RISK ASSESSMENT IN THE CONTEXT OF MODERN METHODS OF CONSTRUCTION LEVERAGING DATA ANALYTICS

by

Ali Tatari

B.S, Civil Engineering, Prince Sultan University, 2019

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Civil Engineering

Boğaziçi University

2022

DEDICATION

Dedicated to my mother, who was always there for me, even on the tough days.

ACKNOWLEDGEMENTS

I would like to express my gratitude to Boğaziçi University faculty members for granting me a splendid opportunity of becoming a member of a prestigious research institute and for welcoming me to a supportive, collaborative, and inspiring academic environment. Maintaining high standards of principles and techniques while being exposed to various obstacles and challenges prepared me in becoming a successful elite in my discipline. I perceive this opportunity as a significant milestone in my career development as it improves upon existing knowledge and establishes a foundation for greater future opportunities.

I would like to convey my sincere gratitude to my thesis advisor Prof. Semra Comu for her insights, vision, and guidance throughout the preparation of the dissertation. Also, I would like to express my appreciation to the committee members of my thesis defense, the juror Prof. Beliz Özorhon from the faculty of engineering in Boğaziçi University and the juror Prof. Esin Ergen from the faculty of engineering in Istanbul Technical University, for their supervision and constructive criticism. I also wish to thank my classmates and friends who facilitated the finalization of the research. From outside of my faculty, I especially would like to express my appreciation for the Office of International Relations. Particularly, I would like to thank the senior advisor Nilüfer Kırıcı for her help and support when I needed it the most.

I wish to give my appreciation to the industry professionals who have participated in face-to-face interviews and validated the reduced list of the 25 risk factors. I also express my gratitude to the 149 industry experts who participated in the questionnaire survey and to those who distributed them. Finally, I declare that there are not any known or undeclared conflicts of interest associated with this publication.

I want to express my gratefulness to everyone who has inspired me in finalizing this research. Particularly, from all of my heart, I would like to express my gratefulness to my family for their unwavering support during my life’s most challenging phase when mental and emotional support becomes of utmost importance. They facilitated fulfilling my dream and pursuing my graduate studies. I dedicate this research to them as they represent an inexhaustible source of support and hope for a brighter future.

ABSTRACT

RISK ASSESSMENT IN THE CONTEXT OF MODERN METHODS OF CONSTRUCTION LEVERAGING DATA ANALYTICS

As the construction industry moves towards more innovative, sustainable, and faster construction techniques, Modern Methods of Construction (MMC) has been regarded as a solution for meeting these demands. Nonetheless, under immense pressure, decision-makers' opposition caused by oblivious stakeholders makes for tenuous circumstances for innovation. Yet, the literature lacks a comprehensive approach for cost overruns risk assessment of implementing MMC. The study, therefore, aims to encourage the further take-up of offsite MMC in future projects of the housing sector by primarily soliciting opinions from experienced professionals. In achieving this endeavor, the study (1) identified and prioritized risk factors; (2) revealed the underlying categories; (3) proposed ways to prioritize risks; and lastly (4) developed an AI-based risk assessment model. Through the adoption of the generative adversarial networks, an Artificial Neural Networks (ANN) model will also be developed.

Study findings revealed the significant dissimilar criticality levels of risk factors. The top seven most risky factors are Safety Hazards, Direct Costs, Poor Understanding, Quality Monitoring, Scheduling and Planning, Site Layout, and Machinery and Technology with an overall frequency of occurrence of 0.736, 0.733, 0.730, 0.725, 0.708, 0.705, and 0.702, respectively. The study brings to light the inadequacy of the current industry and indicates that future research opportunities lie in the adoption of MMC. The study adds value to the literature by exploring and capturing hidden trends and patterns related to conditional dependence between risk factors. The results aid the industry stakeholders to prioritize risk factors to develop risk response measures. Accordingly, decision-makers will be capable to distribute the contingency budget on more uncertain events, which will potentially facilitate achieving project objectives and avoid racking up substantial losses.

ÖZET

VERI ANALITIĞI KULLANARAK MODERN INŞAAT YÖNTEMLERI IÇIN RISK DEĞERLENDIRMESI

İnşaat sektörü daha yenilikçi, sürdürülebilir ve daha hızlı inşaat tekniklerine doğru ilerlerken, Modern İnşaat Yöntemleri (MİY) bu talepleri karşılamak için bir çözüm olarak görülmektedir. Bununla birlikte, paydaşların ilgisiz olması ve karar vericiler üzerinde yaratılan baskı ortamı inovasyon (yenilik) için zayıf koşullar yaratmaktadır. Bununla birlikteliteratür, MİY'nin uygulanmasının maliyet aşımları risk değerlendirmesi için kapsamlı bir yaklaşımdan yoksundur. Bu nedenle çalışma, öncelikle deneyimli profesyonellerin görüşlerini alarak konut sektörünün gelecekteki projelerinde saha dışı MİY'nin daha fazla ele alınmasını teşvik etmeyi amaçlamaktadır. Bu bağlamda tez çalışması kapsamında (1) risk faktörleri önceliklendirme ile belirlenmekte 2) temel kategoriler ortaya çıkarılmakta; (3) riskleri öncelik sırasına koymak için çeşitli yollar önerilmektedir; ve son olarak (4) yapay zeka tabanlı bir risk değerlendirme modeli geliştirilmektedir. Generative Adversarial Networks yönteminin benimsenmesiyle, bir Yapay Sinir Ağı modeli de geliştirilmektedir.

Çalışma bulguları, risk faktörlerinin önemli derecede farklı kritiklik düzeylerini ortaya çıkarmaktadır. En riskli ilk yedi faktör, oluşum sıklığı sırasıyla 0.736, 0.733, 0.730, 0.725, 0.708, 0.705 ve 0.736, 0.733, 0.730, 0.725, 0.708, 0.705 ve 0.702 olmak üzere İş Güvenliği Tehlikeleri, Doğrudan Maliyetler, Yetersiz Anlayış, Kalite İzleme, Programlama ve Planlama, Saha Yerleşimi, ve son olarak Makine ve Teknoloji olarak belirlenmiştir. Çalışma, mevcut endüstrinin yetersizliğine ışık tutmakta ve gelecekteki araştırma fırsatlarının MİY'nin benimsenmesinde yattığını göstermektedir. Çalışma, risk faktörleri arasındaki koşullu bağımlılıkla ilgili gizli eğilimleri ve örüntüleri keşfederek ve yakalayarak literatüre değer katmaktadır. Sonuçlar, sektör paydaşlarının risk yanıt önlemleri geliştirmek için risk faktörlerine öncelik vermelerine yardımcı olmaktadır. Buna göre, karar vericiler, potansiyel olarak proje hedeflerine ulaşmayı kolaylaştıracak ve önemli kayıpların önüne geçilmesini sağlayacak daha belirsiz olaylara beklenmedik durum bütçesini dağıtabilecektir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

Ɗ	Data Set
SD	Standard Deviation
α_i^j	The likelihood of the occurrence of risk i as assessed by respondent j
β_i^j	The degree of loss of risk i if it occurs as assessed by respondent j
α^i	Rank of the risk factor i as assessed by the ANN
β^i	Rank of the risk factor i as assessed by the univariate analysis

LIST OF ACRONYMS/ ABBREVIATIONS

AEC	Architecture, engineering, and construction
AI	Artificial intelligence
AMDFR	Adjusted Mean Difference of Factor's Rank
ANN	Artificial Neural Networks
BDA	Big Data Analytics
CSV	Comma Separated Values
CTGAN	Conditional Tabular Generative Adversarial Network
FA	Factor Analysis
GAN	Generative Adversarial Network
GDP	Gross Domestic Product
IBS	Industrialized Building System
IDE	Integrated Development Environment
KDD	Knowledge Discovery from Data
KMO	Kaiser-Meyer-Oklin
ML	Machine Learning
MDFR	Mean Difference of Factor's Rank
MECE	Mutually Exclusive and Collectively Exhaustive
MMC	Modern Methods of Construction
OSC	Off-Site Construction
OSM	Off-Site Manufacturing
OSP	Off-Site Production
PCA	Principal Component Analysis
PAF	Principal Axis Factoring

RSI	Risk Significance Index
R&D	Research and Development
SD	Standard Deviation

1. INTRODUCTION

1.1. Research Motivation

Nowadays, the construction scene marks the development of most countries. The industry plays a key role in generating wealth and developing social infrastructure. Nonetheless, the industry has been known for lagging behind. For instance, the construction productivity rate has stagnated and continually declined over the past 50 years. Meanwhile, its counterpart, the manufacturing industry, witnessed an increase and almost doubled its productivity rate for the same time interval. This status quo is mainly attributed to the increased intricacy in processes, top management perception, and nature of projects. Having the highest potential for occupational injuries in the harsh working environment poses additional obstacles. The Architecture, engineering, and construction (AEC) industry is also fraught with rapidly arising obstacles that can jeopardize meeting project objectives and threaten the successful finalization. Construction projects suffered an alarming rate of increase in postponements, interruptions, and ultimately total abandonment.

The industry is infamous for being overwhelmed with resource planning and logistic challenges which habitually result in cost overruns, delays, and contractual disputes. Hence, innovation strategies are of utmost importance. There is a surge in the number of attempts to innovate the industry processes. Still, most of these attempts never came to fruition. Among the plethora of published research, MMC is a prominent solution and has captured the attention of several scholars and witnessed a growing interest among industry practitioners. MMC is a form of innovation in the industry that opens up several advancement opportunities in processes and technologies. Because of its prominent proven potential benefits, MMC, has been adopted worldwide. Driven by a range of factors, MMC can meet the demands for faster construction and resolve skills shortages.

Adoption of innovative building systems in a harsh environment poses a particular set of challenges. Besides, the industry knowledge and capabilities hinder innovation initiatives and contractors are still reluctant to innovate. The immense pressure placed on decision-makers from situations of strenuous opposition caused by oblivious stakeholders makes for tenuous circumstances for innovation. This is mainly attributed to the fact that innovation in the industry is accompanied by uncertainties that might hinder meeting projects objectives and threaten the financial stability of the projects. Decision-makers must adopt adequate measures to manage arising uncertainties. Though several attempts exist in the body of knowledge, what is lacking is the development of a risk management framework through a constructive methodology to wrestle the inherent obstacles. Fortunately, the Big Data Analytics (BDA) field of the data science discipline makes for a significant resolution.

1.2. Background of the Research on Data Analysis in the Industry

Big Data Analytics borrows from multiple fields, resulting in a rich intellectual tradition. The wide variety of interconnected fields includes [1], Data Mining (1980), Predictive Analytics (1989) [2], Business Analytics (1997), Knowledge Discovery from Data (KDD) (2002), Data Analytics (2010), and now the Big Data era (2012). BDA are a mere broadening of existing data analytics disciplines that fall under the overarching umbrella of BDAs. The omnipresent computational power of computers facilitates saving those data that previously we would have trashed.

To gauge the construction's overall long-term performance and avoid incurred damages, scholars have suggested numerous notable data analysis techniques. [3] identified multiple parameters for selecting appropriate methods, namely data availability, existing correlations, and output requirements. [4] analyzed the contractor project risks contingency allocation using linguistic approximation. Wasserman [5] acknowledged that the real significance of statistics comes from the fact that various disciplines are borrowing techniques from this field. [6] investigated using Factor Analysis (FA) to explore the factors that inhibit the promotion of the skeleton and infill housing system. [7] investigated the significant factors inhibiting Off-Site Construction (OSC) adoption. Despite the scholars' positive attitude and growing interest in the industry, the literature lacks a comprehensive approach. To this end, this study will adopt not only a univariate statistical analysis but also a multivariate analysis.

[8] implemented text mining, numerical data, and ensemble classifiers to perform cost prediction. Still, the industry did not witness a wide-scale adoption of similar data mining techniques for data analysis in the construction industry [9]. [10] categorized Machine

Learning (ML) tasks into one of three categories, namely supervised learning, unsupervised learning, and reinforcement learning. In the past decade, researchers have been intrigued by conducting studies on Artificial Intelligence (AI) in the AEC industry known as AI-in-the-AECI [11]. [12] conducted an in-depth review of available AI applications within the construction manufacturing industry. [11] conducted a comprehensive scientometric study appraising AI-in-the-AECI. [13] developed a prediction model for construction incidents using the latent class clustering analysis. [14] developed two suitable ML models, utilizing decision trees and naïve Bayesian classification algorithms. [15] reviewed major accidents during excavation and conducts a bibliometric analysis of risk assessment methods in recent years.

1.3. Problem Definition

Nowadays, the construction industry plays a vital role in generating wealth for countries and contributes to the development of economic and social infrastructure. The construction scene also marks the nations' economic development. According to (National Bureau of Statistics of China >> Annual Data, n.d.), it is proven that the Chinese construction industry employed about 55.5 million people in 95,400 construction-related enterprises and contributed to the nation's economy, contributing an output value of more than 235,085 (in 100 million yuan) equating to Gross Domestic Product (GDP) of 5.08% in the year 2018. Other studies revealed that the Malaysian construction sector contributed around 2.1% of the total Malaysian GDP with consistent growth of 5.3% in the year 2007 [16]. In addition, the Malaysian construction sector accounts for 8% of the total Malaysian workforce which represents an excess of 800,000 job opportunities for the years 2005 and 2006 [17]. A more recent analysis revealed that the Malaysian construction industry offers an additional 72,455 job opportunities between the years 2017 and 2018 [18]. Yet, the harsh working environment in the construction industry poses many challenges for the industry participants, especially in less developed countries. Besides, the construction industry is expected to reduce total construction costs, reduce total project schedules, meet high-quality standards, and improve the construction work environment [19].

The construction industry is infamous for being overwhelmed with resource planning, logistic challenges, and risk management which habitually result in cost overruns, project delivery delays, design defects, and contractual disputes [20]. The Architecture, Engineering, and Construction industry is also fraught with involved, challenging, and rapidly arising obstacles [11]. It has been argued that among the most pressing issues faced by construction projects are delays, which are attributed to the sector's complexity and inherent interdependence of delay risk sources [14]. Tall building projects, in particular, have been regarded as risky and suffered an alarming rate of increase in postponements, interruptions, and ultimately total abandonment [21]. The construction projects also still have the highest potential for occupational injuries despite the enormous advancement in safety management practices. Those occupational injuries include high severe work events, which result in injuries or fatalities, and low severe work events, which cause near misses or nonserious injuries [13]. This status quo results in a significant decline in productivity rates.

The construction industry has been known for lagging behind other sectors, especially in improving productivity rates. For instance, although the industry marks the nations' economic development, according to the US Dept. of Commerce, Bureau of Labor Statistics [22], the construction productivity rate declined approximately 20% between the years 1964 and 2003. Meanwhile, its counterpart, the manufacturing industry, witnessed a 120% increase in its productivity rate for the same time interval [23]. In addition, over the past 50 years, more recent studies suggest that the productivity rate in the U.S. construction industry has stagnated, whereas manufacturing industries have almost doubled their productivity levels [12]. This status quo is mainly attributed to the increased complexity in construction processes [24], the top management decision-making approach, and the nature of construction projects. Yet, the success parameters of construction projects are in time completion, within budget, and with requisite performance [25]. The construction industry is also expected to, among other things, reduce carbon emissions, reduce environmental impacts of buildings, reduce defects, and eliminate accidents [26].

Scholars have addressed the issue of improving the performance of the construction industry. Several attempts were made to innovate in the industry. Most of these attempts never came to fruition. One prominent solution is in the implementation of the MMC. It was first defined as transferring the day-to-day site activities to a factory where a controlled environment can be attained. Recently, the exploitation of robotics and other novel tools are also considered in this approach. MMC is a form of innovation in the industry, and as with all innovative methods, it opens up several opportunities for advancement in the industry's processes and technologies. MMC has been regarded as a solution to meet the industry's innovative, sustainable, and faster construction demands [27]. In a more literal sense, MMC is known for its ability to reduce total construction costs and total project schedules, meet high-quality standards, and improve the construction work environment [19]. It was also proven that wastage generation could be significantly reduced up to 100% after implementing prefabrication techniques, in which up to 84.7% can be saved on wastage reduction [28].

Yet, the industry's innovative approaches have not been fully exploited to revolutionize the construction process. Despite its indisputable benefits, contractors are still reluctant to embrace IBS [29]. Industry key players restrict MMC adoption to limited structural elements. Although several attempts were made to revolutionize the industry, most of these attempts never came to fruition. This is mainly attributed to the fact that innovation in the industry is accompanied by not only potential benefits but also uncertainties. Adoption of the smart manufacturing approach in the harsh environment of the construction industry poses a particular set of challenges. Just to name a few, among those challenges are the little replication in the configuration of components with the one-off designs [12]. This is mainly attributed to the nature of the project management discipline as it is widely acknowledged that "no two projects are the same". It is also understandable that the industry innovation efforts to enhance the construction process face multiplying concerns as compared to other non-unique projects in other industries.

Ideally, forecasting project future performance and tracking actual performance requires constant project coordination [30]. However, since uncertainties characterize innovation, these uncertainties will impact the overall project cost and, in turn, prevent top management support. Project cost analysis and control have been viewed as the main challenges of a project manager [31]. In a further research attempt, [32] stated that construction projects have uncertainties inherent in every phase of the project life cycle. Besides, for a long time, the success of enterprises for launching an innovation is hindered by risks [33]. However, significant developers have almost a consensus over the need for more innovation and prefabrication adoption in the industry in future undertaking [34]. Yet, MMC, in particular, faces increased uncertainties compared to the traditional conventional construction processes, and managers should be equipped with the necessary tools and techniques to conduct a thorough risk management process [19]. Uncertainty of cost items is a central aspect of complex projects and cost uncertainty analysis aid in performing cost estimation. It allows decision-makers to account for potential funding exposure of construction projects by identifying risks existing in a particular context to prepare risk response measures [31].

Construction managers, real estate developers, and designers are continuously trying to make a profitable business. Still, risks inherent in innovative methods can lead to potentially high risks and could threaten meeting the objectives and the successful finalization of construction projects. Especially for complex projects, cost analysis is characterized by the enormous uncertainties about different project cost items, namely price variations, productivity rates, technological development, severe inflation, and economic and market conditions, just to name a few. Meanwhile, the industry knowledge and capabilities hinder its ability to adopt MMC [35]. Hence, industry professionals will relinquish their interest in MMC and, in turn, limit their participation in the body of knowledge to add or benefit from studies. Besides, the traditional “brick and block” masonry construction technique is still used in the majority of homes.

On the other hand, the data science discipline can tackle common arising uncertainties and afford a better way to understand, categorize, assess, and develop risk response measures by revealing hidden trends in the data. The data science discipline strengthens the capabilities of decision-makers to perform several types of estimation about projects’ success parameters including cost uncertainty analysis. In general terms, cost analysis refers to a discipline that attempts to forecast the total cost of a project. In addition, diagnostic and prescriptive analysis of causes and preventive measures increase the significance of implementing advanced data science techniques. Fortunately, these advancements in data science fields also provide tools and techniques to perform an uncertainty analysis and mitigate risks associated with construction projects in the industry. Still, uncertainties inherent in construction projects threaten the financial stability of projects, and decision-makers must develop adequate measures to manage arising uncertainties. Planning is an essential function of project management [36], and developing a risk management plan is part of the responsibilities of project managers [37]. Moreover, the construction industry's performance is hindered by the stakeholders' awareness and inefficiency in tackling raising problems [38]. Several factors that exist in construction projects can jeopardize meeting project objectives and threaten projects' successful finalization. Construction project delays cause financial losses for stakeholders in the industry [39]. Besides, several cost items can drastically impact the industry, namely labor, material, operation, maintenance, disposal, and project life cycle [40]. Hence, to overcome uncertainties associated with MMC, industry decision-makers are encouraged to implement a risk management framework and use it to the advantage of construction projects.

Scholars have suggested numerous notable data analysis techniques and analytical tools using statistical methods. [41] pioneered one of the first risk management systems to improve the procedure of risk identification, analysis, evaluation, and response management. Twenty-five years later, [30] suggested that one of the critical challenges of a project management discipline is in forecasting the actual project cost. [31] asserted that the increased risk levels in today's construction projects and quantifying a project cost risk are becoming project participants' focus. Furthermore, [30] added that uncertainty in data-driven decision-making in contingency control, cash flow analysis, and timely project financing are hindering project managers from forecasting the actual project cost. This argument supports the prior claim that top management support is essential to conduct a proper risk management process. Several additional risk uncertainty analysis tools and techniques have been introduced and studied by scholars. Stochastic network analysis has been implemented in the literature to model variations in a project and produce more reliable estimations. In addition, multiple techniques were introduced to solve the uncertain nature of the networks, namely PERT (program evaluation and review techniques), MCS (Monte Carlo simulation), NRB (narrow reliability bounds methods), and PNET (probabilistic network evaluation technique).

Simulation techniques are frequently adopted by practitioners to capture the probabilities of cost items and to evaluate the project's overall cost [42]. Other than the simulation techniques, numerous data analysis techniques are commonly adopted for modeling risks and uncertainties in project cost analysis including feature-based method, case-based reasoning, and regression modeling. In addition, in the deep learning discipline, several applications abound in almost every industry in general and in the construction industry in particular. These applications extend to several areas, namely construction cost prediction, site planning and management, and health and safety, which are yet to be explored [20].

Several studies have been conducted on risk assessment of estimating the construction project's overall cost, and there has been a considerable growing interest of scholars about the implementation of a deep learning-based risk assessment approach. Besides, machine learning offers an ideal set of techniques capable of tackling the nature of complex systems [14]. In a more literal sense, Artificial Intelligence represents a powerful tool to assist in addressing estimation challenges of project cost items. In the past decade, therefore, researchers have been intrigued by conducting studies on AI in the AEC industry known as AI-in-the-AECI [11]. Yet, adopting such techniques within the construction domine remains at an initial stage [14].

Still, these approaches fall short of considering the dependence between different factors. In light of the above, developing a methodology to accurately account for conditional dependency and realistically predict project hinderers using risk management analysis is of utmost importance. Besides, none of these methods capture the conditional dependence between project cost items and are consistent with real-world settings [31]. Meanwhile, scholars acknowledged that conventional data analysis techniques have overlooked the significance of conditional dependence including existing deep learning techniques [43].

Even the well-developed Bayesian Networks (which can outperform not only traditional probabilistic risk analysis techniques but also advanced techniques) did not have the potential to outperform most deep learning methods [43]. The literature also lacks a comprehensive approach for cost overruns risk assessment of implementing MMC in housing sector projects in developing countries. The industry professionals also lack a systematic approach to assess the reliability of their decisions. They are prone to heavily rely on their own experiences and knowledge for decision-making on risk assessment [44]. A comprehensive risk uncertainty analysis plan offers several solutions and can aid in encouraging the take up of MMC by altering existing misleading beliefs about the industry.

Though several attempts exist in the body of knowledge, especially project delay analysis which predominate, these efforts are not the end of the field. What is lacking is the development of frameworks through a constructive methodology to wrestle the inherent obstacles (Sanni-Anibire et al., 2020). Nonetheless, this research is intended to make for a significant resolution. Besides, currently, the body of knowledge lacks a comprehensive application for addressing the conditional dependence pertaining to various risk factors via available deep learning techniques [20]. Furthermore, To the best of our knowledge, there are not any reported significant studies about the implementation of AI in assessing the risks of MMC. The architecture of this study is well suited to the fragmented nature of the construction industry's data. Data analysis will be implemented in this study to analyze the data from different aspects and further explain hidden knowledge.

1.4. Aim and Objectives

As previously discussed, in the problem definition section, numerous data analysis techniques from BDAs are commonly adopted for modeling risks and uncertainty. Nonetheless, when used individually, none of these methods can capture all relevant hidden knowledge existing in the data. In light of the above, developing a methodology to accurately account for assessing risk criticality levels, address conditional dependency, and leverage the power of artificial neural networks make for prudent resolution.

To the best of my knowledge, there is currently limited focus on a comprehensive risk identification and assessment framework using multiple Big Data techniques in the context of the construction industry especially Conditional Tabular Generative Adversarial Network (CTGAN). This study, therefore, attempt to address this deficiency and fills a void by presenting wide-ranging interdisciplinary research of the fields including the construction industry, data science, statistics, computer programing, Machine Learning, and BDAs all in the context of the construction industry. To this end, the study will present the current state of harnessing available computer computation techniques in risk management settings as well as discuss the future potential of tools across the multiple domain-specific sub-areas of the construction industry.

The research comprises two parts that are broken into two distinct, yet interrelated, analyses. This section will address those aims along with their objectives. The study aims to conduct a cost overruns risk analysis of implementing MMC in housing sector projects in developing countries to meet the industry needs and fill a gap in the literature. Shortly, the objectives of this study are

- assess risks inherent in the construction industry along with;
- proposing ways to efficiently prioritize risks to develop risk response measures; and finally
- categorize the risk factors and prioritize risk categories.

Study findings are expected to aid in encouraging the further adoption of MMC in the construction industry on future projects. In the interest of brevity of this analysis, the discussion here is confined to the analysis of univariate and multivariate statistics to deal with risk factors.

Lastly, using artificial neural networks, the study simultaneously integrated the two concepts (AI and MMC). The study also implemented the process of updating a posterior belief (knowledge of state about the distribution of variables) of a subset of variables when other variables (evidence variables) are observed. The study, therefore, analyzed the conditional independence of different risk factors and explain how those risks will amplify one another. To encapsulate this endeavor, the additional objectives of this study, among other things, are to (1) Develop an ANN-based deep learning model, (2) initiate a system of relevant metrics to compare the ANN variables, and (3) identify the most influencing factors for each variable in the model. Study findings facilitate and raise the level of awareness of the potential application areas in AI in the industry and find ways to extend our knowledge. Study findings are also expected to support industry practitioners in managing risks of complex projects.

The overall objective of this study is to provide decision-makers with a comprehensive understanding of construction industrialization. The study will lead the industry innovation efforts about the implementation of novel solutions to existing problems. One of the main contributions of the study is that while there have been several available proposed data-driven solutions, the literature lacks a comprehensive data analytics framework covering different aspects of BDAs in the context of the construction industry. The study will lead the industry innovation efforts about the implementation of novel solutions to existing problems by assessing potential risks.

1.5. Research Methodology

1.5.1. Research Procedure

The research procedure used in this study comprises five main steps as shown in Figure 1.1. Firstly, a review of the literature was conducted to identify the research problem in Phase 01. Subsequently, after identifying the research area of interest, a comprehensive literature review about the current adoption of MMC in the housing sector of the construction industry was carried out. Literature sources were journals, conference proceedings, trade journals, book series, books, and reports. Literature search was performed primarily via the Scopus search engine which is a web-based search engine accessible using the institution contact information of the researcher and capable of identifying multi-disciplinary publications.

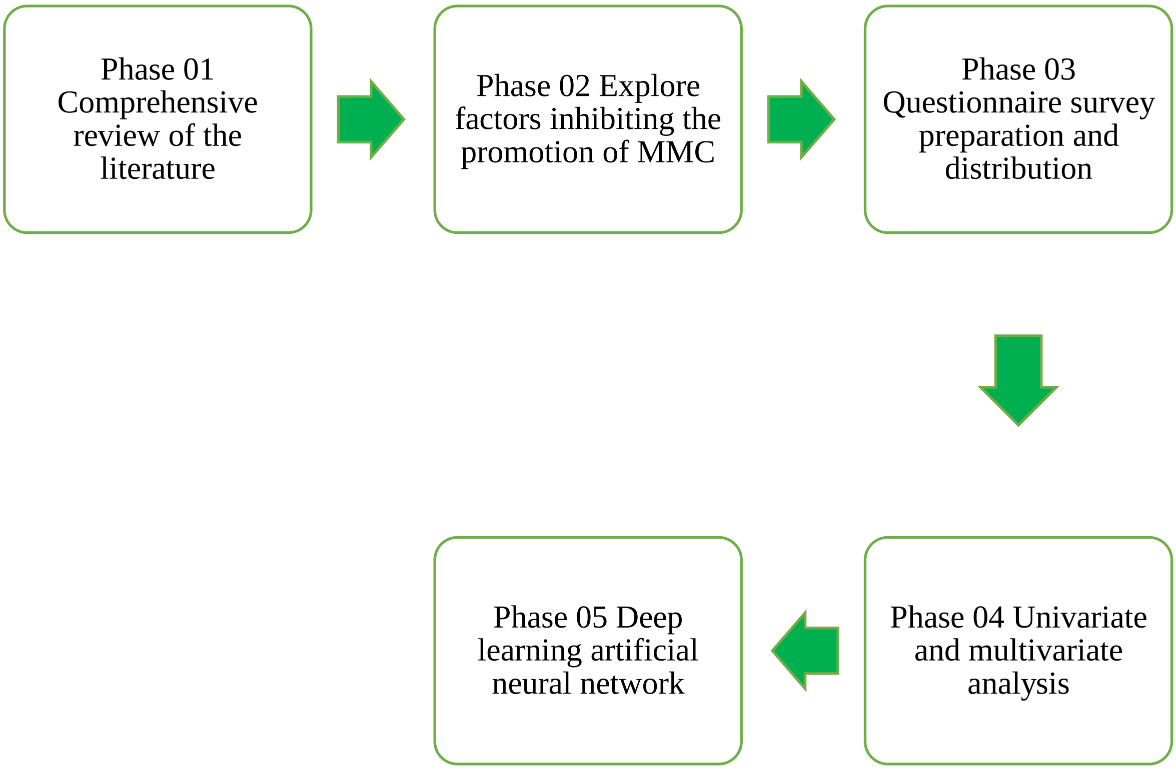

Figure 1.1. Research framework flowchart.

Then, after reviewing the literature, in Phase 02 of the research, the study explored factors that inhibit the promotion of MMC and a list of 25 factors was obtained. Interviews with focus groups were held. Phase 03 consists of the questionnaire survey preparation based on the results of the literature review conducted in the prior phase. The 25 list of risk factors extracted from Phase 2 was included in the questionnaire survey to assess their level of impact and degree of occurrence related to the adoption of MMC in developing countries. The questionnaire survey was then distributed through Google Forms. The study adopts a nonprobability virtual snowball sampling technique to conduct a structured questionnaire survey investigating respondents' attitudes towards risk factors.

The following phase consist of analyzing the data generated from the survey participants using suitable statistical methods. To this end, in Phase 04, Risk Significance Index (RSI) and $Mean\,RSI$ were then adopted to assess the relative significance of risk factors and risk factors categories. To test the existence of underlying categories, FA was adopted using Principal Component Analysis (PCA). Before proceeding to the next phase, using the risk matrix table, each risk factor was categorized into one of three levels of risk criticality represented by R1, R2, and R3 where R1 represents, comparatively, a lower risk level and R3 represents a higher risk level.

Lastly, in phase 05, Results from the first analysis will be adopted here. ANN model was developed to model the conditional dependence between different risk factors using conditional tabular GAN. In this step, the model was optimized using two primary factors, namely the batch size and the epoch number. The results from the evaluation function of the 200 iterations were then graphed in a 3D Cartesian coordinate system via a heat map-like approach. To test the stability of the conditional dependence behavior of the model, data were synthesized to observe the probability of occurrence of R3 for each risk factor. The same procedure is applied in the next step by updating a state of one evidence variable to observe the effect on the rest of the network.

1.5.2. Research Methods

In the body of knowledge, there are several useful data analytics tools that borrows from multiple techniques. This is the key reason that most of the existing work presented has mostly focused on BDA. Hence, this research adopts a triangulated study. A combination of several methods, including a literature review followed by a questionnaire survey to analyze risk factors. The study also adopts qualitative and quantitative research techniques to assess the identified risks. The research will present three different data analysis techniques, namely univariate analysis (descriptive statistics), multivariate analysis (FA), and simulation modeling (GANs). Those enabling technologies will be addressed chronologically.

1.5.3. Research Questions

Some researchers are inclined to believe that "identifying research questions at the beginning is essential to ensure conveying the purpose of the research". While I, to some extent, agree with this mindset, I will be deceiving the readers if I stated I had a clearly predefined list of questions. In the data analysis field, there is a field known as data mining. As the name suggests, it is a process where we search, extract, and capture hidden patterns in the data that we do not even know if it exists. This is referred to "the unknown unknowns." Hence, depending on the nature of data, one analysis tool might be more advantageous than the other. Consequently, by setting a number of questions in advance, I am certainly limiting the potential of the research. Nonetheless, for the sake of clarity, I will address research questions that I would have arisen had I knew there is a potential in this area. I will also exclude questions that I have answers to but were beyond the scope of this research. I will also exclude questions that I still cannot confirm nor deny without additional effort.

The research is intended to explore and present multiple tools and techniques facilitating the take up of MMC. To this end, the arising MMC related techniques will be discussed. Then the discussion will be focused on the literature. From the literature review, the most significant publications will be analyzed. The criterion is the number of citations per article. To understand the body of knowledge commonly occurring trends, the number of yearly published articles will be plotted against the progression of time. Another analysis will be directed to explore which journal and country contributes the most to the field. The number of published articles per journal and per country will be obtained. A more advanced search will analyze the underlying patterns of risk factors. Lastly, risk criticality levels will be discussed to comprehend how they compare between categories and within categories.

1.6. Scope, Limitations, and Significance of the Research

The research scope is divided into two parts. The objectives of the first part are to (1) assess risks inherent in the construction industry along with; (2) proposing ways to efficiently prioritize risks to develop risk response measures; finally (3) categorize the risk factors and prioritize risk categories. The study aims at identifying and assessing risk criticality levels between groups as well as within groups. In a further attempt to analyze the data the second part utilized deep learning-based technique. By simultaneously integrating the two concepts (AI and MMC), the study, therefore, encourage the further take-up of MMC in the construction industry by accounting for conditional dependence of different risk factors. In doing so, five objects were set namely, (1) Develop an ANN-based deep learning model, (2) initiate a system of relevant metrics to compare the ANN variables, and (3) identify the most influencing factors for each variable in the model.

Despite the remarkable potential research promises, this research, too, has its own limitations. The analysis of causality graphs was left for another study. This includes the analysis of the influence of the existence and absence of the back door criterion (back-door path) and front door criterion (front-door path) that will affect the dependency model. Without such analysis I cannot confirm the existence of causation. I am limited to correlations. Besides, in Generative Adversarial Network (GAN) model of risk assessment turned the thinking to the machine. Scholars and industry professionals are advised to never do what is referred to as “move the thinking to a machine”. Practitioners’ intuition should always be present to question the results to interpret better and reason from the data. Although it sounds a little ambitious, future research ideas will be directed towards developing a holistic risk assessment framework that compares between several available frameworks.

The study encourages the further take-up of offsite MMC in the construction industry in future projects by assessing uncertain events. The study facilitates and raises the level of awareness of the potential application areas in AI in the industry and finds ways to extend our knowledge. Study findings are expected to support industry practitioners in managing risks of complex projects. The study will lead the industry innovation efforts about the implementation of novel solutions to existing problems. The findings of the study are expected to lead to cost reduction on construction projects. Besides, scholars have asserted that the analysis of factors, that are relevant to one economy, affecting the uptake of off-site production may also encourage future involvement in other developing economies. Open issues and directions for future work will also be discussed along with possible pitfalls associated with the mere adoption of univariate and multivariate analysis in the industry.

1.7. Organization of the Study

The organization of the study cannot be expressed better without the research framework flowchart. Figure 1.3 briefly depicts the overall framework. The research will begin with an introduction to different MMC techniques. In the 3rd chapter, the research methodology will be introduced along with a comprehensive review of literature on MMC. In the 4th chapter, results and findings will be reported. Scholarly discussion of study findings is another essential part of the study which will be addressed in the 5th chapter. Lastly, conclusions and further recommendations will be drawn in the 6th.

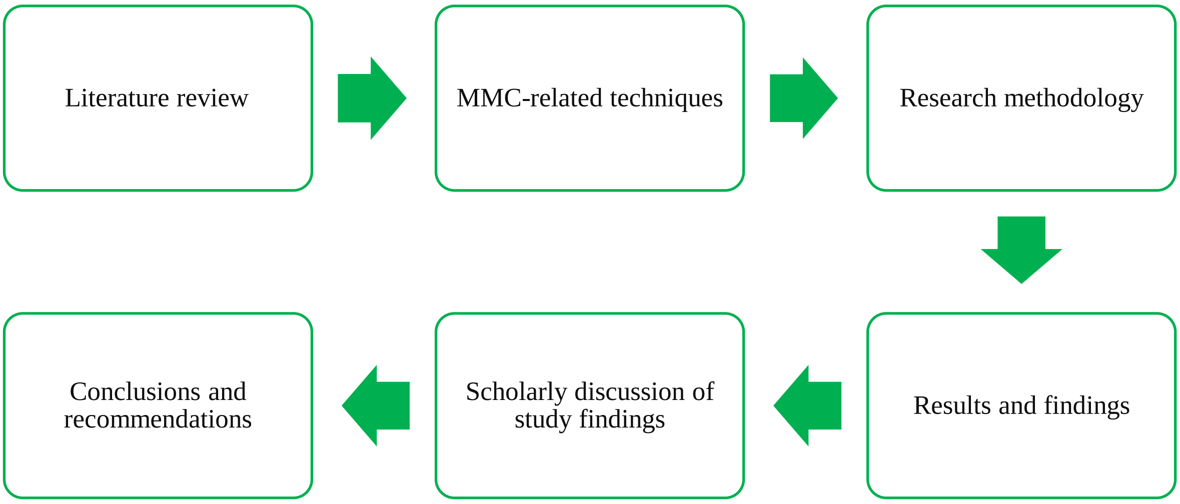

Figure 1.3. Research framework flowchart.

2. INTRODUCTION TO MMC TECHNIQUES

MMC was first pioneered in the UK. Although several UK houses have a brick outer layer and look like traditional houses, several materials are employed for MMC application, the most common being steel, wood, and concrete. In essence, MMC can be thought of as manufacturing house parts offsite in a specially designed factory [45]. Among other things, the two specific products of MMC are (1) panels – including ready-made walls, floors, and roofs and (2) modules – ready-made rooms, also known as pods.

Afterward, the manufacturing panel pods are transported to the site and assembled instantly, often within a day. In addition, electrical and mechanical systems (wiring and plumbing) could be already incorporated inside the panels, making construction even faster. Modules can be pieced together to make a whole house or flat but are most commonly used for restrooms or kitchens, where all the fittings are added in the factory. Other innovative site-based methods could also be part of MMC, such as the use of concrete molds, robotics, and other new technologies.

As documented in previous literature, Off-Site Production (OSP) falls under the overarching umbrella of MMC [35]. Hence, it is necessary to acknowledge that all OSP may be considered MMC, but not all MMC can be regarded as OSP [46]. OSP should be considered with both the product and the process of construction. In this respect, OSP aims to enhance business efficiency, sustainability, environmental performance, the predictability of delivery timescales. Hence, it is more “broadly” based than merely confined to a particular product [35].

In the same sense, it is argued that OSP may be seen as a “more realistic” means of reducing time spent on-site, improving site safety, improving quality, and addressing skills shortages. Nevertheless, for the successful implementation of OSP, strategic planning is required to change the project process from “traditional brick and block masonry” to “manufacturing and installation” [46]. Consequently, OSP is closely interwoven with the “industrialization” concept. Industrialization is defined as a production process that exploits available technologies to reduce the cost associated with manual labor, improve production, and final product quality [47]. This definition is in line with [48], who defined industrialization as involving machines and repetition for mass production and economies of scale.

There are many similarities in terms of prefabrication and pre-assembly of Off-Site Manufacturing (OSM) with off-site construction, off-site fabrication, and off-site production [34]. Nonetheless, as the term was used in the Construction 2020 Report, the term OSM was adopted by the Australian construction industry for consistency purposes. All of Australia, the U.K., and the U.S. have achieved the modular building standard. When it comes to off-site preassembly, the three countries bear many similarities in common, but only the first two have classified off-site preassembly into non-volumetric and volumetric. Thus, Australia and the U.K. share the same similar categorization of off-site systems where most Australian researchers referred to the U.K. [49].

To expand upon our understanding of MMC I will use the MMC definition framework. The framework consists of seven categories of regularized terminologies. It spans multiple types of pre-manufacturing, site-based material, and process innovation. The first five categories are the off-site and pre-manufacturing components. The first of which addresses the 3D primary structural systems. The 3D volumetric units might be a mere basic structure or fully equipped with internal and external finishes and services. The 3D units are either full volumetric units or mini volumetric. The second category is the 2D primary structural systems. The systems consist of flat panel units including floors, walls, and roof structures. [50]

The third category is the non-systemized primary structure. The category includes framed or mass engineered timber, cold/hot rolled steel or pre-cast concrete. The units range from load bearing beams, columns, walls, core structures and slabs. The category focuses more on superstructure elements than sub-structure elements. The fourth category is the additive manufacturing including structural and non-structural. The category involves the use of digital designs and manufacturing techniques to print building elements. The approach might be site-based or remote using various materials. The last category of pre-manufacturing is the nonstructural assemblies and sub-assemblies. This includes non-structural walling systems, roofing finish cassettes or assemblies, non-load bearing mini-volumetric units, utility cupboards, risers, plant rooms, pre-formed wiring looms, and mechanical engineering composites.

The last two categories are the site-based process improvement. The second to last category is the traditional building product led site labor reduction/ productivity improvements. The category includes single building products manufactured in large format, pre-cut configurations, or with easy jointing features to reduce extent of site labor required to install. Lastly, site process led site labor reduction/productivity/assurance improvements. It is the implementation of site-based innovative construction techniques. It encompasses lean construction techniques, physical or digital worker augmentation, workface robotics, exoskeletons and other wearables, drones, verification tools and adoption of technology led plant and machinery.

3. RESEARCH METHODOLOGY

As [51] suggested, triangulated studies can be powerful in gaining insights and obtaining results. This research, therefore, adopts a combination of several methods, including a comprehensive review of the literature followed by a questionnaire survey to assess risk criticality levels of risk factors. In addition, the study also adopts qualitative and quantitative research techniques. In the quantitative part, linguistic terms were used to assign a degree of impact and frequency of occurrence for each risk factor. In the qualitative part, however, the Likert scale was used to prepare the data to be further analyzed and quantify the significance of each factor. Figure 3.1. depicts the overall roadmap of the study methodology.

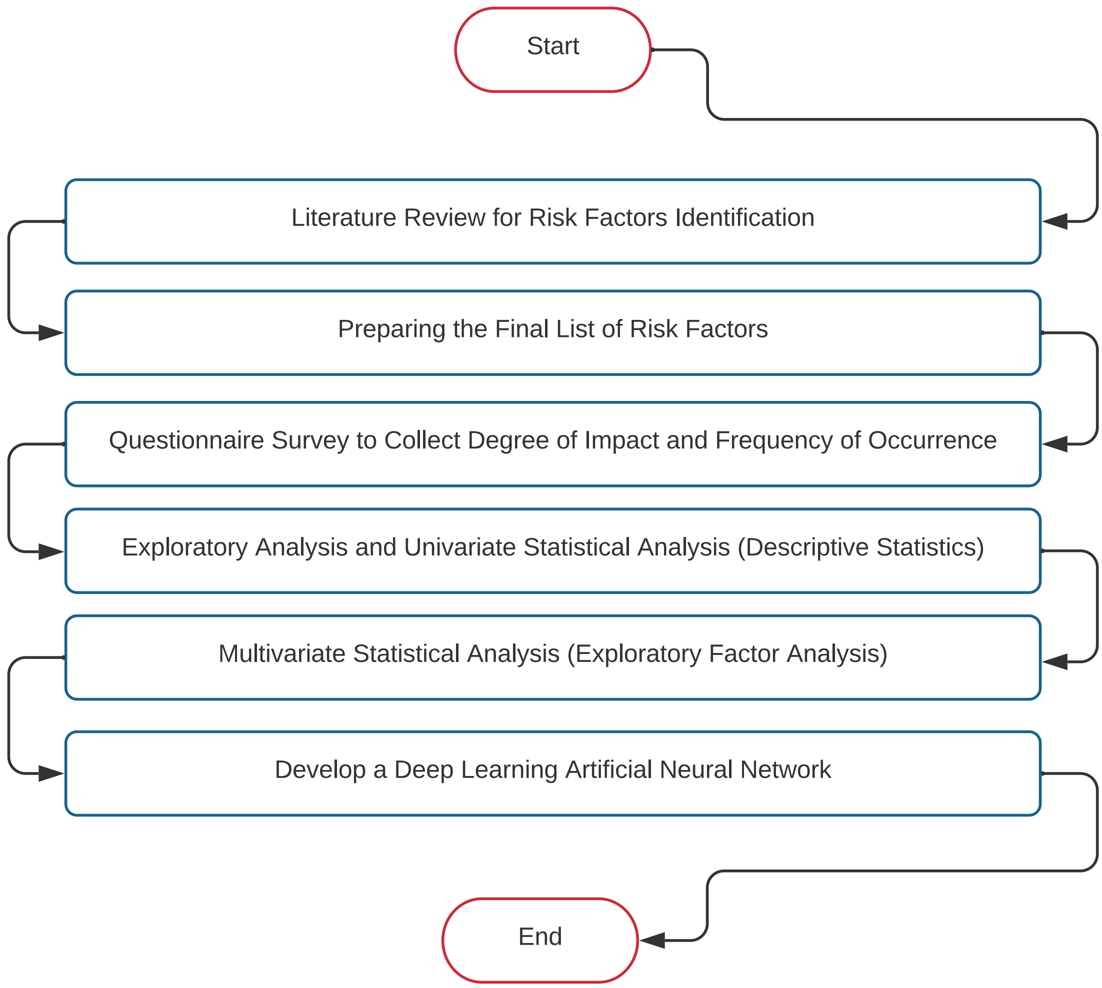

Figure 3.1. Overall methodology roadmap.

3.1. Review of Literature on MMC

The study was initiated with a comprehensive review of relevant literature to understand the industry standards and trends. To this end, the Scopus search engine was utilized to look through various publications from multiple journals and several publishers. The inclusion criteria were the discussion of the implementation of modern methods of construction in general and off-site construction in particular. The keywords list was prepared and the search operators “AND” and “OR” were determined. To be more precise, the search was limited to three parts of the publications, namely document title, abstract, and author keywords. At this stage of the research, there was not any specific consideration for the publication dates. Nonetheless, the earliest reported study dates back to 1972 and the newest dates back to 2022 with 50 years of elapsed time between the two publications. There is a clear trend of interest in the body of knowledge which will be discussed later on in this study. Nonetheless, this step yields a total of 793 publications published in several sources including journals, conference proceedings, trade journals, book series, books, and reports with a total number of publications of 428, 307, 27, 23, 7, and 1, respectively.

Leveraging the power of deep learning and recurrent neural networks, bibliometric analysis was conducted using VOS viewer software. The analysis comprises two stages. The first of which is to visually analyze the scholars’ interests by analyzing the co-occurrence of research keywords. Two networks were obtained, namely the MMC implemented techniques and the adopted methodology to facilitate the implementation. The prior will be discussed later. The latter is depicted in figure 3.2. The figure demonstrates how various tools and techniques are being considered to facilitate the implementation of MMC. The relative importance is captured by the cluster size. By reading the colors and dimensions of factors, we can witness how the industry is shifting its focus to newer techniques, namely sustainability, project stockholders, life cycle, logistics, lean construction, information theory, and finally risk assessment, the gist of the study. Secondly, a more detailed analysis of publications was conducted. Results of this step can aid in answering questions about the publication country of origin, publication date, number of citations, and number of publications per journal.

Scholars have addressed the issue of improving the performance of the construction industry. Several attempts were made to innovate in the industry. Most of these attempts never came to fruition. One prominent solution is in the implementation of the MMC. It was first defined as transferring the day-to-day site activities to a factory where a controlled environment can be attained. Recently, the exploitation of robotics and other novel tools are also considered in this approach. MMC is a form of innovation in the industry, and as with all innovative methods, it opens up several opportunities for advancement in the industry's processes and technologies. MMC has been regarded as a solution to meet the industry's innovative, sustainable, and faster construction demands [27]. In a more literal sense, MMC is known for its ability to reduce total construction costs and total project schedules, meet high-quality standards, and improve the construction work environment [19]. It was also proven that wastage generation could be significantly reduced up to 100% after implementing prefabrication techniques, in which up to 84.7% can be saved on wastage reduction [28].

Because of its prominent proven potential benefits, MMC, or construction industrialization as defined in the Chinese construction industry, has been adopted worldwide [38]. Other studies revealed that the construction industry has started to embrace Industrialized Building Systems (IBS) as a method of reducing risks related to occupational safety and health, alleviating issues for skilled workers and dependency on manual foreign labor, attaining better construction quality and productivity, and achieving the ultimate goal of reducing the overall cost of construction. It is capable of reducing risk-related activities of occupational safety, health, and productivity of craftsmen. It was also argued that it is capable of achieving the ultimate goal of reducing the overall cost of construction [29]. Yet, MMC can aid in reducing the uncertainty that occurs for several reasons, namely uniqueness of project activities; variability resulted from the trade-off between performance measures like time, cost, quality, and safety; and ambiguity from the lack of clarity, data, structure, and biases in estimates [31].

Construction managers, real estate developers, and designers are continuously trying to make a profitable business. Still, risks inherent in innovative methods can lead to potentially high risks and could threaten meeting the objectives and the successful finalization of construction projects. Especially for complex projects, cost analysis is characterized by the enormous uncertainties about different project cost items, namely price variations, productivity rates, technological development, severe inflation, and economic and market conditions, just to name a few. Meanwhile, the industry knowledge and capabilities hinder its ability to adopt MMC [35]. Hence, industry professionals will relinquish their interest in MMC and, in turn, limit their participation in the body of knowledge to add or benefit from studies. Besides, the traditional “brick and block” masonry construction technique is still used in the majority of homes.

On the other hand, the data science discipline can tackle common arising uncertainties and afford a better way to understand, categorize, assess, and develop risk response measures by revealing hidden trends in the data. The data science discipline strengthens the capabilities of decision-makers to perform several types of estimation about projects’ success parameters including cost uncertainty analysis. In general terms, cost analysis refers to a discipline that attempts to forecast the total cost of a project. In addition, diagnostic and prescriptive analysis of causes and preventive measures increase the significance of implementing advanced data science techniques. Fortunately, these advancements in data science fields also provide tools and techniques to perform an uncertainty analysis and mitigate risks associated with construction projects in the industry.

Still, uncertainties inherent in construction projects threaten the financial stability of projects, and decision-makers must develop adequate measures to manage arising uncertainties. Planning is an essential function of project management [36], and developing a risk management plan is part of the responsibilities of project managers [37]. Moreover, the construction industry's performance is hindered by the stakeholders' awareness and inefficiency in tackling raising problems [38]. Several factors that exist in construction projects can jeopardize meeting project objectives and threaten projects’ successful finalization. Construction project delays cause financial losses for stakeholders in the industry [39]. Besides, several cost items can drastically impact the industry, namely labor, material, operation, maintenance, disposal, and project life cycle [40]. Hence, to overcome uncertainties associated with MMC, industry decision-makers are encouraged to implement a risk management framework and use it to the advantage of construction projects.

Simulation techniques are frequently adopted by practitioners to capture the probabilities of cost items and to evaluate the project's overall cost [42]. In the past decade, therefore, researchers have been intrigued by conducting studies on AI in the AEC industry known as AI-in-the-AECI [11]. Yet, adopting such techniques within the construction domine remains at an initial stage [14]. Still, these approaches fall short of considering the dependence between different factors. In light of the above, developing a methodology to accurately account for conditional dependency and realistically predict project hinderers using risk management analysis is of utmost importance. Besides, none of these methods capture the conditional dependence between project cost items and are consistent with real-world settings [31]. Meanwhile, scholars acknowledged that conventional data analysis techniques have overlooked the significance of conditional dependence including existing deep learning techniques [43].

Though several attempts exist in the body of knowledge, especially project delay analysis which predominate, these efforts are not the end of the field. The industry professionals also lack a systematic approach to assess the reliability of their decisions. They are prone to heavily rely on their own experiences and knowledge for decision-making on risk assessment [44]. A comprehensive risk uncertainty analysis plan offers several solutions and can aid in encouraging the take up of MMC by altering existing misleading beliefs about the industry. What is lacking is the development of frameworks through a constructive methodology to wrestle the inherent obstacles [21]. The literature also lacks a comprehensive approach for cost overruns risk assessment of implementing MMC in housing sector projects in developing countries.

There is a gap in the body of knowledge on a comprehensive application for addressing the conditional dependence pertaining to various risk factors via available deep learning techniques [20]. To the best of our knowledge, there are not any reported significant studies about the implementation of AI in assessing the risks of MMC. The architecture of this study is well suited to the fragmented nature of the construction industry’s data. Data analysis will be implemented in this study to analyze the data from different aspects and further explain hidden knowledge.

3.2. Step Two: Risk Identification

This study aims to understand the development status of the implemented MMC in developing countries. For this reason, I explored factors that inhibit the promotion of MMC by following a research framework proposed by Cao et al. [6]. After reviewing the literature, I looked through the top 35 most cited publications. From which, I stumbled upon 15 research articles published between the years 2002 and 2018. The highest cited publication has 242 citations and the lowest has 69. From these publications, I initially obtained a list of 297 risk factors. Refer to Table 12 appendix E. A preliminary risk categorization system was created to categorize the 297 risks. The primary categorization system comprises 10 categories including Site conditions & Environment, Legal concern, Cost & Economy, Quality, Safety, Time, Inbound/outbound Logistics, Knowledge & Skills, Design, and Equipment & Tech with a total number of risk factors of 17, 26, 31, 11, 17, 42, 52, 56, 25, and 20, respectively. This will facilitate comparing and combining similar risk factors. I applied 10 rounds to reduce the list from 297 to 25 risk factors.

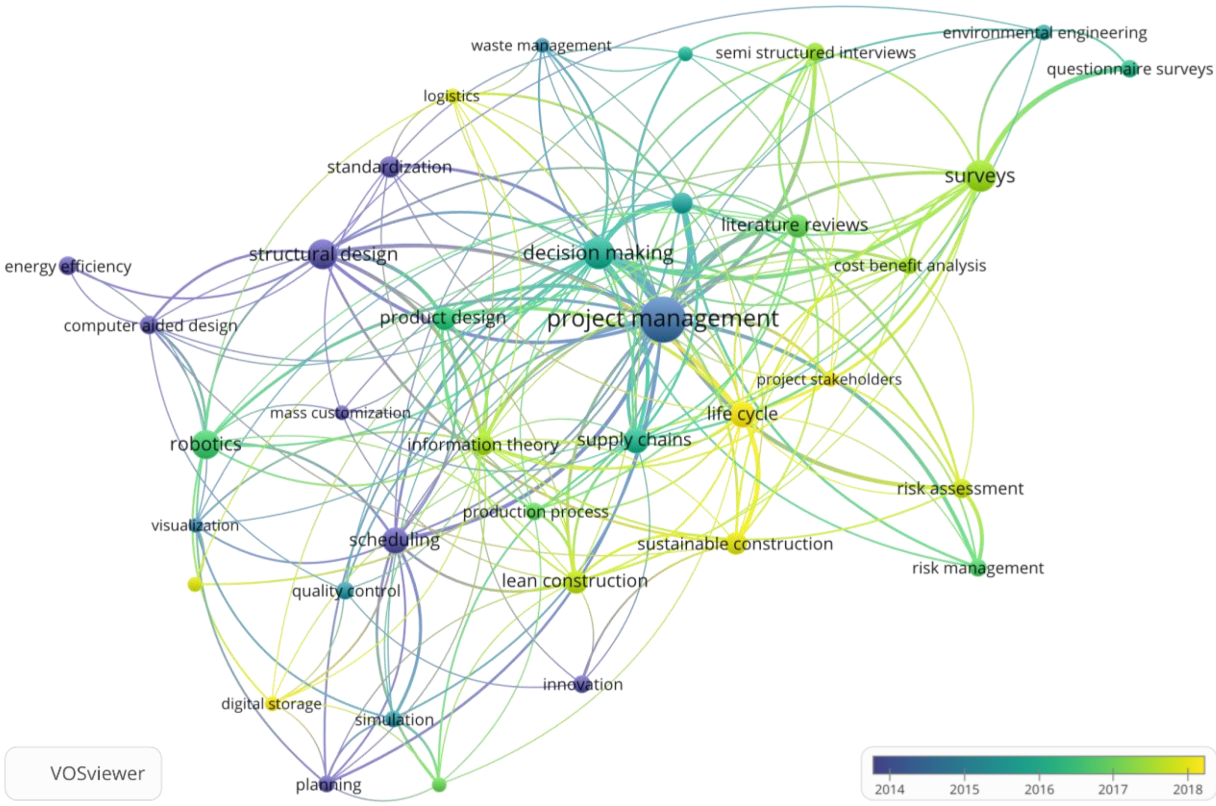

Figure 3.2. Various tools and techniques which facilitate the implementation of MMC.

Face-to-face interviews were then held with focus groups consisting of the industry professionals to validate the final list of risk factors. The three participants of the focus group had work experience of over 25 years in the industry. They are entrepreneurs, investors, and company directors. In addition, they have been engaging in at least two to three construction-related activities in multiple companies in developing countries. Table 3.1 summarizes the demographic information of focus group members.

3.3. Step Three: Conducting Questionnaire Survey

The study adopts a nonprobability sampling technique known as the snowball sampling technique [52]. The snowball sampling technique allows existing study participants to recruit future participants to represent the industry experts from among their acquaintances. The sample group is said to grow like a rolling snowball, and the sample builds up. Since the median through which the samples were collected was virtual social networks, this technique is called virtual snowball sampling [53]. A structured questionnaire survey investigating respondents' attitudes towards risk factors was conducted in several developing countries, representing the overall development of the construction industry in developing countries, especially the development of MMC at the initial stage. The questionnaire survey comprises two sections. The first section was developed to gather information about the respondents' background and experience and check the validity of their contribution to the study. The second section was devoted to collecting the data to be further analyzed and interpreted.

Table 3.1. Demographic information of focus group members.

Interviewee	Years	Countries of experience	Current position
Participant 01	31	Turkey, Saudi Arabia, and Syria	Chief operating officer and investor in real-estate development
Participant 02	29	Saudi Arabia, Egypt, and Turkey	Chief executive officer and investor in pre-cast concrete elements and modules factory
Participant 03	26	China, Saudi Arabia	Corporate relations manager and investor in construction materials supply and equipment

The questionnaire was pre-tested to ensure that questions are well-defined, practical, and not overly burdensome to answer. A total of 149 valid responses were obtained from experts operating in the industry, including 53 from site engineering, 39 from project management and procurement, 21 from quality and safety management, 18 from head office and consultancy, and 18 from real-estate design and development. The study participants had a minimum of 10 years of work experience and up to 28 years of experience. The participants had job positions in five different sectors in the industry, namely construction (contracting), consultancy (consultancy firms), manufacturing (manufacturer & supplier), design (architect, structure engineer, etc.), and real estate business (real estate developer). All of the questionnaire participants from or had worked in the construction industry of a developing country. In addition, they had also worked in some form of off-site construction in the housing sector of the construction industry. Details about the respondents are summarized in Table 3.2.

Table 3.2. Research team background.

Criteria	Category	No. of Respondents
Experience Range (In years)	From 10 years to 15	66
	From 16 years to 20	51
	From 21 years to 25	22
	From 26 years to 30	10
Industry Sector	Construction (Contracting Companies)	102
	Consultancy (Consultancy Firms)	14
	Manufacturing (Manufacturers & Suppliers)	13
	Design (Architect, Structure Engineer, etc.)	12
	Real Estate Business (Real Estate Developers)	08
Respondents' Positions	Site Engineering	53
	Project Management & Procurement	39
	Quality and Safety Management	21
	Head Office and Consultancy	18
	Real-estate Design and Development	18
Respondents by region	Gulf Region	98
	Asia & Pacific	72
	Middle East (Excluding the Gulf countries)	31
	Africa	04

In addition to the six general questions about the participants, the questionnaire had 25 questions to assess the risk level of 25 risk factors. In each question, the participant is asked to identify the level of impact and frequency of occurrence of each risk factor. In doing so, the participants are asked to use a five-point Likert Scale. The level of impact intensity ranges from 1 to 5, representing very low, low, moderate, high, and very high, respectively. The degree of occurrence of Likert scale adopts the following linguistic terms rare (0%-20%), unlikely (20%-40%), possible (40%-60%), likely (60%-80%), and certain (80%-100%) to represent 0.2, 0.4, 0.6, 0.8, and 1.0, respectively.

3.4. Step Four: Conducting Univariate Statistical Analysis

In this step, the Risk Significance Index (RSI) from [54] was adopted to assess the relative significance among the 25 risk factors. RSI can be calculated by the following formula:

$$\mathrm{RSI}^{\mathrm{i}} = \frac{1}{\mathrm{N}} \sum_{\mathrm{j}=1}^{\mathrm{N}} \alpha_{\mathrm{j}}^{\mathrm{i}}\ \beta_{\mathrm{j}}^{\mathrm{i}}\ , \tag{3.1}$$

where $\mathrm{RSI}^{\mathrm{i}}$ denotes the RSI value for risk i, $\alpha_{\mathrm{i}}^{\mathrm{j}}$ is the likelihood of the occurrence of risk i as assessed by respondent j, $\beta_{\mathrm{i}}^{\mathrm{j}}$ is the degree of loss of risk i if it occurs as assessed by respondent j, and N is the total number of effective respondents. Table 4 represents the $\mathrm{RSI}^{\mathrm{i}}$ for each risk factor.

Univariate statistical analysis was also implemented in the study after applying the multivariate statistical analysis and group risks to several constructs. By applying the mean of risk significance index, I can identify the risk significance index not only for certain risk factors but also for the overall risk level of a category. The Mean RSI can be calculated by the following formula:

$$\text{Mean RSI} = \frac{1}{\mathrm{N}} \sum_{\mathrm{i}=1}^{\mathrm{N}} \mathrm{RSI}^{\mathrm{i}}\ , \tag{3.2}$$

where $Mean\ RSI$ denotes the mean risk value of RSI, $\mathrm{RSI}^{\mathrm{i}}$ is the RSI value for risk i, and N is the total number of constructs.

3.5. Step Five: Conducting Multivariate Statistical Analysis

The following step is performed to comprehend how risk factors inherent in the industry might affect the implementation of MMC and how they are related together. A powerful tool to address these concerns is FA. It was pioneered by the English psychologist Charles Spearman in 1904. FA is used for identifying groups of items, which in many cases are survey questions, that are strongly correlated. I assume that the strongly correlated items represent some reflective factor or construct because they move together consistently. In other words, I want to explore the underlying categories. There are three main applications for FA, namely data reduction, exploring data for patterns, and confirming a hypothesis of the factor structure. In this study, I am interested in exploring data for patterns as stated early on in this research. Among the three available extraction methods, Maximum likelihood, Principal axis factoring (PAF), and Principal Component Analysis, I adopted the latter. Accordingly, I utilized FA to group each number of factors and test if these risk factors load together and measure the same construct to form principal components, hence the name PCA.

This can be achieved by comparing their factor loadings. Higher values indicate that the factors stick together and load on the same construct. Despite FA telling us if the items load together, it does not answer what constructs are being measured. This study did not have a hypothesis about how the factors will load, in what ways they will load together, or an idea formed in advance about how they will divide up. Nonetheless, I have some expectations in mind. I expect constructs to be formed to measure industry knowledge, the risk-return trade-off, and conflicts in time and cost of the MMC. Hence, I would find several constructs that some factors would load cleanly on those constructs in these data.

The Statistical Package for the Social Science (SPSS, IBM Corporation, Armonk, NY, USA) was employed in this paper to test the reliability of the collected data and to perform the multivariate calculations. I gathered data about 25 risk factors from 149 respondents. PCA was used for the extraction method with 25 iterations. There is a slight difference between FA and PCA, but often, these terms are used interchangeably. Researchers generally accept PCA as achieving the same goal as FA [55]–[58]. However, if I am interested in reducing the observed variables down to their principal components while maximizing the variance accounted for in the variables by the components, then I should be using PCA [59]. In the first extraction attempt, I chose direct oblimin rotation from the oblique category of rotation methods. That is to say; the constructs are correlated together. The extraction did not converge with this type of rotation. I checked the component correlation matrix to determine the rotation method. When I took the absolute value of the values in the matrix, I found the results not exceeding 0.32. Then, it was clear that the rotation is orthogonal.

In the second attempt, I set the varimax rotation method from the orthogonal category of rotations. In the literature, there is no obvious guidance about how many constructs to choose from. One argues it is enough to retain all factors with eigenvalues exceeding 1.0 and accept it as the default in most statistical software packages [59]. In contrast, scholars showed a broad consensus that this method is among the least accurate methods for determining the number of principal components to retain [57]. Others argue we need to not just rely on the eigenvalue but also observe how cleanly factors are loading on the constructs [57]. Hence, keeping everything constant, I repeated the extraction procedure nine times by only altering the number of constructs. I started with 2 up to 10. From the repetition, I found that eight constructs provide the cleanliest loading among the others and hence represent meaningful factors. The suppression factor was set to be 0.3. There are several theories, but the most common are suppress small coefficients less than 0.3 or 0.4.

Lastly, the Kaiser-Meyer-Olkin (KMO) statistic (adequacy test) is used for measuring sampling adequacy and evaluates the correlations and partial correlations to determine if the variables are likely to merge into components. The measure of sampling adequacy test yields a 0.566 greater than a popular cut-off of 0.5 and is close to 0.6, which is considered perfect. Bartlett’s test of sphericity is carried out to test the null hypothesis that the correlation matrix is an identity matrix, i.e., all terms of the diagonal are equal to 1 and the other terms are equal to zero. A large value of Bartlett’s test statistic for sphericity with a small significance level ($p<0.00$) indicates that the correlation matrix is not an identity matrix and FA can be applied. The data passed Bartlett’s Test of Sphericity with a significance level of 0.001 and an approximate Chi-Square value of 378.291 and a degree of freedom of 300.

3.6. Step Six: Develop the Deep Learning ANN Model

For many years, the nonlinear relation between attributes has been investigated through the implementation of ANN to generate a logic between different variables. ANN and other statistical techniques are used to capture the common causes of hidden conditional dependence between different attributes in the model [13]. Not until the recent technological advancement has finally come in handy, analysis of big data was perceived to be costly. The term costly is used to refer to those operations which demand a lot of computational resources to be used, such as the CPU, GPU, memory cards, etc. The research methodology, however, did not necessitate the use of enormous computational power.

The only part of the methodology where the required computational power was comparatively high is in what I refer to as the model hyperparameters optimization, namely the epoch number and the batch size. This will be addressed in more detail in the subsequent sections. Consequently, the research follows [43] methodology of training an ANN model from existing data. Using the power of deep ANNs, [43] in their previous research developed a tabular generative adversarial network that generates high-quality and fully synthetic tables while simultaneously generating discrete and continuous variables. They then introduced the Conditional Tabular Generative Adversarial Network that enables us to use a conditional generator to address the challenges of imbalanced discrete columns that pose difficulties in the modeling phase.

To encapsulate this endeavor, a computer-based approach that implements the python-based CTGAN software package was used. The PyCharm Integrated Development Environment (IDE) Community Edition 2020.3.3 x64 was used to run the Python interpreter version 3.8.9 and perform the computations. After installing the sdv package using the pip installer package, CTGAN was imported from sdv.tabular package. For the evaluation process, the evaluate software package was imported from the sdv.evaluate package. The seed was set to be a random number of (10) to increase the possibility of reproducing comparable results. The collected data from 149 questionnaire surveys were imported as a Comma Separated Values (CSV) file format and saved as a Pandas Dataframe.

Out of the 149 questionnaire surveys, 100 surveys were used to learn and train the deep learning model. The remaining 49 samples will be then used in the validation process of the developed deep learning model. In the CTGAN learning model, all the parameters including epochs, batch size, dimension of the generator, and the dimension of the discriminator were initially set to be as follows 500, 100, (256, 256, 256), and (256, 256, 256), respectively. None of the primary keys or anonymized fields were assigned to the model. Furthermore, all of the data have been discretized into one of three possible values ranging from L1 (lower risk level), L2 (moderate risk level), and L3 (higher risk level) as described in the previous step. Refer to table 3.3.

In the next part of this step, I tried to optimize the procedure of learning the model. Two primary factors were considered, namely the batch size and the epoch number. For the prior, there are three available types mentioned in the literature including batch mode (the batch size equals the complete dataset), mini-batch mode (the batch size is greater than 1, usually a number that is divisible by sample size), and stochastic mode (the batch size equals 1). In the CTGAN, however, the batch size must be a multiple of 10 in order to use the CTGAN package. Unlike the batch size, the epoch number does not have limitations. Out of the model, 100 thousand samples were synthesized.

Table 3.3. Risk matrix table.

Frequency of Occurrence	5th	L2	L2	L3	L3	L3
	4th	L2	L2	L2	L3	L3
	3rd	L1	L2	L2	L2	L3
	2nd	L1	L1	L2	L2	L2
	1st	L1	L1	L1	L2	L2
		1st	2nd	3rd	4th	5th
		Degree of Impact				

The synthesized data were then evaluated using the 49-testing data with the following metrics. These metrics evaluate the synthesized vs. real data via a score that should be maximized. In the evaluate function, the aggregate was set to be false to not only end up with an average score of all metrics but also to get the score of individual metrics. The evaluate function is based on four well-known metrics, namely chi-squared, inverted Kolmogorov-Smirnov d statistic, Bayesian network log-likelihood, log-likelihood, and discrete Kullback–Leibler divergence.

The code performed 100 iterations of changing epoch size and batch size having 10 possible values each starting from 50 up to 500 with an incremental increase of 50 for epoch number and 10 up to 100 with an incremental increase of 10 for the batch size. After performing the first 100 iterations, from the data, it was clear that an additional 100 iterations are necessary to cover the range starting from 550 up to 1000 with an incremental increase of 50 for epoch number. In both 100 iterations, a sample size of 100 thousand instances was synthesized and evaluated against the 49-testing data as previously illustrated.

This allows us to better understand the model performance when changing its parameters. To automate the evaluation process, for every 100 iterations, a python code was developed. The code comprises two nested for loops that are capable of swapping between several learning parameters and performing learning, sampling, and testing for each one of the 100 iterations.

After swapping between several ANN hyperparameters (a.k.a. learning parameters) and performing learning, sampling, and validating for several iterations, the hyperparameters of the CTGAN model, namely epoch number and batch size were optimized for the values of 300 and 10, respectively. The other learning hyperparameters, namely the dimension of the ANN generator and the dimension of the ANN discriminator were finally set to be as follows (25, 25, 25) for the generator neural network and (25, 25, 25) for the discriminator neural network (a.k.a. critic network). No primary keys or anonymized fields were assigned to the model. After inspecting the scores from the evaluation function, I realized that those hyperparameters will optimize the ANN model.

Lastly, the model was then saved with the filename extension using the .pkl extension to highlight that the serialization protocol used is Pickle. This will make it easier to be shared without sharing the questionnaire survey forms and breach the participants' data privacy agreement. The generated sample was also saved but this time using the CSV file format to be easily imported to the IDE of programming languages.

More detail will be addressed in the discussion section. The results from the evaluation function of the 200 iterations were then graphed in a 3D Cartesian coordinate system against the change in epoch number and the change in the batch size. The scattered points were then connected through the implementation of a special graphing tool from the Matplotlib package for the python programming language.

To understand the model performance while changing the epoch number and the batch size, the scattered data points were plotted against the two axes as mentioned in the methodology section. A heat map-like approach was adopted to represent higher values of the evaluate function with different colors. Figure 4.6. illustrates the plotted scattered data points. Brighter colors indicate a more representative model that has higher scores. If I took a slice of the model to turn it into a 2D graph, I can start to observe that as the batch size increases, the epoch number decreases given that the overall fitness score is held constant. From a ML developer's point of view, this is completely an expected behavior (interaction between the epoch number and the batch size). When the features are fed gradually to perform the training, the model will not need to see the data multiple times as compared to when I feed it all at once. In the first case, the learning computations will take a longer time because the data will be fed in multiple batches. Besides, the epoch number will be increased to result in an even longer training time.

After training the model in multiple iterations, I generally did not feel the need to worry about improving the training time as it did not matter much especially when this is only performed once. On the other hand, I want to maximize the evaluation score while decreasing the epoch number. The increase in both the epoch number and the neural network dimension (nodes in each layer) will result in overfitting. This happens when the model starts to memorize the fed features instead of capturing existing trends in the data. The reason overfitting should be avoided is that when the model is evaluated on the training data, the performance will be maximized. However, this performance cannot be maintained when the model is fed different real-world data. In the study, this is not an issue because the model validating data were not exposed to the model previously. In other words, the performance is determined in a way that overfitting is not common as I compare the fitness of the sampled data against the testing data (a sample of 49 instances). However, I chose to reduce the required training time by reducing the epoch number and the batch size. The final model was saved with the serialization protocol pickle.

After calculating the probability of L3 risk level for each risk factor, I found that some risk factors have significantly higher risk frequency of occurrence. To choose from the list the riskiest risk sources I had two inclusion criteria in mind. The first of which is the relative occurrence of a risk criticality level of 3. Secondly, I looked for the risk that is significantly less than the previous one. As a result, a list of 7 risk factors was developed.

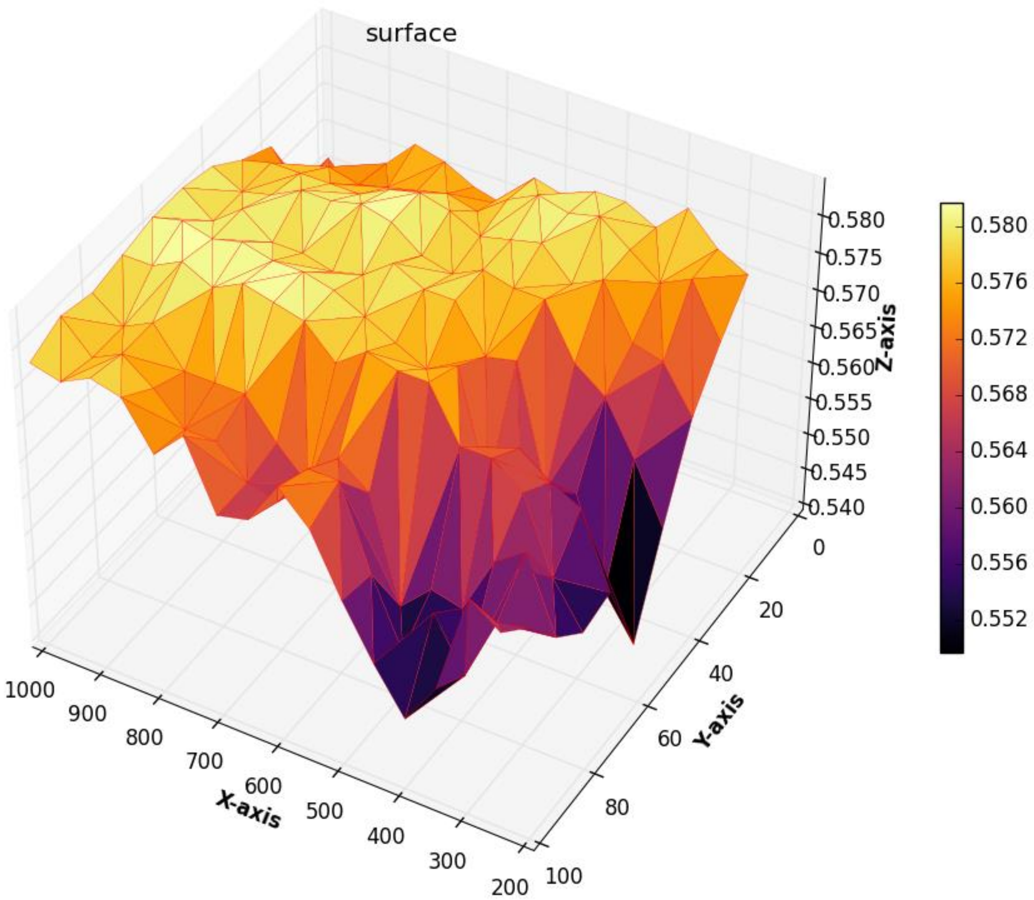

Figure 3.3. Plotted scattered data points of the evaluation function

3.7. Step Seven: Comparing the Risk Ranks of ANN and Univariate Analysis

In addition, a significant analysis is the analysis of comparing the results of ANN and multivariate analysis. Since each measure of risk criticality is different, the rank will constitute the comparison. In the body of knowledge, there are plenty of tools to be utilized. A very remarkable tool is the Kendall Rank Correlation Coefficient. The Coefficient is also referred to as Kendall's tau. The coefficient measures the rank correlation by exploring the similarity of the orderings of the data when ranked [60].

Kendall correlation methods are non-parametric rank-based correlation test to measure rank correlation. The coefficient ranges from 1 and -1, where 1 represents high Kendall correlation between the variables and -1 represents low Kendall correlation between ranks. If the tau coefficient is zero then this indicates there is not any association between the two ranks. Depending on the rank we can conclude whether the factors have similar of dissimilar ranks. In exploring the factor, the R programming language will be used. Unlike some previous analyses, here there is not any randomness. Hence, it was not necessary to set the seed and replication of the results is guaranteed.

4. RESULTS AND FINDINGS

From the body of knowledge, I gathered 793 related publications. From those publications, I analyzed them with VOS viewer to construct a network representing the available MMC techniques in the industry. Figure 4.1. illustrates those techniques. The size, color, and location of the techniques in the network represent the significance, adoption time, and relevance to neighboring techniques. From the graph, it is evident that the industry focus is shifting from traditional prefabrication techniques and IBS to robotics, automation, and lean construction between the years 2014 and 2016. Between the years 2016 and 2018, the industry focus is also shifting towards different techniques. The core of this focus is comprised of off-site construction, sustainable construction, BIM-enabled off-site construction, and modular construction. Modular construction, in particular, has rapidly captured the scholars’ interest. Successful implementation can be found in several industries including the Chinese industry and Singaporean industry.

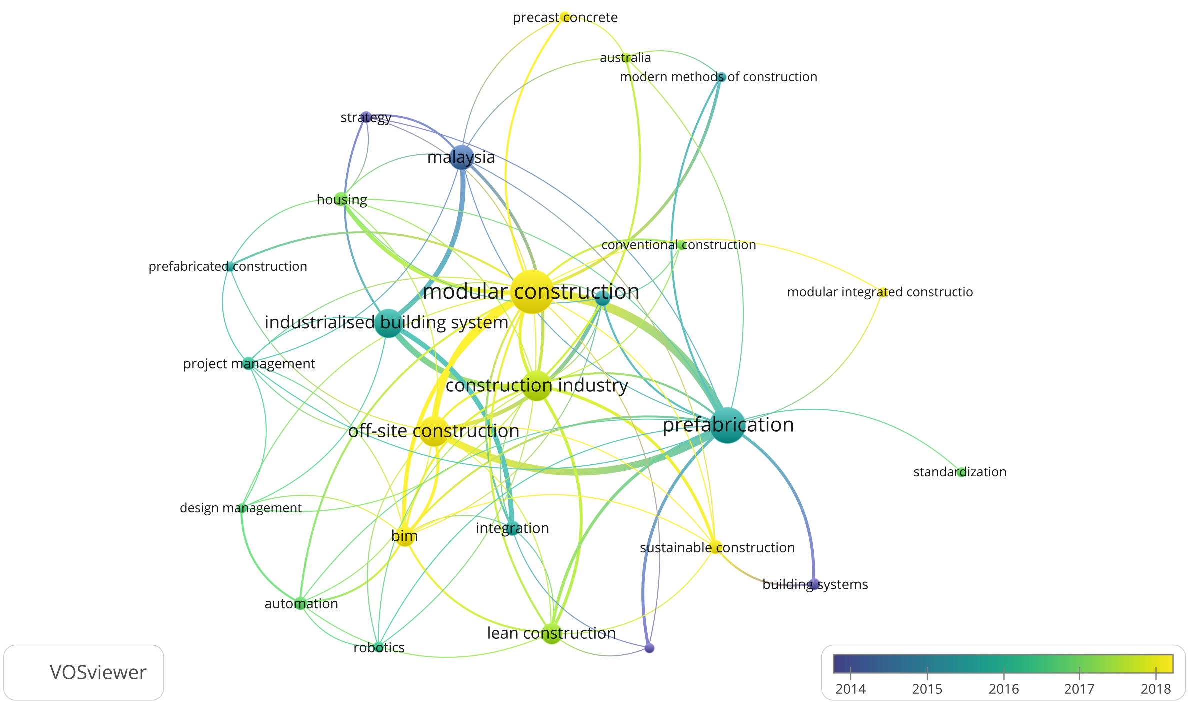

Figure 4.1. MMC techniques in the industry.

Another detailed comparison about the number of citations was made, but this time citations per publication was considered. The three most cited are [61], [62], and [63] with a total number of citations of 242, 201, and 178, respectively. Refer to figure 4.2.

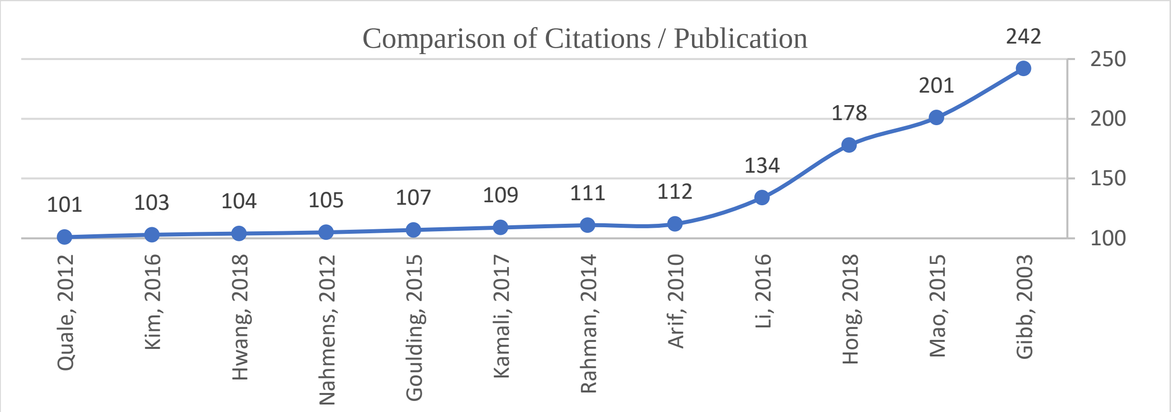

Figure 4.2. Comparison of citations/publication.

To analyze the trend of the number of yearly published publications, the 793 publications were categorized by the publication date. Then a linear regression analysis was performed to understand the underlying relationship between the number of documents and the year of publication. It was found out that there exists a strong linear relationship between the number and year of publications with the model's R^2 of 0.846. This indicates that the model explains about 84.60 % of the variation. More about that will be discussed in the subsequent sections. Refer to figure 4.3.

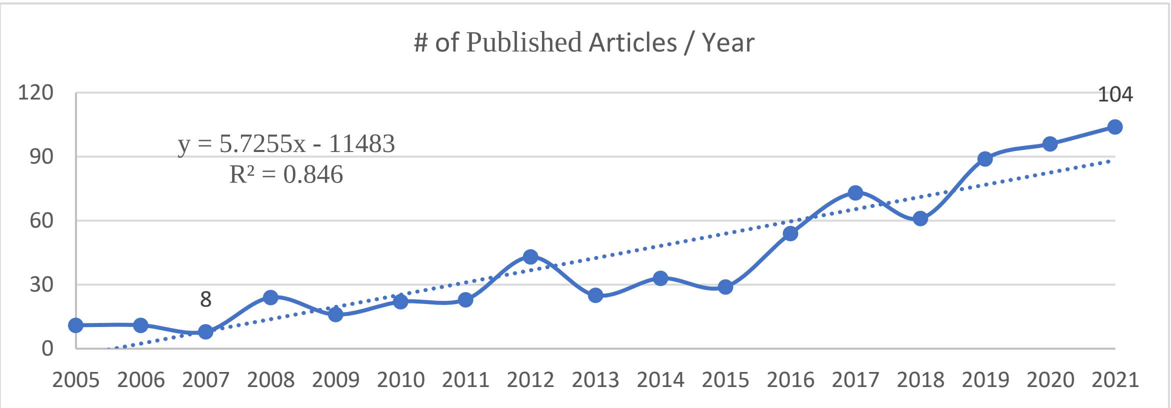

Figure 4.3. Number of published articles/years.

The study review of the literature also revealed additional interesting results. Firstly, a comparison of the number of citations for each academic journal was made. From which, I concluded that the most-cited journals in this area of research are Journal of Cleaner Production, Automation in Construction, Journal of Construction Engineering and Management, Journal of Management in Engineering, and Journal of Architectural Engineering with a total number of citations of 1213, 979, 629, 565, and 441, respectively. Figure 4.4. depicts the results. Lastly, the number of articles and citations per country was analyzed. It was found that the most contributing countries (by the number of citations) are the United States, the United Kingdom, China, Australia, and Malaysia with a total number of citations of 1833, 1807, 1777, 1647, and 1508, respectively. Refer to figure 4.5.

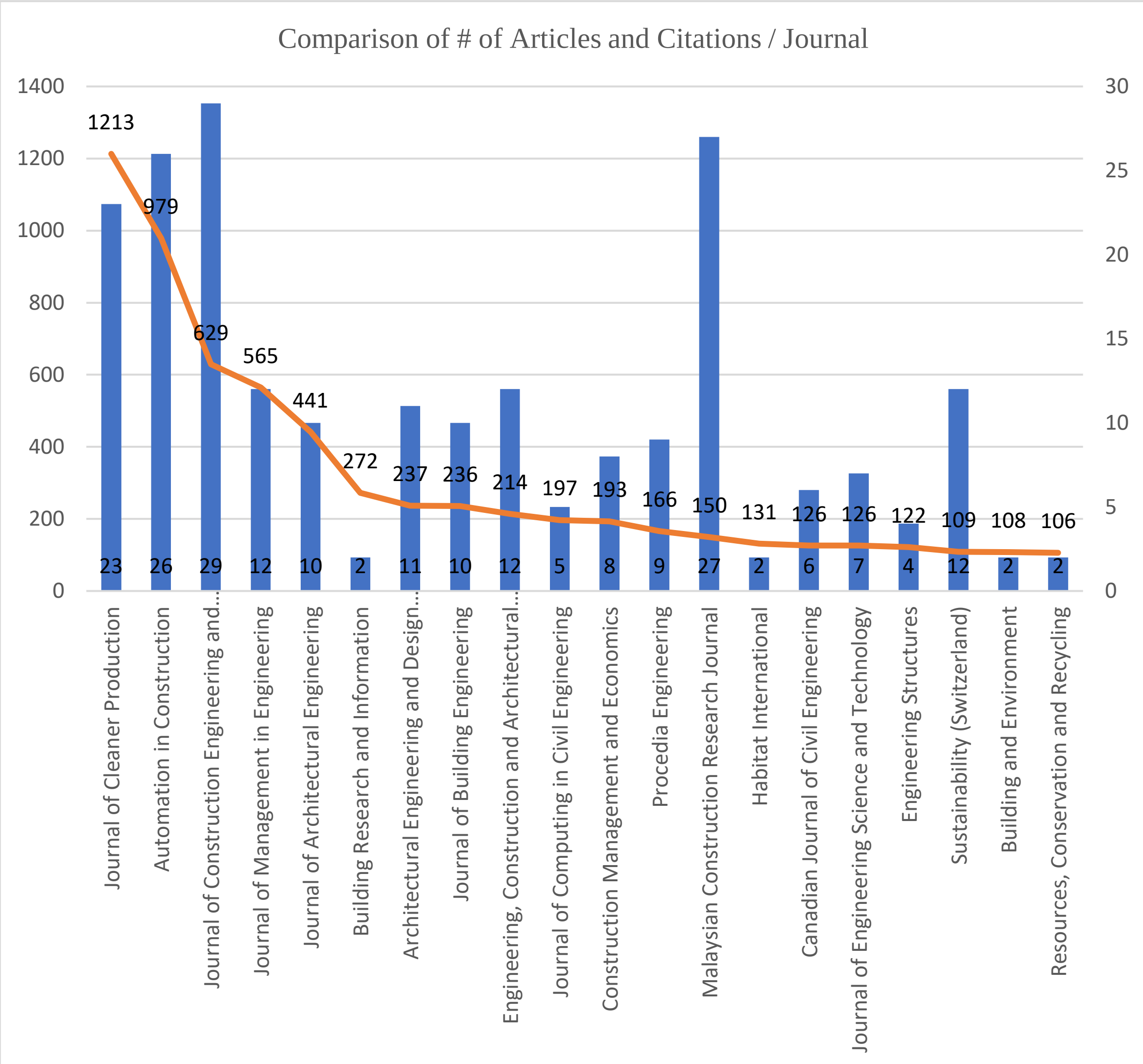

Figure 4.4. Comparison of articles and citations/journal.

After reviewing the literature and looking through the top 35 most cited publications, 15 research articles published between the years 2002 and 2018 were selected. Table 7.1 in Appendix A illustrates detailed information about the articles. The 15 publications were published in Journal of Cleaner Production, Building Research and Information, Journal of Management in Engineering, Journal of Construction Engineering & Management, Journal of Building Engineering, Engineering, Construction & Architectural Management, Journal of Architectural Engineering, and Canadian Journal of Civil Engineering. Refer to Table 7.2 in Appendix B for detailed information about the publications' source. Consequently, an initial list of 297 risk factors was developed. After reducing the list to 25 risk factors, face-to-face interviews were held with focus groups to validate the final list. Refer to Table 7.3 in Appendix C for detailed information about the risks' sources and risk preliminary categories. Risk codes and abbreviations are illustrated in Table 7.4 in Appendix D.

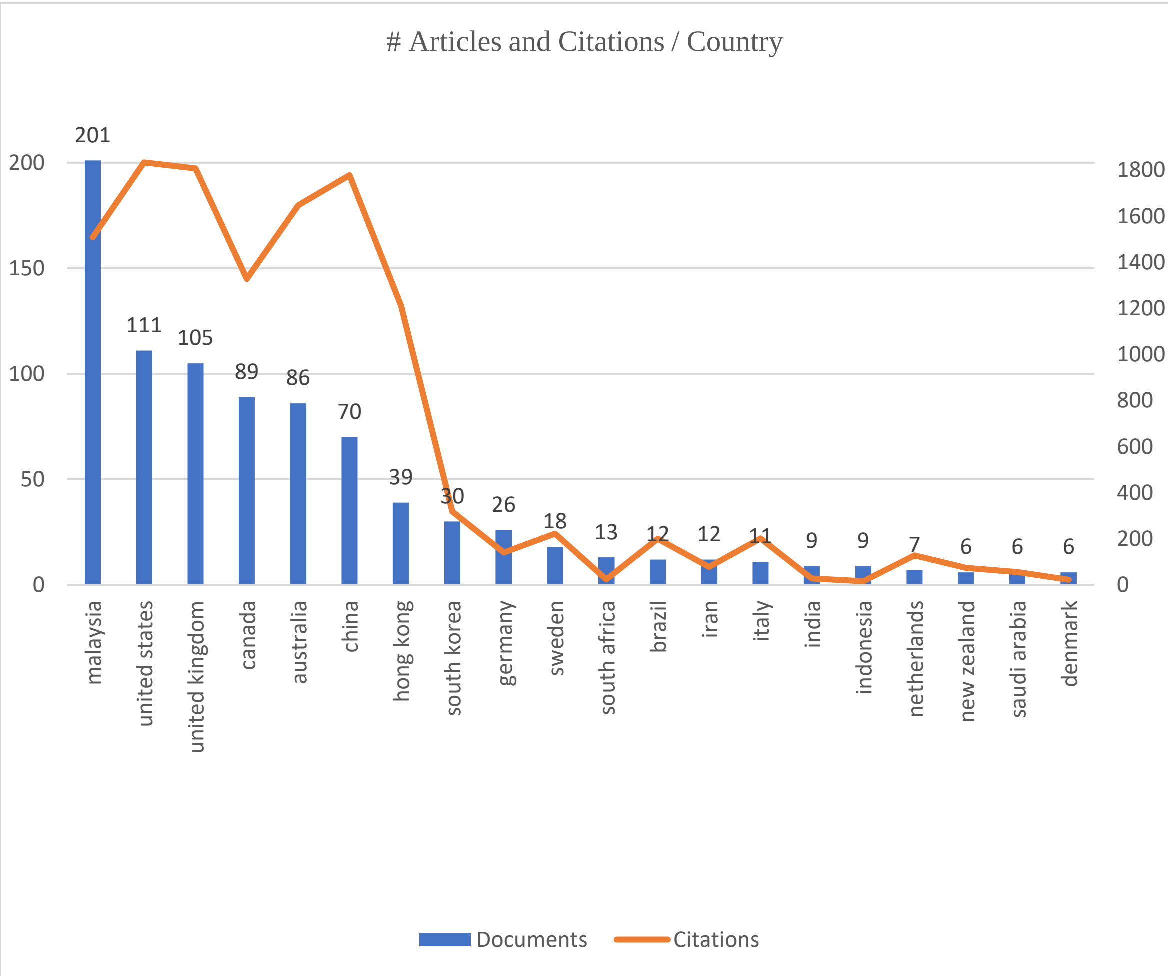

Figure 4.5. Number of publications and citations/country.

There are essential points to be clarified about the multivariant analysis results. First of all, the determinant was found to be 0.066, which is greater than 0.00001. Therefore, I can conclude that there are correlated items, so I can proceed with the analysis knowing that the data is suitable for applying FA. The correlation coefficients also do not fall in the range of 0.8 to 1.0, which means I do not have perfect correlation (singularity) or even multi-collinearity. Hence none of the items should be removed from the analysis. The Kaiser-Meyer-Olkin, the measure of sampling adequacy test, yields a 0.566 greater than a popular cut-off of 0.5 and is close to 0.6, which is considered perfect. The data also passed Bartlett's Test of Sphericity with a significance level of 0.001 and had an approximate Chi-Square value of 378.291 and a degree of freedom of 300.

Table 4.1. Total variance explained

	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Component	Total	Variance	Cumulativ	Total	Variance	Cumulativ	Total	Variance	Cumulativ
01	2.605	10.418	10.418	2.605	10.418	10.418	2.144	8.576	08.576
02	1.766	7.063	17.481	1.766	07.063	17.481	1.748	6.990	15.566
03	1.590	6.360	23.841	1.590	06.360	23.841	1.591	6.365	21.931
04	1.507	6.028	29.869	1.507	06.028	29.869	1.512	6.046	27.978
05	1.479	5.915	35.784	1.479	05.915	35.784	1.492	5.967	33.945
06	1.319	5.275	41.059	1.319	05.275	41.059	1.468	5.871	39.816
07	1.210	4.841	45.900	1.210	04.841	45.900	1.335	5.338	45.154
08	1.143	4.571	50.471	1.143	04.571	50.471	1.329	5.316	50.471
09	1.069	4.274	54.745						
10	1.028	4.112	58.857						
11	0.988	3.954	62.811						
12	0.962	3.850	66.660						
13	0.900	3.598	70.258						
14	0.837	3.349	73.607						
15	0.802	3.208	76.815						
16	0.770	3.080	79.895						
17	0.705	2.819	82.715						
18	0.703	2.811	85.525						
19	0.628	2.511	88.036						
20	0.578	2.313	90.349						
21	0.555	2.220	92.569						
22	0.526	2.103	94.672						
23	0.485	1.940	96.612						
24	0.454	1.815	98.427						
25	0.393	1.573	100.000						

Extraction Method: Principal Component Analysis.

From the communalities table, I found that all of the extraction values exceed 0.3. Hence, I do not have problems with any of the individual questions. However, when the total variance was analyzed, choosing 8 constructs could explain only 50.47% of the variance. Even though this situation is not ideal, it does not threaten the integrity of the research results. Table 4.1 illustrates the eigenvalues and other crucial results. The scree plot did not help choose the number of components because there was a precise elbow shape to set the cut-off number. Therefore, I could finalize the new categorization system by looking at the rotated component matrix presented in Table 4.2 and group risk factors.

Table 4.2. Rotated component matrix

Risk Factors	Components							
	01	02	03	04	05	06	07	08
R18	0.598							
R20	0.574							
R10	0.552							
R25	0.526							
R05	0.442							
R24	0.386							
R14		0.737						
R13		0.620						
R09		0.418						
R03		0.333						
R15			0.732					
R02			-0.581					
R07			-0.516					
R17				0.634				
R08				0.570				
R04				-0.533				
R23					0.709			
R22					0.639			
R11						-0.639		
R01						0.621		
R21						0.478		
R16							-0.736	
R19	0.346						0.576	
R12								0.818
R06								0.441

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 10 iterations.

From the univariate statistical analysis, I found that the most critical risk factor is scheduling and planning. Economies of scale and higher initial expenses came second and third in the list with close results of significant risk index. The survey participants were then concerned about Aesthetics & Tolerance and Staff's Experience, respectively. Standard deviation results of risk factors were also close to the results from [54]. However, there was not any detected correlation between the RSI and its associated standard deviation. Table 4.3 summarizes the findings. The table also summarized the categorization of the risk factors and the risk significance index for each risk category. I came up with an appropriate name for each category as the FA only tells us if the items load together but fail to answer the question of what construct is being measured.

Table 4.3. Descriptive statistics.

Rank	ID	Category	Risk Factor	SD	RSI	Mean RSI
02	R22	Project Initial Phases Rising Importance	Economies of Scale	1.38	3.38	3.30
08	R23		Early Decisions	1.07	3.21	
05	R19	Project Participants' Skills and Experience	Staff's Experience	1.10	3.26	3.25
06	R16		Suppliers' Skills	1.08	3.23	
03	R07	Process Initiation Difficulties	Higher Initial Expenses	1.07	3.35	3.10
09	R02		Site Space & Layout	1.48	3.03	
12	R15		Suppliers' Capacity	1.34	2.92	
01	R14	Time-Cost Conflicting Considerations	Scheduling and Planning	0.93	3.52	2.93
07	R13		Lead Time	1.47	3.23	
10	R09		Direct Costs	1.02	3.01	
23	R03		Climatic Conditions	1.52	1.95	
13	R18	Industry knowledge and Standards	Poor Understanding	1.53	2.83	2.65
15	R20		Authorities Understanding	1.28	2.78	
16	R10		Quality Monitoring Process	1.38	2.56	
18	R25		Consultation Service Co.	1.43	2.45	
21	R05		Legal Framework	1.34	2.37	
20	R24		Machinery & Technologies	1.43	2.37	
04	R21	Meeting Desired Standards and Expectations	Aesthetics & Tolerance	0.93	3.29	2.64
11	R11		Prefabrication Quality	1.44	2.96	
25	R01		Resources Handling	1.03	1.66	
14	R12	Risk-return Worthwhile trade-off	Safety Hazards	1.31	2.83	2.61
19	R06		Preferential Policies	1.40	2.39	
17	R08	Local-global Impact of the Current Economic and Political Systems	Currency Related Issues	1.20	2.55	2.22
22	R17		Supply Chain	1.41	2.21	
24	R04		Policy Modifications	1.17	1.89	

The top seven most risky factors are Safety Hazards, Direct Costs, Poor Understanding, Quality Monitoring, Scheduling and Planning, Site Layout, and Machinery and Technology with an overall frequency of occurrence of 0.736, 0.733, 0.730, 0.725, 0.708, 0.705, and 0.702, respectively. These numbers can give us an overall idea about the riskiness of each factor. The second metric was used to measure the data dispersion is the standard deviation. This metric is essential to measures the accumulated variation (uncertainty) in the data regardless of if the change was positive or negative.

Table 4.4 summarizes the reported results. The top seven most changing factors are Initial Expenses, Safety Hazards, Direct Costs, Poor Understanding, Site Layout, Preferential Policies, and Economies of Scale with corresponding values of 0.006, 0.005, 0.005, 0.005, 0.005, 0.005, and 0.005, respectively. The average change of probabilities is also calculated, and the results were recorded. The top seven highest average change of factors are Safety Hazards, Supply Chain, Legal Framework, Site Layout, Direct Costs, Authorities Understanding, and Preferential Policies with average change values of 0.020, 0.018, 0.016, 0.011, 0.011, 0.011, and 0.010, respectively.

Lastly, the influence is calculated and reported in the table. The table is sorted by the average influence metric from most to least risky factors. The top seven highest average influence of factors are Site Layout, Preferential Policies, Initial Expenses, Poor Understanding, Direct Costs, Authorities Understanding, and Lead Time with average influence values of 0.008, 0.008, 0.008, 0.008, 0.007, 0.007 and 0.007, respectively.

When the Kendall Rank Correlation Coefficient was study for the given rank of risk factors, the results were reported. First of all, there are 25 factors for each analysis technique and there are not any missing data. The reported tau coefficient was 0.08 and a p-value of 0.5948. The test statistics is performed against an alternative hypothesis that true tau is not equal to 0. The two-sided test fails with significance level of 0.95. The meaning of the data will be discussed in the discussion section below.

Table 4.4. The most significant reported results from the research metrics.

No.	Risk	Risk Factor	L3 occurrence	stdev	change idx	influence idx
01	R21	Safety Hazards	0.736	0.005	0.020	0.006
02	R07	Direct Costs	0.733	0.005	0.011	0.007
03	R16	Poor Understanding	0.730	0.005	-0.004	0.008
04	R19	Quality Monitoring	0.725	0.004	0.009	0.005
05	R22	Scheduling and Planning	0.708	0.004	0.010	0.005
06	R23	Site Layout	0.705	0.005	0.011	0.008
07	R14	Machinery and Technology	0.702	0.004	0.003	0.003
08	R18	Preferential Policies	0.663	0.005	0.010	0.008
09	R12	Lead Time	0.660	0.004	0.008	0.007
10	R11	Initial Expenses	0.644	0.006	0.006	0.008
11	R09	Economies of Scale	0.630	0.005	0.002	0.006
12	R13	Legal Framework	0.626	0.005	0.016	0.006
13	R20	Resources Handling	0.625	0.004	-0.006	0.007
14	R02	Authorities Understanding	0.623	0.003	0.011	0.007
15	R08	Early Decisions	0.611	0.004	-0.007	0.005
16	R10	Experience of the Staff	0.602	0.004	0.005	0.007
17	R15	Policy Modifications	0.581	0.005	0.007	0.007
18	R06	Currency Related Issues	0.568	0.003	0.004	0.004
19	R24	Skills of Suppliers	0.558	0.004	0.000	0.006
20	R25	Supply Chain	0.517	0.004	0.018	0.006
21	R05	Consultation Services	0.443	0.003	0.005	0.005
22	R17	Prefabrication Quality	0.421	0.003	0.005	0.003
23	R03	Capacity of Suppliers	0.321	0.003	-0.008	-0.004
24	R04	Climatic Conditions	0.281	0.002	-0.005	-0.004
25	R01	Aesthetics and Tolerance	0.247	0.003	-0.013	-0.007

5. SCHOLARLY DISCUSSION OF STUDY FINDINGS

The construction industry is infamous for being overwhelmed with resource planning, logistic challenges, and risk management which habitually result in cost overruns, project delivery delays, design defects, and contractual disputes [20]. Although several attempts were made to revolutionize the industry, most of these attempts never came to fruition. A prominent solution to the industry's everlasting issue is in the implementation of the innovative approach of MMC. Yet, the literature lacks a framework that uses a constructive methodology to wrestle the inherent obstacles. As I aim to encourage the industry key players to implement available offsite MMC techniques, I conducted a study that adopts a triangulated approach. I started with a comprehensive literature review and then moved to conduct a questionnaire survey to analyze risk factors inherent in MMC finishing off with qualitative and quantitative research analysis techniques.

Although scholars have made several attempts to comprehend the reasons for the industry's reluctance in the adoption of MMC, the body of knowledge lacks a comprehensive approach for cost overruns risk assessment of implementing MMC in developing countries. Besides, the body of knowledge lacks a comprehensive extensive application for addressing conditional dependence using available deep learning techniques. The study, therefore, aims to encourage the implementation of MMC. In achieving this, the objectives of this study are to develop the deep learning ANN model, analyze the convergence behavior of variables, initiate a system of relevant metrics to compare the ANN variables and identify the most influencing factors for each variable in the model. The risk analysis framework of potential risks provides an important tool for establishing an early warning system.

From the literature review part of the research, I came to realize that although the review included publications that have been published with a time elapsed of 16 years (2002-2018), the majority of risk factors lists from different publications shared similar characteristics. That is to say, the industry adoption and learning from the previous undertaking is not fully exploited. Hence, the initial list of 297 risk factors coming out of the 15 publications was reduced to a list of 25 risk factors. The MECE principle (Mutually Exclusive, Collectively Exhaustive) was implemented to ensure that there are no overlaps between the 25 risk factors and every factor is independent of each other. Also, all the list factors must be collectively exhaustive and express the initial list of 297 factors in its entirety. The MECE principle was implemented during each round of reducing the list. Data collection is the primary concern of most researchers, and I was no exception. However, implementing the virtual snowball sampling technique suggested by [53], smoothened the process. A poster was designed that clearly states who we are, what we aim at, and who should participate in the study. The conditions of participating were added to not waste the participants' time and filter out invalid responses. The poster was distributed in a pdf file format and the link to the questionnaire surveys was one click away.

For the univariate analysis, two comments can be made. First of all, the standard deviations of the risk factors were slightly higher than previously reported in other studies [7], [62]. This is acceptable and predictable if I consider that the study includes several developing countries where only one developing country (China) was analyzed in those two publications. The second comment is concerned with the RSI. Out of the 25 risk factors, the value 2.5 of RSI (the mid-point of the RSI scale of 5) was located between the 32^{ed} and 36th percentiles. This poses a question, after conducting a comprehensive literature review and identifying 297 risk factors, did I fail to come up with real risk factors considering that one-third of the identified risk factors scored less than 50% on the RSI scale? Or can I conclude that the industry professional did not take the riskiness of the MMC activities seriously?

Before I can jump to conclusions, I need to understand the RSI scale. In doing so, I need to consult principles from the measurement theory. According to the measurement theory, the scale is an interval scale. This means it fails to satisfy the ration scale conditions. For elaboration, let's take an example where a study participant chooses to select a moderate level of impact with a possible degree of occurrence. According to the RSI, the value of 3 from the level of impact and the value of 0.6 of the level of occurrence yields an RSI value of 1.80. Now, this value is the mid-point of the scale, not the assumed 2.50. If we look again at the table, we find that although the second to last factor is very close to 1.80, only one factor is less than the overall moderate level of RSI. These findings not only refute the previously proposed claims about the selection of risk factors and the seriousness of questionnaire participants but also signify the importance of the list of risk factors because almost all of them have RSI higher than the average. Besides, the values were significantly higher than reported in the literature in similar studies where RSI was adopted [54].

The research findings support previous studies in the literature and fill a research gap in the body of knowledge. Risk attributed to the industry of construction when MMC is used have been identified, categorized, and assessed. By conducting a compare and contrast analysis of the descriptive statistics the study revealed significant results. From the table, I found that compressed project schedule, timely design freeze, advanced project planning, and scheduling has comparable results from the literature. [64] emphasized the importance of prefabrication adoption on constructability that can be reflected by construction time. In their study, they suggested that to optimize the adoption of prefabrication slabs should be semi-prefabricated with in situ toppings, and all of the façade, staircases, balconies, beams, columns, and internal walls should be prefabricated. [65], in their list, prioritized construction time over every other factor with a mean score of 4.64 and Standard Deviation (SD) of 0.665. It was also proven that two of the top five CSFs are timely freeze of scoping and design and due recognition of possible early completion from modularization [66].

Secondly, difficulty to achieve repetition of consistent layout and economies of scale. It matches previous results in the literature. [54] regarded Inappropriate design codes and standards for industrialized buildings as the second most significant factor with a mean of 2.19 and SD of 0.91. [67] not only supported the findings and stated that standardization is considered as significant" or "very significant, but also added that the factor is more significant to developing countries. An average of 4.33 and SD of 0.769 were obtained. However, with established and sophisticated standardization systems in developed countries, this might not be a significant factor. The next factor in the list is higher initial capital cost, higher investment in fixed assets, and speed of return on investment. With a mean of 2.20 and SD of 0.98 in one study, enormous difficulty in achieving a return on high initial investment was identified amongst the top five risk factors [54]. They added, to return the investment market demand for industrialized buildings must not be high as off-site factories require investment in fixed assets of the factory, fabricated molds, and Research and Development (R&D) activities. [19] perceived higher capital cost as the most important factor with 68% of participants' selection.

Our fourth significant factor is complex interfacing between systems, tolerance issues, aesthetics, and monotony of structure. Similarly, in the list of [19], complex interfacing between systems came third. [35] showed that poor aesthetics is a concern of non-contractors; whereas contractors are less concerned about aesthetics than non-contractors. Finally, the lack of understanding of OSM by local authorities with inappropriate design codes and standards came fifth on the list. While the findings indicate that this factor is significant, [68] states that according to the industry surveys, lack of in-house expertise is ranked 10^{th}. Likewise, case study project-oriented surveys ranked this factor in 9^{th} place. Comparably, the lack of experienced contractors was also ranked 9^{th} [62]. Besides, the fragmented nature of the industry is reflected by the lack of prefabrication suppliers as well as experienced collaboration groups.

The three least significant factors reported in the study also produced comparable results with existing reported studies. The least significant factor is Environmental sustainability, resources consumption, waste generation, and waste disposal. Looking at the [67] list, we find three factors related to the factor. If we take the average of the means of each factor, we end up with a mean of 4.34 which is placed after the first 10 significant factors. There is not a great prominence placed on this factor from developed countries, especially from the U.K. [65]. The second to last least significant factor is Unexpected Statutory modifications to existing policies. [19] did not report legal issues to be of any importance. In another study, it was only reported that restrictive regulations are a limiting factor by 32% of the response rate [69]. Change in the governmental regulations was also considered as the least frequent risk from the Malaysian contractors in the industry [70]. In the same study, it was indicated that acts of God the least frequent risk. Likewise, in the study, climatic conditions did not show much significance.

The literature addressed several limitations to deep learning applications. The black box challenge, ethics, cybersecurity, and cost are all limitations that can be expected by industry practitioners and researchers when adopting some of these techniques [20]. I generally did not encounter any significant challenges from this list in the study. However, I did encounter problems in recording and the heterogeneity of incident data. The quality of records also affects the analysis of incidents. [13]. After careful analysis of the data, I tried to reduce the conflicts while analyzing the relationship between attributes. In order to increases the accuracy of risk assessment levels, I tried to collect a large amount of data [15]. I did not implement any special techniques to deal with the heterogeneity problem of the data such as the latent class clustering analysis proposed by [13].

Choosing between different values of epoch number and the batch sizes and trying to understand the relationship between the two variables is not simple at first glance. However, as proposed in the methodology section, the 3D graphical representation of the cartesian coordinate system aided in the understanding. From the graph, it is clear that epoch number and batch size number are negatively correlated given that the score of the evaluation function is held constant. The second observation from the graph is that there are around four local maximum values. In the study, I did not find any particular value that I could consider to be the global maximum value. However, since the objective is to reduce the epoch number in order to avoid losing the ability to generalize the network, I tried to keep the epoch number as small as possible. Hence, I choose the epoch number and the batch size to be 300 and 20, respectively. Figure 5.1. illustrates the score values for different epoch numbers and batch sizes. The circles represent the four local minimum values.

The final issue I want to discuss is related to the network structure when columns in a row do not have a local structure. To capture existing correlations between columns, I used fully connected networks in generator and critic networks. The type of connection and the introduction of the hidden layer(s) makes it possible for the network to exhibit non-linear behavior. I also did not expect the network to learn the training set to perfection. At this stage of research, I am content with the network performance. In the methodology, I did not try to increase the number of nodes in order to secure the ability of the network to generalize the learning I kept it as low as possible. Hence, the network will not be treated as a memory bank that can recall the training set to perfection but does not perform well on testing samples that were excluded from the training set and the generalization will be lousy. It took the platform and the python interpreter around 8 hours to perform the learning, sampling, testing, and validating of the results. It was the longest time required to perform the computations in the methodology and also the longer than what I have experienced in the previous research about Bayesian Network and Bayesian inferences. Because ANN requires a relatively high computational power, there is no wonder why ANN has not been implemented not until recently.

The list of identified seven risk factors shares multiple similarities and dissimilarities in common. To explore the shared characteristics, [62] study results will be utilized. This study was chosen for two reasons. The first of which is that the study analyzes a developing country (China) and developing countries are known to share many similarities. The second reason is that the study is relatively new and can still represent the construction industry which gets developing slowly. The second study that will be considered is the [71]. The main advantage this study presents is the comparison between developing and developed countries (in this context we have the UK and China).

By comparing the first factor in our list, we can conclude that the factor is not even listed in the two considered articles. Safety hazards (as a risk factor) are not considered a risk factor. It is self-evident that for the survey participants to assess the risks they are given a list of risk factors. If the researchers choose to not include a certain risk factor on a list, then it is understandable that the factor will not be included. Given that the survey participants could not (even if they wanted to) state that safety hazards are risky, then why the research did not include it? One possible answer is that they simply thought that by using MMC we can achieve a controlled environment. This is obvious, however, what about the unprecedented new safety-related risks that the industry participants might face? To this discussion, there will not be a definite answer. Future research initiatives might explain the reasons.

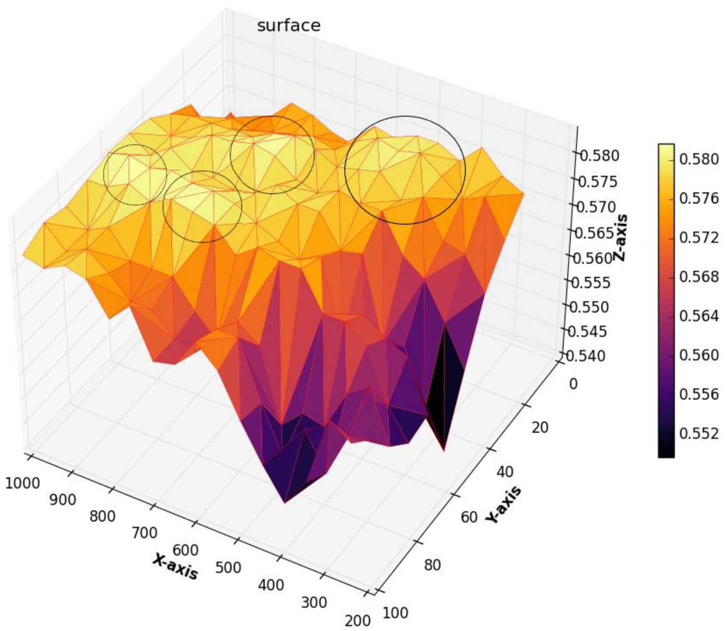

Figure 5.1. The four maximum score values for different epoch numbers and batch sizes.

Moving on to the second risk factor, the direct costs. In [62] direct costs were not explicitly stated in the list. Still, lack of incentives (which might be in the form of cost reduction), higher initial costs, and high-cost pressure came first, second, and fifth on the list. Reflecting on the [71] paper, this factor was ranked first in the UK and seventh in China. From the results, we can conclude that the Chinese construction industry is using MMC without compromising on cost. Still, it is fascinating how developing countries can beat developed countries in taking advantage of MMC even when we know that labor expenses are higher in developed countries.

Thirdly, we have a poor understanding of MMC. In one article it is ranked 8th [62]. Another article ranked the factor 6th in China. The fourth factor is the quality monitoring risk. [62] ranked it fourth in their list. In contrast, [71] ranked it 16th on the list. The fifth factor is scheduling and planning. [62] ranked the factor seventh. In another article, it was ranked fourth in China [71].

The second to last risk factor is site layout. [71] ranked it fifth in the UK and 27 in China. The hypothesis is that developing countries have more space because they construct their project in uninhabited areas while developed countries have more confined construction areas. Lastly, the machinery and technology. [62] research results support the study list. In their results, they ranked the lack of local R&D institutes and services 8th in riskiness.

Although the univariate analysis was conducted to get the overall riskiness of risk categories, a comparison on the factors levels was initiated. When it comes to comparing between the univariate and ANN analysis techniques, a significant aspect is the rank of each risk factor. The riskiest factor Safety Hazards was ranked 14th using the proposed risk significance index. The second riskiest factor, Direct Costs, was ranked 10th using the RIS. Thirdly, Poor Understanding was ranked 13th. Quality Monitoring Process is ranked 14th. Scheduling and Planning was ranked 1st. second to last is the Site Space & Layout. It was ranked 9th. Lastly, Machinery & Technologies Was ranked 21st.

In addition to the one-on-one comparison, the study implemented the Kendall Rank Correlation Coefficient. The reported 0.08 tau coefficient indicates that the ranks of the two analysis methods are in fact not associated. It is closer to zero than 1 (representing similarity) and -1 (representing dissimilarity). Besides, given that the two-sided test statistics failed with a significance level of 0.95 means the true tau is equal to zero.

From the one-on-one comparison and the tau coefficient it is clear that the use of multiple data analytics techniques present different perspective on the risk factors. If we rewind and remembered what each analysis represents, it is self-evident that the ANN focuses more on risk factors that have the tendency to have higher degree of impact and frequency of occurrence. Hence the utilization of ANN in this study brings additional findings that cannot be reached via traditional data analytics techniques.

6. CONCLUSIONS AND FURTHER RECOMMENDATIONS

The construction industry constitutes an important role in economic development. The body of knowledge is rich with studies that address topics relevant to construction management, MMC, risk management, data science tools, etc. The development of MMC is a self-driven process pushed by macro development. Besides, pilot programs are the most effective method to promote MMC, gain recognition by society, and accumulate experience. Hence, MMC is becoming the center of scholars' attention and implemented in several strategies to innovate in the industry. Nevertheless, the introduction of the deep learning available tools and techniques never came to fruition. Especially the ML tools that particularly consider discovering and capturing conditional dependency relationships existing in the data.

Accordingly, I implemented the CTGAN deep learning tool that considers the conditional dependence in the risk assessment data for implementing MMC. An ANN has been developed, validated, and applied in the risk assessment framework. Besides, after conducting a review of the literature, the number of yearly publications was linearly correlated with the progress of time. The downside of the situation is that the industry did not reach the exponential phase of the learning curve. The upside of the situation is that MMC has a great future potential for the industry and scholars alike.

Lately, governments are now more aware of the significant impact of industrialization in the industry and how this will help achieve more sustainable buildings. This approach, however, is facing multifaceted risks. This paper, therefore, identified 15 risk-related research publications from the literature. First, I considered 297 risk factors that have been addressed in the literature out of 15 articles. Then, I shortened the list to a 25 risk factors list and distributed questionnaire surveys. 149 valid responses were collected and analyzed. Linguistic terms helped assign the degree of impact and frequency of occurrence for each risk factor. A qualitative approach was implemented afterward using the risk significance index to quantify and rank the overall risk level of each risk factor. The risk factors were also grouped into eight categories using the FA technique to identify factors measuring the same construct. I named these constructs according to what they are trying to measure.

Afterward, I applied the second round of univariate statistical analysis to each category to get the overall risk significance index within groups and compare it between groups.

Our study also revealed several significant results. First of all, it was found that the epoch number and batch size number are negatively correlated when the evaluation of fitness score is held constant. Besides, a batch size of 10 with an epoch number of 300 optimizes the training process of the ANN model. After calculating the probability of L3 risk level for each risk factor, I found that the top five most risky factors are Safety Hazards, Direct Costs, Poor Understanding, Quality Monitoring, and Scheduling and Planning with an overall frequency of occurrence of 0.736, 0.733, 0.730, 0.725, and 0.708, respectively. I also found that the top five highest average influence of factors are Site Layout, Preferential Policies, Poor Understanding, Initial Expenses, and Lead Time with average influence values of 0.008, 0.008, 0.008, 0.008, and 0.007, respectively. Inspecting the relationship and probability change table, I can observe the significant effect of the Initial Expenses risk factor on individual factors by counting the number of appearances in the table.

The results can be used in the construction industry by stakeholders to identify the most significant risk factors and apply different risk management techniques. Accordingly, the total project cost can be reduced as the top management will distribute the money on more critical risks and only control the least essential risk factors. In addition, the grouping of risk factors will add more assistance to predict how different risk factors can act together and cause a significant risk in each of the eight categories. This research does not aim to create an artificial market for the implementation of MMC. However, the successful implementation will not be possible without the industries’ key players' involvement and awareness of available innovation techniques. The government can efficiently encourage implementing the method without creating an artificial market by applying preferential policies to the industry innovators.

Another pitfall of this study is that it failed to address the relationship between different risk factors and how they can interact. As [41] stated, effective risk management programs must be dynamic and ongoing by nature, not static. Identifying the correlation or causation relationships are beyond the scope of this research. In a cost estimation context, this can be interpreted as follows, if a correlation exists between two variables, it does not always necessarily mean causation exists as well. I also do not want to undermine the importance of descriptive statistics. It is considered as a pillar on which the methodology is based. Nonetheless, to my knowledge, this topic of research has not been addressed in the literature previously. In future research activities, I will build upon these results and implement conditional artificial neural networks to capture hidden conditional dependence between different risk factors.

The uniqueness of this study lies in it being the first to adopt the latest available AI techniques that accounts for hidden conditional relationships in the data. The research has significant contributions to scholars and industry practitioners. The ANN structure can be considered as a model for scholars to adopt. The study approach enables industry practitioners to generate case-based reasoning and provides an in-depth understanding of research questions after modifying the model according to their work conditions. For the world of practice, the research provides a readily available point of reference for R&D entities and for industry professionals, policymakers, and decision-makers. This study would inspire future efforts and provide directions into how to best implement deep learning to numerous intriguing similar industry challenges.

Although it sounds a little ambitious, future research ideas will be directed towards developing a holistic risk assessment framework that compares between several available frameworks. Since the model only identifies and measured the risks with the proposed metrics, future studies might propose necessary preventative measures according to the risk results predicted in the model to avoid racking up losses on construction projects. Finally, practitioners' intuition should always be questioning the existence of conditional dependence between the risk factors.

7. ACKNOWLEDGEMENTS AND DECLARATION OF INTERESTS

The authors wish to give their sincere gratitude to the industry professionals who have participated in face-to-face interviews to validate the reduced 25 risk factor list. The authors also express their gratitude to the 149 industry experts who participated in the questionnaire survey and to those who distributed them. Finally, the authors declare no unreported conflict of interest. Besides,

☐ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.
☒ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

ALI TATARI reports financial support and travel were provided by SELINA INSAAT VE GENEL TICARET LIMITED SIRKETI.

REFERENCES

1. Goldman L., "The Origins of British 'Social Science' : Political Economy, Natural Science and Statistics, 1830–1835," *The Historical Journal*, vol. 26, no. 3, pp. 587-616, 1983.

2. Clarke M. R. B., R. O. Duda, and P. E. Hart, "Pattern Classification and Scene Analysis," *Journal of the Royal Statistical Society*, vol. 137, no. 3, pp. 442-443, 1974.

3. Wasserman L., "All of Statistics : A Concise Course in Statistical Inference Brief Contents," Springer, New York, NY, 2004.

4. Witten H., E. Frank, M. A. Hall, and C. J. Pal, "Data Mining: Practical Machine Learning Tools and Techniques," Morgan Kaufmann, San Francisco, CA, USA, 2011.

5. Darko, A. P. C. Chan, M. A. Adabre, D. J. Edwards, M. R. Hosseini, and E. E. Ameyaw, "Artificial Intelligence in the AEC Industry: Scientometric Analysis and Visualization of Research Activities," *Automation in Construction*, vol. 112, no. 1, p. 1, 2020.

6. Nadim W. and J. S. Goulding, "Offsite Production in the UK: The Way Forward? A UK Construction Industry Perspective," *Construction Innovation*, vol. 10 no. 2, pp. 181-202, 2010.

7. Gibb G., "Off-Site Fabrication: Prefabrication, Pre-Assembly and Modularisation," John Wiley Sons, Scotland, UK, 1999.

8. Warszawski A., "Industrialized and Automated Building Systems: A Managerial Approach," CRC Press, Boca Raton, Florida, United States, 2003.

9. O'Brien M., R. *Wakefield, and Y. Beliveau, "*Industrializing the Residential Construction Site,*" HUD's Office of Policy Development and Research*, vol. 23, no. 1, pp. 12-24, 2000.

10. Goodier I. and A. G. F. Gibb, "Barriers and Opportunities for Off-Site in the UK Barriers and Opportunities for Offsite in the UK," *Abdul Samed Kazi (ed). Systematic Innovation in the Management of Project and Processes, cib Helsinki International Joint Symposium*, vol. 32, no. 1, pp. 148-158, 2005.

11. Azman M. N. A., M. S. S. Ahamad, and W. M. . Wan Husin, "Comparative Study on Prefabrication Construction Process," *International Surveying Research Journal*, vol. 2, no. 1, pp. 45-58, 2012.

12. Ministry of Housing, Communities & Local Government, 2019, "*Modern Methods of Construction Working Group: Developing a Definition Framework*," https://www.gov.uk/government/publications/modern-methods-of-construction-working-group-developing-a-definition-framework, May 2022.

13. Jick T. D., "Mixing Qualitative and Quantitative Methods: Triangulation in Action," *Administrative Science Quarterly*, vol. 24, no. 4, p. 602, 1979.

14. Cao X., Z. Li, and S. Liu, "Study on Factors That Inhibit the Promotion of Si Housing System in China," *Energy Building*, vol. 88, no. 1, pp. 384-394, 2015.

15. Goodman L. A., "Snowball Sampling," *The Annals of Mathematical Statistics*, vol. 32, no. 1, pp. 148- 170, 1961.

16. Baltar F. and I. Brunet, "Social Research 2.0: Virtual Snowball Sampling Method Using Facebook," *Internet Research*, vol. 22, no. 1, pp. 57-74, 2012.

17. Luo L. Z., C. Mao, L. Y. Shen, and Z. D. Li, “Risk Factors Affecting Practitioners’ Attitudes Toward the Implementation of An Industrialized Building System a Case Study from China,” *Engineering, Construction and Architectural Management*, vol. 22, no. 6, pp. 622-643, 2015.

18. Arrindell W. A. and J. Van Der Ende, “An Empirical Test of the Utility of the Observations-To-Variables Ratio in Factor and Components Analysis,” *Applied Psychological Measurement*, vol. 9, no. 34, pp. 165-178, 1985.

19. Guadagnoli E. and W. F. Velicer, “Relation of Sample Size to the Stability of Component Patterns,” *Psychological Bulletin*, vol. 103, no. 2, pp. 265-275, 1988.

20. Velicer W. F. and D. N. Jackson, “Component Analysis Versus Common Factor Analysis: Some Further Observations,” *Multivariate Behavioral Research*, vol. 25, no. 1, pp. 97-114, 1990.

21. Steiger J. H., “Some Additional Thoughts on Components, Factors, and Factor Indeterminacy,” *Multivariate Behavioral Research*, vol. 25, no. 1, pp. 41-45, 1990.

22. Costello B. and J. W. Osborne, “Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting The Most from Your Analysis,” *Practical Assessment, Research and Evaluation*, vol. 10, no. 7, pp. 7-10 2005.

23. Ayhan U. and O. B. Tokdemir, “Accident Analysis for Construction Safety Using Latent Class Clustering and Artificial Neural Networks,” *Journal of Construction Engineering and Management*, vol. 146, no.3, pp. 146-149, 2020.

24. Xu L., M. Skoularidou, A. Cuesta-Infante, and K. Veeramachaneni, “Modeling Tabular Data Using Conditional GAN,” *Advances in Neural Information Processing Systems*, vol. 2, no. 1, pp. 462-467, 2019.

25. Kendall M. G., “A New Measure of Rank Correlation,” *Biometrika*, vol. 30, no. 1, pp. 81-83, 1938.

26. Gibb G. F. and F. Isack, “Re-engineering Through Pre-Assembly: Client Expectations and Drivers,” *Building Research and Information*, vol. 32, no. 2, pp.146-160, 2003.

27. Shen Mao, Q., W. Pan, and K. Ye, “Major Barriers to Off-Site Construction: The Developer’s Perspective in China,” *Journal of Management in Engineering*, vol. 31, no. 3, pp. 31-34, 2015.

28. Hong J., G. Q. Shen, Z. Li, B. Zhang, and W. Zhang, “Barriers to Promoting Prefabricated Construction in China: A Cost–Benefit Analysis,” *Journal of Cleaner Production*, vol. 172, no. 1, pp. 649-660, 2018.

29. Akinosho T. D. et al., “Deep Learning in the Construction Industry: A Review of Present Status and Future Innovations,” *Journal of Building Engineering*, vol. 32, no. 1, p. 32, 2020.

30. Mao, L. Shen, L. Luo, and Z. Li, “Identification of Risk Factors Influencing the Implementation of Industrialized Building System in China,” *Proceedings of the 19th International Symposium on Advancement of Construction Management and Real Estate*, vol. 1, no. 1, pp. 219-230, 2015.

31. Chen Y., G. E. Okudan, and D. R. Riley, “Decision Support for Construction Method Selection in Concrete Buildings: Prefabrication Adoption and Optimization,” *Automation in Construction*, vol 19, no. 6, pp. 665-675, 2010.

32. Blismas N. and R. Wakefield, “Drivers, Constraints and the Future of Offsite Manufacture in Australia,” *Construction Innovation*, vol. 9, no. 1, pp. 72-83, 2009.

33. O'Connor J. T., W. J. O'Brien, and J. O. Choi, "Critical Success Factors and Enablers for Optimum and Maximum Industrial Modularization," *Journal of Construction Engineering and Management*, vol. 140, no. 6, pp. 140-146, 2014.

34. Yunus R. and J. Yang, "Critical Sustainability Factors in Industrialised Building Systems," *Construction Innovation*, vol. 12, no. 4, pp. 447-463, 2012.

35. Pan W., A. F. Gibb, and A. R. J. Dainty, "Perspective of UK Housebuilders on the Use of Offsite Modern Methods of Construction," *Construction Management and Economics*, vol. 25, no. 2, pp. 183-194, 2007.

36. Jaillon L. and C. S. Poon, "Sustainable Construction Aspects of Using Prefabrication In Dense Urban Environment: A Hong Kong Case Study," *Construction Management and Economics*, vol. 26, no. 9, pp. 953-966, 2008.

37. Goodier and A. Gibb, "Future Opportunities for Offsite in the UK," *Construction Management and Economics*, vol. 25, no. 6, pp. 585-595, 2007.

38. Hassim S., M. S. Jaafar, and S. A. A. H. Sazalli, "The Contractor Perception Towers Industrialised Building System Risk in Construction Projects in Malaysia," *American Journal of Applied Sciences*, vol. 6, no. 5, pp. 973-942, 2009.

39. Lin S. S., S. L. Shen, A. Zhou, and Y. S. Xu, "Risk Assessment and Management of Excavation System Based on Fuzzy Set Theory and Machine Learning Methods," *Automation in Construction*, vol. 122, no. 1, pp. 273-297, 2021.

40. Rahman M. M., "Barriers of Implementing Modern Methods of Construction," *Journal of Management in Engineering*, vol. 30, no. 1, pp. 69-77, 2014.

41. Li C. Z., J. Hong, F. Xue, G. Q. Shen, X. Xu, and M. K. Mok, "Schedule Risks in Prefabrication Housing Production in Hong Kong: a Social Network Analysis," *Journal of Cleaner Production*, vol. 134, no. 1, pp. 482-492, 2016.

42. Arif M. and C. Egbu, "Making a Case for Offsite Construction in China," *Engineering, Construction and Architectural Management*, vol. 17, no. 6, pp. 536- 548, 2010.

43. Kamali M. and K. Hewage, "Development of Performance Criteria for Sustainability Evaluation of Modular Versus Conventional Construction Methods," *Journal of Cleaner Production*, vol. 142, no. 1, pp. 3592-3606, 2017.

44. Hwang B. G., M. Shan, and K. Y. Looi, "Key Constraints and Mitigation Strategies for Prefabricated Prefinished Volumetric Construction," *Journal of Cleaner Production*, vol. 183, no. 1, pp. 183-193, 2018.

45. Jiang R., C. Mao, L. Hou, C. Wu, and J. Tan, "A SWOT Analysis for Promoting Off-Site Construction Under the Backdrop Of China's New Urbanisation," *Journal of Cleaner Production*, vol. 173, no. 1, pp. 225-234, 2018.

46. Song J., W. R. Fagerlund, C. T. Haas, C. B. Tatum, and J. A. Vanegas, "Considering Prework on Industrial Projects," *Journal of Construction Engineering and Management*, vol. 131, no. 6, pp. 723-733, 2005.

47. Gan X., R. Chang, J. Zuo, T. Wen, and G. Zillante, "Barriers to the Transition Towards Off-Site Construction in China: An Interpretive Structural Modeling Approach," *Journal of Cleaner Production*, vol. 197, no. 1, pp. 8-18, 2018.

48. Badir Y. F., M. R. A. Kadir, and A. H. Hashim, "Industrialized Building Systems Construction in Malaysia," *Journal of Architectural Engineering*, vol. 8, no. 1, pp. 19-23, 2002.

49. Abanda H., J. H. M. Tah, and F. K. T. Cheung, "BIM in Off-Site Manufacturing for Buildings," *Journal of Building Engineering*, vol. 14, no. 1, pp. 89-102, 2017.

50. Hussein H. X. Li, M. Al-, Z. Lei, and Z. Ajweh, "Risk Identification and Assessment Of Modular Construction Utilizing Fuzzy Analytic Hierarchy Process (AHP) and Simulation," *Canadian Journal of Civil Engineering*, vol. 40, no. 12, pp. 1184-1195, 2013.

APPENDIX A

Table 7.1. Detailed information about selected publications

ID	Authors	Title	Source title	Publisher	Year	Cited
01	(A. G. F. Gibb & Isack, 2003)	Re-engineering through pre-assembly: Client expectations and drivers	Building Research and Information		2003	242
02	(Mao, Shen, Pan, et al., 2015)	Major barriers to off-site construction: The developer's perspective in China	Journal of Management in Engineering	ASCE	2015	201
03	(Hong et al., 2018)	Barriers to promoting prefabricated construction in China: A cost–benefit analysis	Journal of Cleaner Production	Elsevier Ltd	2018	178
04	(C. Z. Li et al., 2016)	Schedule risks in prefabrication housing production in Hong Kong: a social network analysis	Journal of Cleaner Production	Elsevier Ltd	2016	134
05	(Arif & Egbu, 2010)	Making a case for offsite construction in China	Engineering, Construction and Architectural Management		2010	112
06	(Rahman, 2014)	Barriers of implementing modern methods of construction	Journal of Management in Engineering	ASCE	2014	111
07	(Kamali & Hewage, 2017)	Development of performance criteria for sustainability evaluation of modular versus conventional construction methods	Journal of Cleaner Production	Elsevier Ltd	2017	109
10	(Hwang et al., 2018)	Key constraints and mitigation strategies for prefabricated prefinished volumetric construction	Journal of Cleaner Production	Elsevier Ltd	2018	104
22	(Jiang et al., 2018)	A SWOT analysis for promoting off-site construction under the backdrop of China's new urbanisation	Journal of Cleaner Production	Elsevier Ltd	2018	78
24	(Song et al., 2005)	Considering prework on industrial projects	Journal of Construction Engineering and Management		2005	78
25	(Gan et al., 2018)	Barriers to the transition towards off-site construction in China: An Interpretive structural modeling approach	Journal of Cleaner Production	Elsevier Ltd	2018	77
26	(Badir et al., 2002)	Industrialized building systems construction in Malaysia	Journal of Architectural Engineering		2002	76
27	(Abanda et al., 2017)	BIM in off-site manufacturing for buildings	Journal of Building Engineering	Elsevier Ltd	2017	75
30	(H. X. Li et al., 2013)	Risk identification and assessment of modular construction utilizing fuzzy analytic hierarchy process (AHP) and simulation	Canadian Journal of Civil Engineering		2013	71
31	(O'Connor et al., 2014)	Critical success factors and enablers for optimum and maximum industrial modularization	Journal of Construction Engineering and Management	ASCE	2014	69

APPENDIX B

Table 7.2. Sources of the selected 15 publications

No.	Source title	Publisher	Score	Highest percentile	17-20 Cit	17-20 Doc	% Cited	SNIP	SJR	Articles
01	Journal of Cleaner Production	Elsevier Strategy and Management	13.1	98.0%, 07/440	203300	15505	89	2.475	1.937	06
02	Building Research and Information	Taylor & Francis Civil and Structural Engineering	08.2	93.0%, 20/318	1922	0233	85	2.205	1.249	01
03	Journal of Management in Engineering	ASCE Industrial Relations	07.9	99.0%, 01/054	2935	0372	85	2.372	1.646	02
04	Journal of Construction Engineering & Management	ASCE Industrial Relations	06.4	91.0%, 05/054	4506	0706	81	1.875	0.967	02
05	Journal of Building Engineering	Elsevier Architecture	05.5	97.0%, 04/138	8150	1495	76	2.290	0.974	01
06	Engineering, Construction & Architectural Management	Emerald Architecture	04.0	93.0%, 09/138	1841	0466	73	1.210	0.585	01
07	Journal of Architectural Engineering	ASCE Visual Arts and Performing Arts	02.3	97.0%, 13/532	0471	0209	61	1.056	0.410	01
08	Canadian Journal of Civil Engineering	NRC Research Press General Environmental Science	02.0	50.0%, 110/220	0816	0417	62	0.767	0.323	01
		Max	13.1	99.0%, 01/054	203300	15505	89	2.475	1.937	
		Min	02.0	50.0%, 110/220	471	209	61	0.767	0.323	

APPENDIX C

Table 7.3. Risk sources from the selected 15 publications

NO.	Risk Category	Risk Factor	[61]	[62]	[63]	[72]	[73]	[71]	[74]	[75]	[76]	[77]	[78]	[79]	[80]	[81]	[66]
01	Site Conditions and Environment	Resources Handling									0			0	0		
02		Site Space & Layout			0	0	0	0	0	0					0		0
03		Climatic Conditions				0						0				0	
04	Legal concern	Policy Modifications		0		0		0				0	0				
05		Legal Framework		0		0	0	0		0	0	0					0
06		Preferential Policies		0				0									0
07	Cost and Economy	Higher Initial Expenses		0	0	0	0	0	0	0	0		0	0			
08		Currency Related Issues						0				0				0	
09		Direct Costs	0		0			0	0	0		0		0	0		0
10	Quality	Quality Monitoring Process		0		0		0									
11		Prefabrication Quality	0			0				0			0	0	0	0	
12	Safety	Safety Hazards	0			0					0	0			0		0
13	Time	Lead Time	0	0		0	0	0	0	0		0			0		0
14		Scheduling and Planning				0	0	0	0	0	0	0	0	0	0	0	0
15	Inbound/Outbound Logistics	Suppliers' Capacity	0	0		0	0	0		0		0	0				0
16		Suppliers' Skills		0	0	0	0	0	0			0	0				
17		Supply Chain		0	0	0	0	0	0			0					0
18	Knowledge and Skills	Poor Understanding	0	0			0	0					0		0		
19		Staff's Experience		0		0	0	0	0	0	0	0		0	0		0
20		Authorities Understanding					0	0	0	0	0		0				
21	Design	Aesthetics & Tolerance	0	0	0	0	0	0		0			0	0			
22		Economies of Scale		0		0	0	0				0	0				0
23		Early Decisions		0								0					
24	Equipment and Tech	Machinery & Technologies			0	0						0		0			
25		Consultation Service Co.		0		0	0	0		0	0	0					0

APPENDIX D

Table 7.4. Risk codes, factors, and abbreviation list

ID	Risk Factor	Abbreviation
R01	Environmental sustainability, resources consumption, waste generation, and waste disposal	Resources Handling
R02	Restricted or unsuitable site layout or space	Site Space & Layout
R03	Anticipated climatic conditions during construction	Climatic Conditions
R04	Unexpected Statutory modifications to existing policies	Policy Modifications
R05	Insufficient, outdated, or absence of a legal framework	Legal Framework
R06	Few opportunities for obtaining preferential policies on tax, loan, subsidy, etc.	Preferential Policies
R07	Higher initial capital cost, higher investment in fixed assets, and speed of return on investment	Higher Initial Expenses
R08	Currency issues (exchange, inflation, increased loan interest rates, etc.)	Currency Related Issues
R09	Material, labor, maintenance, and operation cost	Direct Costs
R10	Lack of a quality monitoring mechanism for the production process	Quality Monitoring Process
R11	Defective, damaged, or the durability of prefabricated elements are unproven	Prefabrication Quality
R12	Increased possibility of safety adverse events due to large units and heavy loads	Safety Hazards
R13	Increased lead times in design and construction	Lead Time
R14	Compressed project schedule, timely design freeze, advanced project planning, and scheduling	Scheduling and Planning
R15	Inadequate capacity of suppliers for quantity, quality, and complexity	Suppliers' Capacity
R16	Inadequate skills of suppliers for transporting and stocking of prefabricated elements	Suppliers' Skills
R17	Lack of integration in the supply chain with a single-point supplier	Supply Chain
R18	Poor understanding with lack of market research (market forecast, valuation, product positioning)	Poor Understanding
R19	Lack of experienced construction teams, managers, and labors (skills, productivity, availability)	Staff's Experience
R20	Lack of understanding of OSM by local authorities with inappropriate design codes and standards	Authorities Understanding
R21	Complex interfacing between systems, tolerance issues, aesthetics, and monotony of structure	Aesthetics & Tolerance
R22	Difficulty to achieve repetition of consistent layout and economies of scale	Economies of Scale
R23	Early decisions and failure to freeze the design early on	Early Decisions
R24	Lack of construction machinery and technologies and their associated costs	Machinery & Technologies
R25	Lack of consultation service Co, resource R&D, and communication channels	Consultation Service Co.

APPENDIX E: ORIGINAL LIST OF RISK SOURCES

Table 7.5. Original list of risk factors

No.	Risk Sources
001	Productivity and sustainability-driven new urbanization
002	Improved environmental sustainability
003	Social sustainability
004	Save in raw material
005	Occupying extra space for accommodation of precast components
006	Civil disturbances
007	Site constraints and logistics
008	Dominate importance of land acquisition in house building
009	Site-specific constraints, e.g., access limitations and space for large loads
010	modules' dimensional constraints
011	Restricted site layout (e.g. lack of storage space for PPVC modules/lack of space to unload and move the modules)
012	Logistics and site operations, Production facility logistics and stock management difficult (e.g. limited access onsite for maneuver, restricted access to site for delivery, size of components)
013	Module envelop Limitations
014	Inclement weather
015	Impacts of weather conditions
016	Temperature
017	Wind Speed
018	Legal issues
019	Lack of design codes and standards for prefabricated components
020	Uncertain governmental policies
021	Land dominant to government
022	Regulatory authorities: not yet included in planning regulations
023	Inadequate policies and regulations
024	Requirements to meet new regulatory or other imposed requirements
025	Potential unemployment issues to workers "Added after in-depth interview"
026	Inefficiency of design approval
027	Excessive approval procedures
028	Imperfect technological specifications on prefabrication
029	Legal issues
030	Fewer codes/standards available
031	Market protection from traditional suppliers
032	Complex code compliance and inspection process
033	Top-to-down policy support
034	Incomplete policies and standards
035	Immature OSC development conditions
036	Labor agreements or jurisdictional issues
037	Local/Regional political considerations
038	Regulatory requirements
039	Operation and maintenance provisions
040	Continuity through project phases
041	Lack of governmental regulations and incentives
042	Lack of incentives
043	Alignment of drivers
044	High initial cost
045	Difficulty of bidding price from contractors "Added after in-depth interview"
046	High initial cost (cost on new machinery, fabricate molds, and factories)
047	Inadequate project funding
048	Higher capital cost
049	Higher initial (capital) cost to traditional approach
050	need for large initial investment to run modular services
051	Higher initial cost to conventional construction method
052	Higher upfront cost

053	Higher initial cost
054	Initial capital investment
055	Difficulty in obtaining finance, because it requires higher initial cost
056	Specific local economic factors
057	Labor productivity
058	Project goals that include financial incentives
059	Economic condition
060	Cost and Duration Risk Simulation
061	Cost
062	Extra labor cost on checking, counting, and sorting raw materials
063	Potentially higher overall cost to traditional approach
064	availability of cheap labor in the area
065	Higher construction costs to the conventional construction method
066	Overall project cost control
067	Overall project cash flows
068	Future reuse value
069	Cost of construction
070	Cost/value, Perceived as expensive when compared to traditional methods. High initial and set-up costs [18]. Cranage costs can be very high [18]. Intercity or county transport can be very high and can negate any advantage [18]
071	Cost issues
072	Cost of transportation
073	Stability of labor cost
074	Multiple shifts of construction workers
075	Durability of prefabricated unproven
076	Custom check
077	Slow quality inspection procedures
078	Lack of quality assessment tools and accreditation
079	Quality
080	Need for additional protection materials for PPVC modules
081	Quality problems
082	Quality of building
083	Quality, The image of off-site manufacturing is coloured by the experiences of the past, especially around 1960s where some prefabricated buildings collapsed [128].
084	Qualitative Factors
085	Installation error of precast elements
086	Special assembly requirements such as “clean room” conditions
087	Electrical system density
088	Electrical system routing requirements
089	Safety
090	Safety accident occurrence
091	Ensured building quality and work safety
092	Unusual site or regional hazards
093	Ongoing facility operations
094	On-site labor density
095	Increased risk from high elevations, confined spaces, known toxic atmospheres
096	Contractual monetary incentives for a better project safety record
097	Reductions in insurance costs
098	Heavy lifts
099	Insurance and warranties during transport
100	People and Occupational safety and health (OHS), The need for cranes for transporting building components or whole buildings has safety issues associated with their use [18]
101	Process, Require more pre-planning on a project which can potentially increase lead times and may nullify any overall time advantages. Generally very low level of IT integration in the construction industry. Not flexible, does not allow changes as too expensive once manufactured [18].
102	Heavy lift and site
103	Time
104	Longer lead-in time during design stage
105	Weak response to design change during construction
106	Unable to freeze the design early on
107	Early design freeze, due to the long lead-in time, and extensive planning
108	hard to make changes later
109	Requirement for early commitment

110 Equipment or materials with long lead-time
111 Availability of key project team members in early project stages
112 Time
113 Preliminary module definition
114 Owner-furnished long lead equipment specification
115 Change in project scope
116 Tight project schedule
117 Design change
118 Inadequate planning and scheduling
119 Delay of the delivery of precast elements to site
120 Nature of the UK planning system
121 Inflexible/not suitable for late design changes
122 need for more pre-project planning
123 customs delays in borders when transporting internationally
124 The need for additional project planning and design efforts
125 Extensive coordination required prior to and during construction
126 Reduced construction time and labour requirement
127 Shortened schedules
128 Planned shutdowns, outages, or turnarounds
129 Late business decisions
130 Early startup benefits
131 Timing of environmental or other project permitting
132 Time limitations related to shipping and transportation
133 Risks associated with schedule penalties
134 Rewards for early project completion
135 Requirements to get product to market rapidly
136 Requirement for early “freezing” of design
137 Speed of construction
138 Productivity
139 Change in design/scope of work
140 Owner's planning
141 Timely design freeze
142 Early completion recognition
143 Owner delay avoidance
144 Transport Delay avoidance
145 Heavy lifts and related planning
146 Complicated management
147 Availability
148 Lack of manufacturers and suppliers of prefabricated components
149 Reapplication of custom declaration
150 Manufacturing capacity
151 Limited capacity of existing manufacturers
152 Inadequate coordination: procurement, supply chain, site management
153 Expensive long-distance transportation for large and heavy loads
154 Increased transportation and logistics considerations
155 Transportation restrictions due to rules and regulations
156 Availability of transportation methods
157 Transportation infrastructure
158 Supplier availability
159 Availability of qualified suppliers
160 Supplier shop capacity
161 Supplier’s availability of on-site representation
162 Limited market demand
163 Module fabricator capability
164 Transport infrastructure
165 Transportation of prefabricated elements and access to the building site
166 Difficulty to the storage of prefabricated elements
167 Additional procurement costs
168 Additional transportation costs
169 Inefficient verification of precast components because of ambiguous labels

170	Serial number recording error
171	Precast components mistakenly delivered
172	Remanufacturing because of quality control and damage during production
173	Misplacement on the storage site because of carelessness
174	Transportation vehicle damage
175	Transportation road surface damage
176	Logistics information inconsistency because of human errors
177	Difficult identification of proper precast components
178	Reluctance of manufacturers to innovate and change to MMCs
179	hard to transport modules far away
180	Local transportation costs
181	Risks of loss during transportation
182	Foundations required for prework items
183	Level of sophistication of supplier's information systems
184	Ineffective Logistics
185	Poor Manufacturing capability
186	Fragmented industry structure
187	Design information gap between designer and manufacturer
188	Fragmented industry structure
189	Poor integration and interface performance with traditional method
190	High fragmentation in the industry
191	Intermanufacturer rivalry and market protection
192	time delays due to late transit permits for oversized components
193	Protection of proprietary technology or methods
194	Supplier/Contractor flexibility to provide a facility that meets owner's performance requirements
195	Permitting
196	Vendor involvement
197	Operation
198	Uncertainty of market demand
199	Lack of awareness of prefabrication by the market and public
200	Lessons and attitudinal barriers due to historic failures
201	Client conservatism and skepticism
202	Dependence of traditional construction method
203	Reluctance to innovation and driven
204	Lack of local R&D institutes and services
205	Risk averse culture
206	Client skepticism
207	Attitudinal barriers due to historic failures
208	Reluctance to innovate
209	Mindset of the industry (cultural problems)
210	Lacking knowledge and expertise
211	Inappropriate business model
212	Industry market culture, A very conservative industry. Professionals very resistant to change.
213	Special material assembly methods ~alloy welding, etc!
214	Lack of practices and experiences from local projects
215	Organizational mechanism and culture
216	Lack of experienced contractors on prefabrication
217	Lack of experienced collaboration groups
218	Lack of experienced technicians of assembly on site
219	Highly skilled workers
220	Labor dispute and strikes
221	Skills shortages
222	Lack of previous experience
223	Lack of experience and skills
224	more engineering effort
225	availability of knowledgeable engineers and designers in the area
226	Lack of PPVC experiences in term of design
227	Lack of PPVC experiences in term of installation
228	Increased organizational requirements (e.g. changing roles of project participants/increased complexity of procurement and contracting issues)

229 Ever-present need for reducing reliance on manpower
230 Overall or peak labor density requirements ~quantity of workers!
231 Local, regional, or national labor availability
232 Availability of skilled labor
233 Project-specific requirements such as licenses for craft workers
234 Sufficiency of labor in a multiple project environment
235 Total number of laborers
236 Unskilled
237 Skilled
238 Expert
239 Lack of knowledge, Most professionals have not embarked on offsite manufacturing because of lack of knowledge about the benefits of off-site manufacturing [128].
240 Contractor leadership
241 Contractor experience
242 Management of Execution risks
243 Unfavorable organizational mechanism
244 Poor public acceptability: suspicion about meeting customer expectations
245 Clients suspicious about performance, but build at good location for higher price
246 negative perception of new construction methods
247 need for an increased and more detailed coordination in all stages of a project - more communication among all stakeholders
248 Lack of awareness of PPVC's benefits among owners/developers
249 Lack of market acceptance
250 Lack of OSC expertise and stakeholder coordination
251 Dominated traditional project process
252 Lacking social climate & acceptance
253 Limitations to design due to transportation restrictions (e.g. modules' size)
254 Design
255 Monotony of structure type
256 Complex techniques
257 Complex design
258 Incomplete design drawing
259 Redesign because of errors in design
260 Complex interfacing between systems
261 Less tolerance between factory made components and on-site assembly
262 Problems with lightweight construction, e.g., overheating
263 Decreased flexibility for design changes later
264 Unsupportive decision made by designers
265 Poor aesthetic performances
266 Flexibility of design
267 High cost pressure without economics scale effect
268 Inefficient design data transition
269 Difficult to achieve economies of scale
270 Unsustainable: less durable/long-lived, so requires frequent refurbishing
271 Limited market demand
272 Not suitable for small projects, as they require bespoke design
273 Replication on other projects
274 Low standardization
275 Data optimization
276 Unable to modify design scheme
277 Flexibility accommodating modifications or expansion
278 Additional use of tower cranes (vertical transportation)
279 Availability of 3D CAD or similar design technology
280 Infrastructure ~hardware & software! for communications
281 Software compatibility for design and for communication
282 Tower crane breakdown and maintenance
283 Size of equipment of assembly
284 Availability of lifting and hauling equipment
285 Heavy equipment
286 Ease of erection
287 Lack of technologies and testing institute to prefab. Components

288	Monopoly of techniques
289	Lack of experienced design consultancy and designers
290	Low information interoperability between different enterprise resource planning systems
291	Inefficient communication between project participants
292	Lack of long-term cooperation between project teams
293	Low IT integration in the industry
294	Require more communication among all stakeholders
295	Lack of research and development practices and motive
296	Project and/or Owner’s organizational structure
297	Investment in studies

APPENDIX F: MMC RISK ASSESSMENT QUESTIONNAIRE

As part of the ongoing efforts to revolutionize the construction industry, we, researchers from Boğaziçi University, would like you to contribute to the research. The main aim is to encourage stakeholders in the construction industry to adopt a new approach towards the implementation of more innovative modern methods of construction. Your contribution to the research will make it possible to achieve this goal by allowing us to accurately construct a clear vision about how risks can be measured, analyzed, interpreted, and, most importantly, managed.

I have developed the methodology based on expert judgments coupled with computer data mining and data analysis algorithms that, to my knowledge, has never been implemented in the research field. The questionnaire consists of 6 general questions and 25 risk-related questions. It is estimated that the questionnaire will not take more than 15 min to be completed. Now, I would like to stress how valuable your contributions are to the research. Your effort will be rewarded by, upon your request, sharing with you a summary of the research findings and conclusions.

For further information, please let us know

- Research Coordinator: Asst.Prof. Semra Çomu Yapıcı
- Master's Candidate: M.Sc. candidate Ali Tatari, Ali.Tatari01@gmail.com

Note: Any personal information about name, age, gender, etc. will not be collected. - The use of the collected research data, findings are excluded, is limited to only the academic domain. Proceeding to the next step means you have read and accepted the terms and conditions mentioned in the consent form, https://drive.google.com/file/d/1GnnfHpqxkNcihqqk_P3zXQGj48EpGROa/view?usp=sharing.

7.1. General Questions

Q1. Please list all the countries that you have worked in

Q2. Please select the sector you are currently working for

- Consultancy
- Manufacturing (Supplier)
- Construction (Contractor)
- Design (Architect, Structure Engineer, etc.)
- Real Estate Business (Real Estate Developer)
- Other:

Q3. Total years of experience

Q4. Most recent job positions

Q5. In your career, have you be involved with off-site construction projects (Steel structures are considered as off-site construction)?

- Yes
- No

Q6. In your career, what are the type of construction projects that you were part of?

- Industrial
- Residential
- Heavy Construction
- Commercial or Institutional

7.2. Please Specify the Degree of Impact on Project Cost and the Probability of Occurrence for the Following Risk Factors

Q1. Environmental sustainability, resources consumption, waste generation, and waste disposal.

- Very Low
- Low
- Moderate
- High
- Very High

- Rare (0%-20%)
- Unlikely (20%-40%)
- Possible (40%-60%)
- Likely (60%-80%)
- Certain (80%-100%)

Q2. Restricted or unsuitable site layout or space.

- Very Low
- Low
- Moderate
- High
- Very High

- Rare (0%-20%)
- Unlikely (20%-40%)
- Possible (40%-60%)
- Likely (60%-80%)
- Certain (80%-100%)

Q3. Anticipated climatic conditions during construction.

- Very Low
- Low
- Moderate
- High
- Very High

- Rare (0%-20%)
- Unlikely (20%-40%)
- Possible (40%-60%)
- Likely (60%-80%)
- Certain (80%-100%)

Q4. Unexpected Statutory modifications to existing policies.

- Very Low
- Low
- Moderate
- High
- Very High

- Rare (0%-20%)
- Unlikely (20%-40%)
- Possible (40%-60%)
- Likely (60%-80%)
- Certain (80%-100%)

Q5. Insufficient, outdated, or absence of a legal framework.

- Very Low
- Low
- Moderate
- High
- Very High

- Rare (0%-20%)
- Unlikely (20%-40%)
- Possible (40%-60%)
- Likely (60%-80%)
- Certain (80%-100%)

Q6. Few opportunities for obtaining preferential policies on tax, loan, subsidy, etc.

- Very Low
- Low
- Moderate
- High
- Very High

- Rare (0%-20%)
- Unlikely (20%-40%)
- Possible (40%-60%)
- Likely (60%-80%)
- Certain (80%-100%)

Q7. Higher initial capital cost, higher investment in fixed assets, and speed of return on

- Very Low
- Low
- Moderate
- High
- Very High

- Rare (0%-20%)
- Unlikely (20%-40%)
- Possible (40%-60%)
- Likely (60%-80%)
- Certain (80%-100%)

Q8. Currency issues (exchange, inflation, increased loan interest rates, etc.).

- Very Low
- Low
- Moderate
- High
- Very High

- Rare (0%-20%)
- Unlikely (20%-40%)
- Possible (40%-60%)
- Likely (60%-80%)
- Certain (80%-100%)

Q9. Material, labor, maintenance, and operation cost.

- Very Low

- Rare (0%-20%)

- Low
- Moderate
- High
- Very High

- Unlikely (20%-40%)
- Possible (40%-60%)
- Likely (60%-80%)
- Certain (80%-100%)

Q10. Lack of a quality monitoring mechanism for the production process.

- Very Low
- Low
- Moderate
- High
- Very High

- Rare (0%-20%)
- Unlikely (20%-40%)
- Possible (40%-60%)
- Likely (60%-80%)
- Certain (80%-100%)

Q11. Defective, damaged, or the durability of prefabricated elements are unproven.

- Very Low
- Low
- Moderate
- High
- Very High

- Rare (0%-20%)
- Unlikely (20%-40%)
- Possible (40%-60%)
- Likely (60%-80%)
- Certain (80%-100%)

Q12. Increased possibility of safety adverse events due to large units and heavy loads.

- Very Low
- Low
- Moderate
- High
- Very High

- Rare (0%-20%)
- Unlikely (20%-40%)
- Possible (40%-60%)
- Likely (60%-80%)
- Certain (80%-100%)

Q13. Increased lead times in design and construction.

- Very Low
- Low
- Moderate
- High

- Rare (0%-20%)
- Unlikely (20%-40%)
- Possible (40%-60%)
- Likely (60%-80%)

- Very High
- Certain (80%-100%)

Q14. Compressed project schedule, timely design freeze, advanced project planning, and scheduling.

- Very Low
- Low
- Moderate
- High
- Very High

- Rare (0%-20%)
- Unlikely (20%-40%)
- Possible (40%-60%)
- Likely (60%-80%)
- Certain (80%-100%)

Q15. Inadequate capacity of suppliers for quantity, quality, and complexity.

- Very Low
- Low
- Moderate
- High
- Very High

- Rare (0%-20%)
- Unlikely (20%-40%)
- Possible (40%-60%)
- Likely (60%-80%)
- Certain (80%-100%)

Q16. Inadequate skills of suppliers for transporting and stocking of prefabricated elements.

- Very Low
- Low
- Moderate
- High
- Very High

- Rare (0%-20%)
- Unlikely (20%-40%)
- Possible (40%-60%)
- Likely (60%-80%)
- Certain (80%-100%)

Q17. Lack of integration in the supply chain with a single-point supplier.

- Very Low
- Low
- Moderate
- High
- Very High

- Rare (0%-20%)
- Unlikely (20%-40%)
- Possible (40%-60%)
- Likely (60%-80%)
- Certain (80%-100%)

Q18. Poor understanding with lack of market research (market forecast, valuation, product positioning, etc.).

- Very Low
- Low
- Moderate
- High
- Very High

- Rare (0%-20%)
- Unlikely (20%-40%)
- Possible (40%-60%)
- Likely (60%-80%)
- Certain (80%-100%)

Q19. Difficulty to achieve repetition of consistent layout and economies of scale.

- Very Low
- Low
- Moderate
- High
- Very High

- Rare (0%-20%)
- Unlikely (20%-40%)
- Possible (40%-60%)
- Likely (60%-80%)
- Certain (80%-100%)

Q20. Lack of understanding by local authorities with inappropriate design codes and standards.

- Very Low
- Low
- Moderate
- High
- Very High

- Rare (0%-20%)
- Unlikely (20%-40%)
- Possible (40%-60%)
- Likely (60%-80%)
- Certain (80%-100%)

Q21 Early decisions and failure to freeze the design early on..

- Very Low
- Low
- Moderate
- High
- Very High

- Rare (0%-20%)
- Unlikely (20%-40%)
- Possible (40%-60%)
- Likely (60%-80%)
- Certain (80%-100%)

Q22. Lack of construction machinery and technologies and their associated costs

- Very Low
- Low
- Moderate
- High
- Very High

- Rare (0%-20%)
- Unlikely (20%-40%)
- Possible (40%-60%)
- Likely (60%-80%)
- Certain (80%-100%)

Q23. Lack of consultation service Co, resource R&D, and communication channels.

- Very Low
- Low
- Moderate
- High
- Very High

- Rare (0%-20%)
- Unlikely (20%-40%)
- Possible (40%-60%)
- Likely (60%-80%)
- Certain (80%-100%)

Q24. Complex interfacing between systems, tolerance issues, aesthetics, and monotony of structure

- Very Low
- Low
- Moderate
- High
- Very High

- Rare (0%-20%)
- Unlikely (20%-40%)
- Possible (40%-60%)
- Likely (60%-80%)
- Certain (80%-100%)

Q25. Lack of experienced construction teams, managers, and labors (skills, productivity, availability, etc.).

- Very Low
- Low
- Moderate
- High
- Very High

- Rare (0%-20%)
- Unlikely (20%-40%)
- Possible (40%-60%)
- Likely (60%-80%)
- Certain (80%-100%)