DEVELOPMENT OF A COST PREDICTION MODEL FOR MASS TIMBER GRAVITY FRAME CONSTRUCTION

By

Bhushan Rajendra Nankar

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Construction Management – Master of Science

2020

ABSTRACT

DEVELOPMENT OF A COST PREDICTION MODEL FOR MASS TIMBER GRAVITY FRAME CONSTRUCTION

By

Bhushan Rajendra Nankar

Construction materials like concrete and steel have been a primary choice for designers. As a result, the construction industry has become one of the biggest sources of greenhouse gas emissions. With increasing global population and climate change, resource scarcity and a need for healthy habitat is a likelihood. Therefore, to address these issues, green building movement escalated, and mass timber was looked upon as an alternative to offer a family of engineered wood products with comparable strength characteristics and desirable environmental benefits.

Despite its benefits, designers are hesitant to adopt mass timber as a reliable alternative due to their lack of familiarity with the material which leads to uncertainty about costs and acts a barrier for adoption. This study attempts to break this link by investigating the cost implications of using mass timber gravity frame through the development of a cost prediction tool. This tool is envisioned to deliver costs to refer to compare mass timber as an alternative, thereby helping designers to make informed decisions for construction material at a conceptual stage of a project.

This study initially developed regression equations using Principal Component Regression (PCR) and improvised to Partial Least Squares Regression (PLSR) to understand cost implications of selected variables and to develop a prediction tool to address the cost barrier. The researcher also believes that greater adoption will ultimately lead to sustainable forest management, reduced wildfires, and an economic base for rural regions.

ACKNOWLEDGEMENTS

This thesis is an outcome of the opportunity offered to me by Dr. George Berghorn - committee chair and my advisor, as a direct outcome of the confidence Harshvardhan Kalbhor - a friend, a peer and a brother, had placed in me. Without their collective faith, support, and help, it would not have been possible for me to pursue my masters. I am forever indebted to these two concrete supports for providing me the essential foundation of my graduate studies.

This project was made possible by a grant from the US Forest Service Wood Innovations Program.

I would like to offer my special thanks to other committee members – Dr. Syal, for his support and guidance in my academic and personal life struggles; and Dr. Mojgan, for her valuable inputs in achieving the goals of this study in an efficient way. The constructive feedback and expertise of the committee has been very helpful in adding value to this thesis.

I would like to take a moment to thank - Evan Thomas, for his invaluable efforts at embryonic stages of my professional career; John Ciepiela, for helping me in getting a foot in the door of mass timber industry; Christopher Evans and Erica Spiritos, for believing in me and providing opportunities to shape my professional career; Kyle Stevens, for his mentorship in shaping my applied skills. I would also like to thank Graham, William, Zach, Luke, Cody, and Sandra for their positive influence.

My deepest gratitude goes to all the professors at MSU – Dr. Mohamed El-Gafy, Dr. Sinem Mollaoglu, Prof. Tim Mrozowski, Dr. Michael Thibideau, Dr. Lyudmila Sakhanenko, and Dr. Dong

Zhao, who shaped my academic base and provided a diverse and rich experience. I wish to pay my regards to the administrative staff – Jill Selke, for making sure all the paperwork is in place.

I would like to acknowledge my friends more like a family – Aviraj, Sujay, Ayush, Prashant, Shubham, Akshay, Nikitha, and Rahul, for providing a valuable support at various times in my life. I am also appreciative of the support provided by my peers and roommates at MSU – Animesh, Nitesh, Yafeng, Nishchhal, and Saad. I am deeply appreciative of the help of Pooja Mayekar for making sure my accommodation is properly secured even before I arrived in the US. Also, I would like to thank Hari for providing a shelter to stay before I could move in into my apartment.

Last but not the least; I want to thank my family for taking efforts in providing me with the quality education during initial stages of my academic career. Had it not been for the education I received; I would never have been able to make informed decisions. I would like to specially thank Rajiv Bose and Ujjwala Joshi for their support in the fight to pursue my masters. I would like to pay my special regards to Vijay Pakhale, for not limiting his role as a brother but guiding me through various roles.

I am forever grateful and indebt to all these people and a thank you isn't just going to be enough for believing in me at every step of the way. I'd like to add my honest gratitude as you all continue to believe in me with new chapters of my career. I apologize if I missed anyone and express my sincere gratitude for being a vital part of my life.

Spartans Will.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

AEC	Architecture Engineering and Construction
AHP	Analytical Hierarchy Process
ANSI	American National Standards Institute
CBR	Case-Based Reasoning
CF	Cubic Feet
CLT	Cross-Laminated Timber
DF	Degrees of Freedom
DLT	Dowel Laminated Timber
GA	Genetic Algorithm
Glulam	Glued-Laminated Timber
MAPE	Mean Absolute Percentage Error
MLR	Multiple Linear Regression
MRA	Multiple Regression Analysis
MSE	Mean Square Error
NLT	Nail Laminated Timber
PC	Principal Component
PCA	Principal Component Analysis
PCR	Principal Component Regression
PLSR	Partial Least Squares Regression
RCSM	Reduced Cross Section Method

RMSE	Root Mean Square Error
SF	Square Feet
SVD	Singular Value Decomposition
US	United States
WH	Worker Hours

CHAPTER 1 INTRODUCTION

1.1. BACKGROUND

1.1.1. World Population and Climate Change

In the upcoming three decades, world population is predicted to reach 9.6 billion from the total of 7.7 billion today (UNDESAP, 2013). With this population increase, demand for our most fundamental needs - food, clothing, and shelter - is likely to increase. As a result, the burden placed on natural resources will further escalate and eventually we will start facing scarcity of these resources. In a study by Matos (2017), a major share of non-renewable sources is found to satisfy non-fuel and non-food sectors of the economy. Since 1900, the use of construction materials mirrors it accurately by increasing the consumption of crushed stone, sand, and gravel from about 35% to 60% of non-fuel and non-food material consumption in totality (Sznopek and Brown, 1998). Moreover, the percentage of non-renewable resources contributing the United States economy increased from 54% to 96% in 2014 (Matos, 2017).

World population is not the only the issue faced by the globe, climate change also requires scientific and policy attention. Greenhouse gas concentrations are rising rapidly in the atmosphere and this increase is becoming a threat to the United States economy and the world. The greenhouse gas emissions stabilization is necessary to create healthy habitat for living organisms. Construction industry is one of the biggest sources of these emissions in the United States which accounts for 39% of CO_2 emissions and is expected to grow rapidly in commercial sector (Kinzey et al., 2002). Concrete and steel dominate the construction industry as they are primary construction materials. The production of cement, which is a main ingredient of concrete, and steel are the prime sources of these emissions (Laguarda-Mallo and Espinoza, 2018). These sources are some of the most difficult emissions to reduce (Winchester and Reilly, 2019). Furthermore, the energy consumption of buildings is significant as they account for 70% of the electricity demand in the United States (EIA, 2008). A good substitute with optimum emissions and energy consumption has become the prerequisite to achieve the stabilization and to protect the environment.

1.1.2. The Forest Products Industry and U.S. Economy

The United States is the fourth most forest-rich country in the world. Forests in the United States occupy about 33.9% of the total land area, of which 14.5% is private corporate timberland. The forest products industry represents manufacturing of pulp, paper-based packaging, wood building products, and own forest lands. This industry comprises about 1.5% of the total U.S. economy and approximately about 6% of the total U.S. manufacturing Gross Domestic Product. Moreover, the wood manufacturing alone produces over $90 billion in products annually (Wali et al., 2010). As of 2017, 937,500 individuals were employed in construction with U.S. forests and 428,500 individuals worked in the manufacturing of wood products. This industry is among the top 10 manufacturing employer in 48 states and provides jobs in all 50 states (AF&PA, 2019). It pays a significant amount of federal, state, and local taxes amounting $7 billion annually. According to some economists, approximately 11 forest sector jobs are created with every million board feet of harvested timber. This industry plays an important role in rural regions and provides an economic base for such regions. In 2015, primary forest products generated approximately one third of the total forest sector jobs in Oregon (Oregon Employment Department, 2016).

1.1.3. Sustainable Forest Management

Sustainable forest management has resulted in a greater number of trees in the United States than there was a century ago (Alvarez, 2007). Despite these efforts, forest management is facing new challenges like increasing size and frequency of wildfires and deteriorating forest health. Between 1984 and 2016, wildfires in the US burned an area about the size of Texas, approximately 160 million acres (Alvarez, 2007). In the last three decades, bark beetle outbreaks have killed billions of trees (Moens et. al., 2003). The relative stocking of Western hemlock, Douglas fir and other small diameter trees that are conducive to wildfires and pest outbreaks has increased in the national forests (Crampton, 2017). Such small diameter trees have very little intrinsic value and are not viable economically, thus they are not typically harvested and stocking increases. An abundance of these trees and amplified requirements of resources for management of these forests are a major reason for such outbreaks and wildfires (Laguarda-Mallo & Espinoza, 2018). These issues are a significant threat to the needs of future generations and can limit the forest resources. The removal of such low-value timber and small diameter trees can be considered a good strategy to improve the forest health and sustainability.

1.1.4. Construction Industry in U.S.

In 2017, the United States produced goods and services worth $19.3 trillion and the construction industry contributed $826 billion, representing about four percent. The construction industry employed approximately 7.3 million workers in August 2018, with an average pay exceeding 10% of the average of all private sector employees. In an industry-wide survey, four out of five construction firms stated that there is a shortage of the craft workers which characterizes major portion of construction workforce (AGC, 2019). Associated General Contractors of America's chief economist said that, labor shortages in the construction industry are a concern and needs to be addressed properly to continue the economic growth.

1.2. MASS TIMBER PRODUCTS IN U.S. CONSTRUCTION INDUSTRY

Traditionally, wood has dominated residential construction, while conventional construction materials like concrete and steel have been used heavily in the other construction markets in the United States. New single-family houses and multi-family housing units in residential construction have consistently provided an essential market for wood. Annually, one out of three wood products consumed in the United States are used in the construction of these units (McKeever & Elling, 2015). Walls represented 37%, roofs 35%, and floors used 22% of all wood used in new residential construction (McKeever & Elling, 2015). Wood has been the single most dominating building material in the residential construction industry.

Even after such strong markets in residential construction, these products are not extensively used in the non-residential markets. Generally, designers have focused on conventional construction materials like concrete and steel. As the green building movement grew, designers started considering timber as an alternative due to its environmental benefits.

A new family of engineered wood products collectively known as "mass timber" was looked upon as an important timber product for commercial construction due its comparable strength characteristics and desirable environmental benefits. Mass timber is a collective term used to represent a family of engineered wood products such as cross-laminated timber (CLT), nail-laminated timber (NLT), glue-laminated timber (glulam), and dowel-laminated timber (DLT). These engineered wood products have broadened the availability of wood products for construction (Werner and Scholz, 2002). Designers, architects, and engineers have begun to explore these materials for construction due to their efficiency in construction, comparable structural properties, and insulation properties, in addition to benefits for environment (FPInnovations, 2013). The following sections will provide detailed information on these products.

1.2.1. Nail Laminated Timber

Nail-laminated timber (NLT) has been used for more than a century; as a part of the new mass timber system, it has gained a new identity with recent modifications. NLT is a solid structural element which is a composition of individual dimensional lumber, placed on edge and then the individual laminations fastened together using nails as shown in Figure 1.1 (BSLC, 2017). It is comprised of dimensional lumber of nominal thickness 2x, 3x, or 4x and width varying from 4 in. to 12 in. width (BSLC,2017).

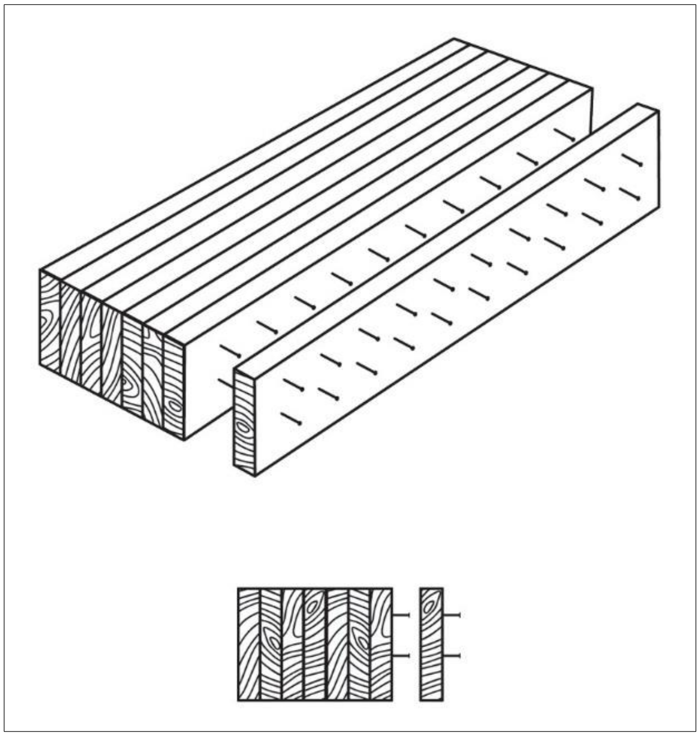

Figure 1.1: Nail-Laminated Timber (Source: Think Wood, 2020)

It can be used for wall, floor, and roof structures. NLT offers a unique opportunity to create both singly curved and freeform panels, due to single direction spanning of individual boards comprising the panel (BSLC, 2017). It has also been used to build elevator and stair shafts in mid-rise wood-frame buildings. It is particularly popular in warehouses where sturdy and solid floors are required, recently it has gained popularity in institutional and residential buildings for exposed slabs and decks to create a unique aesthetic (BSLC, 2017).

1.2.2. Dowel-Laminated Timber

Dowel-Laminated Timber (DLT) also known as "*Dübelholz*" in Europe, it is the first all wood engineered product in North America. DLT does not require use of any chemicals, glues, nails or screws for connecting adjacent panels, rather it uses wood dowels as connectors. DLT is made from softwood lumber panels, which are arranged similar to NLT and then are connected using wood dowels as shown in Figure 1.2.

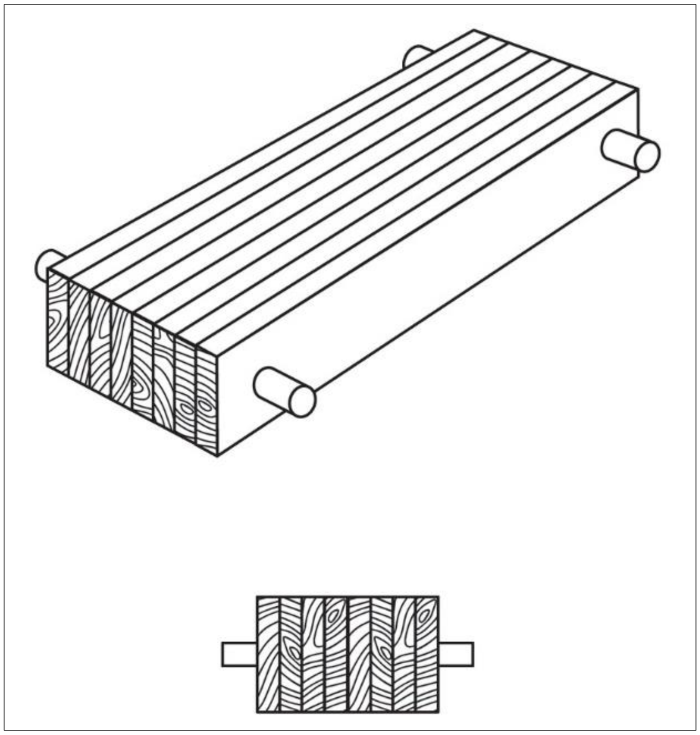

Figure 1.2: Dowel-Laminated Timber (Source: Think Wood, 2020)

These panels use hardwood dowels to friction fit pre-milled boards together on edge, to provide dimensional stability to the panel. DLT panels can be processed through computer numerical control (CNC) machinery to meet the tight tolerance requirements and predrill Mechanical, Electrical, and Plumbing conduits. DLT is available in large panel sizes up to 12’ wide x 60’ long for efficiency on-site along with the flexibility to be made from wide range of wood species (Structure Craft, 2018).

1.2.3. Cross-Laminated Timber

CLT is a comparatively new engineered wood product in the US markets, first developed in the 1990's in Austria. It is also known as "X-lam" or in German "Brettsperrholz" (BSP). It is an engineered wood product made up of layers of solid sawn lumber glued at right angles to the adjacent layer as shown in Figure 1.3. CLT provides a unique opportunity to use low value timber coming from sources such as lesser used species, diseased or infected trees, and comparatively young trees with small diameters (Laguarda-Mallo & Espinoza, 2015).

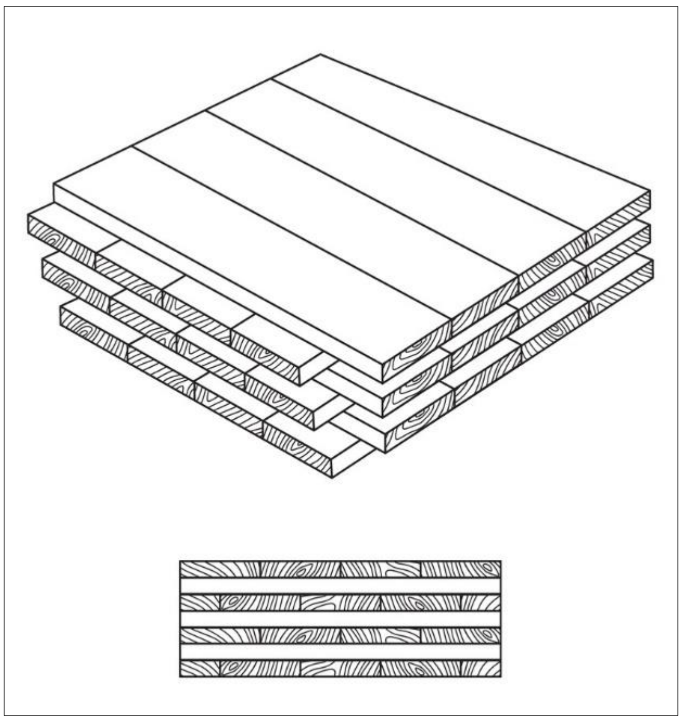

Figure 1.3: Cross-Laminated Timber (Source: Think Wood, 2020)

It can be used in walls, floors, and roof construction. It can be prefabricated with the openings for windows, doors, and service runs before shipping to the site. It is available in three to seven or more layers with varying thickness of layers. Figure 1.4 depicts the different layer combinations used in CLT manufacturing. With thickness variation, structural strength and span length varies which defines panel's use for certain applications. As shown in Figure 4, 3-layer panels can be used in walls and secondary constructions due to their strength characteristics; 5-layer panels of 160 mm thickness can be used in primary construction as walls and floors spanning less than 5.5 m; 11-layer exceptional panels can also be used in heavy applications like bridge decks.

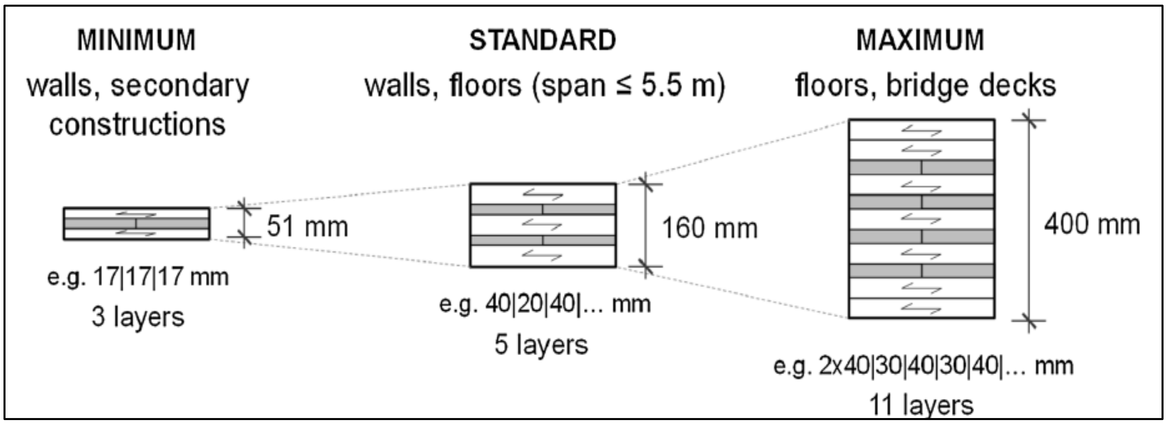

Figure 1.4: Cross-Laminated Timber Layer Combinations (Source: Brandner et al., 2016)

1.2.4. Glued-Laminated Timber

Glue-laminated timber (Glulam) is a combination of individual wood laminations, placed based on their performance characteristics and the application of the member and bonded together using adhesives. The individual laminations are oriented in the same direction, parallel to the length of the member as depicted in Figure 1.5. Glulam is typically used as a part of structural framing members such as beams and columns. It is available with a wide range of properties and appearance grades depending on its application. The glulam manufacturing process allows the application of complex curvature and unique geometry to the members. This provides the use of such members in complex structures without compromising the architectural intent. ANSI Standard A190.1-2012: Standard for wood products-Structural Glued Laminated Timber sets a nationally recognized benchmark for the quality of production. Despite providing stringent standards, it allows the manufacturer the flexibility to select any method which will produce an equivalent or a superior quality product.

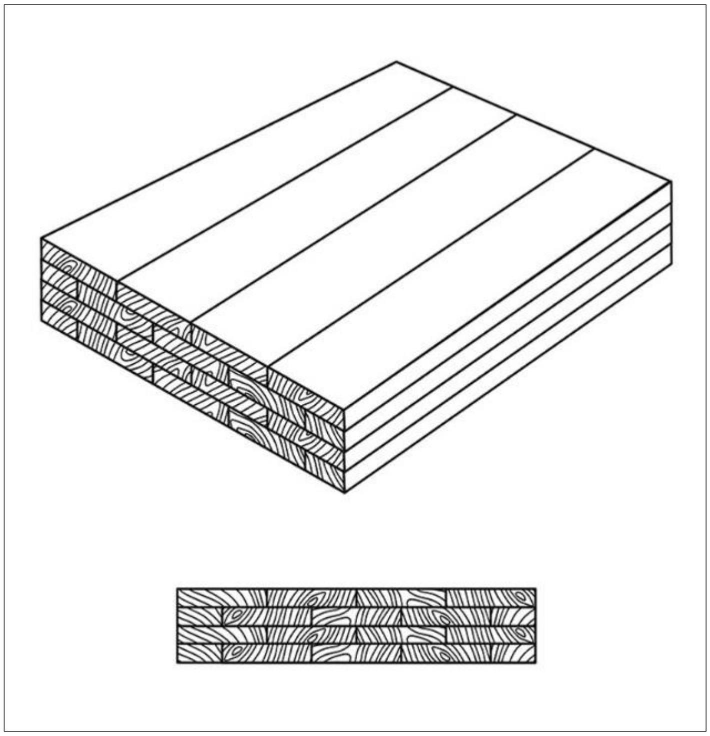

Figure 1.5: Glued-Laminated Timber (Source: Think Wood, 2020)

Glulams are produced based on their application and required performance characteristics, following are the layups serving different purposes and providing different performance characteristics. These layups are a combination of different grades of lumber positioned in such a way that is appropriate for its end use. Based on the function of these members, they are manufactured in three different layup combinations; Unbalanced, Balanced, and Single-grade. These layups are explained with their suitable use and combinations in the following subsections. There is a various grade of lumber which can be used in these layups as shown below.

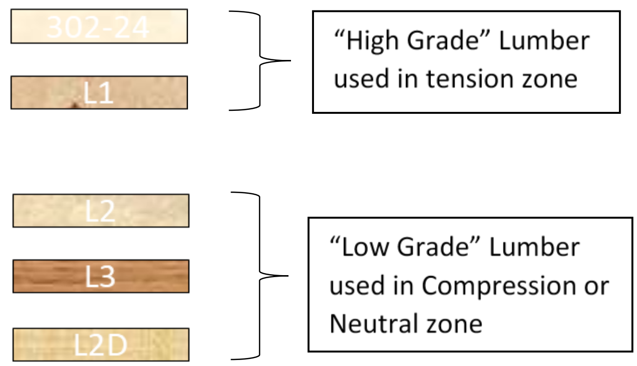

Figure 1.6: Different Grades of Lumber Used in Glulam

The Engineered Wood Association has provided certain percentages for the number of laminations to use in each zone in the layup. These percentages are mentioned in the ANSI 117-2010 and shall be multiplied by the total depth of the member to calculate the number of laminations. Different layup combinations are described below along with the percent tension laminations required.

1.2.4.1. Unbalanced Layup:

These layups are made up of compression laminations at top and tension laminations at bottom (Figure 1.7). This layup provides unequal capacity in positive bending and negative bending and hence are known as unbalanced layup. Usually used as a simple span or as a short cantilever, this layer requires 5% of tension laminations at the bottom of the beam to achieve suitable strength (ANSI 117-2010).

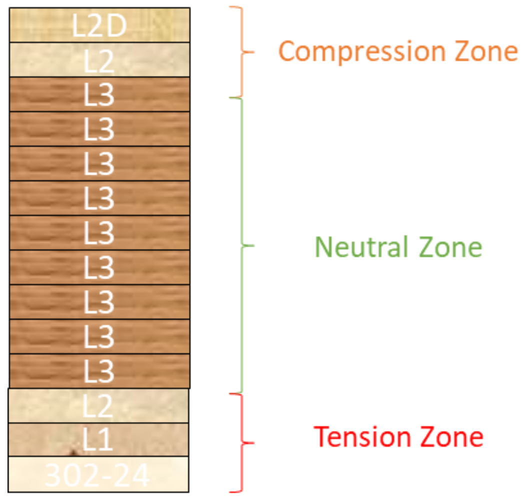

Figure 1.7:Unbalanced Layup

1.2.4.2. Balanced Layup:

These layups are made up of tension laminations at both top and bottom (Figure 1.8). This layup provides equal capacity in positive and negative bending and therefore are known as balanced layup. It is generally used in continuous spans and long cantilevers. It requires 5% of tension lamination at both top and bottom of the beam as per standards as a minimum requirement (ANSI 117- 2010).

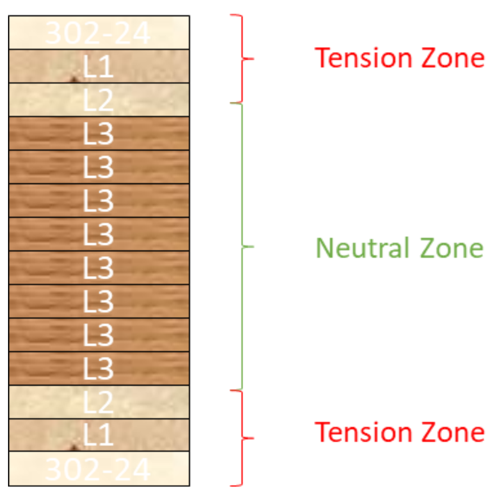

Figure 1.8: Balanced Layup

1.2.4.3. *Single-Grade Layup:*

These are made up of usually either tension or compression laminations (Figure 1.9). This layup uses same grade laminations throughout the depth of beam that is why known as single grade layup. This layup is common in axially loaded members like columns or truss chords (American Wood Council, 2015).

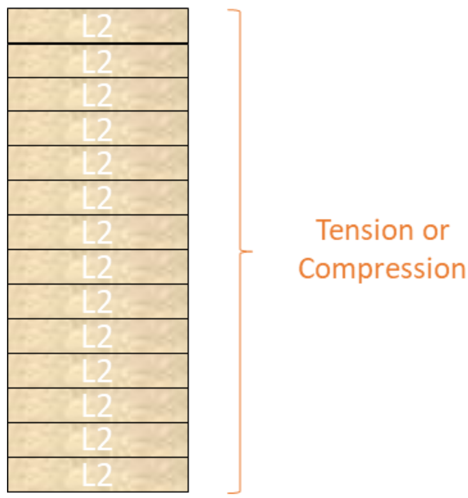

Figure 1.9: Single-Grade Layup

1.3. ADVANTAGES AND BARRIERS OF MASS TIMBER PRODUCTS

With the rising concerns on limited resources, population increase, climate change, and wildfires, there is a need to find alternatives to manage these issues. Engineered wood products have numerous advantages and can contribute to mitigate some of these problems. The following sections elaborate current issues and show how the use of engineered wood products can be useful in managing such issues.

1.3.1. Wildfire Reduction

In recent years, size and frequency of wildfires have increased significantly. The three-year average costs incurred by federal agencies for fire suppression have increased from $0.5 billion to $2 billion. Small-diameter and diseased trees are major contributors to such wildfires, particularly, small diameter trees like a young Douglas-fir has a thin bark & low hanging, needle covered branches which usually is in contact with the ground creating a surface hazard. These small-diameter trees and other surface bush creates a surface hazard and often provides a starting point for wildfires. Following bark beetle outbreaks, trees are more vulnerable to fire as the dead needles remain on the tree. Dead trees provide a fine dry fuel that can catch fire quickly in favorable weather conditions. These fires are then escalated into the tree crowns after the dead trees fall and provide ladder fuels for the spread.

Cross Laminated Timber (CLT) provides a unique opportunity to use low value timber coming from sources such as lesser used species, diseased or infected trees, and comparatively young trees with small diameters (Laguarda-Mallo & Espinoza 2015). It can be manufactured using "junk" trees with diameters as small as 4 inches. The extraction of such low value timber has been considered a good strategy to increase the value and health of the forests (Quesada-Pineda et.al 2018). CLT is manufactured from spruce, although pine species, larch, and fir species can be used, which are more vulnerable to fire. Creating strong markets for such timber will help to reduce the fire hazards to some extent and help manage forests.

1.3.2. Advantages in Construction

In addition to its benefits for forests, CLT has numerous other advantages in the construction industry. CLT is a promising construction material that has proven to be more energy efficient, environmentally friendly, a better material to build comparatively dense built environment, and a material to improve construction speed. CLT acts as a thermal mass that stores heat during the day and releases it at night which lowers overall building energy use (Laguarda-Mallo & Espinoza, 2014). CLT is a renewable material and sequesters carbon instead of emitting carbon like concrete (Laguarda-Mallo & Espinoza, 2014). One of the most important characteristics of CLT is its high strength-to-weight ratio which leads to smaller building foundations for comparable structural capabilities and translates into additional floors with the same weight of structure (Laguarda-Mallo & Espinoza, 2014). CLT's can be prefabricated in a controlled environment and therefore can provide fast paced construction (Schwarzmann et al., 2017).

Faster construction speed means a shorter construction phase, which ultimately adds to the indirect cost and construction time savings compared to the reinforced concrete and steel buildings. Therefore, providing strong markets for CLT will not only reduce the harmful low value timber for forest health, but also will help the construction industry to build more efficient built environment.

1.3.3. Modular Construction

Cost and time overruns account for a significant portion of the challenges faced by the construction industry. Several attempts and techniques have been developed to reduce overruns threatening the success of projects. However, construction projects continue to suffer these problems. Modular construction is an emerging construction approach to provide a solution to such difficulties. It moves the construction process away from the job site into a controlled factory environment (Jiang et al., 2018), eliminates weather delays and provides safer construction (MBI, 2016). It enables delivery of a building as an assembly of a set of modules manufactured offsite (Salama et al., 2017). As a result of its concurrency of offsite and onsite construction operations, modular construction provides higher schedule control (Salama et al., 2017).

In addition to its advantages for schedule control, this technique has numerous advantages for the construction industry. Firstly, it allows the reduction of construction schedules by the parallel scheduling of offsite and onsite schedules to save 30 to 50 percent duration of the project (MBI, 2016). Secondly, it delivers consistent quality and offers accuracy due to its manufacturing in a controlled factory environment (Lawson et al., 2012). Thirdly, it has contributed to the sustainability of the process by reducing the construction disposal of wastes (Moghadam et al., 2012).

Cross laminated timber emphasizes hybrid modular construction, which is the combination of modular and on-site construction. The advantage of this construction is that both types of construction can take place simultaneously and provides the luxury of scheduling it parallelly (Schoenborn, 2012). The shorter project duration reduces indirect project costs and thus reduces the overall cost of the project (Salama et al., 2016). In a survey, 67 percent of firms reduced project schedules using modularization and 35 percent of firms experienced a decrease of four weeks or more (MGH, 2011).

1.3.4. A Case Study for Mass Timber Construction

The following case study demonstrates the use of CLT and a glulam frame and shows how prefabrication can improve construction process.

- Project Name: First Tech Federal Credit Union, Oregon Corporate Office
- Location: Hillsboro, OR.
- Completion Date: June 2018
- Architect: Hacker Architects
- General Contractor: Swinerton Builders
- Manufacturer: Structurlam

The First Tech Federal Credit Union is a five-story, 156,000 SF office building, constructed with CLT roof and floors with a glulam frame (figures 1.10 & 1.11). As of today, it is the largest CLT structure in the US, which took 14 months in total, with only 12 weeks for timber erection (APA, 2018). With wood prefabrication and an integrated approach with the project team, a four percent cost savings was observed, and four months were saved when compared with the use of steel. Wood presented benefit for erection time, fireproofing, and foundation systems.

Figure 1.10: First Tech Federal Credit Union Project (Source: Swinerton Mass Timber)

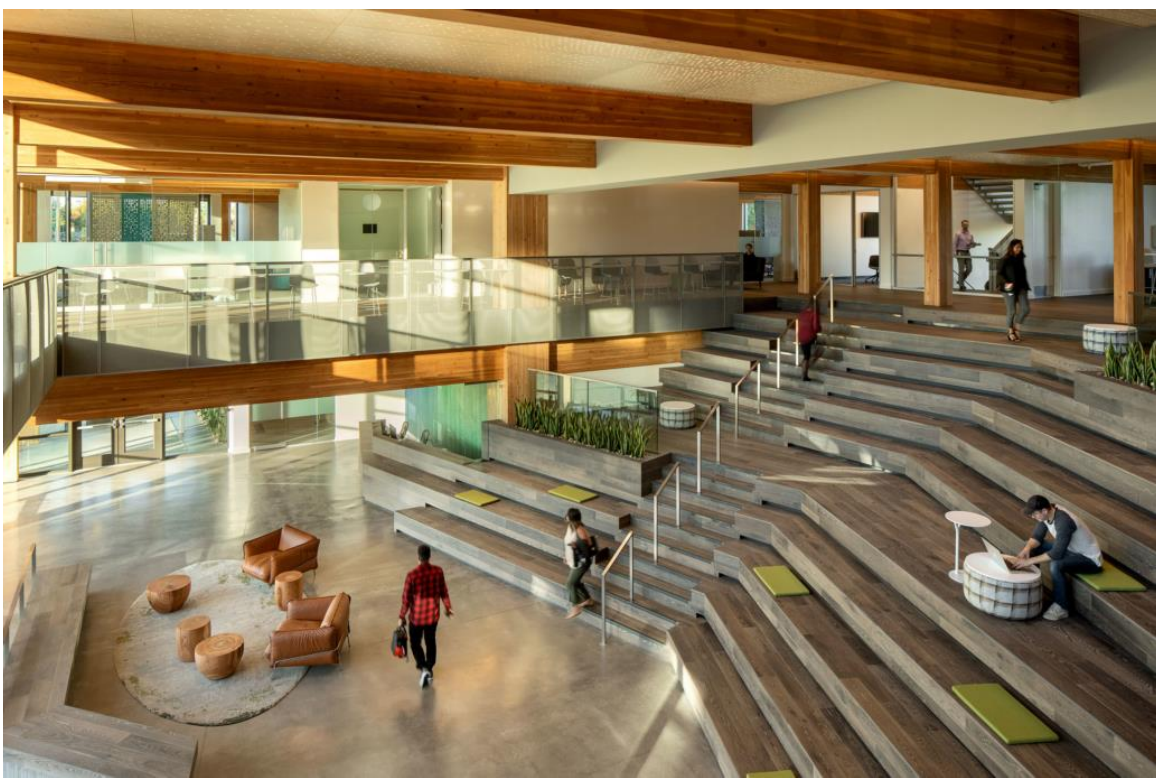

Figure 1.11: First Tech Federal Credit Union Project (Source: Swinerton Mass Timber)

1.3.5. Barriers and Need Statement

The selection of structural materials happens early in the design process and one of the important factors influencing the decisions is the cost of the material in addition to structural performance, fire performance, availability in the market & compatibility with the building code (Laguarda-Mallo and Espinoza, 2016). Furthermore, economic performance of the construction material is one of the highest rated attributes that engineers look for in a material. In a national survey of U.S. structural firms, 88.5% of respondents rated this attribute as “extremely important”, “very important” or “important” (Laguarda-Mallo and Espinoza, 2018). In addition, the results of the survey in multifamily housing highlighted construction costs as the number one factor for choice of a structural system (Schmidt and Griffin, 2013). In order to make an informed choice about building materials, architects and engineers should be well aware of the new materials & the factors involved in the decision making of the structural material.

Specific to the case of CLT, the situation is exactly contradicting. The results of a countrywide survey indicate that the level of awareness of CLT is low to moderate in U.S. engineering firms. The initial costs and the building code compatibility issues are the largest perceived barriers for the adoption of CLT (Laguarda-Mallo and Espinoza, 2018). But, CLTs have proved to be cost competitive by saving the costs associated with onsite labor using less construction time up to 30% (Laguarda-Mallo & Espinoza, 2015). This shows an uncertainty about the cost competitiveness of CLT which leads to its unfamiliarity in the industry.

To counter the initial cost barrier, manufacturing representatives from Europe suggested the evaluation of the life cycle costs for the cost competitiveness (Laguarda-Mallo and Espinoza, 2014). A prediction tool emphasizing on construction costs will help the AEC industry to take informed decisions at the conceptual stages of the project. This tool will provide the solution to one of the biggest barriers for its acceptance i.e. cost, which will predict the cost required for the proposed gravity frame & will give a clear idea of its feasibility.

In summary, the decision of a construction material is majorly dependent on the costs of the material. Given the unfamiliarity about mass timber and the material itself, there is an ambiguity about the costs of the material. This creates a loop of unfamiliarity leading to ambiguity in costs and resulting in the avoidance of mass timber as an alternative for construction material. Therefore, a cost prediction tool is required at the conceptual stages of a construction project, which will provide an idea about the costs during the decision-making process of a structural material. This will result in the understanding of the costs and will provide a reference to evaluate mass timber as an alternative to make informed decisions.

There are various methods for developing such a tool to predict costs, including multiple regression, case-based reasoning, genetic algorithm, artificial neural networks. These methods are explained through their application in prediction tools in chapter 2. With the goal to develop a cost prediction tool, these methods will help in building the statistical model which then can be then used as a core working foundation for this tool. Given the fact that there are small number of mass timber projects across US, a small sample size was expected. Method selection is planned to be based on the accuracy, ease of use, and small sample size.

1.4. RESEARCH GOAL, OBJECTIVES, AND METHODOLOGY

As discussed previously, the cost barrier has received a limited attention and has proved to be a major barrier for the greater adoption of mass timber. The main goal of this research is to understand cost implications of using CLT and glulam in construction. This goal is envisioned through the development of a cost prediction model. Objectives of this research have been aligned towards achieving the research goal, and they are described below.

1.4.1. Objective 1: Understand the current state and background of CLT and glulam in construction

<u>Step 1: Literature Review</u>

The intent of this step is to explore and study previous contributions to the body of knowledge and learn about the developments and shortcomings in the knowledge base. A comprehensive overview is provided in chapter 2, to include manufacturing processes, construction techniques, system advantages, and prediction model development methods.

<u>Step 2: Field Observation</u>

As a part of the review of the current state of the system, field observation was performed to understand the construction process and to study the cost implications of various parameters in the process. Field observation is a part of building a foundation for deciding the input variables of the model. Also, this step assists in discovering limitations that can occur in the field during installation.

1.4.2. Objective 2: Build a foundation for development of the predictive cost model

Step 3: Determining Variables

Variables were identified from the literature review and field observation, and with the assistance of expert knowledge. These variables were used to create a spreadsheet for data collection which enlisted these variables for inputs from industry.

Step 4: Data Collection Process

The intent of this step was to gather the data for the model development and to help in soliciting appropriate analysis method. The spreadsheet prepared in Step 3 was circulated as a part of this step to collect values for these variables. This spreadsheet was updated with incoming information and after the process of data collection was used as a datasheet. Specific to the information received, a suitable statistical method search initiated.

Step 5: Technique Selection

Literature review and field observation helped in outlining the variables and preparing a spreadsheet. After data collection process, obtained sample size was a factor of consideration for the method selection. Statistical experts were consulted for appropriate method selection and data analysis started.

1.4.3. Objective 3: Development and testing of the model

Step 6: Development of the model

This step included preparation of the model and analysis of the possible data and variables. The model was created to satisfy the statistical intent and provide the cost outputs with the help of available sample size. The major scope of this step was to include all the obtained data in the model without loss of any important information.

Step 7: Testing of the model

A sample project was planned to validate the model and to check the accuracy of the model.

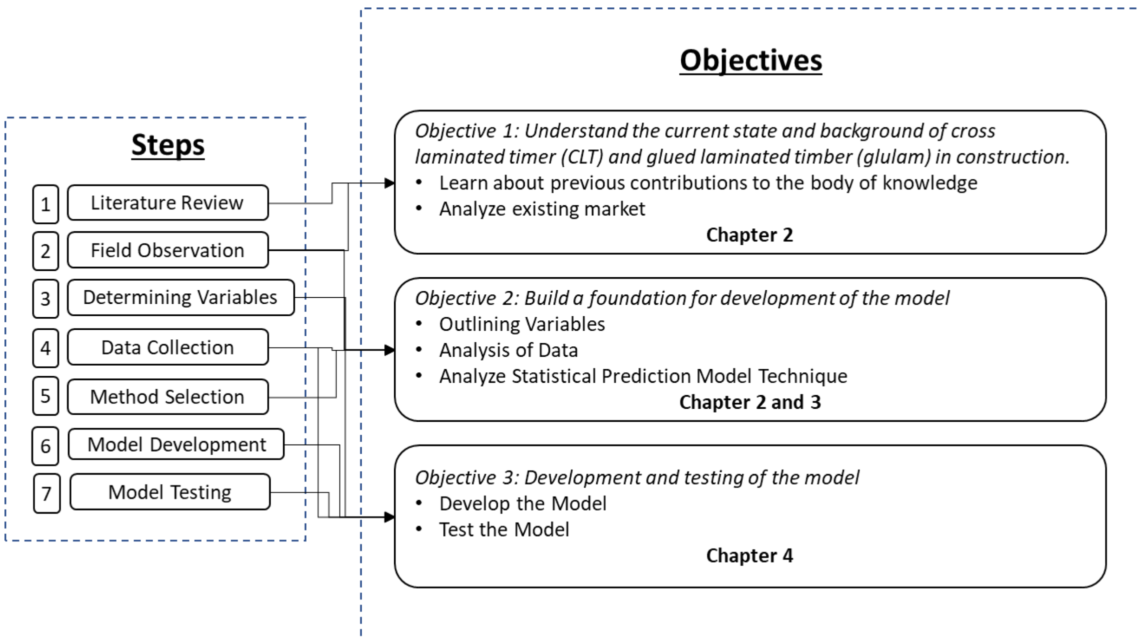

Figure 1.12: Research Objectives and Steps

1.5. RESEARCH SCOPE AND LIMITATIONS

The cost prediction model is developed especially for structures constructed using a glulam gravity frame system. The gravity frame is the structural frame constructed using CLT and glulam to transfer gravity loads. The planned prediction model is designed specifically to consider the costs associated with this timber frame. Costs apart from timber frame including other areas of building involving concrete, steel or any other materials will not be considered as a cost estimated through this tool. As a small sample size is expected, considering costs associated with the timber frame prevents differences in the costs of different building types, such as office, multifamily, civic, educational, etc. It provides a common reference of comparison for all the samples by giving a common gravity frame reference which is essentially a part of every building type. This benefits the model by making the most out of the data at this initial stage of this market, where a few numbers of projects are actually built. This approach allows this study to substantially increases the sample size by boiling down all the projects to a timber gravity frame instead of dividing it into clusters of different type.

This tool is primarily focused on the construction costs of the building and does not involve overall costs involved with the project lifecycle.

This research covers a small number of projects constructed using CLT and glulam. As a result, the sample size is relatively small and therefore the predictive modelling technique must be selected based on the sample size. The technique should be able to generate accurate results with the small sample size.

1.6. RESEARCH OUTCOMES AND FUTURE SCOPE

The primary goal of this research is to increase the cost knowledge base of the construction using CLT and glulam and to break the cost barrier for its adoption by increasing awareness through a predictive cost tool. The following are planned outcomes of this study:

1. This study is intended to provide the cost implications of mass timber based on various parameters related to construction costs.

2. This study will result in delivery of a prediction tool to provide a clear cost estimate at the conceptual stage of projects.
3. The cost prediction tool will increase the awareness of CLT by overcoming the barrier of construction cost uncertainty.

CHAPTER 2 LITERATURE REVIEW

The objective of this chapter is to review the background of the area of study and provide a comprehensive review of information about it. This is aimed to result in greater understanding of the area of study and to outline suitable methodologies which can be used for the prediction model. A topical outline of this chapter is provided in Figure 2.1.

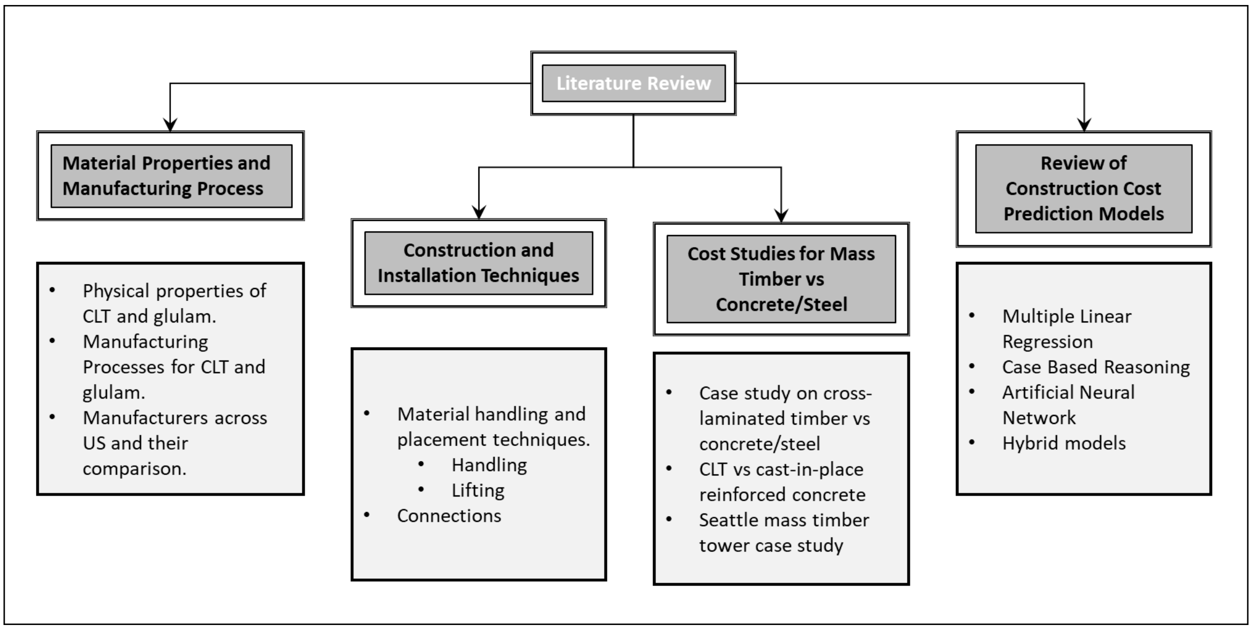

Figure 2.1: Literature Review Outline

This chapter is designed in such a way that it covers material, labor, and equipment aspects of mass timber construction. The review begins with background information about mass timber materials, their associated manufacturing processes, and the current state of the market. The review then shifts to labor and equipment aspects of mass timber buildings, with a focus on installation techniques. Following this groundwork, the review provides cost and time studies on mass timber and finally, a review of methods used in prior cost prediction studies. Therefore, this chapter is supported on four piers: 1) material properties, manufacturing processes, and current state of the market, 2) construction and installation techniques, 3) cost studies for mass timber vs concrete/steel construction, and 4) a review of methods used in building cost prediction models.

The first section of this chapter will focus on the basic properties of wood to demonstrate strength characteristics of glulam and CLT. With wood being questioned for its fire capabilities, fire resistance characteristics will also be explained as a part of this review. After establishing its strength characteristics, the review will proceed into the manufacturing process of these two materials. It will compile the processes involved in the manufacturing of these materials and will provide standards for manufacturing. Mass timber is an emerging construction market, as a result, there are a limited number of manufacturers and knowing about their locations and capabilities will provide an idea about current state of this market. This is envisioned through providing maps, showing locations, and comparing capabilities of a few known manufacturers.

The second section of this review emphasizes resource intensive construction and installation techniques. Labor-dominant onsite connections, and equipment-dominant material handling and lifting are explained in this section. The purpose of this section is to provide some background on field operations and idea about activities utilizing resources.

The third section focuses on cost studies performed to investigate feasibility and economic performance of CLT. These studies compare mass timber construction with concrete/steel, cast-in-place concrete, and post tension concrete structure. It provides an overview of cost breakdowns and potential cost savings.

The fourth section of this review is dedicated towards reviewing previously developed construction cost prediction models with a focus on methods used. It involves cost models developed using multiple regression analysis, case based reasoning, and artificial neural network.

2.1 BACKGROUND, MANUFACTURING PROCESS AND CURRENT STATE OF THE MARKET

This first section of the review starts with unfolding the characteristics and manufacturing process of CLT and glulam. After building the foundation about the background of these materials, this section provides the current state of the market by providing an overview of currently active manufacturers in the United States.

2.1.1 Background of CLT and Glulam

This section is presented to provide background information on CLT and glulam materials by introducing their properties for their applications in construction. As a result, this section will summarize structural properties, fire resistance, seismic behavior, and durability of wood in brief. The next subsection will then highlight on their manufacturing process and the current state of this market in the US.

Wooden building materials are classified into three categories; 1) solid wood, 2) engineered wood, and 3) reconstructed panels. Reconstructed panels are typically used as wall panels and finishes (Asdrubali et al., 2017). Limiting the scope of this study to timber gravity frame, solid wood and engineered wood products are explained below which are typically used in framing and following that properties of wooden materials are explained.

Solid Wood: As its name indicates, this is a solid piece of wood which is derived from timber with the best characteristics in size, growth, and dimensions (Asdrubali et al., 2017). Solid wood does not require any adhesives to hold fibers together. Disadvantage of this material relate to the presence of natural knots and limitations on dimensions based on the size of the tree.

Engineered Wood: With limitations imposed on the size and quality of solid wood, engineered wood products can be manufactured from a wide range of characteristics of trees. They are assembled from small pieces of wood with the help of adhesives to create desirable size, shape, and predictable strength characteristics. This provides much better utilization of the natural resources, and also provides commercial value for small-diameter and otherwise low-grade lumber. Engineered wood products, glulam and CLT are relevant to this study, which can be manufactured in the desired sizes with the help of finger joints, which has the connection efficiency of 100% which means the strength of the connection is equal to the strength of the member it connects (Asdrubali et al., 2017). Table 2.1 summaries common sizes for these wood products and their applications.

Table 2.1: Common Sizes and Applications of Solid Wood, Glulam, and CLT (Source: Asdrubali et al., 2017)

Product	Elements	Function	Common Size	Applications
Solid Wood	Unidimensional: beams, columns, header beams	Load bearing	Length: Up to 5.4m	Structural frames, floors, roofs
			Width: 25-75 mm	
			Depth: up to 250 mm	
Glulam	Unidimensional: beams, columns, trusses	Load bearing	Length: no theorotical limit (40 m)	Bridges, halls, industrial buildings, arenas, distribution centers,
			Width: 60-250 mm	schools, commercial buildings, supermarkets and residential
			Depth: from 180 up to 2000 mm	buildings.
Cross	Bi-dimensional: walls, floors, roofs	Load bearing	Length: up to 20 m	Residential and tall buildings, schools, auditoria, exhibition
Laminated		and shear	Thickness: 50 - 300 mm	places, places of workship, sports halls, theatres and
Timber		walls	Depth: up to 4.8 m	commercial buildings
(CLT)				

<u>Properties of Wooden Materials for Building Applications</u>

Structural Properties:

In addition to the benefit of carbon sequestration, wood as a building material shows equivalent properties for both compression and friction to that of concrete, but with a major advantage of a fifth of the total weight of similar volume of concrete (Asdrubali et al., 2017). This makes wood an excellent choice for applications where weight of the material is a major factor for design. This high strength-to-weight ratio also helps in seismic design and foundations. Wood performs poorly in the direction perpendicular to the grain direction and with presence of knots in its structure. Moisture content of wood plays an important role in defining its strength and stiffness characteristics. With an increase in moisture content below the fiber saturation point, strength and stiffness of wood decreases (Gerhards, 1982). When compared to concrete, Young's modulus of wood is a point of concern, which is nearly thirty percent of concrete and can create issues related to vibrations, buckling, and deformability.

Seismic Behavior of Wooden Structures:

High strength-to-weight ratio, viscoelastic properties, and deformability makes wood appropriate for seismic-resistant construction. Despite these advantages, non-ductile stress-strain behavior of wood is a major area of concern, hence the ductile behavior of wooden structures is focused on the steel connections (Fragiacomo et al., 2011). The failure mechanism is thus governed by the design of steel connections based on the capacity design to compensate for the brittle nature of wood (Asdrubali et al., 2017).

Fire Resistance of Mass Timber:

Fire resistance in simple terms can be explained as the capacity to resist the detrimental effects of fire without losing structural integrity. Wood is assumed as a poor performer in fire due to its combustible nature. Wood has a low thermal conductivity and specific gravity which makes it a slow transmitter of heat (White, 1983; Ritter, 1990). Even if wood burns, while burning it forms a char layer which insulates the unburned wood, as shown in figures 2.2 and 2.3. The char layer of wood itself acts as a fire protection and the wood inside the pyrolysis zone can be considered intact to carry structural loads. Formation of char layer depends on the species of wood used, Douglas-fir, which is commonly used for engineered wood products, chars at a rate of 1.5 inches per hour. When combined with its low conductivity makes wood a better fire resistive material than other conductive materials like steel with the rise in temperatures (Ritter, 1990).

To protect wood structures from fire-related failure, Eurocode 5 (2004) suggests the use of the reduced cross section method (RCSM). In this method, member sizes are defined in such a way that after burning, the remaining intact cross section of the member can carry the structural loads successfully and prevent the structure from failure. Figure 2.3 illustrates the cross section of a burnt wood beam.

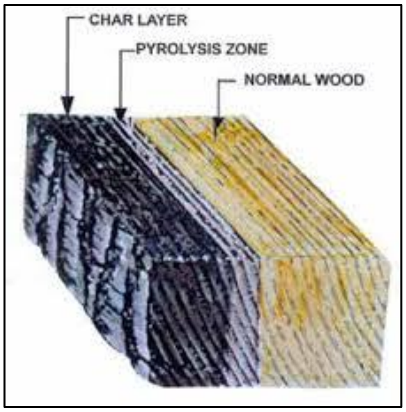

Figure 2.2: Charring Behavior of Wood
(Source: Ritter, 1990)

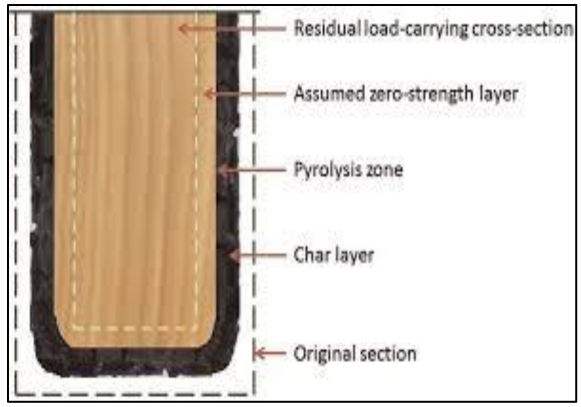

Figure 2.3: Section of a Fire Damaged Wood Beam to Illustrate Char Layer and Zero Strength Layer
(Source: White et al., 2013)

Durability :

Moisture and biological agents are the primary influences on the durability of the wood used as building materials. Due to its interaction with water, wood material volume changes as a result of the swelling and contraction which can cause cracks in wood; additionally, humidity creates a favorable condition for fungi growth (Eaton, 1993). The biological agents that attach to the wooden building materials majorly depend on the type of wood and not on the relative humidity. Therefore, wood used in buildings is usually preserved through either impregnation or through an adequate design of architectural details (Greg et al., 2002).

Other important properties of wood include renewability, carbon sequestration, and embodied energy. Steel is an alloy and concrete is a composite material, which require substances that are mined and heated by fossil fuels at very high temperatures (Sarthre & O'Connor, 2010). These are energy intensive processes compared to wood processing, that makes it more efficient and desirable material.

2.1.2 Manufacturing Process and Standards for CLT and Glulam

The high precision of the engineered wood products distinguishes mass timber from other construction types. Mass timber installation processes are highly influenced by the tolerances in the member sizes, because CLT and glulam arrive ready-to-install onsite. Therefore, manufacturing standards have been developed which are represented in Table 2.2. Figure 2.4 illustrates the manufacturing process for CLT, and Figure 2.5 demonstrates manufacturing process for glulam.

Table 2.2: Manufacturing Standards for CLT and Glulam

Tolerances		Cross-laminated timber	Glued-laminated timber
		ANSI/APA PRG 320-2017	ANSI A190.1-2017
Dimensions	Thickness/Depth	± 1/16 inch (1.6 mm) or 2% of CLT thickness, whichever is greater	+ 1/8 inch (3 mm) per ft of depth. - 3/16 inch (5 mm) or 1/16 inch (2 mm) per ft of depth, whichever is larger
	Width	± 1/8 inch (3.2 mm)	± 1/16 inch (2 mm)
	Length	± 1/4 inch (6.4 mm)	up to 20 ft: ± 1/16 inch (2 mm) Over 20 ft: ± 1/16 inch (2 mm) per 20 ft of length or fraction thereof
Squareness		The length of the two panel face diagonals measured between panel corners shall not differ by more than 1/8 inch (3.2 mm)	± 1/8 inch (3 mm) per ft
Straightness/Camber		Deviation of edges from a straight line between adjacent panel corners shall not exceed 1/16 inch (1.6 mm)	Up to 20 ft: ± 1/4 inch (6 mm) Over 20 ft: ± 1/8 inch (3 mm) per additional 20 ft of length or fraction thereof, but not to exceed 3/4 in. (19 mm)

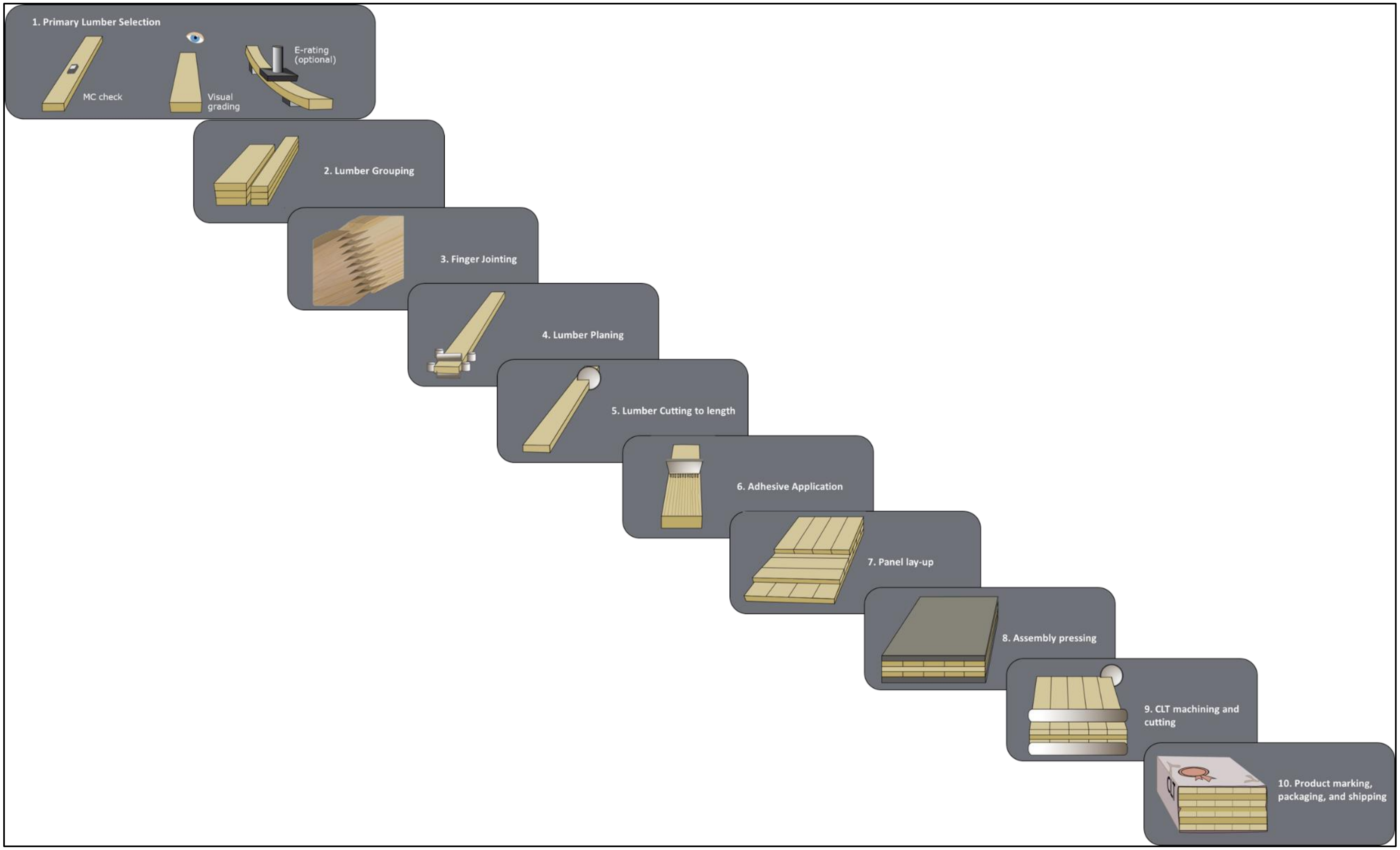

Figure 2.4: CLT Manufacturing Process (Source: CLT Handbook, 2013)

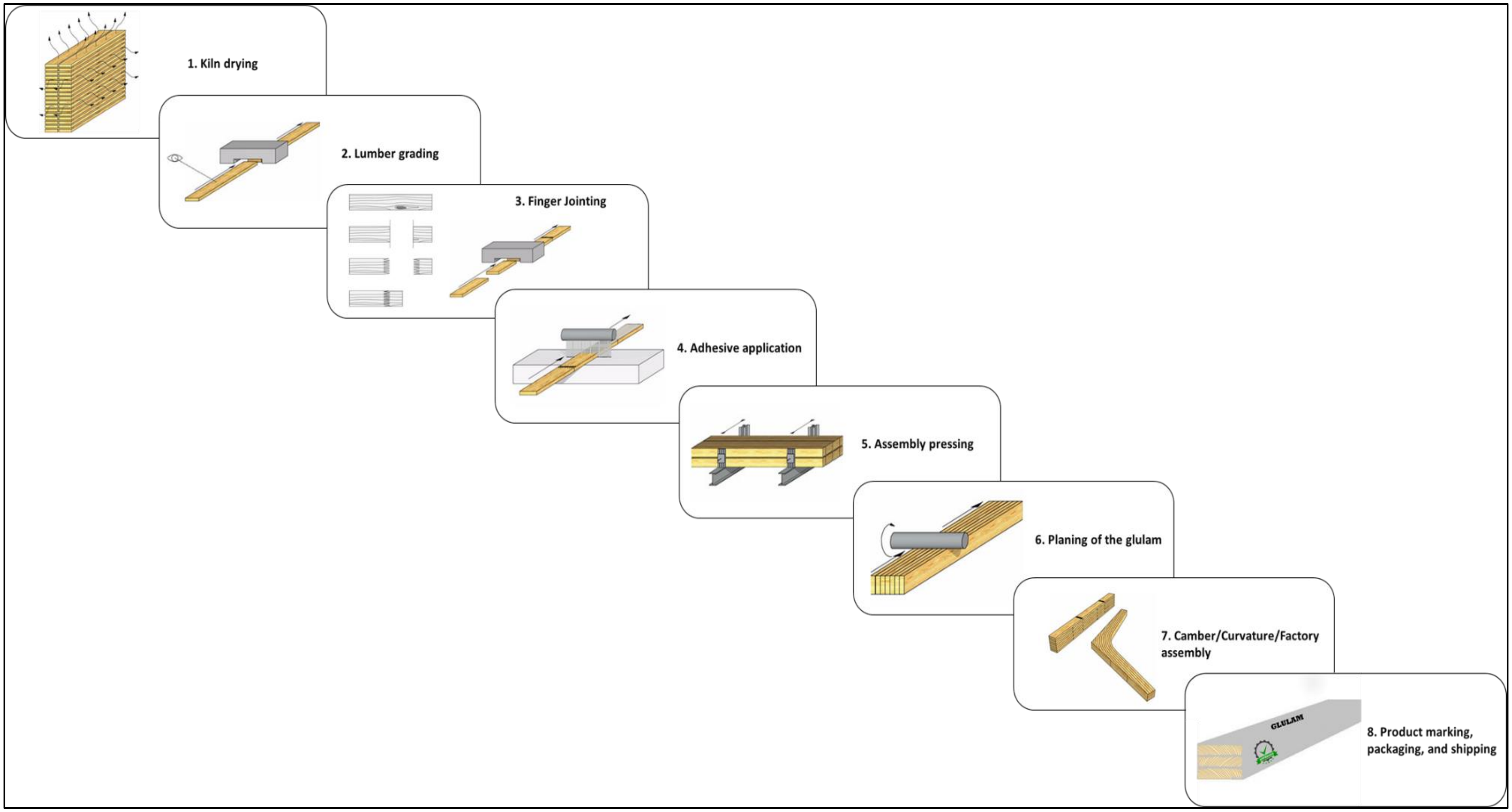

Figure 2.5: Glulam Manufacturing Process (Source: Brettschichtholz, 2013)

2.1.3. Current state of the market

The United States mass timber construction market is still considered at its initial stages. With that said, there are a limited number of manufacturers in the United States with expanding capabilities. Figures 2.6 and 2.7 represent locations of active glulam and CLT manufacturers in the US. Table 2.3 represents manufacturers in the US, their services, and abilities.

Figure 2.6: Location of Glulam Manufacturers Across US

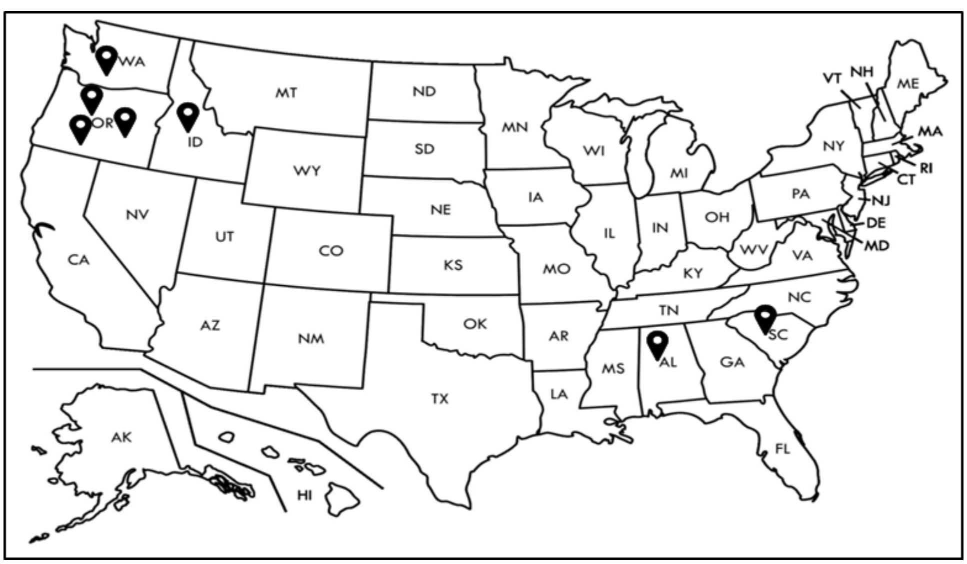

Figure 2.7: Location of CLT Manufacturers Across US

Table 2.3: Manufacturers Across US and Their Capabilities

	Manufacurer	Location	Certification	CLT Dimensions	Glulam Dimesions	Species	CNC Capability
CLT MANUFACTURERS	D R Johnson	Riddle OR	APA/ANSI Certified 3-lam Certified 5-lam Certified 7-lam Certified	Maximum Panel Size 10' X 24'	Maximum Glulam size 20" X 9' X 130' long	Douglas fir, Alaskan Yellow Cedar	NA
	Katerra	Spokane, WA	3 layers 5 layers 7 layers 9 layers	Max. Length 60'-3" Width 9'-9" to 11'-9"	NA	Spruce-pine-fir	Yes
	Vaagen	Colville, WA	NA	12" thk , 4' wide, and 60' long	Straight glulam beams up to 60 ft. long	Spruce-pine-fir, Douglas fir larch	Yes
	Smartlam	Columbia Falls, MT	APA Certified	Upto 41' long but new plant will be able to produce panels 12' wide and 64' long	NA	Douglas Fir, Western Larch, and SPF Grade 2	Yes
GLULAM MANUFACTURERS	Calvart	Vancouver, WA	APA & FSC Certified Length: Up to 88' Width: 1/2" to 30" Depth: Up to 60" Camber: Zero, 2000',3500'	xxx	Can manufacture most custom curved shapes including S-curves, side curves, parabolic arches, double compound curves and curves with radius as tight as 2'-0" 3500' and 4500' radius with no Camber	Douglas Fir and Alaska Yellow Cedar, other species upon request	NA
	American Laminators	Swisshome, OR	APA Certified	xxx	Length: 134' Depth: 73 1/2" Width: 14 1/4" Others are available on request	Douglas Fir, Port Orford Cedar, Alaskan Cedar, Southern Pine	NA
	Western Structures (Custom Glulam)	Veneta, OR	APA Certified	xxx	Length: 8' through 100' Depth: 3" through 53" Widths: 11/2" through 161/4" 0" Camber up to 40' inside radius	Douglas fir, Alaskan Yellow Cedar	NA
	Rosboro	Springfield, OR	APA Certified	xxx	Length: 60' Width: 3 1/2", 5 1/2", 6 3/4", 8 3/4"	Douglas Fir	Yes
	QB Corp	Salmon, ID	APA Certified	xxx	Length: up to 130' Depth: 3" to 96" Width: 1.5" to 20" Ability to curve the beams with inside radius radius of 7'	Douglas Fir, Alaskan Yellow Cedar, Wetsern Red Cedar, Southern Yellow Pine	NA

2.2. INSTALLATION TECHNIQUES

The first part of this chapter focused on the materials aspect of the construction; this second section emphasizes the labor and equipment aspects. Mass timber construction is unique, as CLT panels and glulam members arrive ready-to-install onsite, therefore the installation process becomes a critical activity. Hence, building a groundwork to understand installation provides an evaluation of the costs contributing to the overall project costs. To provide a background, this section starts with different construction systems used in mass timber and following that it explains field operations which is divided into two parts based on the dominant resources: 1) lifting and handling of materials (equipment dominant) and 2) connections (labor dominant).

There are various construction systems adopted using wood as a material to achieve desired quality and efficiency (Asdrubali et al., 2017). Relevant to the area of this study, mass timber framed structures are mainly used in following two systems:

- *X-Lam Structures:* In this construction system, CLT is used as a load bearing element which is used for constructing walls, floors, and roofs. Figure 2.8 depicts this construction system.

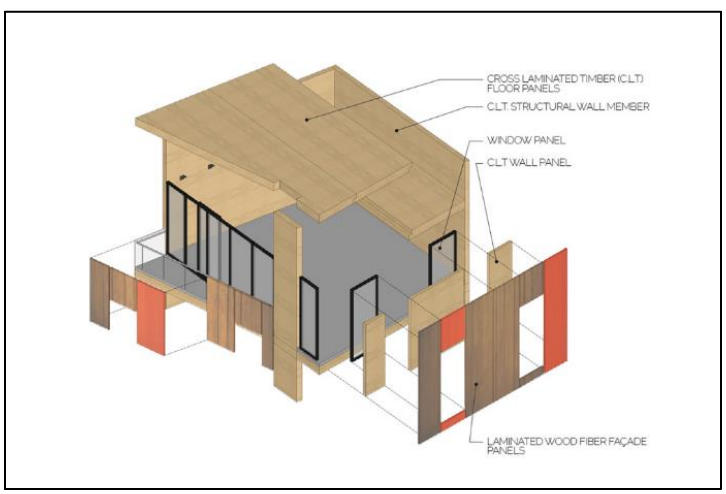

Figure 2.8: X-Lam Structure (Source: Jasontran, 2020)

- *Tall Building Structures:* This system is very common and currently used to build mass timber projects, such as Candlewood Suites, First Tech Federal Credit Union, Albina Yard etc. This method utilizes a glulam gravity frame which is tied with internal core of either CLT or reinforced concrete to resist lateral loads; CLT panels are used for floors or in the perimeter as shear walls.

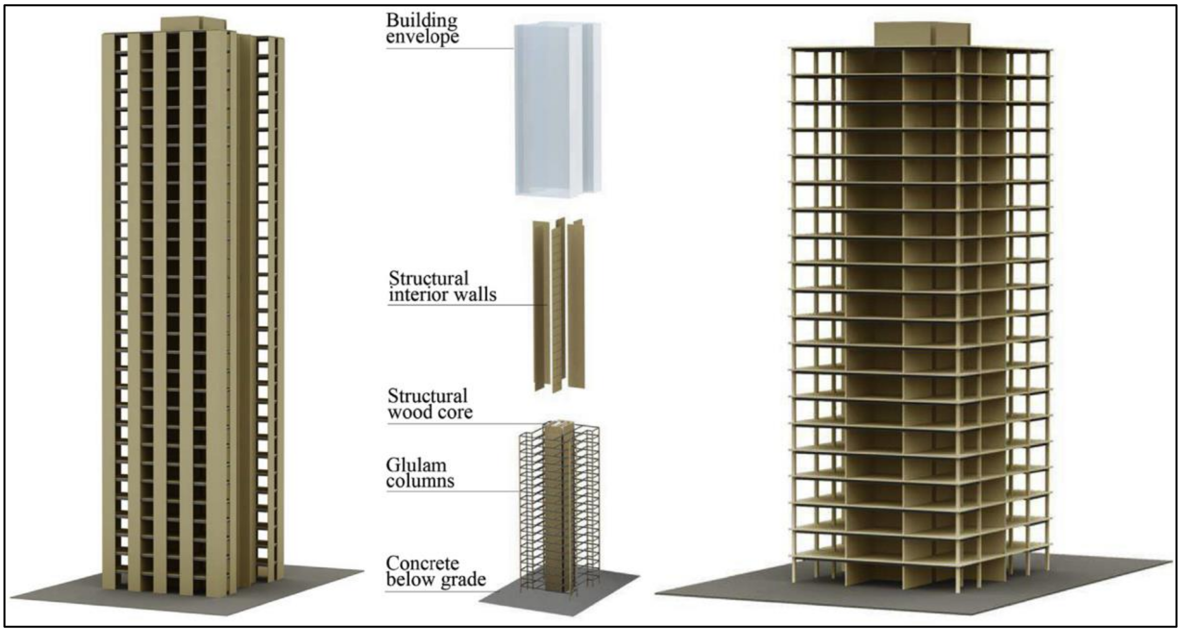

Figure 2.9: Tall Building Structures (Source: Asdrubali et al., 2017)

2.2.1. Lifting and Handling of Materials

Cross Laminated Timber

Cross laminated timber panels are transported from the manufacturing facility or dockyard to the site location on flatbed trucks. The best practice to load CLT is to arrange panels in the order of installation sequence and to stack them without damaging the visual side. These panels are stacked in such a way that large panels are placed individually and separated with dunnage material. Dunnage is typically used to ease the process of unloading panels using forklift or crane. A proper staging area is required to store CLT onsite, with additional dunnage used to keep the panels off of the ground. Forklifts with long forks, typically eight feet long, are required to support the width of the panels when unloading. After unloading and storing materials, the next step is to lift panels and install them at their desired location. To lift these massive panels, lifting devices such as Rampa inserts with lifting loops (figures 2.10 and 2.11) and yokes with straps are used. Rampa inserts and lifting loops are used to lift horizontal panels weighing less than 5,000 pounds.

Figure 2.10: Rampa Insert
(Source: MyTiCon Timber Connectors, 2019)

Figure 2.11: Lifting Loop
(Source: MyTiCon Timber Connectors, 2019)

Depending on the lifting system, lifting devices can also be designed using steel plates, known as "yokes". There are five lifting systems and each one has unique requirements for lifting devices. Figure 2.12 depicts lifting with a spreader bar and Figure 2.13 illustrates lifting with bridle assemblies.

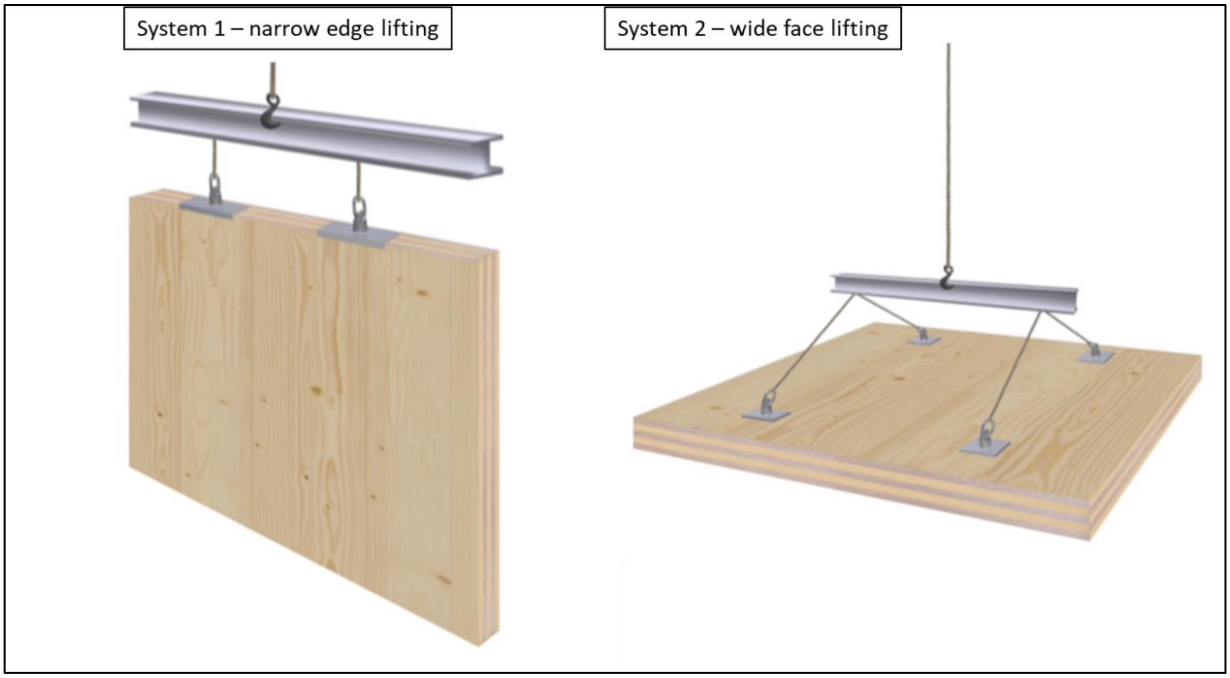

Figure 2.12: Lifting with a Spreader Bar (Source: MyTiCon Timber Connectors, 2018)

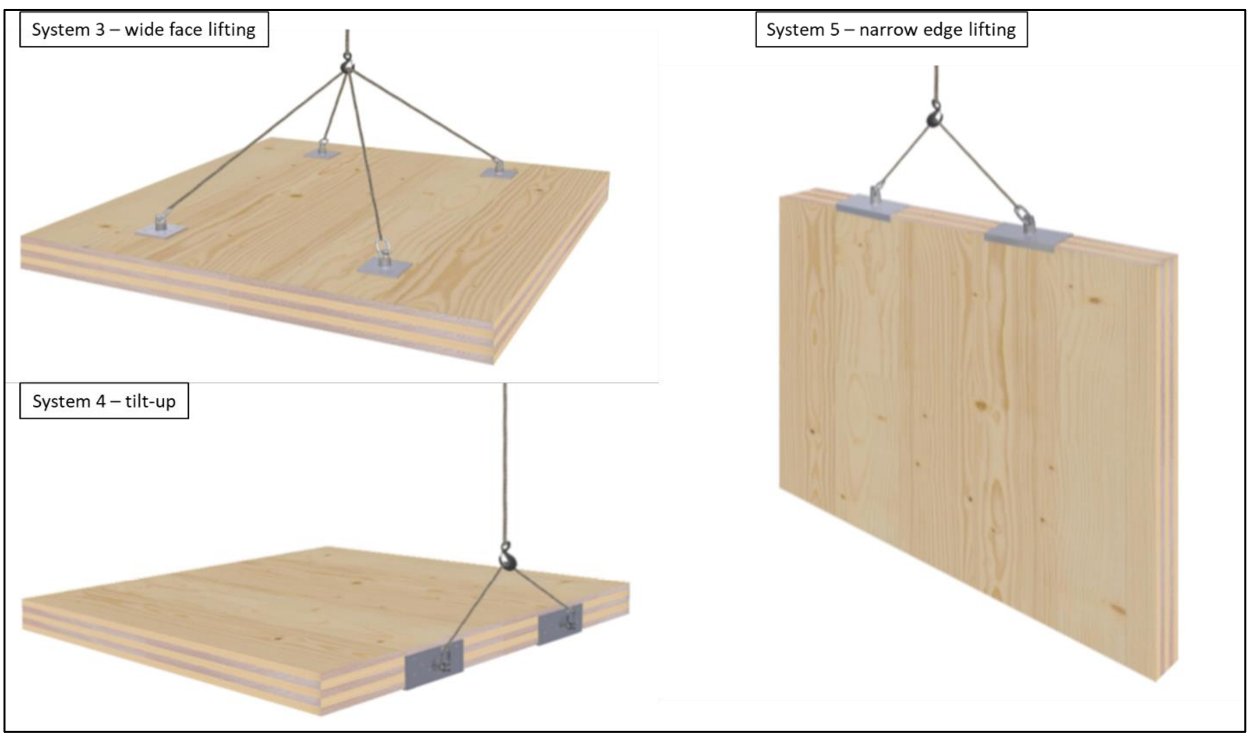

Figure 2.13: Lifting with Bridle Assemblies (Source: MyTiCon Timber Connectors, 2018)

Based on the edge of lifting and the capacity required, yokes are selected. Figure 2.14 portrays different yokes used for rigging panels.

Figure 2.14: Different Types of Yokes (Source: MyTiCon Timber Connectors, 2019)

Yoke 1T is used for floor/roof panels up to 3500 lbs., Yoke 5T is used for panels weighing less than 11,900 lbs., and Yoke XL has the largest capacity to lift panels up to 18,500 lbs.

Glued-laminated timber

Glulam beams are shipped typically in rail cars or trucks and are stacked on lumber blocking or skids. Loads are secured with the help of straps to provide stability and safety. After arrival onsite, glulam beams are unloaded with the help of forklifts or cranes, with wrapping left intact if necessary. Beam sides are placed flat on the forks for better stability. For long beams, multiple forklifts can be used in unison. If a crane is used for unloading, sufficient protection is required at the edges of beam, between slings and edges. Storage of beams if possible is preferred to be in a covered area, elevated off of the ground using blocking or skids. Glulam beams are typically installed using cranes, and nylon or fabric slings are highly recommended compared to steel chokers for lifting with adequate protection. Figures 2.15 and 2.16 illustrates unloading and lifting of the glulam beams.

Figure 2.15: Unloading of Glulam Beams
(Source: APA, 2013)

Figure 2.16: Lifting of Glulam Beams (Source: unknown)

Handling and lifting processes for these materials are equipment intensive. Most of the equipment-related hours are used in these processes, which makes it important to develop understanding of these processes.

2.2.2. Connections

Assembling and installing connections is a labor-intensive activity in mass timber construction, which involves fastening different members together. Some connections are prefabricated and some of the connections are performed onsite. There are some connections which are project specific and some typical connections. This section focuses on such typically used connections related to the timber gravity frame system.

Typical Connections

CLT floor panel to CLT floor panel:

These connections can be found in most of the mass timber buildings using CLT as a floor element. These connections are performed onsite and the main activity for laborers is to connect panels together with structural screws. The CLT Handbook explains these connections in a detailed manner (FPInnovations, 2013) and relevant details are only captured in this section of the report. These are usually used in two types: 1) spline connection and 2) lap joint. Figures 2.17, 2.18, and 2.19 depict different types of spline connections and Figure 2.20 illustrates a typical lap joint.

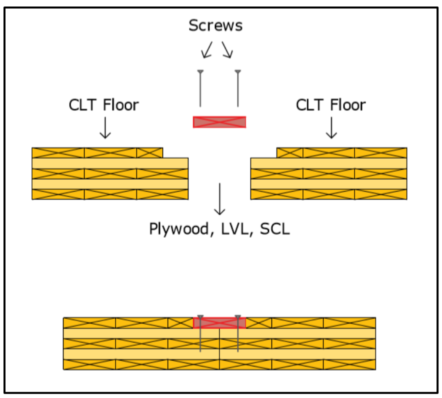

Figure 2.17: Single Surface Spline
(Source: FPI, 2013)

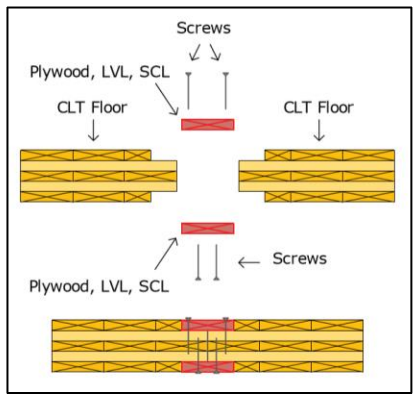

Figure 2.19: Double Surface Spline
(Source: FPI, 2013)

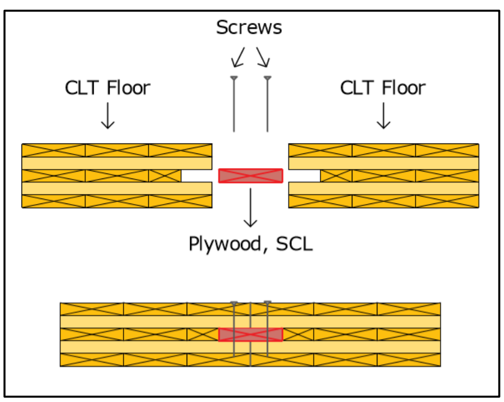

Figure 2.18: Single internal spline connection
(Source: FPI, 2013)

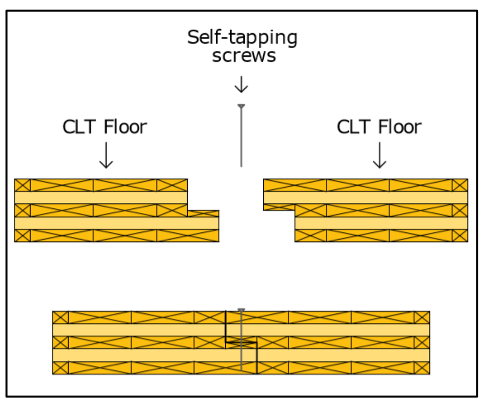

Figure 2.20: Half lapped joint
(Source: FPI, 2013)

CLT floor panel to glulam beam:

This is one more typical connection observed in mass timber construction, used to attach CLT floor panels to glulam beams using structural screws. These connections are also performed in the field during building erection. Figure 2.21 illustrates such a typical connection connecting a 5-ply CLT to a glulam beam.

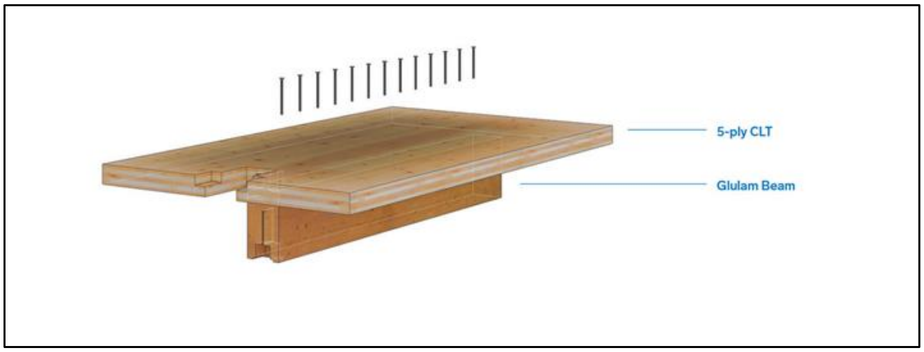

Figure 2.21: Typical CLT Floor Panel to Glulam connection (Source: LEVER Architecture, 2020)

<u>*Project specific connections*</u>

Other connections involve project specific designed connections, which are custom made steel connections used as column to column, beam to column, beam to beam, and column to foundation connections. Important observations to note while designing these custom-made, prefabricated steel connections is that they should be concealed inside wooden members to achieve desired fire resistance. An example is illustrated in Figure 2.22, in which steel connections are concealed inside wood members. An additional advantage of concealed connections is that they are not visible and maintain all wood surface aesthetics. Another array of connections is depicted in figures 2.23 and 2.24 which represent another type of prefabricated post to beam connection.

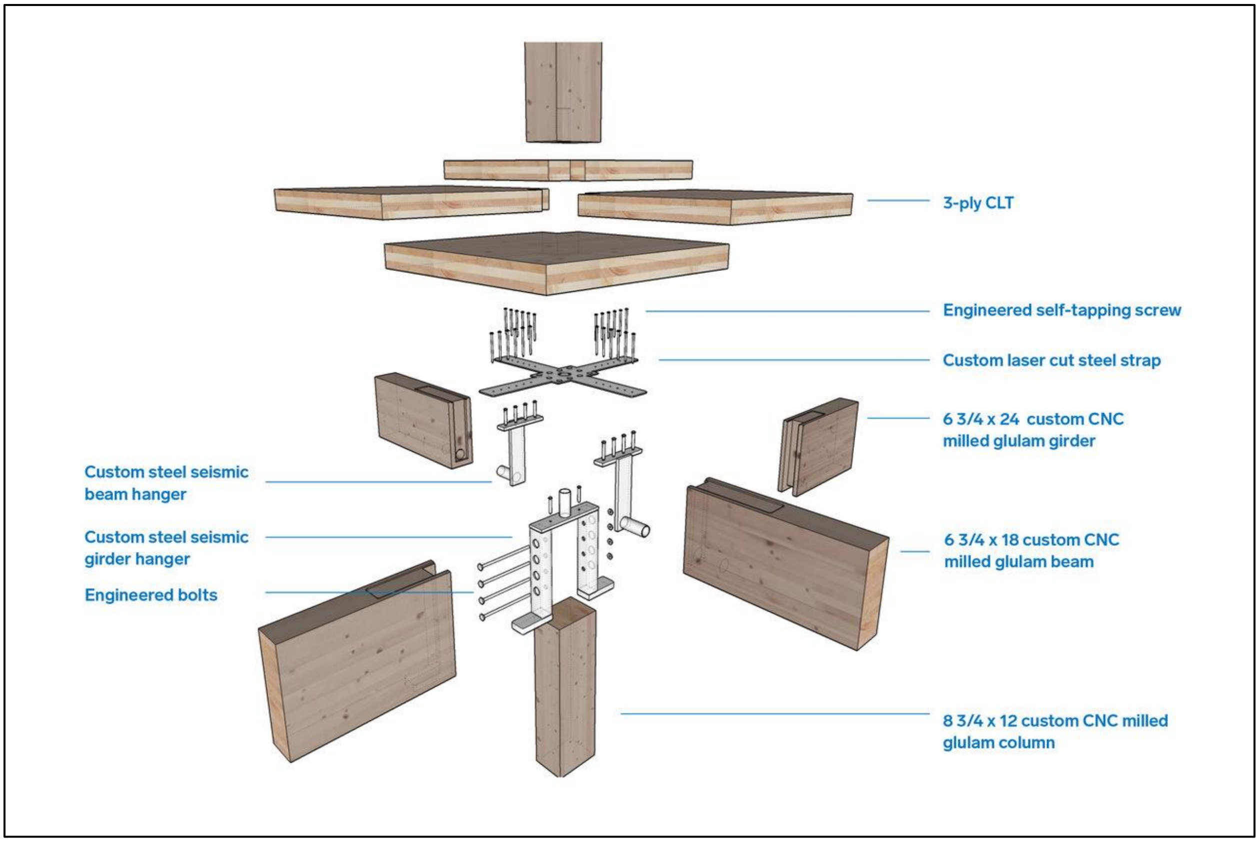

Figure 2.22: Custom Column, Beam, Floor Assembly (Source: Albina Yard, LEVER Architecture, 2020)

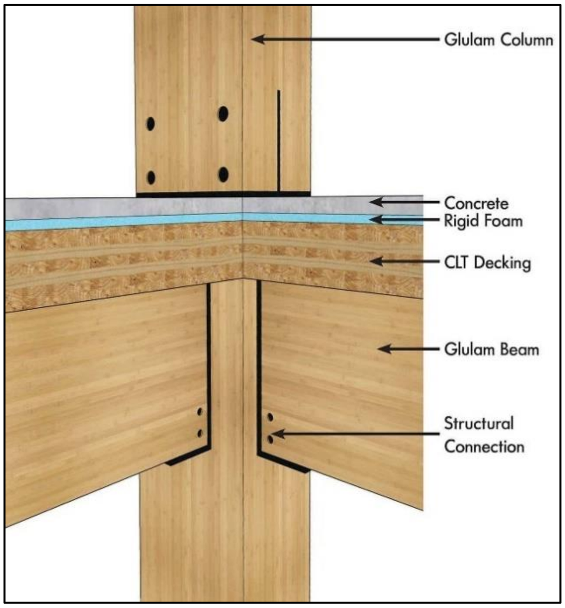

Figure 2.23: Column, Beam, Floor Connection Assembly
(Source: Unknown)

Figure 2.24: Prefabricated Post-to-Beam Connection
(Source: MyTiCon Timber Connectors, 2019)

2.3. COST STUDIES FOR MASS TIMBER VS CONCRETE/STEEL

Cross-laminated timber has received a good response in North American market and is still at its initial stages to be adopted nationwide (Laguarda-Mallo and Espinoza, 2015). There is an ambiguity about costs associated with its adoption and this uncertainty is considered as a major barrier to further use (Laguarda-Mallo and Espinoza, 2014; Laguarda-Mallo and Espinoza, 2015). CLT has a great potential as a construction material due to its benefits for the environment, construction speed, and precision. A study performed by Laguarda-Mallo (2016) found that decisions for structural materials are made at initial design stages and cost was found to be the driving factor. Therefore, it is necessary to perform cost studies for CLT in comparison with traditional concrete and steel construction systems. A few such studies have been performed to illustrate economic performance of CLT.

Laguarda-Mallo (2016) conducted a feasibility and economic performance study on CLT for a performing arts center project and compared it with a concrete and steel construction system. The result showed that CLT had great potential for cost and time savings. This study evaluated five design options which included 1) concrete, structural steel, and light-steel frame, as originally-specified, 2) CLT panels replacing concrete walls and roof, materials from first manufacturer 3) CLT panels replacing concrete walls and roof, materials from second manufacturer, 4) a hybrid of CLT panels, glulam beams, and wood-frame construction, materials from first manufacturer, and 5) a hybrid of CLT panels, glulam beams, and wood-frame construction, materials from second manufacturer. Results showed a cost reduction up to $9 per square feet depending upon manufacturer for the basic CLT option and up to $15 per square feet for the hybrid timber option, depending upon the CLT manufacturer. These results do not include cost reductions due to construction time savings; hence, total construction cost savings could be better when indirect costs are considered. Four-month time savings were observed for the basic CLT option and a sixty-one percent for the hybrid option. The cost breakdowns for all the five options are shown in Table 2.4.

Table 2.4: Summary of Results (Source: Laguarda-Mallo, 2016)

Element	Concrete/ Steel option	CLT options			
		Basic CLT option 1	Basic CLT option 2	Green option 1	Green option 2
	(Concrete walls/roof, steel beams, light-steel frame)	(CLT walls/roof, steel beams, light-steel frame)		(CLT walls/roof, glulam beams, wood-frame)	
Structural Walls	$1,071,680	$624,417	$414,901	$624,417	$414,901
Concrete Slab	$256,416	$256,416	$256,416	$256,416	$256,416
Roof System	$600,975	$427,809	$289,339	$427,809	$289,339
Interior Walls*	$155,304	$155,304	$155,304	$297,666	$297,666
Steel Beams	$506,575	$506,575	$506,575	-	-
Glulam Beams	-	-	-	$29,022	$29,022
Extra CLT Walls	-	-	-	$115,407	$84,977
Extras for CLT**	-	$595,241	$595,241	$654,768	$654,768
TOTAL $	2,590,950	2,565,763	2,217,777	2,405,506	2,027,091
SQFT	40,065	40,065	40,065	40,065	40,065
Cost per sqft	$64	$64	$55	$60	$50

* Interior walls for concrete and basic CLT options are in light-steel frame construction. Interior walls for CLT Green options are in wood-frame construction.

** Extras for CLT includes labor cost and connectors for CLT

Fanella (2018) also performed a cost study to compare cast-in-place reinforced concrete structures with CLT. This study compared costs of the structural frame for a hypothetical 10-story residential building situated in the Pacific Northwest. The CLT option utilized a bearing wall system to resist gravity and lateral loads, whereas the cast-in-place reinforced concrete option used a flat plate system (slabs and columns) to resist gravity loads and a concrete shear wall core to resist lateral loads. The cost of the structural frame for the CLT option was found to be $48 to $56 per gross square feet with an additional $2 to $6 per square feet for acoustical dampening and fire protection. The cost of the structural frame for the concrete option was in the range of $42 to $46 per gross square foot with an additional cost of $1 to $2 per square foot for acoustical dampening. The study showed an incremental cost increase of 16 to 29% for the CLT structural frame option in comparison with the reinforced concrete option. The relative cost information for the structural frame of CLT projects in North America is shown in Figure 2.25.

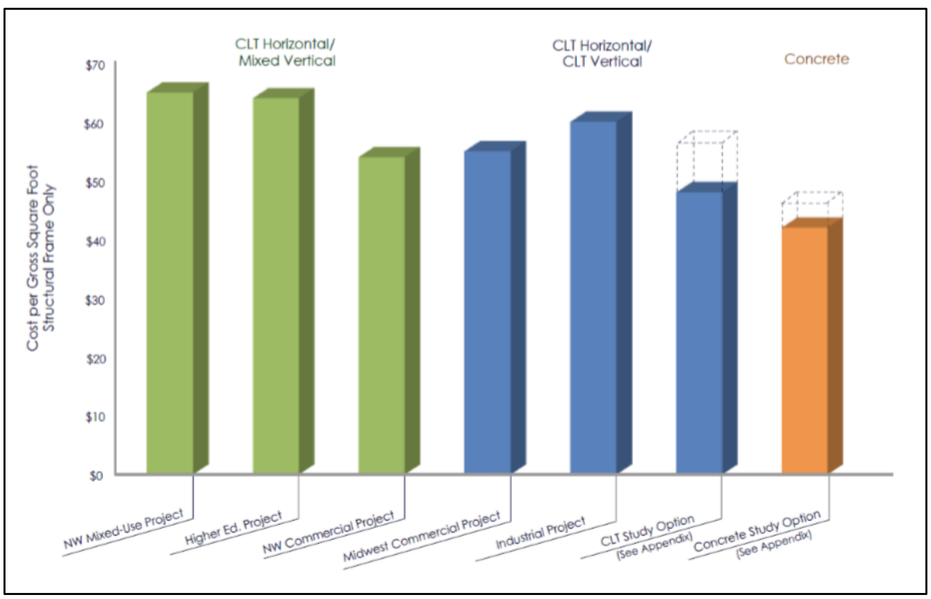

Figure 2.25: Relative Cost Comparison for CLT Projects
(Source: CRSI, 2018)

The previous study comparing mass timber and concrete structural frames did not consider schedule benefits of mass timber. Moreover, only a CLT bearing wall structural system was compared instead of optimizing the construction system. To provide in depth, well rounded perspective of using an optimized mass timber construction system, the most recent feasibility study was produced for a 12-story mixed-use, mass timber tower for Seattle (DLR group, 2018). This detailed study compared a baseline post-tensioned concrete structure with the proposed mass timber tower. The cost calculations are based on 2018-numbers and for an all office-use with street level retail building.

Schedule Analysis

The proposed mass timber design provides a 5-month savings in the total duration compared to the baseline post-tensioned concrete structure (DLR group, 2018). Time to build the structure, lag time to start the exterior, lag time to start interior rough-in trade work, and the finish of the elevator system were outlined as drivers of the project schedule. The mass timber frame saved 13 weeks relative to the post-tensioned concrete frame. These time savings were primarily observed due to prefabricated panels, glulam beams, and glulam columns. Unlike the post tensioned concrete frame, as soon as mass timber members are locked into the lateral system, following trades can instantly start working on the floor below. This minimizes the lag for exterior and interior work start, and in the case of mass timber, it saved a minimum of six weeks of the total project duration. Collectively, this resulted in total of 5-month savings and a total project duration equivalent to 3/4 of the time required for traditional concrete framed structure. Figure 2.26 shows a comparative schedule for the two options from the study.

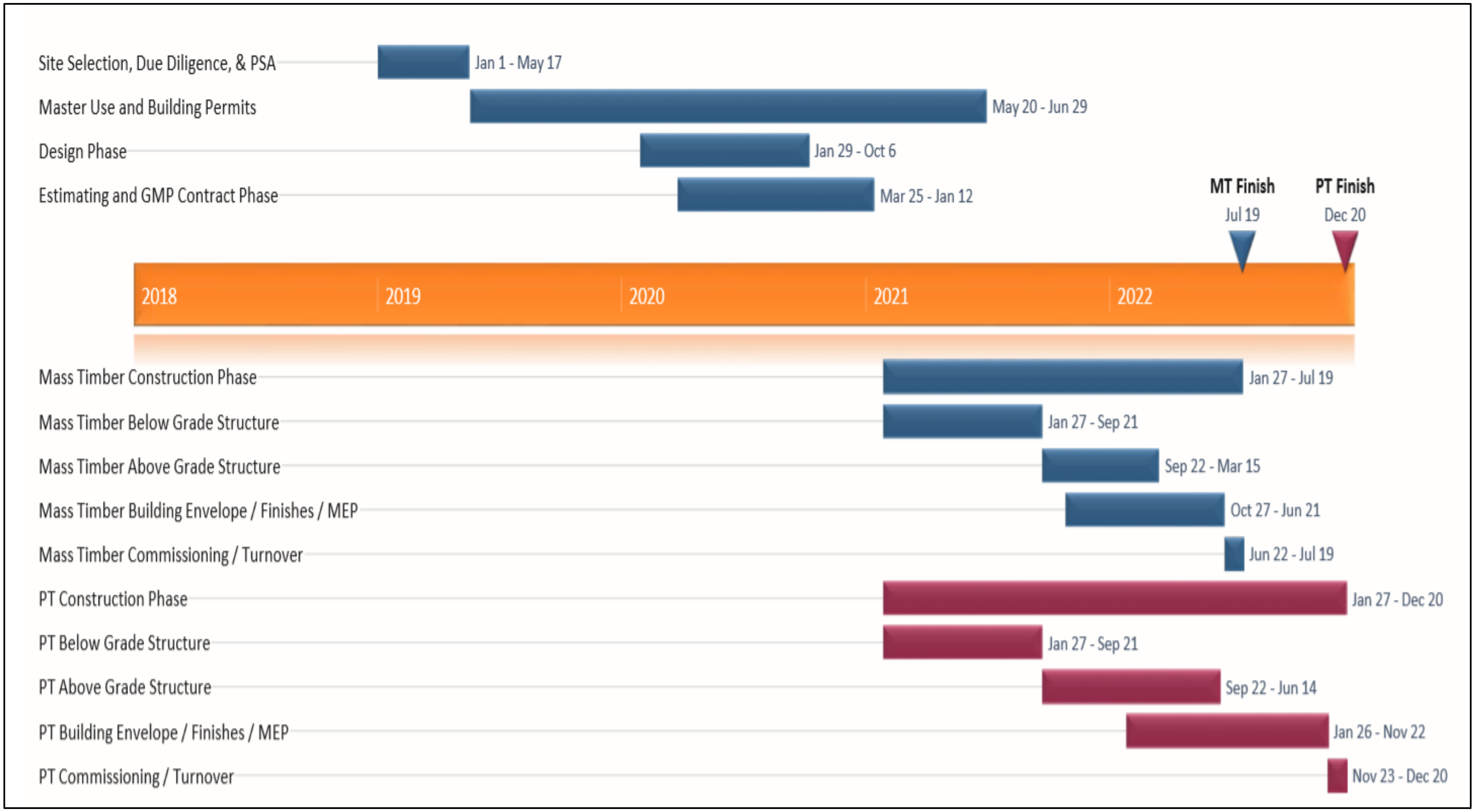

Figure 2.26: Development schedule (Source: DLR Group, 2018)

Cost Analysis

The goal of this study was to compare the complete construction project costs with respect to time for mass timber frame and post tensioned concrete frame design alternates. Hence, this cost analysis is not limited to the costs of the structural frame, it takes into account the benefits of a lighter structural frame, schedule benefits, and aesthetic finishes. After including all the factors, the direct cost of the mass timber system was estimated 2.2% higher than the post-tensioned concrete system. In contrast, the indirect costs were much lower, nearly 20%, relative to the post-tensioned concrete system, which resulted in overall project cost savings of 0.5%. These breakdown costs and costs per square foot are collectively depicted in Table 2.5.

Table 2.5: Cost Comparison Chart and Cost Per Square Foot Chart (Source: DLR Group, 2018)

System	Mass Timber Design	PT Concrete Design	Mass Timber Savings
Direct Cost of Work	$86,997,136	$85,105,091	2.2%
Project Overhead	$ 9,393,750	$11,768,750	-20.2%
Add-Ons	$ 8,387,345	$ 8,429,368	-0.5%
Total	**$104,778,231**	**$105,303,209**	**-0.5%**

Cost Per Square Foot	
Mass Timber Tower:	PT Concrete Tower:
$ 104,778,230	$ 105,303,208
424,175 GSF w/parking	424,175 GSF w/parking
$10,582,601 WSST	$10,635,624 WSST
$115,360,831	**$115,938,832**
$271.97/SF	**$273.33/SF**

2.4. PREDICTIVE COST MODELS

Cost has been observed as a major driver for the selection of structural materials, which typically happens early in the design process (Roos et al., 2010; Bragança et al., 2014; Laguarda-Mallo & Espinoza, 2016). Due to the uncertainty of economic performance of mass timber, it is necessary to understand costs at the conceptual design stage to make informed decisions. In an attempt to find precise costs, numerous studies have been performed to develop predictive cost models. This section reviews relevant such studies.

Lack of information at the earliest conceptual design stages makes estimation a complicated and a vague task. To have better control over the costs of the project, it is important to develop accurate estimates. Therefore, many cost-prediction models have been developed using statistical, probabilistic, and artificial intelligence-based techniques. Initial attempts for cost modeling involved the use of regression analysis (McCaffer, 1975; Skitmore and Patchell, 1990; Trost and Oberlander, 2003; Lowe et al., 2006), as it is a well-known and a powerful statistical tool. The benefit of this method is that it can be useful for examining the variables contributing to the overall costs and to predict costs (Kim et al., 2004). The main disadvantage of using regression analysis is that it does not perform well with multidimensional and multivariate datasets. This is due to the possibility of multicollinearity between large number of independent variables. Despite its limitations, it has been extensively used to develop predictive cost models.

Lowe et al. (2006) presented a regression cost model for prediction of the project costs to assess the feasibility of the project and to provide accurate construction costs to help client make decisions about the project budget and manage the design to meet the budget. They performed forward and backward stepwise analysis on 286 projects to provide six models consisting of eight to fourteen variables each. Initially, he identified forty-one potential independent variables; gross internal floor area, function, duration, mechanical installations, and piling inclusion, were found in all the models due to their dominating linear relationship with costs. He performed forward and backward stepwise regression on log of cost, cost/m^2, and log of cost/m^2 as dependent variables, to produce all six models. The log of the backward cost model was found as the best regression model with an R^2 of 0.661 and mean absolute error percentage of 19.3%, which was comparatively better than previous regression models.

With technology advancements and the arrival of artificial intelligence, case-based reasoning (CBR) and artificial neural networks were identified as useful techniques for accurate cost estimation (Li, 1995; Perera et al., 1998; Bode, 1998). A comparison of these techniques with the traditional statistical methods was a pre-requisite to study benefits and drawbacks. Kim et al. (2004) presented a comparison of regression analysis, neural networks, and case-based reasoning. He examined 530 historical cases, developed cost models and found neural networks as the most accurate technique and case-based reasoning as the second most accurate. Despite the accuracy of neural networks, this method is a black box which does not require representation of any mathematical relationship between the independent variables and cost relationship. Additionally, updating of neural networks was found as time consuming process due to its necessary retraining procedures. Therefore, Kim et al. (2004) found CBR models as the most effective model based on the accuracy, clarity of explanation, and the ease of updating.

Case based reasoning (CBR) is based on finding and retrieving a case stored in a database that is most similar to the test case and applying the solutions of that similar case to the new problem to find the solution and then retain that solution for future use (Kim and Kim, 2010; Zima, 2015). Cases in CBR are represented by attributes; weights are assigned to these attributes and a similar case is the one which carries the maximum weight. Expertise elicitation, intuition, judgement, Analytical Hierarchy Process (AHP), and gradient descent are a few methods which are used to calculate these weights (Arditi and Tokdemir, 1999; Morcous et al. 2002; An et al. 2007). All these methods use human inputs which has the potential to introduce ambiguity and error (Arditi and Tokdemir, 1999; Kim and Kim, 2010).

Kim and Kim (2010) suggested the use of genetic algorithm (GA) to determine weights of attributes with a goal to minimize the construction cost prediction error and the similarity scores of attributes which have a nonlinear relationship with the cost. They developed a cost estimation model for bridge construction projects by collecting data on 585 projects. With the help of this hybrid CBR and GA model, they found mean absolute error of 7.62% with requiring only two attributes required at the early stages; number of lanes and length of the bridge. This model performed better than previously developed models using CBR.

Jin et al. (2012) developed a notable model using CBR and multiple regression analysis (MRA) to predict costs of construction projects at the early stage. He used 41 business facilities and 99 multi-family housing projects with an objective to improve the prediction performance of existing case-based reasoning models. He suggested an MRA based revised CBR model to improve the accuracy of existing models. Results indicated performance improvements in both the models; revised model for business facilities improved by 17.23% and revised model for multi-family housing projects by 4.39%. Therefore, it can be observed that there are ample opportunities to develop cost models with higher accuracy, and with appropriate methods required accuracy can be achieved.

CHAPTER 3 RESEARCH METHODOLOGY

3.1. INTRODUCTION

The background of this study and introduction of the research area is represented in Chapter 1 and Chapter 2 provides a thorough literature review on manufacturing processes, construction process, and a review of cost prediction models. This chapter builds up on the foundation created by both the chapters, by providing research approach and methodology. It further elaborates the objectives and defines the connection and contribution of those chapters to this research.

Objective 1 provides a groundwork for the research by explaining its background and includes literature review and field observations. The Chapter 2 summarizes manufacturing process for CLT and glulam, current state of manufacturers in the United States, construction process, and a review of cost prediction models.

Objective 2 starts building up on the substance, and the information was used to create a spreadsheet with variables impacting construction costs and initiate the process of data collection. Along with the process of data collection, analysis of the methods was started. As a part of methodology selection step, multiple regression, genetic algorithm, artificial neural networks, and case-based reasoning were analyzed. The data collected was used to approach the final selection of the methodology.

Objective 3 develops on the information learnt and variables are analyzed, selected variables were then used to create a regression prediction model. This model is planned for testing as a final step of this research, to check for its ease of use and accuracy.

3.2. RESEARCH OBJECTIVES

3.2.1. Objective 1 – Understand the Current State and Background of Cross-Laminated Timber and Glued Laminated Timber in Construction

The purpose of the groundwork provided by this objective was to understand previous studies in mass timber and provide a clear representation of the current state of this industry. Gathering previous studies and understanding the current state provided a vision towards the need of this research and concrete reasons to study the cost implications of using CLT and glulam in construction. These studies also showed that necessary attention was paid to study prediction tools for concrete and steel structures. On the contrary, mass timber construction lacks the essential attention for such a prediction tool and a limited attention has been paid to study cost implications. This research makes an attempt to study the cost implications of mass timber construction through a prediction tool. Figure 3.1 shows the research outline of methods and outputs with interoperability of objectives.

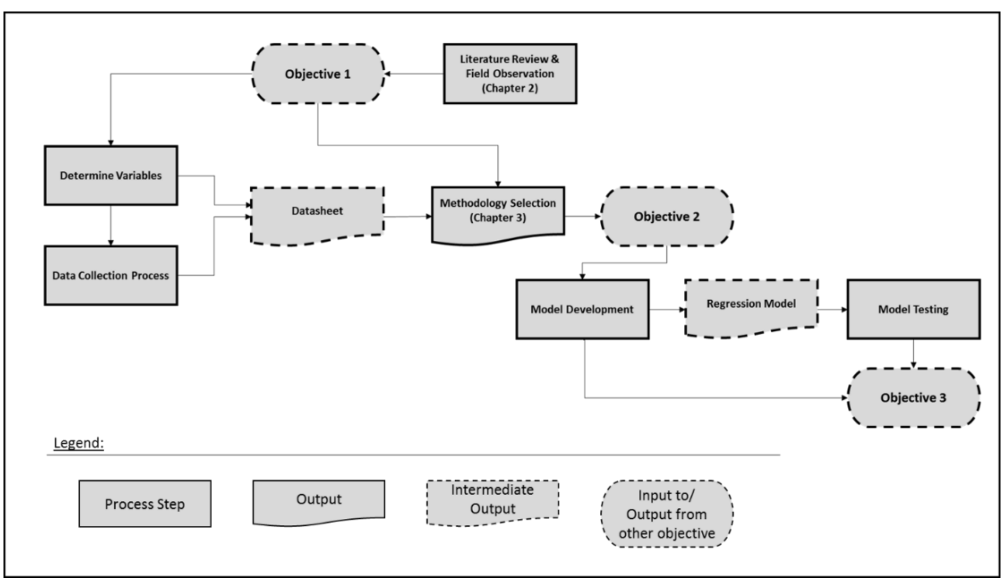

Figure 3.1 Research Outline of Methods and Outputs

The contribution of field observation was to create a deeper understanding of the construction process and to analyze current market conditions. This observation provided familiarity with the current manufacturer capabilities and is represented in Table 2.3 in Chapter 2. Study of cost prediction models provided necessary inputs for technique selection step. Furthermore, literature on various prediction models previously developed for traditional construction projects, provided insight for variables to be determined in Objective 2.

3.2.2. Objective 2 – Build a Foundation for Development of the Predictive Cost Model

The primary purpose of this objective was to create a base required for the development of prediction model which includes three steps; 1) determining variables, 2) data collection, and 3) technique selection. This objective advances on the groundwork performed in Objective 1 by identifying variables and sets a stage for Objective 2 by addressing data collection process and a suitable technique for prediction model development.

3.2.1.1. Objective 2a: Determining Variables

The purpose of this objective was to create a list of study variables contributing to the cost of the construction projects. This list was then planned to serve as an input for data collection to create a spreadsheet, which after data collection served as a datasheet. The list of study variables is represented in Table 3.1 and description of these variables is presented below.

<u>Independent Variables</u>

Due to the lack of previous work in studying cost implications for mass timber construction, some of the independent variables were determined with the help of studies performed for other construction types, with the elicitation of expertise of professionals. The following variables were considered, and their description is as follows:

<u>Gross Area of the System (SF):</u> It is the total floor area covered by the mass timber system. Saxena (2018) and Ji et al. (2011) stated increased duration and material costs as the size of building increases. This variable was selected to reflect this relationship in the prediction model which was also used by Saxena (2018) for his prediction tool. Regardless of the construction system, gross area of the system has an impact to the cost of the project.

<u>Location:</u> An et al. (2007) and Koo et al. (2010) recognized location as one of the necessary variables to estimate the direct cost of a building. The location of the project was selected as an independent variable for two reasons: 1) with a limited number of manufacturers, location of the project can be a critical factor for the material supply and 2) with different zip codes, there are different costs for locally available resources, which is expressed by an equivalent parameter as cost indices.

<u>Number of Wood Stories:</u> This represents the number of timber levels of a building. An et al. (2007), Ji et al. (2011), Koo et al. (2010), Yau et al. (1998), and Saxena (2018) identified number of levels as a prime factor contributing to the costs prediction, which resulted in the selection of this independent quantitative variable.

Table 3.1: List of Variables and their type

		Variable	Type	Category	References
Dependent Variables		Cost of the System	Continuous	Number	NA
Independent Variables	Quantitative Variables	Gross Area of System (SF)	Continuous	Number	An et al. (2007), Ji et al. (2011), Koo et al. (2010)
		Location	Continuous	Cost Indices	An et al. (2007), Koo et al. (2010)
		Number of Stories	Continuous	Number	An et al. (2007), Ji et al. (2011), Koo et al. (2010), Yau et al. (1998)
		General Requirements	Continuous	Number	Carr R.I. (1989)
		Worker hours (WH)	Continuous	Number	Yau et al. (1998), Kaming et al. (1996)
		Equipment hours (WH)	Continuous	Number	Yau et al. (1998)
		Total Material Qty. (CF)	Continuous	Number	Yau et al. (1998), Lee-Hoai et al. (2011), Kaming et al. (1996)
		Labor Cost	Continuous	Number	Yau et al. (1998), Lee-Hoai et al. (2011)
		Material Cost	Continuous	Number	Yau et al. (1998), Lee-Hoai et al. (2011), Kaming et al. (1996)
		Equipment Cost	Continuous	Number	Yau et al. (1998)
	Qualitative Variables	Building Type	Categorical	Office/Multifamily/Civic/Student Residence	Akinsola et al. (1997)
		Construction Type	Categorical	I,II,III,IV,V	NA
		Complexity	Categorical	Low/High/Medium	Kaming et al. (1996)
		CLT Thickness (mm)	Categorical	139, 145, 154, 195, 210	NA
		Number of Ply	Categorical	3, 5, 7	NA
		Level of Design	Categorical	SD, DD, CD	Lee-Hoai et al. (2011), Kaming et al. (1996)
		Gravity System	Categorical	Glulam, Steel, Concrete	Koo et al. (2010), Yau et al. (1998)
		Hoisting Inclusion	Categorical	Yes/No	Abdelmegid et al. (2015), Peng et al. (2018)

General Requirements: General requirements are considered to represent overhead costs of the project. The prime contractor's estimate is the sum of direct costs, project overhead, and general office overhead and profit (Carr, 1989). To characterize the overhead costs, general requirements was represented as a variable in the datasheet.

Worker hours (WH): Labor costs are a direct function and productivity is also a function of worker hours. This variable thus impacts both construction costs and construction time. It is also an equivalent schedule parameter instead of duration of the project. Therefore, this variable was considered to reflect the duration of project at fairly early stages of the construction when data was collected.

Equipment hours (WH): Similar to the worker hours, equipment costs and productivity is also a function of equipment hours, therefore impacts both construction cost and time. To consider its impact and compensate for the duration of the project, this variable was considered as a duration characteristic of the project.

Total Material Quantity (CF): This independent variable, represents the total quantity of CLT and glulam required for the project, expressed in Cubic Feet (CF). This variable was selected as a result of underlying relationship between the material quantities and the total cost of the project. As stated in Chapter 2, Kaming et al. (1997) identified inaccurate quantity takeoff as one of the first three causes of cost overruns, which indirectly represents the underlying relationship between the cost of the project and total material quantity, as total material quantity is a direct derivation of the quantity takeoff.

Labor Cost: Increased labor costs were recognized as a cause of cost overruns by Kaming et al. (1997) and labor resource was considered as one of the few input variables for cost-time prediction model by Lee-Hoai et al. (2011). This represents an indirect connection with the project costs and as a result this variable was considered for this study.

Material Cost: Material cost represents costs associated with CLT and glulam, it does not include costs for steel connections. In a study by Kaming et al. (1997), cost overruns were observed as a consequence of increased material costs and provided a hint of the relationship with the project costs. This quantitative variable was considered for the same purpose.

Equipment Cost: Yau et al. (1998) considered equipment resources as a part representation of the project costs. This variable was considered as it directly impacts the project direct costs (Carr, 1989). Regardless of the construction materials equipment costs are major contributors to the project costs, even though different construction materials demand different crane specifications and requirements.

Building Type: Building type is a formal representation of the primary purpose of the building, structure or a part of it. They were classified into Office, Civic, Educational, Mixed-Use, and student residence. Akinsola et al. (1997) identified project type as a significant influencing parameter to the total value of changes to design after awarding the contract. Which illustrates an indirect impact of the project type to the cost of the project regardless of the construction material used for the project.

Construction Type: Per IBC (2018), five types of construction are established in which each building must be categorized. Each construction type is associated with fire resistance rating requirements for different building elements and is represented in Table 3.2. Fire rating of a wood member is a function char depth (White and Nordheim 1992), which decides the size of members in the building. With different sizes of a member, costs might differ, which reflects the need for selection of this variable.

Table 3.2 Fire- resistance for rating requirements for building elements (hours), Type III, IV, and V
(Source: IBC Section 602, Table 601 and Section 2304.11, McLain R. and Breneman S. (2019))

Construction Type	III-A	III-B	IV	V-A	V-
Exterior wall materials	FRTW	FRTW	FRTW or CLT	Any wood including mass timber	Any wood including mass timber
Exterior bearing wall FRR	2-hour	2-hour	2-hour	1-hour	0-hour
Interior framing materials	Any wood including mass timber	Any wood including mass timber	Heavy timber including mass timber	Any wood including mass timber	Any wood including mass timber
Primary frame, floor & roof construction FRR	1-hour	0-hour	HT	1-hour	0-hour

Complexity: Kaming et al. (1997) identified complexity of the project as one of the most important factors for the cost overruns of the project. As use of mass timber in construction is a fairly new concept in the US, lack of experience and familiarity with construction methods increases the complexity of projects. Kaming et al. (1997) had identified uncertainties with construction methods, technological knowledge and historical data as significant factors for cost overruns and delays.

CLT Thickness: It represents the thickness of CLT panels used in the project. Different panel thicknesses can be used in combinations to optimize costs of the project. With different thicknesses cost of the panels also vary, which ultimately contributes to the costs of the project.

Number of Ply: CLT's are available as 3, 5, 7, and 9 ply combinations. With the increase in ply number, purpose of CLT changes and costs also change with application. These changes in the costs affect the material costs of the project.

Level of Design: This independent categorical variable represents the design stage of the project, which is important as the data was received mostly from preconstruction stage. It is categorized into Schematic Design, Design Development, and Construction Documents. Kaming et al. (1997) identified design changes as one of the first five causes of delays of the project, which states the need to consider it for this study.

Gravity System: Gravity System defines the structural system to transfer the gravity load of the structure to the foundation. Ji et al. (2011) and Saxena (2018) stated with different structural frames productivity will vary which will result in different project costs. Specific to this study all the projects have a Glulam gravity system, this variable served as a filter for the data selection.

Hoisting: With the increase in the demand of economical, resilient, and safe construction of facilities in the shortest possible time, there is often a need to transport materials in a short period of time. Cranes play a vital role in transporting materials and components to the labors onsite and are considered as a costly resource to use in the projects (Peng et al., 2018). Selection of cranes according to the project needs and its efficiency are considered as critical factors significantly impacting the project costs and time (Peng et al.,2018) and (Abdelmegid et al., 2015).

Dependent Variables

Cost of the System: The cost required to build the mass timber system is expressed as "cost of the system" and considered as a dependent variable for this study. The cost is specifically related to the Mass Timber System and does not include costs related to the other parts of structure containing concrete, steel or any other material type. With a few Mass Timber buildings in the United States, small sample size was expected, and restricting costs related to the frame helped in considering different building types to be a similar timber gravity frame. This cost does not include contractor's fees, insurance, taxes, and contingency.

3.2.1.2. Objective 2b: Data Collection

To achieve the goal of creating a prediction model for mass timber construction projects, data collection of such projects was necessary. Construction projects in design, under construction/completed using mass timber construction technique was considered as a population for this study. As this technique is comparatively new in the United States, there are a limited number of projects. As the projects are limited and the market is new, there are a very few contractors having expertise in this area which results into a barrier for sharing such data.

During the data collection process, the population of mass timber projects using CLT and glulam frame was researched through public resources. A total of 150 projects either under construction or already built and 281 projects in design were together considered as the population from which the study sample would be drawn (Woodworks, 2020). As there is an uncertainty about the population characteristics and a very little information is available about the population, Slovin's formula (Stephanie, 2003) was used to calculate the sample size and to better understand the implications on the expected accuracy. Slovin's formula is as follows;

$$n = N / (1 + Ne^2)$$

n = Sample size

N = Total population

e = Margin of error

According to AbouRizk et al. (2002); conceptual estimates are generally accepted within the error range of +/- 30%. Furthermore, Jong (1992) provided 20% error as a criterion to decide validity of a conceptual estimate. From a study performed by Langat (2018), it was evident that 95% confidence level generate statistically significant results. Therefore, to calculate the sample size Slovin's formula was used with the acceptable 20% margin of error. A sample size of twenty-three was required to achieve the acceptable level of accuracy.

The data was obtained from various sources such as general contractors in the United States. After data collection process, a total of 28 projects were obtained which accounts for about six and a half percent of the population and exceeds the required number of samples. As it is an emerging market, there are a limited number of completed projects, but comparatively significant number of projects in the preconstruction stage. As a result, the data collected was from projects in preconstruction stage and projects under construction. Figure 3.2 shows the state wise distribution of these projects across United States.

The process of data collection was initiated by reaching out to professionals in this industry, followed by that an Excel spreadsheet was circulated as represented in Table 3.3. After receiving these Excel spreadsheets, all the data was collated in one sheet and was referred as a "datasheet". This datasheet was then used for technique selection as a final step of Objective 2. Furthermore, this datasheet also provided required input for Objective 3.

3.2.1.3. Objective 2c: Methodology Selection

After obtaining the datasheet, the immediate goal was to commence data analysis, but with the sample size obtained and the number of variables, the prediction methods reviewed in Chapter 2 offered a challenge for the accuracy and suitability of the methods for the model. In order to overcome that challenge, experts were consulted and principal component regression (PCR) was suggested as a better suited methodology for the study. This objective explores some background for selecting PCR and how can it be performed to achieve the goals of this study.

Specific to this study, the sample size obtained was relatively small given that mass timber construction is relatively a new approach in the US construction industry. The initial approach was to use multiple regression for this study, but for using multiple regression, a large dataset is required and there should not be any collinearity among the independent variables (Chan, 2005). With a handful of observations and the possibility of multicollinearity among the independent variables, there existed a need to find a better suited approach for the development of the model. With 18 independent variables, the statistical power was also a factor of consideration (Gnanadesikan, 1997), thus there was also a need to reduce the dimensionality of the data.

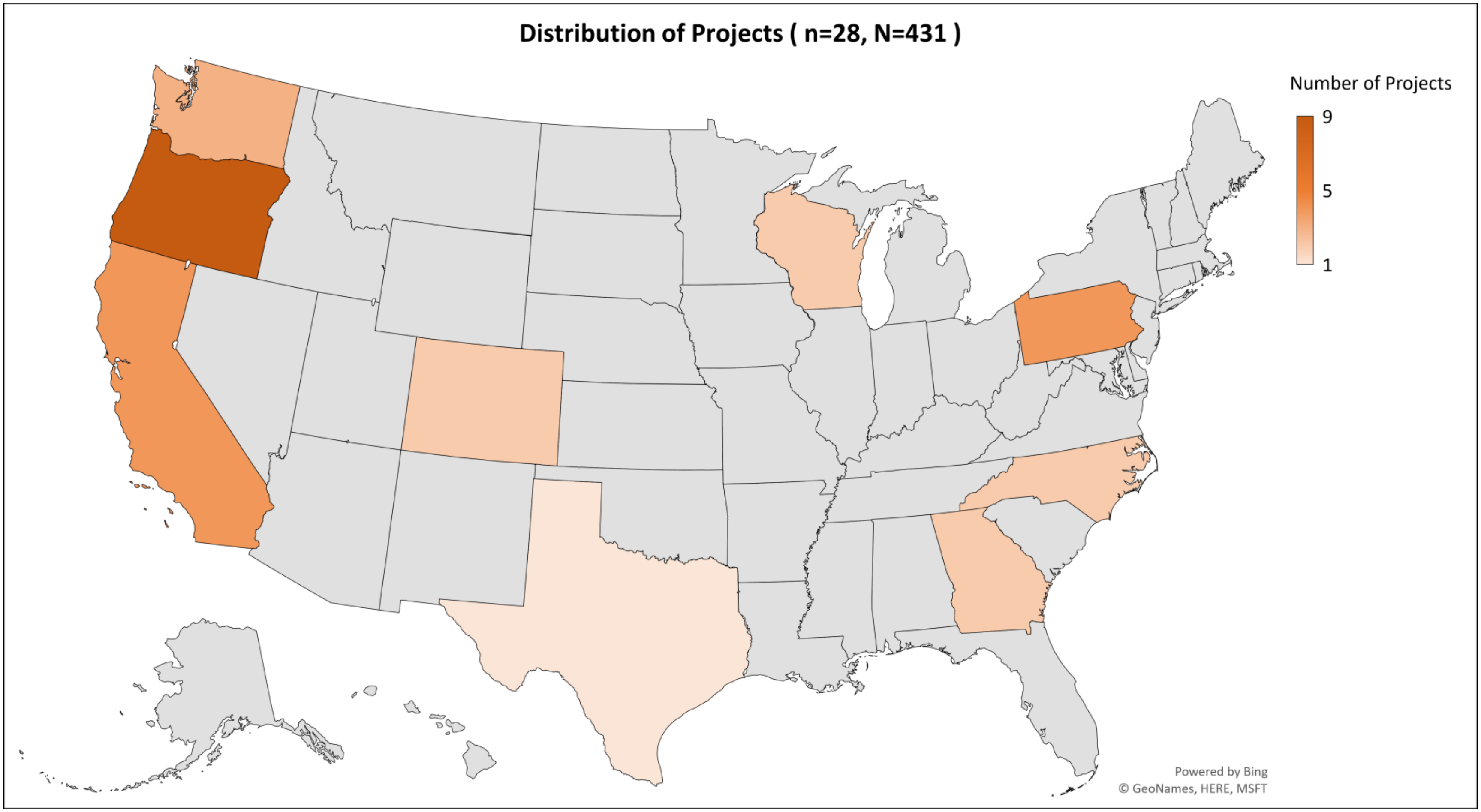

Figure 3.2 Distribution of projects of collected data

Table 3.3 Spreadsheet Circulated During Data Collection Process

Project Name	Project 1	Project 2	Project 3	Project 4	Project 5	Project 6	Project 7	Project 8	Project 9	Project 10
Cost of the System										
Worker Hours (WH)										
Equipment Hours (WH)										
Building Type										
Stories of Wood										
Area (SF)										
Construction Type										
CLT Thickness										
Number of Ply										
Level of Design										
Total Material Qty (CF)										
Labor Cost										
Material Cost										
Equipment Cost										
General Requirements										
Year										
State										
City										
County										
Zip Code										
Complexity										
Gravity System										
Hoisting Included										

Principal component regression (PCR) was considered afterwards due to its advantages of reducing dimensionality and handling multicollinearity of the data. The prerequisite for this regression is to perform principal component analysis (PCA), which extracts a small number of principal components that explain the most variation in the data. The next stage is to perform regression on these selected principal components against the dependent variable. As PCA generates a new set of variables (principal components), which are uncorrelated, and dimensions are of manageable size, it was a better suited approach for this study.

Principal Component Analysis (PCA)

PCA is a non-parametric method which in many ways formulates the basis of multivariate data analysis (Wold et al., 1987). PCA is used frequently in nearly all the forms of analysis, which include but not limited to neuroscience, chemistry, geology, engineering, and computer graphics. It is known by various names in different fields; singular value decomposition (SVD), Karhunen-Loéve expansion, and Hotelling transformation are a few widely known alternate names (Golub et al., 1983, Mandel, 1982, Karhunen, 1947, and Loéve, 1948). With the help of PCA, complex datasets containing large number of correlated variables can be reduced to a manageable dimension. It helps in extracting relevant information from confusing datasets and sometimes reveal the basic hidden patterns that underlie it (Shlens, 2005). The key idea of PCA is to retain the maximum variation possible present in the data while reducing the dimensionality of a dataset containing large number of interrelated variables (Joliffe, 1986). It attains this goal by transforming the original variables to a new set of variables, known as the principal components

(PCs). These principal components are uncorrelated and ordered in a way that first few components show the maximum variation in the data set.

History of PCA

Pearson (1901) formulated PCA in statistics and was more concerned with finding the lines and planes of closest fit to the set of points in *p*-dimensional space. Additionally, the geometric optimization problems he considered also lead to principal components (Joliffe, 1986). After Pearson until Hotelling, according to Joliffe (1986) there was a very little relevant material was published. Hotelling (1933) then further developed PCA and provided his paper in two parts. The first, most important part, stated the problem in detail and presented a method of analysis. Furthermore, he showed its geometric meaning and illustrated methods of solution. The second part discussed certain derivative problems which included; determination of principal components for individuals, iterative solution of normal equations, tests as samples of a larger aggregate of tests, principal components with perfect weighting and the "sand" and "cobblestone" theories of the mind. Hotelling's motivation was that there is a possibility that a smaller number of 'fundamental set of variables… which determine the values' of the original *p* variables. He states that such variables have been called "factors" in the psychological literature and introduces "components" as a new term to avoid confusion. He chose these components such that they maximize their successive contributions to the total of the variances of the original variables and called them as principal components. Then, the method to find these principal components was called as "method of principal components". Hotelling (1933) in his paper also showed a way to find components using power method and then provided an accelerated version of the same method (Hotelling, 1936). Girshick (1936) also provided some alternative ways to determine PCs and presented an idea that sample PCs are a maximum likelihood estimate of the population PCs. Before Rao (1964) presented numerous ideas concerning uses, interpretations, and extensions of PCA, there was a very limited amount of work that was published. Gower (1966) then provided links between PCA and other statistical methods and some geometric insights. After Gower (1966), the practical side of the PCA was demonstrated by Jeffers (1967) with the help of case studies which showed the use of PCA more than just a simple dimension reduction tool.

Advantages of using PCA

The important points of consideration while analyzing the obtained multivariate data were; 1) Large number of dimensions of the data, 2) Interrelationships of the independent variables, and 3) Small sample size. These challenges are explained as follows,

Dimensionality Reduction

The basic problem with the analysis of a large multidimensional dataset is that it is too big to handle, visualize, understand, and interpret. Also, while dealing with multidimensions, it is difficult to simply visualize and catch the hidden patterns in the data, because of the nature of its size. On the contrary, with the two-dimensional or three-dimensional data, it is comparably interpretable and understandable without putting in too much efforts as we can visualize that data. Therefore, we use different techniques with a vision to reduce the dimensions and size of the data. But with the reduction in the size of the data, there is a loss of information and it is necessary to protect the important information in the data especially when small size is involved.

Multicollinearity

With multidimensional data, there is almost always a possibility of interrelationships between the independent variables at hand. It is important to understand those relationships in order to then use such data for prediction models. Specific to this study, the data was planned to be used further for regression and the elementary assumption of performing regression is that the independent variables should be uncorrelated. In such cases it is important to identify the interrelationships with a goal to either discard them or to prepare a strategy to manage them.

Small Sample Size

With small sample size obtained it was very important to protect all the information necessary and to use all of it for the predictions. With dimensionality reduction and tacking multicollinearity there was a possibility of information loss and was an important criterion and a challenge while selection of a method. PCA showed strength in all these areas and demonstrated suitability for this study as follows:

The issues with dimensionality reduction in a multivariate data is finding balance between achieving simplicity for understanding, visualization, and interpretation, on one side, and minimal loss of information for adequate representation on the counter side. PCA is the most widely used multivariate method for linear dimension reduction (Gnanadesikan, 1997). The central idea of this technique is to reduce the dimensions by providing principal components in order that first few components explain maximum variation in the data. As a result, loss of information can be managed by selecting appropriate number of components that retain most of the information in the data, given that the sample size was small retaining most of the information was necessary.

PCA generates these components in a way that they are uncorrelated with each other regardless of the multicollinearity in the data. Which was the most crucial factor for this study, as the goal was to generate a prediction model using regression and one of the major difficulties with the usual least squares estimators is the multicollinearity. PCA is helpful in dimensionality reduction to visualize, interpret, and understand the data without much loss of information. To tackle these issues PCA was selected in order to obtain maximum possible accuracy from the data.

Limitations of PCA

An important limitation of using PCA is that it is only suitable for continuous data set (Kolenikov et al., 2004). As mentioned by Anderson (2003), Hotelling (1933), and Mardia et al. (1980), it was developed for the samples from the multivariate normal distribution and most of the theoretical methods were derived with the normality assumption. Discrete dataset can be used in PCA by implementing Filmer-Pritchett PCA procedure of generating dummy variables. But there are a variety of implications of using a discrete dataset in the standard PCA and the problems related to discrete dataset are listed below;

Firstly, with using discrete dataset, normality assumptions of PCA are violated as discrete data do not follow a density function. Furthermore, even with finite range, the discrete data has high kurtosis and skewness, when the majority of data points are concentrated in a single category. PCA only addresses approximating the real data with the normal distribution of the same mean and covariance matrix (Kolenikov et al., 2004).

Second, using discrete variables in the dataset may not mirror the true covariances or correlations of the underlying variables. According to an example provided by Kolenikov et al. (2004), these discrete variables tend to be biased towards zero. If the covariance structure of the observed variables is not consistent with the theoretical model, then the estimated principal component weights will be inconsistent and will be biased (Kolenikov et al., 2004).

Third, PCA using discrete variables reduces the share of variance that is represented by the first few components. These distortions are greater with categories less than five per categorical variable, high skewness and kurtosis, and opposite skewness of different categorical variables (Kolenikov et al., 2004).

Steps to perform PCA

This section will explain the steps required to perform a PCA on a set of data. The analysis is can be simplified and performed by following the six steps described below.

Step 1: Start with the data for n observations on p variables

The first step to perform any statistical analysis is to collect a dataset, similarly, to perform PCA a dataset is needed. The dataset for this study had p = 18 variables and n = 28 observations. Therefore, the matrix obtained was a 28 x 18 matrix. Out of these eighteen variables, only ten variables were continuous and as PCA is an orthogonal linear transformation, discrete variables were discarded. This data contained variables with various units and have different percentage of contributions to the dependent variables. In order to take care of these different scales, the next step standardizes the data so that each one of them contributes equally to the analysis.

Step 2: Form a matrix of size n x p with deviations from mean for each of the variables

This step is performed for the purpose to prevent variables with large ranges from dominating over variables with small ranges. For example, the stories of wood ranges from 1-16, but the cost ranges from \$1,000,000 to \$20,000,000, so the cost variable will dominate over the stories of wood. To account for this, standardization is performed to transform the data to comparable scales.

Mathematically, standardization can be performed by subtracting the mean and dividing by the standard deviation of each value of each variable.

$$\textit{Transformed Data Point} = \frac{Original\ Data\ Point - Mean}{standard\ deviation}$$

After performing this step, all the variables will be transformed to the same scale and differences in the units and ranges will not be an issue.

Step 3: Calculate the covariance matrix (p x p)

Covariance is nothing but the degree to which two variables are linearly associated, in other words it is the measure of how changes in one variable are associated with changes in a second variable. To recall, it is always measured between two variables and for a n-dimensional dataset, $\frac{n!}{(n-2)! * 2}$ different covariances can be calculated (Smith, 2002). The goal of this step is to understand the relationship between variables and find any redundant information if they are highly uncorrelated. So, in order to identify these correlations, we compute the covariance matrix. Covariance matrix for a dataset with ‘n’ dimensions can be defined as follows,

$$C^{nxn} = (C_{i,j}, C_{i,j} = \mathrm{cov}\ (Dim_i, Dim_j)),$$

Where,

C^{nxn} is a covariance matrix with n rows and n columns

Dim_x is the x^{th} dimension

In summary, the equation above states that for a n-dimensional dataset, covariance matrix has *n* rows and *n* columns.

If the covariance is positive, then both the variables are correlated and if it is negative, both the variables are uncorrelated. Also, covariance matrix is a symmetrical square matrix as cov (a,b) = cov (b,a).

For the obtained dataset, ten dimensions were selected, and a 10 x 10 covariance matrix was obtained.

Step 4: Calculate the eigenvalues and eigenvectors of the covariance matrix

This is the step where the concept of dimension reduction is initiated. As covariance matrix is a square matrix, eigenvectors and eigenvalues can be calculated to determine the principal components. The significance of computing eigenvectors is to extract lines that characterize the data and eigenvalues is to determine the lines showing maximum variation. Each eigenvector corresponds to a principal component and each eigenvalue corresponds to the information captured by that component.

Step 5: Choose principal components and form a feature vector

As explained in section 3.2.2.3, principal components are new variables which are uncorrelated and first few principal components show most of the information. So, with ten-dimensional data, ten principal components can be obtained, but PCA will try to represent maximum possible information in the first component, then maximum remaining information in second and so on. In this step, eigenvectors are arranged in the order of highest to lowest eigenvalues to represent the eigenvectors corresponding to the principal components capturing most of the information. Out of these principal components, first few components showing maximum information are selected. Typically, the principal components are selected with Rencher's (1995) Criterion which extracts PCs whose eigen-values are more than the average of the eigenvalues.

$$(\sum_{i=1}^{n} \lambda i)/p$$

Where,

p: the number of principal components extracted from the data

λ: the eigenvalue of component i

After selecting the eigenvectors corresponding to the principal components in order of their significance, feature vector is formed. A feature vector is simply a matrix that represent the selected eigenvectors as columns of the matrix. This matrix solidifies the dimensionality reduction process, as the feature vector represents a few selected components, the final dataset will have only selected number of dimensions with minimal loss of information. So, for a ten-dimensional data, if we select only first four eigenvectors then we will have a four-dimensional dataset as an outcome of PCA.

Step 6: Derive the new data set

This is the final step of PCA, the goal is to transform the data from original axes to the axes represented by principal components using feature vector formed with the help of eigen vectors. Until step 5, apart from standardization, there were no changes to the data, as a result the data points were with respect to the original axes. In order to analyze data points with respect to new components, they are transformed into new data points which are oriented with respect to new trend lines. Now, with the help of these new values of data points, we can interpret that how those points arranged with respect to new components. This transformation is performed as follows,

$$\textit{Final Data} = (\textit{Feature Vector})^{T} \times (\textit{Standardized Original Data})^{T}$$

This new dataset generated by PCA was used to perform regression against the dependent variable.

3.2.3. Objective 3: Development and Testing of the Model

The primary goal of Objective 2 was to build a foundation for the model by determining variables, collecting data, and selecting methodology. This foundation was a necessary input for the next objective so as to achieve the goal of developing the model. To achieve the same goal, this objective was subdivided into two objectives; 1) development of the predictive cost model and 2) testing of the model.

3.2.3.1. *Objective 3a: Development of the predictive cost model*

As discussed in Objective 2c, Principal Component Regression (PCR) was selected to develop the cost prediction model. This objective expands on that process to develop the model and the linkage between PCA and PCR is represented in figure 3.4. It also elaborates the stepwise procedure for implementing PCR.

Principal Component Regression (PCR)

The problem with using multiple regression is multicollinearity (Gunst, 1983), which states some kind of a linear relationship between independent variables. Due to the multicollinearity, inflated variances of the coefficients estimated using regression can occur. These inflated variances could result into unstable and potentially misleading regression equation. To overcome this issue of multicollinearity, well known approach is PCR (Joliffe, 1986), which uses principal components of the predictor variables in place of predictor variables itself. The reasoning behind using PCs instead of predictor variables is that PCs are uncorrelated, so there are no multicollinearities and the calculations can also be simplified by reducing dimensionality. Figure 3.3 comparing multiple linear regression with PCR and showing dimensionality reduction in PCR.

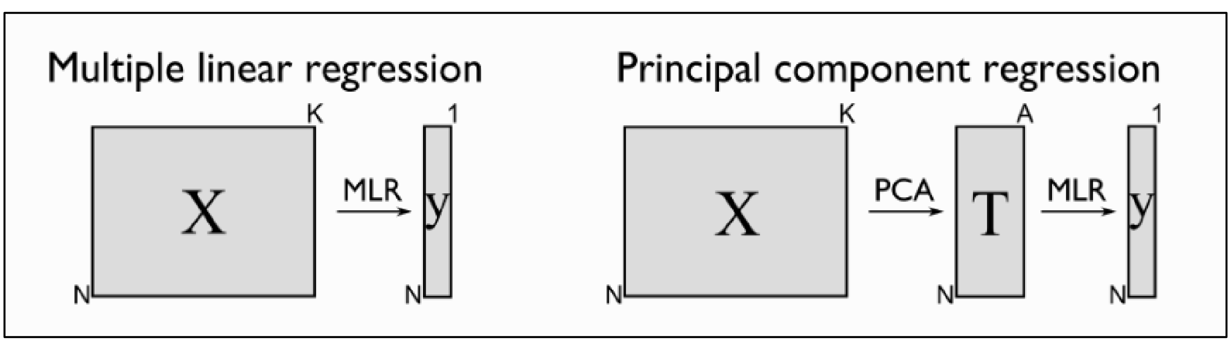

Figure 3.3: Dimensionality Reduction in PCR (Source: Kevin Dunn, 2019)

Where,

X : It is a N X K matrix

N: Number of observations

K: Number of variables

Y: Dependent variable represented as N X 1 matrix

T : It is a feature vector represented as a N X A matrix

A : Number of selected principal components

Stepwise procedure to perform PCR

This subsection will elaborate steps involved in the process of PCR.

Step 1: Centralization of dependent variable

Before performing PCR, the data is pre-processed, and it involves centering the dependent variables. This step is very crucial as PCR uses PCA for the independent variables and PCA is sensitive to the centering of the data. The following equation can be used for centering the dependent variables,

$$y_i = Y_i - \bar{y}$$

Where,

y_i : Centered value of i^{th} observation of y

y : Dependent variable

Y_i : Original value of i^{th} observation of y

$\bar{y}$: Mean of a dependent variable

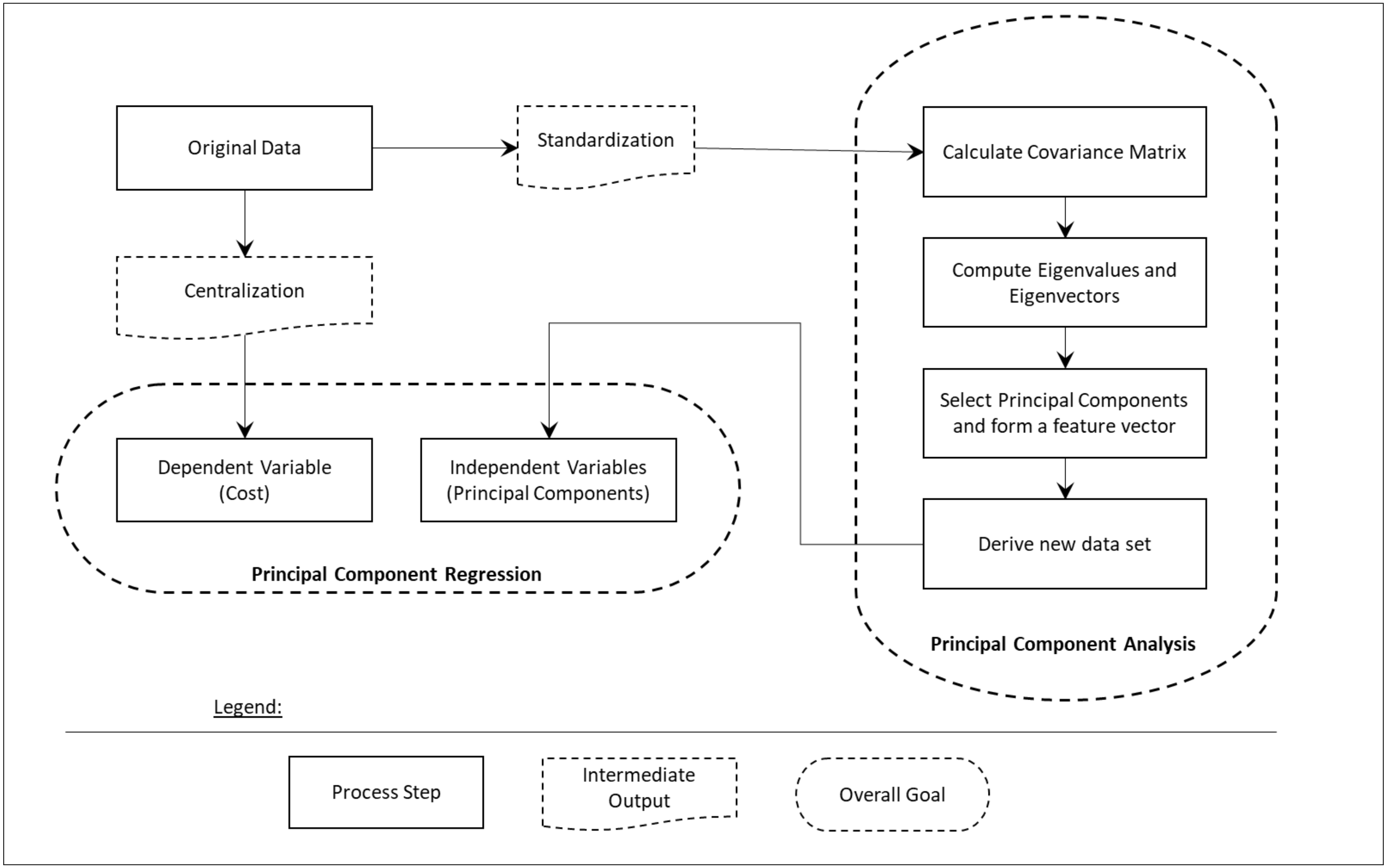

Figure 3.4: Principal Component Regression Process Diagram

Step 2: Perform PCA

After centering the data, the next step is to perform PCA and to obtain the uncorrelated principal components. To perform this step, it is important to understand PCA and the procedure to perform PCA is mentioned in the section 3.2.2.3. As PCA is performed and transformed data is obtained, as mentioned in the step 6, this data then provides an input for the regression in the next steps. This transformed data serves as a new set of independent variables and are then regressed against the dependent variable.

Step 3: Perform Linear Regression on selected components against dependent variable

This is the final step to perform PCR and it requires input from PCA and centered data as shown in figure 3.3. As the goal to perform regression is to find the best estimator of the regression coefficients, this step focuses on finding those and building an equation. This equation can be used afterwards to predict the values of dependent variables. The regression model can be given as below,

$$Y = ZB$$

Where,

Y is n x 1 vector of n observations of the centered dependent variable

Z is n x p matrix of n values of transformed data of p variables

B is a p x 1 vector of unknown parameters

With the help of this equation the regression coefficients are identified, and the regression model can be developed.

3.2.3.2. Objective 3b: Testing of the predictive cost model

The aim of this objective is to implement the model on a project and to check the feasibility of the model. Furthermore, to see the accuracy of the results of the model and to check the usefulness along with the applicability of the model. In order to achieve so, a new observation will be collected and then this observation will be used as an input for the cost prediction. The results obtained using this model, will then be compared with the actual cost of the project to check the accuracy and applicability of the model. This step will help researcher in providing future research insights and discussion on the results.

3.3. CHAPTER SUMMARY

This chapter is focused on developing a foundation for the cost prediction model by determining variables and selecting a suitable prediction method. To achieve that goal, this chapter provides discussion on the variables determined. It also provides a detailed review on the selected prediction method and its suitability for this study. This piece of the study breaks down all three objectives outlined in the first chapter and represent how they are interconnected. As a result, principal component analysis was selected for data analysis and principal component regression was considered as the most suited method to develop the cost prediction model. Chapter 4 will focus on data analysis and the developed cost prediction model results.

CHAPTER 4 : COST MODEL DEVELOPMENT

4.1. INTRODUCTION

The previous chapter focused on elaborating the research outline and research methodology to be used. It also provided in depth understanding of the objectives and the discussion about selected variables. In summary, Chapter 3 explained the design phase of this study and Chapter 4 will explain the execution phase by implementing the selected methodology for the cost model development.

This chapter will start with the analysis of the data and cleaning the data. The data will be visualized with the help of the univariate box plots. Following this, to tackle the multicollinearity and reduce the dimensions, PCA will be performed to find the principal components which are perpendicular to each other and explain most of the variation in the data. These principal components will then be regressed against the dependent variable to produce the regression model. Analysis with the help of PCA and regression model using PCR will be shown in this chapter and start building up on disadvantages of PCA specific to this study.

After this, an alternative method to develop a cost prediction model which will be user friendly and more lucid to understand and improve will be introduced. Partial-Least Squares Regression (PLSR) will be explained as a suitable method for this study and provide the final developed model. The following chapter will conclude with results and discussion.

4.2. DATA VISUALIZATION AND ANALYSIS

In the data collection process, a total of 28 observations were obtained and for all these observations, a total of eighteen variables were collected, resulting in a relatively complex dataset to clean and visualize. This section focuses mainly on this analysis and visualizes data. Chapter 3 contained a detailed discussion on the determined variables and the list is represented in Table 3.1. As shown in Table 3.1, there are ten quantitative variables and eight qualitative variables, and given the limitation of PCA as explained in section 3.2.2.3, that it is only suitable for continuous data set (Kolenikov et al., 2004), ten quantitative variables were selected for further model development. Out of these ten quantitative variables, location was also filtered as there are no specific cost indices for mass timber construction. These nine variables; gross area of the system (SF), number of stories, general requirements, worker hours, equipment hours, total material quantity (CF), labor cost, material cost, equipment cost were then analyzed and visualized.

To visualize and clean the data, univariate box plots were plotted with the help of XLSTAT software and the results are shown in following figures.

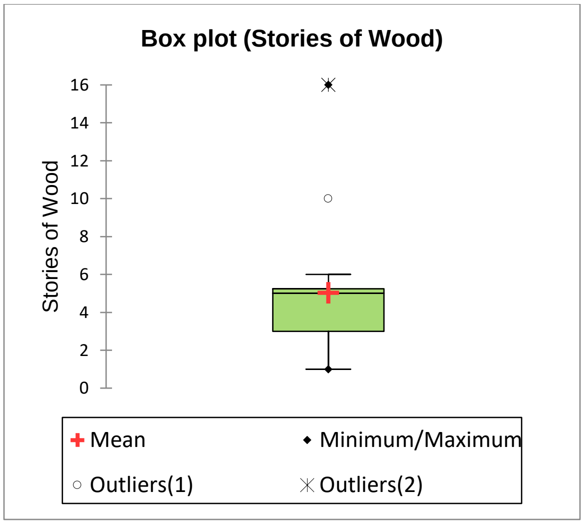

Figure 4.1: Box plot for stories of wood

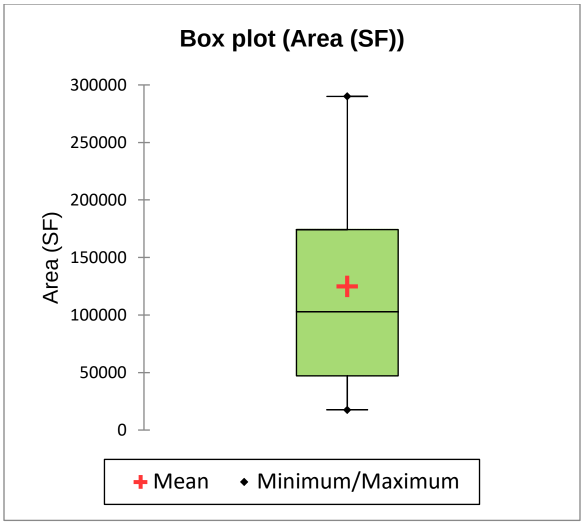

Figure 4.2: Box plot for gross area of the system (SF)

As shown in the figures 4.1 and 4.2 above, there are a total three outliers in the stories of wood and no outliers in the gross area of the system (SF).

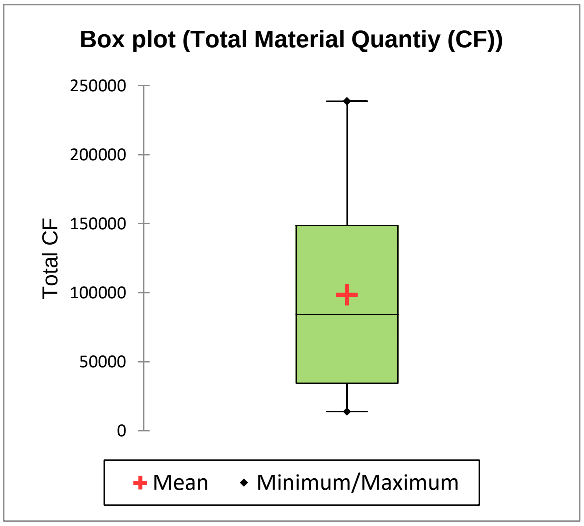

Figure 4.3: Box plot for total material quantity (CF)

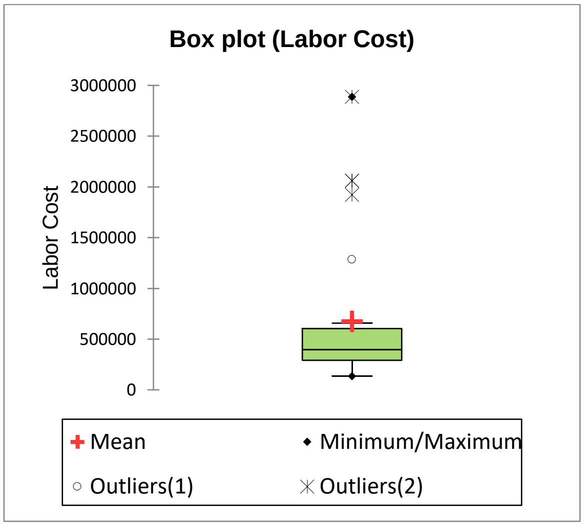

Figure 4.4: Box plot for labor cost

Total material quantity (CF) has no outliers as illustrated in the figure 4.3 and there are three outliers present in the labor cost data as shown in the figure 4.4.

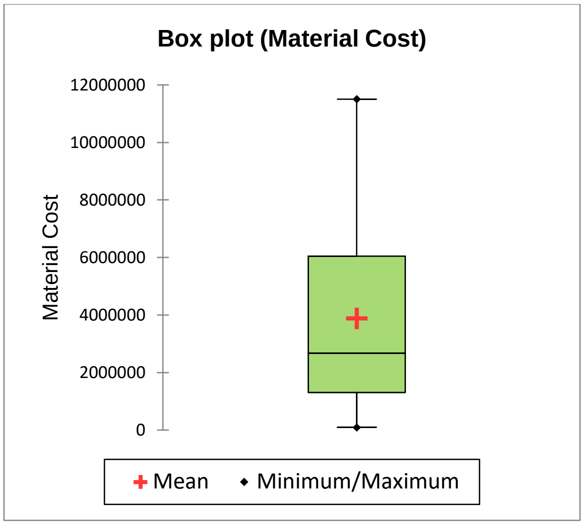

Figure 4.5: Box plot for material cost

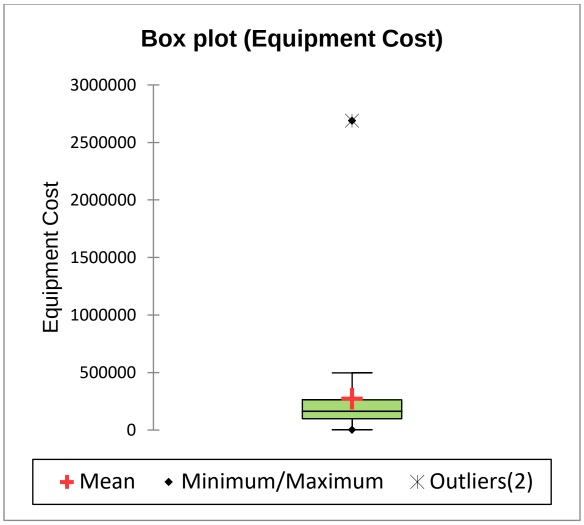

Figure 4.6: Box plot for equipment cost

The box plot in figure 4.5 clearly shows that there are no outliers in the material costs but figure 4.6 represents that there are two outliers in the equipment costs.

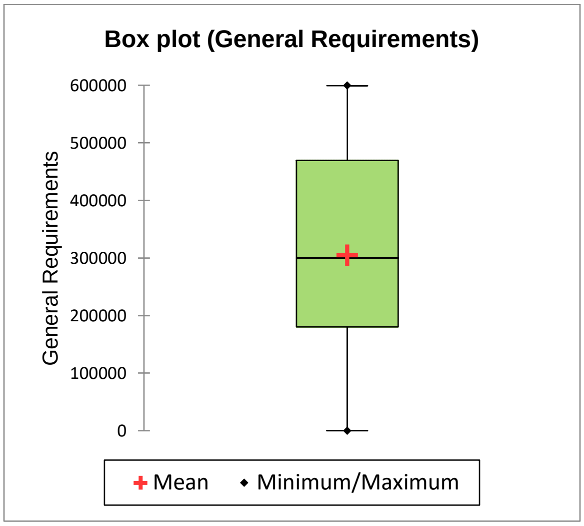

Figure 4.7: Box plot for general requirements

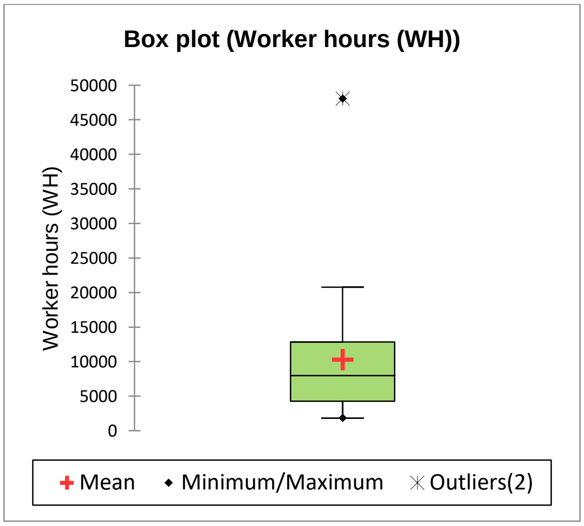

Figure 4.8: Box plot for worker hours (WH)

Figure 4.7 illustrates that there are no outliers in the general requirements, but two outliers were found in the worker hours (WH) box plot as shown in figure 4.8.

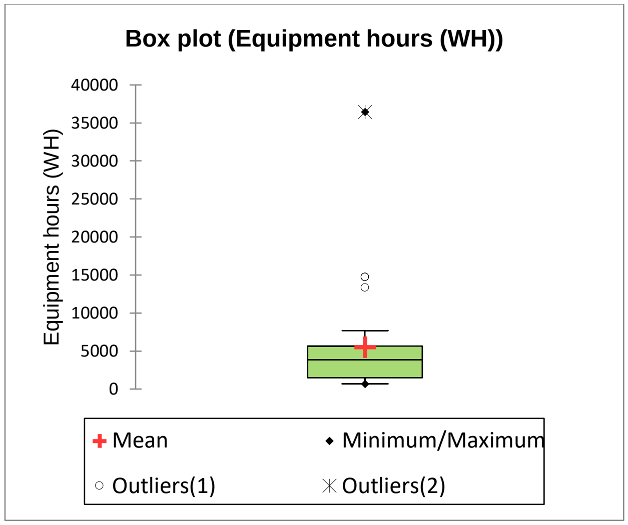

Figure 4.9: Box plot for equipment hours (WH)

The box plot for equipment hours represent that there are three outliers for this particular variable as shown in the figure 4.9.

After performing the analysis, PCA was performed with and without including these outliers. Missing data was not accepted in the XLSTAT software, as a result, observations with missing data were removed and then analysis was performed on 24 observations out of 28. After identifying the outliers using univariate box plots, the observations with outliers were removed and the analysis was performed on 16 observations. A summary of the results is illustrated below;

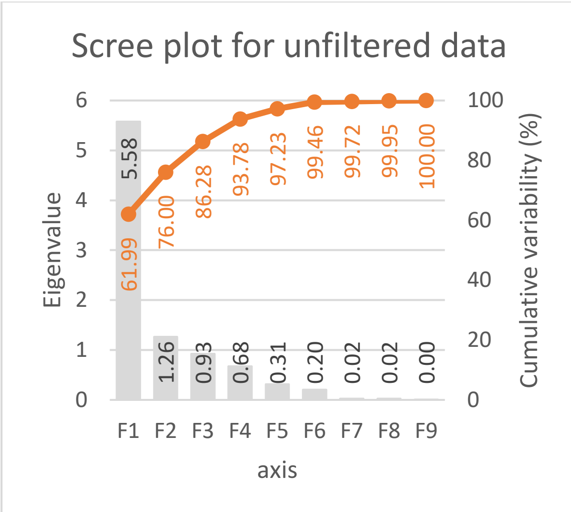

Figure 4.10: Scree plot for data with outliers

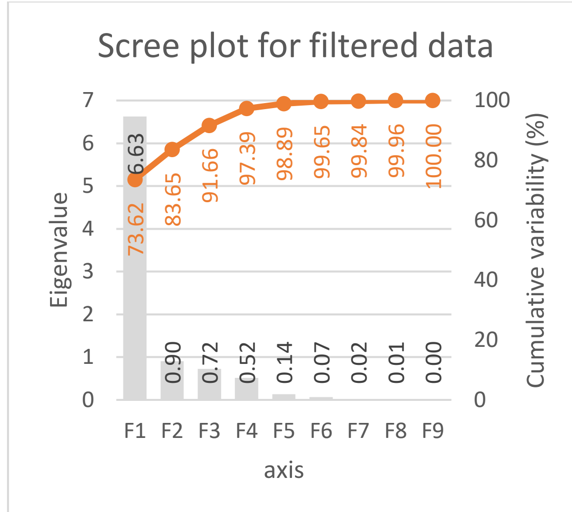

Figure 4.11: Scree plot for data without outliers

Figure 4.10 and figure 4.11 represent scree plots for the unfiltered and filtered data. Scree plot represent eigenvalues on one hand and variance explained by the axes on the other. Eigen values helps to choose principal components in order of their significance and the variance explains the information extracted by these components from the data. As shown in Figure 4.10 (data with outliers), axes F1 and F2 explain 76% variability in the data and two components can be selected based on the Rencher's (1995) Criterion. Whereas, in Figure 4.11 (data without outliers), axes F1 and F2 explain 83.65% variability and only first component satisfy Rencher's Criterion.

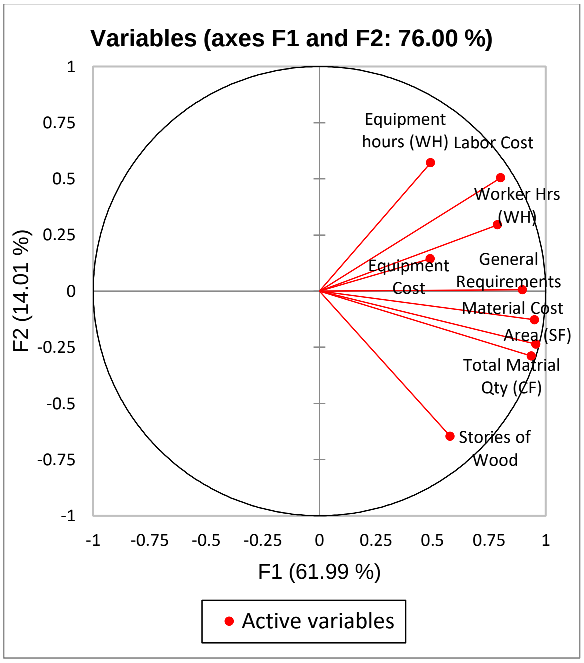

Figure 4.12: Correlation circle for data with outliers

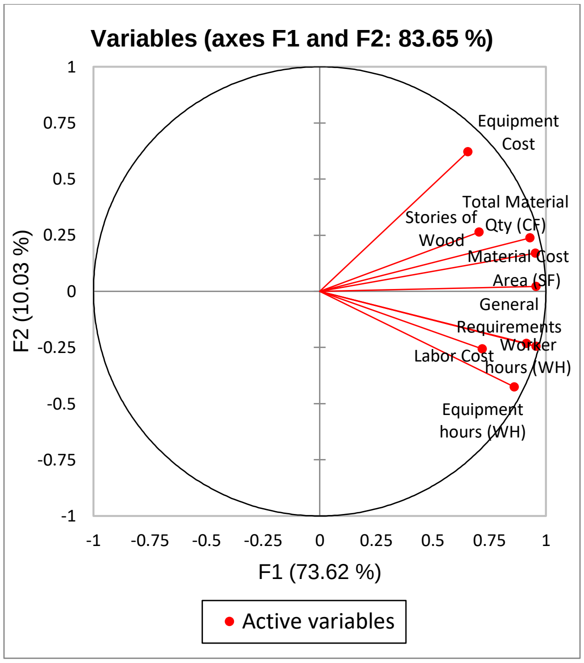

Figure 4.13: Correlation circle for filtered data

Figures 4.12 and 4.13 illustrate the correlation circle, which can be used to interpret the correlation between variables and the information captured by corresponding axes. When variables are far from the center, then, most of the information respective to those variables is carried by the represented axes. In contrast, if they are close to the center, some information of these variables might be carried by other axis (such as equipment cost in Figure 4.12), and any interpretation can be hazardous in such cases. For correlation, if two variables are at an acute angle with each other, they are significantly positively correlated such as material cost and total material quantity (CF) as shown in Figure 4.13. When two variables are orthogonal, they are not correlated and if they are on the opposite side of the center, they are negatively correlated. With the analysis shown above in figure 4.13, area (SF), general requirements, total material quantity, material cost, worker hours (WH) are highly correlated.

4.3. PRINCIPAL COMPONENT REGRESSION MODEL DEVELOPMENT

As a result of this analysis, it can be concluded based on Rencher's Criterion that for the analysis with outliers and analysis without outliers, two principal components and one principal component can be selected respectively. With the help of these uncorrelated components we can safely perform regression against the dependent variable i.e. cost of the system. Equation 4.1 depicts the regression model for cost of the system which includes outliers. This regression was performed on two selected principal components (PC_1 and PC_2) having eigenvalues greater than 1 (Rencher's Criterion). This regression showed a strong model with a R^2 (coefficient of determination) value of 0.933 but with a very high value of Mean Absolute Percentage Error (MAPE) equal to 127.03.

Cost of the system *= 647925.863066667+1772297.84973879*PC*$_1$*+15071.507622778*PC*$_2$ *..............(4.1)*

In contrast, when regression was performed on the filtered data, only one principal component was selected, and the model developed showed a strong R^2 value of 0.936 with relatively low MAPE of 53.80. Equation 4.2 illustrate the regression model for the filtered data.

Cost of the system *= 8.66525016473902E-11+780496.348259816*PC*$_1$*...(4.2)*

Figures 4.14 and 4.15 illustrate the plot of actual cost of the system against predicted cost of the system for both the models. Tables 4.1 and 4.2 represent goodness of fit statistics for both the models.

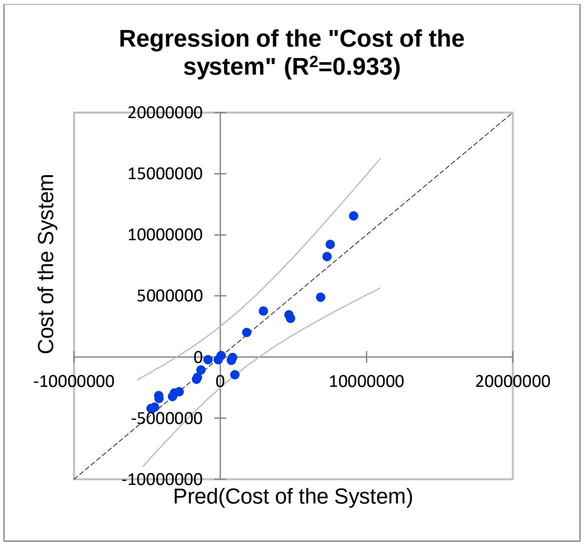

Figure 4.14: Predicted vs actual cost plot (including outliers)

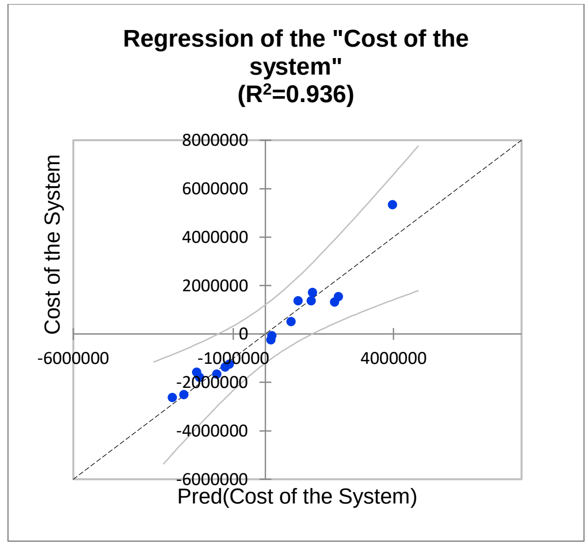

Figure 4.15: Predicted vs actual plot (excluding outliers)

Table 4.1: Goodness of fit statistics (including outliers)

Observations	24.000
Sum of weights	24.000
DF	21.000
R^2	0.933
Adjusted R^2	0.927
MSE	1379195330350.590
RMSE	1174391.472
MAPE	127.031
PC	0.086

Table 4.2: Goodness of fit statistics (excluding outliers)

Observations	16.000
Sum of weights	16.000
DF	14.000
R^2	0.936
Adjusted R^2	0.932
MSE	295285253182.005
RMSE	543401.558
MAPE	53.805
PC	0.082

4.4. LIMITATIONS OF PCR MODEL

Regression models developed with the help of PCR, explained 93% of the variability of the dependent variable, solved the problem of multicollinearity and produced a reliable model. However, considering the scope of this study, limitations were observed with these models for the desired outcome of a cost prediction tool. The major disadvantage that was observed with these models was that the equation represented independent variables in terms of the principal components. Due to this nature of the model, to predict costs, it was essential to perform PCA to generate values for these independent variables and that creates a time-consuming process and threatens the long-term usability of the model. Furthermore, it also produces a model which is complicated and difficult for a new user to operate. Given a broad perspective of this study and the audience, it was necessary to build a model which can be used by almost everyone and is user friendly. To resolve these limitations an alternative method was researched, almost similar to PCR but with the advantages of satisfying the goals of this study. Partial Least Squares Regression (PLSR) was considered as a better alternative to PCR to achieve the goals of this research. Following limitations were observed for the PCR and are explained below.

1) *Ease of use:*

With a goal of developing a cost prediction model which can be used by everyone and is easy to use, various methods were reviewed. But, with the nature of this study, state of the market, and received sample size and its nature challenged us to investigate an untraveled path. To deliver a reliable and a credible model, a method was selected which was accurate and deals with the sample size and its nature. In this process, ease to use this model and interpret the interrelationships was compromised. Given the outcome of this study to understand the relationships between independent and a dependent variable, developing a PCR model does not satisfy the intent of this study and forms a limitation of using this model.

2) Time consuming process:

After developing a model, it was observed that the model is represented in terms of principal components which ultimately needs values for these principal components to perform prediction. These values can only be obtained if a PCA is performed on the data, which adds an extra effort to predict the costs. This step makes the whole process time consuming. Moreover, for the people using this tool, it creates a limitation of knowing this method to perform the prediction. Which makes this model unreliable for the audience not familiar with this method.

3) Complicated model to understand relationships between independent and dependent variable:

As shown in equations 4.1 and 4.2, the dependent variable is expressed in terms of the principal components and not in terms of the independent variables. With this nature of the model, it is very difficult to explain the interrelationships between various independent variables with the dependent variable. To understand the cost implications of the variables listed in this study, it is necessary to have a mathematical representation of relationships which is not satisfied through these models.

Given these limitations and to fulfill the goal of this research, an alternative method Partial Least Squares Regression (PLSR) was used for developing the model and to resolve these limitations observed in PCR models.

4.5. PARTIAL LEAST SQUARES REGRESSION (PLSR)

As discussed in Section 3.2.2.3, PCA can be used to reduce the dimensions and to address the multicollinearity problems. The drawback of that approach is that it only captures the characteristics of the independent or predictive variables. The relationship of independent and dependent variables is not considered during the process of dimension reduction. Therefore, PCA is considered as an unsupervised dimension reduction technique. When the goal is to perform multivariate regression, considering the relationship between predictors and target variables in addition to capturing most of the information in predictors, can significantly improve the predictive ability of the model. Partial least square provides an alternative approach to achieve this balance (Maitra and Yan, 2008). PLSR is a technique that generalizes and combines features from principal component analysis and multiple linear regression (Abdi, 2007). Its goal is similar to the goal of regression, specifically to predict a dependent variable from a set of independent variables. This prediction is performed by extracting from the predictors a set of orthogonal factors called latent variables which have the best predictive power. PLSR is useful when we need to predict a set of dependent variables from a very large set of independent variables (predictors).

This method originated in the field of social sciences specifically economics, but has been popular in the areas of chemical engineering and sensory evaluation (Maitra and Yan, 2008; Martens & Naes, 1989). H. Wold pioneered PLS in the field of economics in late sixties, following that S. Wold and H. Martens pioneered its chemical applications in late seventies (Geladi & Kowalski, 1986). Initially, PLS was demonstrated as an algorithm similar to the method used for computing eigenvectors but was interpreted in a statistical framework in no time (Abdi, 2007; Phatak, & de Jong, 1997; Tenenhaus, 1998; Ter Braak & de Jong, 1998). Articles published by Wold et al. (1984) and Otto et al. (1985) demonstrate PLS is more robust and is a good alternative to the traditional multiple linear regression and `principal component regression approach. PLS was considered as a robust method due to the stability of its model parameters with the addition of new calibration samples. Geladi and Kowalski (1985) have illustrated a tutorial on PLS regression including the steps involved in it. They have also provided an understanding of MLR and PCR with some practical suggestions to perform PLSR.

Maitra and Yan (2008) provided a comparison of principal component analysis and partial least squares regression as two-dimension reduction techniques. They have also provided an underlying algorithm used in these methods for dimension reduction. Table 4.3 provides a comparison of multiple linear regression, principal component regression, and partial least squares regression corresponding to this study. These methods are compared based on dimensionality reduction, multicollinearity, mathematical base and solution, ease of use, supervised approach for dimension reduction, and ease of interpretation. These criterions were compared on the scale of poor, good, and excellent with 'NA' to show inability of the method.

Table 4.3: Comparison of regression methods to develop cost model

Criteria	**MLR**	**PCR**	**PLSR**
Dimensionality Reduction	NA	Good	Excellent
Multicollinearity Handling	Poor	Excellent	Excellent
Mathematical Solution	Excellent	Good	Excellent
Ease of use	Good	Good	Excellent
Supervised Dimension Reduction	NA	Poor	Excellent
Ease of interpretation	Excellent	Poor	Excellent

As it is clear from the table above that PLSR is a better version of PCA and is very beneficial specific to this study. Based on the points discussed in section 4.4, PCA was able to perform computations, but could not fulfill deliverables of this study. PLSR overcomes all those limitations and generates results which are comparatively easily interpretable, provides a user-friendly method, and delivers an efficient approach. Due to these benefits, PLSR was performed and computed results and developed models are explained in the following section.

4.6. MODEL DEVELOPMENT

This section demonstrates models developed with the help of PLSR and Table 4.4 summarizes the results of these models. A total of five models were developed, three on a broader scale to include most of the contributing variables and two for the development of a prediction tool involving variables that are more likely to be defined at conceptual stage of a construction project. The first broader scale model included all nine variables; the following two models were optimized to include variables that contributed the most to the dependent variable value. The two models developed for use in the conceptual construction stage included area of the system, stories of wood, and total material quantity. Models were optimized with the help of outlier analysis to filter outliers and by selecting the most important components to improve the model accuracy.

Cost Model 1

The initial model was developed based on complete dataset, to check the variance explained by the components, number of components contributing the most, contributing variables, and to assess outliers. The following cost models were optimized based on the results obtained to improve the accuracy and to include only contributing variables. The initial model showed a high R^2 value of 0.975 and a relatively close Q^2 of 0.957. The difference of 0.018 between these two values represent a good quality of the data. Model parameters are given in Table 4.4 and goodness of fit statistics are illustrated in Table 4.5. Figure 4.16 shows the plot of actual vs predicted cost of the system. The correlation circle illustrated in Figure 4.22 and variable contributions in Figure 4.18 clearly show that equipment cost, and stories of wood have relatively low contribution to the cost of the system.

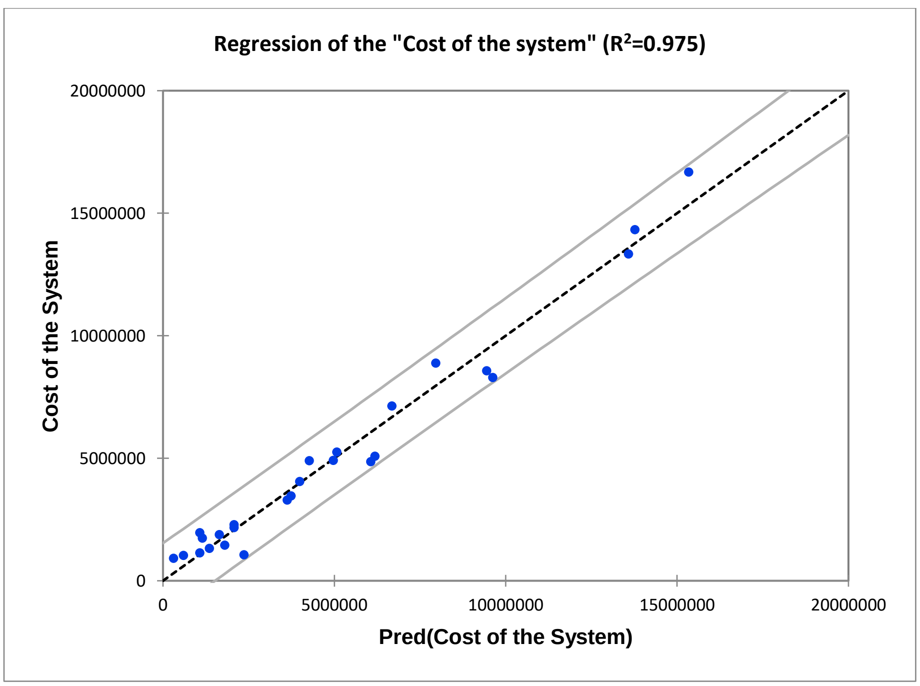

Figure 4.16: Plot of actual vs predicted cost of the system

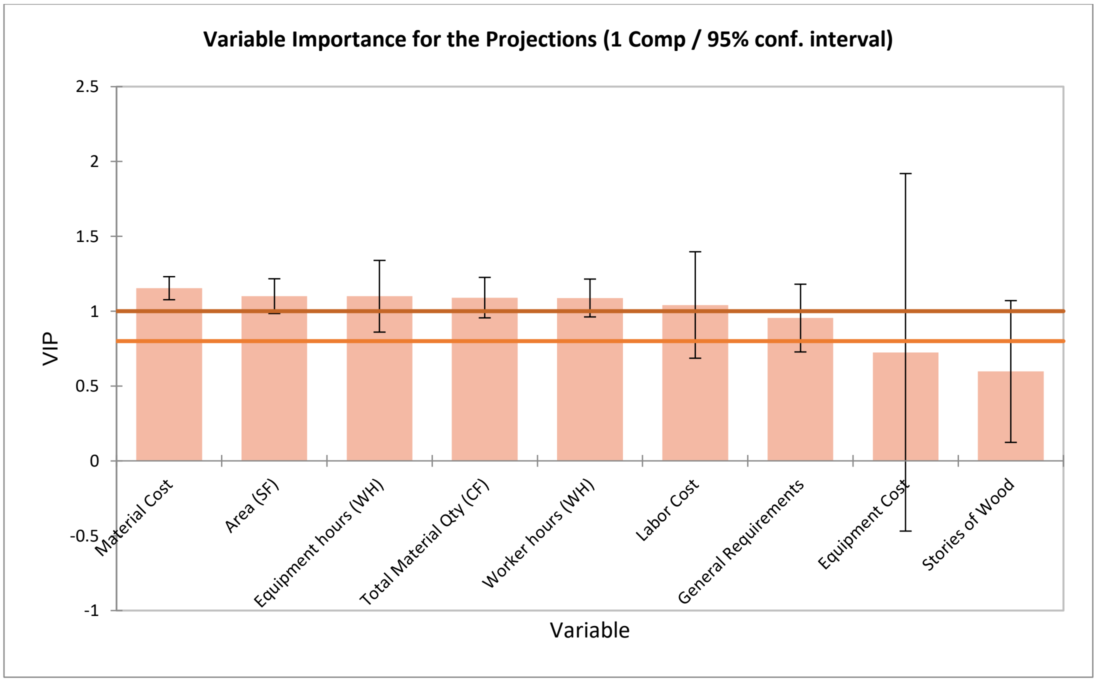

Figure 4.17: Variable Importance for the Projections (VIP) for each explanatory variable

<u>Cost Model 2 and 3</u>

Cost model 2 was modified based on the results of cost model 1; it excluded the two least contributing variables (stories of wood and equipment cost). The model parameters and results are demonstrated in tables 4.4 and 4.5. After removing both the variables, the remaining variables showed significant contribution to the model as shown in Figure 4.19. The obtained R^2 value of 0.961 was slightly less than cost model 1 but a very close Q^2 value of 0.954. The difference between R^2 and Q^2 value reduced significantly after removing equipment cost and stories of wood, which represent a high-quality data for the remaining variables. After looking at the results carefully, the outliers analysis in PLSR showed the presence of outliers and an increased error value. Figure 4.18 illustrate the plot of actual vs predicted cost of the system which mirrors the error value by showing significant distortions from the regression line. The next model was therefore developed by removing outliers noticed in outliers analysis.

Cost model 3 was performed on 23 observations and seven variables. The resulting model was the most accurate and explained maximum variance in the data. As given in Table 4.5, the obtained R^2 value of 0.985 was higher than both the models developed previously and much lower values of mean square errors. Furthermore, obtained Q^2 value of 0.979 was higher than previous two models and the least difference with the R^2 value. Figure 4.20 shows the plot of actual vs predicted cost of the system which illustrate much smaller and a narrow range of distortions from the regression line. Figure 4.21 shows the contribution of all the seven variables and all the variables are contributing significantly to the model. This model was considered as a reliable model based on the results obtained and no new models were developed.

Models for web-based tool

The part goal of this study was to develop a conceptual prediction tool that can be readily used by the Architecture, Engineering, and Construction (AEC) industry. To achieve this goal three variables, 1) area (SF), 2) stories of wood, and 3) total material quantity (CF) were selected which are typically identified during schematic design. Using these variables models were developed and optimized to predict the cost of the system. This section discusses these two models and results obtained for these two models.

The first developed model used 26 observations for training the model and three variables mentioned above. After performing PLS on the data, the iterations stopped on three components out of which the first two components explained variance equivalent to all the three components. The Q^2 value of 0.938 and an R^2 value of 0.959 was obtained shown in table 4.5. Model parameters and the goodness of fit characteristics are explained in tables 4.4 and 4.5. Figure 4.24

shows a plot of actual vs predicted cost of the system. Area of the system and total material quantity (CF) contributed significantly to the first component, whereas stories of wood lacked in making a significant impact. The second component solved this problem by capturing the remaining information and balanced the contribution of these three variables, making a substantial impact to the model. The third component did not contribute noticeable changes to the model as can be seen in figures 4.25, 4.26, and 4.27. The goal to develop the next model was to reduce the computation time of the model. Therefore, to achieve this goal, the iterations of the components were fixed to two components as the third component did not contribute significantly. The second component was selected to capture impactful contribution of the variable; stories of wood. The second model was developed by removing the outliers obtained in PLSR outliers analysis. The results improved significantly with a R^2 value of 0.995 and a very close Q^2 value of 0.991. As can be seen, initially with first model the difference obtained was 0.018 and with removing outliers it reduced to 0.004. This small gap shows that the data obtained was reliable and the gap was further reduced when outliers were removed. The Root Mean Square Error (RMSE) dramatically reduced from a value of $ 865,361.16 to $ 300,250.80 after excluding the outliers. Figures 4.28, 4.29, and 4.30 illustrate the results obtained for variable contribution and the plot of actual vs predicted cost of the system. Table 4.4 and 4.5 illustrate the parameters and goodness of fit characteristics.

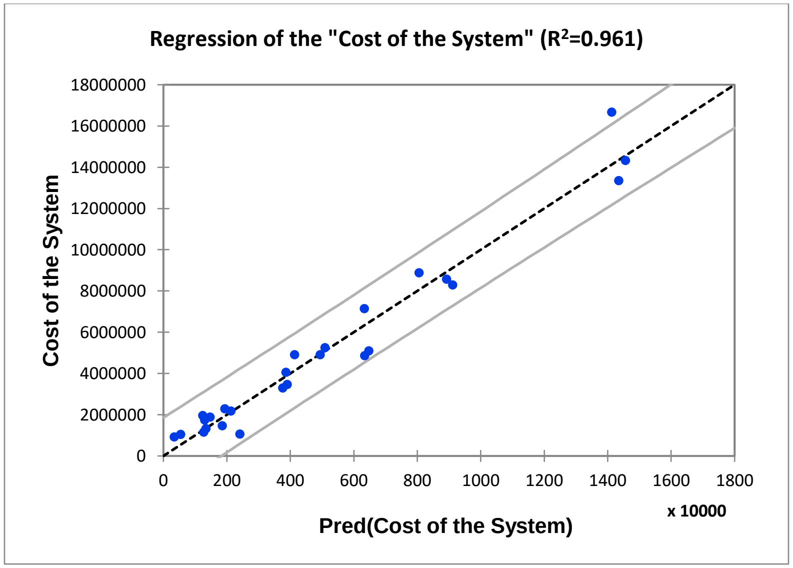

Figure 4.18: Plot of actual vs. predicted cost of the system (Cost model 2)

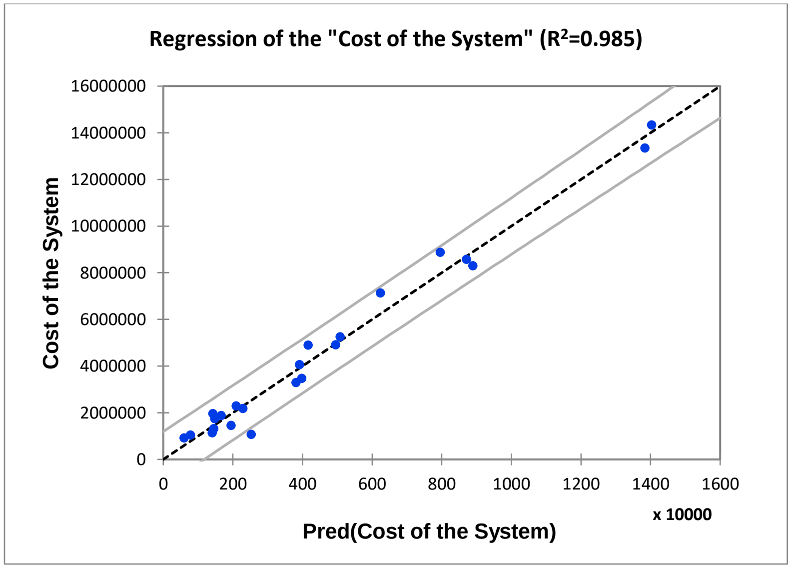

Figure 4.20: Plot of actual vs predicted cost of the system (Cost model 3)

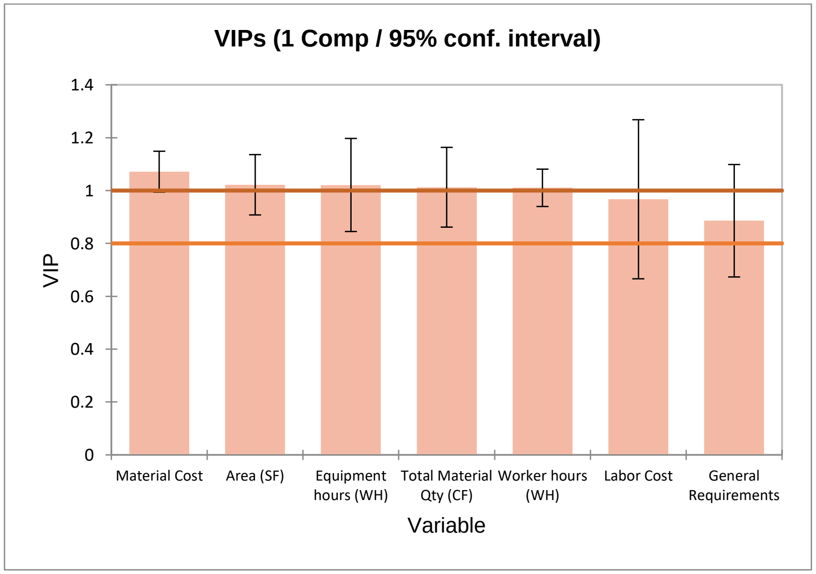

Figure 4.19: Variable Importance for the Projections (VIP) for explanatory variable

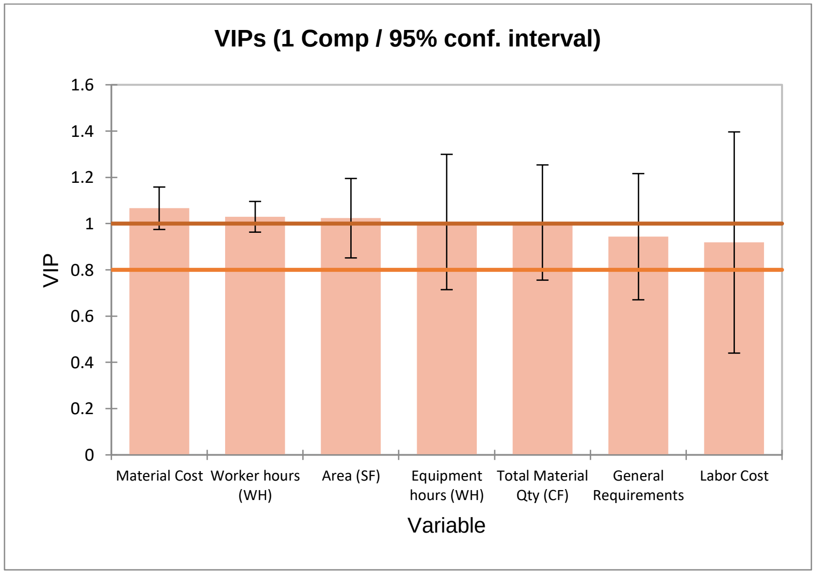

Figure 4.21: Variable Importance for the Projections (VIP) for explanatory variable

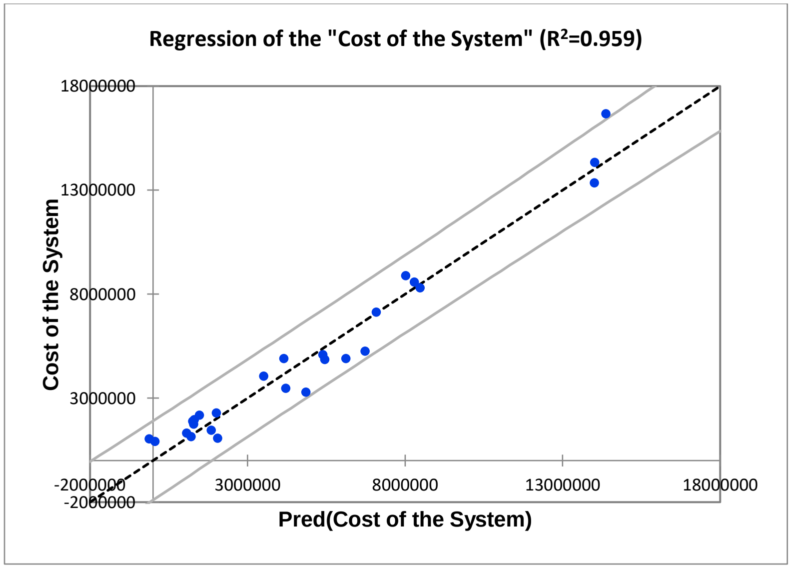

Figure 4.22: Plot of actual vs predicted cost of the system (Web-based model 1)

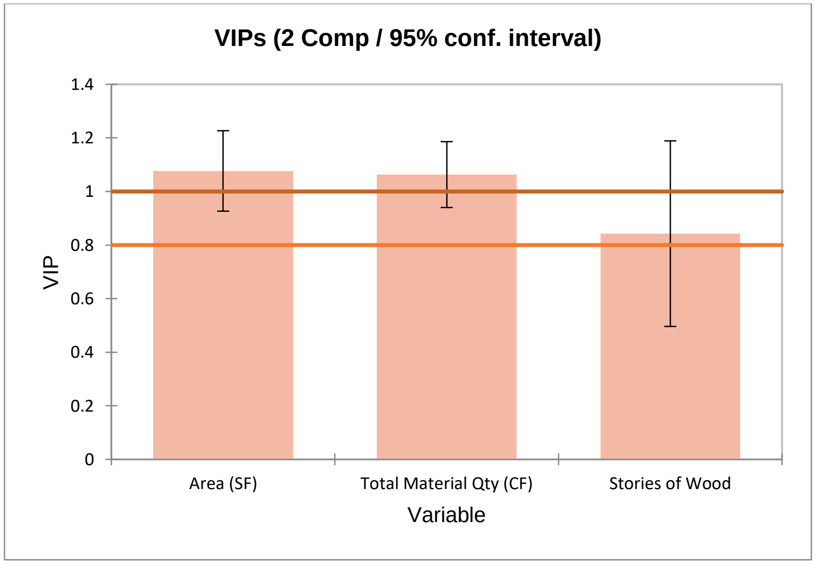

Figure 4.24:Variable Importance for the Projections (VIP) for explanatory variable

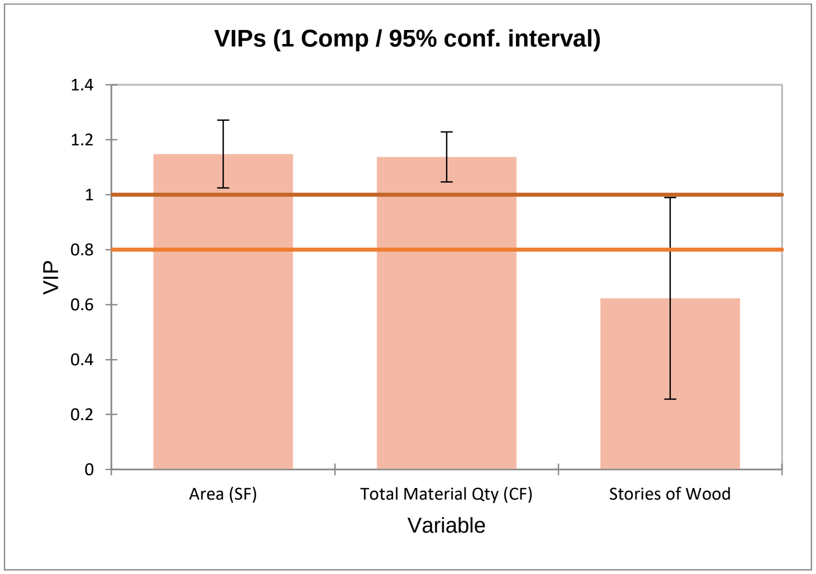

Figure 4.23: Variable Importance for the Projections (VIP) for explanatory variable

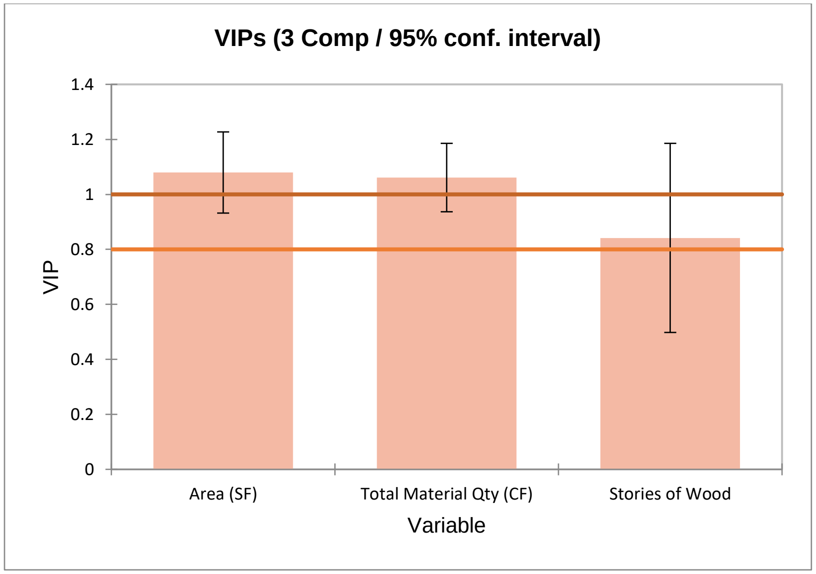

Figure 4.25: Variable Importance for the Projections (VIP) for explanatory variable

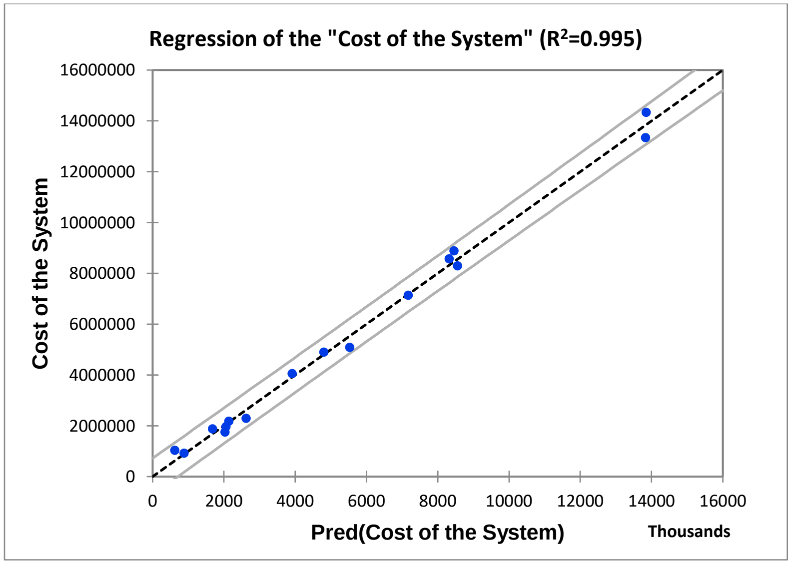

Figure 4.26: Plot of actual vs predicted cost of the system (Web-based model 2)

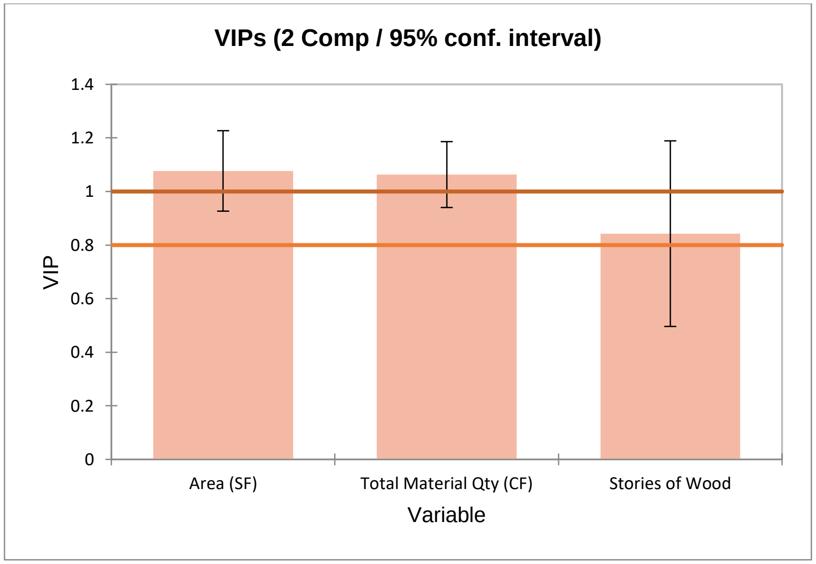

Figure 4.28:Variable Importance for the Projections (VIP) for explanatory variable

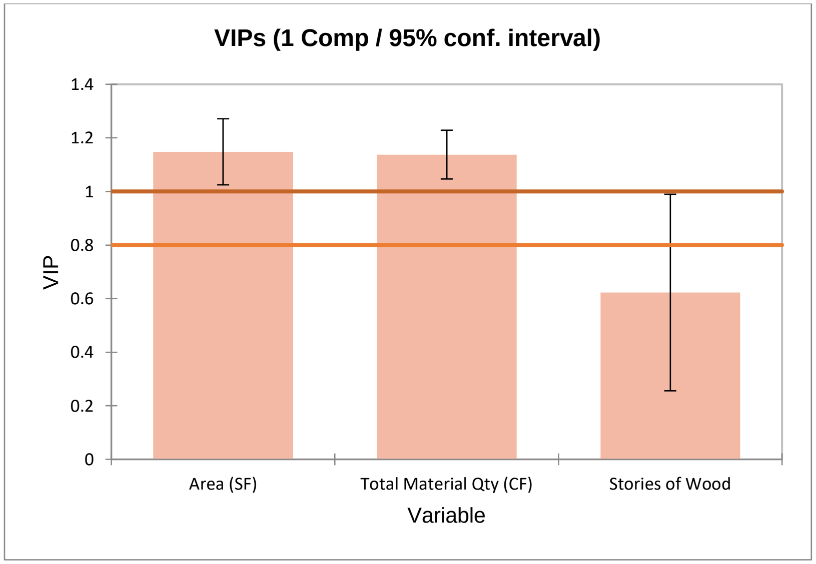

Figure 4.27: Variable Importance for the Projections (VIP) for explanatory variable

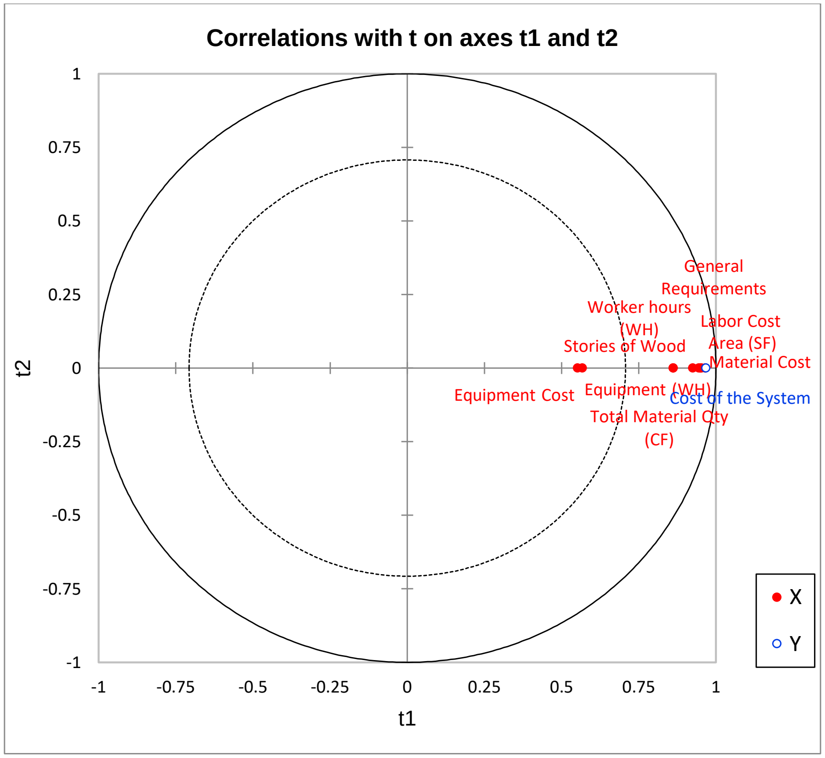

Figure 4.29: Correlation circle of the developed cost models

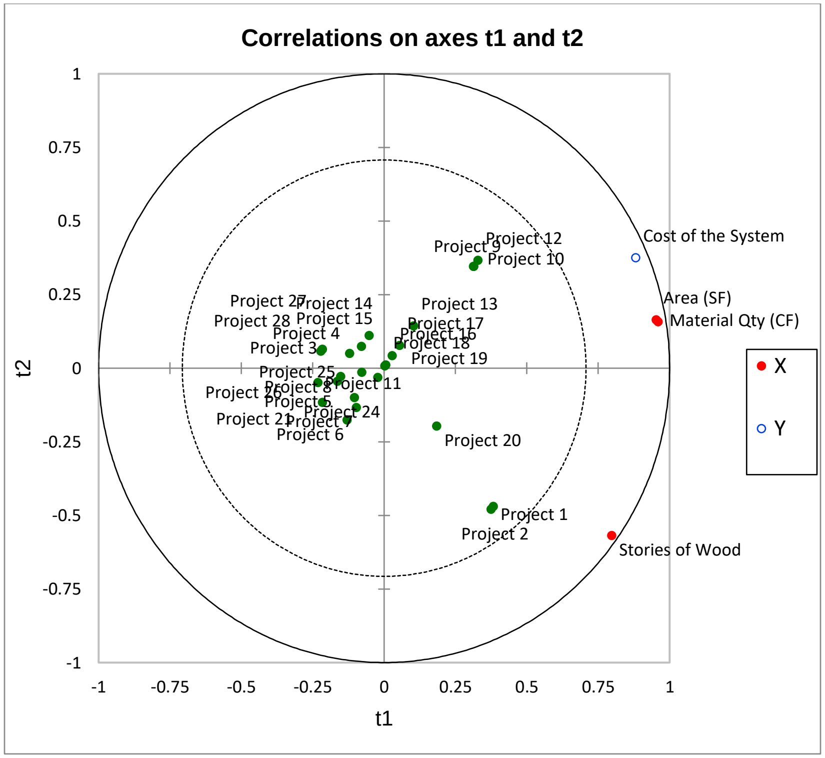

Figure 4.30: Correlation circle of the model developed for web-based tool

Table 4.4: Model Parameters for "Cost of the System"

	Cost Model 1	Cost Model 2	Cost Model 3	Web-Based Tool 1	Web-Based Tool 2
Variable	**Cost of the System**	**Cost of the System**	**Cost of the System**	**Cost of the System**	**Cost of the System**
Intercept	-863692.814	-1064149.889	-768676.688	-5169034.971	602321.968
Stories of Wood	86272.306	0	0	-163851.293	-371031.823
Area (SF)	6.659	7.431	6.844	52.245	27.593
Total Material Qty (CF)	7.898	8.919	8.138	45.747	33.136
Labor Cost	1.026	1.145	1.065	0	0
Material Cost	0.184	0.205	0.195	0	0
Equipment Cost	0.782	0	0	0	0
General Requirements	2.916	3.529	3.600	0	0
Worker hours (WH)	102.269	114.135	109.504	0	0
Equipment hours (WH)	146.921	163.966	155.420	0	0

Table 4.5: Goodness of fit characteristics of models developed with PLSR

	Cost Model 1	Cost Model 2	Cost Model 3	Web-Based Tool 1	Web-Based Tool 2
Variables	Stories of Wood	*Stories of Wood*	*Stories of Wood*	Stories of Wood	Stories of Wood
	Area (SF)	Area (SF)	Area (SF)	Area (SF)	Area (SF)
	Total Material Qty (CF)	Total Material Qty (CF)	Total Material Qty (CF)	Total Material Qty (CF)	Total Material Qty (CF)
	Labor Cost	Labor Cost	Labor Cost		
	Material Cost	Material Cost	Material Cost		
	Equipment Cost	*Equipment Cost*	*Equipment Cost*		
	General Requirements	General Requirements	General Requirements		
	Worker hours (WH)	Worker hours (WH)	Worker hours (WH)		
	Equipment hours (WH)	Equipment hours (WH)	Equipment hours (WH)		
Observations	26.000	26.000	23.000	26.000	16.000
Sum of weights	26.000	26.000	23.000	26.000	16.000
DF	24.000	24.000	21.000	22.000	13.000
R^2	**0.975**	**0.961**	**0.985**	**0.959**	**0.995**
Std. deviation	706337.091	875354.250	481334.934	940746.766	333098.360
MSE	460534233095.717	707303135746.117	211536942952.731	748849942818.762	90150545614.176
RMSE	678626.726	841013.160	459931.455	865361.163	300250.805
Number of components	1	1	1	3	2
Q^2	0.957	0.954	0.979	0.938	0.991

4.7. FINAL REGRESSION MODEL AND DEVELOPMENT OF A PREDICTION TOOL

With the models developed and optimized in the previous section, the final regression model developed are to be used in the prediction tool. The final models are divided in two categories; 1) A model with all significantly contributing independent variables 2) A model with a subset of independent variables which are available at the conceptual design stage of a construction project. This section provides these models developed which is the output of previous section, leading towards achieving the goal of this study. Below are the developed and optimized final regression equations of this study.

1) *Partial Least Squares Regression equation with all significantly contributing independent variables*

This equation was a direct outcome of the cost models that were developed and optimized in the previous section. As discussed previously, this equation was obtained after eliminating stories of wood and equipment cost due to their relatively small contribution and removing the outliers in the data. This is the final regression equation involving all the significantly contributing selected variables. This model is a baseline model to understand implications of the variables on the cost of the system and provides a core of this study.

Cost of the System *= -768676.688348047 + 6.84454584637333 * Area (SF) + 8.13842116105593 * Total Material Quantity (CF) + 1.06466461251464 * Labor Cost + 0.19460095714582 * Material Cost + 3.59979840832418 * General Requirements +109.503571981154 * Worker hours (WH) + 155.419740238789 * Equipment hours (WH) ....(4.3)*

2) *<u>Partial Least Squares Regression equation with a subset of independent variables to predict costs at the conceptual design stage of a construction project</u>*

This equation was developed keeping in mind the goal of this study to develop a cost prediction tool to predict costs at the conceptual stage of a construction project. With that said, the list of variables used in the baseline model above, requires a thorough detail of the project to determine the quantify the variables. Therefore, a subset of these variables that is a more reasonable list based on the data that most project teams will have during the conceptual design stage was selected. This list includes stories of wood, area of the system (SF), and total material quantity (CF). The regression equation obtained is illustrated below.

Cost of the System *= -602321.967589678 - 371031.822957286 * Stories of Wood + 27.5926124219158 * Area (SF) + 33.1362949877177 * Total Material Quantity (CF) ....(4.4)*

This equation above, was used further to develop a user-friendly prediction tool using Excel. The user interface for this tool is shown in Figure 4.31. The inputs for the variables, are to be entered in the yellow cells. The above equation was used to predict the cost of the system as an output which is obtained in the green cell as shown in Figure 4.31. The output cell was formulated with this equation and the yellow cells were used in the equation as inputs as can be seen in the formula bar of Figure 4.31.

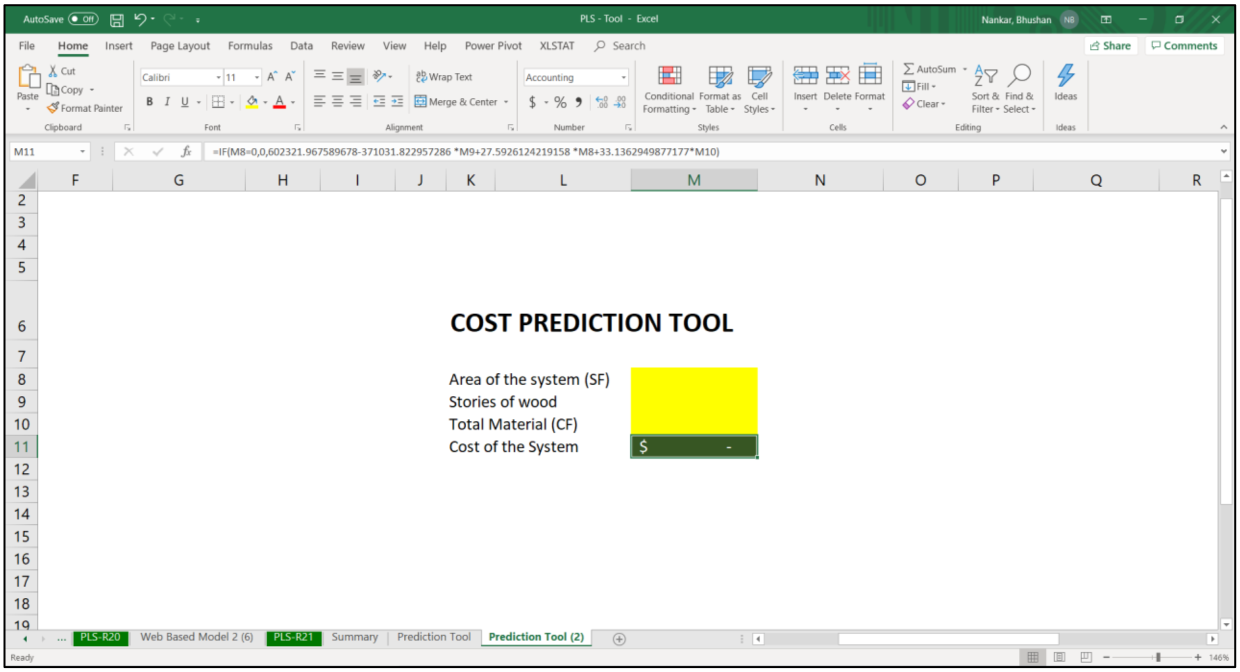

Figure 4.31: Screenshot of the cost prediction tool

4.8. MODEL TESTING

4.8.1. Case Study 1

The model developed for the application of predicting costs at the conceptual stages of a construction project was tested on a project which was not used to train the model out of the twenty-eight observations obtained. This section provides results of the model and accuracy of the model. This project is located in the Midwestern United States refer Figure 4.32 and is a six-story office building, with a total gross floor area of approximately 215,000 SF. Using this information model was tested for its accuracy and results are given below.

Table 4.6: Cost prediction results for case study 1

COST PREDICTION TOOL	
Area of the system (SF)	215,000
Stories of wood	6
Total Material (CF)	150,000
Cost of the System	$ 9,278,986.95

Actual cost of the system for this case study was $ 10,017,152.00. After predicting the costs, it was observed that the predicted costs vary by seven percent compared to the actual costs. This difference in cost could be a result of the fact that this project cost involved a mass timber package which was a combination of steel bracing and a mass timber frame. The costs for steel bracing are not accounted in the costs predicted by this tool. Also, this project was discarded as an outlier due to high labor costs and that difference is may be the reason for this variation. Despite these limitations, the cost prediction was in the acceptable range of variation according

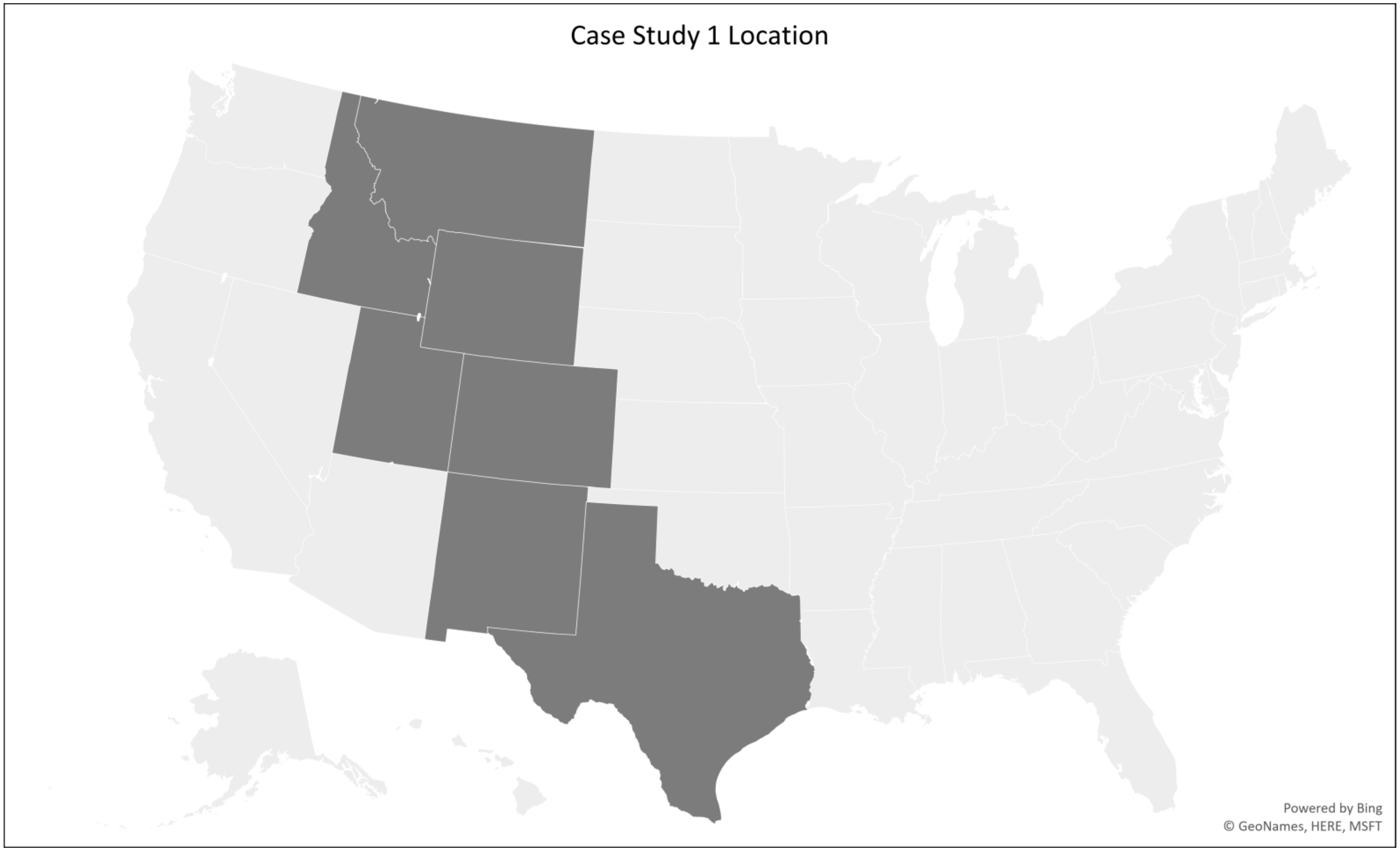

Figure 4.32*:* Case study 1 location

of margin of error of twenty percent for the given sample size.

4.8.2. Case Study 2

Another project was tested using the same model, in order to better understand the cost differences observed in the predicted and actual cost of the system. As shown in Figure 4.33, the project is located in the highlighted region, the Midwestern part of the United States. This is a four-story mass timber structure used for scientific laboratories, office, and gathering space. It features approximately 121,280 SF of mass timber area which accounts for 50,000 CF of CLT and approximately 52,000 CF of glulam. This project is currently under construction and is expected to be complete in Fall 2020.

Table 4.7: Cost prediction results for case study 2

COST PREDICTION TOOL		
Area of the system (SF)		121,280
Stories of wood		4
Total Material (CF)		102,590
Cost of the System	$	5,864,079.21

This project's mass timber package also included steel bracing similar to case study 1; however, cost data was available which allowed us to deduct the costs for steel bracing and understand the actual differences between actual and predicted costs. The initial package construction cost of $ 6,700,000 included the costs of the steel bracing, due to which the model showed a variation of nearly twelve percent from actual costs. To better understand the accuracy of the model, the actual cost of this structure excluding steel bracing was adjusted to account for the cost of the steel bracing and the cost difference reduced from twelve percent to two percent. This solidifies the accuracy of the model within twenty percent error.

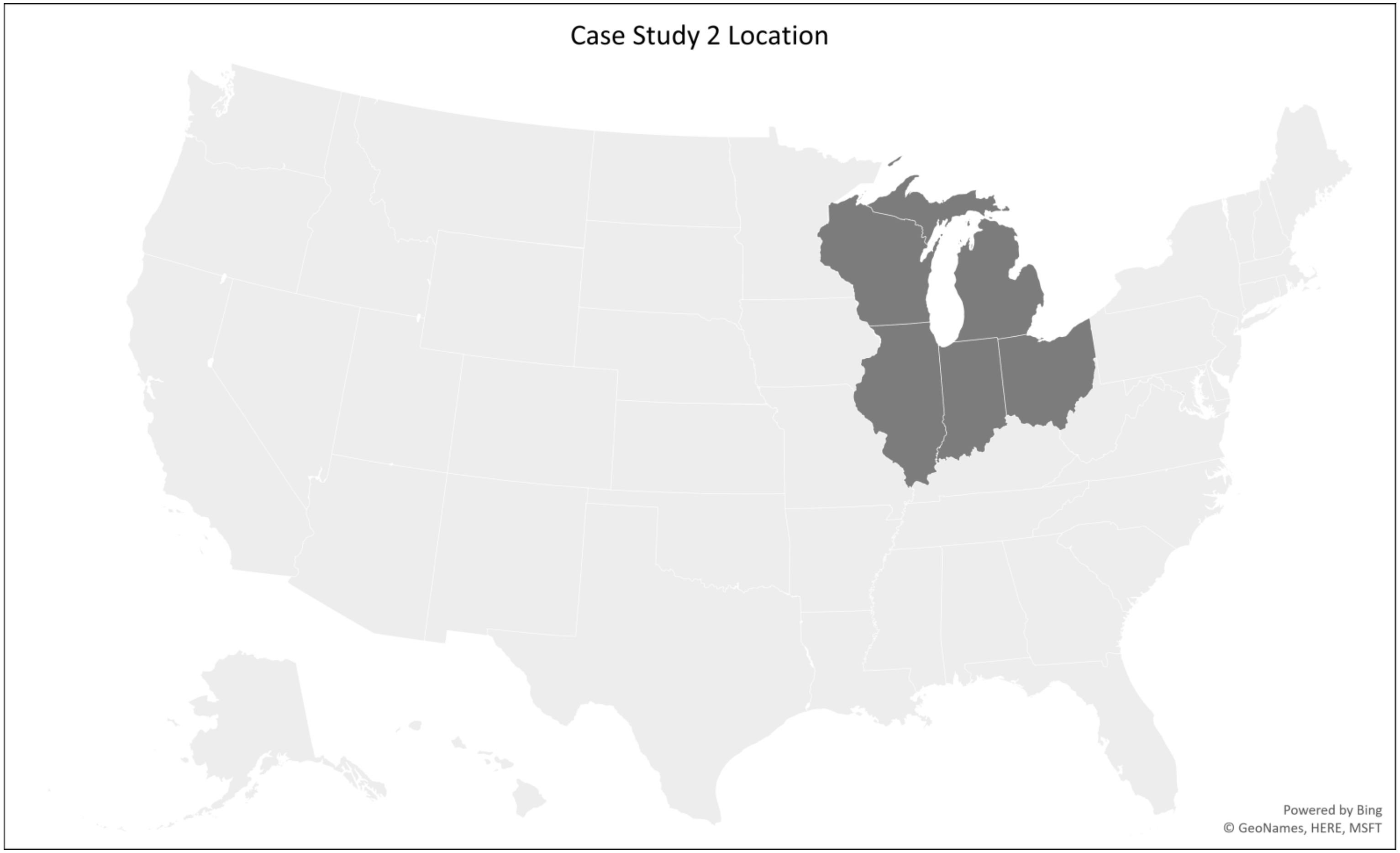

Figure 4.33*: Case study 2 location*

4.9. SUMMARY

This chapter is majorly focused on the application of the literature and using it to develop the cost model and to fulfil the goal of this study. This chapter started with data analysis and then using PCR to develop initial models. Observing the results and the model, it was necessary to improve the model for ease of use and interpretation, an alternative methodology was researched and PLSR was selected due to its benefits for this study. This chapter provides succinct literature on PLSR and progresses towards building the cost models and provides results for the developed cost models. The last section tests the model for its accuracy and ease of use. The next chapter will focus on the conclusions and future areas for improvement.

CHAPTER 5 : SUMMARY, CONTRIBUTIONS, AND FUTURE RESEARCH

5.1. OVERVIEW

Chapter 4 summarized the execution phase of this study by providing results for the cost models and improving on the process for better results, as well as demonstrating the application of the model to a case study. This chapter will provide summary of this study, conclusions based on the results obtained, and future research scopes.

The main outcome of this study was to build a cost prediction model and study various parameters impacting the costs. To achieve this, Chapter 1 discussed the background of this study, introduced the term mass timber, and defined the scope and plan of this study. Chapter 2 elaborated the available literature in detail which included material properties, manufacturing processes, construction techniques, previous cost studies, and a review of cost prediction models. Chapter 3 focused on the research methodology, a discussion of variables used in the study, and literature on select statistical prediction methods. Chapter 4 used all this information to build the cost prediction model and improve it to better suit the needs of this study. All these chapters are linked to each other and contribute towards the final goal. The following section will summarize outputs of the set objectives.

5.2. SUMMARY OF OUTPUTS

The goal of Objective 1 (Understand the current state and background of CLT and glulam in construction) was to understand the relevant background for this study and increase understanding of the current state of the mass timber construction market. This was achieved with the help of literature review and field observation. Literature review mainly helped in understanding the methods used in prediction models and provided a list of potential methods which can be used for prediction model. Field observation in combination with literature review laid out eighteen variables contributing to the cost of the system. These lists of methods and variables were used in Objective 2 (Build a foundation for development of the predictive cost model) as a necessary input.

This objective can be considered as a conception stage of this study, which concluded cost as one of the major barriers for its adoption. Building on this conclusion it translates into a design stage where research goal was determined and with the literature review scopes were outlined. This objective also concluded the list of eighteen variables as shown in Table 3.1 and the possible methods for model development as mentioned in Section 2.4.

Objective 2 was divided into three steps. The first step involved determining appropriate variables for the study and was used to produce a spreadsheet of projects that were candidates for inclusion in the study along with their variable values. This spreadsheet was circulated in the second step of data collection. During the data collection process, simultaneously a method for predictive cost model development was researched and Principal Component Regression (PCR) was selected. These two outputs data and methodology served as an input for the next objective.

This objective can be considered as a preconstruction stage, where groundwork required for the model development was created. This objective concluded that with a small sample size and a large number of variables with a possibility of multicollinearity, Principal Component Regression is a suitable choice.

Objective 3 (Development and testing of the model) was divided into two parts as model development and testing of the model. With data and methodology as an input from Objective 2, models were developed. With the obtained models, it was observed that the models were not satisfying the goals of this study. With that said, an alternative method, Partial Least Squares Regression (PLSR) was implemented and the step was performed again. The obtained models were appropriate to this study and satisfied the goals of this study. The next step being testing of this model, a pilot project was tested on the model and the predicted cost varied with the actual cost by nine percent.

This was the actual execution stage of the research where the chosen methodology was implemented, and developed models were analyzed to suit the goals of this study. Based on the PCR models, it was concluded that PLSR is better suited for this research and was utilized to develop final models.

5.3. CONCLUSIONS

- *Coefficient of determination (R^2):* The R^2 value is nothing but the measure of the predictive accuracy of the model. Its value ranges from 0 to 1, where 1 represents 100% predictive accuracy. The literature suggests 0.75, 0.50, 0.25 as substantial, moderate, and weak levels of predictive accuracy as a rough rule of thumb of an acceptable R^2 value (Hair et al., 2011; Henseler et al., 2009). The final equations developed in this study showed a high R^2 value of 0.985 for an equation including all the significantly contributing variables and 0.995 for a conceptual model. These values indicate that the equations developed have substantial prediction accuracy and are reliable.

- *Fit of the Model and Data Quality:* Q^2 is a way to measure the predictive relevance and the difference between R^2 and Q^2 provides an idea about the model fit (Xiaosong and Fujun, 2012). Also, the difference should not be greater than 0.3 which shows overfitting (Leach 2001; Veerasamy et al., 2011). The obtained difference of 0.006 and 0.004 shows that the model is a good fit and does not show overfitting. Additionally, the difference between R^2 and Q^2 did not show much variation (from 0.016 to 0.004) when the model included outliers and excluded outliers, which indicates that good quality data was obtained.

- *Variables impacting the cost of the system:* According to PLSR analysis, a total of seven variables contributed significantly out of which the material cost was the most contributing variable to the cost of the system. Whereas, stories of wood and equipment cost did not contribute significantly.

5.4. RESEARCH CONTRIBUTIONS

The primary goal of this study was to increase awareness about the costs associated with mass timber construction in buildings that use a CLT and glulam gravity frame. To achieve this goal, a cost prediction model was developed to improve the cost knowledge base and provide understanding of the variables impacting the costs. This tool was envisioned to break the cost barrier for the adoption of these materials by predicting costs at the conceptual stages of the project when the decision for structural material is taken. This section will discuss the notable contributions of this study.

5.4.1. Cost Prediction Tool to Predict Costs at Conceptual Stages of a Construction Project

The major physical contribution of this study is a cost prediction tool to predict costs at the conceptual stages of a construction project. This tool was developed with the model using the variables stories of wood, area of the system (SF), and total material quantity (CF) as input variables. The values for these variables are available at the conceptual stages of the project and using this tool costs associated with the gravity frame can be computed. This will provide a premise for an economic comparison with other structural frame construction options. The regression model on which this tool is built has a coefficient of determination of value 0.954 which is considered as a good representation of the variance of the dependent variable.

5.4.2. Implications of Variables on Construction Costs

Another important outcome of this study is the variables that impact the cost of the system and how they impact the costs. As shown in correlation circles in Chapter 4, out of a total nine variables seven variables have a high impact on the costs which are 1) labor cost, 2) worker hours (WH), 3) equipment hours (WH), 4) total material quantity (CF), 5) area (SF), 6) material cost, and 7) general requirements. Equipment costs and stories of wood were unable to represent significant variation in the cost of the system. An important observation is that all these variables are highly correlated with each other and depicts that measures should be taken while using these variables for building such models.

5.4.3. A Regression Model for More Accurate Cost Predictions

In the process of developing a conceptual cost prediction tool, a regression model using all the nine variables was developed. This model was optimized to include the variables which are contributing significantly to the model. The optimized model is a regression model developed based on the dataset obtained to predict costs for the audience who wants to check costs much further in the project lifecycle. This can be a preconstruction stage and can be helpful for the estimators to improve the costs based on the variables contributing the most. Based on the optimized model, it can be seen in Figure 4.21 that material costs, worker hours (WH), area (SF) are the most dominating contributors. Following these three variables equipment hours (WH) and total material quantity (CF) are contributing significantly and labor cost is the least significant contributor to the model. This information is a significant addition to the body of knowledge which can be used to develop more robust cost models and estimates.

5.4.4. Comparison of Principal Component Regression with Partial Least Squares Regression

This study improves from using PCR to PLSR due to the nature of outcomes required. Issues encountered with PCR specific to this study are addressed using PLSR which are enlisted in section 4.4. Studies where there is an uncertainty on how the dependent variables and independent variables are related and if independent variables are contributing to the dependent variables or not, PLSR should be the method of choice.

5.4.5. Engine for Development of a Widely Available Cost Prediction Model

Even though the scope of this study is to focus on the gravity frame and not on the overall project costs, this tool provides an engine for development of existing models which are used to predict overall costs. This model also provides a background working of some of these existing cost prediction models which are presented as an unpredictable black box. Methodology used in this tool development can be easily followed to modify and develop new models.

5.5. FUTURE RESEARCH AREAS

Despite being at relatively embryonic stages in the United States, mass timber construction has shown a great deal of potential in construction sector. There is a need to spread awareness about this construction type and address the identified practice barriers through research. While this study addresses the cost barrier, there are other important areas which needs to be addressed and this section focuses on recommending future research areas and improvements that can be done to prepare a more robust prediction model.

5.5.1. Increasing sample size and adding categorical variables to include more variation

The developed regression model was built on a small sample size and despite obtaining good results and a credible model, it needs to be tested for a wide sample size to include more variation. With small sample size it is possible that the model is developed is focused on a cluster and can produce inaccurate results if sample from a different cluster is tested. As more projects are built and this market grows, more projects should be included in the sample and model should be improved. According to Slovin's formula (Stephanie, 2003), to achieve an accuracy of

95% confidence level and a ten percent margin of error, at least 79 samples should be included in the model. Additionally, categorical variables can be included in the model with advanced PLSR approaches. Simultaneously, more variables can be added into the model like geography of the projects, material source, and travel distance from the mass timber manufacturer.

5.5.2. Use of Computer-Based Methods for Prediction Model Development

The primary goal of this study was identifying the cost implications of various mass timber project-specific variables, it was necessary to develop a mathematical equation representing the relationship. Now that the equation has been developed, more advanced methods can be used to optimize this model. In this era of artificial intelligence, Artificial Neural Network, Genetic Algorithm, and Case-Based Reasoning can be used to predict the costs and optimize the model. To achieve this goal in future, more data needs to be collected and using similar variables models based on these new techniques can be built.

5.5.3. A Model to Predict Complete Project Costs

This model was developed to predict costs of the system and not the overall costs, due to a small sample size and the need for variable reduction. As more project data is available, a more robust model can be developed to predict overall project costs. Also, with more data, projects can be divided based on building use and occupancy classification and models can be built to include that as a variable. This will enable the user more advanced options while predicting costs and can be achieved with inclusion of more variables corresponding to overall costs of the project and not just frame of a structure.

5.5.4. Construction Time Model Development

A similar approach can be used to develop construction time prediction model, as time is a major factor contributing to the costs of this construction type. Construction time can either be calculated with an equivalent schedule indicator such as worker hours or project duration in days. A similar regression model can be developed to understand the difference in the project duration for different construction materials.

5.5.5. Address Other Barriers for the Adoption of Mass Timber Construction

This study majorly focuses on addressing the cost barrier for the adoption of mass timber construction, there are other important barriers which needs attention. Studies contributing to the body of knowledge regarding code changes, performance of mass timber, technical concepts, and performance-based design can result in greater adoption of mass timber.

5.6. DISCUSSION ON THE RESEARCH IMPACT

The conception stage of this study discussed benefits of using mass timber as a construction material to reduce the intensity of issues like scarcity of non-renewable resources, climate change, and wildfires. This further emerged into perceived barriers for the adoption of mass timber in the construction industry - construction cost being one of the most important factors for construction material selection, with attendant lack of familiarity leading to cost ambiguity of mass timber served as a disadvantage for its adoption. Cost ambiguity is a combined outcome of the preconception of high initial costs and unavailability of cost information for mass timber. This perception about the costs forms a hesitation amongst designers to discuss mass timber as an alternative construction material. However, mass timber has potential environmental benefits and comparable strength characteristics. Therefore, this study focused on addressing the cost barrier by providing a tool to predict costs at the conceptual design stages of a construction project. Which was envisioned to address the cost ambiguity by providing a dollar amount for using mass timber which can then be compared with the costs of using other construction materials. Providing a number for the costs will lead to change the perception of unavailability and uncertainty of cost information and promote discussions on the cost differences of using mass timber against different materials. This tool will provide a foundation for the consideration of mass timber as an alternative construction material and help designers to take its other benefits into account while analyzing the cost differences. Even though the initial material costs could appear high, the further analysis will lead to the consideration of construction duration, sustainability, and biophilic design advantages of using mass timber. This in turn will result into greater adoption of mass timber as a construction material and as discussed in Chapter 1, will lead to control of CO_2 emissions, conservation of non-renewable resources, and help in managing wildfires.

In summary, the cost prediction tool helps mass timber in getting a foot in the door and opens further discussions of its additional advantages over other construction materials like concrete and steel. This will help in increasing adoption of mass timber and result into a sustainable built environment.

5.7. SUMMARY

This chapter provides an overall summary of this study, contributions to the body of knowledge, and potential future areas of research. Each objective was summarized, outputs of these objectives, and the conclusions were drawn. Given the goal of producing a cost prediction model, research contributions and future areas for improvement and expansion of this study were recommended.

To address the cost barrier, a cost prediction tool and variables impacting cost were developed. The researcher believes that if this tool is utilized and optimized in future, an economic comparison can be made with steel or concrete at the initial stages of a construction project. Ultimately, this will help the construction industry to make informed decisions and address the ambiguity related to the costs.

REFERENCES

REFERENCES

Abdelmegid, M.A., Shawki, K.M., Abdel-Khalek H. (2015). "GA optimization model for solving tower crane location problem in construction sites." Alex. Eng. J. 54 (3) (2015) 519–526, https://doi.org/10.1016/j.aej.2015.05.011.

Abdi, H., (2007). "Partial least square regression (PLS regression)." In: Salkind, N. J. (ed.), Encyclopedia of measurement and statistics. Sage.

AbouRizk, S.M., Babey, G.M., & Karumanasseri, G. (2002). "Estimating the cost of capital projects: an empirical study of accuracy levels for municipal government projects". Canadian Journal of Civil Engineering, 29, 653-661.

Akinsola, A.O., Potts, K.F., Ndekugri I., Harris, F. C. (1997). "Identification and evaluation of factors influencing variations on building projects, International Journal of Project Management." Volume 15, Issue 4,1997, Pages 263-267,ISSN 0263-7863, https://doi.org/10.1016/S0263-7863(96)00081-6.

Alvarez, M. (2007). "The State of America's Forests." Society of American Foresters, 1–76.

American Forest and Paper Association (AF&PA) (2019). Economic Impact Infographic. Available at: https://afandpa.org/images/default-source/infographics/infographics-2019/econ-impact-graphic-2020.jpg (Accessed 01/17/2020).

American Wood Council (2015). "Today's Glulam: What Design and Building Professionals need to know for Code Conformance." Available at: https://www.awc.org/pdf/education/mat/AWC-MAT230-Glulam-2hr-150530.pdf (Accessed on: 01/18/2020).

An Sung-Hoon, Kim Gwang-Hee, Kang Kyung-In (2007). "A case-based reasoning cost estimating model using experience by analytic hierarchy process." Building and Environment, Volume 42, Issue 7, 2007, Pages 2573-2579, ISSN 0360-1323, https://doi.org/10.1016/j.buildenv.2006.06.007.

An, S. H., Kim, G. H., and Kang, K. I. (2007). "A case-based reasoning cost estimating model using experience by analytic hierarchy process." Build. Environ., 42, 2573–2579.

Anderson, T. W. (2003). "An Introduction to Multivariate Statistical Analysis." 3rd edition, John Wiley and Sons, New York.

ANSI 117-2010, (2010). American National Standard for Standard Specification for Structural Glued Laminated Timber of Softwood Species. APA – The Engineered Wood Association, Tacoma, Washington.

APA – The Engineered Wood Association (2018). "APA Case Study – First Tech Federal Credit Union: The Building That Wanted to be Mass Timber." Retrieved 23 February 2020, from https://www.structurlam.com/wp-content/uploads/2018/07/First-Tech-CU-case-study-1.pdf

APA – The Engineered Wood Association (2017). "ANSI/APA PRG 320: Standard for Performance-Rated Cross-Laminated Timber." Tacoma, WA.

APA – The Engineered Wood Association (2017). "ANSI A190.1-2017: Standard for Wood Products – Structural Glued Laminated Timber." Tacoma, WA.

APA (2013). "Proper Storage and Handling of Glulam Beams." Available at: http://www.apawood.org/buildertips/pdfs/R540.pdf

Arditi, D., and Tokdemir, O. B. (1999). "Comparison of case-based reasoning and artificial neural networks." J. Comput. Civ. Eng., 13(3), 162-169.

Asdrubali F., Ferracuti B., Lombardi L., Guattari C., Evangelisti L., Grazieschi G. (2017). "A review of structural, thermo-physical, acoustical, and environmental properties of wooden materials for building applications." Building and Environment, Volume 114, 2017, Pages 307-332, ISSN 0360-1323, https://doi.org/10.1016/j.buildenv.2016.12.033.

Associated General Contractors (AGC) of America (2019). "2019 Workforce Survey Results: National Results." Available at: https://www.agc.org/sites/default/files/WorkforceDevelopment_2019_National_Final_1.pdf (Accessed 12/23/2019).

Bode J. (1998). "Neural networks for cost estimation." Cost Engineering 1998;40(1):25–30.

Bragança, Luís, Vieira, Susana M., Andrade, Joana B. (2014). "Early Stage Design Decisions: The Way to Achieve Sustainable Buildings at Lower Costs." The Scientific World Journal 2356-6140. https://doi.org/10.1155/2014/365364

Brandner, R., Flatscher, G., Ringhofer, A., Schickhofer, G., & Thiel, A. (2016). Cross laminated timber (CLT): overview and development. European Journal of Wood and Wood Products, 74(3). https://doi.org/10.1007/s00107-015-0999-5

Carr, R. I. (1989). “Cost-estimating principles,” Journal of Construction Engineering and Management, vol. 115, no. 4, pp. 545–551, 1989.

Chin, W.W., (1998). “The partial least squares approach to structural equation modeling.” In: Marcoulides, G.A. (Ed.), Modern Methods for Business Research. Lawrence Brlbaum Associates, Mahwah, NJ, pp. 295–336.

Concrete Reinforcing Steel Institute – CRSI (2018). “Cost Comparison of Cross Laminated Timber (CLT) and Cast-in-place Reinforced Concrete Structures”, CRSI Technical Note ETND-5-18, Schaumburg, Illinois, 4 pp.

Crampton, A. (2017). “Cross-laminated Timber: An Innovative Building Material Takes Hold in Oregon.” Metroscape, Winter 2017, pages 6-12.

DLR Group (2018). “Tall With Timber: A Seattle Mass Timber Tower Case Study”, Available at: https://www.dlrgroup.com/media/735797/seattle-mass-timber-tower-book_print-5319update-spreads-reduced.pdf Accessed on: Jan 25, 2020.

Dunn, K., G., (2020). “Process Improvement using Data.” Available at: https://learnche.org/pid/PID.pdf?ddf4ec (Accessed 20 January 2020).

Eaton, R. A., and Hale, M. D. C. (1993). “Wood: Decay, Pests and Protection.” Chapman and Hall Ltd. http://www.cabdirect.org/abstracts/19940606334.html;jsessionid¼770A55932A7376C4D023914856569567?freeview¼true.

EN 1995-1-1 (2004) (English): Eurocode 5: Design of timber structures - Part 1-1: General - Common rules and rules for buildings [Authority: The European Union Per Regulation 305/2011, Directive 98/34/EC, Directive 2004/18/EC]

Environmental Information Administration (2008). “EIA Annual Energy Outlook.” DOE/EIA-0383(2008) (Washington, DC, June 2008).

Fanella, D. (2018). “Cost Comparison of Cross Laminated Timber (CLT) and Cast-in-place Reinforced Concrete Structures.”

First Tech Federal Credit Union – Swinerton Mass Timber. (2020). Retrieved 22 February 2020, from https://swinertonmasstimber.com/project/first-tech-federal-credit-union/

Fragiacomo M., Dujic B., Sustersic I., (2011). “Elastic and ductile design of multi-storey crosslam massive wooden buildings under seismic actions.” Engineering Structures, Volume 33, Issue 11, 2011, Pages 3043-3053, ISSN 0141-0296, https://doi.org/10.1016/j.engstruct.2011.05.020.

Gdub G. and VanLoan C. (1983). “Matrix Computations.” The Johns Hopkins University Press, Oxford, 1983.

Geladi P., Kowalski B. R. (1986). “Partial least-squares regression: a tutorial.” Analytica Chimica Acta, Volume 185, 1986, Pages 1-17, ISSN 0003-2670, https://doi.org/10.1016/0003-2670(86)80028-9.

Gerhards, C. C. (1982). “Effect of moisture content and temperature on the mechanical properties of wood: An analysis of immediate effects.” Wood and Fiber 14(1): 4-36.

Girshick, M.A. (1936). “Principal components.” J. Amer. Statist. Assoc., 31, 519–528.

Glued laminated timber - Production - brettschichtholz.de. (2020). Retrieved 10 April 2020, from https://www.glued-laminated-timber.com/glued-laminated-timber/production/mn_43528

Gnanadesikan, R. (1997). “Methods for statistical data analysis of multivariate observations.” New York: J. Wiley.

Golub, G., VanLoan C., (1983). “Matrix Computations.” The John Hopkins University Press, Oxford.

Gower, J. C. (1966). “Some distance properties of latent root and vector methods used in multivariate analysis.” Biometrika, 53, 325–338.

Greg C., Robert H., Chi-hsiang W. (2002). “Durability design for wood construction.” For. Prod. J. 52 (2002) 10-19.

Gunst, R. F. (1983). “Regression analysis with multicollinear predictor variables : Definition, detection and effects.” Commun. Statist. – Theor. Meth., 12, 2217-2260.

Hair, J.F., Ringle, C.M. and Sarstedt, M. (2011), “PLS-SEM: indeed a silver bullet”, Journal of Marketing Theory and Practice, Vol. 19 No. 2, pp. 139-151.

Hammond, P. G., Jones, C. (2011). “Inventory of Carbon & Energy (ICE).” http://dx.doi.org/10.1680/ener.2008.161.2.87.

Henseler, J., Ringle, C.M. and Sinkovics, R.R. (2009), “The use of partial least squares path modeling in international marketing”, Advances in International Marketing, Vol. 20, pp. 277-320.

Hotelling, H. (1933) "Analysis of a Complex of Statistical Variables into Principal Components." *Journal of educational psychology* 24.6 (1933): 417-441. *ProQuest.* Web. 24 Nov. 2019.

Hotelling, H. (1933) "Analysis of a Complex of Statistical Variables into Principal Components." *Journal of educational psychology* 24.7 (1933): 498-520. *ProQuest.* Web. 24 Nov. 2019.

Hotelling, H. (1936). “Simplified calculation of principal components.” Psychometrika, 1, 27–35.

Jaadi, Z. (2019). “A step by step explanation of principal component analysis.” Last accessed on 21 November 2019. https://builtin.com/data-science/step-step-explanation-principal-component-analysis.

Jasontran (2020). Retrieved 10 January 2020, from https://cargocollective.com/jasontran.

Jeffers, J.N.R. (1967). “Two case studies in the application of principal component analysis.” Appl. Statist., 16, 225–236.

Ji, Sae-Hyun (2011). "Cost estimation model for building projects using case-based reasoning." *Canadian Journal of Civil Engineering*, vol. 38, no. 5, 2011, p. 570+. *Gale Academic Onefile*, https://link.gale.com/apps/doc/A268310270/AONE?u=msu_main&sid=AONE&xid=a94bab02.

Jiang, R., Mao, C., Hou, L., Wu, C., Tan J. (2018). “A SWOT analysis for promoting off-site construction under the backdrop of China's new urbanization.” J. Clean. Prod., 173 (2018), pp. 225-234.

Jin, R., Cho, K., Hyun, C., & Son, M. (2012). “MRA-based revised CBR model for cost prediction in the early stage of construction projects.” Elsevier. doi:10.1016/j.eswa.2011.11.018.

Joliffe, I. T. (1986). “Principal Component Analysis.” Springer-Verlag New York.

Jong, S. (1992). “Object-oriented cost estimation system.” Master thesis, Department of Civil Engineering, National Taiwan University, Taiwan, ROC (in Chinese).

Kaming, P. F., Olomolaiye, P.O., Holt, G. D. & Harris, F. C. (1997). “Factors influencing construction time and cost overruns on high-rise projects in Indonesia, Construction Management and Economics.” 15:1, 83-94, DOI: 10.1080/014461997373132

Karacabeyli, E., Douglas, B., Forest Products Laboratory (U.S.), FPInnovations (Institute), & Binational Softwood Lumber Council. (2013). “CLT handbook: Cross-laminated timber.” Pointe-Claire, Québec: FPInnovations.

Karhunen K. (1947). “Uber lineare Methoden in der Wahrscheinlichkeitsrechnung.” Annales Acudemiae Scientiarum Fennicae, Series A, 137.

Kim Gwang-Hee, An Sung-Hoon, Kang Kyung-In (2004). “Comparison of construction cost estimating models based on regression analysis, neural networks, and case-based reasoning.” Building and Environment, Volume 39, Issue 10, 2004, Pages 1235-1242, ISSN 0360-1323, https://doi.org/10.1016/j.buildenv.2004.02.013.\

Kim, G.-H., An, S.-H., and Kang, K.-I. (2004). “Comparison of construction cost estimating models based on regression analysis, neural networks, and case-based reasoning.” Build. Environ., 39, 1235–1242.

Kim, Kyong Ju and Kim, Kyoungmin (2010). “Preliminary Cost Estimation Model Using Case-Based Reasoning and Genetic Algorithms.” Journal of Computing in Civil Engineering, 11/2010, Volume 24, Issue 6, pp. 499–505.

Kinzey, B., Kim, S. & Ryan, J. (2002). “The Federal Buildings Research and Development Program: A Sharp Tool for Climate Policy.” ACEEE Buildings Summer Study 2002, Pacific Grove. 11 vols.

Kolenikov, S., & Angeles, G. (2004). “The use of discrete data in PCA: theory, simulations, and applications to socioeconomic indices.” Measure Evaluation October 20th.

Koo, Choong-Wan (2010). "A study on the development of a cost model based on the owner's decision making at the early stages of a construction project." *International Journal of Strategic Property Management* 14.2 (2010): 121-37. *ProQuest.* Web. 14 Nov. 2019.

Laguarda-Mallo, M. F., & Espinoza, O. (2015). "Awareness, perceptions and willingness to adopt Cross-Laminated Timber by the architecture community in the United States" Journal of Cleaner Production, 2015(94), 198–210.

Laguarda-mallo, M. F., & Espinoza, O. (2018). "Awareness , Perceptions and Willingness to Adopt CLT by U . S . Engineering Firms" BioProducts Business 3(1), 1-14.

Laguarda-Mallo, Maria & Espinoza, Omar. (2016). “Cross-Laminated Timber Vs. Concrete/Steel: Cost Comparison Using A Case Study.”

Langat daisy chelangat rono. (2018); “Microcredit and its Relationship to The Growth of Small and Medium Enterprises in Konoin Subcounty, Kenya.” Int. J. of Adv. Res. 6 (4). 961-968] (ISSN 2320-5407).

Lawson, M., Ogden, R., and Bergin, R., (2012). “Application of Modular Construction in High-Rise Buildings.” J. Archit. Eng., 2012, 18(2): 148-154.

Leach, A.R. (2001). “Molecular modeling: Principles and applications.” Pearson Education Ltd. Harlow, England, 2001.

Le-Hoai, L., Lee, Y. D., & Lee, J. Y. (2008). “Delay and cost overruns in Vietnam large construction projects: A comparison with other selected countries.” doi:10.1007/s12205-008-0367-7.

LEVER Architecture, Albina Yard, (2020). Retrieved 10 March 2020, from https://leverarchitecture.com/projects/albina_yard.

Li H. (1995). “Neural networks for construction cost estimation.” Building Research and Information 1995;23(5):279–84.

Lindsay, S. (2002). “A tutorial on principal component analysis.” Details available at <csnet.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf>, last accessed on 20 November 2019.

Loéve M. (1948). “Fonctions aleatoires de seconde ordre.” in P. Levy (Editor), Processes Stochastiques et Mouvement Brownien, Hermann, Paris, 1948.

Lowe, D. J., M. W. Emsley, and A. Harding (2006). “Predicting construction cost using multiple regression techniques.” J. Constr. Eng. Manage. 132 (7): 750–758. https://doi-org.proxy1.cl.msu.edu/10.1061/(ASCE)0733-9364(2006)132:7(750).

Maitra, S. & Yan, J. (2008). “Principle Component Analysis and Partial Least Squares: Two Dimension Reduction Techniques for Regression.”

Mallo, M. F. L., & Espinoza, O. (2014). "Outlook for cross-laminated timber in the United States" *BioResources*, *9*(4), 7427–7443.

Mallo, M. F. L., & Espinoza, O. A. (2016). “Cross-Laminated Timber vs. Concrete/steel: Cost comparison using a case study.” In WCTE 2016 - World Conference on Timber Engineering (WCTE 2016 - World Conference on Timber Engineering). Vienna University of Technology.

Mandel J. (1982). “Use of the singular value decomposition in regression analysis.” American Statistician, 36 (1982) 15-24.

Mardia, K. V., Kent, J. T. & Bibby, J. M. (1980). “Multivariate Analysis.” Academic Press, London.
Martens, H, & Naes, T. (1989). “Multivariate Calibration.” London: Wiley.

Matos, G., R. (2017). “Use of raw materials in the United States from 1900 through 2014: U.S. Geological Survey Fact Sheet 2017–3062.”, 6 p., https://doi.org/10.3133/fs20173062. [Supersedes Fact Sheet 2012–3140.]

McCaffer, R. (1975). “Some examples of the use of regression analysis as an estimating tool.” Quant. Surveyor, 32, 81–86.

McLain, R., Breneman, S. (2019). “Fire Design of Mass Timber Members Code Applications, Construction Types and Fire Ratings.” WoodWorks Wood Products Council. Available at: https://www.woodworks.org/wp-content/uploads/Wood_Solution_Paper-Fire-Design-of-Mass-Timber-Members-WoodWorks-Apr-2019.pdf.

McGraw-Hill (MGH), (2011). “Smart Market Report: Prefabrication & Modularization at World of Modular, 2011.”

McKeever, D. B.; Elling, J. (2015). “Wood Products Other Building Materials Used in New Residential Construction in the United States.” Tacoma, WA: APA - The Engineered Wood Assoc. 2015; 131 p.

Modular Building Institute (MBI), (2016). “Modular Advantage for the commercial modular construction industry.” 59 pages, Virginia, USA.

Moens, Y., Lanz, F., Doherr, M. G., & Schatzmann, U. (2003). "A comparison of the antinociceptive effects of xylazine, detomidine and romifidine on experimental pain in horses" Veterinary Anaesthesia and Analgesia, 30(3), 183–190.

Moghadam, M., Alwisy, A., and Mohamed, A. L. (2012). “Integrated BIM/Lean Base Production Line Schedule Model for Modular Construction Manufacturing”, Construction Research Congress 2012 Construction Challenges in a Flat World, 1271-1280.

Morcous, G., Rivard, H., and Hanna, A. M. (2002). “Case-based reasoning system for modeling infrastructure deterioration.” J. Comput. Civ. Eng., 16(2), 104–114.

MyTiCon Timber Connectors Inc. (2018). “Canadian lifting device design for Cross Laminated Timber (CLT).” Surrey, BC. Available at: https://www.cascadejoinery.com/wp-content/uploads/2018/03/cross-laminated-timber.pdf

MyTiCon Timber Connectors Inc. (2019). “Timber Connection Solutions.” Available at: https://mtcsolutions.com/wp-content/uploads/2019/06/Product_Brochure_11102019.pdf.

Nail-Laminated Timber U.S. Design and Construction Guide v1.0. (2017). Binational Softwood Lumber Council (BSLC), Surrey, BC.

Oregon Employment Department, based on North American Industry Classification System (NAICS) and U.S. Census data, in cooperation with the Oregon Department of Forestry and the Oregon Forest Resources Institute. October 2016.

Otto M. and Wegscheider W. (1985). “Anal. Chem.” 57 (1985) 63.

Pearson, K. (1901). “On lines and planes of closest fit to systems of points in space.” *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *2*(11), 559–572. https://doi.org/10.1080/14786440109462720

Peng Bo, Forest Lee Flager, Jiaao Wu (2018). “A method to optimize mobile crane and crew interactions to minimize construction cost and time, Automation in Construction.” Volume 95, 2018, Pages 10-19, ISSN 0926-5805, https://doi.org/10.1016/j.autcon.2018.07.015.

Peng, Xiaosong & Lai, Fujun. (2012). “Using Partial Least Squares in Operations Management Research: A Practical Guideline and Summary of Past Research.” Journal of Operations Management. 30. 10.1016/j.jom.2012.06.002.

Perera S, Watson I. (1998). “Collaborative case-based estimating and design.” Advances in Engineering 1998;29(10):801–8.

Phatak, A., & de Jong, S. (1997). “The geometry of partial least squares.” Journal of Chemometrics, 11, 311–338.

Quesada-Pineda, H., Smith, R., & Berger, G. (2018). “Drivers and barriers of cross-laminated timber (CLT) production and commercialization: A case of study of Western Europe’s CLT industry.” BioProducts Business, 3(3), 29–38.

Rao, C.R. (1964). “The use and interpretation of principal component analysis in applied research.” Sankhya A, 26, 329–358.

Rencher, A.C. (1995). “Methods of Multivariate Analysis.” Wiley, New York.

Ritter, M. A. (1990). “Timber bridges.” Forest Products Laboratory, Forest Service, Madison, USA, 1990.

Ritter, Michael A. (1990). “Timber Bridges: Design, Construction, Inspection, and Maintenance.” Washington, DC: 944 p.

Roos, A., Woxblom, L. & McCluskey, D. (2010). “The influence of architects and structural engineers on timber in construction – perceptions and roles.” Silva Fennica 44(5): 871–884.

Salama, T., Salah, A., and Moselhi, O. (2016). “Alternative Scheduling and Planning Processes for Hybrid Offsite Construction.” Proceedings of the 33rd International Symposium on Automation and Robotics in Construction (ISARC 2016), Auburn, Alabama, USA.

Salama, T., Salah, A., and Moselhi, O., (2017). “Integration of Offsite and Onsite Schedules in Modular Construction.” Proceedings of the 34th International Symposium on Automation and Robotics in Construction (ISARC 2017), Taipei, Taiwan.

Sarthre, R. and O’Connor J., (2010). “A Synthesis of Research on Wood Products and Greenhouse Gas Impacts.” FPInnovations. 117p (Technical report TR-19R)

Saxena, R. (2018). “Predictive model for cost estimation of cross laminated timber construction.” M.S. report, East Lansing, MI: Michigan State University.

Schmidt, J., & Griffin, C. (2013). “Barriers to the design and use of cross-laminated timber structures in high-rise multi-family housing in the United States.”

Schoenborn, J. M. (2012). “A Case Study Approach to Identifying the Constraints and Barriers to Design Innovation for Modular Construction.” Master of Science in Architecture, Faculty of the Virginia Polytechnic Institute and State University.

Schwarzmann, G., Hansen, E., & Berger, G. (2017). "Cross-laminated Timber in North America : What can we learn ?" BioProducts Business 3(7), 81-91.

Shlens, J. (2005). “A tutorial on principal component analysis.” Available: http://www.cs.cmu.edu/∼elaw/papers/pca.pdf. Last accessed on November 21, 2019.

Skitmore, R., M. and Patchell, B., R., T. (1990). “Developments in contract price forecasting and bidding techniques.” In: Skitmore RM, Marston VK, editors. Cost modelling. London: E& FN Spon; 1990. p. 53–84.

Smith, L.I. A tutorial on principal components analysis. Cornell Univ. USA 2002, 51, 52.

Stephanie, E. (2003). “Slovin's Formula Sampling Techniques.” Houghton-Mifflin, New York, USA.

Structure Craft (2018). “Dowel Laminated Timber - The All Wood Panel - Mass Timber Design Guide.” Available at: https://structurecraft.com/blog/dowel-laminated-timber-design-guide-and-profile-handbook (Accessed: October 20, 2019).

Swee Lean Chan Corresponding author & Moonseo Park (2005). “Project cost estimation using principal component regression, Construction Management and Economics”, 23:3, 295-304, DOI: 10.1080/01446190500039812.

Sznopek, J., Brown, W. (1998). “USGS Fact Sheet FS-068-98, June 1998.” Available at: https://pubs.usgs.gov/fs/1998/0068/report.pdf (Accessed August 12, 2019).

Tenenhaus, M. (1998). “La régression PLS.” Paris: Technip.

Ter Braak, C. J. F., & de Jong, S. (1998). “The objective function of partial least squares regression.” Journal of Chemometrics, 12, 41–54.

Think Wood (2020). “Designing Sustainable, Prefabricated Wood Buildings” Available at: <https://www.thinkwood.com/our-ceus/designing-sustainable-prefabricated-wood-buildings> (Accessed: 15 March 2020).

Trost, S., M. and Oberlander, G., D. (2003). “Predicting accuracy of early cost estimates using factor analysis and multivariate regression.” Journal of Construction Engineering and Management 2003;129(2):198–204.

UNDESAP, United Nations, Department of Economic and Social Affairs, Population Division (2013). “World Population Prospects: The 2012 Revision, Highlights and Advance Tables.” ESA/P/WP.228. (Accessed 10/12/2019).

Veerasamy, R., Rajak, H., Jain, A., Sivadasan, S., Christapher, P. V., & Agrawal, R. (2011). “Validation of QSAR Models - Strategies and Importance.” Int J Drug Design and Discov. 2. 511-519.

Wali, M., Evrendilek, F., Fennessy, M. (2010). “The Environment: Science, Issues, and Solutions.” Boca Raton: CRC Press, https://doi.org/10.1201/9781420007336

Werner, F. & Scholz, R. (2002). “Ambiguities in decision-oriented Life Cycle Inventories: The Role of mental models.” International Journal of Life Cycle Assessment. 7. 330-338. 10.1007/BF02978680.

White, R. H. (1983), "Use of Coatings to Improve Fire Resistance of Wood," Fire Resistive Coatings: The Need for Standards. ASTM STP 826, Morris Lieff and F. M. Stumpf, Eds., American Society for Testing and Materials, 1983, pp. 24-39.

White, R.H., Nordheim, E.V. Charring rate of wood for ASTM E 119 exposure. Fire Technol 28, 5–30 (1992). https://doi.org/10.1007/BF01858049.

White, R. H., and Woeste, F. E. (2013). "Post-fire analysis of solid-sawn heavy timber beams." Artic. Struct. Perform. (2013) (accessed January 22, 2020), http://www.structuremag.org/.

Winchester, N., and Reilly, J., M. (2019). "The economic and emissions benefits of engineered wood products in a low-carbon future." Energy Economics, online first (doi: 10.1016/j.eneco.2019.104596) (https://doi.org/10.1016/j.eneco.2019.104596).

Wold S., Esbensen K., Geladi P. (1987). "Principal component analysis, Chemometrics and Intelligent Laboratory Systems." Volume 2, Issues 1–3, 1987, Pages 37-52, ISSN 0169-7439, https://doi.org/10.1016/0169-7439(87)80084-9.

Wold S., Ruhe A., Wold H. and Dunn W. (1984). "SIAM" J. Sci. Stat. Comput., 5 (1984) 735.

Wold, H. (1966). "Estimation of principal components and related models by iterative least squares." In P.R. Krishnaiaah (Ed.). Multivariate Analysis. (pp.391-420) New York: Academic Press.

WoodWorks Wood Products Council (2020). "Mass Timber Projects In Design and Constructed in the US (March 2020)." Available at : https://www.woodworks.org/publications-media/building-trends-mass-timber/ (Accessed on April 17, 2020).

Yau, N. J., & Yang, J. B. (1998). "Case-Based Reasoning in Construction Management." *Computer-Aided Civil & Infrastructure Engineering*, *13*(2), 143. https://doi-org.proxy1.cl.msu.edu/10.1111/0885-9507.00094

Zima, K. (2015). "The Case-based Reasoning Model of Cost Estimation at the Preliminary Stage of a Construction Project." Procedia Engineering. 122. 57-64. 10.1016/j.proeng.2015.10.007.