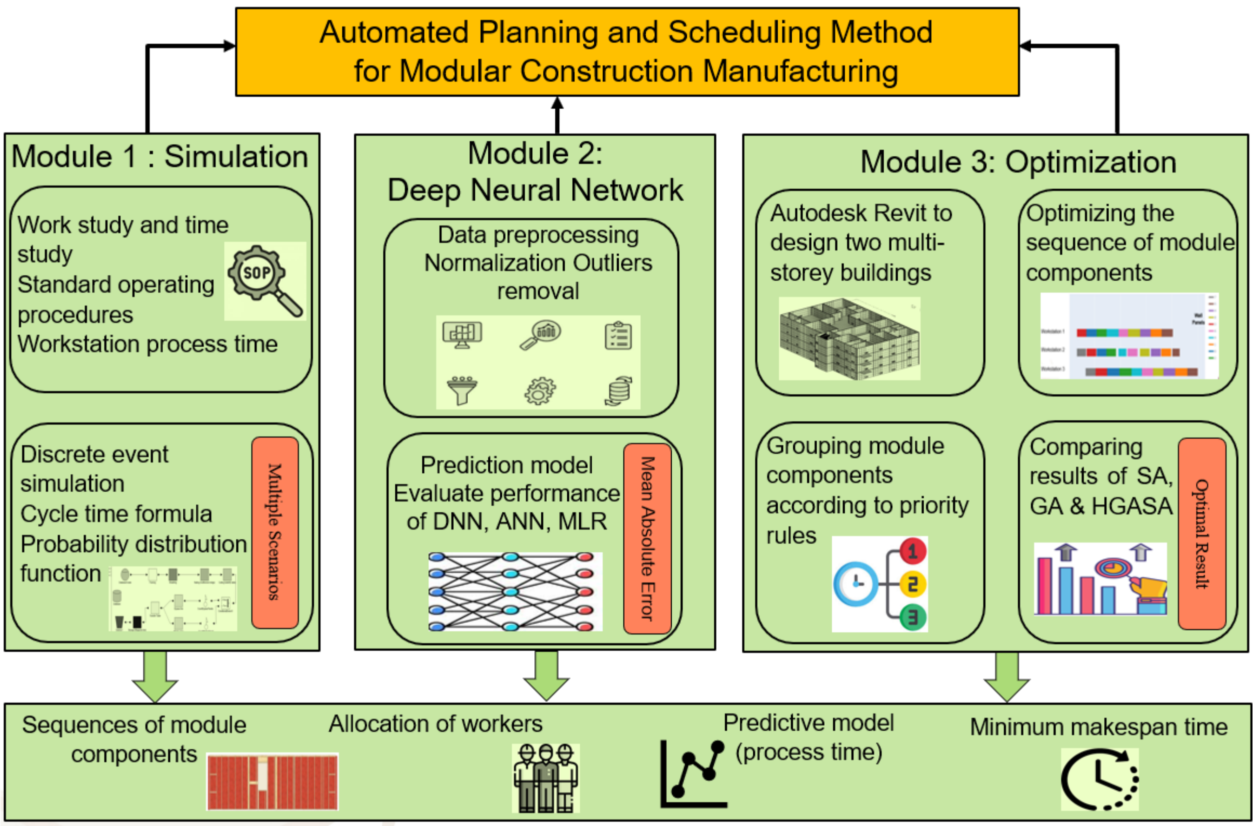


Figure 3.1: Main components of the developed method

3.2 Module 1: Simulation- Based Planning

Figure 3.2 depicts the components of the developed method and its simulation-based planning process for MCM lines using historical and near real-time data. The process encompasses three phases: (i) data collection; (ii) data analysis; and (iii) simulation-based planning. Input parameters such as process times of workstations, design specifications of modules, and the number of workers at workstations are housed in a central database. The criteria are workflow of the production line, availability of resources, working hours, and the capacity of workstations in terms of the module length they can accommodate (e.g., the capacity of a framing workstation may be a length of 20 ft). In the data collection phase, work and time studies are performed in order to gain understanding of the SOPs at workstations as well as collect historical and near real-time workstation production data. The data collected includes the start and finish times of modules, design specifications, and the number of workers assigned to various workstations. This data is stored in a database via a cloud-based time-track application called "C-track". The database is used in: (i) the data analysis phase to identify SIFs, develop the probability distribution functions and cycle time formula (CTF) using statistical method, and select the best predictive method by comparing the performance of the cycle time formula with that of the probability distribution functions; and (ii) the simulation of the production line developed in Simphony.NET. Both cycle time formula (CTF) and probability distribution functions are used in the simulation as a way of capturing the unique nature of production lines (e.g., facility layout and number of workstations). In this respect, the outputs (productivity) from the simulation method using cycle time formula (CTF) and probability distribution functions are compared with historical productivity data in order to determine which is most accurate. The output, it should be noted, lists the sequence of modules and allocation of workers at workstations, and can be used to determine the total duration for completing a module in the multiple scenarios generated by the simulation. The simulation method also helps to balance the production line by assigning an optimal sequence of modules and allocation of workers.

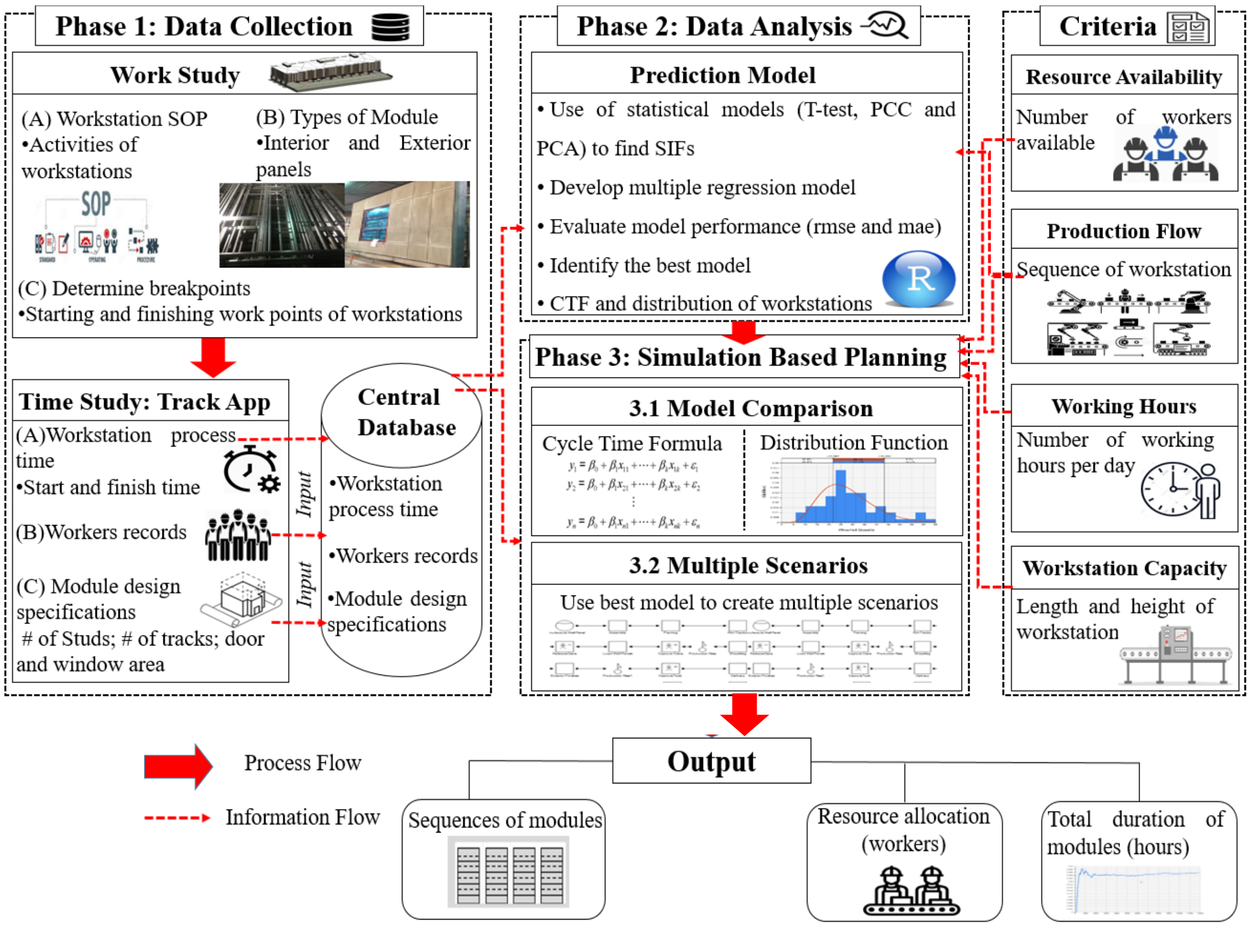
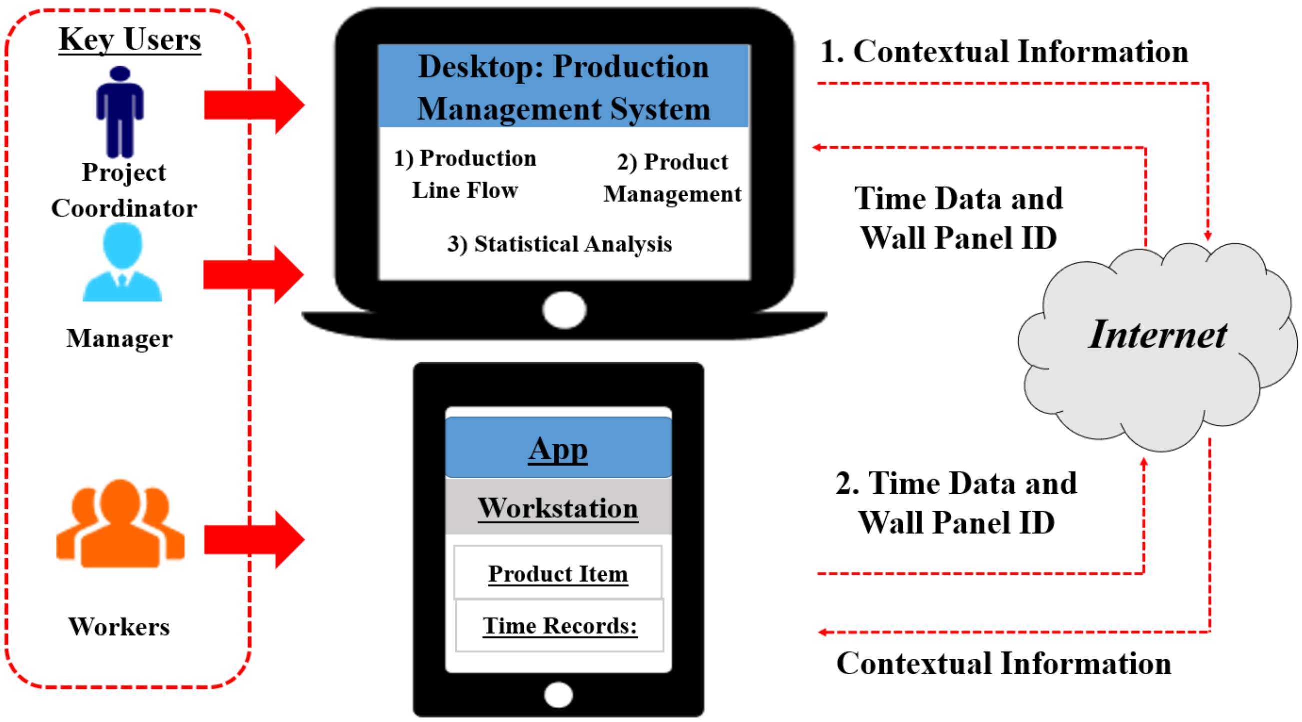


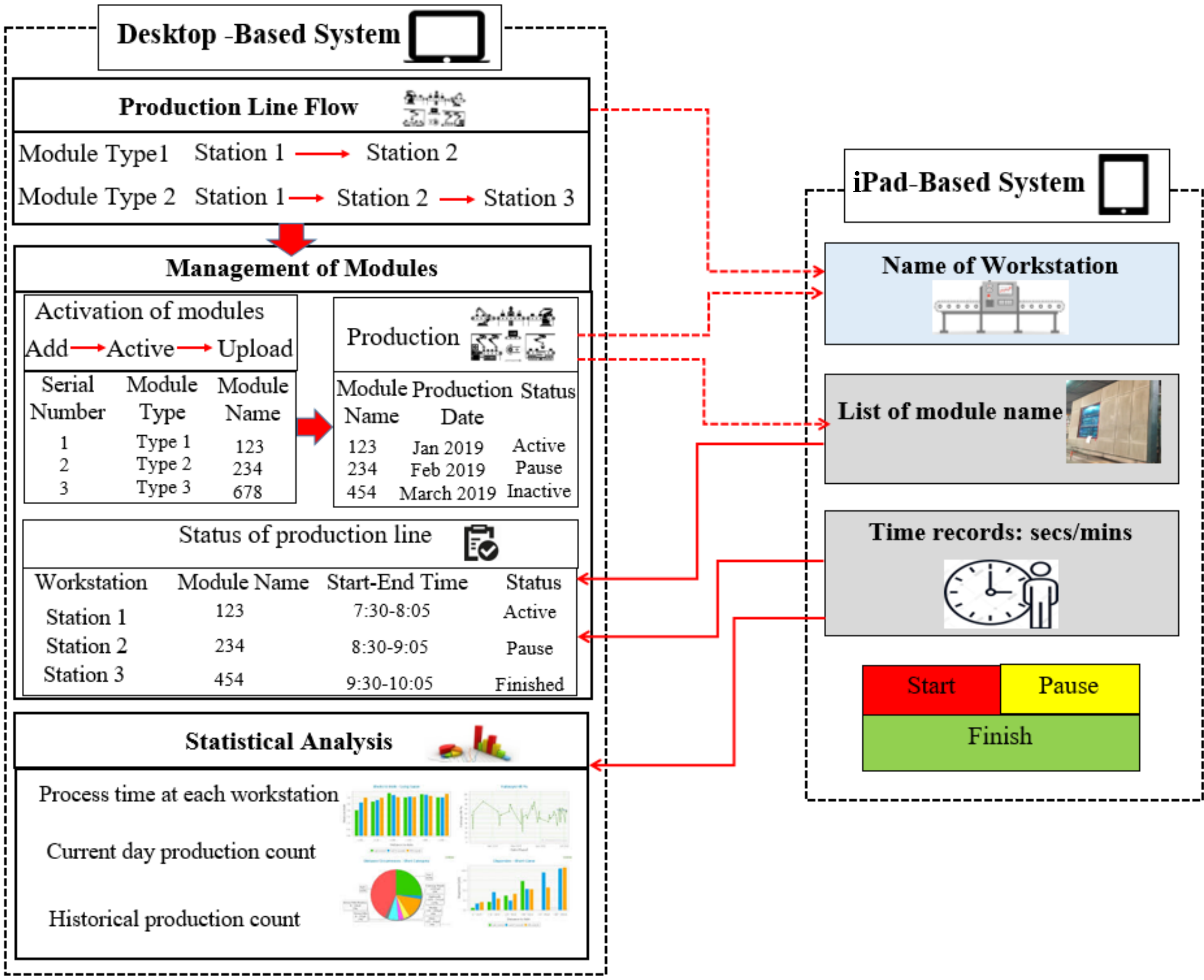
Figure 3.2: Main components of the simulation-based statistical method

3.2.1 Data Collection

The purpose of the data collection phase is to understand the performance of production lines by conducting work and time studies. The work study involves: (i) reviewing the SOPs of workstations in order to gain a high-level understanding of the sequences of activities at workstations prior to direct observation; (ii) classifying the types of modules (e.g., interior and exterior wall panels) and their components (e.g., number of window openings per module) using the shop drawings; and (iii) determining the breakpoints, i.e., the start and finish points for the various work processes at the workstations, in order to ensure that the process times are collected efficiently and accurately. In a typical time-study, a series of time data in the production line is collected and recorded manually using a stopwatch and timesheets, but this approach is time-consuming and inaccurate. Therefore, to achieve more accurate results in a more efficient manner, a C-track app is used in the present study in order to: (i) improve the accuracy of the data collection process by recording near real-time data of modules at workstations; and (ii) improve communication by providing efficient information transfer between the production line and the production planning department. Figure 3.3 (a and b) illustrates the details of the C-track app that consists of a desktop-based production management system (used in the production planning department) and an iPad-based system (installed at the workstations). The key users in the desktop-based system are the project manager and the project coordinator. The desktop-based system encompasses: (i) representation of the production line for the purpose of defining the workflow; (ii) management of the assignment of modules to workstations (i.e., sequence); and (iii) data organization (e.g., productivity of specific workstations and of the production line overall) based on the historical and near real-time recorded by the iPad-based system. The iPad-based system receives the information related to modules and their workflow from the desktop-based system. The workers at workstations follow the information concerning sequencing of modules assigned under the list of module names in order to record the process times. For collecting the time records, the 'Track App' features start, pause, and finish buttons. The worker first selects a 'module name' from the list, then presses 'start' to record the process start-time. In case of a disruption due to an error in the drawings or a work stoppage for a scheduled break, the worker uses the 'pause' button to stop the time record. After completing a module, the worker selects the 'finish' button. The timestamp is then recorded and transferred automatically in the 'production management system' and the database. Along with this, the 'module name' of the finished module is updated automatically in the 'product item' list of the next workstation. Although workers are still required to manually start/stop the time recordings, this semi-automated data collection application efficiently and accurately tracks modules at workstations and allows the productivity of the production line to be monitored in real-time. Based on the app's data organization feature, production managers can monitor bottlenecks on the production line in near real-time by comparing current and historical production rates.



(a)



(b)

Figure 3.3: (a) C-track application; and (b) detail flow of information in cloud-based tracking application

3.2.2 Data Analysis

Figure 3.4 provides a flowchart of the data analysis used to develop the method for predicting process times. To address the effect of missing data and outliers on the accuracy of the predictive method, the first step of the data analysis is to identify any missing values in the dataset due to human error (e.g., incorrect keystrokes on the app.) or technical issues such as temporary internet disconnection. The most common method to deal with missing values in the data is substitution through linear interpolation or regression (Piryonesi and El-Diraby, 2019). Furthermore, outliers, a set of data points that follow a pattern inconsistent with the rest of the data points, must be removed to improve the accuracy of the predictive method. Outliers can be identified and removed either by drawing scatter plots or by using residuals, which measure differences between observed and estimated durations. Residual analysis proceeds with: (i) fitting a regression to the dataset; (ii) finding the estimated durations; and (iii) calculating the residuals and adjusted residuals. In this research, the standardized residuals (individual residual divided by the standard deviation of residuals) are found to fall within the range of ±1.64, where data points outside of this range are considered outliers, as per Cottrell (2006). After detecting the outliers, normalization is implemented using Equation 3.1 to: (i) reduce the sizes of variables and thereby reduce computation time for calculating the process times; and (ii) improve the accuracy of the predictive method.

$$V' = (V - minA) / (maxA - minA) \qquad \text{Equation 3.1}$$

where minA and maxA are the minimum and maximum values, respectively, of the independent variable, A, and V represents the original value of A.

The dataset having been cleaned, statistical techniques such as PCA, t-test, and PCC are applied in order to identify the SIFs affecting the process times. These feature selection techniques enhance the performance of the predictive method and provide deeper understanding of the underlying process (Guyona and Elisseeff., 2003; Chanmeka et al., 2012). These techniques, it should be noted, are selected for the present study by virtue of (i) their wide use in investigating prediction and variable selection problems in manufacturing and pipe fabrication; and (ii) their simplicity, empirical accuracy, and generic applicability. To avoid overfitting, machine learning algorithms such as random forest and decision tree, as alternative methods for variable selection, are not considered because of the small size of the dataset, i.e., less than 1,000 observations. According to

Makridakis et al. (2018), the use of such methods for small datasets can yield a "black box solution", which may not be acceptable to industry practitioners. Below is a brief description of statistical techniques used:

1) t-test: The sample t-test is a statistical analysis technique used to determine the probability value of rejecting the hypothesis, which, in this research, is the statistical significance of the selected factors that affect the process times at MCM workstations. In this research, a sample t-test (Gerald, 2018) is deployed based on Equation 3.2, where n is the sample size, $\bar{x}$ is the mean of the sample data, μ is the population mean, and σ is the standard deviation. The SIFs are selected by evaluating the p-value—defined as p-value ≤ 0.01 in this research.

$$t = (\bar{x} - \mu)/(\sigma/\sqrt{n}) \qquad \text{Equation 3.2}$$

2) Pearson correlation coefficient (PCC): PCC (Yu and Liu, 2003) is represented by Equation 3.3, where $\overline{xij}$ is the mean of independent variable x j and $\overline{yi}$ is the mean of dependent variable y (duration). $xi\ j$ and yi represent the original values of variables xi and y. The range of R is ±1, where values close to 1 are indicative of a strong correlation between the independent variables, while values close to 0 signal a weak correlation between independent variables.

$$R = \frac{\Sigma\left(x_i^j - \overline{x_i^j}\right)(y_i - \bar{y}_i)}{\sqrt{\Sigma(x_i^j - \overline{x_i^j})^2}\sqrt{\Sigma(y_i - \bar{y}_i)^2}} \qquad \text{Equation 3.3}$$

where i is the number of observations and j is the number of variables.

3) Principal component analysis (PCA): PCA is a widely used data reduction technique that extracts the small set of variables that accounts for maximum variance in the original dataset. In this method, typically the principal components (PCs) accounting for 90% of the dataset's total variation are selected for further analysis (Rocchi et al., 2004). These components represent a linear combination of original independent variables, where the first PC has the most significant variance, and the succeeding PCs are built by reducing the variances of the preceding PCs. The

variation is expressed as the factor scores, which are calculated by representing the data into a form of a matrix (X) consisting of the $J \times I$ matrix where J is a number of observations and I is a number of independent variables. The singular value decomposition (SVD) or Eigen-decomposition is used to decompose the X into two orthogonal matrices and one diagonal matrix. Equation 3.4 represents the SVD of the X, where P is the eigenvectors of matrix P^TP, Q is the eigenvectors of Q^TQ matrix, which computes the factor scores and Δ is the diagonal matrix. As a result, the SIFs are identified by selecting the smallest factor score at each of the PCs.

$$X = P\Delta Q^T \qquad \text{Equation 3.4}$$

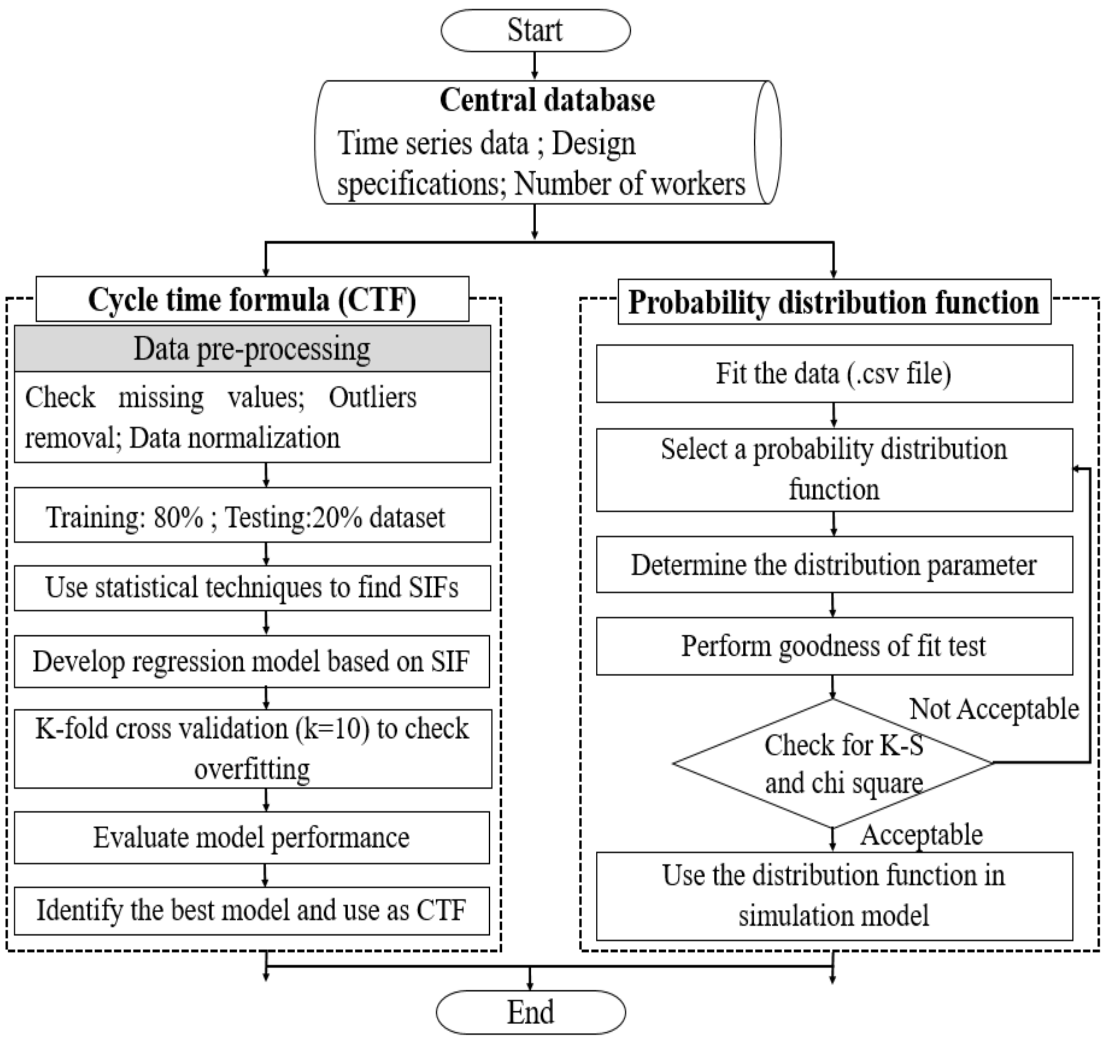


Figure 3.4: Flowchart of data analysis

Before developing the regression, the dataset is divided into training (80%) and testing (20%) subsets. Based on the SIFs and the training subset, multiple linear regression (MLR) are developed to predict the process times at workstations. To identify and mitigate overfitting, which diminishes the generalizability and accuracy of prediction, cross-validation, as a technique that identifies overfitting by testing the accuracy of the predictive method (Mahmood and Khan, 2009), is applied. The present study adopts K-fold cross-validation, as this is one of the most robust approaches for validating and testing the performance of a model, given that it utilizes unseen data. The dataset is divided into K folds, where, in turn, (K−1) folds are used to train the predictive method, and the remaining fold is used to test the accuracy. This process is repeated for K iterations, where the accuracy of the predictive method is determined by calculating the average of these K iterations.

Based on the results of the t-test, PCC, and PCA using different lists of SIFs, there are a few different MLR models capable of predicting the process times. There are four evaluation indices that can be used to determine which predictive approach is most accurate: (i) R-square (R^2); (ii) adjusted R-square, (adj. R^2); (iii) root-mean-square error (RMSE); and (iv) mean absolute error (MAE) (Willmott and Matsuura 2005, Elmousalami 2019). These evaluation indices have been shown to be reliable and are used widely in various disciplines such as manufacturing and construction (Martinez et al., 2020).

Based on the historical time data stored in the database, a probabilistic model is developed to predict the process times at workstations. Various functions, including Gamma, Weibull, Uniform, and Triangular, are fitted in order to generate the process times based on the parameters of the associated distributions, such as moment matching, maximum likelihood, and least squares. A given function and associated parameters having been determined, goodness-of-fit tests, including

Kolmogorov–Smirnov (K–S) and chi-squared, are performed in order to assess the goodness-of-fit between the observed and expected distributions (AbouRizk et al., 2016). If the predicted and actual process times are in close correlation, the given function is accepted for use as an input in the simulation.

3.2.3 Simulation Method

In practice, different types of modules, such as interior and exterior wall panels, may be produced in the same production line but following different SOPs at the workstations. In this respect, MCM typically follows a mixed-production line model, leading to varying process times and, in turn, imbalanced production flow and inefficient utilization of resources. Given this, it is important to reduce: (i) waiting time of modules between workstations; and (ii) idle time of workers at workstations. In this respect, continuous workflow can be achieved by planning efficient sequencing of modules and labor allocation through predictive modeling based on historical time data. In this regard, in the present study production planning is implemented using a simulation method developed in Simphony.NET. Figure 3.5 illustrates the process flow of the simulation-based planning method, which consists of (i) developing a simulation by mimicking the workflow of the production line (i.e., number of workstations and their sequence) based on the work study; (ii) importing inputs such as cycle time formula (CTFs), probability distribution functions, labor allocation, and module design specifications (e.g., heights of framing components and number of frames) from the database into the simulation using structured query language (SQL); (iii) comparing the results (productivity) between the cycle time formula (CTFs) and probability distribution functions using historical productivity data in order to identify the most accurate predictive approach ; (iv) using the most accurate prediction to develop multiple scenarios in the simulation by considering various sequences of modules and various labor allocation cases. For example (i) wall panels (exterior and interior) were prefabricated randomly in the production line considering floor levels, (ii) exterior wall panels of a given unit were prefabricated first, followed by the corresponding interior walls for the unit, (iii) the interior wall panels of a given unit were manufactured first, followed by the exterior wall panels, (iv) the interior wall panels of all units were produced first, followed by the exterior wall panels of all units and (v) the exterior wall panels of all units were produced first, followed by the interior wall panels of all units; and (v) selecting the best scenario, that is, the scenario with the minimum total duration to complete the project (i.e., to produce the given modules).

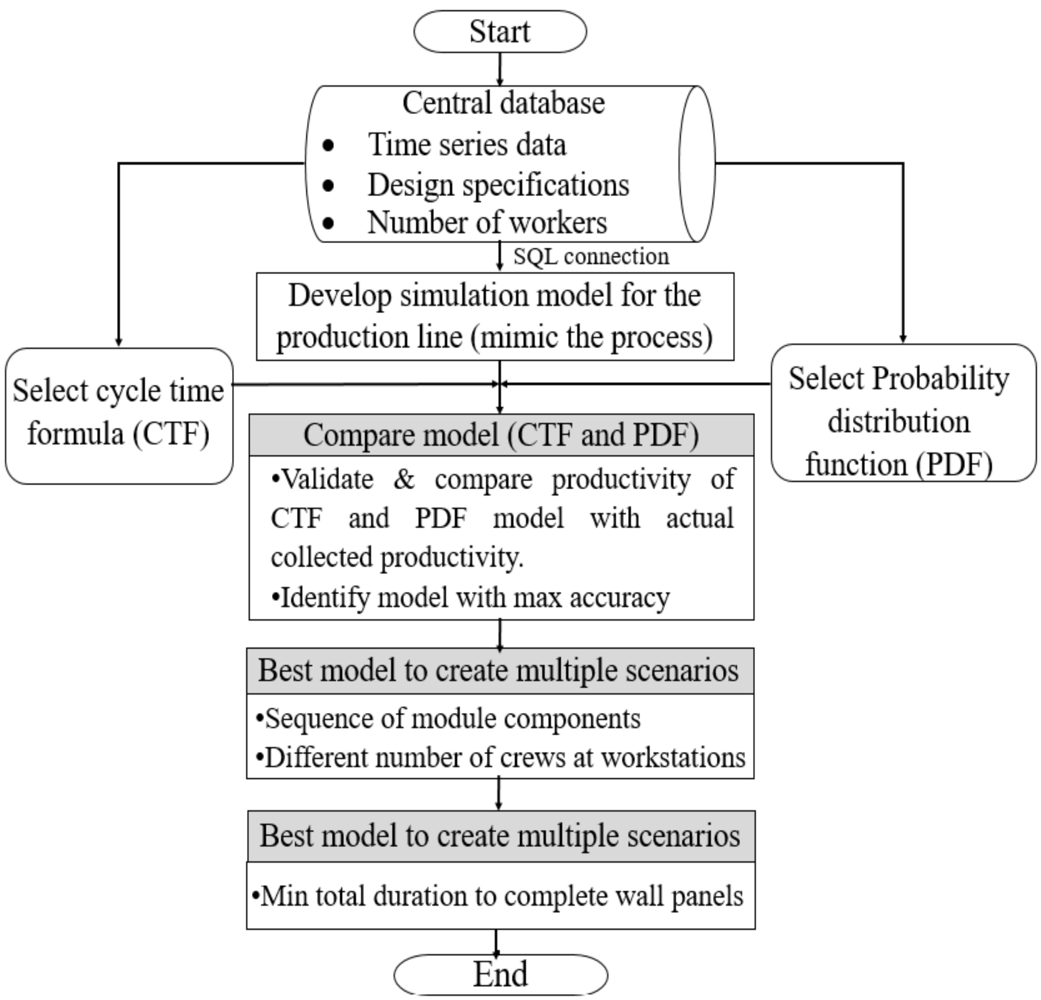


Figure 3.5: Flowchart of simulation-based planning

3.3 Module 2: Deep Neural Network Based Predictive Method

Figure 3.6 presents main components of the developed method for time prediction of prefabricating modules components in the production line. The method consists of three modules: (i) the first is the data input module, designed to facilitate and organize the input data pertinent to the unique characteristics of each module component for fabrication. It includes data description in order to specify the module components design parameters (e.g., number of studs and doors) and timestamps (i.e., start and finish time) from the collected RFID data; (ii) the second is the data preprocessing module. In it, the captured data is cleaned and prepared, combining attributes with the similar properties (e.g., windows + large windows are combined as 'windows') and identify and removal of outliers; and (iii) the third is the time prediction module, which houses three tools for time prediction and enables a comparative study to select the most suitable tool for predicting fabrication process time at each workstation. In this study ANN, DNN and MLR are utilized and compared. Later, GA is also used in conjunction with DNN to fine tune the architecture of the network (i.e., number of hidden layers, neurons known also processing elements, momentum and learning rate). For the ANN it is considered one hidden layer and processing layer with five neurons. It is noteworthy that a trial-and-error method is used in this study to determine the number of neurons in the single hidden layer of the ANN. The number of neurons is varied (1, 2, 3, 4, 5, and 6) with 100 epochs, and the resulting output (mean absolute error) was monitored. This procedure is repeated for each neuron count, and the network configuration with the lowest error rate (i.e., one hidden layer and five neurons) is chosen. Additionally, multiple linear regression (MLR) is selected due to its successful application in previous studies for predicting process time and its ease of result interpretation. In MLR, a linear relationship is established between the independent variables and the dependent variable. This relationship is a combination of the independent variables with predetermined weights and a bias coefficient, as illustrated in Equation 3.5 (Hasan and Lu 2022).

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots\ldots\ldots\ldots\ldots \beta_k x_k \qquad \text{Equation 3.5}$$

where Y is the dependent variable (process time); X_k = independent variables (number of studs and door); β_0 = bias value; and β_k = weighting coefficients for the independent variable.

The data used in this study is provided by the case study company from July 2015 and August 2018. The data captured utilized RFID system, with passive RFID tags attached to each module component. In total 416950 timestamps (i.e., start and finish time) for module components along the production line is gathered. The captured dataset, after preprocessing, is randomly divided into training (80%) and testing (20%) subsets. The criteria in the developed method include the availability of materials in the factory, which does not affect the process time, and capacity of workstations defined by the maximum module component length they can accommodate (e.g., the capacity of a framing workstation may be a length of 40 ft).

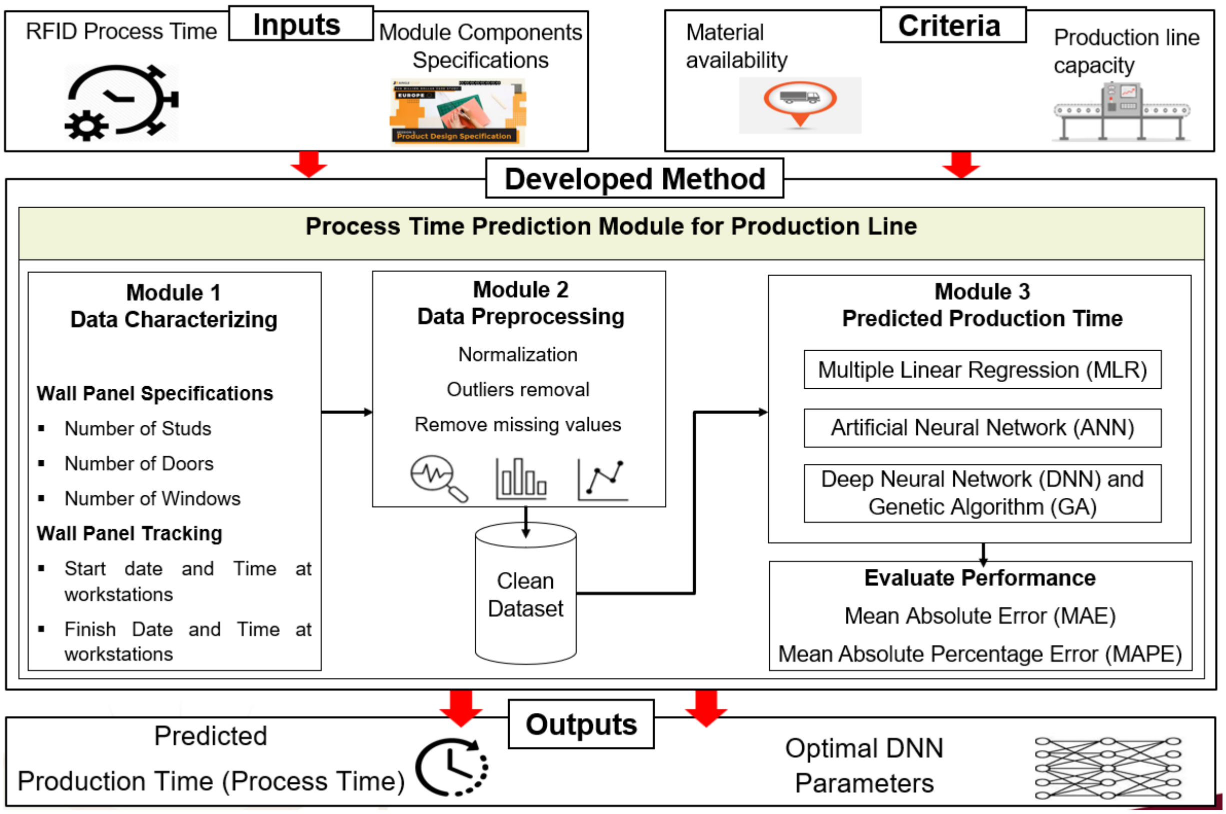


Figure 3.6: Overview of the Developed Method

3.3.1 Data Characterizing and Data Preprocessing

Data characterization involves two steps: (i) understanding the production line, and (ii) extracting process time. In the first step, the standard operating procedures (SOPs) of each workstation are reviewed to gain a comprehensive understanding of the production line before direct observation (e.g., identifying which workstations are fully automated and which are manual). Additionally, shop drawings are used to classify module components (e.g., the number of studs and windows in interior and exterior wall panels). In the second step, before extracting process time, breakpoints (i.e., the start and finish points for the various work processes at the workstations) are identified to align RFID antennas and readers with the corresponding workstations. The RFID-based system (Bardareh and Moselhi 2022; Altaf et al. 2018) is utilized to collect production line data in order to develop predictive method. The system consists of: (i) RFID printer that generates the ID tag number for module components; (ii) RFID tags that are attached to each module component; (iii) RFID antennas that are installed at each workstation (i.e., read zone) and picks the tag signals thereby ensuring to capture the movement of module components along the production line; and (iv) RFID readers to which these Antennas are connected and transfers the timestamps (i.e., initial and last read time) into the database. The next critical step is to extract the process times and relevant attributes of the module components (e.g., number of studs and doors) at each workstation from the RFID raw data file provided by the industrial partner. The process times (i.e., the time required to complete one module component at each workstation) is extracted using Equation 3.6:

$$PT_{m,w} = IRT_{m,w+1} - IRT_{m,w} \qquad \text{Equation 3.6}$$

Where PT is the process time of module component ID 'm' at current workstation 'w', 'w+1' is the next workstation and IRT represents the initial read time at workstations.

In this respect, the next essential step is to perform data pre-processing and ensure that the dataset is cleaned for prediction purposes. Data pre-processing includes: (i) discarding missing values; (ii) removing outliers using data visualization techniques (i.e., pie charts) and 'Mean ± 1.5 SD' (i.e., data points that are above and below are considered as possible statistical outliers); (iii) combining attributes with the similar properties (e.g., doors + large doors are combined as 'Doors' and different types of studs such as studs+ Dstud + Lstud + Mstud are combined as 'Studs'); and (iv) data normalization are implemented using Equation 3.7 in order to reduce the sizes of independent variables and model computation time.

$$v' = \frac{(v - min_A)}{(max_A - min_A)} \qquad \text{Equation 3.7}$$

where min_A is the minimum value and max_A is the maximum value of the independent variable A, and V represents the original value of A.

3.3.2 Predicted Production Time

Figure 3.7 presents a flowchart of the developed method for predicting production time, which integrates a deep neural network (DNN) and a genetic algorithm (GA) to predict the process times of module components (e.g., wall panels) at workstations along the production line using historical data. A DNN is characterized by having two or more hidden layers (Aggarwal, 2018). In this thesis, 'deep' refers to the presence of multiple hidden layers, enabling the network to learn complex representations from the input data. It should be noted that this study used a deep neural network (DNN) due to its features of having multiple hidden layers (i.e., two or more), which can efficiently perform complex non-linear transformations. Each layer entails several neurons representing the input, transfer, and output variables. The dataset, having been pre-processed, is divided into training (80%) and testing (20%) subsets and based on the training subset, a predictive method is developed. In this thesis, the rectifier activation function is selected, and the range searched for upper bound/lower bound is 3-10 for hidden layers and 6-100 for the number of neurons. Additionally, the cross-validation technique in order to prevent overfitting and obtain a better evaluation of the predictive method is applied. The present study adopts K-fold cross-validation for testing the performance. The dataset is divided into K groups, where, in turn, the predictive method is trained using (K−1) groups, and the remaining fold is used to test the accuracy of the predictive method. The process is then iteratively repeated, holding a different group for validation and using the remaining ones for training. The overall performance is then measured as the average performance of each iteration. The parameters of the prediction algorithms were automatically tuned based on an optimization procedure aiming to minimize the mean absolute error resulting from the 10-fold cross-validation.

It should be noted that the parameters of the DNN (i.e., the number of hidden layers, nodes, learning rate and momentum) have a significant impact on the model's performance. For instance, if there are a small number of nodes, the model cannot be trained well, and with a large number of nodes, performance can be enhanced, but a large number of connections will increase the computational time. Therefore, it is critical to establish hyperparameters for neural networks which can be trained in a reasonable computational time and provide errors within the tolerance limit. Therefore, hyperparameter tuning is performed, to determine optimum DNN parameters (i.e., number of nodes and hidden layers). This research implements the GA for hyperparameter tuning because it is faster and more efficient while finding optimal solution for prediction problem. The optimization objective is to minimize the prediction error (i.e., mean absolute error (MAE)) as show in Equation 3.8.

$$MAE = \frac{\sum_{i=1}^{n} |A_i - P_i|}{n} \qquad \text{Equation 3.8}$$

Where A_i is the actual process time of a module component; and P_i is the predicted process time of a wall panel.

Additionally, error percentage is calculated using the symmetric mean absolute percentage error (SMAPE), which is a modified MAPE and involves dividing the absolute error by the average of the actual observation and the predicted process time as shown in Equation 3.9. According to Makridakis, (1993), MAPE is asymmetric, as it imposes a greater penalty on predictions that surpass the actual values compared to those that fall short (i.e., a small actual value in the denominator lead to very high percentage errors). To address this limitation, this thesis uses SMAPE, which is symmetric and treats overestimation and underestimation errors more equally by incorporating both actual and predicted values in the denominator. This ensures a more balanced evaluation of both over-predictions and under-predictions.

$$SMAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{A_i - p_i}{(A_i + p_i)/2}\right| \quad \text{Equation 3.9}$$

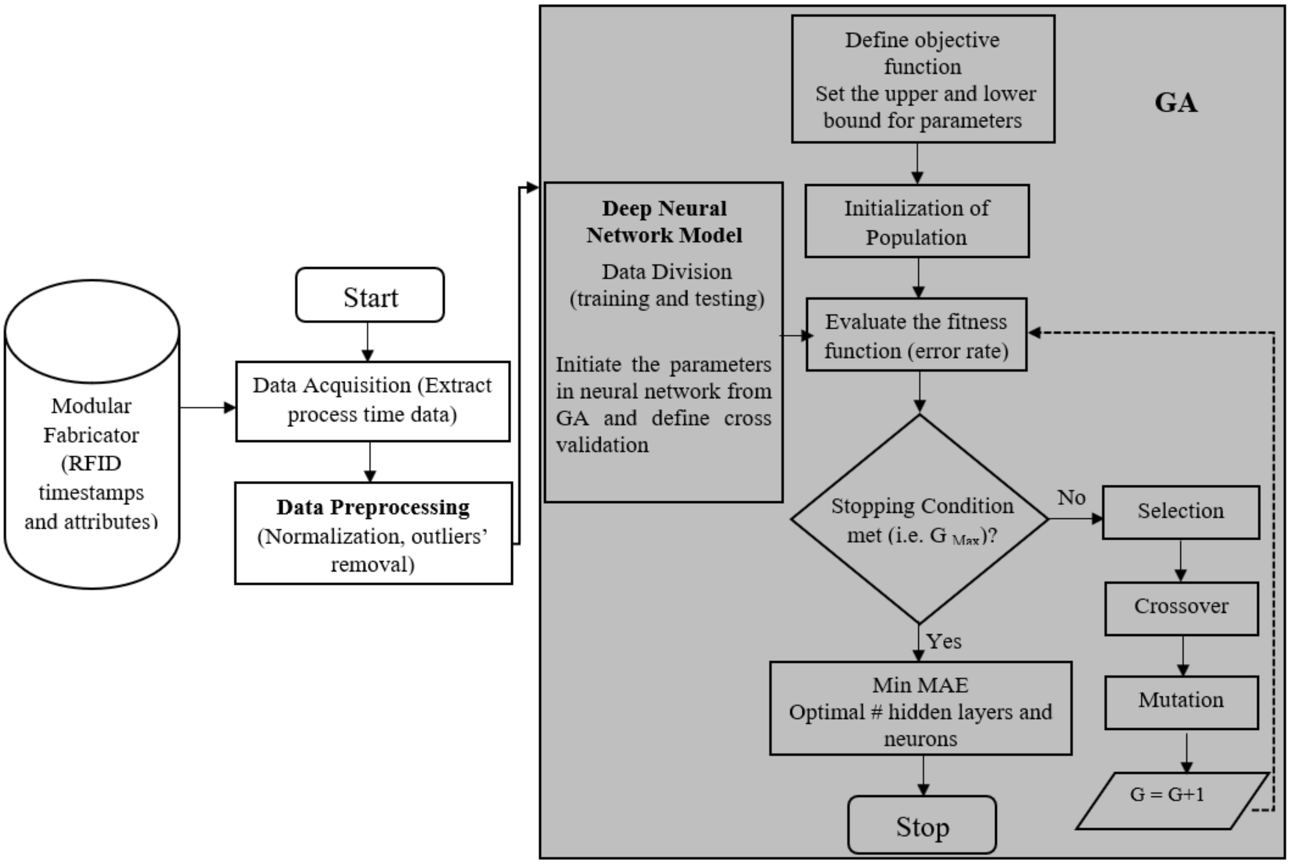


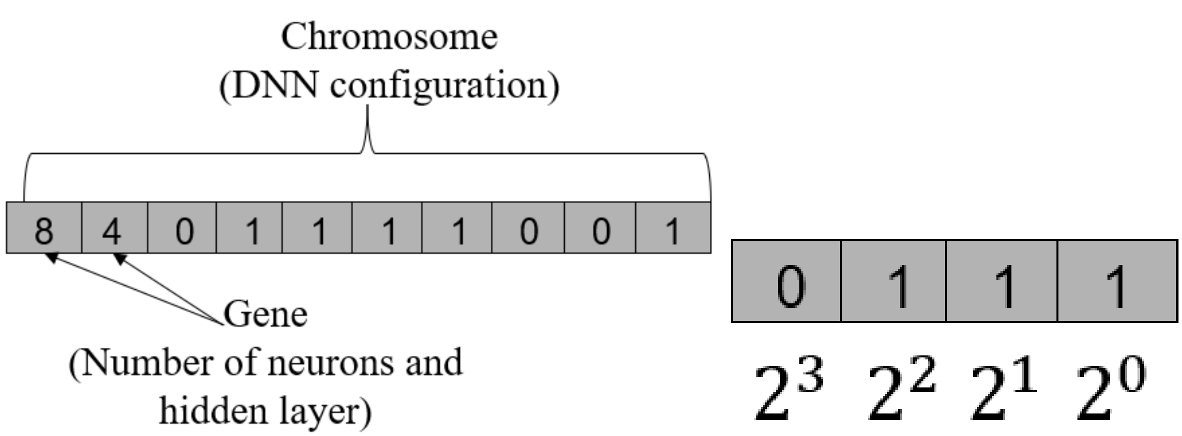
Figure 3.7: Flowchart for Deep Neural Network Optimization

The optimization process involves: (i) creating an initial population, which consists of chromosomes that represent various configurations of the number of neurons, hidden layers, learning rate, and momentum. This initial population is generated randomly, with the number of chromosomes determined by the population size. Each chromosome contains multiple genes, each representing a specific aspect of the DNN configuration (i.e., neurons, hidden layers, learning rate, and momentum). Notably, the number of neurons and hidden layers are represented as discrete values, chosen randomly from the combinatorial range of their upper and lower bounds, whereas learning rate, and momentum are represented as continuous values. In this study, neurons have a step size of 2, and hidden layers have a step size of 1. Figure 3.8 (a) illustrates the structure of a generated chromosome, which represents a potential solution. For example, in the chromosome 8,4,0,1,1,1,1,0,0,1, '8' indicates the number of neurons in each hidden layer (a discrete value), '4' denotes the number of hidden layers (also a discrete value), and the binary numbers represent the learning rate and momentum (continuous values). In this binary chromosome, the first four digits correspond to the learning rate, and the remaining digits represent momentum. These binary strings are decoded using Equation 3.10, a crucial step for converting the encoded parameters back into their real-world values.

$$\mathrm{D} = \left(\Sigma bit \times 2^{i}\right) \times P + \alpha \qquad \text{Equation 3.10}$$

$$\mathrm{P} = \frac{b-a}{2^{l}-1} \qquad \text{Equation 3.11}$$

Where D represents the decoding process, and P is the parameter (as shown in Equation 3.11). Variables a and b are the lower and upper bounds of the parameter (e.g., learning rate), and l is the length of the chromosome. For instance, to decode the variable learning rate, which has lower (a) and upper bounds (b) of 0.01 and 0.3 respectively, and a chromosome length of 4 (Figure 3.8 b), the decoded value (D) will be 0.15, which falls between 0.01 and 0.3.



(a) (b)

Figure 3.8: (a) Structure of the chromosome; and (b) Decoding of chromosome (learning rate)

(ii) performing the process of evaluation, where fitness values of each chromosome are evaluated based on the optimization criteria (i.e., minimum error rate); (iii) the selection process involves choosing the best chromosomes (i.e., potential parents) from the population as the best solution (i.e., the best combination of the number of hidden layers, neurons, learning rate, and momentum) based on their fitness value. This thesis uses tournament selection for this purpose. Tournament selection is chosen due to its efficiency, ease of implementation, and relatively low computational time compared to methods like roulette wheel selection, as it requires fewer comparisons (Razali and Geraghty 2011). In this approach, k (i.e., tournament size) number of chromosomes are randomly selected from the population, and the chromosome with the best fitness score among the k chromosomes is selected as parent 1. This process is repeated until the desired number of parents for crossover is selected. For example, in a population of 10 with a tournament size (k) of 3, chromosomes 1, 2, and 3 with fitness values of 0.75 min, 0.82 min, and 0.68 min are randomly picked. The chromosome with the best fitness (i.e., minimum error rate), in this case, chromosome 3 (fitness 0.68), is selected as the first parent, and the process continues to find the second parent; and (iv) after selecting two parents, crossover is performed to create new chromosomes for the next generation. The number of hidden layers and neurons are exchanged between parent 1 and parent 2, while for continuous variables (i.e., learning rate and momentum), a two-point crossover is used. This method is preferred due to its effectiveness and successful application in various optimization problems (Murata 1996). For each pair of parents, a crossover point is chosen randomly. Figure 3.9 demonstrates the crossover process, where offspring one inherits the number of neurons (8), number of hidden layers (5), and continuous values for learning rate and momentum from parent 2. Later mutation process is performed, and this process (i.e., the sequence of selection, crossover, and mutation) continues, with each DNN configuration being trained, tested, and evaluated using the fitness function, until the termination criteria (i.e., maximum number of generations) are met. The developed method uses: (i) a population size of 20; (ii) a maximum of 50 generations; (iii) a mutation probability of 0.1; (iv) a crossover probability of 1; (v) a tournament size of 3; (vi) momentum ranging from 0.01 to 0.99; and (vii) a learning rate ranging from 0.01 to 0.3.

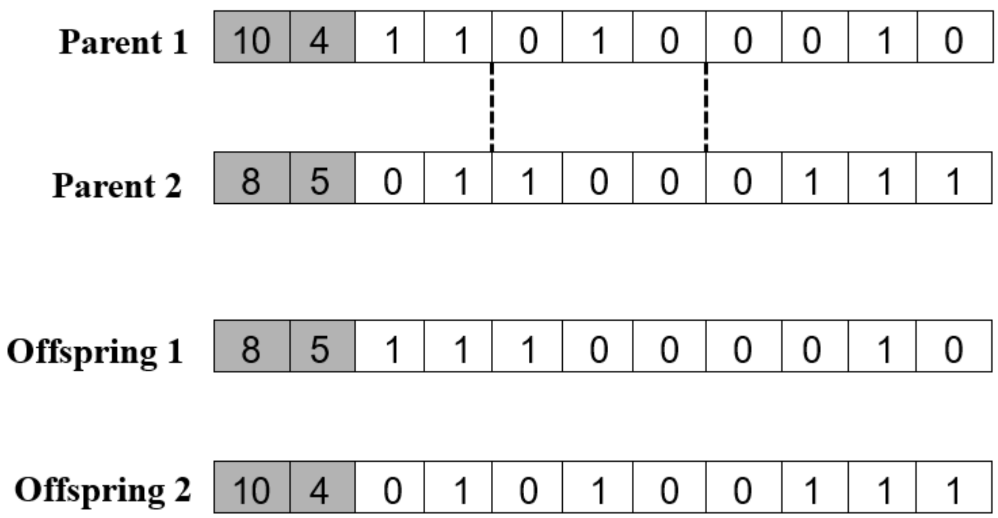


Figure 3.9: Two-point crossover

The DNN results are compared with those of ANN and MLR to determine the best tool for predicting the process time of prefabricating module components at each workstation of the production line. The input layer of the neural networks (both ANN and DNN) comprises a number of independent variables, which vary according to the standard operating procedures at the workstations. For example, as indicated in Table 3.1, the framing station has five variables (number of studs, doors, windows, length, and height), whereas the window/door installation station includes four variables (number of doors, windows, length, and height). It is worth mentioning that process time at workstations largely depends on the type of module components (interior and exterior wall panels) and its design specifications (e.g., number of doors, windows and studs) (Altaf et al. 2018). Therefore, this research primarily focuses on considering module components design specifications as input variables into the prediction model. The number of hidden layers and neurons in the DNN is chosen using the GA optimization technique. Each hidden layer and number of neurons in hidden layer is formed as a weighted sum of all input features based on the connection weights. During this feed-forward step, the output layer receives the values from neurons and calculates their weighted sum to produce a predicted result. The output layer provides the predicted dependent variable, which is process time (duration) in this study. The ReLu (Rectified linear unit) is used as an activation function to introduce non-linearity into the network. It should be noted that input factors (i.e., independent variables) used to train the MLR, ANN and DNN varies according to workstations.

Table 3.1:Input variables at workstations

Input Layer Variables	**Framing Station**	**Sheathing Station**	**Nailing Station**	**Butterfly Table**	**Window Door Installation**
Number of Studs	✓	X	X	X	X
Length (ft)	✓	✓	✓	✓	✓
Height (ft)	✓	✓	✓	✓	✓
Number of Doors	✓	✓	✓	✓	✓
Number of Windows	✓	✓	✓	✓	✓
SheetPartial	X	✓	✓	X	X
SheetFull	X	✓	✓	X	X
Area	X	✓	✓	X	X
Nailcount	X	X	✓	X	X
Nailline	X	X	✓	X	X

It should be noted that MAE is selected as a measure of performance, which according to various studies is a better alternative to R square while evaluating the performance of model in respect to non-linear data. According to Lseth (1983), utilization of R square for evaluating model performance in non-linear data, leads to misinterpretations and produces misleading conclusions. Moreover, in the mathematical literature, it has been concluded that the R-square generally do not increase even for better non-linear predictive models (Spiess and Neumeyer 2010).

3.4 Module 3: Optimized Planning and Scheduling Method

Figure 3.10 presents the components of the developed method and its scheduling process for modular construction manufacturing (MCM) production lines. The developed procedures include three steps: *(i)* creating a work planning structure (WPS), which involves grouping panels by type and location (i.e., precedence relationships for sequencing and installing panels based on types and locations within the building), namely work-packages that serve as fundamental elements of the developed scheduling method; *(ii)* developing optimal sequences of panels in the production line using optimization algorithms; and *(iii)* validating the effectiveness of the developed hybrid optimization algorithm by comparing the results between GA, SA, and HGASA. The input parameters include the process times of panels at workstations and the layout of the production line. Additionally, input parameters include the specifications of panels (e.g., length (ft), height (ft), number of studs, number of doors and windows) collected by extracting quantities of panels from BIM to support panel component-centric planning and scheduling by providing essential information (e.g., total number of wall panels) for downstream factory production lines. At this juncture, it should be noted that this thesis uses the process times of panels at workstations generated by a predictive method, which is a result of previous study (Bhatia et al. 2023). The criteria are the working hours and the capacity of workstations (i.e., the maximum length of panels the workstation can accommodate).

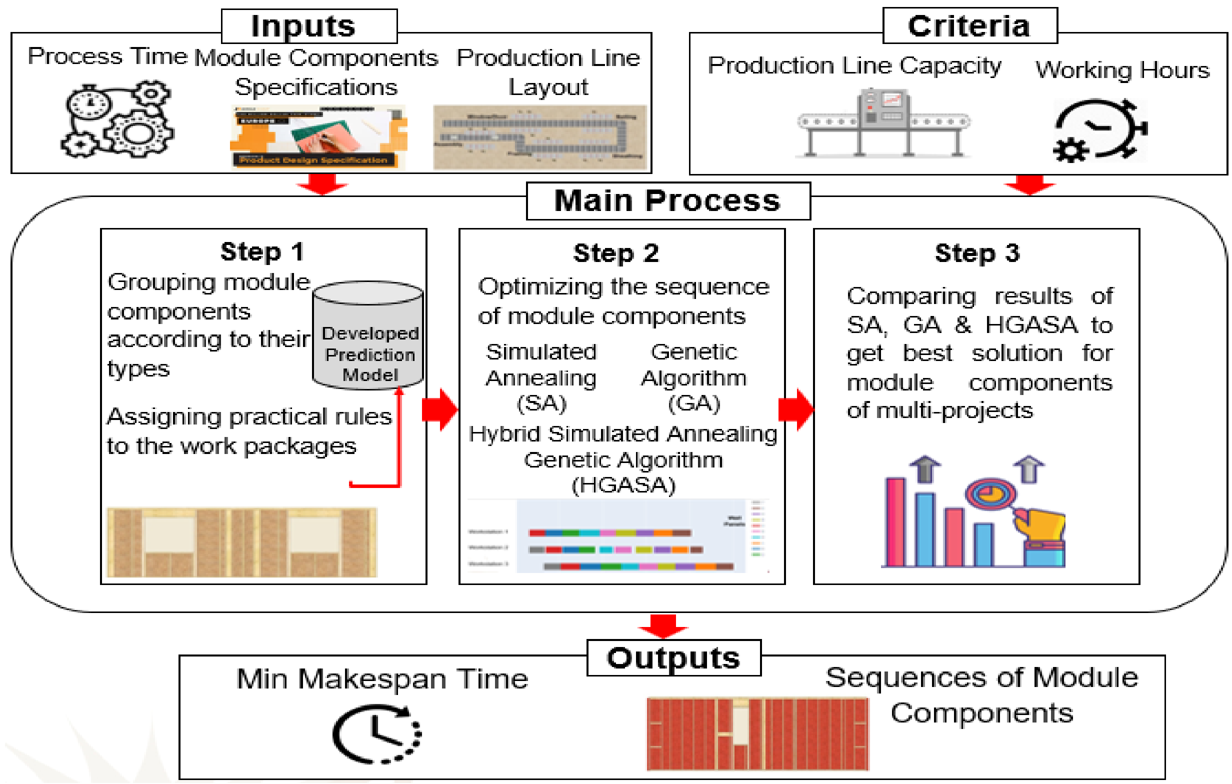


Figure 3.10: Developed method for optimized planning and scheduling

3.4.1 Work Planning Structure

Although the aforementioned study (Zhang et al. 2020) introduced a production line breakdown structure (PBS) that supports production managers in analyzing, diagnosing, and solving problems to enhance the performance of MCM. However, it is not suitable for planning and scheduling the sequences of panels to be fabricated in the production line, since it does not support: (i) grouping of panels according to types (e.g., interior versus exterior wall panel) and locations of the panels on the building (e.g., 1st storey versus 2nd storey); and (ii) the requisite fabrication processes (i.e., number of workstations required to prefabricate the panels in the production line can be different in accordance with panel types). For example, the required number of tasks varies for interior wall panels with and without openings. Three workstations (i.e., framing, sheathing and nailing) are required for interior wall panels without openings, while, for those with openings, four workstations (i.e., framing, cutting, sheathing and nailing) are required. This information is essential for planning and scheduling the sequences of panels for multiple projects. In practice, based on work dependencies and engineering logic, interior wall panels need to be installed on site before exterior wall panels on any given floor. Therefore, interior wall panels on the 1st storey must be prioritized over exterior wall panels for prefabrication in the production line. To efficiently plan and schedule the sequencing of panels across multiple projects: (i) these panels (e.g., interior and exterior wall panels) should be grouped according to their type, location and required number of tasks (i.e., workstations); and (ii) assigning priority to these groups of panels (i.e., work-packages) based on the practical rule of installing wall panels on-site. In this respect, as shown in Figure 3.11, this thesis develops a WPS that involves five levels: (i) level 1 represents the project type (e.g., residential multi-storey or single storey); (ii) level 2 involves total number of floors; (iii) level 3 contains the different types of panels (e.g., wall, floor and/or roof panels) and total number of panels at each floor; (iv) level 4 represents the type of panel component, such as interior versus exterior wall panels with or without doors/windows; and (v) level 5 is the number of workstations required to prefabricate these panels in the production line. Accordingly, this WPS provides the following benefits: *(i)* grouping the panels with practical rules as work-packages leads to the development of precise project schedule estimates (i.e., makespan) for multiple projects since these work packages are manageable components; and *(ii)* information regarding the amount of work required (i.e., total number of panels) facilitates the clear identification of work dependencies (e.g., on-site installation sequence) and enables effective tracking of the project progress. It is worth mentioning that, in this thesis, the practical rules are defined as heuristicbased guidelines to assign a priority to the work packages (e.g., group of module components needed for initial on-site installation should be prioritized in production line over other groups of module components).

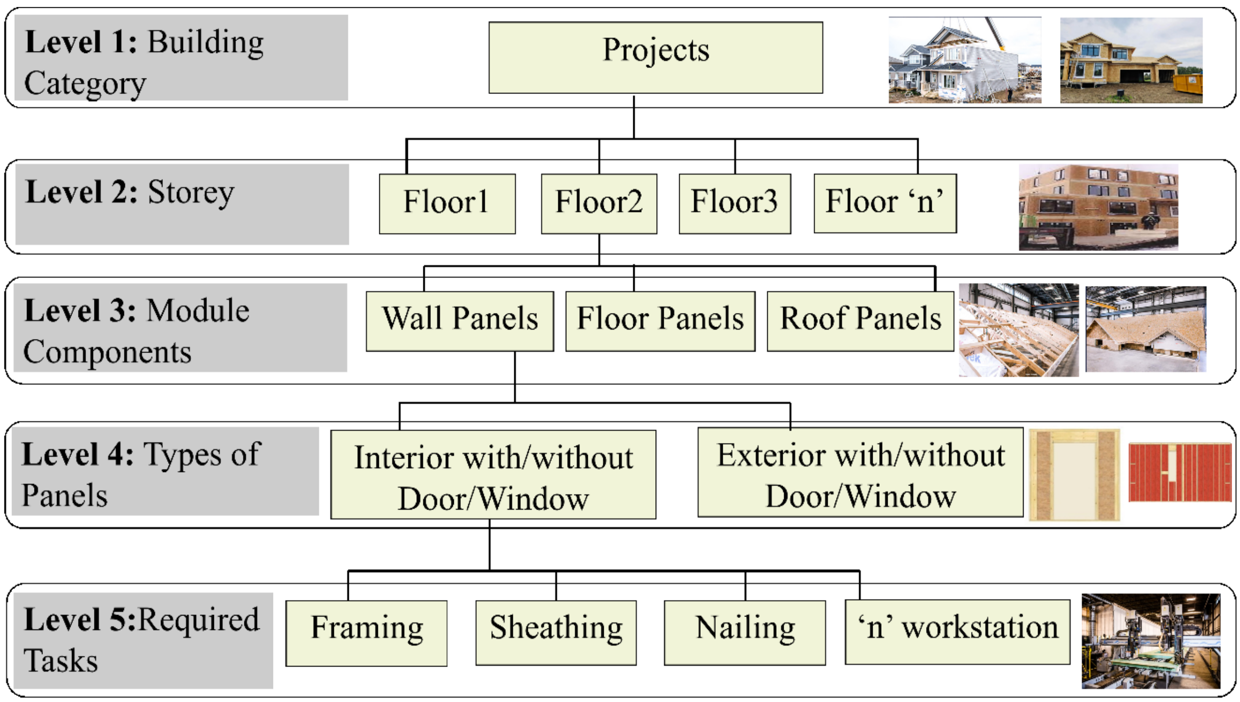


Figure 3.11: Work Planning Structure

3.4.2 Scheduling Method

3.4.2.1 Modular Construction Manufacturing Scheduling Problem (MCMSP)

In the workstations of the production line, the panels are prefabricated in a sequence predefined by production managers. As such, the MCMSP is modeled as an operation sequencing optimization problem, where M = {1, 2…… m} and W = {1, 2…… w}, 'm' being the number of panels to be scheduled at a set of 'w' workstations. Based on the sequencing process, constraint 1, as outlined in Equation 3.12, ensures that the start time of a given panel at a workstation will be greater than the completion time of the previous panel at the workstation (i.e., 'm−1'). The completion time of the previous panel is calculated as the sum of the start time (S) of the previous panel and the process time (P) of the previous panel at the workstation.

$$S_{m,\,w} \geq S_{(m-1),\,w} + P_{(m-1),\,w};\ \forall\, m \in M,\ w \in W \qquad \text{Equation 3.12}$$

Constraint 2, represented in Equation 3.13, ensures that the start time of a given panel at a given workstation will be greater than the completion time of that panelat the previous workstation (i.e., 'w−1'). For instance, if the start time of panel 2 at the framing workstation is 8:30 a.m., then panel 2 must be finished at the previous workstation before 8:30 a.m. It should be noted that $S_{m,\,w}$ is the maximum value in constraints 3.12 and 3.13.

$$\mathrm{S}_{\mathrm{m,\,w}} \geq \mathrm{S}_{\mathrm{m,\,(w-1)}} + \mathrm{P}_{\mathrm{m,\,(w-1)}};\ \forall\, \mathrm{m} \in \mathrm{M},\ \mathrm{w} \in \mathrm{W} \qquad \text{Equation 3.13}$$

Equation 3.14 represents Constraint 3, which has to do with workstation capacity (*X*), ensuring that each workstation can process a maximum of one single panel at a time.

$$Max\ X_w = 1 \qquad \text{Equation 3.14}$$

In practice, when bottlenecks appear, temporary manual workstations (i.e., parallel workstations) are added to not only eliminate the bottleneck but also meet the on-site delivery date requested by sites. Therefore, to reflect this practical case of capturing parallel workstations (i.e., a workstation that can prefabricate two wall panels simultaneously) in the optimization algorithms, Equation 3.15 outlines the 4th constraint concerning workstation capacity (X), which stipulates that a maximum of two wall panels can be prefabricated at parallel workstations during any given time.

$$Max\ X_2 \leq 2;\quad \forall\, \mathrm{S}_{\mathrm{m,2}} \qquad \text{Equation 3.15}$$

Equation 3.16 outlines the 5th constraint, which stipulates that the start time of a wall panel (i.e., panel 'm') at parallel workstation must be equal to the completion time of the same wall panel at previous workstation. This can reduce the waiting times of wall panels between workstations, and allows the parallel workstation to work on two wall panels simultaneously. It is important to be noted that if the workstation is still processing the previous two wall panels (i.e., max two wall panels can be prefabricated simultaneously at parallel workstation) and the next wall panel is finished at the earlier workstation, the next wall panel will have to wait in the queue. In this situation, it is crucial to identify which of the two wall panels at the parallel workstation completed first. This is determined using Equation 3.17, which states that the start time of the next wall panel at the parallel workstation should be equal to the earliest completion time of the previous two wall panels (i.e., m-1 and m-2) at the parallel workstations.

$$S_{m,2} = C_{m,1} \qquad \text{Equation 3.16}$$

$$S_{m,2} \geq \min\left(C_{m-1,2};\ C_{m-2,2}\right) \qquad \text{Equation 3.17}$$

Based on these constraints, the objective of the optimization problem examined in this study is to find a near-optimal schedule (i.e., sequences of panels) in order to minimize the makespan (i.e., the minimum time to complete all panels in the production line), represented by Equation 3.18:

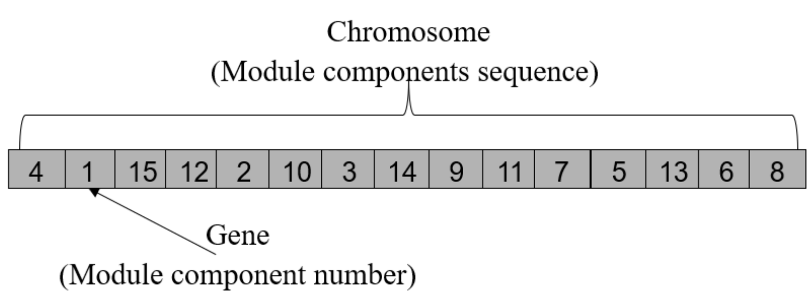
$$Min\ (Max\ C_{m,w};\ \forall\ m \in M,\ w \in W) \qquad \text{Equation 3.18}$$

$$C_{m,w} = S_{m,w} + P_{m,w} \qquad \text{Equation 3.19}$$

where $C_{m,w}$ is the completion time of a module component at a workstation; S is the start time of a module component, and P is the processing time of a module component. It is worth mentioning that $Max\ C_{m,w}$ involves determining the maximum completion time among all module components, which is computationally faster than calculating and summing up the completion times for every wall panel. This reduces the number of operations needed during optimization.

3.4.2.2 Genetic Algorithm Optimization

As shown in Figure 3.14, GA is selected as a fundamental algorithm for the development of the optimized scheduling method in this research, as it has been widely adopted in optimization problems in the construction industry. The process begins with extracting and pre-processing the RFID-based data provided by the modular fabricator. This data is used to develop a predictive method in order to forecast the process times of module components. Additionally, the module component information from a BIM is used to categorize module components based on type (e.g., interior versus exterior wall panels) and the required fabrication process (i.e., number of workstations required). This involves during the process of production initialization to determine precedence relationships for installing module components according to on-site locations. For example, during production initialization, 'job type' is checked. If the component is identified as 'Int wall panel' on a floor, the work-package associated with this 'Int wall panel' type is prioritized for prefabrication ahead of the work-package for the 'Ext wall panel' type. Consequently, the 'Ext wall panel' remains queued until all int wall panels are completed. Prior to the implementation of the optimization algorithms, the matrix that contains the predicted process times of module components at each workstation of the production line for the given work-package is developed. This is used as an input in the optimization algorithm. The critical elements of the GA are initialization, selection, crossover, and mutation. These elements function as follows: (i) First, an initial population (i.e., comprising chromosomes representing a list of module components) is generated randomly. Figure 3.12 (a), shows the structure of the generated chromosome, which represents a possible solution (i.e., sequence of module components). This study uses a permutation of module components (i.e., arranging the number of module components in different orders) as a chromosome, where a primary module component is assigned first, followed by the succeeding module component. The number of chromosomes in a population is determined based on the population size. Each chromosome in the population contains the number of genes, and each gene represents a module component number (these being linked to the workstations of the production line). For instance, the permutation 4, 1, 15, 12, 2, 10, 3, 14, 9, 11, 7, 5, 13, 6, 8 is a chromosome that represents a sequence in which module component # 4 is prefabricated first on all of the workstations, followed by module components #1 and #15. (ii) Second, the process of evaluation is performed, where fitness values of each chromosome are evaluated based on the optimization criteria (i.e., minimum makespan). (iii) Third, the best chromosomes (i.e., potential parents) from the population are selected as the best solution (i.e., optimal sequence of module components) based on their fitness value. In this research, the selection process is performed using roulette wheel selection. The rationale for this choice is that it is the most suitable selection strategy for flow-shop scheduling problems due to its ease of implementation and lower computational time compared to other selection strategies (Anand et al. 2015; Chen et al. 2009). Moreover, roulette selection ensures that individuals with higher fitness have a higher likelihood of being chosen, while still allowing individuals with lower fitness to have a chance of selection. This helps to maintain diversity in the population and prevents premature convergence of solutions (Yadav and Sohal 2017). In this respect, chromosomes are chosen according to their fitness score relative to the entire population (i.e., a chromosome with a higher score has a greater chance of selection). Each chromosome is assigned a probability (p) of being selected, where p equals the fitness of the chromosome divided by the sum of the fitness scores of all chromosomes in the population (Xei 2001). It should be noted that the goal of the selection process is to retain the favorable and remove the unfavorable chromosomes in order to produce offspring (i.e., new schedules) and, ultimately, the next generation of the population. (iv) Fourth, following the selection of two parents, crossover is implemented in order to generate new chromosomes for the next generation. In order for each pair of parents to be mated, a crossover point is chosen using a two-point crossover. The two-point crossover is selected due to its effectiveness and successful implementation in various flow-shop problems (Murata 1996). Figure 3.12 (b) illustrates the two-point crossover process, where a set of module components between two randomly selected points is inherited from parent 1 to the offspring and the other module components are placed in order of their appearance from parent 2. For example, offspring 1 is obtained by taking genes 1, 2, and 5 from parent 1 and the rest (i.e., 6, 3, 4, and 7) from parent 2. In the mutation process, to introduce variability and diversity, some of the genes in the offspring are exchanged. In the present study, swap mutation is implemented, where mutation involves randomly selecting two genes in a chromosome and swapping their positions. Figure 3.12 (c) shows the positions of genes 12 and 3 being swapped. It should be noted that mutation occurs only when the mutation rate is met. This selection, crossover, and mutation process continues until the termination criteria (i.e., the maximum number of generations) are met. The developed method considers a generation size of 200, a population size of 30 chromosomes (i.e., schedules), and crossover and mutation rates of 0.8, and 0.2, respectively. These values are determined by experimenting with various values and selecting the one yielding the best results.



(a)

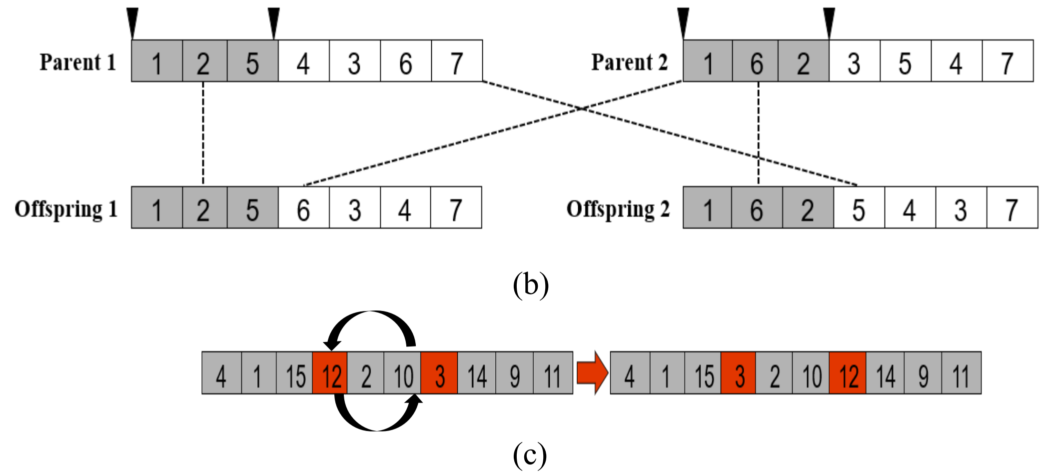


Figure 3.12:(a) Chromosome for module components; (b) two-point Crossover; and (c) swap mutation

3.4.2.3 Simulated Annealing Optimization

In this research, as illustrated in Figure 3.14, SA is implemented to find global minima for the MCMSP within a reasonable computational time. The SA process begins with an initial solution 'S' (i.e., random sequence of panels). Prior to the process of generating neighbor solution, the initial temperature (T) and cooling rate (C) are determined. It should be noted that determining an appropriate initial T and C are critical to the performance of the SA algorithm. A higher initial temperature causes the SA algorithm to spend a long period of time evaluating inferior solutions; on the other hand, a lower initial temperature leads the SA algorithm to a local optimal value. In this respect, the temperature and cooling rate value employed in the present study (1,000 and 0.95, respectively) are selected based on experimentation with various parameter values. The neighbor solution is created by randomly swapping the positions of two distinct panels in the solution string. This means that two components are randomly selected, and their sequence order is exchanged, thus forming a neighbor solution by making a small adjustment to the current solution. For example, Figure 3.13 shows the swapping of module component 2 and 1 to generate the neighbor solution.

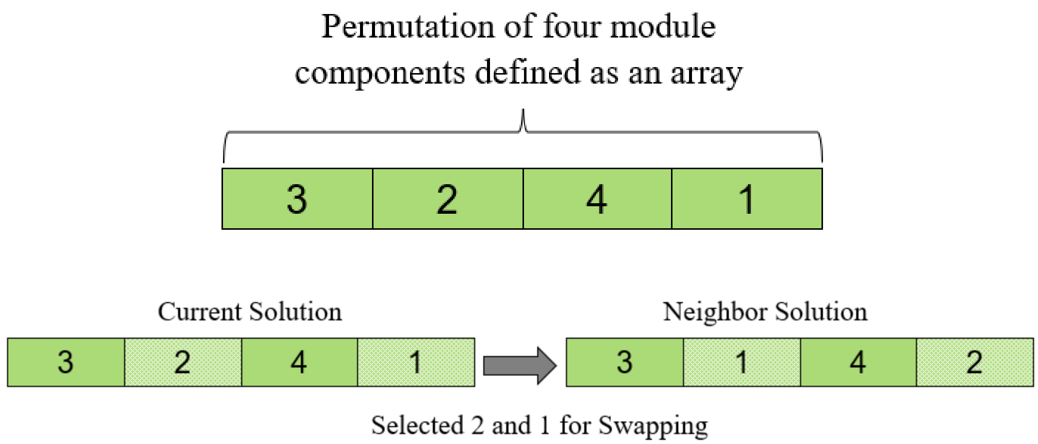


Figure 3.13: Module components defined as an array and neighborhood generation.

A neighbor solution having been generated by making a small change to the current solution, the difference in the objective function (i.e., makespan) between the current (S) and neighbor (S_n) solutions (i.e., delta = $S_n - S$) is calculated. When S_n is found to be superior to S, it is designated as the current best solution. However, if neighbor solution 'S_n' is not better than the current solution (S), S_n is not immediately rejected to allow the algorithm to explore a broader solution space. Instead, the algorithm uses the probability of acceptance (P) criterion to determine whether to accept this inferior solution (i.e., the neighbor solution), this being expressed in Equation 3.20:

$$\text{P (accept)} = exp - \frac{(S_n - S)}{T}\, w \qquad \text{Equation 3.20}$$

As an illustration, we suppose that S has a sequence of panels, 3, 2, 4, 1, resulting in a makespan of 20 min. After randomly swapping panels 2 and 1, S_n, with a sequence of 3, 1, 4, 2, has a makespan of 23 min and a T-value of 100. Although S_n has a longer makespan than S, however, it is not rejected, and the probability (P) of accepting this S_n is calculated to be 0.97. After calculating P, its value is compared to a randomly generated number between 0 and 1. It should be noted that this number is generated using a pseudo-random number generator, which provides a uniformly distributed value between 0 and 1 (i.e., every number in this range is equally likely to be produced).

This random number is crucial for the acceptance criterion step, since it is used to determine whether to accept an inferior solution (i.e., neighbor solution) based on the probability. The purpose of comparing the acceptance probability to a randomly generated number is to introduce a probabilistic acceptance mechanism that allows the algorithm to escape local optima. By occasionally accepting inferior solutions, the algorithm explores the solution space more thoroughly and avoids becoming stuck in local minima. If the acceptance probability is superior to the random value, S_n updates the current solution. It is worth noting that the temperature (T) is updated after every iteration, where an inferior solution can be accepted. The temperature is updated based on a cooling rate, which is crucial for the algorithm's performance, as it determines how quickly the temperature decreases over time, in turn influencing the probability that an inferior solution will be accepted. At higher temperatures, the probability of an inferior solution being accepted is higher, whereas, as the temperature decreases, the probability an inferior solution being accepted diminishes. The most common method of updating the temperature is to multiply the current temperature by a cooling factor (in the present study, 0.95), which controls the rate at which the temperature decreases over the number of iterations. For example, with a cooling factor of 0.95 and a current temperature (T) of 100, then the updated temperature (UT) is 95 for subsequent iterations and the probability of an inferior solution being accepted decreases. The process continues until the specified number of iterations is met. The number of iterations is finalized after experimenting with a number of different iteration values (e.g., 1,000, 1,500, and 2,000) and selecting the value that provides the best solution.

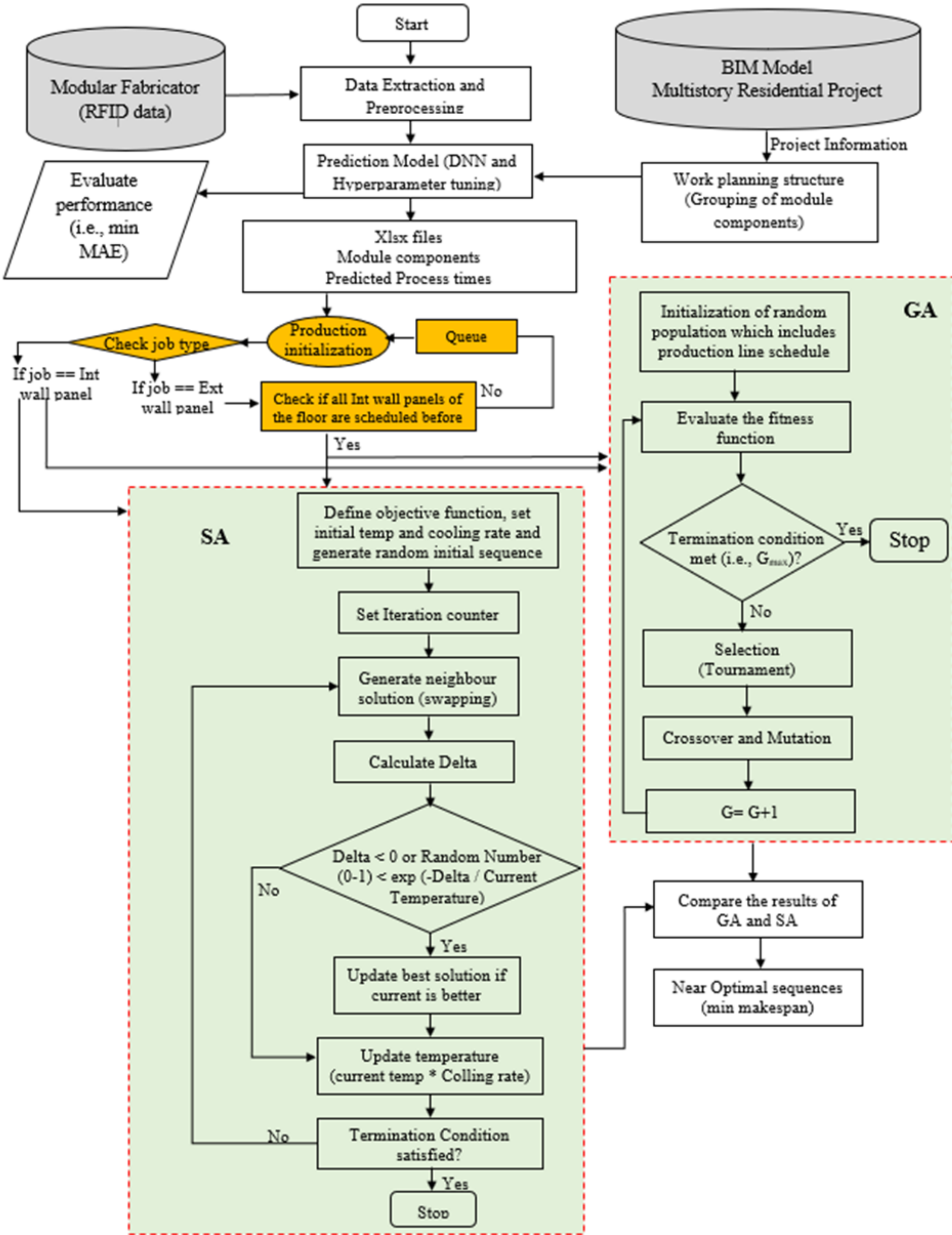


Figure 3.14: Flowchart for scheduling based on GA and SA

3.4.2.4 Hybrid GASA Optimization

GA can reserve excellent individuals for the next generation in the genetic operation process and maintain the population's diversity. However, GA is liable to converge prematurely (i.e., trapped in local optimal solutions). On the other hand, while the SA algorithm has strong local search ability, it also has a greater computational burden. As such, a hybrid algorithm combining the two (i.e., HGASA) is developed to solve the MCMSP. As shown in Figure 3.15, a sequential approach of hybrid optimization is developed in which GA and SA are used in a sequential manner (i.e., the best solution from the final population of GA is selected as the starting point for the SA phase) to solve the MCMSP. The sequential approach leverages the strengths of both algorithms GA providing diverse solutions and SA providing further refinement (i.e., fine-tuning the scheduling by exploring various small adjustments to the sequences of module components) in order minimize the makespan. First, a global search is performed by GA to explore the wide solution space. Global search, it should be noted, refers to the algorithm exploring a wide and diverse range of potential solutions across the entire solution space. GA randomly generates the initial solution, as explained in section 3.3.2, then evaluates the initial population and operates on the population using three genetic operators (i.e., selection, crossover, and mutation) to produce a new population, which allows the algorithm to select better solutions. Once the termination condition is met (generation size is defined as 200 in the present study) and the GA has identified an optimal solution, it sends the best individual solution (sequence of module components) to SA for further improvement (i.e., reducing the makespan duration). In other words, a local search is carried out through SA in order to fine-tune the solution provided by GA. Local search in the context of the present study refers to the process of exploring and refining the solutions within the immediate neighborhood of a current solution. A probabilistic acceptance criterion, described in detail in section 3.3.3, is used to decide whether to move to a new solution. It is worth noting that the initial steps of the HGASA are similar to those in GA (i.e., generating a random initial solution and using selection crossover and mutation to find the optimal solution). The main difference is that SA, rather than starting with a random initial solution, starts with the optimal solution generated by GA and performs a local search on this solution rather than searching all individuals. This process continues until the termination condition of the SA algorithm (i.e., the maximum number of iterations) has been met.

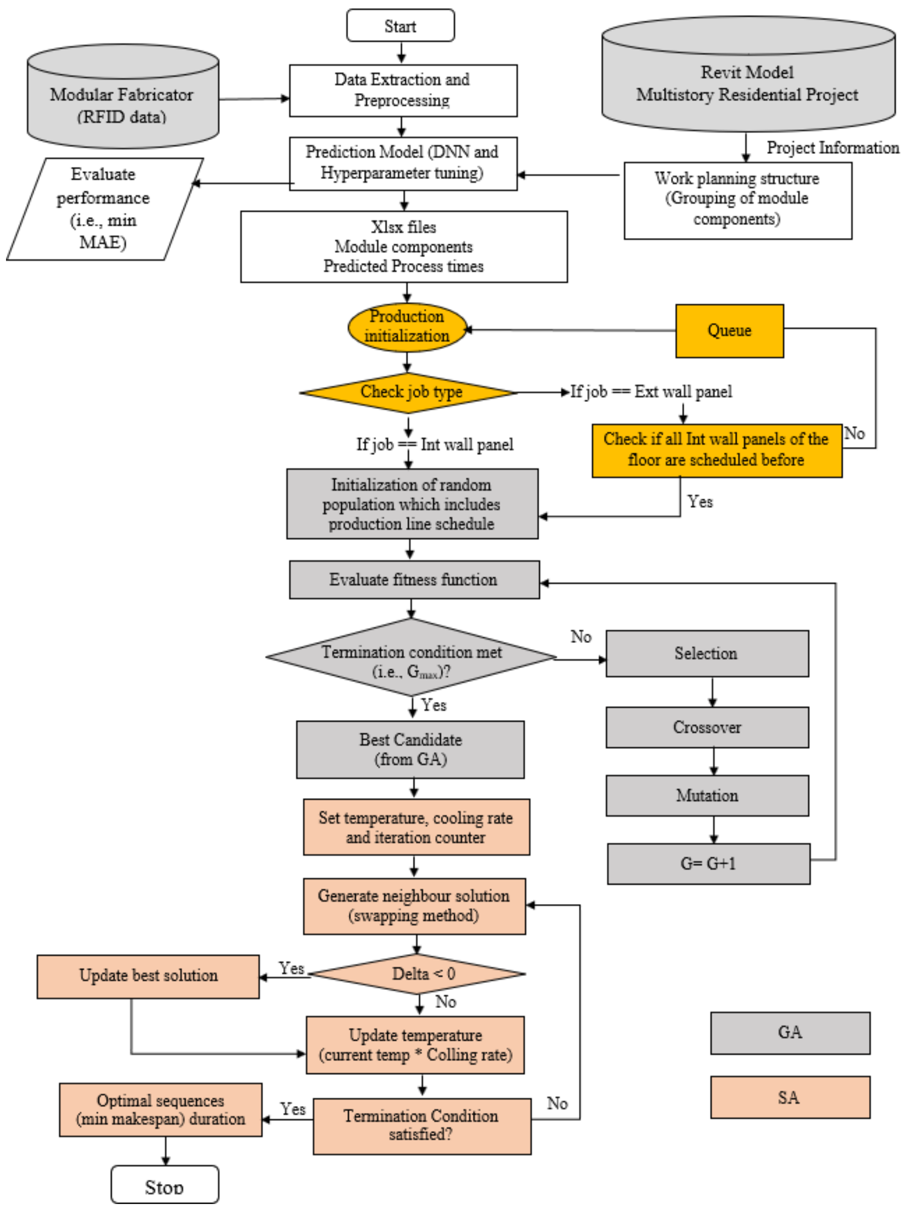


Figure 3.15: Flowchart for scheduling based on Hybrid GASA

3.4.3 Comparison Matrix

A comparison matrix is developed to evaluate and compare different algorithms based on specific performance metrics. These matrices help in understanding the effectiveness of GA, SA and Hybrid GASA, making them useful in selecting the best optimization algorithm for planning and scheduling in panelized construction manufacturing. In this respect, this thesis used the following metrics (Shavari et al. 2022; Jain and Meeran 1999): (i) makespan, which is the total time required to complete all panels of the multiple projects; (ii) computational speed, which shows convergence time of the algorithm in order to provide optimal solution, which can be defined as slow, medium and fast; *(iii)* complexity, which is ease of implementing the algorithm and depends on the number of parameters in the algorithm, the less number of parameters the ease in implementation*; (iv)* production load balancing (P_{LB}) represented in Equation 3.21, which reflects how the algorithm handles the work packages consisting of different number of wall panels across the production line and a ratio closer to 1 indicates better load balancing. It is worth mentioning that P_{LB} developed in this thesis is modified based on the load balancing formula available in the literature (Pinedo 2016); and *(v)* parallel workstation (P_w) represented in Equation 3.22, which measures the percentage reduction in makespan after adding parallel workstation, for the GA, SA and hybrid algorithm.

$$P_{LB} = \frac{Max\ (WP_{mt})}{Min\ (WP_{mt})} \qquad \text{Equation 3.21}$$

where WP_{mt} = makespan of a work package

$$P_w = \frac{(I_{mt} - P_{mt})}{I_{mt}} * 100 \qquad \text{Equation 3.22}$$

where I_{mt} and P_{mt} = makespan before and after adding parallel workstation, respectively.

CHAPTER 4: CASE STUDY AND RESULTS

4.1 Introduction

This chapter outlines the implementation of the developed method in two modular construction factories, each using different material (i.e., wood and light gauge steel) for their module components. Additionally, the production line in these prefabrication factories consist of a varying number of workstations. These differences introduce complexities due to distinct design specifications required to meet client demands, necessitating different standard operating procedures (SOPs) at each workstation. This variation results in different process times, leading to various types of lean manufacturing wastes (e.g., waiting and overproduction) in the production line. Table 4.1 presents two types of case studies: *(i)* case study I focus on a production line that prefabricates light gauge steel-based wall panels featuring 4 workstations (i.e., assembly, framing, sheathing and exterior finishing); and *(ii)* case study II involves a production line that manufactures wood-based wall panels with 6 workstations (i.e., framing, sheathing, nailing, cutting, window/door and wall magazine). Given the changing SOPs and design specifications, implementing the developed method on just one type of prefabrication factory and production line is insufficient. In this respect, to generalize the developed method, two different type of production lines are utilized. This chapter presents an overview of the two-wall production process, followed by details of data collection process, development of predictive method, discrete event simulation method creation and implementation of optimization algorithms for the purpose of planning and scheduling.

Table 4.1: Production line and module components

Production line and material type	**Number of workstations**	**Module components (wall panel)**
Case study I- Wall panel production line Light gauge steel-based wall panel	4 workstations (assembly, framing, sheathing and exterior finishing)	
Case study II- Wall panel production line Wood-based wall panel	6 workstations (framing, sheathing, nailing, cutting, window/door and wall magazine)	

4.2 Case Study I

4.2.1 Description of the Case Study

The developed method was implemented on a wall panel production line of a modular fabricator in Edmonton, Canada as shown in Figure 4.1. The industry partner produces both interior and exterior light gauge steel (LGS) wall panels on a production line consisting of three main workstations: (i) assembly station; (ii) framing station; (iii) sheathing station and; (iv) exterior finishing. Additionally, the production line consist of a rim tracks sub-station where track studs are installed on the panels to secure the wall studs. The industry partner produces various types of LGS wall panels for six-storey residential buildings, each comprising 120 units and more than 1,500 wall panels with varying design specifications. Workstation process times (200 series of time data for each three workstation) were collected and stored in a database using the C-track app. The names of wall panels, the number of workers at workstations, and module design specifications such as the number of studs and total area of windows and doors (in sq. ft) were stored manually into the database.

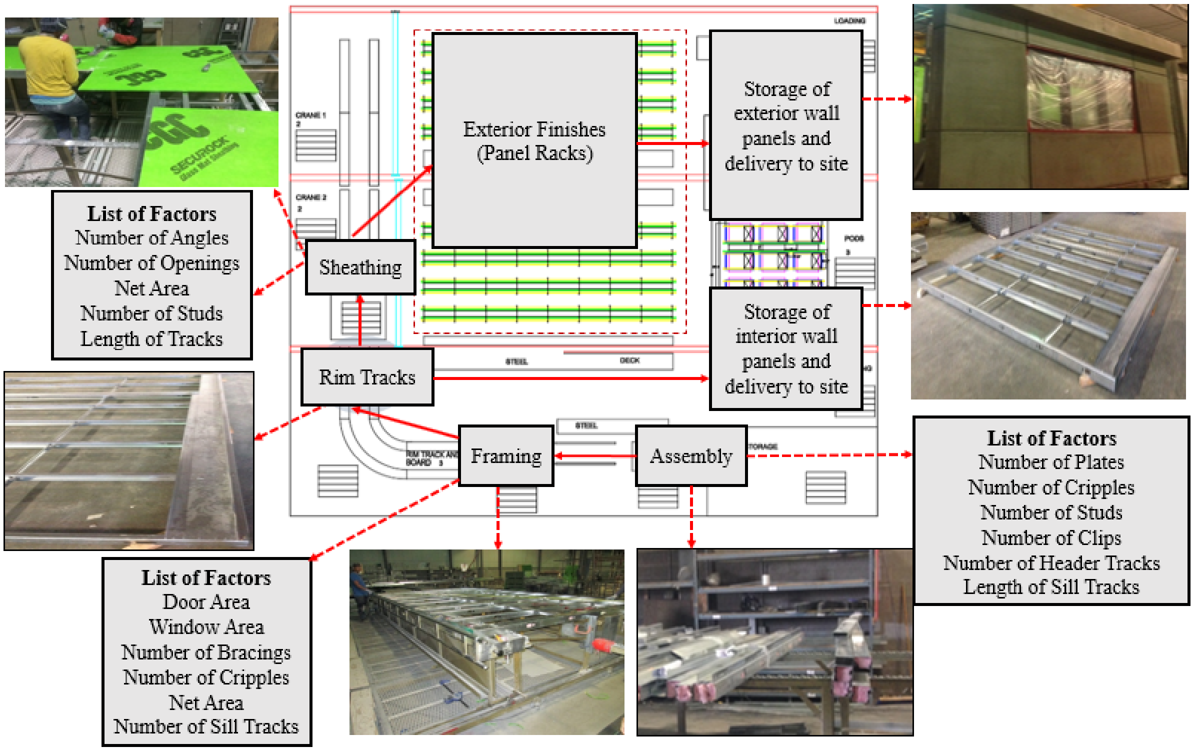


Figure 4.1: Light gauge steel (LGS) wall panel production line

Assembly workstation

The wall panel production starts at the assembly workstation (Figure 4.2). Workers begin by reviewing shop drawings to verify wall panel components such as types and numbers of studs, cripples, and tracks. Next, they take wall panel components from the floor stock and place them on the workstation table. The process includes marking, aligning, installing clips, cutting studs for connections, and assembling components for king studs, king tracks, sills, and cripples as specified in the shop drawings. These prepared components are marked with their wall panel numbers (e.g., Int 301), wrapped, and moved as bundles to the framing station. The assembly workstation also features a cutting machine for cutting cripples and bracings to the sizes specified in the shop drawings.

Figure 4.2: Assembly workstation

Framing workstation

At the framing station, workers manually fasten wall components together to form wall panels (Figure 4.3). They start by using measuring tape and a triangular ruler to mark the locations of vertical studs on the top and bottom tracks. The tracks and studs are placed on a framing table, and the tracks are closed after positioning the studs to align with the marked locations. Workers measure the dimensions from the top to the bottom tracks to ensure the studs are correctly positioned. Studs are then erected between the top and bottom tracks, and nail guns are used to secure the studs to the tracks. Since the work is done manually, workers use leveling tools to check for any fastening errors.

Figure 4.3: Framing workstation

Sheathing workstation

Figure 4.4 illustrates the sheathing workstation where gypsum drywall is manually cut and installed on exterior walls, secured using nail guns. For openings such as doors and windows, a 6-inch clearance is maintained between the tracks of the opening area and the drywall. Workers ensure the drywall is leveled around the corners of the panel and the opening spaces. Additionally, any exterior wall panels requiring angles are prepared at the sheathing workstation before the drywall installation begins.

Figure 4.4: Sheathing workstation

Exterior Finishes workstation

At the exterior finishing workstation, finishing tasks are performed on the exterior wall panels (Figure 4.5). The process includes: (i) applying a thin layer of cement paste for waterproofing and installing electrical back boxes; (ii) fixing foam panels, cutting out holes for utilities, milling grooves, and applying spray foam to seal gaps between foams; (iii) applying a base coat and primer; and (iv) applying the final coat and allowing it to dry for 3-4 hours. This workstation also functions as a temporary storage area where wall panels are held before being delivered to construction sites.

Figure 4.5: Exterior finishing and storage workstation

4.2.2 Analysis and Discussion of Results

The 200 series of time data and impact factors of each of the three main workstations (i.e., the assembly station, the framing station, and the sheathing station) were used for the data analysis. The time-series data collected, though it would be considered a relatively small dataset, provided useful insights about the production line, as illustrated below. Looking at the results of the data analysis, the workstation process times were found to vary depending on the wall panel design specifications, even when the allocation of tasks and workers at workstations did not change. For example, as shown in Figure 4.6 (a), the process times at the assembly workstation ranged from 7 to 99 minutes depending on the number of wall panel design factors (e.g., the number of studs, plates, and clips). The data in the figure also indicates that at this particular workstation all the listed wall panel design factors affected the process times. For example, the number of plates in the wall panels ranged from 5 to 80, and this factor, as the figure clearly shows, strongly affected the process times. However, as shown in Figure 4.6 (b) and Figure 4.6 (c), it was found that some design factors (e.g., number of tracks and window area) did not affect the process times at some workstations. In other words, these factors do not have a significant relationship with the workstation process times. For instance, as shown in Figure 4.6 (b), the number of tracks required was 2 to 3 for all types of wall panels; however, the process times were found to vary from 11 to 102 minutes. As illustrated in Figure 4.6 (d), there was found to be a high level of variance in the process times at workstations due to the influence of these design factors. For example, although there was an SOP at the assembly workstation, the process times ranged from 23 to 100 minutes, leading to reduced productivity due to an imbalanced production line. As described above, this imbalance can be reduced by planning more efficient sequencing of modules and allocation of labor based on historical production data. In this respect, the next critical step in the case study would be to identify the SIFs among the design factors that highly influence the process times.

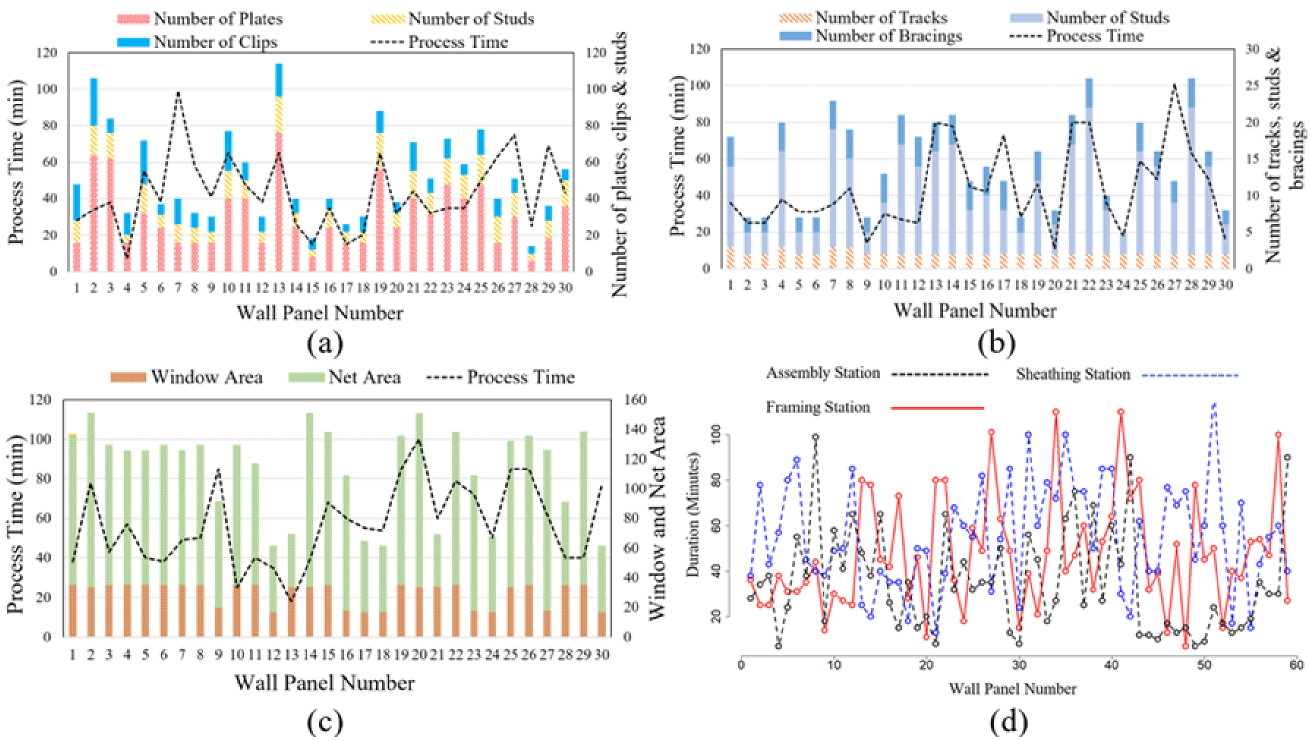


Figure 4.6: (a,b,c): Effect of design specifications on process time, and (d) average productivity of workstation

Prior to identifying the SIFs, though, data pre-processing was implemented to identify and remove the outliers from the raw dataset based on the scatter plot and the standardized residual test. Based on the expert opinion of the industry partner's production manager, process times >80 minutes were removed from the scatter plot, given that such data points are indicative of a work disruption such as a delay due to errors in the shop drawings. After this, the standardized residuals test was implemented to identify hidden outliers. The data points with standardized residuals outside the range of ±1.64 were considered outliers and were removed from the dataset. As a result of the pre-processing task, the datasets numbered 178, 180, and 178 for the assembly, framing, and sheathing workstations, respectively. These data were normalized using a min–max normalization technique to transform the integer values into values ranging between 0 and 1. A t-test, PCC, and PCA were implemented to find the SIFs, with these, in turn, serving as the primary input in the development of the MLR. The results of the t-test and PCC are represented in Table 4.2. With the significance level (p-value) defined as 1%, the different workstations were found to have different SIFs. For example, for the assembly workstation, six design factors (number of: header foam, studs, stud's foam, plates, clips, and openings) were identified as SIFs. However, the SIFs at the sheathing station were track length, number of angles, and window and door area. In terms of PCC, the design factors were selected by examining the correlation coefficient, which represents the relationship between the dependent variable (duration) and the independent variables. Design factors with a correlation coefficient >0.65 as defined by the experiments were deemed to be SIFs. Based on this, for the assembly workstation, four factors (number of: studs, studs' foam, cripples, and openings) had a coefficient >0.65 and thus were identified as SIFs.

Table 4.2: Results of t-test and Pearson correlation coefficient

Assembly Station			Framing Station			Sheathing Station		
Factors	P value	Corr.	Factors	P value	Corr.	Factors	P value	Corr.
# of HeaderTrack	0.02977	0.57	# of HeaderTrack	2.08e-09	0.77	TrackLength	1.88e-11	0.71
# of SillTrack	0.05807	0.59	# of SillTrack	2.78e-12	0.58	# of Studs	0.697	0.73
# of HeaderFoam	0.00212	0.40	# of Studs	0.434	0.76	WindowArea	2.17e-07	0.42
# of Studs	1.40e-05	0.82	# of Bracings	0.007	0.71	Door Area	0.0002	0.05
# of StudsFoam	0.00141	0.74	Net Area	6.99e-05	0.18	Net Area	0.993	0.68

# of Plates	0.00393	0.62	# of Openings	0.69	0.70	# of Angles	1.61e-09	0.48
# of Clips	0.00086	0.36	-	-	-	# of Openings	0.941	0.39
# of Cripples	0.9381	0.66	-	-	-	-	-	-
# of Openings	3.53e-05	0.69	-	-	-	-	-	-

In addition to the t-test and PCC, PCA was applied. The first step in the PCA was to determine the percentage of variance for PCs, selecting for further analysis the set of PCs that cumulatively accounted for 90% of the dataset's total variation. For example, for the assembly station, the minimum set of PCs accounting for ≥90% of the cumulative variance was the set of PC1, PC2, PC3, and PC4, representing 61.5%, 14.1%, 9.56%, and 6.9%, respectively, of the variation, for a cumulative variation of 92.1%. These components were used as the basis for identifying SIFs, where PC1 was found to be highly correlated with the number of openings, PC2 with the number of clips, PC3 with the number of studs, and PC4 with the number of plates as shown in Table 4.3.

Table 4.3: A result of principle component analysis for assembly station

Factors	PC1	PC2	PC3	PC4
No. of Header Tracks	0.138	0.334	0.108	0.303
No. of Sill Tracks	0.177	0.377	0.066	0.282
No. of Header Foam	0.073	0.348	0.181	0.376
No. of Studs	0.355	0.288	0.676	0.207
No. of Studs Foam	0.391	0.113	0.374	0.329
No. of Plates	0.317	0.098	0.324	0.428
No. of Clips	0.075	0.535	0.318	0.339
No. of Cripples	0.305	0.391	0.041	0.019

No. of Openings	0.681	0.198	0.381	0.083
No. of Workers	0.033	0.196	0.022	0.478

Based on the training dataset (80%) and the lists of SIFs generated by the different analyses, MLR models were developed and validated using a K-fold cross-validation. The training dataset was randomly split into 10 folds, with one-fold used for the testing set and the remaining 9 folds used as a training set. The cross-validation process was repeated against all 10 folds in the dataset, and the average evaluation indices were calculated. The best predictive model was then selected based on four performance indices: R^2, adj. R^2, RMSEs, and MAE. As observed in Table 4.4, for the assembly station, the R^2, adj. R^2, RMSE, and MAE values in the training and testing datasets were 80.1%, 74.33%, 79.08%, 71.01%, 7.93 min, 9.83 min, 5.82 min, and 7.16 min, respectively, when the SIFs were the number of: header foam, studs, stud's foam, plates, clips, and openings identified by the t-test. The R2 value for testing was 80.1%, meaning that the model was found to predict 80% of the outcomes. The RMSE value for testing depicted the deviation of 9.83 minutes between predicted and actual duration. However, for the framing station, the SIFs identified by the correlation test were selected instead of those identified in the t-test, since, for this station, the RMSE and MAE values generated by the correlation test were lower than those generated by the t-test.

Table 4.4: Results of multiple regression at workstations

Model	R square (%)		Adj. R square (%)		RMSE (min)		MAE (min)	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing
			Assembly Station					
t-test	80.1	74.33	79.08	71.01	7.93	9.83	5.82	7.16
Correlation	79.86	71.38	78.83	70.07	7.97	9.88	6.32	7.64

PCA	74.35	71.23	73.27	69.79	9.00	10.59	7.033	7.95
				Framing Station				
t-test	78.09	76.56	77.36	74.31	7.09	8.10	5.78	6.01
Correlation	79.13	74.55	78.33	73.48	6.89	6.99	5.56	5.89
PCA	64.74	57.01	63.87	51.64	8.37	10.20	6.58	8.46
				Sheathing Station				
t-test	74.90	73.35	72.37	71.26	9.04	9.54	7.56	8.19
Correlation	61.91	49.66	60.56	46.93	11.77	13.80	9.66	11.20
PCA	57.59	51.7	55.7	50.01	10.88	12.40	8.32	10.66

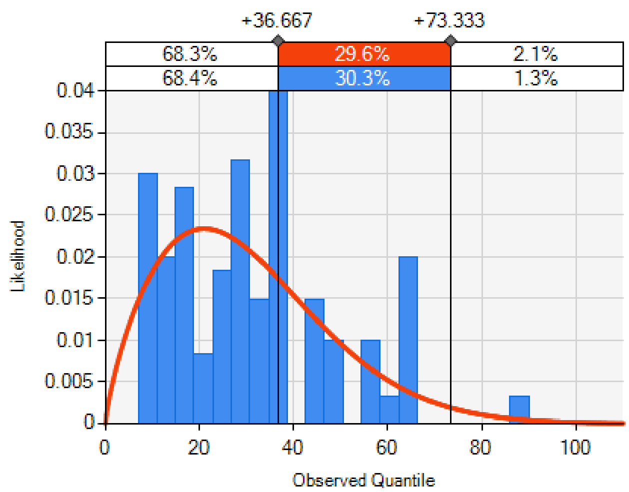
Table 4.5 represents the cycle time formula (CTF) of workstations, which involves the coefficient values of SIFs to predict the process times of workstations. As can be seen, some of the coefficients (Coef.) of SIFs had positive values while others were negative. For example, for the assembly station, the XHF, XS, and XC coefficients indicate that the process times increased when the XHF, XS, and XC in the wall panel increased. Similarly, the coefficient of XW means that the process times decreased whenever additional workers were allocated to the station. To improve accuracy, probability distribution function-based models were developed based on the time data used in the statistical analysis. To identify the most suitable probability distribution functions, as shown in Table 4.5, a goodness-of-fit method, consisting of both a Kolmogorov–Smirnov (K–S) test and a chi-square test, was performed. Weibull, gamma, and triangular distributions were selected as the best fit for the assembly, framing, and sheathing stations, respectively. These distributions, along with cycle time formula (CTF), were used as inputs in the simulation model.

Table 4.5: Cycle time formula and summary of the goodness-of-fit test

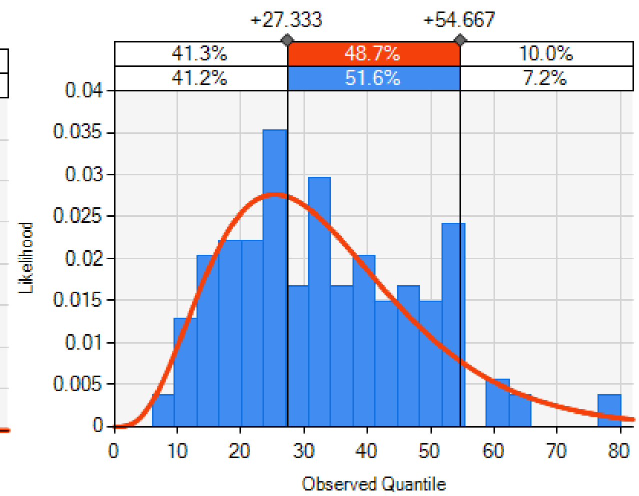
Cycle time formula					
Assembly Station	Coef.	Sheathing Station	Coef.	Framing Station	Coef.
Number of Header Foam (X_{HF})	20.54	Track Length (X_{TL})	70.32	Number of Header Track (X_{HS})	15.14
Number of Studs (X_S)	38.36	Number of Angles (X_A)	15.63	Number of Studs (X_S)	20.78
Number of Plates (X_P)	−21.49	Window Area (X_{WA})	−36.47	Number of Bracings (X_B)	7.43
Number of Clips (X_C)	8.74	Door Area (X_D)	−20.72	Number of Openings (X_o)	0.05
Number of Workers (X_W)	−16.19	Number of Workers (X_W)	−7.27	Number of Workers (X_W)	−11.3
Number of Studs foam (X_{SF})	27.75	-	-	-	-

Goodness-of-fit test								
Assembly Station			Framing Station			Sheathing Station		
Distribution	K-S	X^2	Distribution	K-S	X^2	Distribution	K-S	X^2
Weibull	0.0818	45.47	Gamma	0.0543	26.80	Triangular	0.0752	27.59
Triangular	0.0915	64.26	Triangular	0.0558	30.83	Uniform	0.0849	16.29
Gamma	0.0916	41.97	Weibull	0.0571	27.35	Normal	0.0891	31.17
Normal	0.1052	64.26	Lognormal	0.0678	28.82	Weibull	0.0958	22.57
Lognormal	0.1063	60.21	Uniform	0.0701	35.04	Gamma	0.1074	42.47

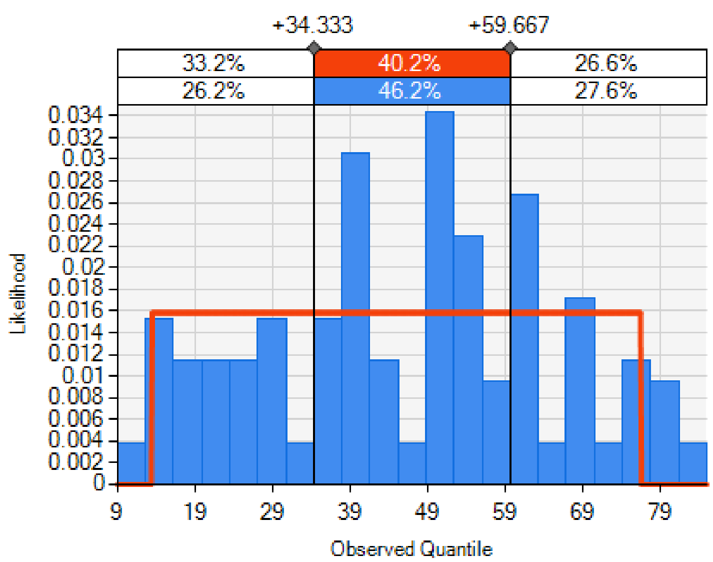
Figure 4.7 shows the probabilistic density charts of the panel processing time at the assembly, framing and sheathing workstations. For example: (i) Weibull distribution at assembly workstation (4.7 a) with parameters: shape (1.75) and scale (33.87); (ii) Gamma distribution at framing workstation (4.7 b) with parameters: shape (4.25) and scale (7.79); (iii) Uniform distribution at sheathing workstation (4.7 c) with parameters: low (13.36) and high (76.46).



(a)



(b)



(c)

Figure 4.7: Probabilistic density chart (a) assembly; (b) framing and (c) sheathing workstation

Based on the process flow of the production line, the simulation model illustrated in Figure 4.8 was built in Simphony.NET. The simulation uses the 'database' element to update wall panel information from a central database using SQL. In the simulation model, 'Database Create,' 'task,' and 'conditional branch,' 'resource', and 'destroy' elements from the general template are used to mimic the process flow of the actual production line. Depending on the type of wall (interior or exterior), different process flows were required. In the model, the 'Database Create' element generates the simulation entities (wall panels) and passes them into the next task elements, which are the assembly station and framing station. After completing the tasks at the framing station, the wall panel components proceed to the 'conditional branch', where they are classified as components of either interior or exterior walls (since these types of walls require different tasks). From the 'conditional branch element' pertinent were related to types of wall panels, interior wall panel components are directed to 'storage' while exterior wall panel components are directed to the 'sheathing station'. Once the process is complete at each workstation, the given entity is destroyed using the 'destroy' element. The simulation was run using cycle time formula (CTFs) and probability distribution functions as inputs, and the results in terms of productivity were compared with the historical actual productivity data for validation purposes. It should be noted that 2 and 3 workers were allocated to the assembly and framing stations, respectively, for interior wall panels, while 3, 2, and 3 workers were allocated to the assembly, framing, and sheathing stations, respectively, for exterior wall panels.

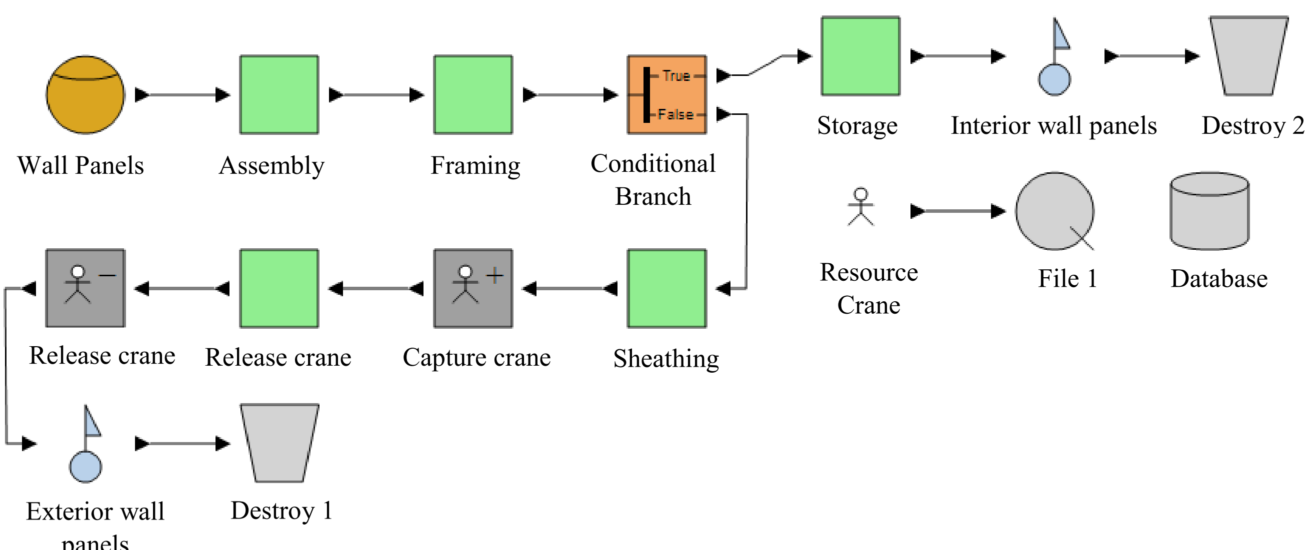


Figure 4.8: Simulation of a production line

Figure 4.9 shows a comparison of the actual(historical) and simulated (cycle time formula and probability distribution functions) cycle times for producing interior wall panels. As can be seen, the actual cumulative cycle time for the manufacture of 200 interior and exterior wall panels was 119 hr, compared to 106.38 hr for the cumulative cycle time formula and 90.1 hr for the cumulative probability distribution functions. It should be noted that cumulative cycle time is the total sum of the process times of all wall panels, while cumulative probability distribution functions is the total process time to complete the wall panels with probability distribution as an input in the simulation.

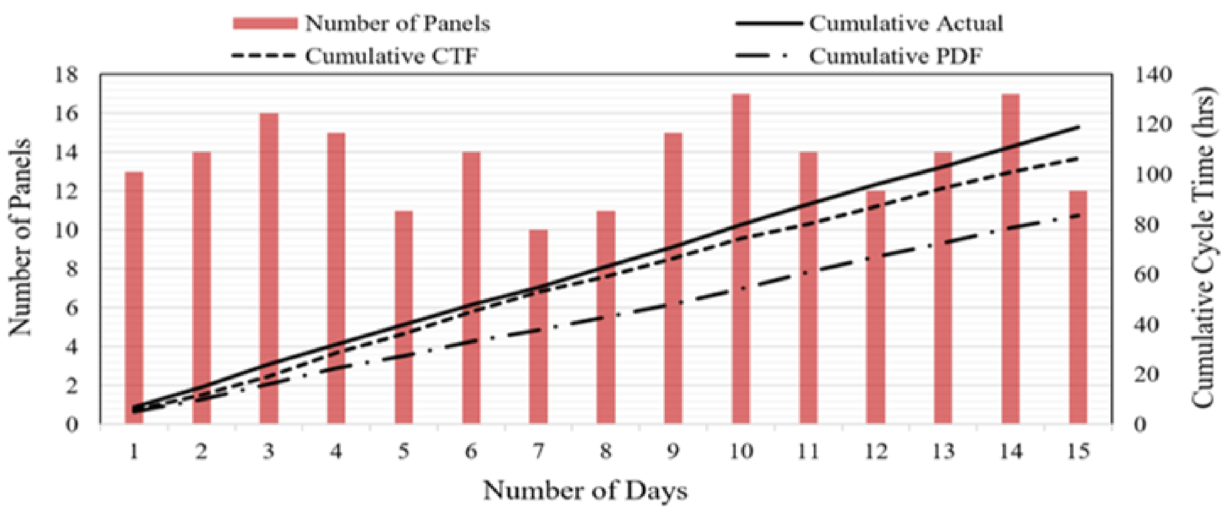


Figure 4.9: Comparison between cumulative actual and simulated cycle time

These results are indicative of general agreement between the actual and simulated cumulative process times generated by the CTF (i.e., 89.39% accuracy). In this respect, the simulation method developed and implemented in conjunction with cycle time formula (CTF) in this case study was deemed to be a reliable predictive method. After validating the production rate generated, multiple simulation-based scenarios were developed and analyzed using cycle time formula (CTF).

To provide better planning scenarios, the fabrication of 309 wall panels in five different panel-sequencing scenarios, each with the same labor allocations at the various workstations, was simulated. In the first scenario, the units of a residential project were selected in a clockwise direction and manufactured unit-by-unit, and their wall panels (exterior or interior) were prefabricated randomly in the production line, which represents as the current state of the case study company. In the second scenario, the units of the residential project were again selected in a clockwise direction and manufactured unit-by-unit, but this time the exterior wall panels of a given unit were prefabricated first, followed by the corresponding interior walls for the unit. In the third scenario, the interior wall panels of a given unit were manufactured first, followed by the exterior wall panels. In the fourth scenario, the interior wall panels of all units were produced first, followed by the exterior wall panels of all units. In the fifth scenario, the exterior wall panels of all units were produced first, followed by the interior wall panels of all units. Based on these scenarios, the simulation method provided the cumulative duration to prefabricate all wall panels in each of the five scenarios, yielding cumulative durations of 93.52 hr, 51.92 hr, 69.76 hr, 59.84 hr, and 60.64 hr for the five respective scenarios. Thus, Scenario 2 was found to outperform the other scenarios. In terms of production planning, finding the optimum allocation of labor to workstations is crucial in that it is a critical factor in (i) synchronizing process times; (ii) reducing waiting times along the production line, and (iii) increasing productivity. In this respect, as shown in Table 4.6, in the case study we experimented with different numbers of workers as an additional decision variable in the simulation method built based on Scenario 2. Workers were allocated to workstations according to the type of wall panel, and Scenario 2.3 was found to outperform the other labor allocation scenarios, with a cumulative duration of 44.42 hr. It is notable that the duration in Scenario 2.4 was found to be longer than that in Scenario 2.3 even though the number of workers at the framing and sheathing station was increased in Scenario 2.4. This may have been attributable to space congestion at the framing station disrupting the coordination between the workers completing the work. In this regard, Zhang et al. (2020) have demonstrated that space congestion interrupts the workflow in MCM, thereby reducing productivity at the workstation level. As such, this simulation-based statistical method for production planning in MCM is a significant tool for analyzing the effect of different crew sizes, especially in that it demonstrates that it is not always ideal to increase the number of workers at a workstation to reduce process times.

Table 4.6: Comparison of different crew sizes at workstations

Scenario	Duration (hours)	Number of workers				
		Interior Wall Panels		Exterior Wall Panels		
		Assembly	Framing	Assembly	Framing	Sheathing
2.1	51.92	2	3	3	2	3
2.2	45.44	2	3	3	3	3
2.3	44.42	2	4	3	2	2
2.4	45.42	2	4	3	3	3
2.5	51.04	2	2	3	2	3
2.6	51.04	2	2	3	3	3

The case study demonstrates that optimal sequencing of modules and allocation of workers is critical to improving productivity. The method described in this study is also capable of identifying the SIFs affecting fabrication process times, thereby removing the guesswork from production planning. However, it should be noted that the SIFs in this study are a function of the given product design specifications and tasks performed at the workstations of the case study production line. For other cases, practitioners would need to modify the data analysis phase based on the given design specifications in order to identify the SIFs. In modular construction, due to unpredictable demand, production managers must frequently alter their plans to accommodate change orders. As an alternative to this challenging and error-prone approach, the method implemented in the case study can be deployed to devise different production line scenarios in terms of labor allocation and sequencing to streamline the MCM planning process. However, this method does not provide the optimal solutions and the data used to train the prediction method consist of 200 series of time data, which is not considered as a large dataset. Also, there is a need to manually test the simulation scenarios. Therefore, future work will seek to enhance the performance of the developed method. This will include, for one, applying optimization to identify sequencing of modules more efficiently and rapidly. Second, to develop generalized predictive method for forecasting process time at workstations using large dataset for the industrialized construction. In this respect, the deep neural network and optimization-based method was developed and implemented in the wood-based wall panel production line.

4.3 Case Study II

4.3.1 Description of the Case Study

The developed method was implemented on a wood frame wall panel production line operated by a modular fabricator in Edmonton, Canada. The workstations of the production line are equipped with computer numerical control (CNC) machines and some workstations require manual works. Figure 4.10 represents the wall panel production line consists of framing, sheathing, nailing, window installation and loading. The details of operations regarding each workstation is as follows: (i) framing station where the wall components such as studs, cripples, and sill plates are fastened together to form an interior and exterior wall panel frame using CNC machine; (ii) sheathing station where drywalls are cut manually and placed on wall panels; (iii) multifunction bridge where drywalls are nailed using CNC machine and moved to next workstation using transfer cart at butterfly station for other activities. It should be noted that interior wall panels are moved from the multifunction bridge to the window bypass line to store them at the wall magazine line and exterior wall panels are transferred to the window/door line; and (iv) window/door installation lines, where windows/doors are installed on the wall panels and transferred to the storage area (i.e., wall magazine line) as they await delivery to on-site using trucks and trailers.

The process times of wall panels at each workstation were collected by the modular fabricator company using an RFID system (i.e., RFID printer, tag, antenna and reader). The worker at the first workstation (i.e., framing workstation) of the wall panel production line attaches the passive tags to each wall panel and the RFID antennas that are located at the entrance of the workstations captures the movement of wall panels through the read-zone (i.e., antennas captures the tag signal as the given wall panel passes through the workstation). The antennas were connected to the RFID reader, which transfers the captured timestamps automatically into the database. The timestamps data for the workstations was received from the case study company in the form of 'RFID raw data' file. The data covers the period between July 2015 and August 2018. The data file contains: (i) 416950 timestamps (i.e., start and finish time) for wall panels along with the workstations. This includes total 10 attributes such as tagID, panel number, antenna description; and (ii) design attributes of each wall panel (e.g., number of cripples, number of doors, and width of wall panels) with total of total 39703 records. This includes total 37 attributes such as floor, Dstud, Mstud and Drillhole. Considering the RFID data, the next critical step in the case study was to extract the process times of wall panels at the workstation from the 'RFID raw data' file based on difference between starting time of the wall panel at the consecutive antenna locations (i.e., workstations) as expressed in Eq. (3.6).

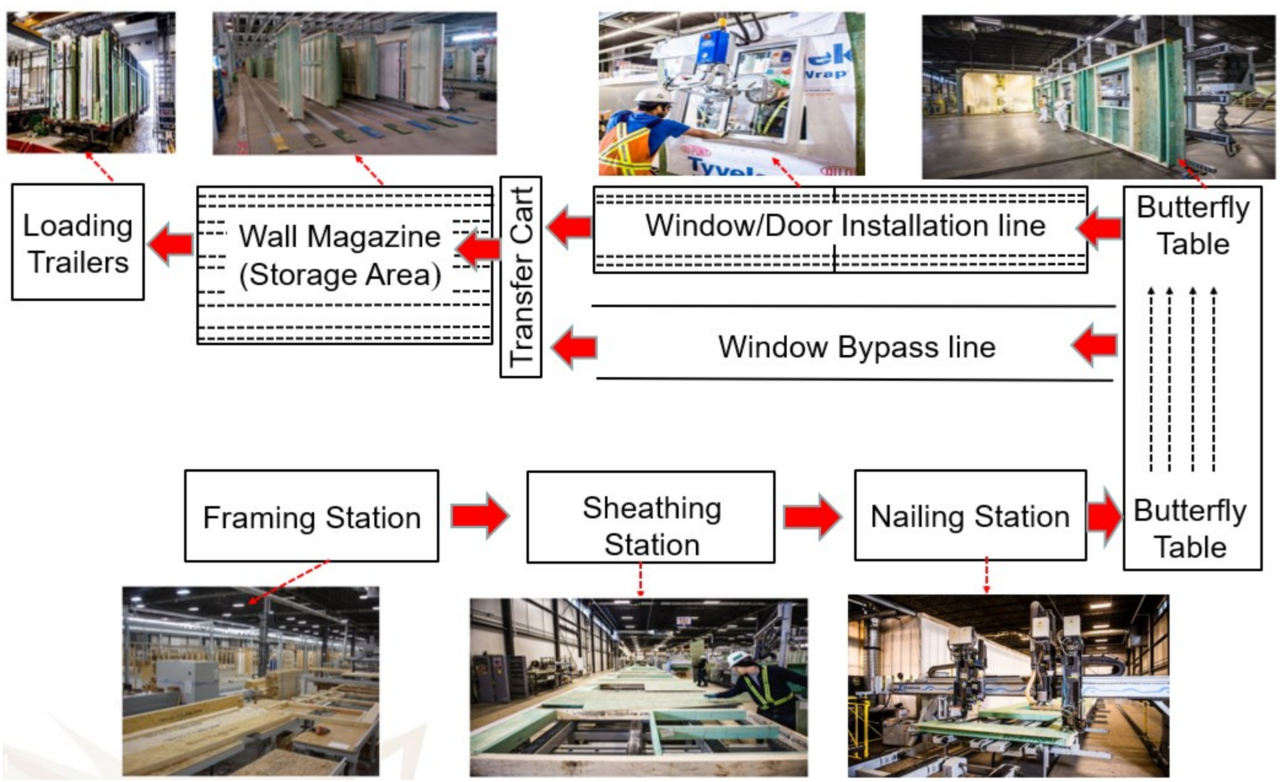


Figure 4.10: Wood-based wall panel production line

Framing workstation

Wall panel production starts at the framing station, Figure 4.11, where workers load the studs into the machine in sequence, after which the machine nails the studs to the top and bottom plates of the exterior and interior wall panels.

Figure 4.11: Framing workstation case study II (picture provided by Dr. Sadiq Altaf)

Sheathing workstation

After leaving the framing station, wall panel is moved to sheathing station, as shown in Figure 4.12, where sheathing boards are placed, correctly positioned, and manually nailed. Additionally, here one or more workers perform several manual tasks. These tasks include correcting any errors from the framing machine, installing backing and other support materials, and marking the wall panel with its name. Finally, the wall panel is moved to the nailing workstation for machine nailing.

Figure 4.12: Sheathing workstation case study II (picture provided by Dr. Sadiq Altaf)

Nailing workstation

At nailing workstation shown in Figure 4.13, sheathing boards are automatically and securely fastened to the wall panel studs by a CNC machine equipped with nail guns mounted on a moving bridge. The process requires just one worker to bring the panel from the sheathing station, start the nailing process, and move the wall panel to next workstation.

Figure 4.13: Nailing workstation (picture provided by Dr. Sadiq Altaf)

4.3.2 Analysis and Discussion of Results

The process times of wall panels at each workstation were collected by the modular fabricator company using an RFID system (i.e., RFID printer, tag, antenna and reader). The worker at the first workstation (i.e., framing workstation) of the wall panel production line attaches the passive tags to each wall panel and the RFID antennas that are located at the entrance of the workstations captures the movement of wall panels through the read-zone (i.e., antennas captures the tag signal as the given wall panel passes through the workstation). The antennas were connected to the RFID reader, which transfers the captured timestamps automatically into the database. The timestamps data for the workstations was received from the case study company in the form of 'RFID raw data' file. The data covers the period between July 2015 and August 2018. The data file contains: (i) 416950 timestamps (i.e., start and finish time) for wall panels along with the workstations. This includes total 10 attributes such as tagID, panel number, antenna description; and (ii) design attributes of each wall panel as shown in table 4.7 (e.g., number of studs, number of doors, and width of wall panels) with total of total 39703 records. This includes total 37 attributes such as floor, Dstud, Mstud and Drillhole. Considering the RFID data, the next critical step in the case study was to extract the process times of wall panels at the workstation from the 'RFID raw data' file based on difference between starting time of the wall panel at the consecutive antenna locations (i.e., workstations).

Table 4.7: Summary of wall panels data from the production line

Panel	Type	Job	Length	Width	Stud	LStud	MStud	DStud	Windows	Door	LargeDoor	Duration
E-16_30DES-17-10-11_10	EXT	30DES-17-10-11	11919	2467	22	5	0	0	0	0	0	13
E-17_30DES-17-10-11_10	EXT	30DES-17-10-11	12175	2467	23	5	0	0	0	0	0	10
E-18_30DES-17-10-11_10	EXT	30DES-17-10-11	12168	2467	14	5	0	0	3	0	0	9
I-2_10GLR-17-0016_00	INT	10GLR-17-0016	12035	2467	24	1	0	0	0	2	0	6
I-3_10GLR-17-0016_00	INT	10GLR-17-0016	12195	2467	25	1	0	0	0	3	0	6
I-4_10GLR-17-0016_00	INT	10GLR-17-0016	7329	2467	14	1	0	0	0	2	0	5
E-6_10GLR-17-0016_00	EXT	10GLR-17-0016	11576	2467	6	7	0	0	3	0	0	8
E-7_10GLR-17-0016_00	EXT	10GLR-17-0016	11475	2467	20	2	0	0	0	0	0	14
E-8_10GLR-17-0016_00	EXT	10GLR-17-0016	12192	2467	24	0	0	0	2	0	0	6
E-28_10GLR-17-0016_00	EXT	10GLR-17-0016	4407	2467	1	2	0	0	1	0	0	4

In addition, initial data visualization was performed in order to enhance the understanding of operating procedures of the workstations and their effect on the process times. Table 4.8 briefly summarizes the statistics of the design attributes for wall panel. For example, the maximum number of studs in a wall panel can be 68 and minimum are 2 studs. This extreme difference is resulted due to highly customized nature of wall panels, which effected the processing time of wall panels at workstations. Table 4.9 presents the statistical details (i.e., mean and standard deviation) for each workstation on the production line. For example, the mean process time at the framing workstation is 8.15 minutes, with a standard deviation of 2.98 minutes.

Table 4.8: Summary of design attributes for wall panels

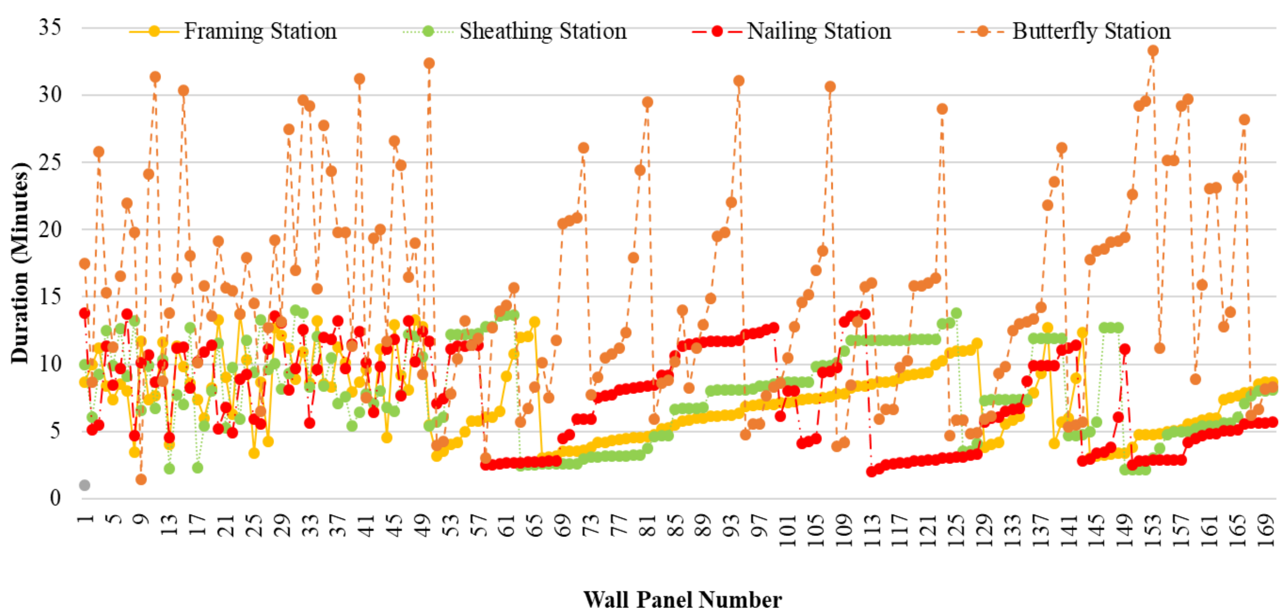
Design Attributes	**Min**	**Max**	**Mean**	**Std.**
Length (ft)	1.93	40.02	12.10	9.91
Width (ft)	2.55	17.58	8.28	0.79
Area (sq. ft)	2.69	473.23	100.31	81.63
Door (number)	0	4	1	1
Window (number)	0	6	1	1
Sheathing (number)	0	21	3	3
Studs	2	68	17	7

Table 4.9: Mean and standard deviation for each workstation

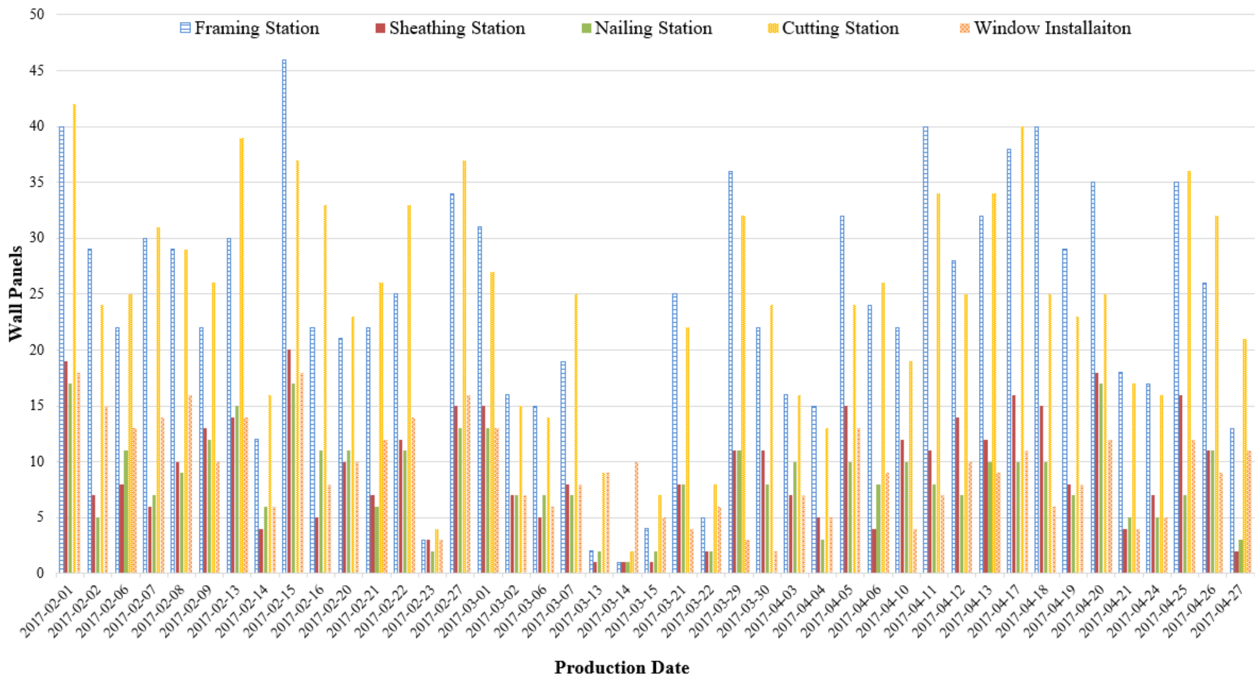
Workstation	**Mean Process Time (minutes)**	**Standard Deviation (minutes)**
Framing	8.15	2.98
Sheathing	9.39	2.79
Nailing	8.35	3.02

Butterfly	6.18	2.48
Window Door	18.5	5.54

Figure 4.14 (a) shows a high level of variance in the process times at workstations. For example, at the sheathing station, the processing time of wall panel 3 was 9 minutes, wall panel 4 was 12 minutes, and wall panel 6 was 2 minutes. Moreover, the process times of cutting station is very high compared to other workstations. This is due to the influence of the design factors (i.e., number of windows, panel length, and number of studs), which is required to fulfil the requirements of clients. This variation in process times affects the daily productivity of the workstations as shown in Figure 4.14 (b), which for example illustrates that the daily production on March 30 was 12 exterior panels at the sheathing station, 22 exterior panels at the framing station, and 6 exterior panels at the nailing station, respectively, causing an imbalanced production line.



(a)

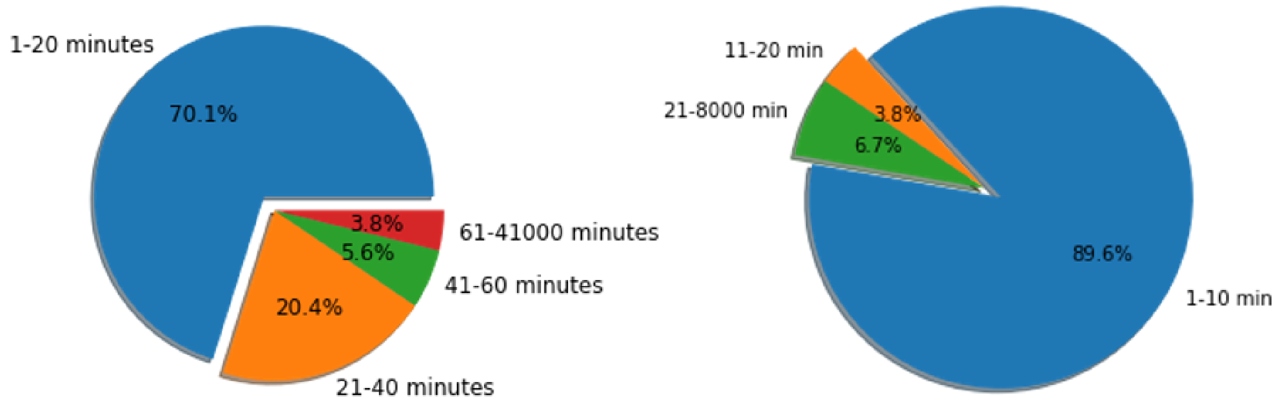


(b)

Figure 4.14: (a) Process times of wall panels; and (b) Daily production of exterior wall panels

Prior to the development of a predictive method, data pre-processing was implemented: (i) keeping initial reading of the wall panels at the workstation and multiple readings of the specific wall panels at that workstation were removed. For example, in Figure 4.15 (a) wall panel E29 initial read time (i.e., 2:34 pm) at wall magazine line station (i.e., A12) was kept and other readings of wall panel E 29 at wall magazine line station (i.e. 2:35, 2:36, 2:37 and 2:39 pm) were removed; (ii) wall panels with missing timestamps and wrong timestamps of wall panels (i.e., wall panels with negative process times) were discarded from the dataset; and (iii) TagID, antenna description, location source antenna, first and last read date, backing, floor, siding, weight, model and parent unit information were also removed before input in the prediction method due to their irrelevance to predict the process times of wall panels. Additionally, outliers were removed based on data visualization (i.e., pie chart), which helps to visualize the distribution of process times that are inconsistent from the data set. For example, Figure 4.15 (b) presents the distribution of process times at cutting workstation (i.e., butterfly workstation), and transfer table. In this respect, the process times above 60 minutes (butterfly workstation), and 10 minutes (transfer table) were removed. The reason for removing these points is that for example at the butterfly workstation around 4% of the wall panel's processing times have excessive times (i.e., 61-410000 minutes). Such data points indicated waiting between the workstations resulted due to: (i) errors in the shop drawings causing work disruption; or/and (ii) wall panels prefabrication that started on preceding day and finished on the succeeding day. In addition, data points above and below 'Mean ± 1.5 SD' were marked as possible statistical outliers. As a result of the pre-processing tasks, the datasets numbered 7256, 2885, 3035, 19998, and 1868 for the framing, sheathing, nailing, butterfly table, and window door, workstations respectively. Moreover, similar physical attributes of a wall panel (e.g., studs, window and sheetfull) were combined into a single attribute in order to reduce data dimensions and computational time. For example: (i) DStud, LStud, and MStud were combined into a 'stud'; (ii) window and large window were combined as a 'window'; (iii) door and large door were combined to 'door'; and (iv) Sheetfull and Sheetpartial were combined as 'sheet'. The min-max normalization technique was applied in order to transform values ranging between 0 and 1 before developing the deep neural network based predictive method. It should be noted that the steps in data pre-processing are specific to the case study (i.e., function of given wall panels design attributes). For other cases, practitioners require to modify the data pre-processing phase according to the given module component design attributes.

PanelNumber	AntennaDes	LocationSourceAı	LocationTagID	InitialReadDateTime
BUMPOUT-2-BLQ-14-24-25-WO_24	A12	A13	E2009A4110029AF0	2:57:08 PM
BUMPOUT-2-BLQ-14-24-25-WO_24	A12	A13	E2009A4110029AF0	2:49:34 PM
BUMPOUT-2-BLQ-14-24-25-WO_24	A12	A13	E2009A4110029AF0	2:41:41 PM
E-29_30DES-17-60-59_U59	A12	A13	E2009A4110029AF0	2:39:19 PM
E-29_30DES-17-60-59_U59	A12	A13	E2009A4120009AF0	2:37:06 PM
E-29_30DES-17-60-59_U59	A12	A13	E2009A4120009AF0	2:36:16 PM
E-29_30DES-17-60-59_U59	A12	A13	E2009A4100005AF0	2:36:14 PM
E-29_30DES-17-60-59_U59	A12	A13	E2009A4110029AF0	2:36:10 PM
E-29_30DES-17-60-59_U59	A12	A13	E2009A4030009AF0	2:35:48 PM
E-29_30DES-17-60-59_U59	A11			2:34:59 PM
E-29_30DES-17-60-59_U59	A12	A13	E2009A4110029AF0	2:34:27 PM
E-14_30DES-17-60-59_U59	A12	A13	E2009A4110029AF0	2:34:13 PM



(i) (ii)

Figure 4.15: (a) RFID data of wall panels at workstations; and (b) Process times distribution of wall panels at: (i) butterfly station; and (ii) transfer table

The variation in the process times at workstations due to differences in design attributes of customized wall panels leads to an imbalanced production line. In this respect a predictive method by considering the wall panel design factors was developed in order to provide the optimal estimates for the process times for using in the planning and scheduling of production line. The DNN in this research was selected due to their successful implementation in construction problems as discussed previously. Additionally, the relationship between the independent (e.g., studs, doors) and dependent variables (process time) does not follow a straight line; instead, it fluctuates, indicating a non-linear pattern. In this respect, DNN are selected as it is well-suited for modeling non-linear data. Based on the training dataset (i.e., 80%) and the independent variables (e.g., panel length, number of regular studs, number of doors, and number of windows) which were used as inputs, the DNN was developed and validated using a 10-fold cross-validation. The developed DNN consists of input layer, multiple hidden layers, and output layer (i.e., process times of wall panels) and each node element was connected and layered with neurons of the next layer. In addition, ANN and MLR were developed to compare the results with DNN. The independent variables such as panel length, number of studs, doors, and windows were used as inputs, while the process times of wall panels in minutes were the output variable.

Table 4.10 presents a comparison of the Mean Absolute Error (MAE) for various numbers of neurons in a single hidden layer of the ANN for each workstation. The number of neurons tested ranges from 1 to 6, and the best result for each workstation (i.e., the minimum MAE achieved) is selected for comparison with the results from DNN and MLR. For instance, at the framing workstation, 5 neurons resulted in the lowest MAE of 2.21 minutes.

Table 4.10: Mean absolute error measurement for ANN

Neurons	Framing (MAE)	Sheathing (MAE)	Nailing (MAE)	Butterfly (MAE)	Window Door (MAE)
1	2.48	5.72	4.93	4.12	8.03
2	2.40	2.30	2.51	3.07	7.76
3	2.42	2.32	2.48	3.10	6.01
4	2.37	2.29	4.46	2.58	5.90
5	2.21	2.36	2.40	2.18	5.12
6	2.39	2.38	2.56	2.93	5.88

In this thesis, the objective is to minimize the error (i.e., MAE) of the predictive method. The MAE is selected to measure the performance of the predictive method based on the relevant literature analysis performed on various performance evaluation techniques. The GA optimization algorithm was selected to minimize the MAE by identifying the optimal number of hidden layers and neurons, which assist in developing a better predictive method. In this respect, for the DNN the rectifier activation function was selected, and the range searched for upper bound/lower bound was 3-10 for hidden layers and 6-100 for the number of neurons. Additionally, the momentum was varied between 0.01 and 0.99, and the learning rate was adjusted between 0.01 and 0.3. In this thesis, the optimization parameters were assigned as follows: (i) population size of 20; (ii) the maximum number of generations was 50; (iii) mutation probability of 0.1; (iv) crossover probability of 1 and: (v) number of tournaments were 3. As observed in Table 4.11, most of the workstations (i.e., framing, sheathing, and nailing station) had MAE of less than 2.50 minutes, respectively. As compared to the works of Mohsen et al. (2022) where the MAE of the prediction algorithms ranges from 4.4 min – 9.2 min, this research provides better predictive method. Additionally, percentage error of the prediction method for the workstations are consistent with the previous prediction methods reported in the literature (Mohsen et al., 2022 and Alsakka et al., 2023). The DNN results were also compared with Artificial Neural Network (ANN) and Multiple linear regression (MLR). As these results show, the DNN method was found to have better MAE values except for the Nailing workstation, where ANN provide better MAE of 2.40 min, respectively. Regarding mean percentage error, Elmousalami (2020), Aydin (2015), and Lewis (1982) classify mean absolute percentage error as follows: less than 10% indicates an excellent prediction, between 10% and 20% is considered a good prediction, between 20% and 50% is seen as acceptable (reasonable) forecasting, and more than 50% is deemed an inaccurate prediction in the industrial context. In this respect, based on the SMAPE, the prediction method of this study can be considered as reasonable prediction method. Additionally, in the context of scheduling and planning, knowing the absolute error in duration is often more important. In this respect, accurate prediction of the exact number of minutes it takes to prefabricate the wall panels at each workstation is more crucial than the percentage error in respect to this study, as it directly impacts production planning.

The predictive method exhibits a SMAPE between 22-28% due to the absence of key factors that can impact the process time of wall panels at the workstation. For example, according to standard operating procedures (SOPs) file provided by industrial partner: (i) workers need to set up the CNC machine after completing each wall panel, but the current data lacks any variables related to this setup time; and (ii) at the sheathing workstation, workers spend time installing guard wrap on the top plate of a wall panel, cutting OSB (oriented strand board) sheets, labelling the wall number on the top plate, and conducting a quality check (e.g., ensuring end studs are aligned with the plate); these activities are not included as variables in the current dataset. Additionally, the dataset does not include variables for worker idle time at workstations due to material shortages or machine breakdowns. The RFID system records only the timestamps (start and end times) of wall panels at each workstation. However, it doesn't capture the activities that occur between these timestamps, such as additional factors affecting process time beyond the design factors previously mentioned. In this respect, the current prediction method can be improved by collecting data related to activities such as cutting OSB (oriented strand board), inspection and labelling the wall number on wall panels.

Table 4.11: Comparison of DNN, ANN and MLR

Workstations	DNN Selected Value		DNN MAE (min)	DNN SMAPE (%)	ANN Value		ANN MAE (min)	ANN SMAPE (%)	Reg. MAE (min)	Reg SMAPE (%)
	Hidden Layers	Neurons			Hidden layers	Neurons				
Framing	3	74	2.17min	25.78%	1	5	2.21	27.94%	2.24	28.86%
Sheathing	3	72	2.11min	22.49%	1	4	2.29	24.87%	2.28	24.81%
Nailing	3	14	2.41min	28.42%	1	5	2.40	28.31%	2.43	28.68%
Butterfly	7	70	2.05min	28.87%	1	5	2.18	31.59%	2.24	32.12%
Window Door	8	42	4.48min	27.62%	1	5	5.12	31.56%	4.61	28.42%

Prior to the planning and scheduling phase, the building was designed in Autodesk Revit 2022 and comprised two multi-storey residential buildings (Project A: 4 floors and Project B: 2 floors) consisting of two units at each floor (i.e., a total of 8 units and 4 units for project A and B, respectively). It is noteworthy that wall panels were designed according to the design specifications of the case company. For instance, in the wall panels, the stud spacing was 2 ft, and header/sill tracks and jack studs (i.e., double studs) were added for the wall panels with door/window. The wall panels were of various lengths (e.g., 10.74 ft, 9.4 ft, 10 ft, 12.3 ft, and 6.25 ft). Based on the wall panel information from the design model, the walls were grouped according to type (i.e., exterior versus interior) and used to develop the practical rules (i.e., precedence relationships for sequencing and installing wall panels based on type and location) governing the prefabrication of wall panels for multiple projects in the production line. Based on the two projects, total 390 wall panels need to be prefabricated in the production line—258 wall panels for project A and 132 wall panels for project B. These panels are grouped by floor (f) levels to create work packages. For example, in project A, 1st, 2nd and 4th floors include 66 wall panels each and, 3rd floor includes 60 wall panels. These wall panels were further categorized into interior and exterior types. Specifically, project A includes 141 interior panels (f 1: 34, f 2: 39, f 3: 31 and f 4: 37) and 117 exterior panels (f 1: 32, f 2: 27, f 3: 29 and f 4: 29) while project B comprises 73 interior panels (f 1: 34 and f 2: 39) and 59 exterior panels (f 1: 32 and f 2: 27). Additionally, number of workstations (i.e., tasks) needed to prefabricate these panels on the production line was used to further categorize the interior and exterior wall panels. The rationale behind grouping these panels based on the number of workstations is that panels requiring fewer workstations should be grouped together. For instance, interior wall panels pass through four workstations- framing, sheathing, nailing and butterfly before moving to the storage area via the bypass line. In contrast, exterior wall panels go through five workstations: framing, sheathing, nailing, butterfly and window/door installation before reaching storage area. It's important to note that the number of interior and exterior wall panels within the group did not change from the previous step (i.e., types of wall panels) since all panels follow same process (i.e., interior wall panels need four workstations and exterior wall panels requires five). As represented in Table 4.12, the second floor of project A consists of 39 interior and 27 exterior panels. Since the production line handles multiple projects simultaneously, interior and exterior wall panels, which are located at the same floors of project A and B, are combined into one single work package. For instance, project A and B have 34 interior panels on the first floor, respectively. Therefore, 1st work package has 68 interior wall panels and 2nd work package has 64 exterior wall panels. As a result, there are total eight work packages for project A and B which should be scheduled in sequence. These group of panels (work packages) were assigned priority according to practical rule of installing wall panels on-site (i.e., interior wall panels should be installed first before exterior wall panels of 1st floor).

Table 4.12: Number of work packages for project A and B

Work package	Floor	Project A		Project B		Total
		Interior	Exterior	Interior	Exterior	
1	1	34	-	34	-	68
2	1	-	32	-	32	64
3	2	39	-	39	-	78
4	2	-	27	-	27	54

5	3	31	-	-	-	31
6	3	-	29	-	-	29
7	4	37	-	-	-	37
8	4	-	29	-	-	29

Once the work packages were created, these work packages were used in the optimization algorithm in order to find optimal sequences of prefabrication in the production line. Prior to the start of an optimization, predictive method was developed to forecast the process times of wall panels and, the details being reported in a previous study (Bhatia et al. 2023). The predictive process times of exterior and interior wall panels were in turn used as inputs in the optimization method in the form of a matrix. Figure. 4.16 represents the process times of exterior wall panels at workstations for 6th work package, where the first column represents the wall panels Id's.

	Workstation 1	Workstation 2	Workstation 3	Workstation 4	Workstation 6	Workstation 7
WallPanel 1	7.2198215	8.021043	7.511536	15.596252	18.5	2.8000617
WallPanel 2	7.4876328	7.5772123	7.5491457	15.984384	18.5	3.2997293
WallPanel 3	7.4127507	8.674869	7.494323	15.982462	18.5	3.2993722
WallPanel 4	8.016617	8.790088	7.587493	16.139832	18.5	3.2994926
WallPanel 5	8.014654	7.4042625	8.69894	11.843401	18.5	3.2997293
WallPanel 6	7.2352366	8.7715845	7.4623394	11.791236	18.5	3.2997293
WallPanel 7	7.3726606	8.010055	8.557127	16.99746	18.5	3.9827423
WallPanel 8	7.570148	7.91331	9.353164	17.191147	18.5	3.2281199
WallPanel 9	7.930071	7.802701	9.478735	16.481155	18.5	3.298924
WallPanel 10	7.2387724	9.142428	7.4972625	17.067968	18.5	3.1623058
WallPanel 11	8.149931	7.4410324	7.4775376	14.995985	18.5	3.2997293
WallPanel 12	8.903113	9.3082075	7.4750266	10.053232	18.5	3.532126
WallPanel 13	7.510326	9.060581	7.5104775	15.703651	18.5	3.9153986
WallPanel 14	7.668431	7.656288	7.4599285	16.314507	18.5	3.9183934
WallPanel 15	7.849856	8.173088	7.478407	11.781427	18.5	2.4395328
WallPanel 16	7.141686	8.993142	7.503044	15.561876	18.5	3.304953
WallPanel 17	7.2160316	8.048378	7.578043	16.340176	18.5	3.9849076
WallPanel 18	8.381582	8.010055	7.537537	14.318943	18.5	3.304953

Figure 4.16: Examples of process times in workstations for 6th work package

Once all input data is ready, GA, SA, and HGASA were implemented in order to schedule wall panels of each work package by minimizing the makespan of wall panel prefabrication. The GA optimization parameters used in this case study were 200 generations, each generation containing 30 populations with mutation and crossover rates of 0.2 and 0.8, respectively. These values are determined by experimenting with various values and selecting the one yielding the best results. Figure 4.17 (a) shows the makespan value (i.e., duration) of 2nd work package which has total 64 exterior wall panels. The GA started with an initial solution of a makespan of 1,114 min and converged to the best solution which has 1,108 min (18.46 hr) in 120 generations. The SA parameters used in this case study were 2,000 iterations, a temperature of 1,000, a cooling rate of 0.95, and the swapping method as the method of choice for neighbor generation. In this respect, as shown in Figure 4.17 (b), SA started to schedule 5th work package with an initial solution of a makespan of 484 min, converging to the best solution in 255 iterations with a makespan of 479 min (7.98 hr). The HGASA parameters used in this case study were similar to that were defined in the GA and SA. The reason is to ensure that HGASA perform under similar parameter settings and the differences in the makespan can be attributed to the algorithm structure rather than the influence of different parameter settings. Figure 4.17 (c) shows the respective optimal sequences of wall panels in 5th work package generated by the GA, SA, and HGASA. Intuitively, the different sequences of wall panels generated by these algorithms resulted in different makespan. For example, in the case of GA, the sequence starts with wall panel 24 and requires 13.58 hr, compared to starting with wall panel 12 and requiring 7.98 hr (as in the case of SA), and starting with wall panel 3 and requiring 7.26 hr (as in the case of HGASA).

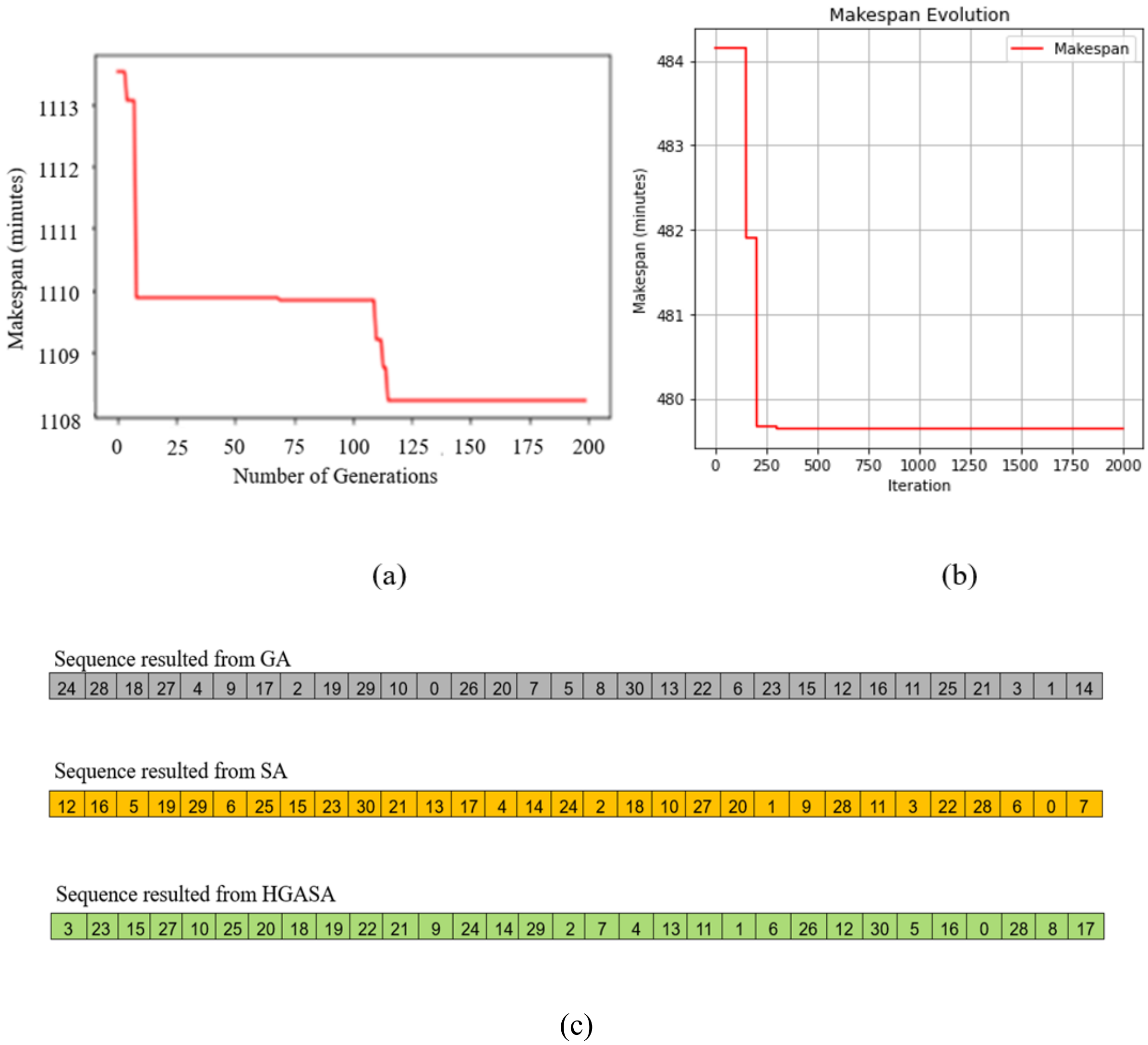


Figure 4.17: (a) Makespan for 2nd work package using GA; (b) Makespan for 5th work package using SA; (c) Optimal sequences of wall panels at 5th work package.

Table 4.13 represents makespan and computation times of work packages in accordance with GA, SA, and HGASA. The results show that HGASA provides the best optimal sequences of wall panels since total makespan of all work packages is 105.63 hr which is minimum value comparing with 108.06 hr in SA and 138.08 in GA. At this junction, it should be noted that HGASA generates higher makespan than ones in GA or SA in some work packages. For example, HGASA provides higher makespan, which are 19.73 hr and 16.65 hr in 2nd and 4th work packages, respectively, than ones in GA which are 18.46 hrs and 16 hrs. However, these makespan are less than ones resulted in SA. In a view of computational time, HGASA requires the highest computation time and SA provides the shortest computation time.

Table 4.13: Makespan of work packages using GA, SA and HGASA

Work Package	GA		SA		HGASA	
	Production time (hr)	Comput. Time (s)	Production time (hr)	Comput. Time (s)	Production time (hr)	Comput. Time (s)
1: Int. (68 wall panels)	26.65	40	16.76	22	16.60	61
2: Ext. (64 wall panels)	18.46	40	20.03	20	19.73	58
3: Int. (78 wall panels)	29.8	46	18.76	22	18.61	64
4: Ext. (54 wall panels)	16	34	16.95	19	16.65	52
5: Int. (31 wall panels)	13.58	23	7.98	15	7.26	37
6: Ext. (29 wall panels)	9.38	23	9.25	15	8.93	37
7: Int. (37 wall panels)	14.88	27	9.08	17	8.91	42
8: Ext. (29 wall panels)	9.33	24	9.25	14	8.94	36
Total	138.08	217	108.06	144	105.63	387

The optimization algorithms described above were used to schedule optimal sequences of wall panels which have minimum makespan in work packages. However, there might be bottlenecks in the wall panel sequence due to reasons, such as change orders requested by sites and variation of process times in workstations. For example, in the process times of sheathing workstation for 5th work package, the most wall panels took process times between 7 min and 8 min, but some panels (e.g., wall IDs 3, 7, 9, 20, 23, and 29) required between 9 min and 9.5 min. In other words, the longer process times of these panels compared to the majority of the panels caused bottlenecks at the sheathing workstation. In this respect, in practice, especially industrial partner represented in the case study, they have operated an additional sheathing workstation to eliminate the bottleneck and improve productivity. As a result, there are parallel workstations which are operated in manual and machine-based sheathing stations concurrently. The process times for this manual sheathing station were generated based on a triangular distribution provided by the case company. The parameters of the triangular distribution were a minimum of 0.72 min, a maximum of 15.92 min, and a mean of 7.69 min. To capture parallel sheathing stations (i.e., sheathing workstation that can prefabricate two wall panels simultaneously) in the optimization algorithms, constraints (Eq. (3.15), (3.16) and (3.17) discussed in section 3.4.3) was developed and applied into optimization algorithms. Table 4.14 represents makespan values of various work packages using GA, SA, and HGASA. The results show that the optimal sequences of wall panels generated by HGASA provides better (minimum) makespan, since total makespan of all work packages was 103.55 hr comparing with 105.92 hr in SA and 136.34 hr in GA. To identify the effectiveness of the parallel workstation over a single sheathing workstation, the makespan values were compared. For example, 2nd work package was 20.03 hr for SA, 18.46 hr for GA and 19.73hr for HGASA. However, with the parallel sheathing workstation, the makespan was reduced to 19.68hr, 18.18 hr and 19.34 hr for SA, GA and HGASA, respectively.

Table 4.14: Makespan of work packages using GA, SA and HGASA after adding parallel workstation

Work Package	GA	SA	HGASA
	Production time (hr)	Production time (hr)	Production time (hr)
1: Int. (68 wall panels)	26.43	16.53	16.27
2: Ext. (64 wall panels)	18.18	19.68	19.34
3: Int. (78 wall panels)	29.3	18.38	18.25
4: Ext. (54 wall panels)	15.83	16.58	16.31
5: Int. (31 wall panels)	13.49	7.81	7.12
6: Ext. (29 wall panels)	9.21	9.04	8.76
7: Int. (37 wall panels)	14.72	8.86	8.74
8: Ext. (29 wall panels)	9.18	9.04	8.76
Total	136.34	105.92	103.55

To identify the most effectiveness of optimization algorithm in scheduling sequences of wall panels, as represented in Table 4.15, a comparison metrics with one linear fashion of the production line and parallel workstations (e.g., sheathing workstation can work on two wall panels simultaneously) were calculated in GA, SA, and HGASA. The Hybrid algorithm achieves the best makespan of 105.63 hr for prefabricating wall panels for two projects; however, its implementation is significantly more complex compared to GA and SA. Additionally, the Hybrid algorithm shows a 1.96% reduction in makespan after incorporating a parallel workstation, compared to a 1.98 % reduction for the SA algorithm.

Table 4.15: Result of comparison metrics

Algorithm	Makespan for multiple projects (hr)	Comput. Time (min)	Complexity	Production load balancing	Parallel workstation

GA	138.08	4.28	Medium	3.19	1.21 %
SA	108.06	2.41	Low	2.54	1.98 %
Hybrid GASA	105.63	6.45	High	2.71	1.96 %

The present study is subject to some limitations to be addressed in future research. First, the optimization approach presented herein overlooks the logical relationships governing the on-site installation of interior walls. Consideration of the logical relationships underlying the installation of wall panels within the units will facilitate the efficient implementation of the JIT lean principle (i.e., the idea that the factory should fabricate wall panels in the order they are required for on-site installation in each unit). Doing so will: *(i)* prevent overproduction; and *(ii)* reduce the storage space requirements within the factory (and associated costs). In this respect, future research will focus on establishing practical rules for work-packages in order to maintain the optimal sequence of interior wall panels within each unit, extending beyond just floor-level considerations. Second, the hybrid optimization method presented herein employs a sequential hybrid approach, and therefore lacks continuous refinement of GA solutions. Rather than waiting until the GA phase is complete to begin refining the solution via SA, improvements can be made iteratively, ensuring that solutions are progressively enhanced throughout the optimization process, leading to more effective exploration of the solution space. In this respect, future research can explore an embedded hybrid GASA approach, where SA is used within the GA process to continuously refine individual solutions, potentially enhancing optimization outcomes.

CHAPTER 5: CONCLUSION

5.1 Research Summary

Modular construction provides several advantages, such as improved productivity, quality and safety. However, customization arising from owners' requirements results in different processing times for module components at workstations. This poses a challenge for production line managers to accurately predict the process times of module components at workstations, leading to inefficient production line performance.

To solve these problems, this thesis introduces a framework composed of three modules. The first includes a simulation-based production line planning method that uses near real-time and historical data to assist production managers in achieving better productivity and control. This method integrates a C-track app, statistical analysis, and simulation for production planning in MCM. As an alternative to the experience-based approach implemented in traditional MCM production planning, the developed method collects historical and near real-time data using the C-track app to enhance decision making; identifies the SIFs affecting fabrication process times using statistical techniques that increase the accuracy of the predictive method; and improves productivity by using simulation as a production planning tool. It also introduces the concept of developing and evaluating multiple production sequencing and labor allocation scenarios using two types of input, cycle time formula and probability distribution functions in the simulation process. The second module houses a newly developed method that utilizes historical time data of the manufacturing factory as an input in the Deep Neural Network (DNN), Artificial Neural Network (ANN) and Multiple Linear Regression (MLR) to predict the process times at workstations. Subsequently, the GA optimization is implemented to optimize the architecture of the DNN and, as such, finds a near optimum number of hidden layers and nodes in each layer. The third module focuses on the development of an optimized planning and scheduling method for MCM that provides a near optimum sequencing of module components at workstations of the production line. The method consists of: *(a)* extracting quantities of module components (e.g., number of studs) from BIM of residential projects and then inputting this data into the trained predictive method in order to forecast the process time at each workstation; *(b)* establishing a work planning structure (WPS), which involves categorizing module components by type and assigning practical rules to the work-packages; and *(c)* utilizing a hybrid algorithm that integrates GA and SA, and validating the effectiveness of this hybrid algorithm by comparing their respective results (i.e., the optimal sequences of wall panels generated) with those of GA and SA.

To demonstrate the effectiveness and test the performance of the developed methods, two case studies were analyzed: (i) light gauge steel (LGS) wall panel production line, which mainly operates manually; and (ii) wood-based wall panel production line, where certain workstations are operated by CNC machines. The simulation-based planning method developed in the first module, was implemented in the LGS case and demonstrates that this approach leads to improved productivity (i.e., reduced durations with the same labor input) and control. In particular, scenario 2.3 (i.e., exterior wall panels of a given unit were prefabricated first, followed by the corresponding interior walls for the unit along with allocation of labor to workstations) was found to outperform the other scenarios, with a 44.42 hr duration to produce 309 wall panels. As demonstrated by the case study, this method can assist production managers in understanding the effects of proposed changes to the production line before implementing them in reality. In this way, production managers can plan effectively and reduce project costs. Next, the results of the DNN-based process time predictive method (second module) after implementing in the second case, demonstrates that the method predicted the process times with a MAE of less than 2.50 minutes for most of the workstations, respectively. It is worth noting that in most cases (workstations) the DNN provided better results compared to ANN and multiple linear regression, however in case of Nailing workstation ANN performed better. As compared to the works reported in literature where the MAE of the prediction algorithms ranges from 4.4 min – 9.2 min, this study reports better MAE for the predictive method. The optimized planning and scheduling method of the third module shows that it can efficiently generate near optimum production schedules and calculate makespan. In particular, the optimal sequences of wall panels generated by HGASA are found to result in a reduced makespan compared to those generated by GA and SA with hybrid optimization reduces makespan to 105.63 hr from those generated by GA (138.08 hr) and SA (108.06 hr). In particular, the optimal sequences of wall panels generated by HGASA resulted in reduced makespan compared to those generated by GA and SA.

5.2 Research Contributions

The contribution of this research to the body of knowledge include:

1) The developed simulation-based planning method assist in evaluating the scenarios of: (i) sequencing the modules and (ii) allocating resources along the production line, which can reduce idle time of workers, waiting time of modules at workstations and duration of prefabricating module components. Also, identification of significant impact factors (SIFs) influencing fabrication process times at workstations provides deeper understanding of the underlying production line process, in this way, production managers can plan effectively.

2) The contributions of the DNN based process time predictive method include: (i) implement GA to determine the optimal hyperparameters of DNN considering number of hidden layers, neurons, momentum and learning rate to reduce the prediction error; (ii) develop process time predictive method considering all production line workstations, rather than experience-based approach for

estimating process times, which is error prone. This assist in making data-driven decisions with respect to efficiently plan and schedule sequences of module components; and (iii) helping production managers gain insights into the production line, enabling them to shorten the prefabrication time of module components at workstations

3) The contributions of the optimized planning and scheduling method includes: (i) develop a hybrid optimization approach that capitalizes on the strengths of both GA and SA, ultimately minimizing the makespan for multiple projects; (ii) provide a practical function in the optimization algorithms which can operate parallel workstations as an effort to eliminate the bottleneck in the production line; and (iii) enables production line managers to choose an optimal schedule and effectively implement frequent hourly or daily adjustments, adapting to the dynamic nature of the production line and leading to continuous improvements in the MCM production lines.

5.3 Research Limitations

The main limitations of this research are:

1) The developed method lacks consideration for optimizing the allocation of workers at workstations.

2) The current hybrid optimization method employs a sequential hybrid approach, which lacks continuous refinement of GA solutions. Rather than waiting until the GA phase is complete to begin refining the solution via SA, improvements can be made iteratively, ensuring that solutions are progressively enhanced throughout the optimization process, leading to more effective exploration of the solution space.

3) Findings of the optimization algorithms can be considered as domain specific (i.e., according to case studies utilized in this thesis). Additionally, the on-site installation of interior wall panels first

followed by exterior are according to the multi-storey building projects and can be considered as case specific.

4) The developed methods are applicable to other production line cases; however, practitioners would need to modify the input factors based on the given design specifications of module components.

5) The RFID system records only the process time of wall panels at each workstation. However, it doesn't capture the activities that occur between these process time, such as additional factors affecting process time beyond the design factors, which reduces the accuracy of the prediction model.

6) This research primarily emphasizes the development of predictive method with module design specifications and number of workers, without focusing on factors such as work shift, material availability and on-site change orders that can affect process time.

5.4 Future Work

Although this research has investigated the potential for enhancing planning and scheduling method using optimization, discrete event simulation and data analytics, there are aspects that might need additional study:

1) Optimizing both the sequence of module components and allocating number of workers at each workstation.

2) Future studies can explore multi-objective optimization for production line schedules, considering factors such as cost and makespan simultaneously.

3) In the optimization process, the number of neurons in each hidden layer remains constant. To introduce more diversity, future research could investigate optimization processes where the number of neurons varies in each hidden layer based on optimal results.

4) Optimizing the design of wall panels to minimize the impact of customization so that project completion time can be reduced.

5) Quality metrics can be integrated into the optimization process by quantifying the rework percentage of prefabricating module components at each workstation.

REFERENCES

AbouRizk, S. M., and D. W. Halpin. 1992. Statistical properties of construction duration data. Journal of construction engineering and management 118 (3): 525–544.

AbouRizk, S. and Mohamed, Y., 2000. Simphony-an integrated environment for construction simulation. In Winter Simulation Conference Proceedings, 2, 1907-1914.

AbouRizk, S., 2010. Role of simulation in construction engineering and management. Journal of construction engineering and management, 136(10), 1140-1153.

AbouRizk S., Hague S. and Ekyalimpa R. 2016. Construction simulation: An introduction using Simphony, Edmonton, AB, Canada: University of Alberta.

Aggarwal, C.C., 2018. Neural networks and deep learning. Cham:Springer. 10, 978, 3.

Aghajamali, K., Suliman, A., Alattas, A. and Lei, Z., 2022. Data-driven cycle time prediction of fitting and welding stations in steel fabrication. Modular and Offsite Construction (MOC) Summit Proceedings, 122-129.

Akpinar M, Akpinar KO, Aktas E, Senocak OF, Donmez H, Cavdar V. 2020. Duration estimation of vehicle-based production: case study of assembly line. Int. Conf. Data Anal. Bus. Ind. W. Towar. a Sustain. Economy

Alsakka, F., Yu, H., Hamzeh, F. and Al-Hussein, M., 2023. Data-and Knowledge-Driven Cycle Time Estimation in Off-site Construction.

Alsakka, F., Yu, H., El-Chami, I., Hamzeh, F. and Al-Hussein, M., 2024. Digital twin for production estimation, scheduling and real-time monitoring in offsite construction. Computers & Industrial Engineering, 191,110173.

Altaf, M.S., Bouferguene, A., Liu, H., Al-Hussein, M. and Yu, H., 2018. Integrated production planning and control system for a panelized home prefabrication facility using simulation and RFID. Automation in construction, 85, 369-383.

Ambrogio, G. and Gagliardi, F., 2013. Design of an optimized procedure to predict opposite performances in porthole die extrusion. Neural Computing and Applications, 23,195-206.

An, L., Yang, P., Zhang, H. and Chen, M., 2014. Multi-objective optimization for milling operations using genetic algorithms under various constraints. International Journal of Networked and Distributed Computing, 2(2), 108-114.

Anand, S., Afreen, N. and Yazdani, S., 2015. A novel and efficient selection method in genetic algorithm. International Journal of Computer Applications, 129(15), 7-12.

Assad, A. and Bouferguene, A., 2022. Data mining algorithms for water main condition prediction—comparative analysis. Journal of Water Resources Planning and Management, 148(2), 04021101.

Aydin, G.Ö.K.H.A.N., 2015. Forecasting natural gas production using various regression models. Petroleum Science and Technology, 33(15-16),1486-1492.

Azimi, R., Lee, S., AbouRizk, S.M. and Alvanchi, A., 2011. A framework for an automated and integrated project monitoring and control system for steel fabrication projects, Automation in Construction, 20(1), 88-97.

Backus, P., Janakiram, M., Mowzoon, S., Runger, C. and Bhargava, A., 2006. Factory cycle-time prediction with a data-mining approach. IEEE Transactions on Semiconductor Manufacturing, 19(2), 252-258.

Baker, K.R. and Trietsch, D., 2013. Principles of sequencing and scheduling. John Wiley & Sons.

Ballard, H.G., 2000. The last planner system of production control, Doctoral dissertation, University of Birmingham.

Bardareh, H and O. Moselhi., 2022. An integrated rfid--uwb method for indoor localization of materials in construction. Journal of information technology in construction, 27.

Barkokebas, B., Martinez, P., Bouferguene, A., Hamzeh, F. and Al-Hussein, M., 2023. Digitalization-based process improvement and decision-making in offsite construction. Automation in Construction, 155, 105052.

Barkokebas, B., Khalife, S., Al-Hussein, M. and Hamzeh, F., 2021. A BIM-lean framework for digitalisation of premanufacturing phases in offsite construction. Engineering, Construction and Architectural Management, 28(8), 2155-2175.

Barkokebas, B., Ritter, C., Li, X. and Al-Hussein, M., 2018. Predictive Model for Siding Practice in Panelized Construction. Modular and Off-site Construction (MOC) Summit Proceedings.

Bhatia, A., Moselhi, O. and Han, S.H., 2023. Optimized Production Scheduling for Modular Construction Manufacturing. In ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction, 40, 270-277.

Bhatia, A.P.S., Han, S. and Moselhi, O., 2022. A simulation-based statistical method for planning modular construction manufacturing. Journal of Information Technology in Construction, 27.

Boyd, N., Khalfan, M.M. and Maqsood, T., Off-site construction of apartment buildings. 2013. Journal of architectural engineering, 19(1), 51-57.

Callens, A., Morichon, D., Abadie, S., Delpey, M. and Liquet, B., 2020. Using Random forest and

Gradient boosting trees to improve wave forecast at a specific location. Applied Ocean Research, 104, 102339.

Chanmeka A., Thomas S.R., Caldas C.H. and Mulva S.P. 2012. Assessing key factors impacting the performance and productivity of oil and gas projects in Alberta, Canadian Journal of Civil Engineering, 39, 3, 259–270.

Chen, J., Wang, M., Kong, X.T., Huang, G.Q., Dai, Q. and Shi, G., 2019. Manufacturing synchronization in a hybrid flowshop with dynamic order arrivals. Journal of Intelligent Manufacturing, 30, 2659-2668.

Chen, H., Ding, G., Zhang, J. and Qin, S., 2019. Research on priority rules for the stochastic resource constrained multi-project scheduling problem with new project arrival. Computers & Industrial Engineering, 137,106060.

Cheng, Y., Chen, K., Sun, H., Zhang, Y. and Tao, F., 2018. Data and knowledge mining with big data towards smart production. Journal of Industrial Information Integration, 9, 1-13.

Chen, J.S., Pan, J.C.H. and Lin, C.M., 2009. Solving the reentrant permutation flow-shop scheduling problem with a hybrid genetic algorithm. International Journal of Industrial Engineering: Theory Applications and Practice, 16(1), .23-31.

Cottrell D.S. 2006. Contractor process improvement for enhancing construction productivity, Journal of Construction Engineering and Management, 132, 2, 189–196.

Dragicevic N. and Riaz K. 2024. seizing the modular construction opportunity, CSA public policy centre (assessed Feb, 2024).

El-Sawah, H. and Moselhi, O., 2014. Comparative study in the use of neural networks for order of magnitude cost estimating in construction. Journal of Information Technology in

Construction (ITcon), 19(27), 462-473.

ElFiky, H., Owida, A. and Galal, N.M., 2020, March. Resource constrained multi-project scheduling using priority rules: application in the deep-water construction industry. In Proceedings of the International Conference on Industrial Engineering and Operations Management, 353-363.

Elmousalami H.H. 2019. Artificial intelligence and parametric construction cost estimate modeling: State-ofthe-art review, Journal of Construction Engineering and Management, 146, 1, 03119008.

Elmousalami, H.H., 2020. Comparison of artificial intelligence techniques for project conceptual cost prediction: A case study and comparative analysis. IEEE Transactions on Engineering Management, 68(1), 183-196.

Employment and Social Development Canada (ESDC), Government of Canada promotes in-demand skilled trades as a first-choice career path. < https://www.canada.ca/en/employment-social-development/ news/2022/01/skills-trade.html > (accessed March, 2024).

Framinan, J.M., Leisten, R. and García, R.R., 2014. Manufacturing scheduling systems. An integrated view on Models, Methods and Tools, 51-63.

Garusinghe. G, Perera. B, Weerapperuma. U. 2023. Integrating circular economy principles in modular construction to enhance sustainability, Sustainability. 15, 15, 11730.

Ganesh, K. and Punniyamoorthy, M., 2005. Optimization of continuous-time production planning using hybrid genetic algorithms-simulated annealing. The International Journal of Advanced Manufacturing Technology, 26, 148-154.

Gann, D.M., 1996. Construction as a manufacturing process? Similarities and differences between industrialized housing and car production in Japan. Construction Management & Economics, 14(5), 437-450.

Gerald B. 2018. A brief review of independent, dependent and one sample t-test, International Journal of Applied Mathematics and Theoretical Physics, 4, 2, 50.

Guyon, I. and Elisseeff, A.,2003. An introduction to variable and feature selection. Journal of machine learning research,1157-1182.

Hegazy, T. and Kamarah, E., 2022. Schedule optimization for scattered repetitive projects. Automation in Construction, 133, 104042.

Han, S.H., Al-Hussein, M., Al-Jibouri, S. and Yu, H., 2012. Automated post-simulation visualization of modular building production assembly line. Automation in construction, 21, 229-236.

Halpin, D.W., 1977. CYCLONE-method for modeling job site processes. Journal of the construction division, 103 (ASCE 13234 Proceeding).

Hajjar, D. and AbouRizk, S., 1999. Simphony: an environment for building special purpose construction simulation tools. In Proceedings of the 31st conference on Winter simulation: Simulation---a bridge to the future-Volume 2, 998-1006.

Hasan, M. and Lu, M., 2022. Variance Analysis on Regression Models for Estimating Labor Costs of Prefabricated Components. Journal of Computing in Civil Engineering, 36(5), 04022019.

Hong, T., Cho, K., Hyun, C. and Han, S., 2011 Simulation-based schedule estimation model for ACS-based core wall construction of high-rise building, Journal of Construction Engineering and Management, 137(6), 393-402.

How Modular Construction Will Get You a Better Return on Investment. <

https://www.modular.org/2021/07/14/how-modular-construction-will-give-you-a-better-return-on-investment/ >(accessed Nov, 2023).

Hernández-Ramírez, L., Frausto-Solis, J., Castilla-Valdez, G., González-Barbosa, J.J., Terán-Villanueva, D. and Morales-Rodríguez, M.L., 2019. A Hybrid Simulated Annealing for Job Shop Scheduling Problem. International Journal of Combinatorial Optimization Problems & Informatics, 10(1).

Hyun, H., Yoon, I., Lee, H.S., Park, M. and Lee, J., 2021. Multiobjective optimization for modular unit production lines focusing on crew allocation and production performance. Automation in Construction, 125, 103581.

Ji, M., He, M. and Wang, Q., 2009, August. Application of a novel simulated annealing in flow shop scheduling problem. In 2009 Fifth International Conference on Natural Computation, 6, 339-342.

Jain, A.S. and Meeran, S., 1999. Deterministic job-shop scheduling: Past, present and future. European journal of operational research, 113(2), 390-434.

Koskela, L., 1992. Application of the new production philosophy to construction (Vol. 72). Stanford: Stanford university.

Kolisch, R., 1994. Efficient Priority Rules for the Resource-Constrained Project Scheduling Problem Kiel.

Koc, K., Ekmekcioğlu, Ö. and Gurgun, A.P., 2021. Integrating feature engineering, genetic algorithm and tree-based machine learning methods to predict the post-accident disability status of construction workers. Automation in construction, 131, 103896.

Kva Lseth, T.O. 1983. Note on the R 2 measure of goodness of fit for non-linear models. Bulletin of the Psychonomic Society, 21, 79-80.

Krüger, D. and Scholl, A., 2010. Managing and modelling general resource transfers in (multi-) project scheduling. 32(2), 369-394.

Lee, J. and Hyun, H., 2019. Multiple modular building construction project scheduling using genetic algorithms. Journal of Construction Engineering and Management, 145(1), 04018116.

Lee, J. and Kim, J., 2017. BIM-based 4D simulation to improve module manufacturing productivity for sustainable building projects. Sustainability, 9(3), 426.

Lewis, C.D., 1982. International and Business Forecasting. Methods Butterworths: London, 144.

Lee, D.E., 2005. Probability of project completion using stochastic project scheduling simulation. Journal of construction engineering and management, 131(3), 310-318.

Liu, H., Al-Hussein, M. and Lu, M., 2015. BIM-based integrated approach for detailed construction scheduling under resource constraints. Automation in construction, 53, 29-43.

Lu, M., and S. M. AbouRizk. 2000. Simplified CPM/PERT simulation model. Journal of Construction Engineering and Management. 126 (3): 219–226

Mahmood Z. and Khan S. 2009. On the use of k-fold cross-validation to choose cutoff values and assess the performance of predictive models in stepwise regression, The International Journal of Biostatistics, 5, 1.

Makridakis S., Spiliotis E. and Assimakopoulos V. 2018. Statistical and machine learning forecasting methods: Concerns and ways forward, PLoS ONE, 13, 3, 0194889.

Makridakis, S., 1993. Accuracy measures: theoretical and practical concerns. International journal of forecasting, 9(4), 527-529.

Martinez P., Mohamed E., Mohsen O., and Mohamed Y. 2020. Comparative study of data mining models for prediction of bridge future conditions, Journal of Performance of Constructed Facilities, 34, 1, 04019108

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C. and Hung Byers, A., 2011. Big data: The next frontier for innovation, competition, and productivity. (assessed Feb, 2024).

Mattsson S, Li D, Fast-Berglund Å, Gong L 2017.. Measuring operator emotion objectively at a complex final assembly station. Adv Intell Syst Comput; 488:223–32.

Mohsen, O., Petre, C. and Mohamed, Y., 2023. Machine-Learning Approach to Predict Total Fabrication Duration of Industrial Pipe Spools. Journal of Construction Engineering and Management, 149(2), 04022172.

Mohsen. O, Mohamed. Y, Al-Hussein. M, 2022. A machine learning approach to predict production time using real-time RFID data in industrialized building construction. Advanced Engineering Informatics, 52, 101631.

Mohsenijam, A., Siu, M.F.F. and Lu, M., 2017. Modified stepwise regression approach to streamlining predictive analytics for construction engineering applications. Journal of Computing in Civil Engineering, 31(3), 04016066.

Mohsenijam, A. and Lu, M., 2019. Framework for developing labour-hour prediction models from project design features: case study in structural steel fabrication. Canadian Journal of Civil Engineering, 46(10), 871-880.

Moon, S., Hou, L. and Han, S., 2023. Empirical study of an artificial neural network for a manufacturing production operation. Operations Management Research, 16(1), 311-323.

Moselhi, O., Hegazy, T. and Fazio, P., 1991. Neural networks as tools in construction. Journal of construction engineering and management, 117(4), 606-625.

Moselhi, O. and Siqueira, I., 1998. Neural networks for cost estimating of structural steel buildings. AACE International Transactions, IT22.

Mohamed, B. and Moselhi, O., 2022. Conceptual estimation of construction duration and cost of public highway projects. J. Inf. Technol. Constr, 27(29), 595-618.

Majdoub Hassani, Z., El Barkany, A., Jabri, A., EL Abbassi, I. and Darcherif, A.M., 2021. A Comparative Analysis of Metaheuristic Approaches (Genetic Algorithm/Hybridization of Genetic Algorithms and Simulated Annealing) for Planning and Scheduling Problem with Energy Aspect. SAE International Journal of Materials and Manufacturing, 14(4), 363-374.

Mehdi, M., 2011. Parallel hybrid optimization methods for permutation-based problems (Doctoral dissertation, Université des Sciences et Technologie de Lille-Lille I).

Modular Building Institute, Permanent Modular Construction Report 2023 < mbimodularbuildinginstitute.growthzoneapp.com/ap/CloudFile/Download/LwYRWRzp > (accessed March, 2014).

Modular Building Institute, Permanent Modular Construction Report < https://www.deluxemodular.com/hubfs/Third%20Party%20Press%20Downloads/MBI2018PermanentModularConstructionReport.pdf > 2018, (accessed July, 2019).

Moghadam, M., Alwisy, A. and Al-Hussein, M., 2012, May. Integrated BIM/Lean base production line schedule model for modular construction manufacturing. In Construction Research Congress 2012: Construction Challenges in a Flat World, 1271-1280.

Mohsen, O. M., P. J. Knytl, B. Abdulaal, J. Olearczyk, and A.-H. Mohamed, 2008. Simulation of modular building construction, Proceeding, Winter Simulation, 2471–2478.

Marzouk, M., 2003. Optimizing earthmoving operations using computer simulation (Doctoral dissertation, Concordia University).

Murata, T., Ishibuchi, H. and Tanaka, H., 1996. Multi-objective genetic algorithm and its applications to flowshop scheduling. Computers & industrial engineering, 30(4), 957-968.

Mizrak, P. and Bayhan, G.M., 2006. Comparative study of dispatching rules in a real-life job shop environment. Applied Artificial Intelligence, 20(7), 585-607.

Nick.B, Steffen.F, Jan.M, Robert.P, Gernot.S, and Jonathan.W,. Modular construction: From projects to products, < modular.org/documents/document_publication/mckinseyreport-2019.pdf>, (assessed August 2019)

Nik-Bakht. M, An. C, Ouf. M, Hafeez. G, Dziedzic. R, Han. S.H, Nasiri. F, Eicker. U, Hammad. A, Moselhi. O. 2021. Value Stream Mapping of Project Lifecycle Data for Circular Construction. Proceedings of the International Symposium on Automation and Robotics in Construction. 38, 1033-1042.

Piryonesi S.M. and El-Diraby T.E. 2019. Data analytics in asset management: Cost-effective prediction of the pavement condition index, Journal of Infrastructure Systems, 26, 1, 04019036.

Pickardt, C.W., Hildebrandt, T., Branke, J., Heger, J. and Scholz-Reiter, B., 2013. Evolutionary generation of dispatching rule sets for complex dynamic scheduling problems. International Journal of Production Economics, 145(1), 67-77.

Pinedo, M.L. and Pinedo, M.L., 2016. Design and Implementation of Scheduling Systems: More

Advanced Concepts. Scheduling: Theory, Algorithms, and Systems, 485-508.

Rahman, M., and Han, S.H. 2024. Linear scheduling method-based multi-objective optimization for off-site construction manufacturing. Automation in Construction, 167, 105712.

Rashid, K., Louis, J. and Swanson, C., 2020, December. Optimizing labor allocation in modular construction factory using discrete event simulation and genetic algorithm. Winter Simulation Conference (WSC), 2569-2576.

Rashid, K.M. and Louis, J., 2020. Activity identification in modular construction using audio signals and machine learning. Automation in construction, 119, 103361.

Razali, N.M. and Geraghty, J., 2011. Genetic algorithm performance with different selection strategies in solving TSP. In Proceedings of the world congress on engineering, 2, 1, 1-6. Hong Kong, China: International Association of Engineers.

Rocchi L., Chiari L. and Cappello A. 2004. Feature selection of stabilometric parameters based on principal component analysis, Medical and Biological Engineering and Computing, 42,1, 71–79.

Steinhardt, D.A. and Manley, K., 2016. Adoption of prefabricated housing–the role of country context. Sustainable cities and society, 22, 126-135.

Salama, T., Salah, A. and Moselhi, O., 2021. Integrating critical chain project management with last planner system for linear scheduling of modular construction. Construction innovation, 21(4), 525-554.

Salama T., Salah A. and Moselhi O.(2018. Integration of linear scheduling method and the critical chain project management, Canadian Journal of Civil Engineering, 45, 1, 30–40.

Shafai, L, Simulation based process flow improvement for wood framing homebuilding production lines., MSc thesis, University of Alberta, 2012

Shahvari, O., Logendran, R. and Tavana, M., 2022. An efficient model-based branch-and-price algorithm for unrelated-parallel machine batching and scheduling problems. Journal of Scheduling, 25(5), 589-621.

Smith, S. D. 1999. "Earthmoving productivity estimation using linear regression techniques." J. Constr. Eng. Manage., 10.1061/(ASCE) 0733-9364(1999)125:3(133), 133–141.

Senghore, O., Hastak, M., Abdelhamid, T. S., Abuhammad, A., and Syal, M. G, 2004. Production process for manufactured housing, Journal of Construction Engineering and Management, 130(5), 708–718.

Spiess, A.N. and Neumeyer, N. 2010. An evaluation of R2 as an inadequate measure for non-linear models in pharmacological and biochemical research. BMC pharmacology, 10, 1, 1-11.

Taiwo, R., Hussein, M. and Zayed, T., 2022. An Integrated Approach of Simulation and Regression Analysis for Assessing Productivity in Modular Integrated Construction Projects. Buildings, 12(11), 2018.

Uslu, M.F., Uslu, S. and Bulut, F., 2022. An adaptive hybrid approach: Combining genetic algorithm and ant colony optimization for integrated process planning and scheduling. Applied Computing and Informatics, 18(1/2), 101-112.

Varela, L.R., Alves, C.F., Santos, A.S., Vieira, G.G., Lopes, N. and Putnik, G.D. 2022. Analysis of a Collaborative Scheduling Model Applied in a Job Shop Manufacturing Environment. Machines, 10,12,, 1138.

Wei, Y., Lei, Z. and Altaf, M.S., 2024. Simulation-based comparison of push-and pull-based planning in panelized construction. Automation in construction, 158, 105228.

Willmott C.J. and Matsuura K. 2005. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance, Climate Research, 30, 1, 79–82.

Xie, H., Chowdhury, M.M., Issa, R.R. and Shi, W., 2018. Simulation of Dynamic Energy Consumption in Modular Construction Manufacturing Processes. Journal of Architectural Engineering, 24(1),04017034.

Xie, H., 2001, August. A genetic algorithm approach to job shop scheduling problems with a batch allocation issue. In Proceedings of the 17th International Workshop on Artificial Intelligence, Seattle, 126-131.

Xu, C., Zhang, Y. and Xu, B., 2016. Research of Construction Safety Factors Based on the Factor Analysis Method. In ICCREM BIM Application and Off-Site Construction, 1348-1354. Reston, VA: American Society of Civil Engineers.

Yadav, S.L. and Sohal, A., 2017. Study of the various selection techniques in genetic algorithms. International Journal of Engineering, Science and Mathematics, 6(3), 198-204.

Yu, H., Al-Hussein, M., Al-Jibouri, S. and Telyas, A., 2011. Lean transformation in a modular building company: A case for implementation, Journal of management in engineering, 29(1), 103-111.

Yu, H., Tweed, T., Al-Hussein, M. and Nasseri, R., 2009. Development of lean model for house construction using value stream mapping, Journal of construction engineering and management, 135(8), 782-790

Yu, H., Al-Hussein, M., Al-Jibouri, S. and Telyas, A., 2013. Lean transformation in a modular building company: A case for implementation. Journal of management in engineering, 29(1), 103-111.

Zaalouk, A., Moon, S. and Han, S., 2023. Operations planning and scheduling in off-site construction supply chain management: Scope definition and future directions. Automation in construction, 153, 104952.

Zhang, Y., Lei, Z., Han, S., Bouferguene, A. and Al-Hussein, M., 2020. Process-oriented framework to improve modular and offsite construction manufacturing performance. Journal of Construction Engineering and Management, 146(9), 04020116.