Modeling Productivity losses Due to Change Orders

Ali Emamifar

A Thesis

In the Department

of

Building, Civil and Environmental Engineering

Presented in Partial Fulfillment of the Requirements

For the Degree of

Master of Applied Science (Civil Engineering) at

Concordia University

Montreal, Quebec, Canada

March 2019

CONCORDIA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: Ali Emamifar

Modeling Productivity Losses Due to Change Orders

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

MASTER of APPLIED SCIENCE (Civil ENGINEERING)

(Department of Building, Civil and Environmental Engineering)

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

Signed by the final examining committee:

Dr. F. Nasiri	Chair
Dr. A. Youssef	External to Program
Dr. S. H. Han	Examiner
Dr. F. Nasiri	Examiner
Dr. O. Moselhi	Thesis Supervisor

Approved by Dr. A. Bagchi, Chair .

Graduate Program Director

Dr. A. Asif, Dean .

Faculty of Engineering and Computer Science

Date: 11/03/2019

Abstract

Modeling Productivity Losses Due to Change Orders

Ali Emamifar, 2019

Change orders are an integral part of construction projects regardless of project size or complexity. Changes may cause interruption to the unchanged scope of work and working conditions and, if poorly managed, may be detrimental to project success. Many studies have been carried out to quantify the impact of change orders on construction labour productivity, with varying degrees of accuracy and variables considered. These studies reveal that quantifying loss of productivity due to change orders is not an easy task and requires a comprehensive and holistic method.

There are several methods for quantifying loss of productivity, such as measured mile analysis (MMA) and the total cost method (TCM). Although measured mile analysis (MMA) is a well-known and widely accepted method for quantifying the cumulative impact of change orders on labour productivity, it is not readily applicable to many cases. In this research two models were developed to quantify losses arising from change orders. The first model does not account for the timing of change orders, but the second model considers the timing of change orders on labour productivity. Two models were developed and tested utilizing artificial neural networks and two sets of data collected by others in that field.

The two datasets were statistically analyzed and preprocessed in order to transfer the data to normal distribution form and eliminate insignificant variables considered in their development. Using best subset regression, a total of seventeen variables were reduced to nine variables accordingly. Also, the study datasets were categorized into three types of timing periods; early change, normal change and late change to create the timing model. This was implemented to enable a comparison with models developed by others.

Three types of artificial neural network techniques were experimented with and evaluated for possible use in the developed models. These three types are Feed Forward Neural Network, Cascade Neural Network, and Generalized Regression Neural Network. Candidate techniques were evaluated and analyzed by neural network parameters and analysis of variance (ANOVA) to select the most efficient type of neural networks, and subsequently using it to develop two models; one considers timing and the second does not. The analysis performed led to the selection of the cascade neural network for the development of the two models productivity losses due to change orders.

The developed models were tested and validated utilizing several actual cases reported by others. The models were applied to a number of cases and the results were compared to those generated by frequently cited models to demonstrate their accuracy. The comparison outcome showed that the developed models can generate more accurate and satisfactory results than those of reported in previous studies.

Acknowledgement

I would like to first thank GOD for granting me the health, peace, and wellness required to complete this course of study.

I thank my supervisor, Professor Osama Moselhi for his supervision and support in guiding me through the Master's program and providing me with his valuable insights.

I thank my colleagues and friends, especially Dr. Mohamed Elmasry and Dr. Ali Zaker for their companionship and support. Especially I would like to express my gratitude to Mr. Sasan Golnaraghi for his patience and support. He is the best person one could ask for, and he is truly showed someone can be supportive, helpful and patient.

I deeply appreciate my cousin Dr. Mohammad Tavallaeinejad and Ms. Mehrnaz Tavallaei for all the love, patience and encouragement in this adventure.

Finally, it is my deepest pleasure to thank the two persons who always strived me an educated professional. I will always be indebted to my beloved parents. Without your support and blessing, I would have never been able to attain this accomplishment.

Contents

List of Figures

List of Tables

Chapter.1 Introduction

1.1. Overview

The construction industry has played a distinguished role in the North American industry. The construction industry is an essential contributor to Canada's economy, accounting for approximately 7% of the nation's gross domestic product (GDP) in 2017 (Statista, 2018). Around 1.5 million people in Canada were employed in the construction field in 2017 (Statistics Canada, 2018). Furthermore, in 2016, private construction in the United States made up roughly $899 billion of construction industry spending. It is estimated that new construction investments will reach over $1.4 trillion by 2021 (Statista, 2016).

Essentially, every construction contract contains a "changes clause" that defines the process for identifying and documenting changes. Change orders can challenge owners and contractors when construction begins before design completion, and the scope of the project is adjusted along the way. Change orders can often cause owners and contractors serious problems such as cost overruns and costly disputes. These can get worse if owners and contractors have a minimal understanding and appreciation of the impact of changes on project productivity (Moselhi, Ihab Assem, & El-Rayes, 2005). Changes themselves might not cause productivity losses, but during the change orders procedure, the changed work can affect the unchanged work, and as a result, a loss of productivity may occur (Thomas & Napolitan, 1995).

This study intends to focus on estimating loss of productivity caused by change orders. The proposed methodology leads us to find accurate solutions by processing the selected variables with a selected technique.

1.2. Problem Statement

Change orders are frequently encountered in construction projects. Contractors often face extra expenses to their projects arising from change orders. They have to negotiate and sometimes struggle with owners and designers to compensate for these additional expenses which frequently leads to disputes and lengthy litigations. Several methods have been proposed to help the relevant parties to resolve their conflicts. However,when compared to new techniques for quantifying impacts of change orders on construction productivity, these methods are obsolete and inefficient. They cannot accurately predict loss of productivity due to the difficulty in their models.

Previous research has attempted to establish several method or techniques for quantifying loss of productivity caused by change orders. These methods have aimed to identify the direct effects of change orders on construction productivity, define some change order variables which impact productivity, and include statistical methods, regression, and neural networks. Chapter two will look at the limitations of these methods in detail. For effective prediction of loss of productivity due to the change orders, new techniques must be taken into consideration. As such, there is a need for an optimum model for quantifying loss of productivity due to change orders and minimizing the the absolute and average error.

1.3. Research Objectives and Scopes

The overall goal of this research is to quantify loss of productivity caused by change orders using artificial intelligence (AI) modeling techniques. To achieve this objective, several sub-objectives were considered:

1. Reviewing previous studies which considered loss of productivity, change orders and the impacts on productivity caused by change orders;

2. Distinguishing between significant variables which have more effects on construction labour productivity;

3. Developing a model to predict loss of productivity caused by change orders; and

4. Validating the developed model by employing real case studies.

This research focuses on mechanical, architectural, electrical and civil projects with lump sum, unit price, fixed price and cost plus contracts. The minimum value of the original contract included in this data sets was $80000, while most have an original amount of more than $1 million, which could be a practical sum for small and heavy projects. The characteristics of the data sets are discussed in more details in Chapter 4.

1.4. Research Methodology

In order to achieve the aforementioned objectives, the methodology of this research includes the following four main phases (Figure 1):

1. Data collection;

2. Data normalization and variable selection;

3. Model development; and

4. Model validation.

Step 1: The data sets used in the data collection stage of this research are collected from two previous studies, namely the Leonard (1988) and Assem (2000) models. A total of 123 data sets are generated by combining these two data sets. A datasets consists of the input parameters and their associated output. This can be considered adequate data volume for developing a model to quantify loss productivity.

Step 2: In the data normalization and variable selection stage, the data sets are normalized so they can be transferred to a normal distribution. Some of the variables have a large variance and others have a small variance.This study employs best subset regression to eliminate those variables from datasets that are insignificant and select the most significant variables.

Step 3: To reach an optimal model, this research employs three artificial neural network (ANN) techniques: the feedforward neural network (FNN), cascade neural network (CNN) and generalized regression neural network (GRNN).

Step 4: Developed model is then validated against real case studies to evaluate performance and accuracy and the predicted datasets are compared with actual data sets and previous studies such as Ibbs' model (2005) and Leonard's model (1988).

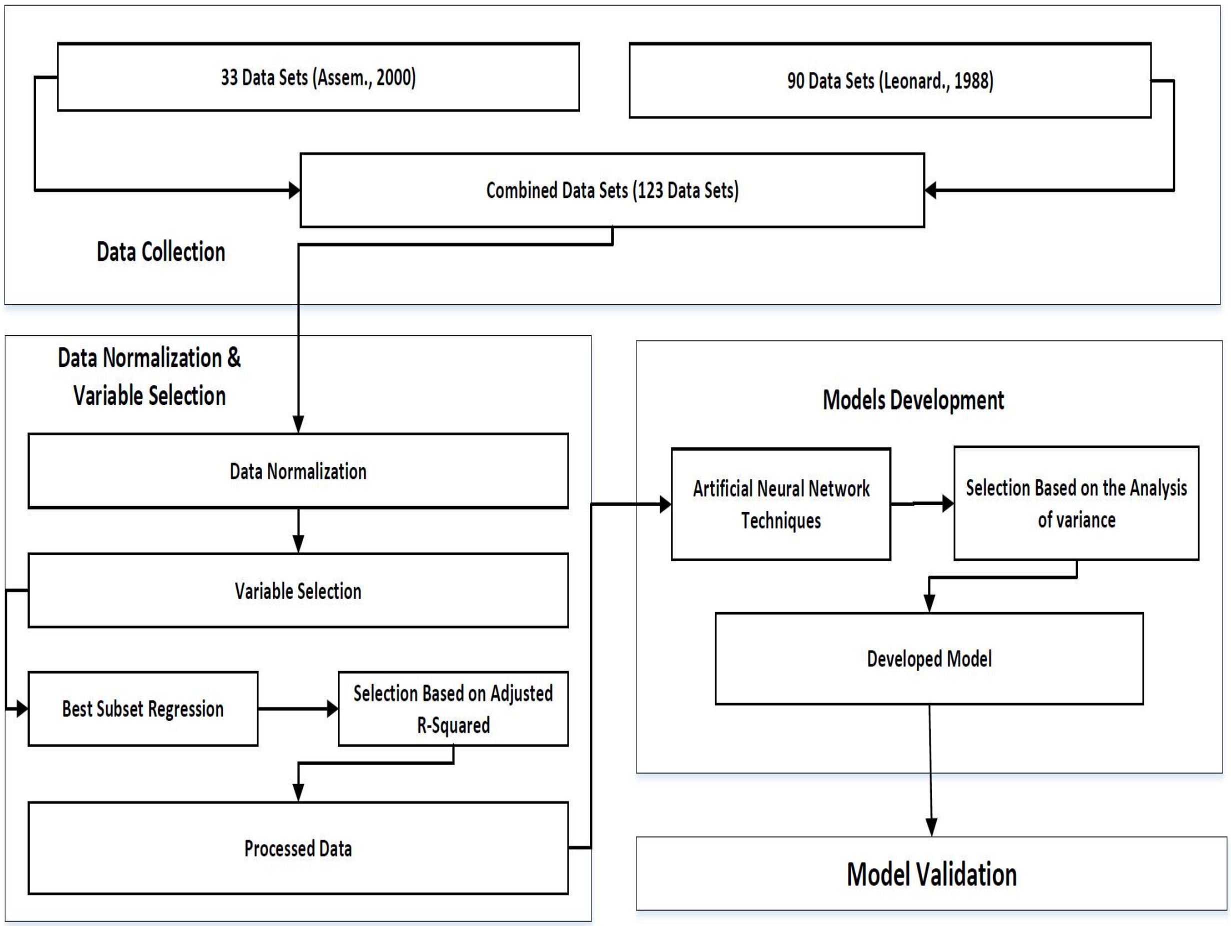


Figure 1. Developed Model General Overview.

1.5. Thesis Organization

This research is organized into seven chapters. Chapter 1 explains the basic principles of change orders and loss of productivity. Chapter 2 presents a literature review of previous work conducted in the areas of change orders and loss of productivity. Chapter 3 outlines the research methodology. It provides details regarding the data normalization, variable selection and model development to predict loss of productivity caused by change orders.

Chapter 4 illustrates the data collection. In this chapter, a brief description of the architectural, civil, mechanical and electrical data sets included is presented. Also, it highlights the data normalization and variable selection used to implement the three models. Statistical analysis is used to determine variable significance or insignificance to improve model accuracy. These variables are used to quantify loss of productivity caused by change orders. Furthermore, in order to develop the timing model to compare the current model with Ibss timing model (2005), the combined datasets are divided into the three separate timing datasets. The combined datasets are devided into early change datasets, normal change datasets and late change datasets.

Chapter 5 explains the model development and describes the three developed models, namely the feedforward neural network (FNN), cascade neural network (CNN) and generalized regression neural network (GRNN). Analysis of variance (ANOVA) is used to select the superior technique.

Chapter 6 outlines the model validation. In this section, the developed model is compared with those of previous studies, such as Ibbs (2005) and Leonard's (1988) models in terms of the absolute error, average error and actual loss of productivity. In the final chapter, conclusions are made, followed by remarks on the application of results and recommendations for further research.

1.6. Summary:

This research reveals that change order has played a prominent role in the construction industry. Change orders may create lengthy disputes among the parties involved in a construction contract. This research considers 123 data sets and 17 variables drawn from two previous studies. It implements the following three artificial neural network techniques: the feedforward neural network (FNN), cascade neural network (CNN) and generalized regression neural network (GRNN). The results of the model development show that the developed model is more precise and accurate than those of previous studies.

This study consists of seven chapters. Chapter 1 includes an introduction to the research, highlighting the research problem, scope, objective and methodology. Chapter 2 includes a literature review of past studies concerning change orders and loss of productivity quantification. Chapter 3 outlines the methodology. Chapter 4 shows the data collection. Chapter 5 details the model development. Chapter 6 outlines the model validation. The final chapter includes the conclusion and recommendations.

CHAPTER 2. LITERATURE REVIEW

2.1. Introduction

This chapter discusses previous studies pertinent to change orders as well as issues related to construction labour productivity. Even though this study includes projects where change orders were the main causes of loss of labour productivity, other productivity issues are meticulously discussed before proceeding to the next chapter. In order to comprehend the impacts of change orders on labour productivity, this research aims to present in-depth knowledge of all factors related to change orders. A review of recent studies reveals that numerous factors, including delays and disruptions, have asignificant effects on labor productivity. It also shows that change orders are one of the main causes of disruptions in construction productivity. The comprehensive literature review presented in this chapter is divided into two parts. First, productivity and change orders as concepts are intrduced and discussed. This is followed by a presentation of previous methods for quantifying change order impacts on labour productivity.

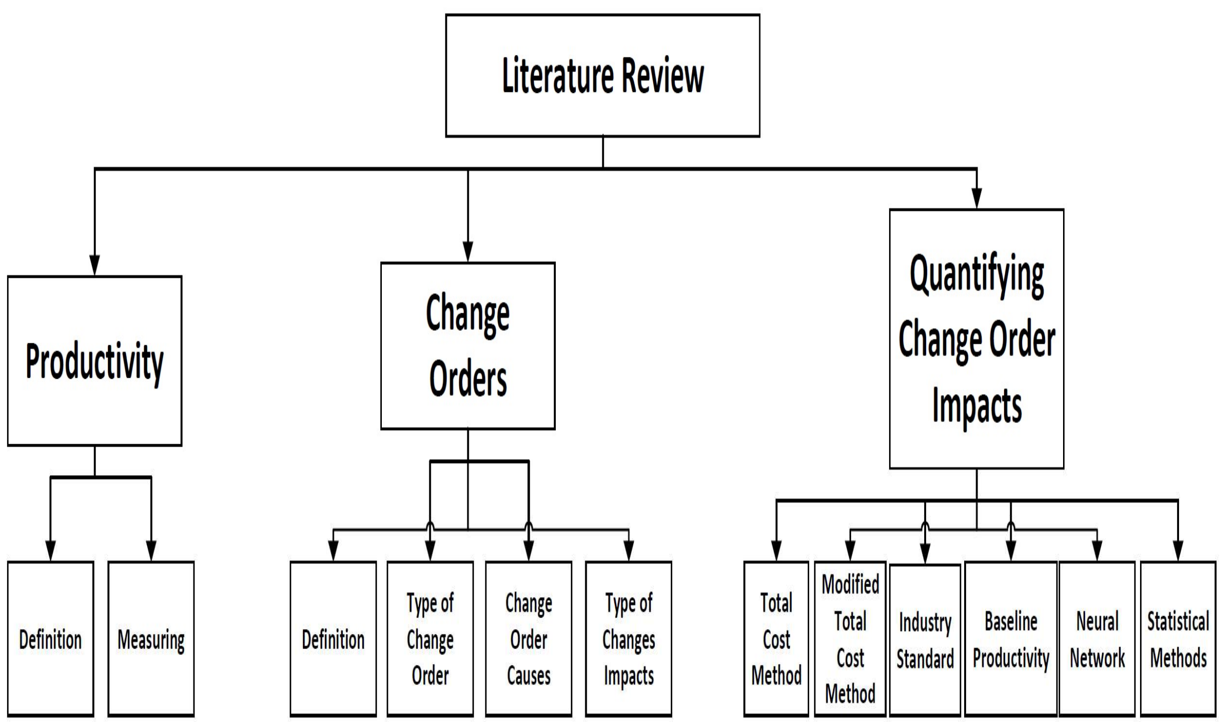


Figure 2. Chapter 2 Flow Chart

2.2. Productivity

Up to now, there has been no exact definition of productivity in studies of construction management. The Concise Oxford Dictionary offers three definitions for productivity as follows:

1. The power of being productive is the force behind production itself;

2. Efficiency is a measure of how well a given set of factors are utilized; and

3. Rate is a measure of the output of the factors of production over a defined period of time. The term "productivity" is mainly used to indicate a relationship between the outputs and their associated inputs in a production process (Yi & Chan, 2014).

Another definition of productivity provided by the American Association of Cost Engineers (AACE) is based on a relative measurement of efficiency: productivity is a measure of production output relative to labour input (AACE, 2013).

2.2.1. Measuring Productivity

According to Equation 2.1, measuring productivity losses is accomplished using production rate and unit rate. This can easily be compared to the standard rate for calculating productivity and productivity losses percentage (Lee, 2007).

$$\text{Productivity Ratio} = \frac{(\text{A}'\text{s Production Rate})}{(\text{Standard Production Rate})} = \frac{(\text{A}'\text{s} \frac{\text{Output}}{\text{Input}})}{(\text{Standard } \frac{\text{Output}}{\text{Input}})} = \frac{(\text{A}'\text{s Units})}{(\text{Standard Units})} \qquad \textbf{(Equation 2.1)}$$

2.3. Change orders

According to the Oxford Dictionary of Construction, a change order is defined as "an alteration to that which was planned, contracted, priced, or proposed. Changes to work are normally authorized by the client representative or project administrator using a change order" (Oxford Dictionary of Civil Engineering, 2012). Schwartzkopf defines a change order as "a formal contract modification incorporating a change into the contract" (Schwartzkopf, 1995). Additionally, Hester defines a change order as "any written alteration in the specification, delivery point, rate of delivery, contract period, price, quantity or other contract provisions of an existing contract, accomplished by mutual action of the parties to the contract" (Hester et al. 1991).

Every construction contract contains a “changes clause” which defines the process for identifying and recording changes. Owners and contractors might face disagreements regarding the quantification of change orders concerning things such as cost, scope, delay, differing site conditions, time of performance, etc. This is called a dispute. In such a case, the disagreement regarding time, money or both has not yet been formalized into a request for a contract adjustment or lawsuit. (Serag, 2006).

In this study, scope changes are considered to be change orders based on the data set of this research. Scope change is defined as customer-directed changes that requires the alteration of a project’s cost or schedule. The types of scope change include:

1. Engineering change;

2. Quantity change;

3. Support change; and

4. Schedule change (BusinessDictionary, 2018).

2.3.1. Different Types of Change Orders

Change orders can be classified into five categories as follows: (Brams and learner, 1996; O’Brien, 1998)

1. Bilateral: A type of change order which is approved by construction parties and thus reducing the chance of disputes or claims between them;

2. Unilateral: A type of change order which is requested by owner and executed by contractor in accordance with the relative contractual clauses. In this type of change order, the dispute causes job noncompletion and a risk of claims;

3. Formal: A written document to ensure the contractor executes work changes within the general scope and appeals for equitable adjustment;

4. Informal: An oral format which is given to the contractor mostly as the result of a defective specification. This type is also referred to as a constructive change order; and

5. Cardinal: a change order or series of change orders which are beyond the scope of the contract. Failure to perform them does not legalize violation of contract.

2.3.2. Change Orders Causes

Hasegawa mentioned the following seven reasons for the implementation of change orders (Hasegawa, 1995):

1. New or revised functional requirements or desires may lead customers or clients to request permission for change orders;

2. Criteria change: Any alteration in building or design code after the award or after the construction has begun may necessitate change orders;

3. Design deficiency for which the designer was not responsible;

4. Design error or omission for which the designer was held liable;

5. Additive bid item;

6. Unforeseen conditions; and

7. Initiated value engineering change.

In addition, the National Electrical Contractors Association reported the following reasons for the occurrence of change orders (NECA, 2000):

1. Capital shortage: The parties of the contract are faced with a shortage of capital and a high-interest rate on borrowed money. A very small margin is set for the contingency allowance in the bids or budget;

2. Challenges in permits and raising money: Owners face difficulties in raising money to finance the construction project and get construction permits, resulting in delays in starting the design process. As a result, they rush the design, causing errors and omissions;

3. Tight schedules: Owners may set tight construction schedules to complete projects in less than standard time to make the facility a profit-producing asset instead of a liability under construction;

4. Overestimating and underestimating the bid: Some owners may take bids on incomplete or inadequate plans rather than wait until the designs are complete, assuming it will cost less to settle the claims. These owners assume that cost increase due to inflation during a long design phase will be greater than the cost of the claims and change orders that result from an incomplete design and

earlier start. Similarly, some contractors may underestimate the bids and believe they will get their profit from the change orders. In other words, they view claims as their profit center;

5. Complexity of work: New products, assemblies and construction techniques add to the complexity of coordinating the work of multiple contractors, which might result in delays, restricted access and the stacking of trades; and

6. Changing market conditions: Changes in population and markets often result in alterations to the owner's needs between the conceptual stage of the project and its completion. This factor may lead not only to change orders to accommodate the owner's needs but also to the acceleration or deceleration of payments if the owner decides they do not need the facilities as soon as they had expected.

2.3.3. Types of changes impacts

When projects face change orders, the impacts of scope changes can be direct or indirect and cause loss of productivity and cumulative impact. In the worst-case scenario, if several change orders occur during a project, it might be difficult or impossible to individually calculate and measure their impacts.

2.3.3.1 Direct

Direct impacts of change orders can be defined as impacts which affect unchanged work and are, to a significant degree, measurable and predictable by experienced professionals. Unchanged work is defined as contract work which is not covered by a specific contract change order (Jones, 2001).

Adding or eliminating certain activities may cause changes in an established sequence of work. Therefore, the new sequence of work may take more time to complete than the original. Such impacts should be considered as thoroughly as possible when forward-pricing a change, so that time extensions and increased contract costs can be requested and granted (Lee, 2007). The U.S Army has suggested questions to help identify the direct impacts of changes as follows:

1. Has any activity been moved from a favorable to an unfavorable weather season?

2. Are there now more activities in progress at a given time than before the revision?

3. Have any activities slipped to the extent that significant phases of the work will not be accomplished before overriding factors (such as winter, high water stages, unavailability of the site) prevent its completion, thus making it necessary to defer a portion of the work until the next favorable season? (US Army Corps of Engineers, 1979).

As shown in Figure 3, the direct cost of change order is as follows: (Moselhi, 2003);

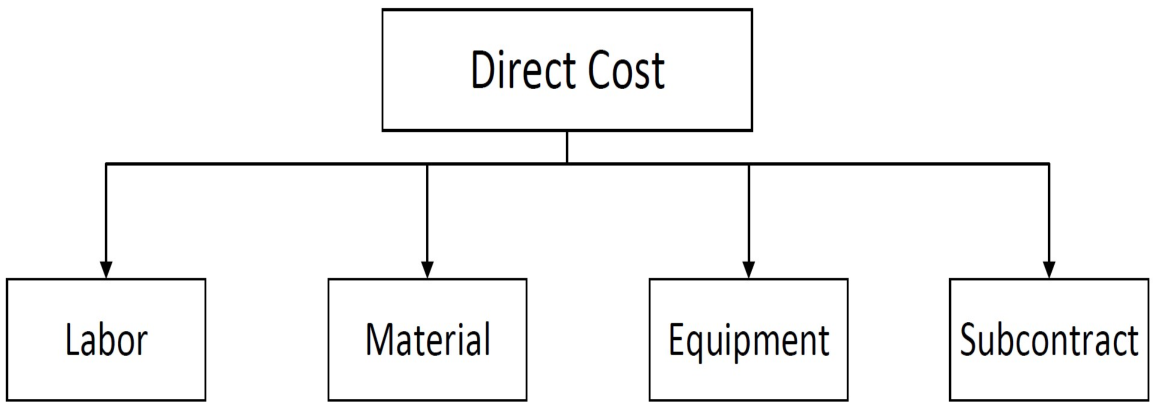


Figure 3. The direct cost of change orders

2.3.3.2 Indirect Impacts

Indirect impacts are defined as changes to indirect costs such as job site overhead, interest, and profit. A change might extend the duration of work items which in turn extends the overall duration of the project. This requires the project superintendent and project manager to spend more time on the project (Lee, 2007). As shown in Figure 4, a change order has indirect costs which are as follows (Moselhi, 2003):

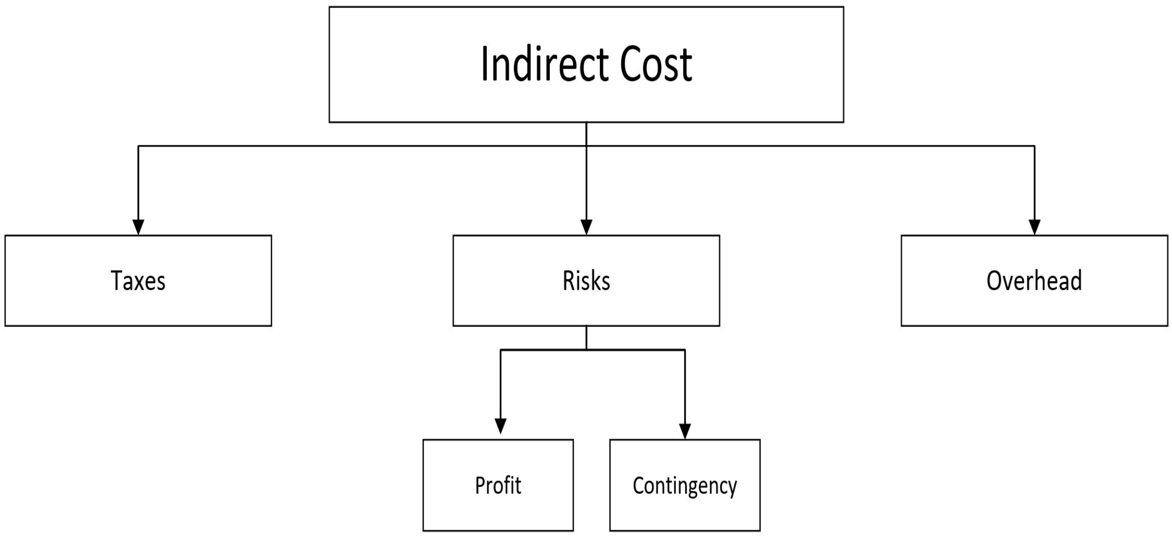


Figure 4. The indirect cost of change orders

2.3.3.3 Productivity Loss

Productivity is vulnerable to change, which is very frequent in a project. Severe financial losses can be caused if actual productivity is lower than anticipated productivity. These losses are referred to as loss of productivity. Loss of productivity is defined as the extra time spent per unit work. This relationship between productivity and loss of productivity is:

PI= $\frac{E}{A}$ **(Equation 2.2)**

In which:

PI= Productivity Index,

E= Earned Hours,

A= Actual Hours.

Loss of Productivity = $(1 - PI) \times 100$ **(Equation 2.3)**

In which:

LP= Loss of Productivity(%),

A= Actual contracts Hours,

E= Earned Contract Hours, and

PI= Productivity Index.

2.4. Previous Methods for Quantifying Change order impacts

Considerable studies and methods are available for quantifying change order impact. Quantification methods of change orders are as follows:

1- Total Cost Method,

2-Modified Total ost Method,

3- Industry Standards,

4- Measured Mile,

5- Baseline Productivity,

6- Statistical Methods,

7- Neural Network.

2.4.1. Total Cost method

In this method, the actual cost of the project is subtracted from the estimated cost. However, this approach has some disadvantages and inefficiencies. A major inefficiency of this method is that in the case of a project where there are multiple causes of productivity loss, the contractor will not be able to separate and divide the impacts of each cause. This method is not widely accepted in court and not recommended for claims (Pinnel, 1998).

This approach is applicable under the following conditions:

1. The actual damages and nature of loss cannot be determined with reasonable accuracy;

2. The estimated cost of the project was realistic;

3. The contractor’s actual costs were reasonable; and

4. The contractor was not responsible for the added costs (Schwartzkopf and Mcnamara, 2000).

2.4.2. Modified Total Cost Method

The modified total cost method can be defined as the modified version of the total cost method and is applicable if other methods are not applicable. This method is applied when the owners are no longer responsible for the errors and inefficiencies of the contractors' performance in the bid estimate (Schwartzkoph et al. 1992). However, when using this method, it is still difficult to separate the effects of changes contributed by owner (Serag, 2006).

2.4.3. Industry Standards

Some industries use methods to study the impacts of change orders on construction productivity.

2.4.3.1 Mechanical Contractors Association of America (MCAA)

Mechanical Contractors Association of America (MCAA) published Bulletin No. 58 in 1976, rewritten into PD-2 in 1994, a guideline for quantifying the loss of productivity. It is meant to assist the members in quantifying the effect of different variables in percent loss with each factor as minor, average, severe condition and includes 16 different factors. This guideline was both rejected and accepted by US courts, though it includes some qualitative damage degree identifications (Hanna, 2004).

According to (Serag, 2006), these Guidelines have some disadvantages as follows:

1- These guidelines did not indicate whether multiple factors should be summed, weighted, or combined in some other way. Also, these guidelines did not illustrate how to handle multiple or overlapping factors which affec labour productivity;

2- The factors used are redundant.Factors such as stacking of trade and joint occupancy are the result of other situations and not actual causes of inefficiency; and

3- The factors which are applied in a claim can only be implemented in the mechanical contracting industry.

In addition, This model has some disadvantages as follows:

1- The number of contractors that replied the survey is unknown. Only 88 data sets are given.

2- This model ignores the different types of data sets and trades have on the level of overtime impact.

3- Model accuracy is weak. The Coefficient of Determination (R-Squared) is 49.6% (Lee, 2007).

2.4.4. Measured Mile Approach

One quantification method which is widely accepted by researchers is the measured mile method. This method compares the impacted period of a data set with the unimpacted period. Once the difference between the impacted and unimpacted rates is found, loss of productivity can be measured by multiplying the number of units or amount of work performed during the impacted period (Schwartzkopf, 1995).

Loss of Productivity = (Impacted Rate - Unimpacted Rate)× Number of Units or Amount of Work During Impacted Period **(Equation 2.4)**

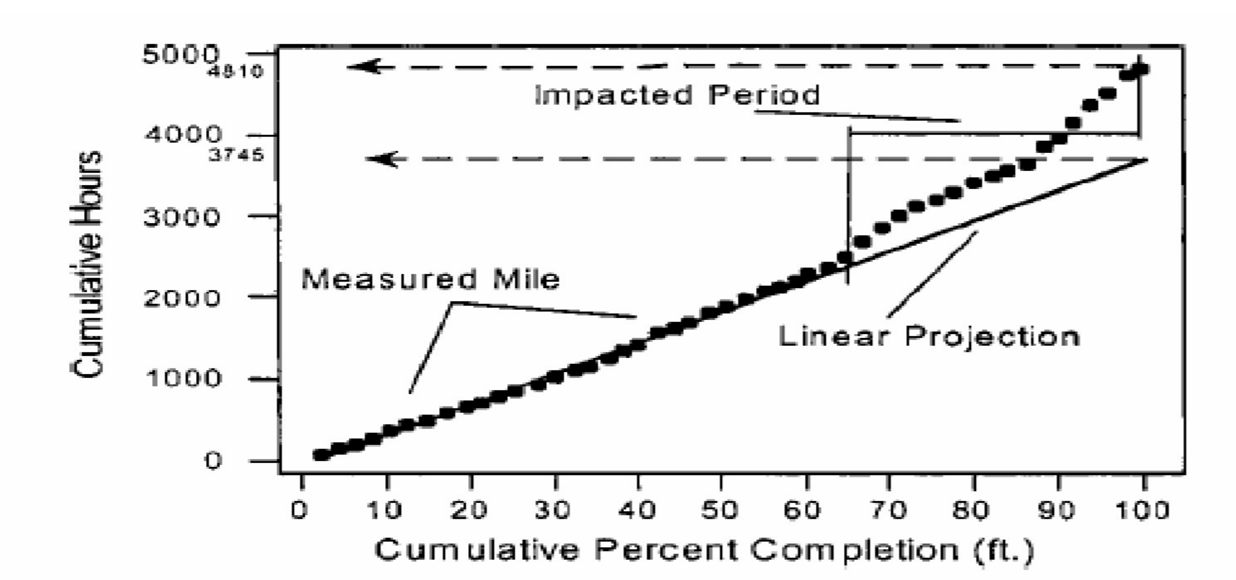


Figure 5. Measured Mile Approach (Gulezian & Samelian, 2003).

The production information needed to effectively track production efficiency and support the measured mile method includes the following:

1. Defining the work activity or cost account for the work performed;

2. Logging accurate work-hours used to perform the work; and

3. Logging accurate quantities of work completed for the period;

4. Briefly defining any condition or event that prevented optimum production, such as material deliveries, insufficient design information, field directives or changes to the original work scope (Presnell, 2003).

The advantages of this method are as follows:

1. It relies on data obtained during actual contract performance;

2. Labour productivity levels for both normal and affected periods are derived from such data sets records as job cost reports, daily logs and inspection reports; and

3. It avoids the shortcomings of industry studies and estimation guidelines because it is tied to the measured mile approach in probing claims for labour inefficiencies (Loulakis & Santiago, 1999).

However, the disadvantages are as follows:

1. The method does not provide any logical explanation of how the impacts of changes might lead to additional work;

2. It cannot segregate the effects of changes on productivity and assumes that all losses of productivity are related to owner (Serag, 2006); and

3. It assumes that the unimpacted period starts at the beginning of the project (Eden, 2003).

2.4.5. Baseline Productivity

Some researchers have developed a new approach to overcome the shortcomings of the measured mile analysis approach. Thomas and Završki (1999) and Thomas and Sanvido (2000) proposed the baseline productivity method for identifying the unimpacted measured mile approach when it is not applicable (AACE, 2004)(Thomas & Završki, 1999). Their method considers the period in which the contractor's performance is at its best to be the baseline period for quantifying productivity loss regardless of continuity or the presence of other disruption events (Thomas & Sanvido, 2000).

Table 1. Differences between the Baseline period productivity and Measured Mile (Thomas & Sanvido, 2000).

Measured Mile Approach	**Thomas Baseline Productivity Approach**
The negative Impacts should be limited to those caused solely by the contractor	The baseline period need not be free of owner impacts
The measured mile time frame should be several or more consecutive reporting periods.	The baseline time frame need not be consecutive reporting periods.
The focus is on finding periods of time where there are no owner caused impacts.	The focus is on finding the best performance the contractor could achieve.

In 2012, Ibbs highlighted a major BPM drawback (Thomas & Sanvido, 2000). In BPM, the baseline productivity period is considered to represent ten percent of project duration, though no logical basis supports that percentage as a realistic representation of baseline productivity. A baseline productivity period is highly dependent on project characteristics. For example, in 1993, AbuHijeh and Ibbs showed that 20% is sometimes required to accurately represent baseline productivity periods, while for other projects, 3% might be satisfactory (Ibbs, 2012).

In 2005, Ibbs and Liu developed a method using the statistical clustering technique to determine unimpacted productivity as the baseline for comparison. This method relies on the separation of data into different groups. The datasets are first divided into K groups and then moved between the clusters, which are between each data sets and K cluster centers. The process of iteration is then completed until there is no more change in cluster means (Ibbs & Liu, 2005).

However, the proposed method requires a complicated calculation process, which makes it less desirable. In addition, there are some drawbacks associated with K-means clustering which can impact results. Results generated using K-means clustering are highly dependent on initial cluster centroid choice. Data sets are classified in advance into unimpacted and impacted groups (k=2) on the assumption that these two groups can include all datasets (Zhao & Dungan, 2014).

2.4.6. Statistical Methods

2.4.6.1 Leonard Study

In 1988, Leonard performed research based on 90 case studies from 57 different data sets to quantify loss of productivity caused by change orders for electrical, mechanical, civil and architectural work using linear regression analysis. This study attempted to clarify the relationship between the variables and their effects on labour productivity. These variables are as follows (Leonard, 1988):

1- The Frequency of Change Orders = $\frac{Number\ of\ Chang\quad Orders}{Contract\ Duratio\ (Month\ \ )}$ **(Equation 2.5)**

2- The Average Size of Change Orders = $\frac{Total\ Chage\ Orders\ Hours}{Number\ of\ Chang\quad Orders}$ **(Equation 2.6)**

3- The Percentage of Change orders hours = $\frac{Total\ Chang\quad Orders\ Hours}{Act\qquad Contract\ Hours} \times 100$ **(Equation 2.7)**

This study illustrates that the number of change orders and their average size does not precisely reflect loss of productivity due to the low correlation of the coefficients. However, this study did not show that the change orders percentage was an accurate reflection of loss of productivity due to the high correlation of the coefficients (Serag, 2006).

In this study, Leonard produced six curves to predict loss of productivity. The data was divided into two main categories: electrical/mechanical and civil/architectural contracts. According to Leonard, Fig. 5 shows the following three causes of productivity loss:

Type 1: Change orders are the only major cause of productivity loss;

Type 2: Change orders and one additional major cause result in loss of productivity; and

Type 3: Change orders and more than one additional major cause result in productivity loss.

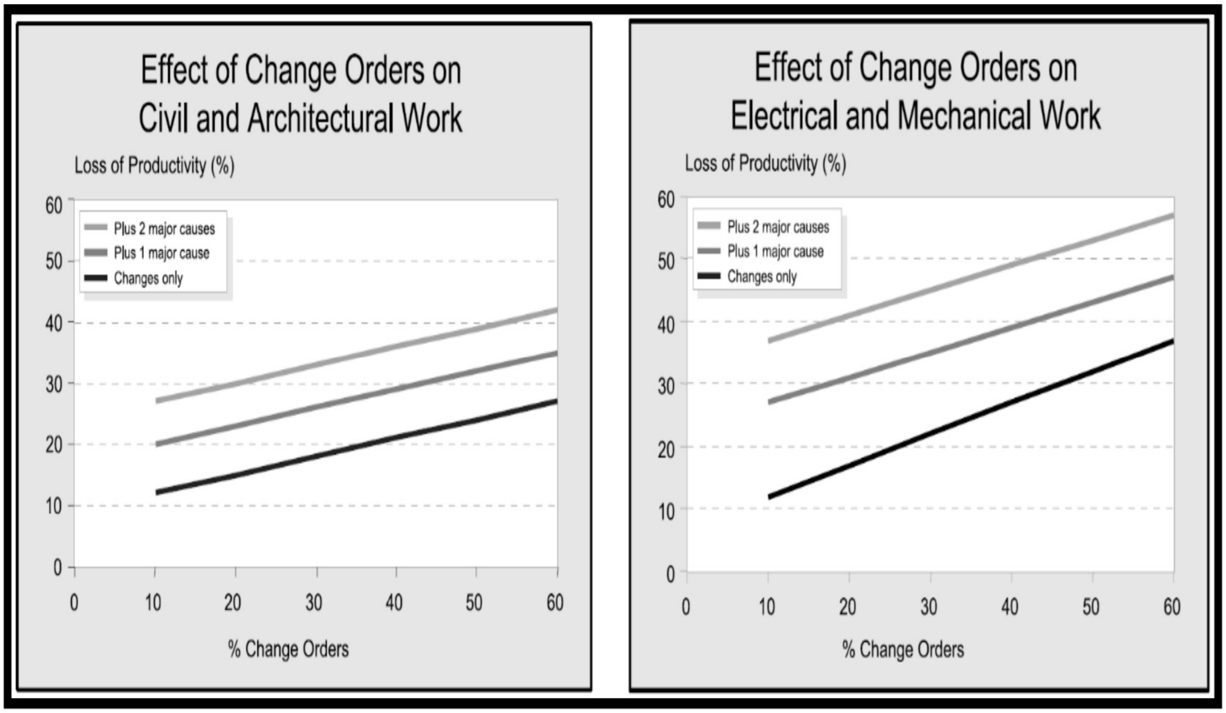


Figure 6. Leonard Curves for M/E and C/A Work (adopted from Revay, 2003)

In this study, additional major causes included acceleration, poor scheduling and coordination, increased complexity and late delivery (Moselhi, Leonard, & Fazio, 1991).

Due to its simplicity and acceptable results, this model can function as a substitute for measured mile analysis when accurate data is not available. However, some studies have criticized the model for the following reasons: it is based on a biased project that went to the claim phase, it does not compare impacted and unimpacted projects and it does not consider other variables such as the timing of change orders (A. S. Hanna, Russell, & Vandenberg, 1999).

2.4.6.2 Ibbs Study

In 1995, a study by Ibbs and Allen identified the quantitative relationships between volume and timing of change orders and the consequences of such changes based on statistical analysis. The study contains data from 104 public and private data sets with different project delivery systems. The average and median project value are $80 million and $44 million, respectively. One of the advantages of this study is that it does not categorize productivity into civil/architectural and mechanical/electrical. This study evaluates three main hypotheses, which are as follows:

1. Changes in the later stages of the project are carried out less efficiently than changes that occur early in the project;

2. The more changes there are during a project, the more negative the impact on labor productivity; and

3. Hidden or unforeseen costs of change increase with more changes to the project (Ibbs & Allen, 1995).

Although this study shows changes have a considerable impact on productivity, it has several disadvantages, such as:

1. Poor statistical performance indicators, such as the coefficients of correlation and determination values;

2. This study cannot completely prove the fact that changes which occur late in a project are implemented less efficiently than those which occur at earlier stages in the project (A. S. Hanna et al., 1999).

In 2005, Ibbs studied the impact of change's timing on labour productivity. He illustrated that late changes are more disruptive to project productivity than early changes. In this model, datasets were collected from 162 construction data sets over a nine-year period. The project size ranged from \$3.9 million to \$14.5 billion. Thirty-five Percent of the projects were heavy/highway, 16% were commercial and 49% were industrial. Three curves were generated to illustrate the impact of early, normal and late changes on productivity. Figure 7 shows that projects with early and normal timing have shallower curves and can actually tolerate a small amount of change prior to reaching a productivity value below 1.00 (Ibbs, 2005).

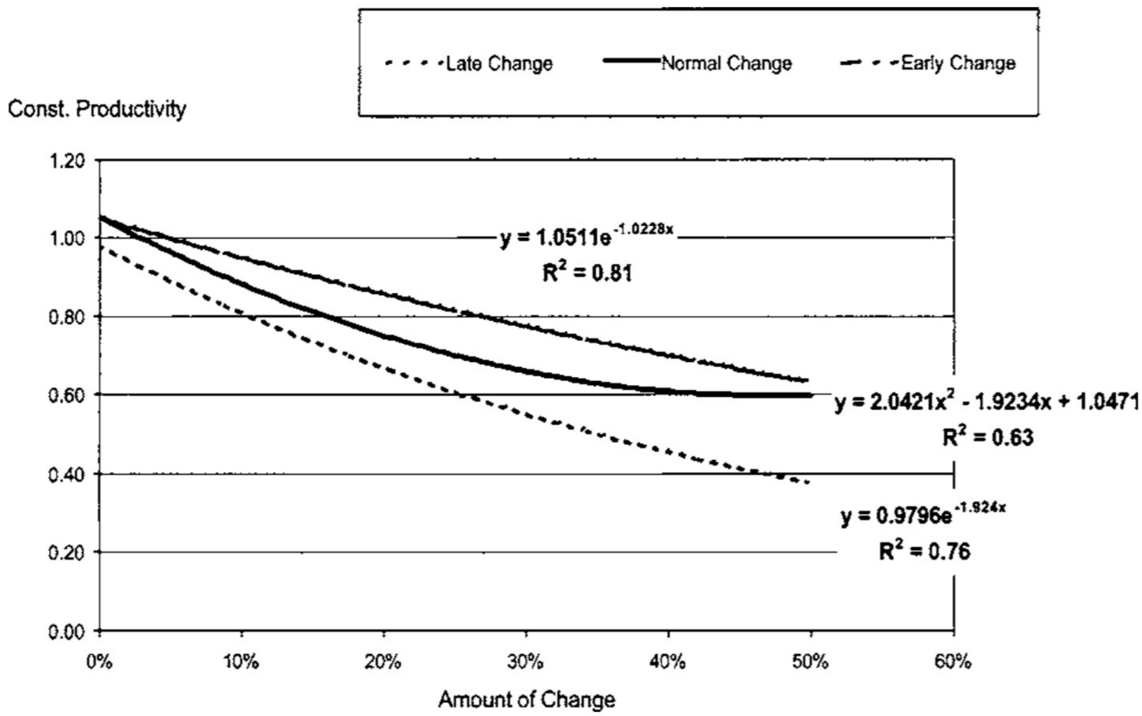


Figure 7. Construction Productivity versus Timing of Change (Ibbs, 2005).

In 2008, Ibbs and McEniry published a study comparing the Ibbs (2005) and Leonard (1988) models. Their comparison explains that the Ibbs study has a higher R-squared percentage, which can be considered a better predictor (Ibbs & McEniry, 2008).

2.4.6.3 Hanna's Study

Since 1990, Hanna has conducted several studies on electrical and mechanical projects in order to determine productivity loss using "Delta" as a factor. According to Equation 2.8, a positive delta value implies that the actual productivity is lower than the estimated productivity. However, a negative delta value indicates that a project has a higher productivity than estimated, as shown in Equation 2.8 (A. S. Hanna et al., 1999).

$$\%\ \text{Delta} = \frac{(Total\ Actual\ Labor\ Hours) - (Estimated\ Hours + Change\ Order\ Hours)}{Total\ Actual\ Labor\ Hours} \times 100 \qquad \textbf{(Equation 2.8)}$$

According to Hanna (1999), one of the following criteria will decide whether the project is impacted or unimpacted:

1. Planned versus actual loading curve: Coffman (1997) indicated that a project is considered impacted by change orders when the actual and planned cumulative work hours vary substantially;

2. Time extension: Projects that are impacted by changes tend to take longer than the originally planned duration;

3. Timing of change orders: Changes issued in the latter part of projects tend to have a more negative impact than changes issued when the project is <50% complete (Ibbs & Allen, 1995);

4. Lead time: If the lead time, or the time available between making a change and the actual completion of the work, is small, the loss of productivity tends to be much higher than if the lead time is adequate; and

5. Ripple effect: There is a strong correlation between projects impacted by changes and schedule compression, stacking of trades and overmanning. This indicates that change orders can create another set of productivity-related problems such as schedule compression.

To determine if there is a connection between the amount of change and the degree of impact on labour productivity, the amount of change is calculated in the following two ways:

1. The total number of change orders that occur on a project; and

2. The amount of change is measured as a percentage of the project size.

This is achieved by taking the estimated change order hours as a percentage of the estimated base hours and total actual hours. Results show that there is a statistically significant relationship between the amount of change, measured both as the number of changes and as change order hours/estimated base hours, and the degree of impact on labor productivity (A. S. Hanna et al., 1999).

This study uses the timing effect of the change order and its effect on labor productivity. In order to develop the model, variable weighted timing (WTIMING) is used to determine when the change orders occurred. After the WTIMING is measured, a comparison is carried out for both the impacted and unimpacted projects. The results indicate that there is a relationship between WTIMING and labour productivity amount. Delta, as a percentage of total labour hours, is selected to be the dependent variable (A. S. Hanna et al., 1999).

In 2002, Hanna published another study to quantitatively determine if there is a relationship between electrical or mechanical projects and change orders. The research also contains logistic regression analysis to develop the model. The authors define the project as impacted or unimpacted based on a survey question regarding whether the project is over budget or within budget based on the budget labour hours, which are defined as the total estimated hours that the contractors used to allocate the labour resources. Impacted data sets with less than 5% delta are considered to be unimpacted (A. Hanna, Camlic, Peterson, & Nordheim, 2002).

In 2017, Hanna and Iskandar published a study which quantitatively predicts the cumulative impact of change orders using the delta approach. This study uses 68 data sets, with the scope of the research being limited to those projects built under lump sum contracts and delivered under the traditional design-bid-build approach. The author uses owner-initiated change order percent, productivity, turnover, project management time and overmanning. In this model, forward stepwise regression is used to ensure model appropriateness. Contrary to previous studies, this study validates its models using cross-validation and prediction error sums (Hanna & Iskandar, 2017).

Based on the Farbarik study of Hanna's research, delta percentage is not an appropriate indicator of whether the contractor suffered a labor hour overrun because of the changes, as the owner takes on the risk and might be responsible for compensation (Farbarik, III, Hanna, Moselhi, & Hassanein, 2004). In addition, the stepwise regression technique, which was used to develop the model, has several limitations, as follows:

1. The stepwise regression technique does not necessarily produce the best model if there are redundant predictors; and

2. The amount of data used in the study is too small to develop an appropriate model.

2.4.6.4 Decision Tree

In 2004, Lee published research which performed unbiased classification of interaction selection and estimation. It is used to develop a model that can classify data sets impacted by change orders. Lee states that a decision tree is composed of nodes and connections between nodes. This study contains 142 case studies of which 69 datasets were impacted and 73 unimpacted. The output feature of this research has a binary case of 1 if the project was impacted by a change and 0 if it was not impacted (Golnaraghi et al., 2018).

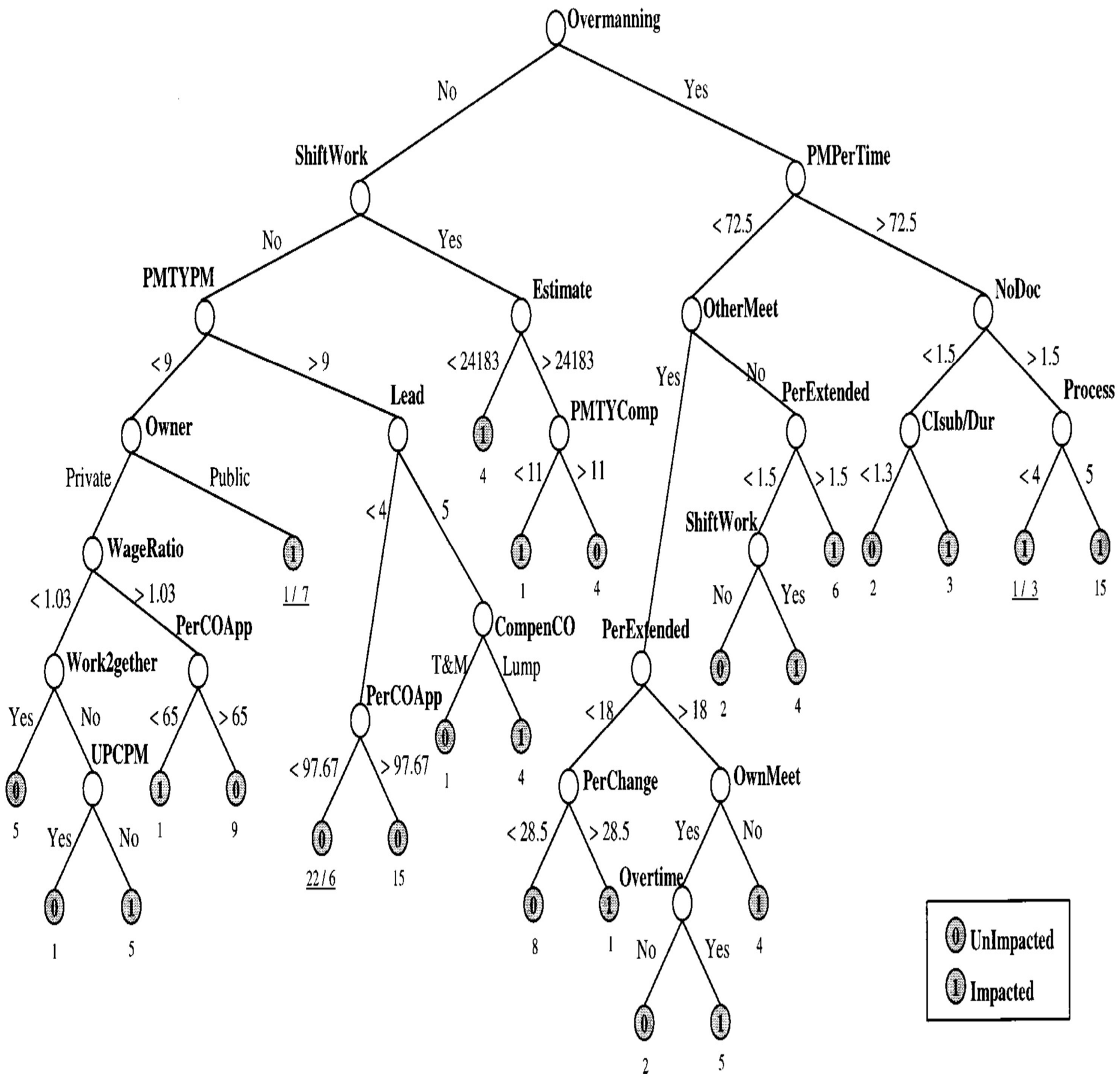


Figure 8. Impact Classification Tree (Lee et al., 2004).

The regression tree is implemented to develop a quantification model that predicts productivity loss due to multiple change orders. The regression is used when it is proven that the project is impacted. Figure 8. illustrates a regression tree for impacted projects. The lee model was developed by utilizing stepwise regression (Lee et al., 2004).

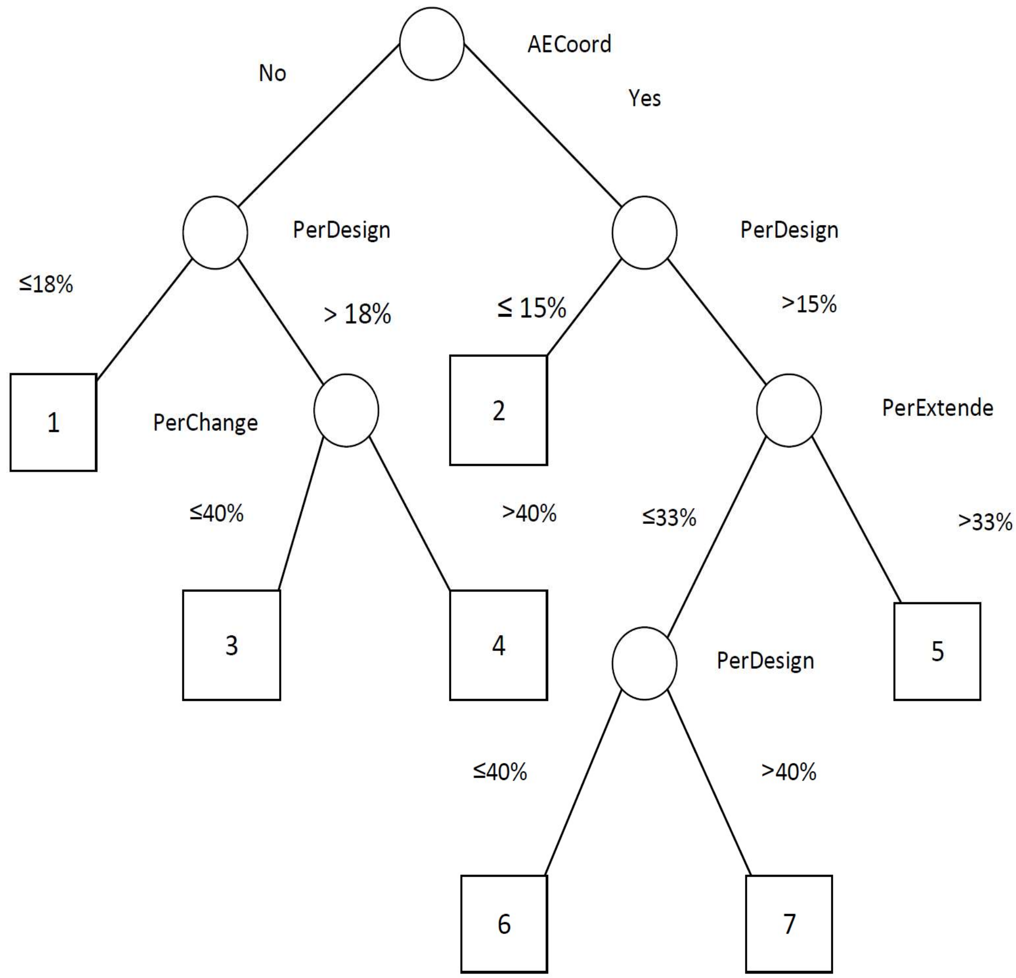


Figure 9. Quantification Tree Model (Lee et al., 2004).

This study suffers from some disadvantages, such as the small number of cases examined. This limitation causes the model to become weak in the downstream branches. It is also assumed that the developed model can be applied to electrical and mechanical trades. However, the model development procedure is complex, and concerns exist regarding an imperfect understanding of the principles, theories, and techniques (Lee, 2007).

2.4.7. Building Information Modeling (BIM)

In 2017, a study by Moayeri, Moselhi and Zhu, building information modeling was implemented to visualize the ripple effects of owners' design changes. The study measures the impact of design changes and their ripple effect on a project's total duration and updates the project schedule accordingly. It also predicts the impact of design changes, and their ripple effect, on project cost. The cost data provided by the developed model would be the updated project total cost including direct cost, indirect cost and impact cost (Moayeri, 2017).

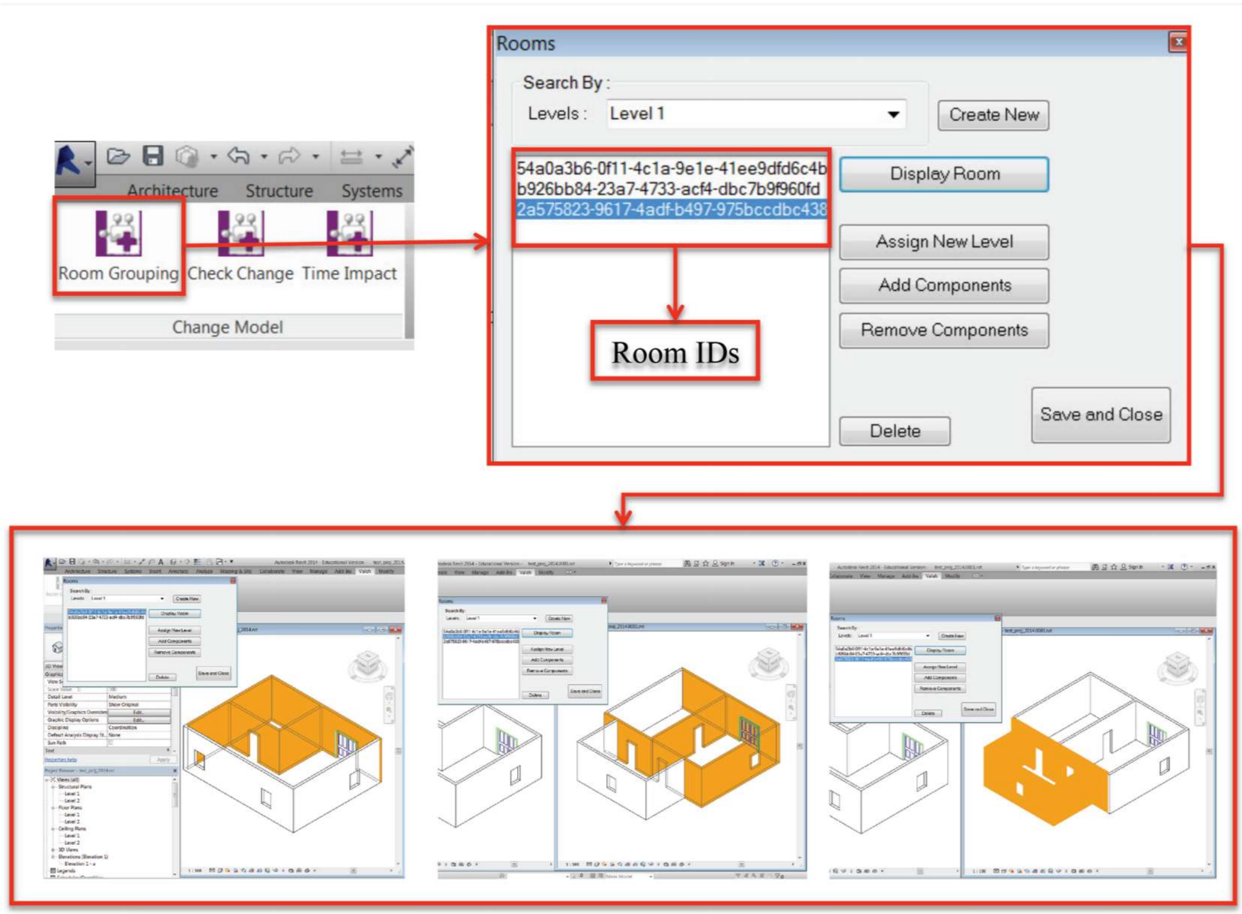


Figure 10. Room Grouping Command and Generated Group IDs (Adopted from (Moayeri et al., 2017).

This model was developed in several separate add-ins, and using Revit Application Programming Interface (API) coded in C#. To ascertain the ripple effect of owner-requested design change on other project components, a BIM-based add-in named Check-change was developed. These add-ins are as follows:

1- Check Change: To compare the "as-planned" building information model with "as-changed".This add-in considers and introduces changes on the quantities of impacted building components and respective dependencies;

2- Ripple Effect: To generate a detailed report for each contemplated design change and visualize the ripple effects;

3- Time Impact: To quantify the ripple effect of owner-requested design change on project time. This add-in calculates the impact of design changes and their ripple effect on project' total duration;

4- Room Grouping: this add-in was developed to link the impacted components' duration to project original schedule to create as-planned schedule;

5- Cost Impact: this add-in measures the impact of design changes and their ripple effect on a project's total duration; and

6- Data Filtering: this add-in was developed to assist the user browse through data faster by filtering unnecessary and unwanted data (Moayeri et al., 2017).

Finally, the proposed model was validated utilizing Autodesk Revit Architecture 2014 to develop the case study for as planned and as changed bim models. The results of model validation show that the developed bim-based software has several advantages over other related softwares (Moayeri, 2017).

2.4.9. Neural Network

Artificial neural networks are inspired by their biological equivalents, such as the brain and the nervous system. The biological brain is completely different from the conventional digital computer in terms of structure and information processing. Essentially, the biological brain (or its most perfect example, the human brain) is superior to and much more sophisticated than conventional computers. The most important feature of a biological brain is the ability to learn and adapt. However, the conventional computer does not have such features. The basic structure of neural networks is a "neuron which can be alleged as a processing unit. In a neural network system, neurons are connected with each other through weights. Each neuron receives weighted information through the connections from the neurons that it is connected to and produces an output by passing the weighted sum of those input signals through an 'activation function' (Sazli, 2006).

Another feature of neural networks is learning. This feature distinguishes neural networks from conventional computers. It can assist neural networks in their environment and improve their performance through learning. Based on Hykin's research, this is the defining feature of a neural network, as its free parameters are adapted through a process of stimulation by the environment in which the network is embedded. The type of learning is determined by the manner in which the parameter changes take place (Haykin, 1999).

One of the most popular and widely-used learning algorithms is the backpropagation algorithm. As shown in Fig. 12, the artificial neurons are organized into two or more layers, the signals are sent forward and, later, the errors between the actual results and the predicted outputs can propagate backward. Subsequently, the neurons in the input layer receive the input data and the neurons in the output layer will create the output of a given database. In particular, the hidden layer enables the networks to calculate complex links among the layers. The purpose of the backpropagation of the ANN model can be achieved by providing the algorithm with inputs and ideal target results through training and self-learning. This causes the errors to be minimized (Zhang, 2017).

The Basic concepts of Backpropagation Algorithm are common in the previous literatures. Based on the Equation 2.9, the initial stage of ANN is on the Activation Function of ANN which is (Gershenson, 2003):

$$Aj(\bar{x}, \bar{w}) = \sum_{i=0}^{n} XiWji \qquad \textbf{(Equation 2.9)}$$

Where Xi and Wji are the inputs and their weights respectively.

When the output function is the same as the activation function, it would be called linear, which has several limitations. To avoid the limitations, a sigmoidal function which is representing the output has been used:

$$Oj(\bar{x}, \bar{w}) = \frac{1}{1+e^{Aj(\bar{x},\bar{w})}} \qquad \textbf{(Equation 2.10)}$$

The outputs of this equation were very close to zero for large negative numbers, for Zero the output is 0.5, and for the large positive numbers, the output is close to Zero (Gershenson, 2003).

Since the purpose of the training process is to obtain the desired output obtain from a given input and to minimize the error, the error function of each neuron can be defined as Euation 2.11- 2.14 (Gershenson, 2003; (Ai & Zsaki, 2017):

$$Ej(\bar{x}, \bar{w}, d) = (Oj\ (\bar{x}, \bar{w})\ -\ dj)^{\ 2} \qquad \textbf{(Equation 2.11)}$$

Thus, the error of the system in the output layers can be calculated as :

$$E(\bar{x}, \bar{w}, \bar{d}) = \sum_j (\mathrm{Oj}\ (\bar{\mathrm{x}}, \bar{\mathrm{w}})\ -\ \mathrm{dj})\ 2 \qquad \textbf{(Equation 2.12)}$$

The Backpropagation method implements the Gradient Descendent method to adjust the weights as follows:

$$\Delta Wji = -\eta\ \frac{\partial E}{\partial Eji} \qquad \textbf{(Equation 2.13)}$$

Based on the equation 2.13, each weight(ΔWji) can be calculated by multiplying the negative constant eta (η) with $\frac{\partial E}{\partial Eji}$ (is the derivative of E in resepect to Wji). In this equation ΔWji is the adjustment fro a specific weight of Wji.

Equation 2.14 is used until appropriate weights(Minimal Error) is found. Consequently, by considering the equations 2.12, 2.13 and 2-14, the adjustment of each weight can be calculated as follows:

$$\Delta \mathrm{Wji} = -\eta\ \frac{\partial E}{\partial Eji} = -\ 2\ \eta\ (Oj - dj)Oj(1 - Oj)Xi \qquad \textbf{(Equation 2.14)}$$

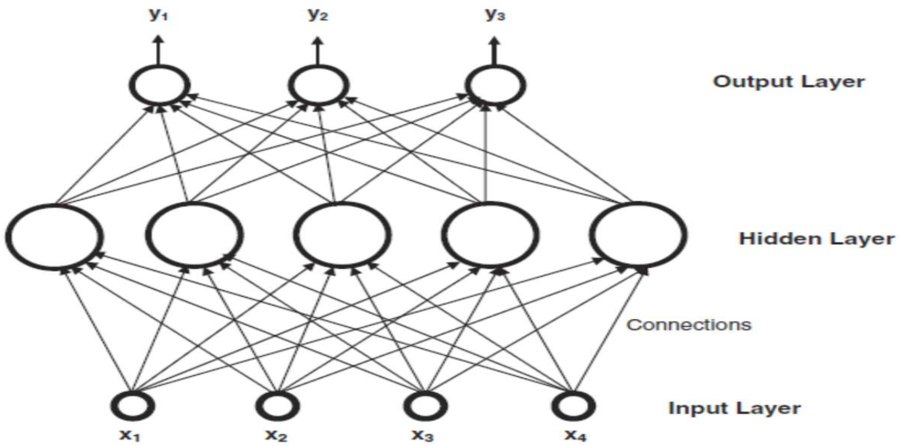


Figure 11. A Typical Structure of Artificial Neural Network (Goh & Kulhawy, 2005)

The quantification of the impact of change orders on productivity is a good candidate area for implementing a neural network. One of the major studies which has been conducted in this field is that of Moselhi (2005). This model was completed in three stages: identifying change order factors which cause the loss of labor productivity, modeling the timing impact and developing a neural network model. According to this study, the following factors affect labor productivity:

1. Intensity: this factor can be expressed as the number of change orders, the frequency and the ratio of change order hours to contract hours;

2. Timing in relation to project duration: this factor assumes that high labor productivity losses occur at the end of the project duration and does not consider the ripple effect of change orders on the remaining unchanged work;

3. Work type: the type of work affects the impact on productivity. This impact is due to the difference in the level of skill and the interdependence which varies among work types and from one type of work to another;

4. Type of impact: change orders do not solely cause productivity loss. Rather, due to the changes, other variables cause losses of productivity, including overtime, overmanning and congestion;

5. Project phase: this factor differentiates between changes introduced during the design phase and construction phase; and

6. On-site management: this factor is related to the project manager's experience (Moselhi et al., 2005).

The timing impact of change order was measured by equation 2.15 (Moselhi et al., 2005):

$$TP(t)= \frac{HCO}{PH} \qquad \textbf{(Equation 2.15)}$$

Where TP= timing impact of a change order in the period (t), HCO= Actual change order hours during (t), PH= Planned Hours during period (t).

The model was therefore developed using NeuroShell2 to quantify the impact of change orders on productivity. The proposed model's average estimation error for eight case studies was 17.7%, which is outstandingly low compared to those of A. S. Hanna et al., 1999 and Moselhi et al., 1991.

2.4.10. Summary

In this chapter, previous studies were discussed and some of their drawbacks highlighted. The previous studies in this area suffer from some disadvantages that limit their use in estimation and prediction of productivity loss. The main drawbacks of the previous studies are small data sets, types of implemented techniques and weak correlations and coefficients of determination between variables and loss of productivity. All of these drawbacks might discourage owners, contractors and courts from quantifying loss of productivity due to change orders.

Chapter 3. Proposed Methodology

3.1. Overview

This chapter provides an overview of the research methodology as shown in Figure 12. It consists of a research flowchart as well as literature review, data collection, data normalization, variable selection, model development and model validation. In order to show the proposed model, each stage is considered as a separate part.

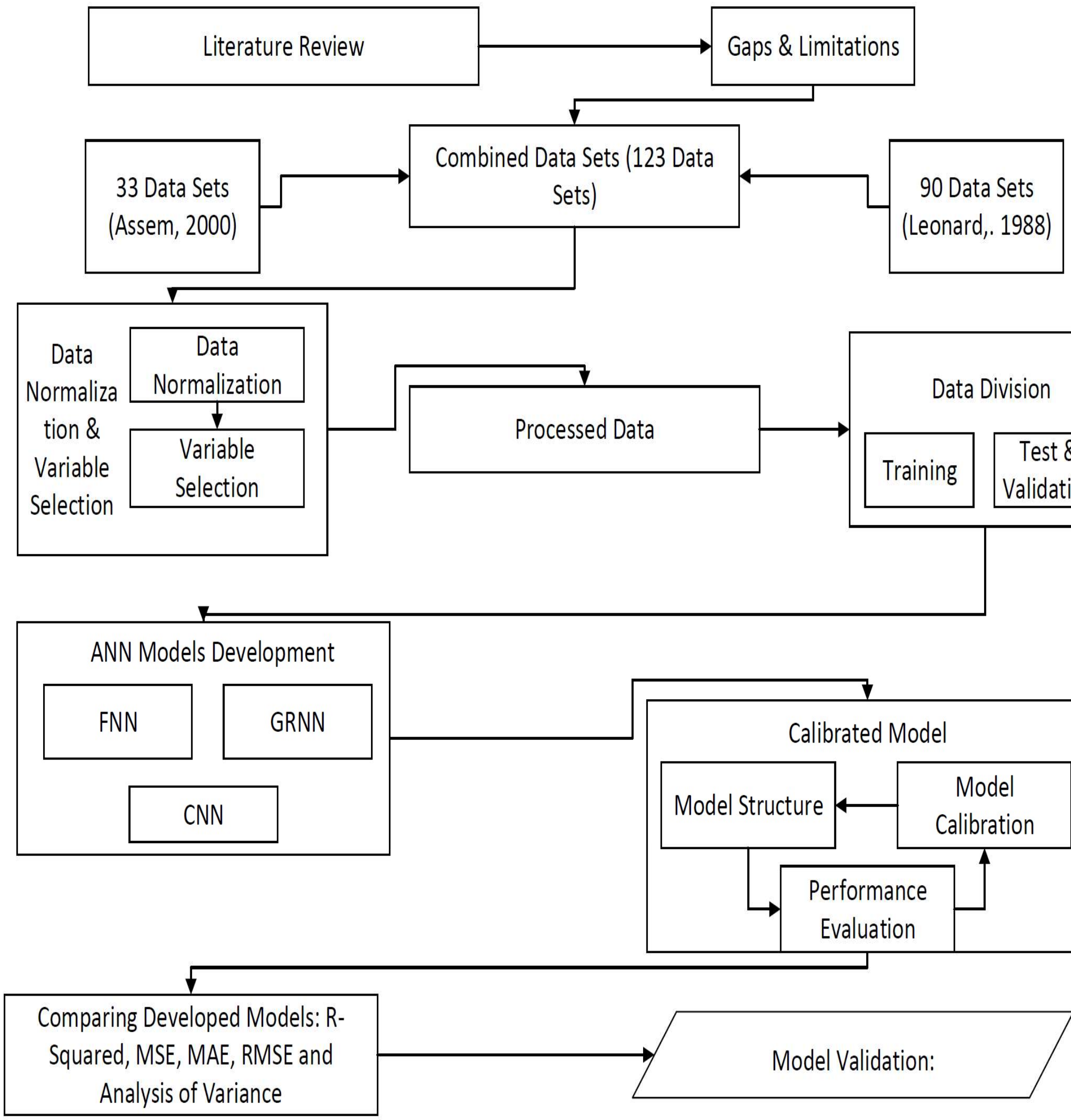


Figure 12. Overall Flowchart of the Study

3.2. Literature Review

To review previous studies and clarify recent methods, this research considers previous studies related to change orders and loss of labor productivity. The literature review consists of 3 parts which simplify recent attempts to quantify the loss of labor productivity due to change orders. This research considers the following previous methods: total cost method, modified total cost method, industry standard, measured mile approach, baseline productivity, statistical methods and neural networks.

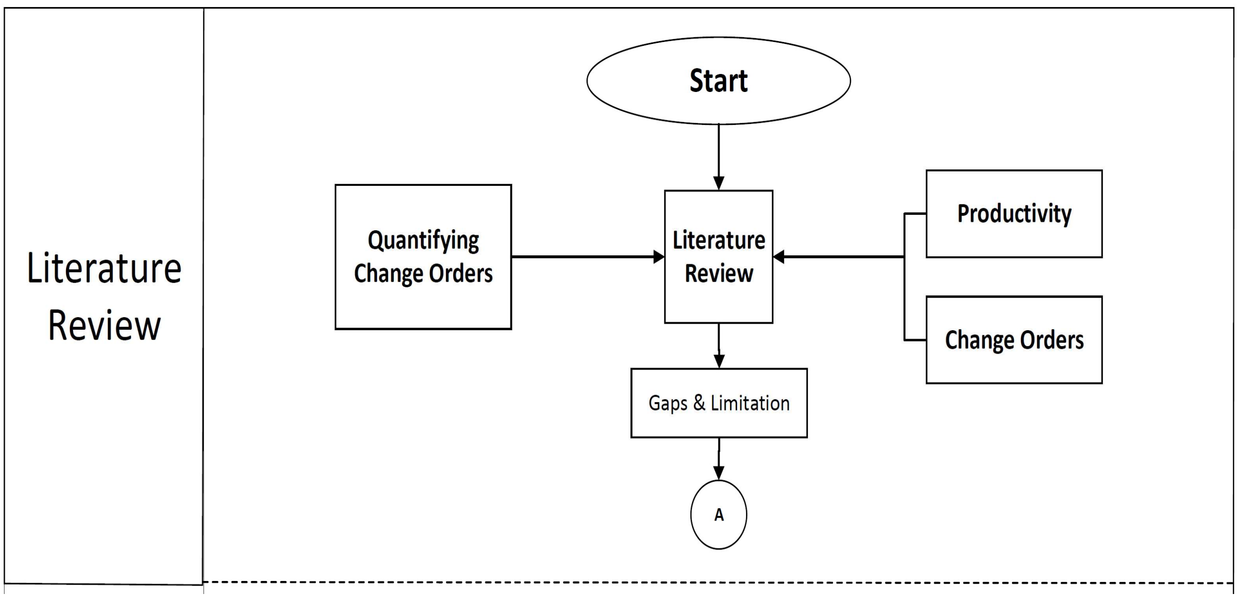


Figure 13. Summary of research Literature Review

3.3. Data Collection

This study included of two datasets which were collected from previous studies of construction projects: namely, Leonard (1988) and Assem (2000). A total of 123 data sets were collected, which can be considered as a satisfactory volume of data to develop a model for quantifying loss of productivity due to change orders. The collected datasets have seventeen unique variables with diverse types and scales.

The combined datasets were normalized since they were not normally distributed For example, the data set range for one of the variables is from \$80,000 to \$23,172,000. However, the range of another variable was from 0.76 to 2.78. Hence, the data sets had to be transformed into a normal distribution. In most of the data sets, the variables were normally distributed. If the datasets is not normally distributed, the chances of the researcher committing either an overestimation or underestimation error increase (Jason, Osborne, & ERIC Clearinghouse on Assessment and Evaluation MD., 2002). In order to transform the data sets into a normal distribution, the range of the data sets had to be transformed from 0 to 1. This transformation structures the data set well and reduces the chances of overestimation and underestimation (Patro & Sahu, 2015).

In order to redistribute the normal distribution, this study utilized the min-max normalization technique. This technique consists of three parameters: original data, min value, and max value (Jain & Bhandare, 2011).

$$Z= \frac{(Ai-Mi\ \ (A))}{(Max(A)-Min(A))} \qquad \textbf{(Equation 3.1)}$$

Where,

Z= Normalized data,

Ai= Original data,

Max (A) = Max range of original data,

Min (A) = Min range of original data.

Furthermore, in order to distinguish insignificant variables from major variables, this study implemented best subset regression to eliminate insignificant variables. Best subset regression is a method for variable selection based on the analysis of MLR (multiple linear regression) models using all combinations of variables and the selection of the combination of variables that gives the best fit (André, Narula, Elian, & Tavares, 2003).

Best subset regression can provide analysis of certain statistical criteria such as R-square, adjusted R-squared and Mallows's Cp. (Ruengvirayudh & P.Brooks, 2016). In order to find the best combinations of variables, this study considered adjusted R-squared, R-squared and Mallows's Cp. Cp is an unbiased estimator for the ordinary least square error (Madigan & Ridgeway, 2004). A small Mallows's Cp value indicates that the model is relatively precise (Gilmour, 1996).

In addition, R-squared and adjusted R-squared values measure the perfection of the model, where R^2=1 represents a perfect model fit. Adjusted R-squared is the number of predictors in the regression model.The maximum value of adjusted R-squared represents the model's ideal combination. (Harel, 2009). Based on the above information, the best subset regression was calculated in Minitab 2018 software. Minitab 2018 is statistics software used for basic statistics, regression and ANOVA, quality tools, design of experiments, control charts, and reliability.

Subsequently, in order to develop the timing model to compare the current model with Ibss timing model, the combined datasets were divided into the three separate timing datasets. The combined datasets were devided into eary change datasets, normal change datasets, and late change data sets. In order to divide Leonard's datasets (1988), Equation (3.2) is used to recognize each change (Eldin, 1989).

Earned Hours = Original Estimated Hours x Percent Complete Work **(Equation 3.2).**

The percent complete of works shows that the change occured at each period. A percent complete work value of less than 25% shows when the change happened at an early stage, the value between 26% to 75% shows that the change happened at middle (Normal) stage and the value between 76% to 100% shows that the change happened at the final (Late) stage (Ibbs, 2005).

Similarly, in order to divide the Assem datasets, each change was recognized by change order direct hours results from the Assem study (2000). The biggest value of five change orders direct hours showed that the change was happening in one of the three timing stages.

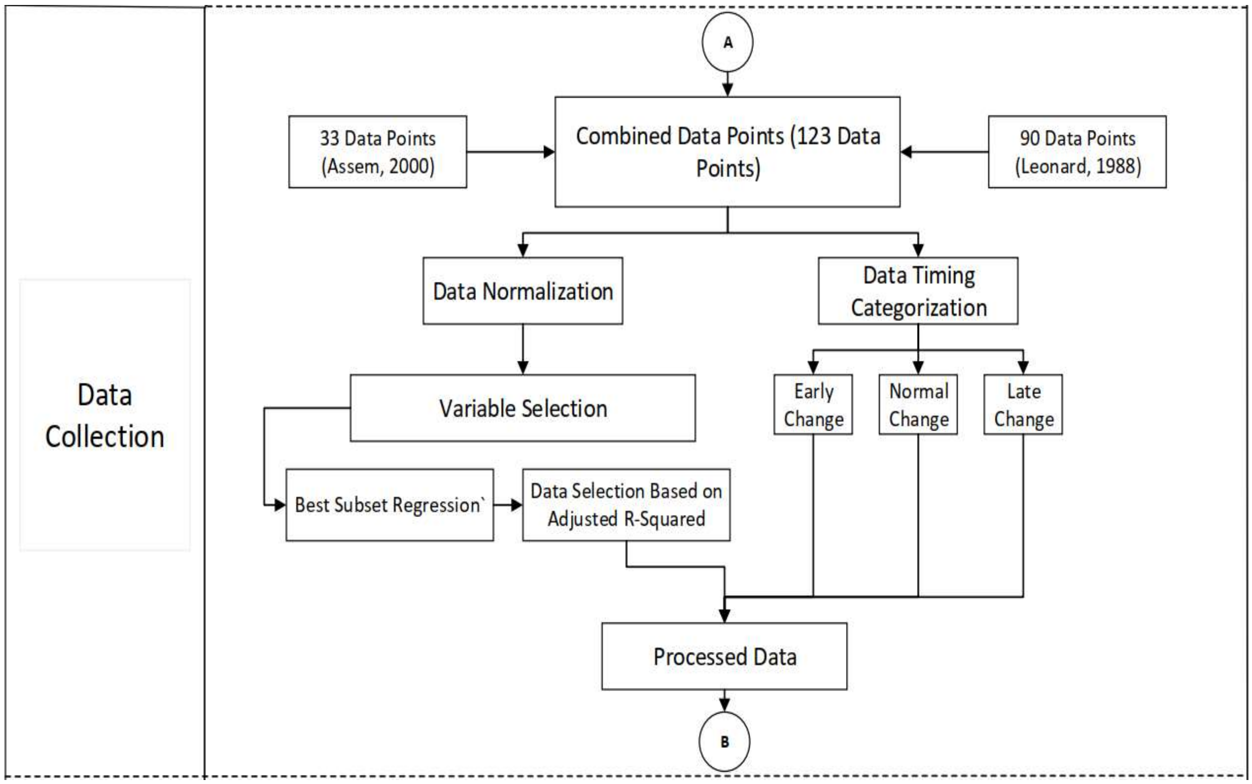


Figure 14. Summary of research Data Collection

3.4. Model Development

In order to develop a model for quantifying loss of productivity caused by change orders, three artificial neural network techniques were implemented. These techniques were the feedforward neural network (FNN), cascade neural network (CNN) and generalized regression neural network (GRNN).

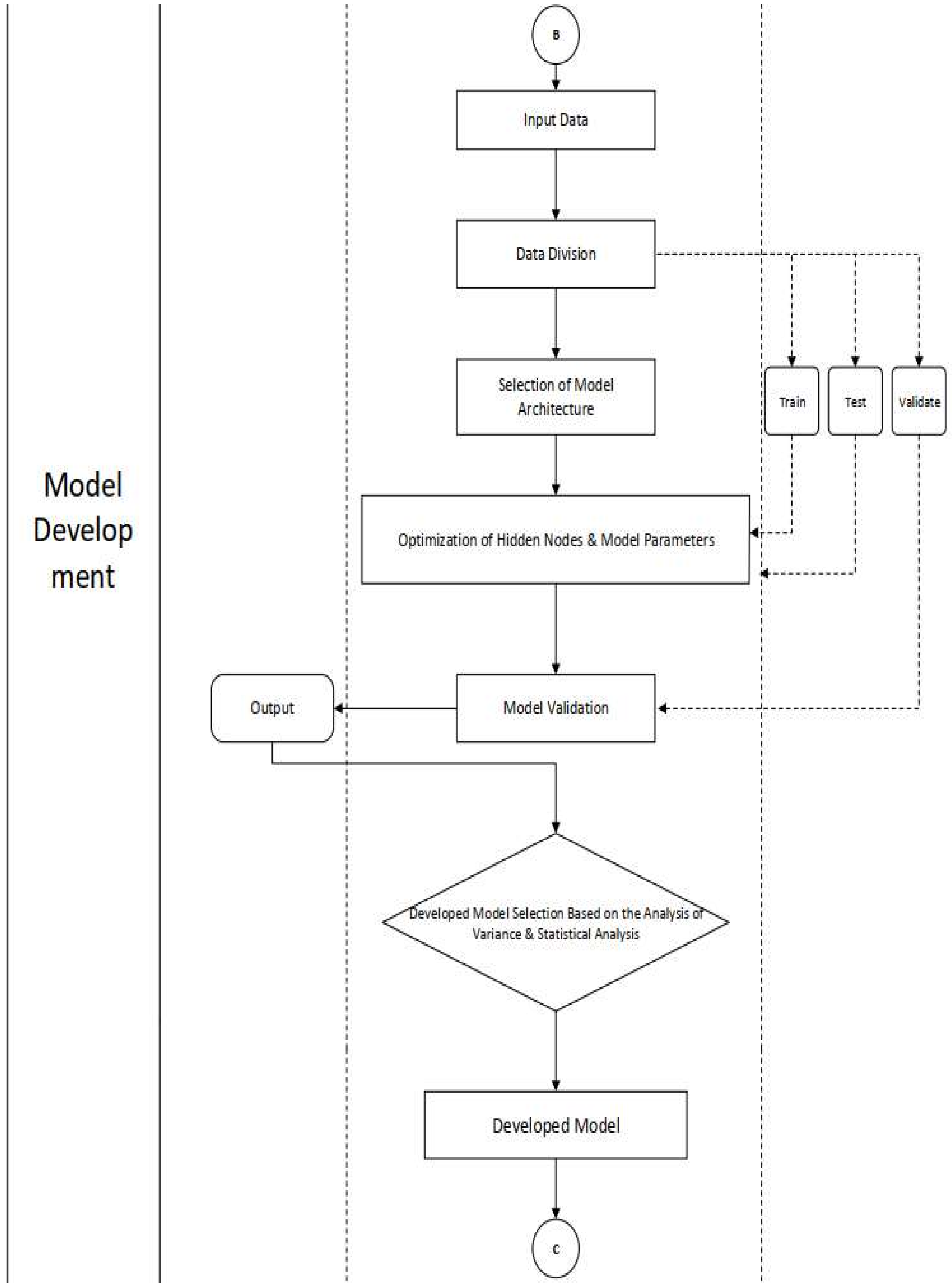


Figure 15. Summary of Model Development

3.4.1 Feed Forward Neural Network (FNN)

When there is no feedback of neurons towards the input, the neural network is referred as a feedforward neural network. Each layer other than the output layer contains a connection to the next layer. This neural network also uses learning algorithms for training information instead of backpropagation. Feedforward neural networks can be categorized as single-layer or multilayer. In contrast with the single-layer type, the multilayer type has at least one layer of hidden neurons between the input and output layers. The existence of one or more hidden layers assists the network in extracting higher-order statistics. There is a direct relationship between the number of hidden nodes and the complexity of the neural network which requires more training data and time (Haykin, 1999).

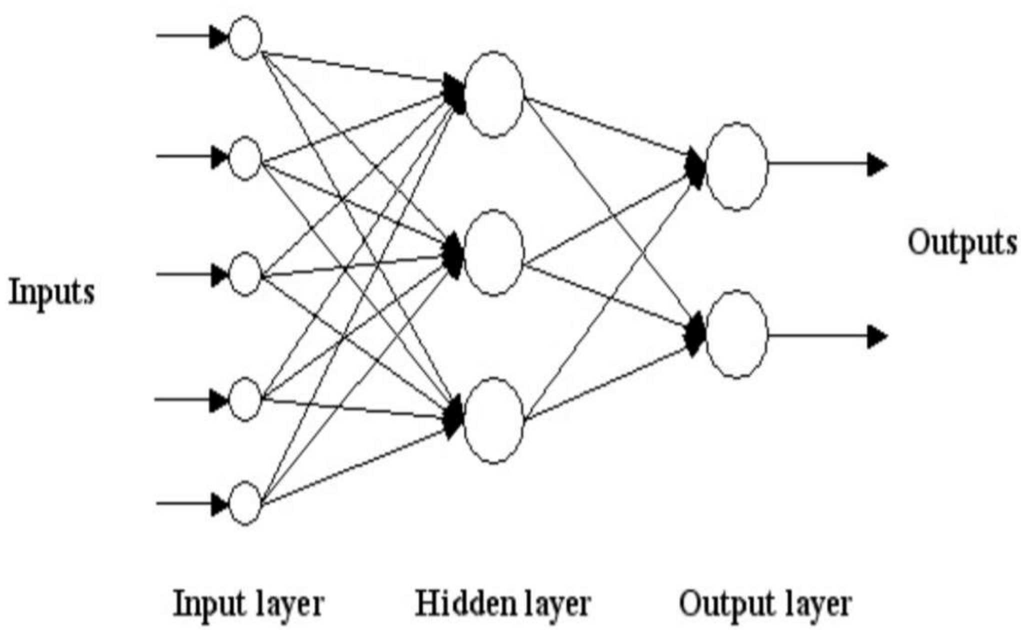


Figure 16. A multilayer Feed Forward Neural Network (Sazli, 2006).

3.4.2 Cascade Neural Network (CNN)

The cascade neural network was developed in 1990 by Lebiere and Fahlmann to boost learning in artificial neural networks (Fahlmann & Lebiere, 1990). Cascade neural network is similar to feedforward neural networks but include a weight connection from the input to each layer and from each layer to the following layers. While a two-layer feedforward neural network can potentially learn virtually any input-output relationship, a feedforward neural network with more layers might learn complex relationships more quickly. A backpropagation algorithm is used in cascade neural networks for weight updating (Badde, Gupta, & Patki, 2009). The fast learning of cascade neural networks is due to the additional connection between the input layer and following layers (Osama et al., 2013).

The training algorithm of cascade neural network was divided into 4 steps as follows:

1. Originally, there are no hidden units in the network, only direct input-output connections which are trained first using the propagation algorithm (Fahlman, 1988);

2. Therefore, when no appreciable error reduction happens in network, a first hidden unit is added to the network from a pool of candidate units, which are trained independently with different random initial weights simultaneously;

3. Once installed, the hidden unit input weights are frozen, while the weights to the output units are retrained;

4. Finally, this process is repeated with each additional hidden unit, which receives input connections from both the inputs and all preceding hidden units, resulting in structure (Nechyba & Xu, 1997).

The performance of cascade neural networks and feedforward neural networks was evaluated using the mean absolute error (MAE), mean square error (MSE) and coefficient of correlation (Badde et al., 2009).

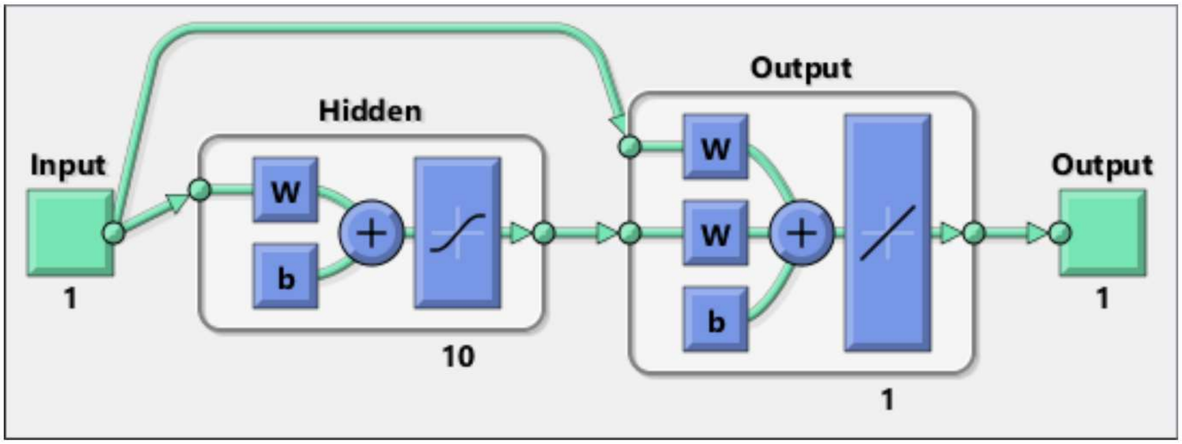


Figure 17. Cascade Neural Network Algorithm (Adopted from MathWorks)

3.4.3 Generalized Regression Neural Network (GRNN)

Generalized Regression Neural Networks (GRNN) is a combination of the feed-forward type of Neural Networks and normalized Gaussian kernels in the hidden layer as activation functions. This technique consists of the input, hidden summation, division layer and output layers. This technique therefore does not require any backpropagation algorithm due to the fact that it memorizes every unique pattern (Specht, 1991).

After training the data with sufficient training algorithms, a GRNN was able to generalize for new inputs. The GRNN output was computed using equations 3.2 and 3.3 (Al-Mahasneh et al. 2018).

$$Di = (X - Xi)^T(X - Xi) \qquad \textbf{(Equation 3.3)}$$

$$Y = \frac{\sum_{I=1}^{N} Ye^{\left(-\frac{Di}{2\sigma^2}\right)}}{\sum_{I=1}^{N} e^{(-Di/2\sigma^2)}} \qquad \textbf{(Equation 3.4)}$$

Where$Di$is the Euclidean distance between the input$Xi$and the training sample input X, Y is the training sample output and $\sigma$is the smoothing parameter of the generalized regression neural network.

This technique has some advantages, such as fast learning and convergence to the optimal regression surface as the number of samples becomes very large. In addition, due to the regression potency, this technique can be used for sparse data in a real-time environment (Cigizoglu, 2005).

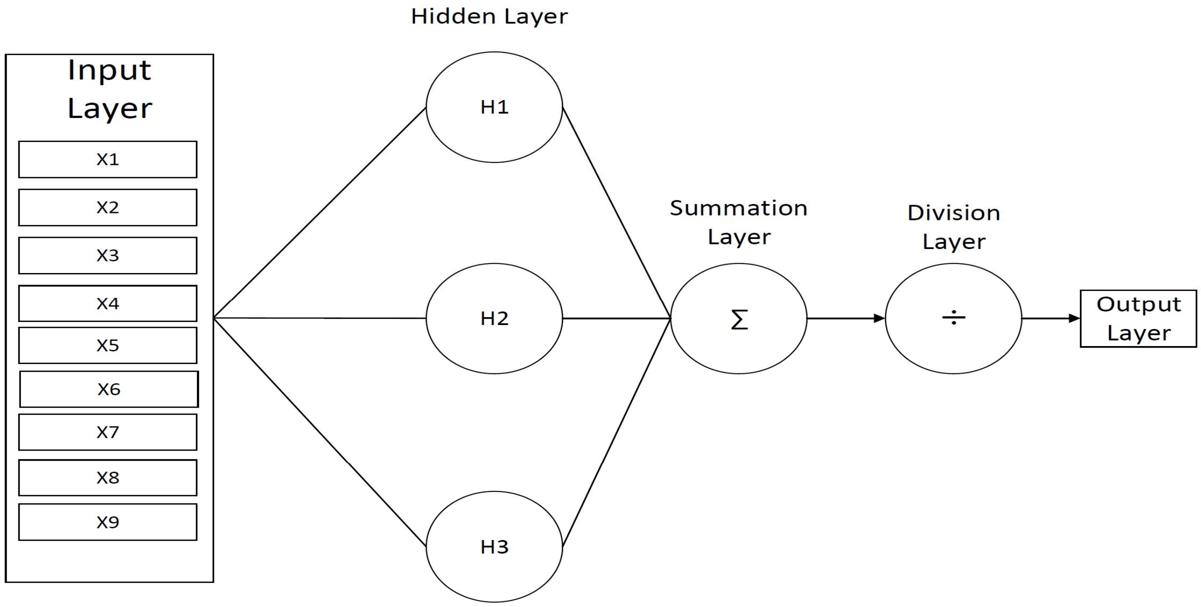


Figure 18. GRNN Structure (Al-Mahasneh et al. 2018).

Consequently, the results of the three techniques were evaluated by analysis of variance to select the developed model. Analysis of variance (ANOVA) is a statistical tool which distinguishes differences between experimental group means and is used for factorial designs. Analysis of variance is acceptable in experimental designs with one dependent variable that is a continuous parametric numerical outcome measure, and several experimental groups within one or more independent (categorical) variables.

In order to distinguish the three techniques based on analysis of variance (ANOVA), the two following parameters were considered:

P-value: the p-value is defined as the possibility of observing the given value of the test statistic which is greater or smaller than the null hypothesis. The null statistical hypothesis is the first step in the analysis of variance and states that there is no small statistical difference between the groups (Ferreira & Patino, 2015). The null hypothesis is rejected when the p-value is less than alpha. Alpha is the probability of incorrectly rejecting the null hypothesis and supposing the group means differ when in fact the groups are from a single population. By convention, alpha is typically set to 0.05 (Sawyer, 2009); and

F-value: the f-value is defined as the ratio of variance between groups to the variance within groups. A larger f-value indicates that the means of the groups differ greatly from each other compared to the variation of the individual observations in each group. If the f-value is equal or close to 1, there is no significant difference between the results and the actual data (Kim, 2014). The one-way ANOVA was implemented in this study to select the developed model.

In addition, this study selects the model based on statistical parameters such as MSE (mean squared error), MAE (mean absolute error) and RMSE (root mean square error).

3.5. Model Validation

The chief purpose of model validation was to examine model efficiency and accuracy. Model validation approves the model's ability to generalize rather than simply process the input-output relationships of the data sets. The most commonly used technique for validating artificial intelligence models compares the results of the developed model with real-world datasets that were not used in model development. The key benefit of this method is that it does not need further resources such as human judgment (Helmy, 2002). This method of validation is selected to assess and evaluate the accuracy and efficiency of the developed cascade neural network model in predicting loss of productivity due to change orders.

In order to test and validate the output compared to real data sets, three case studies from Leonard's data set are used to compare the developed model with other methods for calculating loss of productivity caused by change orders. These datasets were evaluated using the developed model and the regression models of Leonard (1988) and Ibbs (2005). The estimated loss of productivity due to change orders in the developed model, Leonard's model and Ibbs' model were compared to actual loss of productivity. Furthermore, to enrich the validation stage, the average error and absolute error of the three techniques were calculated and compared with each other.

Also, the timing model of change was created to validate the current technique. This model was built based on the value of change percent and productivity in three different time of change in projects. Finally, three real case studies were utilized to compare the developed model with Ibbs (2005) timing model.

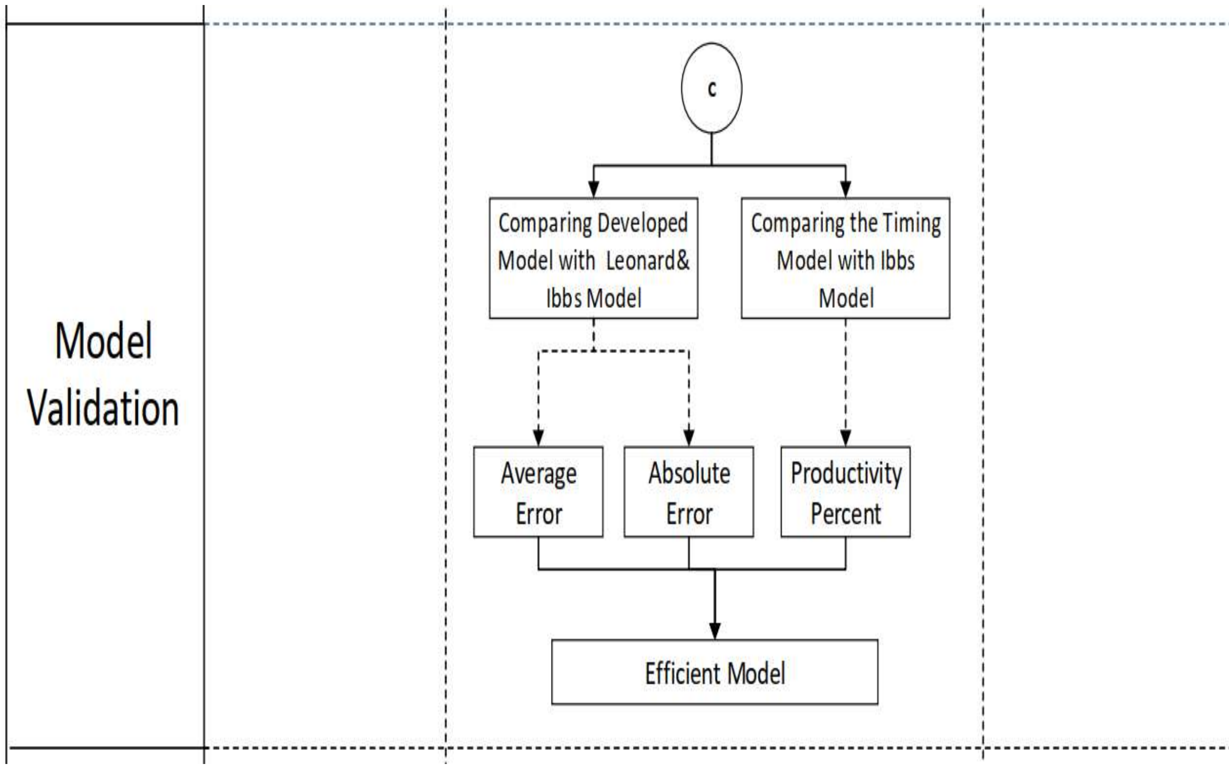


Figure 19. Summary of Proposed Model Validation

3.6. Summary

This chapter offered the research methodology used to develop the model for quantifying loss of productivity caused by change orders. In brief, the research methodology goes through several stages. Starting with the literature review, the components were expanded into eight parts to simplify the recent efforts to quantify loss of labour productivity caused by change orders. The data collection stage consisted of two datasets from previous studies using construction data sets from Leonard (1988) and Assem (2000). This study implemented data enhancement to redistribute the data set to a normal distribution and distinguish the insignificant variables from the major variables. The original data set has seventeen unique variables, and by implementing best subset regression, it is reduced to the nine most important variables.

Three artificial neural networks were used to quantify productivity loss of productivity caused by change orders. In order to select the most efficient model for this research, analysis of variance was applied to select the developed model based on the analysis of variance parameters. Consequently, three case studies from Leonard's datasets (1988) were used to compare the developed model with other methods for calculating loss of productivity due to change orders. Also, the timing model was created based on the three timing of change on project. In order to validate the timing model, the proposed model was compared with Ibbs timing model (2005) and actual productivity.

Chapter 4. Data Collection

This research amalgamates two data sets from Leonard (1988) and Assem (2000). A total of 123 data sets were generated by combining these two data sets, which can be considered an acceptable data volume for developing an artificial intelligence model for loss of productivity quantification, as will be shown in Chapter 5. In this chapter, a brief description of the architectural, civil, mechanical and electrical projects included in the data sets is presented.

This chapter is divided into three steps:

Step 1. Data Characteristics;

Step 2.Data Normalization and Variable Selection; and

Step 3. Data Timing Categorization.

4.1. Data Characteristics

Table 2. Distribution of 123 Data sets based on the type of work.

Number of Projects			**Available Data sets**	**Total of Available Data sets**
Leonard Research (1988)	66 M/E Data Sets	24 C/A Data Sets	90 Data Sets	123 Data Sets
Assem (2000)	30 M/E Data Sets	3 C/A Data Sets	33 Data Sets	

Based on Table 2, 78% of the collected data sets are mechanical and electrical projects, and 22% of the collected data sets are civil and architectural projects.

4.1.1. Datasets Size

In this study, datasets size was defined as the actual work hours employed at project completion, including change order hours. By merging all 123 data sets, it was found that the actual hours of the entire set total about 7,683,646 work hours. Figures 20 and 21 show the histograms of project size for all the architectural/civil and mechanical/electrical data sets, respectively. The mean project size for architectural/civil projects was 35,600 work hours while the mean project size for mechanical/electrical projects was 43,000 work hours.

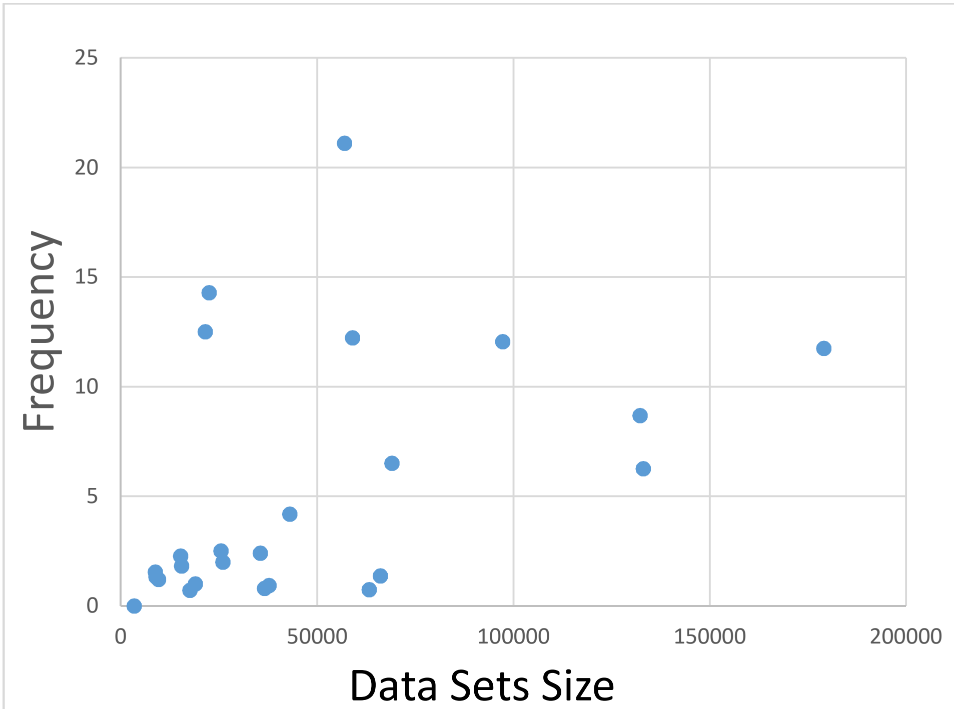


Figure 20. Distribution of Data Sets Size for Architectural/Civil Projects.

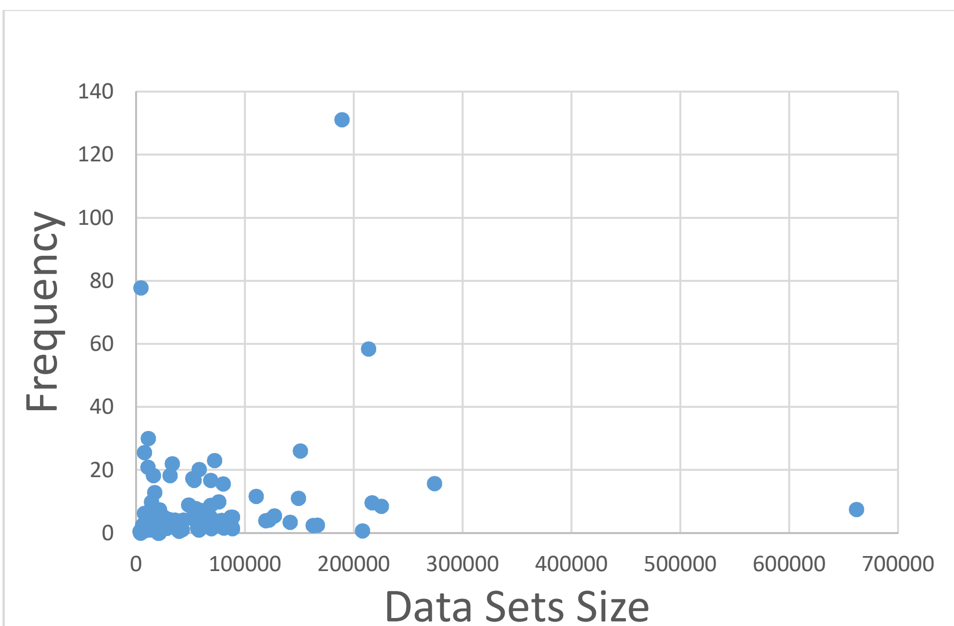


Figure 21. Distribution of Data Sets Size for Mechanical/Electrical Projects.

4.1.2. Data Set Contracts

Figure 22 shows the distribution of projects by four types of project contracts: lump sum, unit price, fixed price and cost plus. The majority of projects were lump sum contracts.

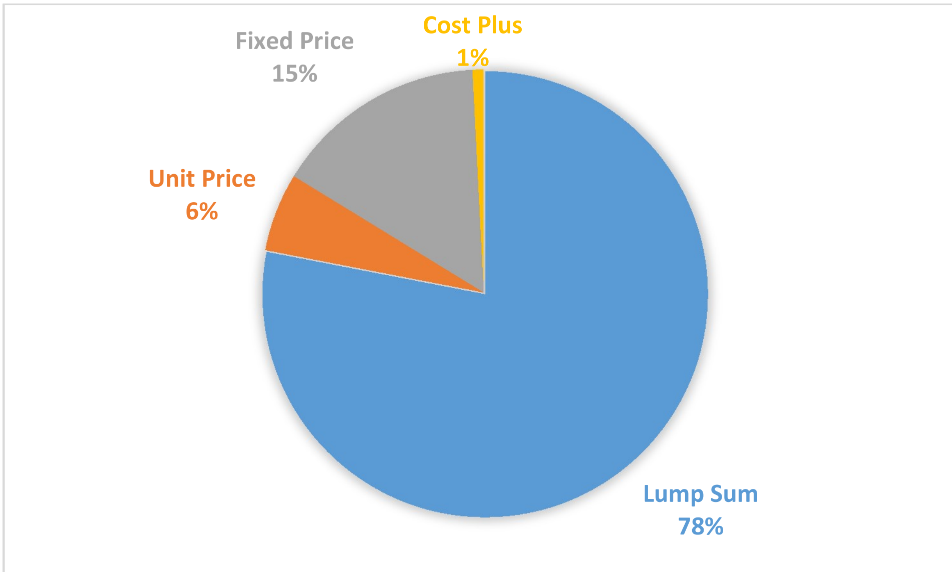


Figure 22. Distribution of Data Sets Contracts

The results show that lump sum contracts make up the largest percentage among contracts in this study.

4.1.3. Change Order Values

In this research, the value of change orders varied from $3,200 to $8,857,000. Figure 23 shows that 18% of the data sets had change order values of more than $1,000,000 and 12% of the combined datasets had change order values of between $500,000 and $1,000,000.

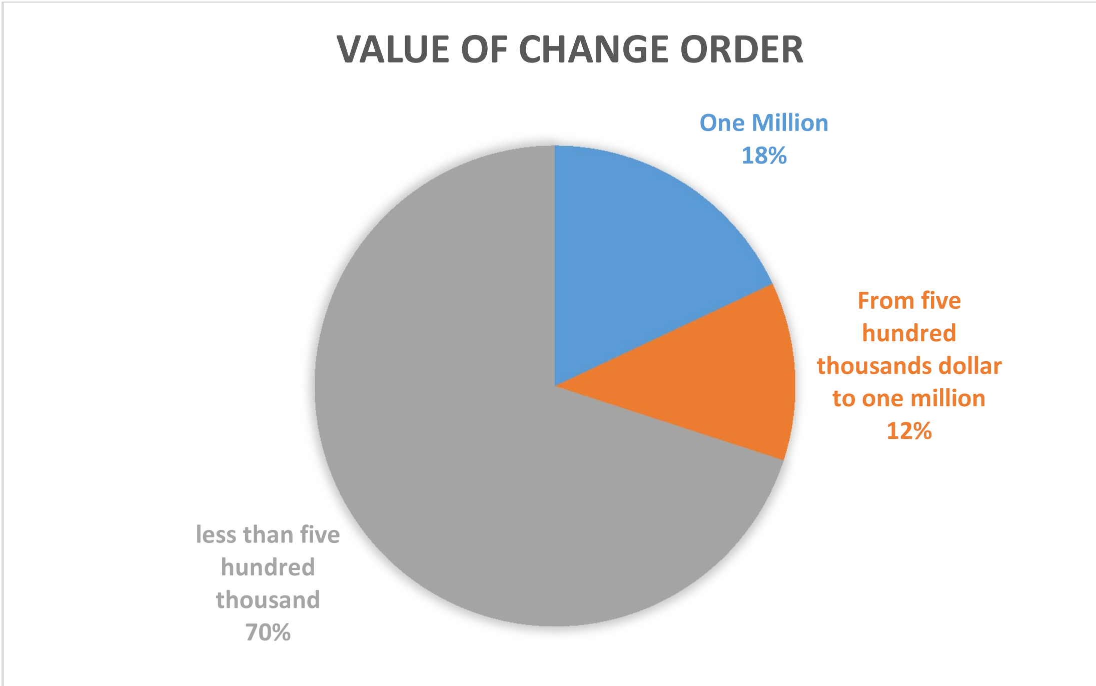


Figure 23. Value of Change Order

4.1.4. Reasons for Change Order

Change orders can occur for numerous reasons. Figure 24 shows the distribution of the data sets included in the combined datasets with respect to the reasons for change orders. The chart illustrates that 56% of change orders were due to design changes, 21% were due to incomplete designs, 6% were due to new technology and the rest were due to design errors, unforeseen conditions, late design completion, and the reworking of defective equipment.

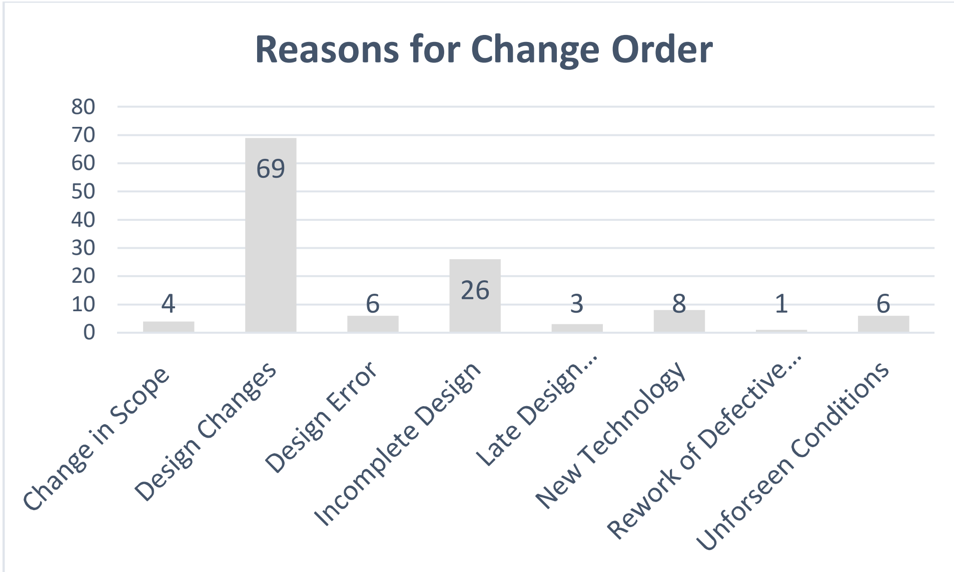


Figure 24. Reasons for Change Order

4.1.5. Decision Variables

The combined data sets have 17 unique parameters with diverse types and scales which are as follows:

1- Type of Impact, 2- Type of Work, 3- Type of Contract, 4- Value of Original Contract, 5- Value of Change Orders, 6- Original Duration, 7-Actual Duration, 8- Extended Duration, 9- Original Estimated Hours, 10- Experience Factor, 11- Actual Hours, 12- Number of Change Orders, 13- Change Order Frequency, 14- Change Orders Hours, 15- Average Size, 16- Change Order Percent, 17- Earned Hours.

The 17 variables in this study have a diverse range and differing values.Therefore, the input variables had to be normalized. Table 3 shows the diverse range and values of the variables.

Table 3. Descriptive Statistics of Combined Data sets.

Variables	Mean	Standard Deviation	Variance	Minimum	Median	Maximum	Range
Type of Impact	1.5528	0.7266	0.5279	0.0000	1.0000	3.0000	3.0000
Type of Work	2.358	1.415	2.002	1.000	2.000	5.000	4.000
Type of Contract	1.5203	1.0739	1.1533	1.0000	1.0000	5.0000	4.0000
Value of Original Contract	2761539	3241535	1.05076E+13	80000	1694000	23172000	23092000
Value of Change Orders	607634	1157541	1.33990E+12	0	250000	8857000	8857000
Original Duration	11.407	6.037	36.440	1.000	10.000	28.000	27.000
Actual Duration	16.485	8.298	68.858	1.800	16.000	44.000	42.200
Extended Duration	0.5569	0.6165	0.3801	-0.1500	0.4067	4.5000	4.6500
Original Estimated Hours	39677	57499	3306157991	1087	25000	557000	555913
Experience Factor	1.2068	0.2816	0.0793	0.7600	1.0900	2.7800	2.0200
Actual Hours	62469	76833	5903297716	3275	43000	661600	658325

Number of Change Orders	101.2	205.7	42312.0	0.0	70.0	2150.0	2150.0
Frequency	8.15	15.14	229.23	0.00	3.89	131.10	131.10
Change Orders Hours	9970	11633	135326276	0	5863	83000	83000
Average Size	158.7	243.1	59096.8	0.0	97.2	1547.6	1547.6
Change Orders Percent	0.1909	0.1416	0.0200	0.0000	0.1393	0.5730	0.5730
Earned Hours	50680	66908	4476711492	1265	33300	557000	554100

4.2. Data Normalization and Variable Selection

This stage is divided into two steps:

Step 1. Data Normalization; and

Step2. Variables Selection.

4.2.1. Data Normalization

The two main objectives for the data normalization and variable selection step were the normalization of the available data sets based on the distribution of the combined datasets and the elimination the insignificant variables based on variable selection parameters.

In this study, the available data sets were not normally distributed. For example, the data set range for one of the variables was from \$80,000 to \$23,172,000 while the range of another is from 0.76 to 2.78. Therefore, to redistribute the available data sets to a normal distribution, the min-max normalization technique was used (Jain & Bhandare, 2011).

$$Z = \frac{(Ai - Min(A))}{(Max(A) - Min(A))} \quad \textbf{(Equation 4.1)}$$

Where,

Z= Normalized data,

Ai= Original data,

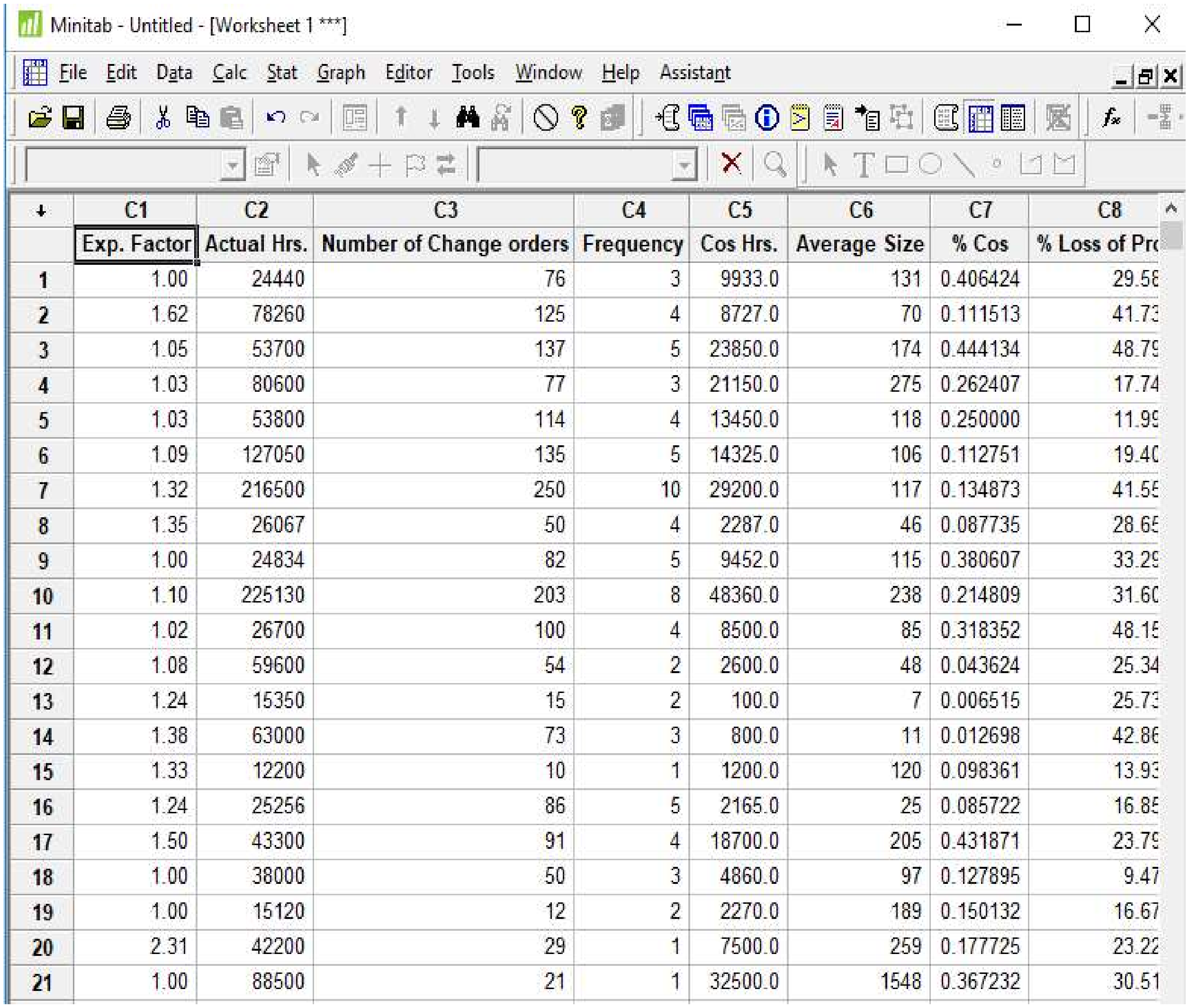
Max (A) = Max range of original data,

Min (A) = Min range of original data.

As shown in Table 18 in Appendix 1 and Table 20 in Appendix 2, the available data sets are not distributed appropriately. Table 19 in Appendix 1 and Table 21 in Appendix 2 show the normalized datasets, which are distributed properly.

Subsequently, in order to distinguish the insignificant variables from the significant variables, this research utilized best subset regression in Minitab software with variable selection parameters. The parameters for selecting the significant variables are as follows: R-squared, adjusted R-squared and Mallow's Cp.

Figure 25 shows the input variables in Minitab Software. Seventeen variables were implemented in Minitab to discriminate the insignificant and significant variables.



↓	C1	C2	C3	C4	C5	C6	C7	C8
	Exp. Factor	Actual Hrs.	Number of Change orders	Frequency	Cos Hrs.	Average Size	% Cos	% Loss of Pr
1	1.00	24440	76	3	9933.0	131	0.406424	29.58
2	1.62	78260	125	4	8727.0	70	0.111513	41.73
3	1.05	53700	137	5	23850.0	174	0.444134	48.79
4	1.03	80600	77	3	21150.0	275	0.262407	17.74
5	1.03	53800	114	4	13450.0	118	0.250000	11.99
6	1.09	127050	135	5	14325.0	106	0.112751	19.40
7	1.32	216500	250	10	29200.0	117	0.134873	41.55
8	1.35	26067	50	4	2287.0	46	0.087735	28.65
9	1.00	24834	82	5	9452.0	115	0.380607	33.29
10	1.10	225130	203	8	48360.0	238	0.214809	31.60
11	1.02	26700	100	4	8500.0	85	0.318352	48.15
12	1.08	59600	54	2	2600.0	48	0.043624	25.34
13	1.24	15350	15	2	100.0	7	0.006515	25.73
14	1.38	63000	73	3	800.0	11	0.012698	42.86
15	1.33	12200	10	1	1200.0	120	0.098361	13.93
16	1.24	25256	86	5	2165.0	25	0.085722	16.85
17	1.50	43300	91	4	18700.0	205	0.431871	23.79
18	1.00	38000	50	3	4860.0	97	0.127895	9.47
19	1.00	15120	12	2	2270.0	189	0.150132	16.67
20	2.31	42200	29	1	7500.0	259	0.177725	23.22
21	1.00	88500	21	1	32500.0	1548	0.367232	30.51

Figure 25. The Input variables in Minitab Software

Afterward, the seventeen input variables were evaluated by Best Subset Regression parameters to select the significant variables. As shown in table 4 the insignificant variables were eliminated in Minitab Software with Best Subset Regression Technique.

Table 4. Best Subset Models for Different Number of Parameters

Model	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
R-Sq	18.9	25.4	28.4	33.6	36	37	38.2	39.3	39.6	39.9	40.1	40.2	40.3	40.3	40.4	40.4
Adj R-Sq	18.2	24.2	26.6	31.4	33.3	33.8	34.4	35.1	34.8	34.6	34.1	33.7	33.2	32.6	32	31.4
Mallows CP	25.2	15.5	12.2	4.9	2.7	2.9	2.9	2.9	4.3	5.8	7.5	9.2	11.1	13	15	17
Type of Impact	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Type of Work						*	*	*	*	*	*	*	*	*	*	*
Type of Contract								*	*	*	*	*	*	*	*	*
Value of Original Contarct									*	*	*	*	*	*	*	*
Value of Change Orders														*	*	*
Original Duration													*	*	*	*
Actual Duration									*	*	*	*	*	*	*	*
Extended Duration															*	*
Original Estimated Hours				*	*	*	*	*	*	*	*	*	*	*	*	*
Earned Hours			*	*	*	*	*	*	*	*	*	*	*	*	*	*
Exprience Factor					*	*	*	*	*	*	*	*	*	*	*	*
Actual Hrs				*	*	*	*	*	*	*	*	*	*	*	*	*
Number of Change Orders			*											*	*	*
Frequency							*	*	*	*	*	*	*	*	*	*
Change order Hours												*	*	*	*	*
Average Size												*	*	*	*	*
Change Order Percent	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*

4.2.2 Determining the Significant and Insignificant Variables

The objective of this section is to determine the significant and insignificant variables to select the optimal model using best subset regression. Table 4 shows that the resulting models pertaining to the different numbers of parameters were simultaneously compared to one another. The selected variables can be revealed based on various criteria, including R-squared, adjusted R-squared and Mallows's Cp. The model with a low Mallows's Cp and relatively high R-squared and adjusted R-squared values was chosen as optimal model for final consideration. Cp is an unbiased estimator of ordinary least square error (Madigan & Ridgeway, 2004). A small Mallows's Cp value indicates that the model is relatively precise (Gilmour, 1996). Although model 16 had the highest R-squared value among all models, the adjusted R-squared value for this model was lower than that of model number eight. In addition, model number five had a lower Mallows's Cp value than other models, its R-squared and adjusted R-squared values were very low compared to those of other models.

Consequently, for model eight, the R-squared and adjusted R-squared values were found to be 39.3 and 35.1, respectively. As mentioned previously, model eight is more desirable than other models because it has higher R-squared and adjusted R-squared values than the other models. Furthermore, the Mallows's Cp value in model number 8 is lower than that of other models. The selected variables in the best subset regression are the type of impact, type of work, type of contract, original estimated hours, experience factor, actual hours, Earned Hours ,change order frequency, and change order percent.

4.3. Data Timing Categorization

In order to compare the developed model with previous studies such as regression, this study created the timing model to compare the results with the Ibbs (2005) model. This model separated the time of change order into three measures. To divide the timing of a change order, the datasets were ranked and ordered into three groups as follows:

1- Early Change: The change which occurs at the 0 to 25 % project's period;

2- Normal Change: The change which occurs at the 26 to 75 % project's period; and

3- Late Change: The change which occurs at the 76 to 100% project's period (Ibbs, 2005).

Therefore, the Leonard (1988) and Assem (2000) datasets were used to create the timing model and compared it with Ibbs (2005) model. According to Equation 4.2, the timing of the change in the Leonard datasets can be recognized by calculating the percent complete work (Eldin, 1989).

Earned Hours = Original Estimated Hours x Percent Complete Work **(Equation 4.2).**

The percent complete of work shows when the change occurs at each period. The percent complete work value of less than 25% shows that the change happened at an early stage, the value between 26 to 75% shows that the change happened at middle (Normal) stage and the value between 76 to 100% demonstrates that the change happened at the final (Late) stage (Ibbs, 2005).

As shown in Table 23 Appendix 4, the Leonard datasets was divided into three separate timing stages. As seen in Table 5, the percent complete work shows that most changes in Leonard datasets happened at the late and normal stage.

Table 5. Number of Available Datasets in Leonard's Datasets

Type of Change	**Number of Datasets**
Early Change	7 Datasets
Normal Change	7 Datasets
Late Change	26 Datasets

However, in Leonard's datasets 50 datasets were not considered due to inappropriate value of percent complete work.

In addition, in order to recognize the time of the change in Assem study, this study categorized the time of the change with change orders direct hours. This study divided the change order hours into five separate times, and as shown in Table 24 in Appendix 4, the highest value of change order hours among five change order hours periods demonstrates that a change occurred at one of the three timing stages.

Table 6 illustrates that the most of the changes in Assem's datasets occurred at normal and late stages.

Table 6. Number of Available Datasets in Assem's Datasets

Type of Change	Number of datasets
Early Change	5 Datasets
Normal Change	20 Datasets
Late change	8 Datasets

4.4. Summary

This chapter describes the data collection, which was divided into three steps. Step 1 provides a detailed description of the 123 datasets included in the combined datasets. The combined data sets were collected from the studies by Leonard (1988) and Assem (2000). The data sets in the current study were relatively large, which allows for the development of an artificial intelligence model for loss of productivity quantification.

Step 2 considered the use of data enhancement to normalize the available data sets and discriminate between the significant and insignificant variables. Available data sets were transferred to a normal distribution using the normalization technique. The significant and insignificant variables were separated using the best subset regression technique in Minitab. This technique eliminated eight insignificant variables based on statistical parameters.

Finally, in step 3 the leonard and Assem datasets were divided into three timing stages based on the percent complete of work and highest value of change order hours. Table 5 and Table 6 illustrate that the 73 datasets were collected by combining these two datasetes, which can be considered a satisfactory data volume for developing timing model.

Chapter 5. Model Development

5.1. Overview

This Chapter outlines how the methods described in previous sections were used to model loss of productivity due to change orders by developing three artificial neural network techniques. The model development included two steps:

Step 1. Artificial neural network techniques; and

Step 2. Model selection.

5.2. Model Development

In this research, three techniques were utilized to quantify loss of productivity caused by change orders. They are the feedforward neural network (FNN), cascade neural network (CNN) and generalized regression neural network (GRNN).

5.2.1. Brief Introduction of Artificial Neural Network in Matlab Software

The three techniques were developed in Matlab Software. Matlab's Neural Network Toolbox was used to create the neural network and aid it in training and testing the impacted data set.

The complete structure of the neural network is shown in Figure 26. In this figure, 1 is the input variables. 10 is the number of neurons in the hidden layer, and the in the output is the number of output results, which is estimated loss of productivity.

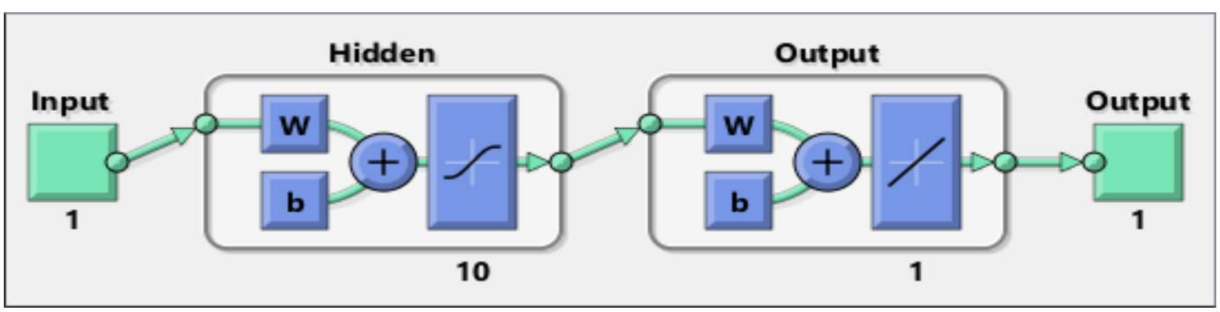


Figure 26. Complete Structure of Neural Network in Matlab (Adopted from Mathworks).

Matlab provides users with three algorithms: the Levenberg-Marquardt algorithm, Bayesian regularization and the scaled conjugate gradient algorithm. Matlab recommends the Levenberg-

Marquardt algorithm for most problems. Conversely, the Bayesian regularization algorithm is better suited to small data sets, but can require more processing time. If the datasets is large, Matlab recommends the scaled conjugate gradient algorithm, which takes less memory and is more efficient than the other two techniques (Ai, 2016). In order to find the best training algorithm for the given datasets all three algorithms algorithms were used:

- Levenberg-Marquardt
- Bayesian Regularization
- Scaled Conjugate Gradient algorithm

As shown in Table7, the Levenberg-Marquardt algorithm was selected due to having the highest processing speed and efficiency despite requiring more memory than the other two algorithms.

Table 7. Different Statistical Performance of the three Algorithms

Techniques	**Levenberg-Marquardt Train R-Sq**	**Bayesian-Regularization**	**Scaled-Conjugate Gradient algorithm**
Feed Forward Neural Network	80.88%	76.07%	36.07%
Cascade Neural Network	83.51%	34.17%	46.95%

Figure 27 indicates the criteria for estimating whether the neural network has or has not been successfully trained. In the regression diagram, R measures the correlation coefficient between the inputs and targets. An R-value of 1 shows that there is a strong relationship between the inputs and targets, while an R-value of 0 means that there is a random relationship between the inputs and targets (Ai, 2016). In addition, the mean squared error (MSE), which represents the average squared difference between the outputs and targets, can measure the neural network. The ideal MSE is 0. The lower the MSE, the higher the prediction accuracy.

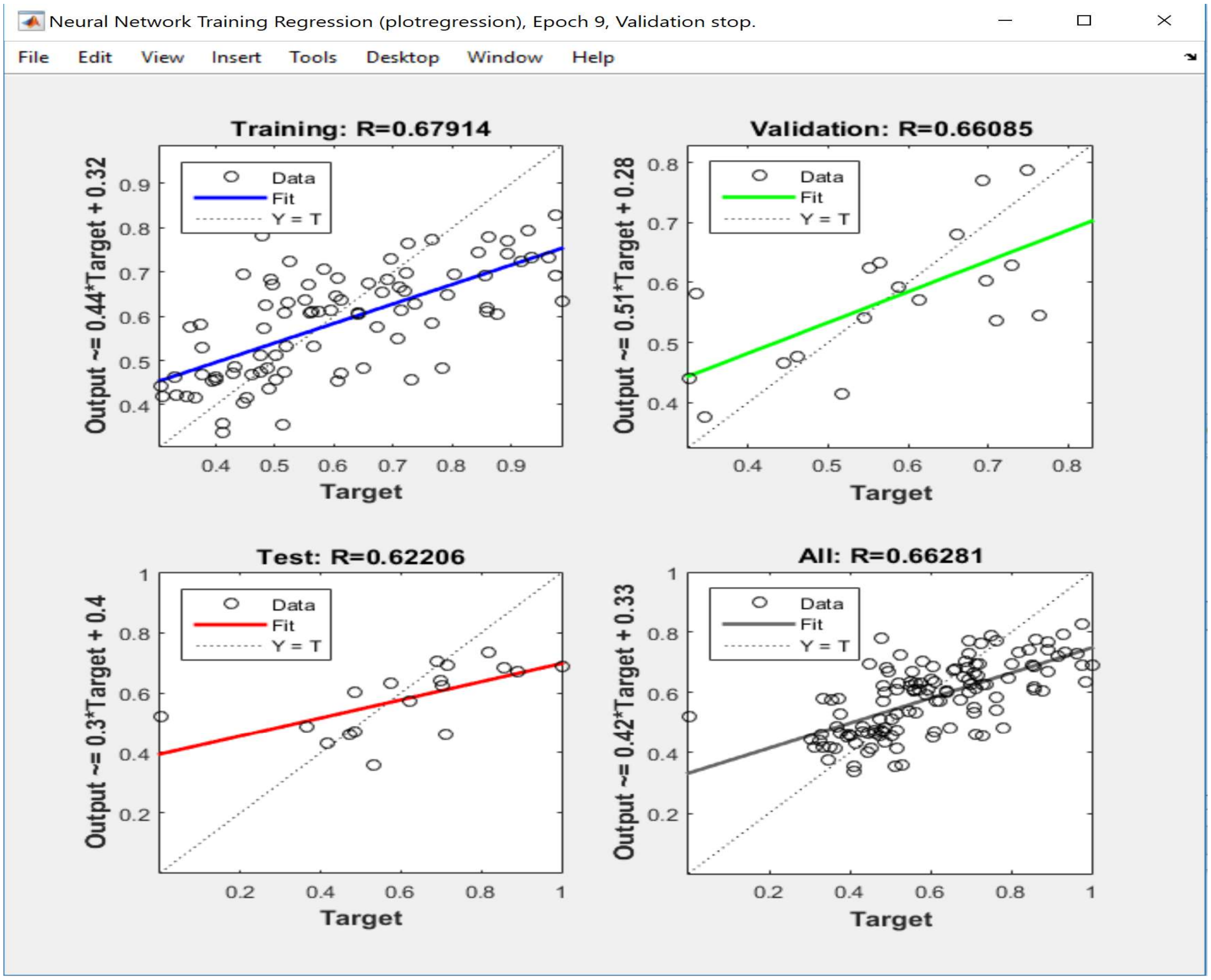


Figure 27. Regression Results in Matlab Software

5.2.2. Feed Forward Neural Network

As mentioned previously, one of the three techniques used in this research is the feedforward neural network. In this technique, the nine neural network inputs are type of impact, type of work, type of contract, original estimated hours, experience factor, actual hours, change order frequency, change order percent, and earned hours. The neural network's output is the estimated loss of productivity.

In the training process, the Matlab software allocated 70% of the 123 samples for training, 15% for validation and 15% for testing. The training samples were given to the network during training, and the network was adjusted according to its error. The validation samples were used to measure network generalization and to stop training when generalization stopped enhancing. The test samples do not affect training and thus provide an independent measure of network performance during and after training (Ai, 2016).

Once the framework of the feedforward neural network was established, the results after running were satisfactory due to the high value of R (correlation coefficient) and low value of the mean squared error (MSE).

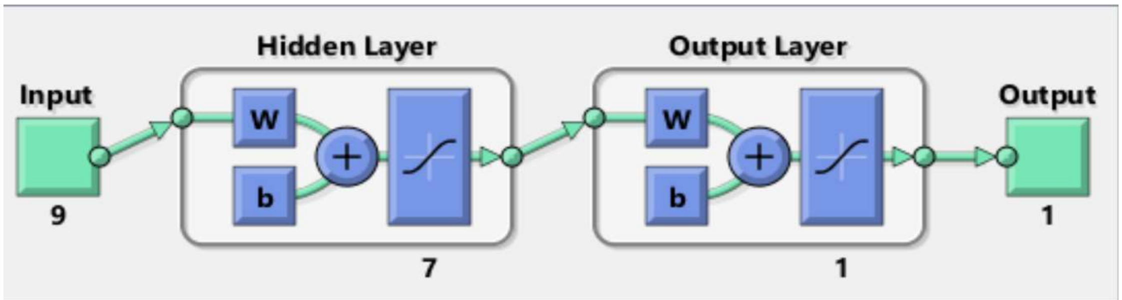


Figure 28. Feed Forward Neural Network Analysis Structure in Matlab Software

The feedforward neural network's structure contains nine input parameters (type of impact, type of work, type of contract, original estimated hours,experience factor, actual hours, frequency, change order percent and earned hours) and one hidden layer composed of seven hidden nodes. The number of neurons within the hidden layer is dependent on the number of input parameters within the input layer. The number of neurons in the hidden layer was determined to be about 75% of the number of input parameters in the input layer (Salchenberger, Cinar, & Lash, 1992).

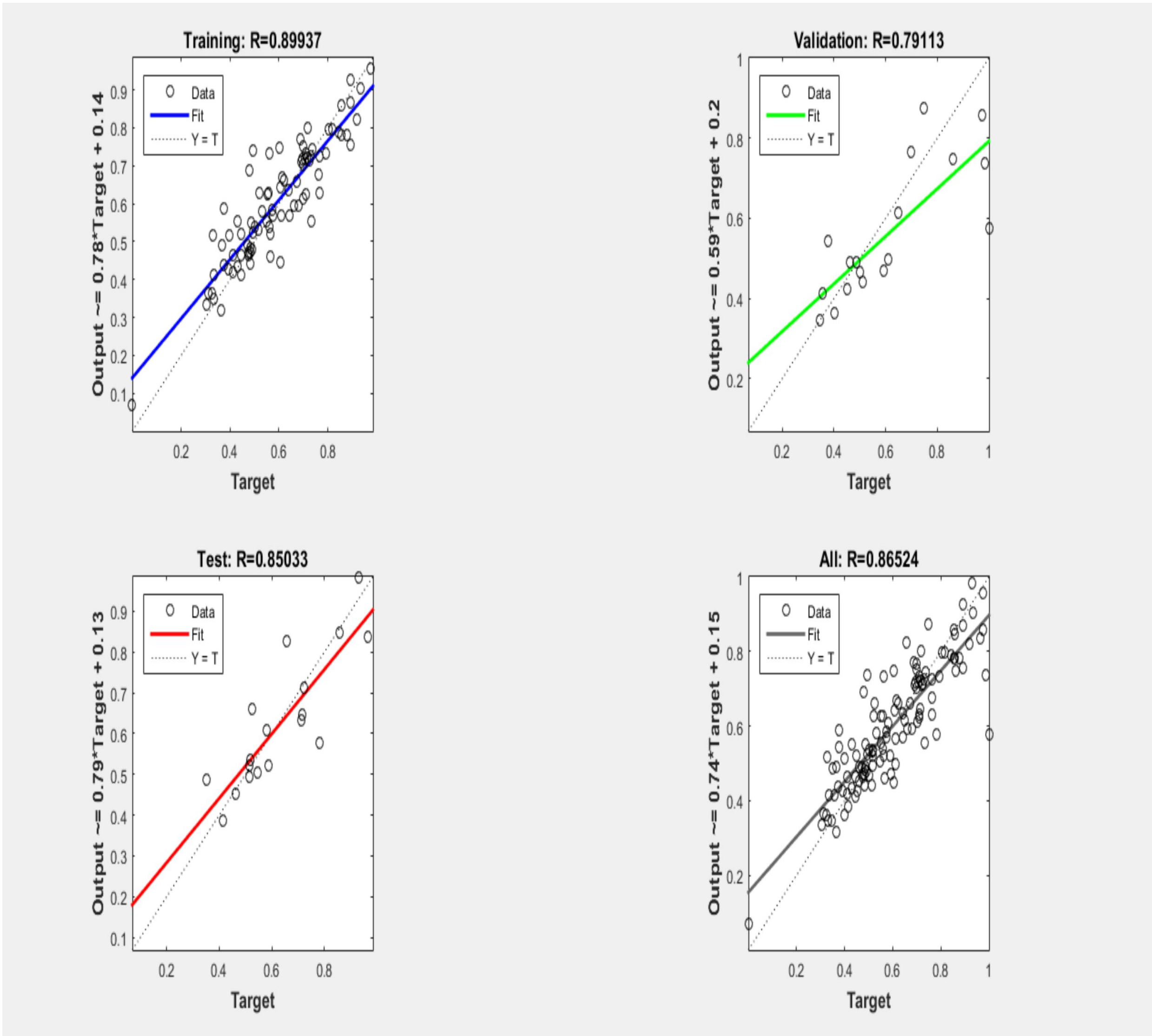


Figure 29. Feed Forward Neural Network Regression Results

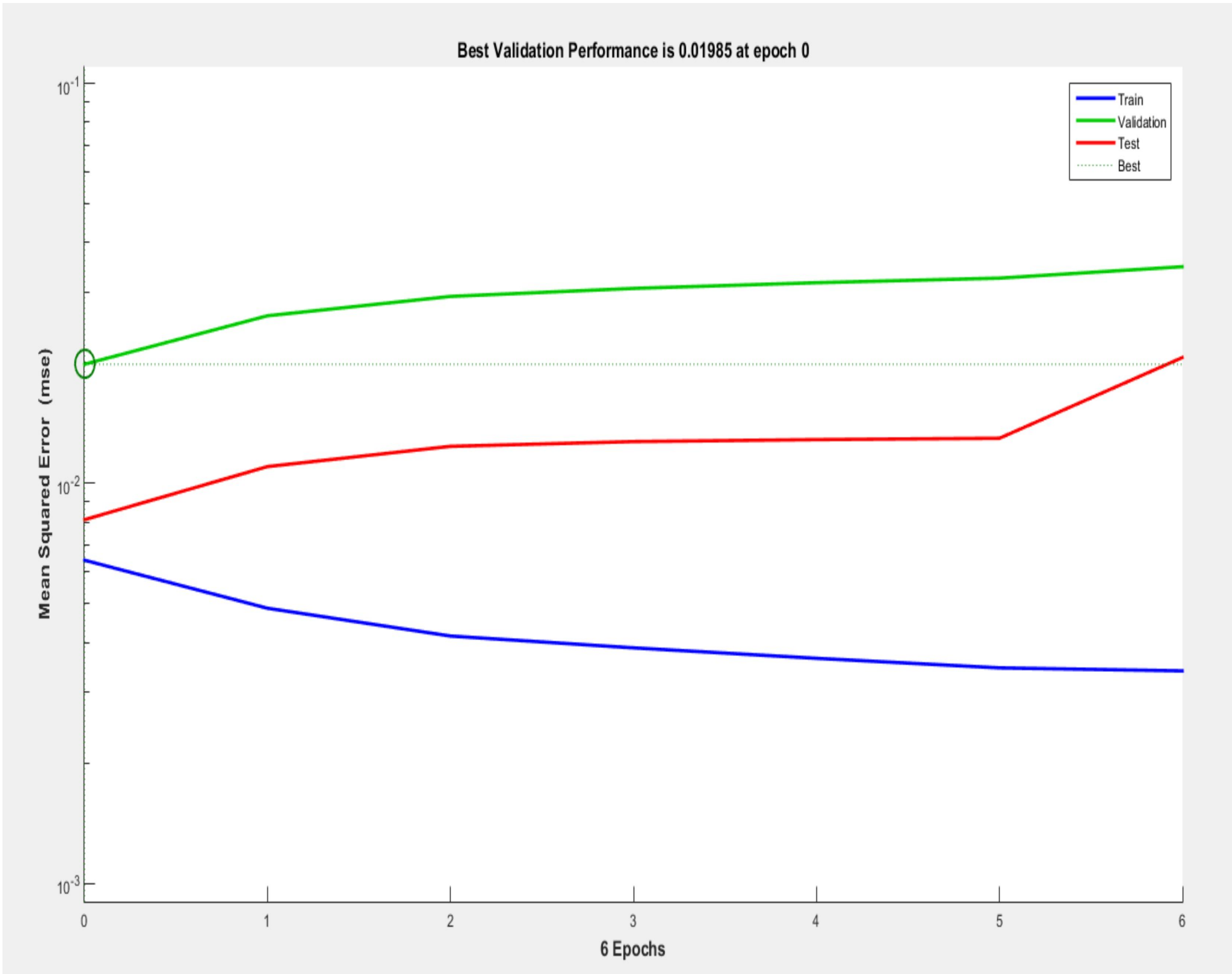


Figure 30. Feed Forward Neural Network Performance

Figures 29 and 30 illustrate that the results of the feedforward neural network were satisfactory due to the high percentage of R (correlation coefficient) in Training (89.935%) and Testing (85%) as well as the low value of MSE in the network.

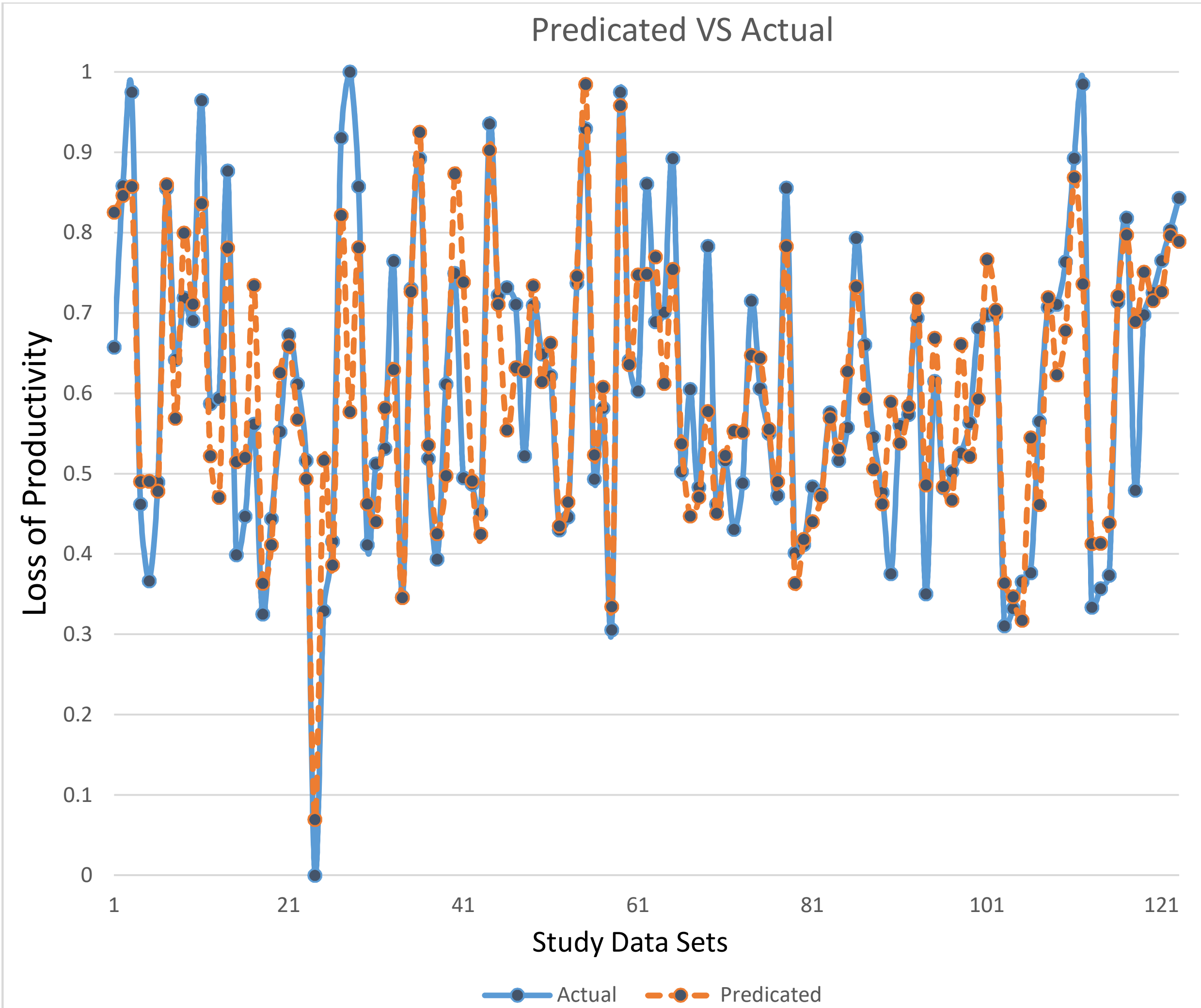


Figure 31. Actual Loss of Productivity VS Predicated Loss of productivity.

5.2.3. Cascade Neural Network (CNN)

The second artificial neural network technique used in this study is the cascade neural network (CNN). Cascade-forward network is similar to feedforward network, but include a connection from the input and every previous layer to the following layers (MathWorks, 2018). This technique implements nine inputs (type of impact, type of work, type of contract, original estimated hours,experience factor, actual hours, frequency, change order percent and earned hours) to estimate loss of productivity.

Cascade neural network starts learning with only one neuron. During learning, the algorithm automatically adds and trains new neurons, creating a multi-layer structure. The number of hidden neurons, or in other words the complexity of the network, increases gradually while the training error decreases. Therefore, the training algorithm develops a neural network of near-optimal complexity which can satisfactorily generalize (Schetintin, 2005). Two algorithms can be used to train this technique: the Levenberg-Marquardt algorithm and the Bayesian regulation backpropagation algorithm (Gaikwad & Thool, 2014).

Due to the similarity between the feedforward neural network and the cascade neural network, the results were referred based on the high R (correlation coefficient) value and the low mean squared error (MSE) value. The results extracted in Matlab showed that this technique is much more effective than the feedforward neural network.

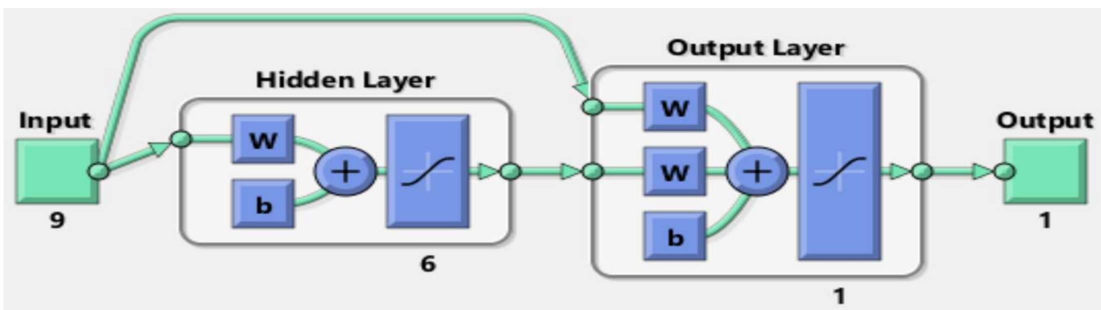


Figure 32. Cascade Neural Network Analysis Structure in Matlab Software

By applying Cascade Neural Network, the network structure encompasses nine input parameters (type of impact, type of work, type of contract, original estimated hours,experience factor, actual hours, frequency, change order percent and earned hours) and one hidden layer composed of six nodes. This technique was implemented with the Levenberg-Marquardt algorithm to estimate loss of productivity.

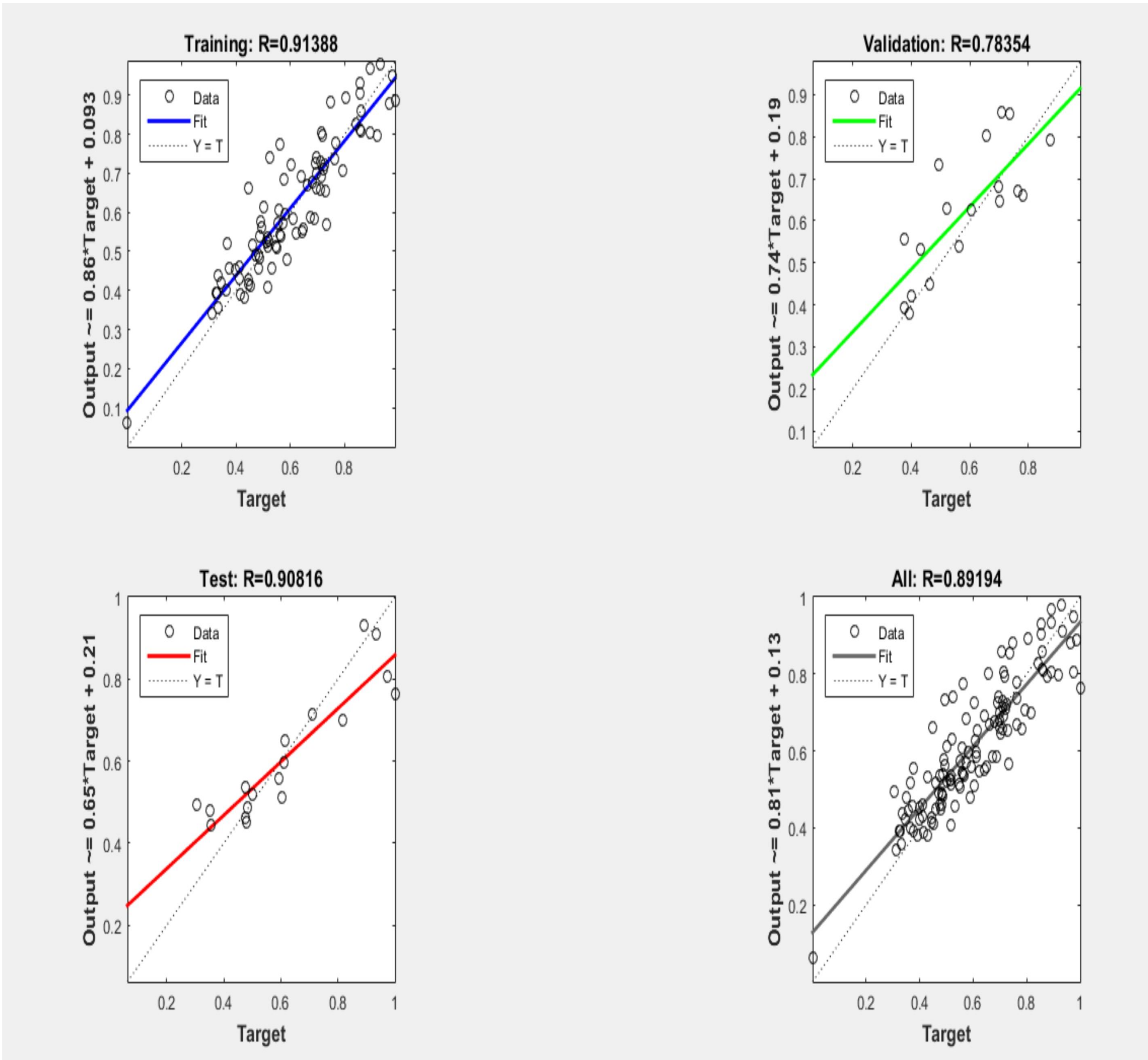


Figure 33. Cascade Neural Network Regression Results

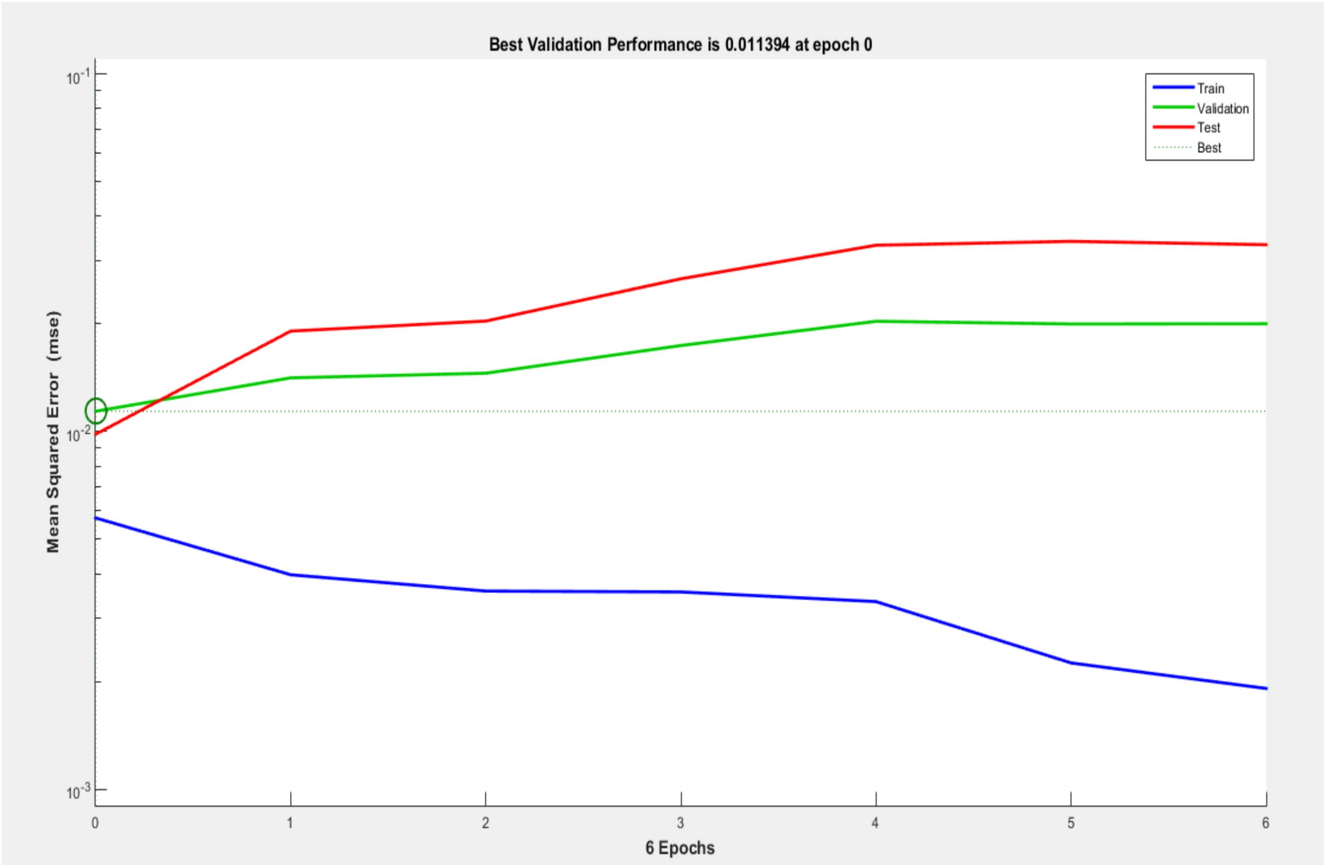


Figure 34. Cascade Neural Network Performance

Figures 33 and 34 clarify that the results of the cascade neural network are much more acceptable than those of the feedforward neural network due to the high percentage of the R (correlation coefficient) in Training (91.38%) and Testing (90.81%) as well as the network's low MSE value.

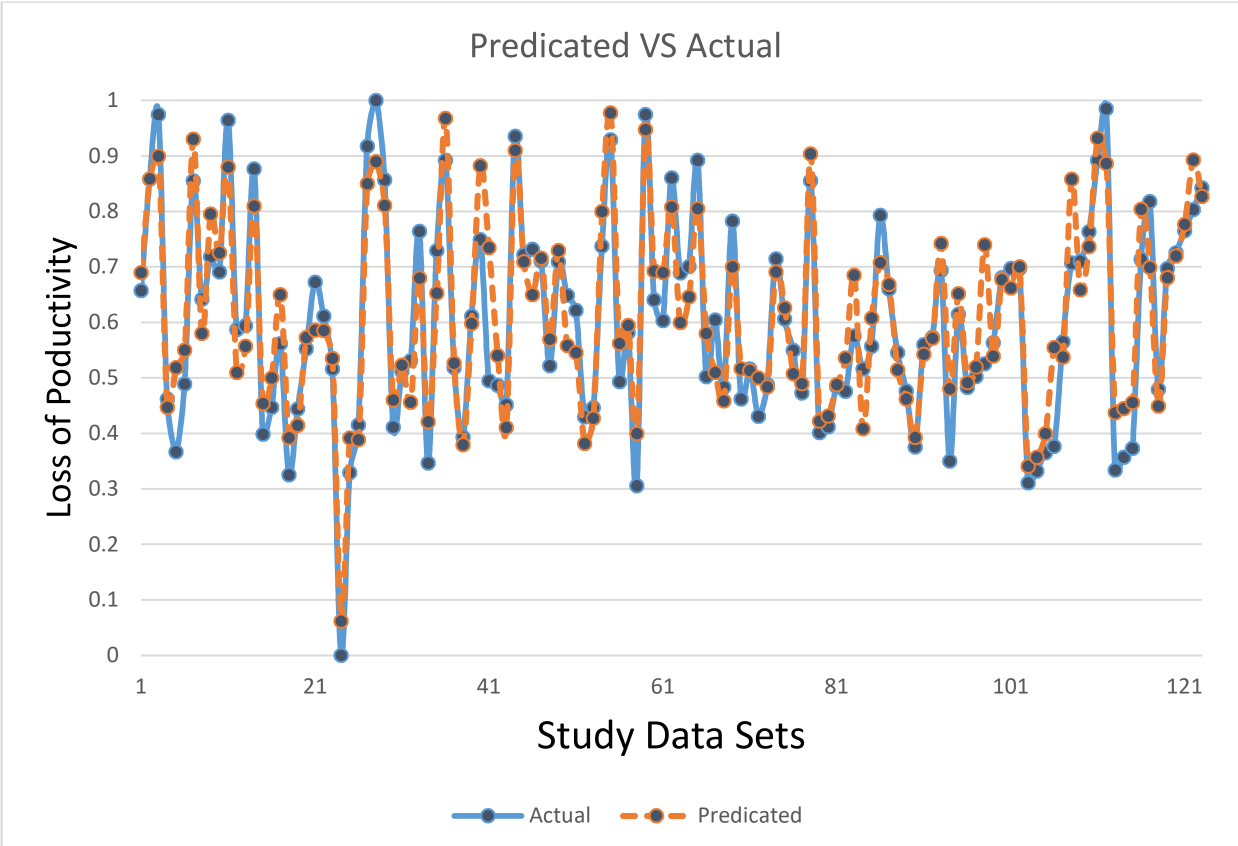


Figure 35. Actual Loss of Productivity VS Predicated Loss of productivity.

Figure 35 illustrates that the predicted loss of productivity is very close to the actual loss of productivity. Furthermore, the mean squared error (MSE) is much lower than those found in the previous literature and techniques.

5.2.4. Generalized Regression Neural Network (GRNN)

The last technique used in this research is the generalized regression neural network. Like previous techniques, this network has nine input parameters (type of impact, type of work, type of contract, original estimated hours,experience factor, actual hours, frequency, change order percent and earned hours) in order to estimate the loss of productivity. This technique was implemented in DTREG predictive modeling software.

DTREG is a robust application that is easily installed on any Windows system. DTREG reads comma-separated value (CSV) data files that are easily created from almost any data source. Once a data file is created, it can be fed into DTREG, which will do the work of creating decision trees, support-vector machines, K-means clustering, linear discriminant functions, linear regressions or logistic regression models (DTREG, 2018).

In the training process, DTREG allocated 80% of the 123 samples for training, 10% for validation and 10% for testing. Once the framework of the generalized regression neural network was extracted from DTREG, the results after running were referred due to the high R-squared (coefficient of determination) value for Training and Testing. The low MSE value is another parameter used to select the finest results (Del Rosario et al., 2016).

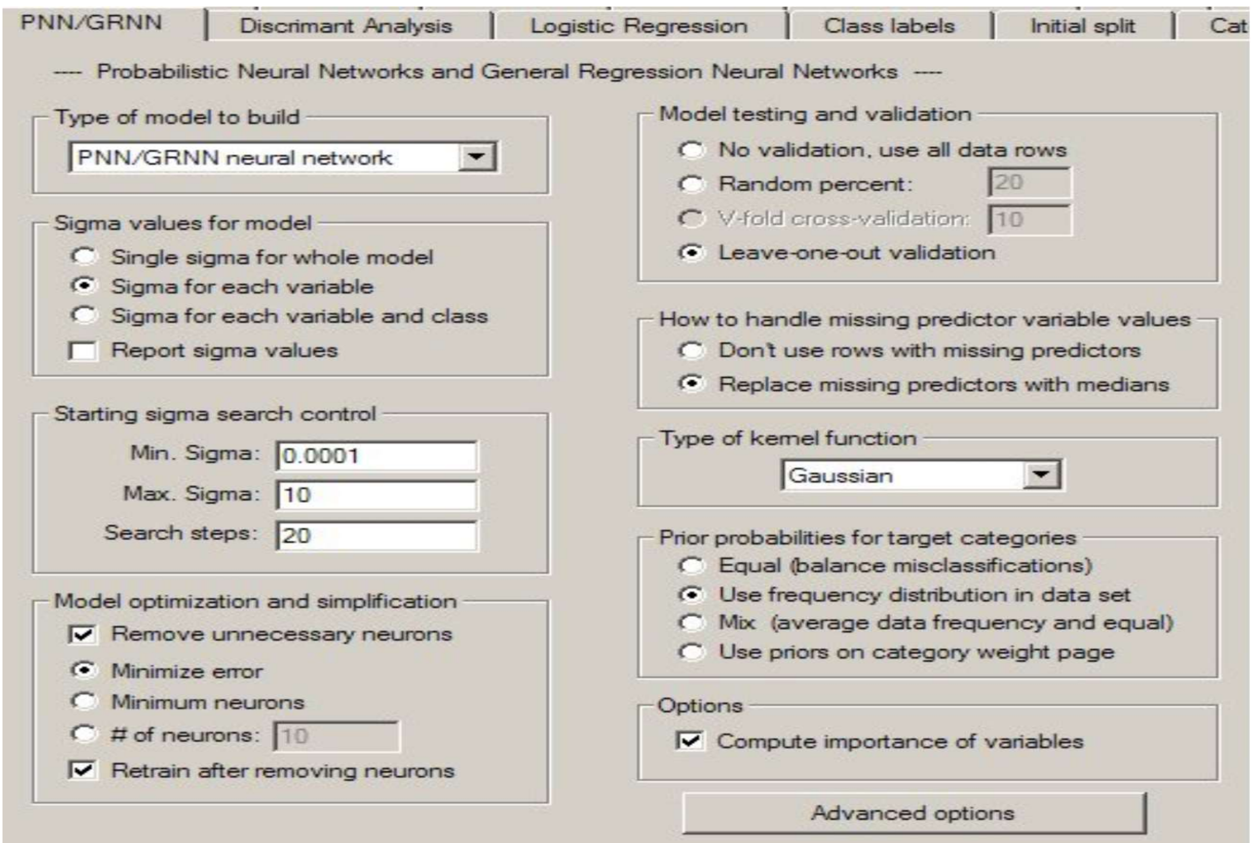


Figure 36. A Sample Screen of DTREG Software

To apply the GRNN, the kernel function must to be taken into consideration. In this study, the Gaussian function is applied as a kernel function. The generalized regression neural network's structure has nine input parameters (type of impact, type of work, type of contract, original estimated hours,experience factor, actual hours, frequency, change order percent and earned hours) and 26 hidden nodes.

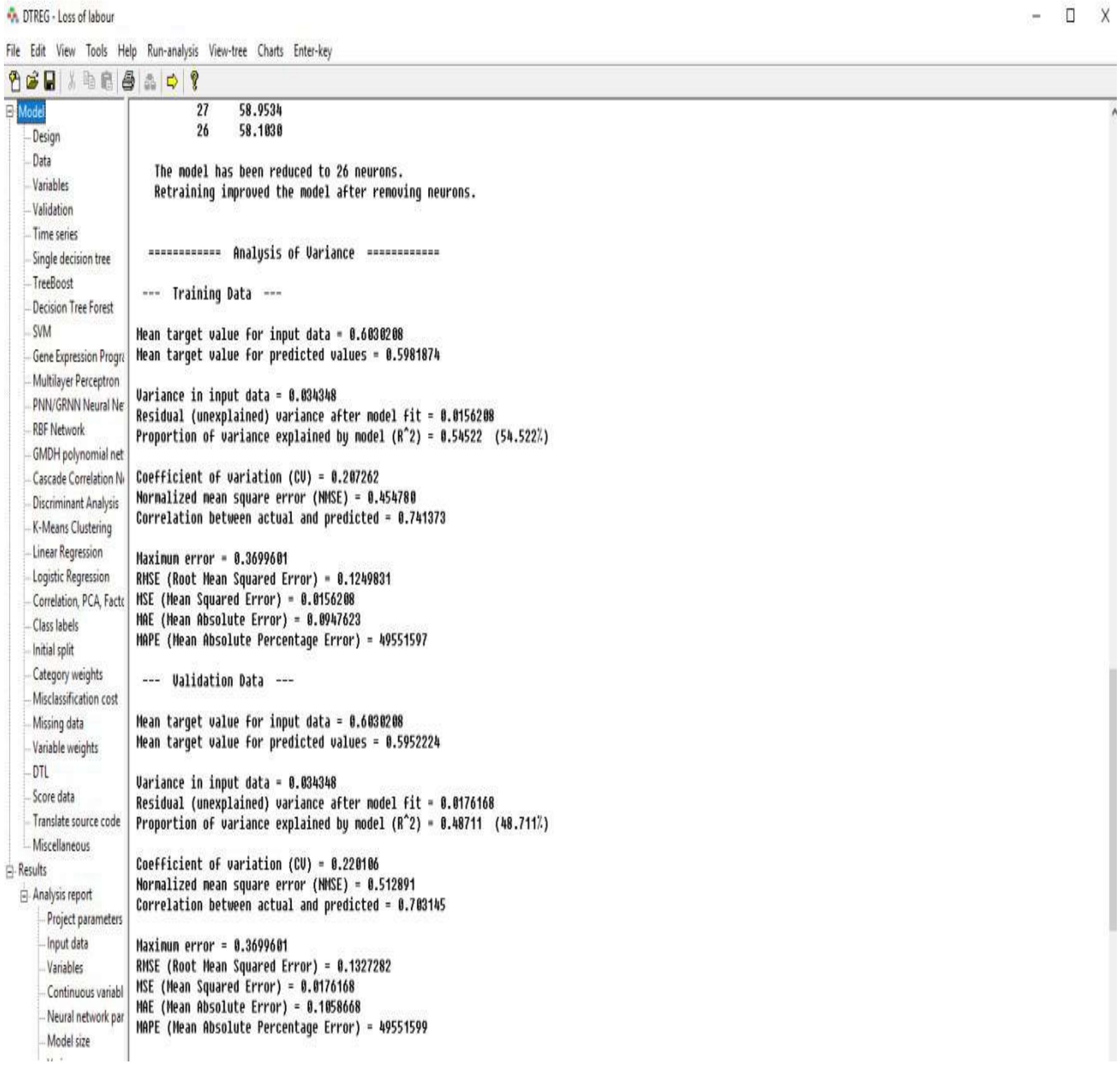


Figure 37. Generalized Regression Neural Network Regression Results

Figure 37 shows that the results of the generalized regression neural network were not satisfactory due to the low R percentage (correlation coefficient) in Training (54%) and Testing (48%). In addition, the MSE value was 0.01, which is not acceptable.

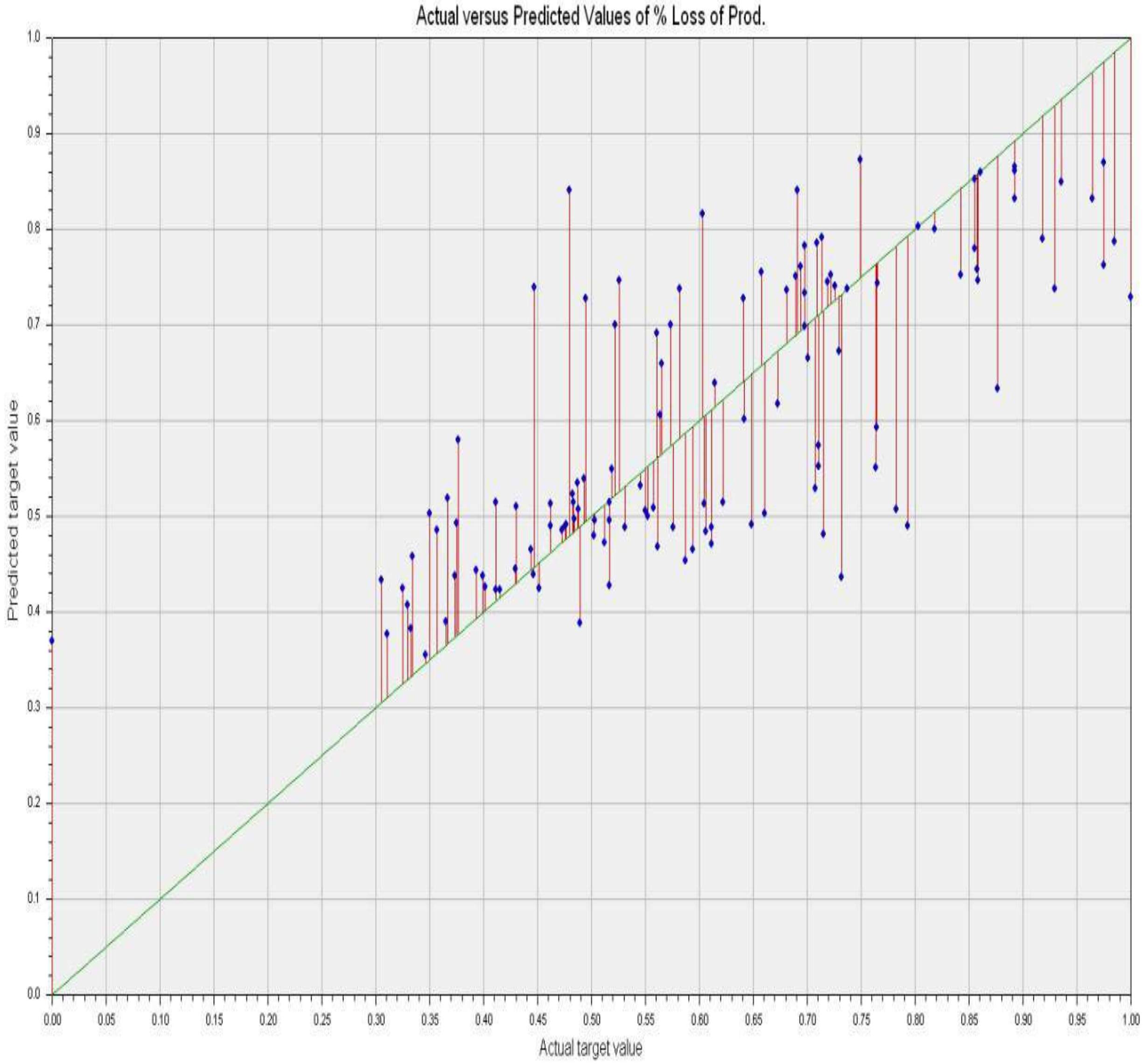


Figure 38. Actual Loss of Productivity VS Predicative Loss of Productivity.

5.3. Model Selection

This study selected the optimal model based on the Training and Testing R-squared values, root mean squared error (RMSE), mean squared error, mean absolute error (MAE), and analysis of variance parameters (p-value and f-value) as shown in equation 5.2-5.4 (Golnaraghi et al., 2019; Goyal, S., & Goyal, G. K. , 2011; Chai & Draxler, 2014).

MSE = $\sum_{1}^{N} \left(\frac{Q\,\exp - Q\,cal}{n}\right)^2$ **(Equation 5.1)**

Where Q exp = Observed Value,

Q cal = Predicted Value

N = Number of Observations.

RMSE = $\sqrt{\frac{1}{n}\sum_{1}^{N}(\frac{Q\exp - Q\,cal}{Q\,exp})^2}$ **(Equation 5.2)**

Where Q exp = Observed Value,

Q cal = Predicted Value

N = Number of Observations.

MAE = $\frac{1}{n}\sum_{i=1}^{n}|ei|$ **(Equation 5.3)**

Where ei = Model Errors,

N= Sampel of Model Errors.

This research selects the model in two steps:

First, the models were compared with the R-squared values of Training and Testing, mean squared error (MSE), root mean squared error (RMSE), and mean absolute error (MAE).

Second, a comparison was conducted with the analysis of variance (ANOVA) parameters (p-value and f-value).

Table 8. Comparison between the Results of the Three Techniques.

Techniques	**Train R-Squared**	**Test R-Squared**	**MSE**	**RMSE**	**MAE**
Feed Forward Neural Network (FNN)	% 80.88	% 72.30	0.009	0.093	0.066
Cascade Neural Network (CNN)	% 83.50	% 82.46	0.0065	0.080	0.061
Generalized Regression Neural Network (GRNN)	% 54.522	% 48.711	0.017	0.1327	0.105

According to Table 8, the comparison shows that the results of the feedforward neural network and cascade neural network are very similar, making it very challenging to select the most efficient model. However, the generalized regression neural network had a lower R-squared value than the other techniques and also had high MSE, RMSE and MAE values. The GRNN results were therefore not satisfactory.

In order to select the most efficient of the two remaining models, this study uses analysis of variance in Minitab.

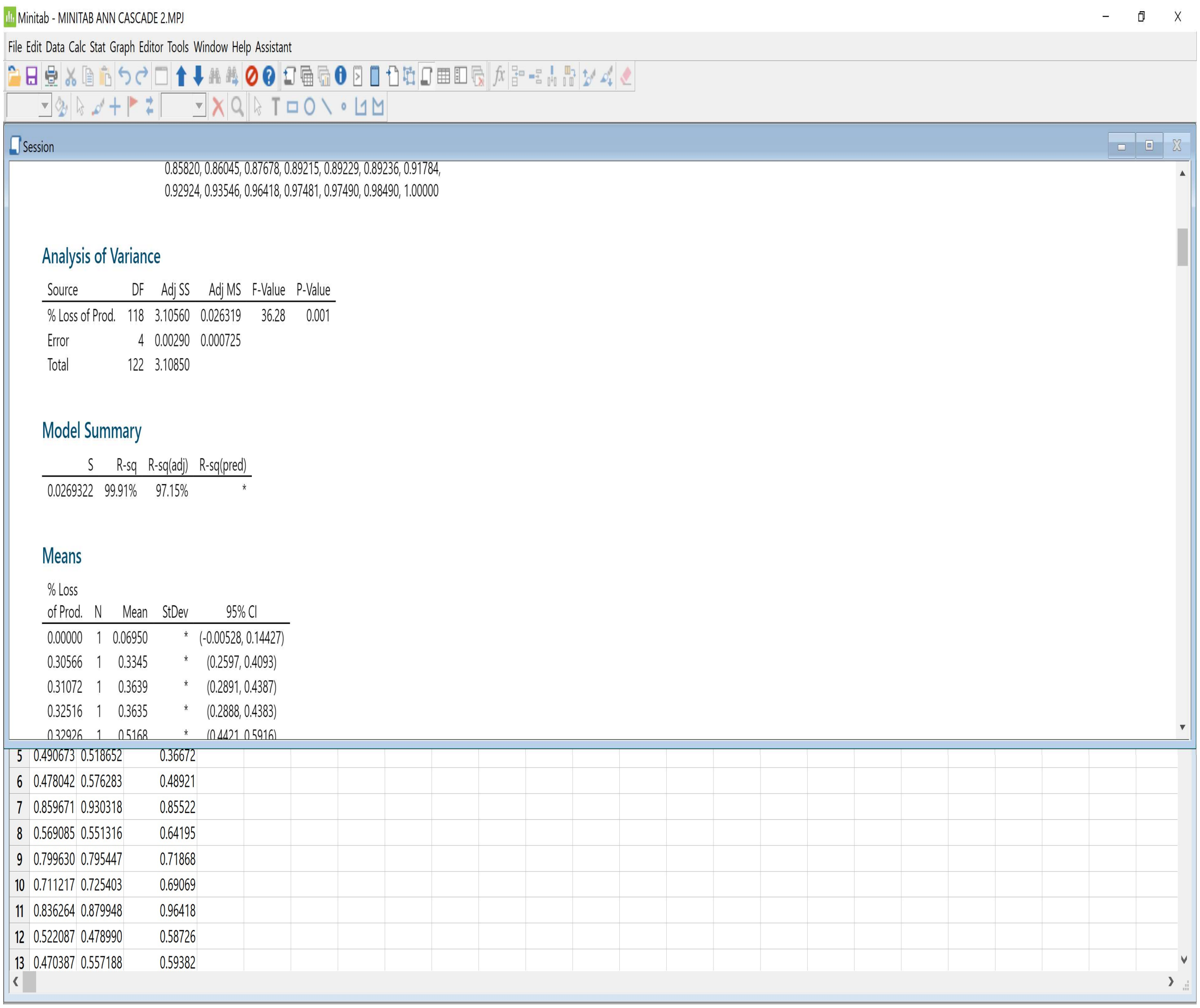


Figure 39. Analysis of Variance in Minitab Software.

Table 9. The Results of Analysis of Variance between CNN and FNN.

Techniques	P-Value	F-Value
Feed Forward Neural Network (FNN)	0.001	36.28
Cascade Neural Network (CNN)	0.0013	11.74

As mentioned previously, the p-value of each technique was less than alpha, in other words, there is a difference between the results of these techniques and actual loss of productivity. Moreover, a comparison of each technique's f-value shows that the f-value of the cascade neural network (CNN) is lower than that of the feedforward neural network, meaning that there is no significant difference between the cascade neural network output results and actual loss of productivity.

Consequently, the cascade neural network (CNN) was selected due to its high percentage of R-squared value, and low MSE, RMSE, MAE, p-value and f-value values.

5.4. Summary

This chapter describes the model development, which was divided into two steps. Step one concerns the model development using three artificial neural network techniques: the feedforward neural network (FNN), cascade neural network (CNN) and generalized regression neural network (GRNN). Afterward, the available data sets were implemented in Matlab and DTREG software to extract the training and testing R (correlation coefficient) values and to find the statistical parameters such as the mean squared error (MSE), root mean square error (RMSE) and mean absolute error (MAE).

In step two, the process of selecting the most efficient model was conducted in two parts. In the first part, the statistical parameters such as training and testing R-squared values, MSE, MAE, and RMSE were set. In the second part, the analysis of variance (ANOVA) was conducted. In this part, the techniques were evaluated using the p-value and f-value.

The cascade neural network was subsequently chosen as the optimal method based on these two steps due to its high R-squared value and low values of MSE, MAE, RMSE, and f-value.

Chapter 6. Model Validation

System validation is considered a complex and critical task. The primary objective of validation is to certify that a program fulfills its requirements and satisfies its end user. Although validation and verification are clearly different techniques, they have been used interchangeably (Jagdev, Browne, & Jordan, 1995).

Validation and Verification can be defined as follows:

Validation is defined as developing the correct system and verification is described as developing a system correctly (O'Keefe & O'Leary, 1993). The validation process usually occurs after verification and varies from one industry to another. Since the expert system plays a prominent role in system development, system validation is a crucial task. System validation is a process to ensure that it precisely signifies an expert's knowledge in a particular problem domain (O'Leary, Goul, Moffitt, & Radwan, 1990).

The developed system was validated using a real case study. To validate and test the effectiveness of the proposed model to quantify loss of productivity due to change orders, three case examples from Leonard's datasets (1988) were used to compare the developed model with other methods for calculating the loss of productivity. The cases were analyzed using the developed model and the regression models of Leonard (1988) and Ibbs (2005). In order to compare the results with Leonard and Ibss model, the case studies results were compared with absolute error, average error and actual loss of productivity. In addition, this study created the timing model to compare the cascade neural network model with Ibbs model.

6.1. Model Validation via Absolute Error and Average Error

In order to calculate the absolute error and average error, this study used the following these equations (Golub & Van Loan, 1996; Chai & Draxler, 2014) :

$$\varepsilon\, abs = \left|\frac{\Delta x}{x}\right| \qquad \textbf{(Equation 6.1)}$$

Where $\Delta x =$ Difference between predicated loss of productivity and actual loss of productivity,

x = Predicated Loss of productivity.

$\varepsilon\ avg = \frac{1}{n}\sum_{i=1}^{n}|ei|$ **(Equation 6.2)**

6.1.1. Case Study 1

The first case is an arena having a value of $5.7 million awarded to a mechanical contractor with a planned duration of 17 months. The type of project was lump sum and design changes was the major cause of change order in this project .The project was delayed for one month also project experienced several change orders for the value of $1.5 million. The planned hours were 34,400, the actual hours of the project were 38,000, and the total hours spent on change orders were 4,860. This project was affected by one major cause of change orders which is design changes.

In order to compare loss of productivity with Leonard (1988) and Ibbs (2005) Models, these models were compared based on the change order percent. Table 10 shows the comparison between the proposed model, Leonard model, Ibbs Model and Actual loss of productivity.

Table 10. Actual vs. Predicated Loss of Productivity for Case Study 1

Developed Model	**Leonard Model**	**Ibbs Model**	**Actual Loss of Productivity**
13.5%	15%	17%	9%

Also, Table 11 illustrates the comparison between absolute error of the developed model, Leonard's model, and Ibbs model.

Table 11. Calculated Absolute error for the three models.

Models	**Leonard's Model**	**Ibbs Model**	**Developed Model**
Absolute Error	40%	47%	33.33%

The results show that the developed model is more accurate and reliable in the first case study.

6.1.2. Case Study 2

The second case is a hotel having a total of $3,218,000 awarded to an electrical contractor. The original duration was 16 months; however, the project was extended for 18 months. The type of contract was lump sum and desihn changes was the major cause of change orders .The project experienced several change orders for the value of $700,000. The planned hours were 39,500, the actual hours were 88,300 and the total hours spent on change orders were 12,300. This project was affected by two major causes.

Table 12 reveals the comparison between the proposed model, Leonard's model, Ibbs model, and actual loss of productivity.

Table 12. Actual vs. Predicated Loss of Productivity in Case Study 2

Developed Model	**Leonard Model**	**Ibbs Model**	**Actual Loss of Productivity**
38.85 %	29 %	18.2 %	42%

In addition, Table 13 shows the comparison between absolute error of the developed model, Leonard's model, and Ibbs model.

Table 13. Calculated Absolute error for the three models in Case Study 2.

Models	**Leonard's Model**	**Ibbs Model**	**Developed Model**
Absolute Error	44%	56.66%	8.1 %

6.1.3. Case Study 3

The third case is a residence having a total of $3,675,000 awarded to a civil contractor. The original duration was 15 months; however, the project was extended for 20 months. The type of contract was lump sum and design changes and incomplete design were the the major causes of change orders. The project experienced several change orders for the value of $1,000,000. The planned hours were 13,166, the actual hours were 179,000 and the total hours spent on change orders were 25,000. This project was affected by two major causes.

Table 14 shows the comparison between the proposed model, Leonard's model, Ibbs model and actual loss of productivity.

Table 14. Actual vs. Predicated Loss of Productivity in Case Study 3

Developed Model	**Leonard Model**	**Ibbs Model**	**Actual Loss of Productivity**
27%	21%	20%	26%

Furthermore, Table 15 shows the comparison between absolute error of the developed model, Leonard's model, and Ibbs model.

Table 15. Calculated Absolute error for the three models in Case Study 3.

Models	**Leonard's Model**	**Ibbs Model**	**Developed Model**
Absolute Error	23%	30%	4%

According to the three case studies, Table 16 illustrates the Average Error of the three absolute error.

Table 16. Calculated Average Error for The Three Models.

Models	**Leonard's Model**	**Ibbs Model**	**Developed Model**
Average Error	35%	44.55%	16%

As it can be noticed, these results indicate that the proposed model can produce more accurate and reliable results when calculating loss of productivity due to change orders. The proposed model has the lowest average error and absolute error among other models.

6.2. Model Validation via Timing of Change

In order to compare the cascade neural network model with previous studies such as regression, this study creates the timing model to compare the results with the Ibbs model. Consequently, the results were compared with three real case studies to validate and test the effectiveness of the proposed model.

6.2.1. Early Change

This stage contains the change which occurred at the 0 to 25% project period. In this stage, the input of the neural network is change order percent. The neural network’s output is the productivity index.

In the training process, the Matlab software allocated 70% of the 12 samples for training, 15% for validation, and 15% for testing. The training samples were given to the network during training, and the network was adjusted according to its error. The validation samples were used to measure network generalization and to stop training when generalization stops enhancing.

Once the framework of the cascade neural network was established, the results after running were satisfactory due to the high R (correlation coefficient) value and low mean squared error (MSE) value.

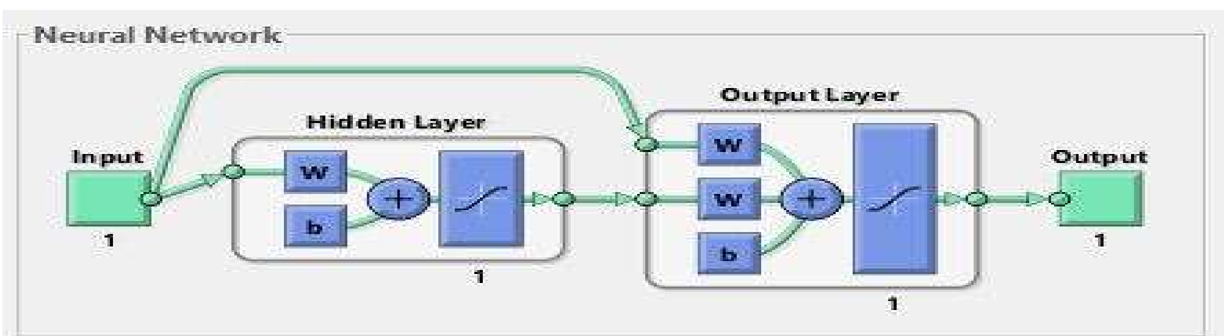


Figure 40. Cascade Neural Network Analysis Structure in Matlab Software

The cascade neural network’s structure contains one input parameters (change order percent) and one hidden layer composed of one hidden nodes. Also, the Levenberg-Marquardt training algorithm is used due to having the highest processing speed and efficiency despite requiring more memory than bayesian regularization and scaled conjugate gradient algorithm.

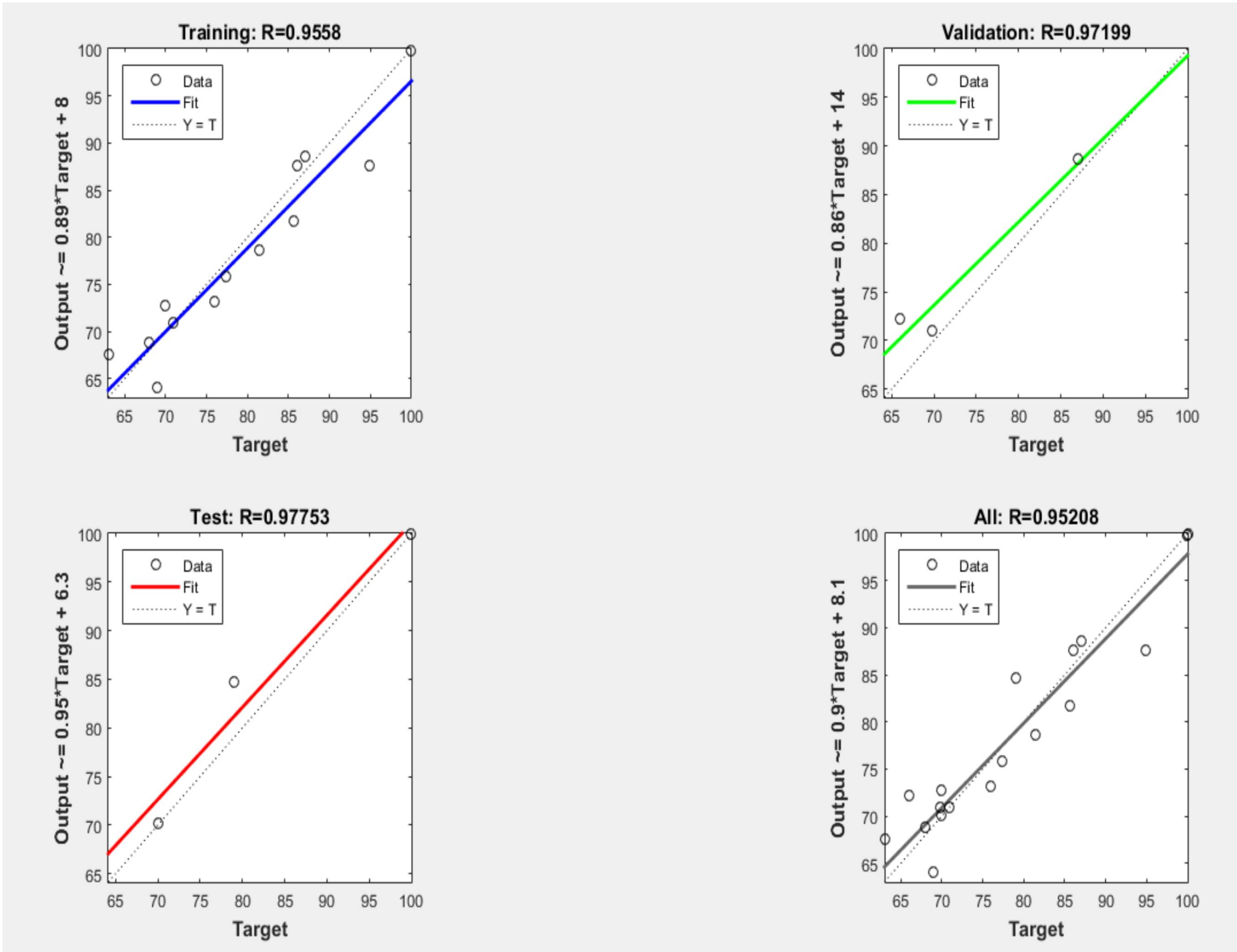


Figure 41. Cascade Neural Network Regression Results for Early Stage

Figure 41 illustrates that the results of the cascade neural network were satisfactory due to the high R-Squared (Coefficient of determination) percentage in Training (90%) and Testing (95%).

6.2.2. Normal Change

This stage includes of the change which occurred at the 26 to 75 % project period. In this stage, the input of the neural network is change order percent. The neural network's output is the productivity index.

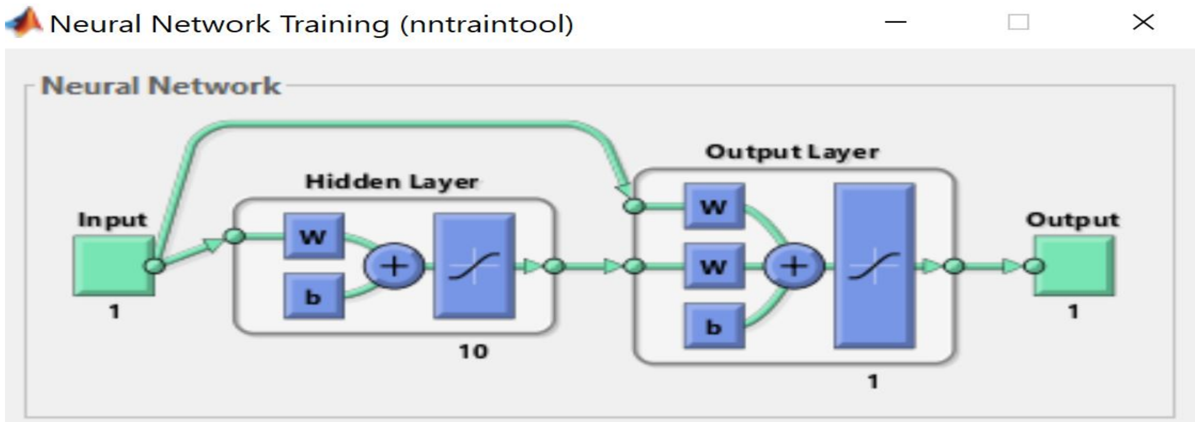


Figure 42. Cascade Neural Network Analysis Structure in Matlab Software.

The cascade neural network's structure contains one input parameter (change order percent) and one hidden layer composed of one hidden nodes. Furthermore, the Levenberg-Marquardt training algorithm is utilized. In the training process, the Matlab software assigned 70% of the 27 samples for training, 15% for validation and 15% for testing. The training samples were given to the network during training, and the network was adjusted according to error. The validation samples were used to measure network generalization and to stop training when generalization stops enhancing.

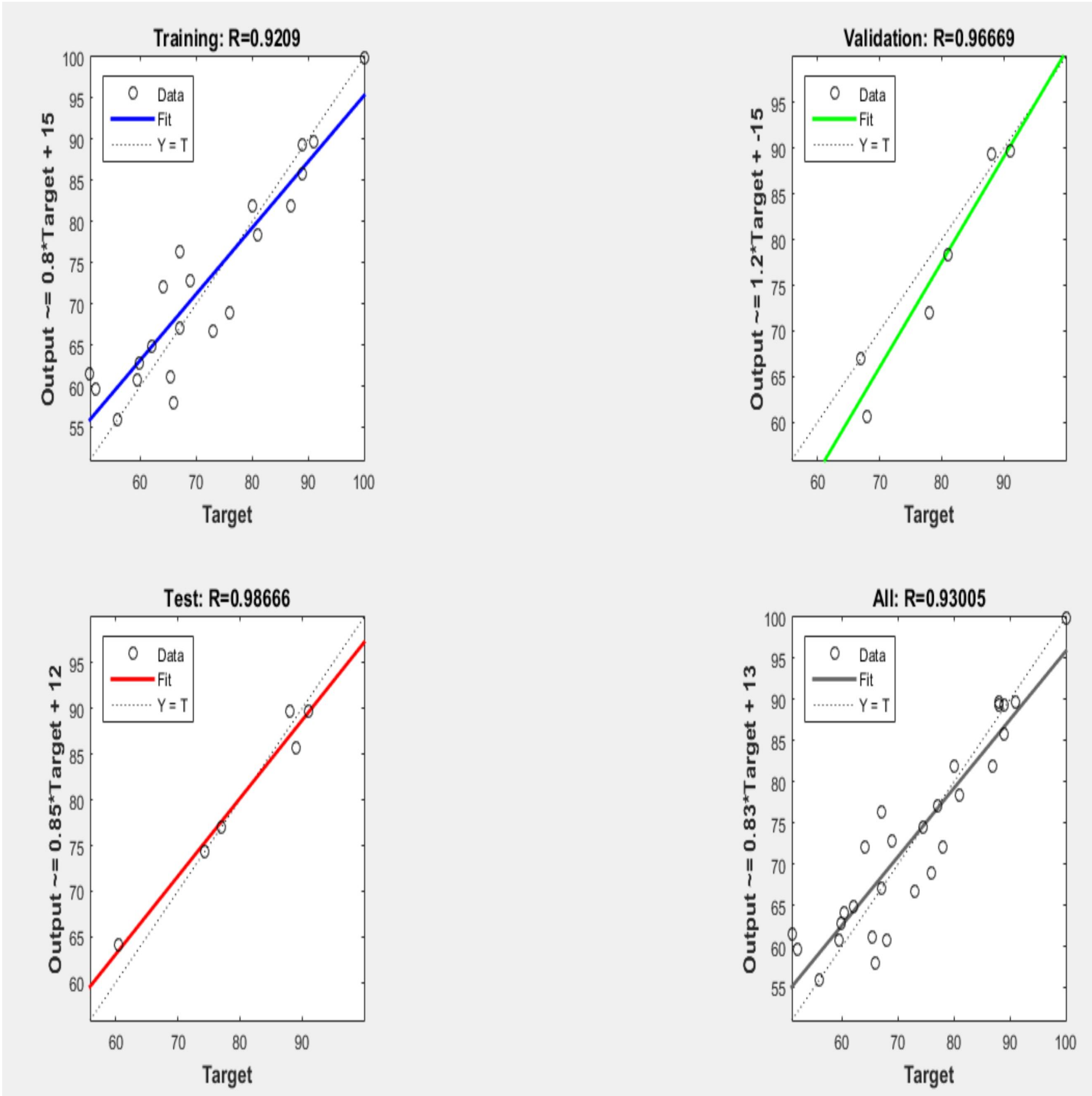


Figure 43. Cascade Neural Network Regression Results for Middle Stage.

Figure 43 shows that the results of the cascade neural network were satisfactory due to the high R-Squared (Coefficient of determination) percentage in Training (85%) and Testing (97%).

6.2.3. Late Change

This stage contains of the change which occurs at 76 to 100 % project period. In this stage, the input of the neural network is change order percent. The neural network's output the productivity index. In the training process, the Matlab software allocated 70% of the 34 samples for training, 15% for validation and 15% for testing. The training samples were given to the network during training, and the network was adjusted according to its error. The validation samples were used to measure network generalization and to stop training when generalization stops enhancing.

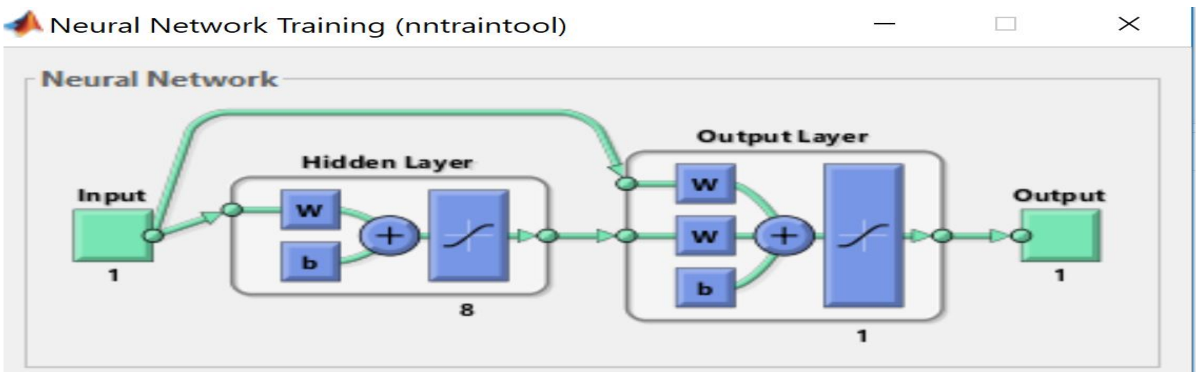


Figure 44. Cascade Neural Network Analysis Structure in Matlab Software.

The cascade neural network's structure contains one input parameters (change order percent) and one hidden layer composed of one hidden nodes. Furthermore, the Levenberg-Marquardt training algorithm is utilized than bayesian regularization and the scaled conjugate gradient algorithm.

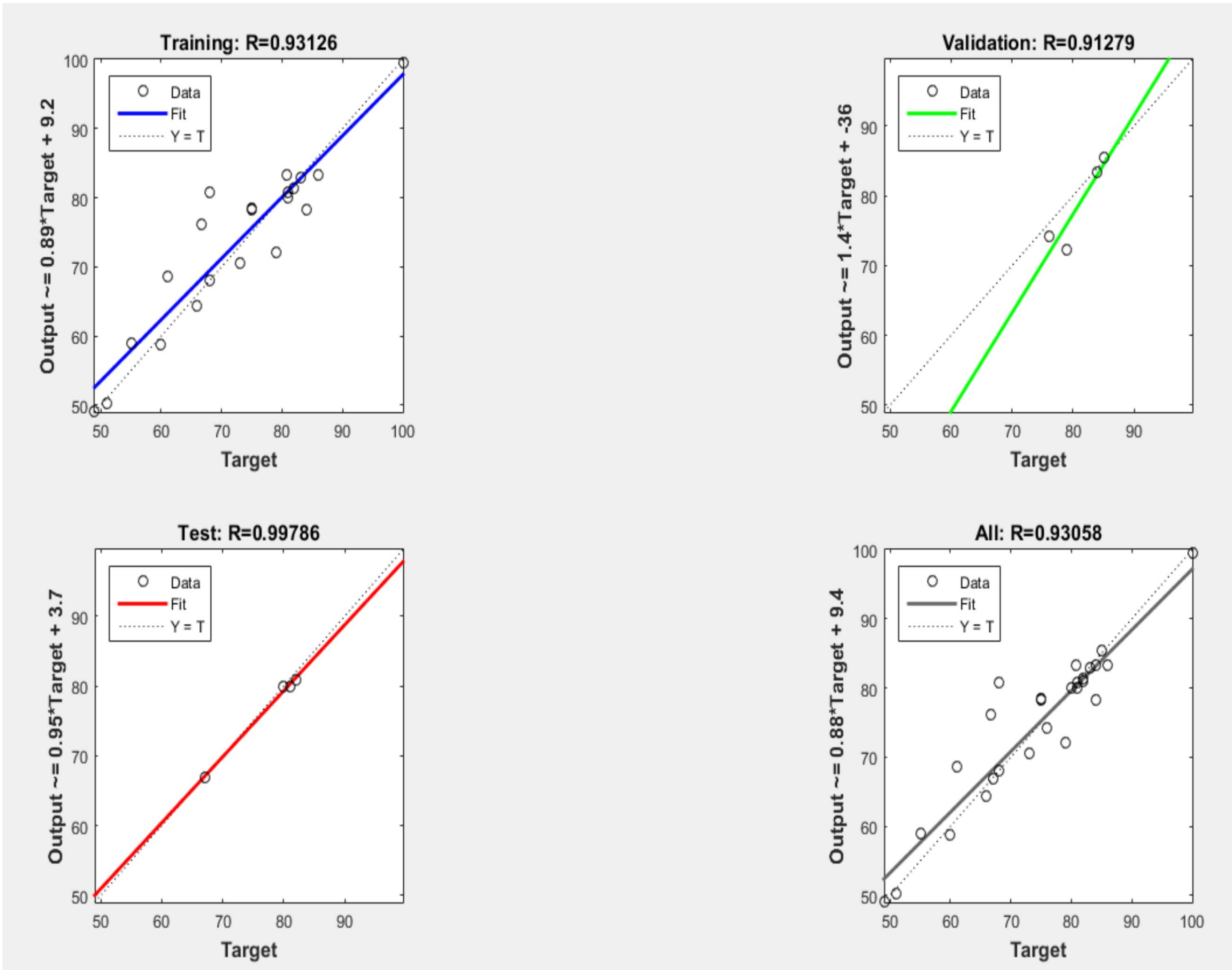


Figure 45. Cascade Neural Network Regression Results for Late Stage.

Figure 45 shows that the results of the cascade neural network were satisfactory due to the high R-Squared (Coefficient of determination) percentage in Training (87%) and Testing (98%).

6.2.4. Current Timing Model

Based on the three stages of the project, the current timing model was built in Matlab software to compare the current model with Ibbs model (2005).

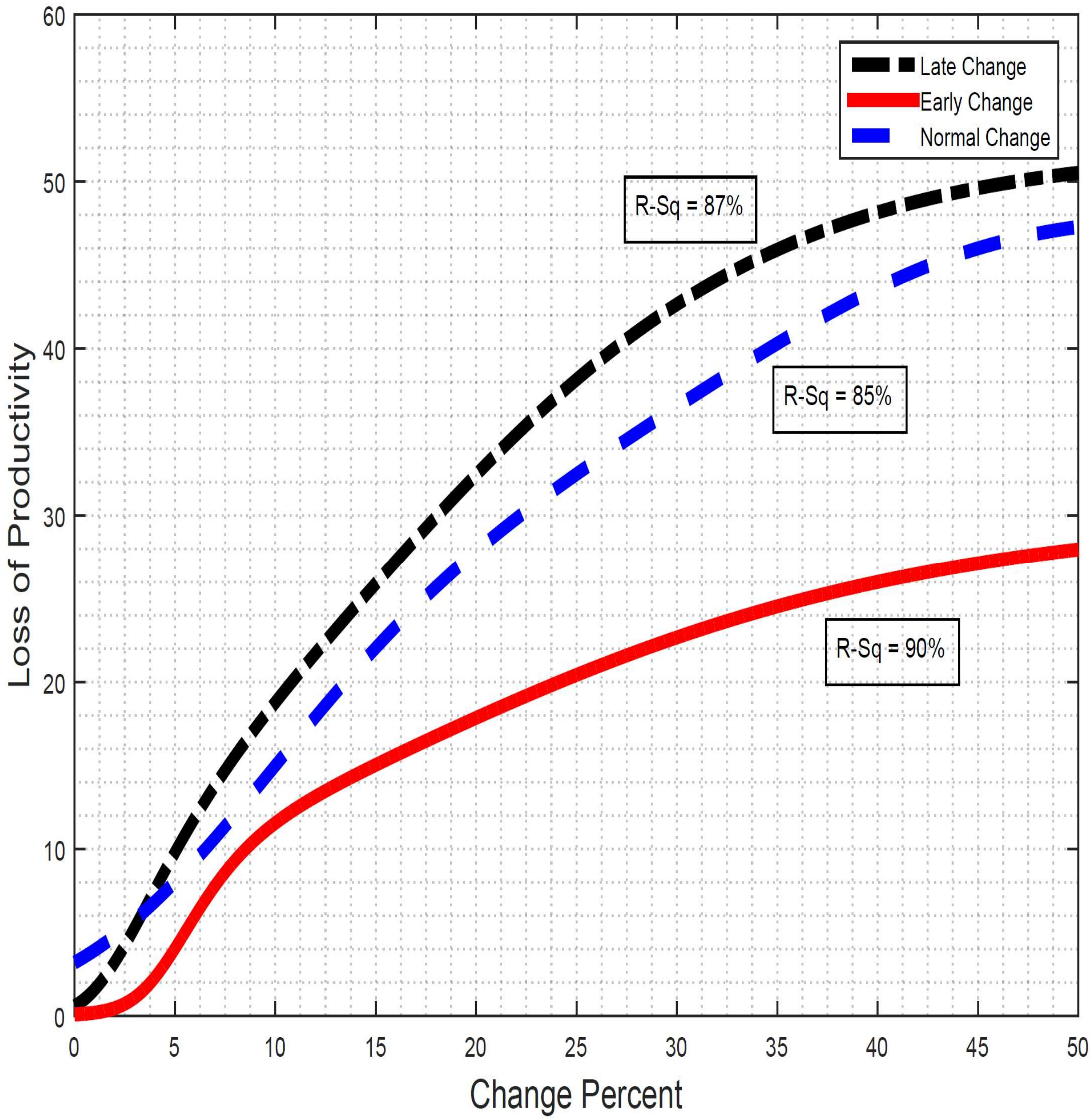


Figure 46. Loss of Productivity versus Timing of Change.

6.2.5. Comparing the Current Timing Model with Ibbs Model

In order to compare and validate the current timing model with Ibbs study, three real case studies were used based on the loss of productivity and change order percent.

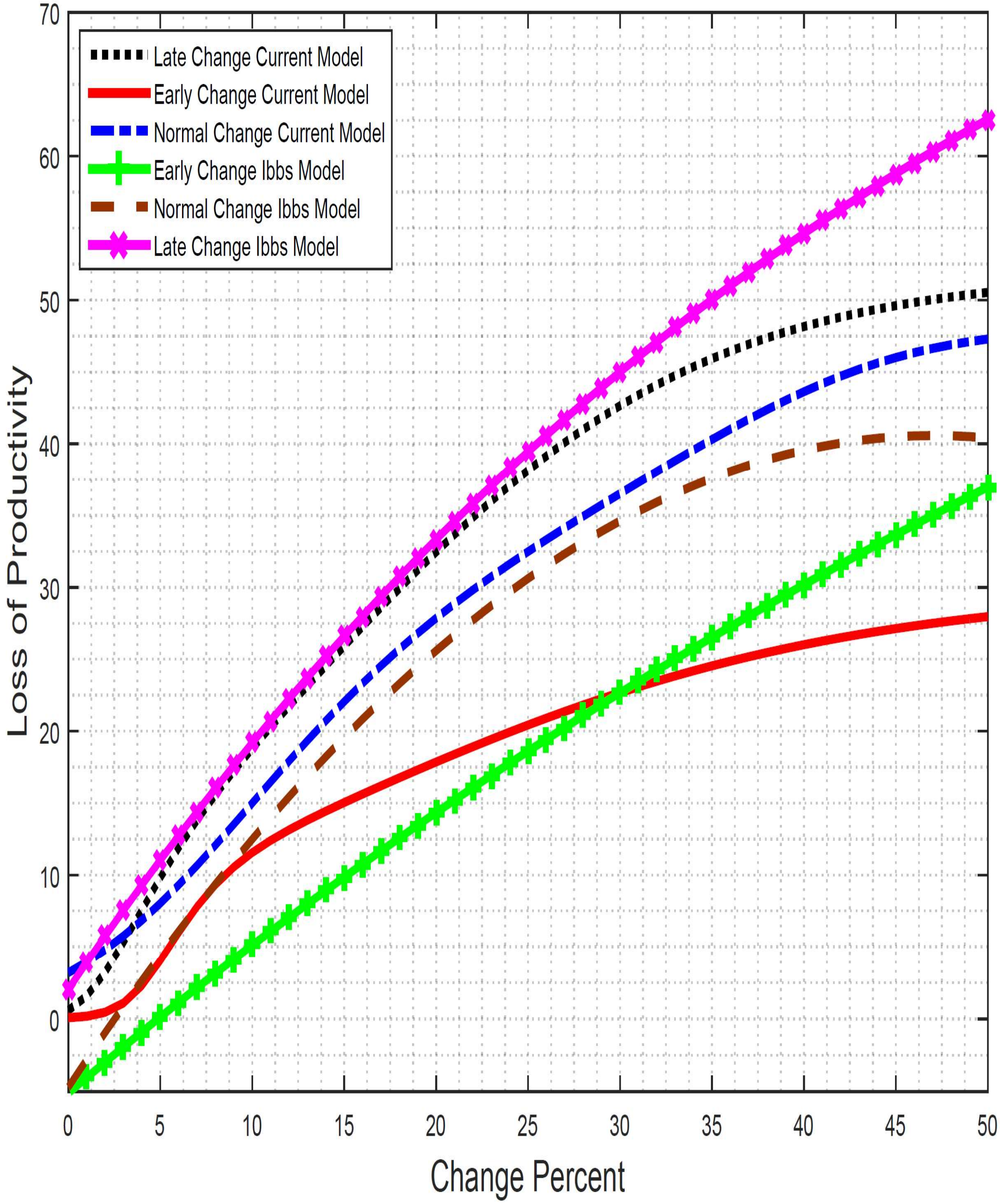


Figure 47. Current Model versus Ibbs Model (2005)

Table 17 illustrates the comparison between Ibbs model (2005) and cascade neural network model. Three real case studies were used to demonstrate the proposed model accuracy.

Case Study 1:

The first case is a mechanical project having a value of $1.948 million awarded to a contractor with a planned duration of 7 months. The type of project contract was fixed price, and incomplete design was the major causes of change order in this project. The project was delayed for two months also project experienced several change orders for the value of $7,074. The planned hours were 11,178, the actual hours of the project were 12,934, and the total hours spent on change orders were 998. This project was affected by one major cause of change orders.

Case Study 2:

The second case is an electrical project having a value of $1.29 million awarded to a contractor with a planned duration of 10 months. The type of project contract was lump sum, and incomplete design was the major causes of change order in this project. The project was delayed for 4 months also project experienced several change orders for the value of $100,000. The planned hours were 17,640, the actual hours of the project were 26,000, and the total hours spent on change orders were 2,287. The project was affected by plus one major cause of change orders.

Case Study 3:

The third case is an electrical project having a value of $1.2 million awarded to a contractor with a planned duration of 2 months. The type of project contract was unit price and design changes was the major cause of change order in this project. The project was delayed for two months also project experienced several change orders for the value of $509,290. The planned hours were 12,249, the actual hours of the project were 13,886, and the total hours spent on change orders were 989. This project was affected by one major cause of change orders.

Case Study 4:

The fourth case is an electrical project having a value of $6.615 million awarded to a contractor with a planned duration of 17 months. The type of project contract was lump sum and design changes was the major causes of change order in this project. The project was delayed for ten months also project experienced several change orders for the value of $4.5 million. The planned hours were 35,260, the actual hours of the project were 53,700, and the total hours spent on change orders were 23,850. This project was affected by one major cause of change orders.

Case Study 5:

The fifth case is mechanical project having a value of $476,000 awarded to a contractor with a planned duration of 2 months. The type of project contract was lump sum and incomplete design was the major causes of change order in this project. The project was delayed for 3 months also project experienced several change orders for the value of $150,000. The planned hours were 9,000, the actual hours of the project were 10,675, and the total hours spent on change orders were 4,000. This project was affected by one major cause of change orders.

Table 17. Actual Loss of Productivity versus Estimated productivity for the Two Models.

Case Study No.	**Type**	**Change Order Percent**	**Current Model**	**Ibbs Model**	**Actual Loss of Productivity**
1	Early	8%	10%	6%	13%
2	Early	9%	11%	7%	14%
3	Normal	7%	12%	7%	11%
4	Late	44%	51%	58%	49%
5	Late	37%	45%	52%	31%

As a consequence, these results indicate that the proposed model can produce more accurate and reliable results when calculating the productivity

6.3. Summary

In this chapter, validation process consisting of two parts.In the first part, three real case studies were used to validate the proposed model with two previous studies: Leonard (1988) and Ibbs (2005) models. The developed model was compared with absolute error and average error to ensure the model reliability and accuracy. The results show that the proposed model has the lowest average error and absolute error which can be considered as a precise model to estimate loss of productivity due to change orders. Also, in the second part the timing model is created to estimate the productivity with change order percent. Three real case studies were utilized to validate the proposed timing model with Ibbs model (2005). The results show that the proposed model is more accurate for predicting loss of productivity.

Chapter 7. Conclusion

7.1. Introduction

This chapter outlines the research findings and contributions and discusses areas for future research and development.

7.2. Research Overview

This study was conducted to quantify the impact of change orders on labour productivity using an artificial neural network model. This research consisted five stages as follows:

1. Literature Review, 2. Data Collection, 3. Model Development and Model Selection, and 4. Model Validation.

To review previous studies and clarify recent methods, this study considered literature related to change orders and loss of labour productivity. The literature review consisted of eight parts which simplified recent attempts to quantify loss of labour productivity caused by change orders. This research reviewed the previous methods, which were as follows: total cost method, modified total cost method, industry standard, measured mile approach, baseline productivity, statistical methods and neural networks.

In the data collection stage, 123 data sets were gathered by combining the data from two previous studies conducted by Assem (2000) and Leonard (1988). Also, in the data collection stage, the collected data sets were not normally distributed. For example, the range of data sets for one of the variables was from $80,000 to $23,172,000, while the range of another variable was from 0.76 to 2.78. In order to redistribute the available datasets to a normal distribution, the min-max normalization technique was used.

In addition, the variables of this study were distinguished to determine which were significant or insignificant. The variables were implemented using the best subset regression method in Minitab. The variables were selected based on the adjusted R-sq, R-sq and Mallows's Cp values. The following nine variables were selected: type of impact, type of work, type of contract, original estimated hours, experience factor, actual hours, change order frequency, change order percent and earned hours. Furthermore, in the data collection stage the 123 datasets were categorized into three timing datasets to compare the model accuracy in model validation stage.

In the model development stage, three artificial neural network techniques were implemented. These techniques were the feedforward neural network (FNN), cascade neural network (CNN) and generalized regression neural network (GRNN). This study selected the optimal model based on the R-squared values of Training and Testing, root mean square error (RMSE), mean squared error, mean absolute error (MAE) and analysis of variance parameters (p-value and f-value). Finally, stage four validated and tested the effectiveness of the proposed model to quantify loss of productivity caused by change orders. Three case examples from Leonard's data set were used to compare the developed model with other methods for calculating the loss of productivity. The cases were analyzed using the developed model and the regression models of Leonard (1988) and Ibbs (2005).

In order to compare the results with the models of Leonard and Ibbs, the case study results were compared with the absolute error, average error and actual loss of productivity. In addition, a timing model was created to estimate the productivity with change order percent. Three real case studies were utilized to validate the proposed timing model with Ibbs model (2005). The results shows that the proposed model is more accurate for predicting productivity.

7.3. Research Conclusion

- The results explained that the best subset regression method is the convenient technique for distinguishing the insignificant variables from the significant variables due to statistical parameters such as adjusted R-squared, R-squared and Mallows's Cp;
- This study quantified loss of productivity caused by change orders using an artificial neural network technique. The results of this research illustrated that the cascade neural network model yields more satisfactory results in comparison to the feedforward neural network (FNN) and generalized regression neural network (GRNN) due to its high R-squared percentage and low value of MSE, RMSE, MAE, P-value, and F-value;
- According to three real case studies, the developed model can more accurately predict loss of productivity than those proposed by previous studies; and
- The developed timing model is more precise in comparison to the existing studies such as Ibbs (2005) model, based on three real case studies.

7.4. Research Contributions

- **Enhancing the loss of productivity prediction via cascade neural network**

This study enhances loss of productivity prediction by implementing a new neural network technique. Previous studies have utilized a backpropagation neural network to estimate loss of productivity. However, this research implemented the cascade neural network technique (CNN) after eliminating other neural network techniques using R-sq, RMSE, MSE, MAE, p-value and f-value. The results demonstrate that there is no difference between predicted loss of productivity and actual loss of productivity in the proposed model. This model also has a lower average error and absolute error in comparison with those of previous studies, which shows that this model is more precise and accurate in predicting loss of productivity.

- **Creating the precise timing model for predicting the productivity**

  One of the major contributions of this research is the use of Cascade Neural Network to develop a timing model to predict productivity. This model divides the change time into three separate stages which are early (0-25% project period), normal (26-75% project period) and late (76-100% project period). The results of this model shows that the value of the estimated productivity is much more precise than the Ibbs timing model (2005).

7.5. Recommendations for Future Work

The model was developed to accomplish the research objectives set in this study. The developed model was implemented and validated through case studies and the results show excellent accuracy. However, the model can be expanded further. Suggested methods for enhancing and advancing the model include:

1. Investigating and implementing the other variables that affect loss of productivity due to the change orders;

2. Considering a large and comprehensive data set from construction sites to develop a more effective model for quantifying loss of productivity, which also ensures contractors and owners can reduce their additional expenses; and

3. Implementing cutting-edge neural network techniques such as convolutional neural networks or wavelet neural networks to improve loss of productivity prediction.

8. References

AACE. (2004). Estimating Lost Labor Productivity In Construction Claims. AACE Recomended Practice No. 25R-03.

Ai, X., & Zsaki, A. M. (2017). Stability assessment of homogeneous slopes loaded with mobile tracked cranes—An artificial neural network approach. Cogent Engineering, 4(1), 1360236.

Ai, Xin. (2016). Stability Assessment of Homogeneous Slopes Loaded with Mobile Tracked Cranes – An Artificial Neural Network Approach," Master of Applied Science Thesis, Department of Building, Civil and Environmental Engineering, Concordia University, Montreal, Canada.

Al-Mahasneh, A. J., Anavatti, S. G., & Garratt, M. A. (2018). Review of Applications of Generalized Regression Neural Networks in Identification and Control of Dynamic Systems. arXiv preprint arXiv:1805.11236.

American Association of Cost Engineers (AACE). (2013). AACE international recommended practice 10S-90, cost engineering terminology, AACE International, Morgantown, WV.

André, C. D., Narula, S. C., Elian, S. N., & Tavares, R. A. (2003). An overview of the variables selection methods for the minimum sum of absolute errors regression. Statistics in medicine, 22(13), 2101-2111.

Badde, D. S., Gupta, A., & Patki, V. K. (2009). Cascade and Feed Forward Back propagation Artificial Neural Network Models for Prediction of Compressive Strength of Ready Mix Concrete. IOSR Journal of Mechanical and Civil Engineering (IOSR-JMCE).

Brams, R. S., & Lerner, C. (1996). Construction claims deskbook: Management, documentation, and presentation of claims (Construction Law Library). New York: Wiley.

BusinessDictionary. (2018). Scope Change." Retrieved on Dec 5, 2018, from http://www.businessdictionary.com/definition/scope-change.html.

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)?. Geoscientific Model Development Discussions, 7, 1525-1534.

Cigizoglu, H. K. (2005). Generalized regression neural network in monthly flow forecasting. Civil Engineering and Environmental Systems, 22(2), 71-81.

Del Rosario Martinez-Blanco, M., Castañeda-Miranda, V. H., Ornelas-Vargas, G., Guerrero-Osuna, H. A., Solis-Sanchez, L. O., Castañeda-Miranda, R., ... & Martínez-Fierro, M. (2016). Generalized regression neural networks with application in neutron spectrometry. In Artificial Neural Networks-Models and Applications. InTech.

DTREG (2018)." About The Software.' Retrieved on Nov 20, 2018, from https://www.dtreg.com/

Eden, Colin, Williams, Terry, and Franck A. (2003).Issues with the Measured Mile. European Academy of Management (EURAM).

Eldin, N. N. (1989). Measurement of work progress: Quantitative technique. Journal of

Construction Engineering and Management, 115(3), 462-474.

Fahlman, S. E., & Lebiere, C. (1990). The cascade-correlation learning architecture. In Advances in neural information processing systems (pp. 524-532).

Fahlman, S.E.(1988).An Empirical Study of Learning Speed in Back-Propagation Networks. Technical Report.CMU-CSTR-88-162.Carnegie Mellon University.

Farbarik, J. J., III, G. O., Hanna, A. S., Moselhi, O., & Hassanein, A. (2004). Discussion of "Quantitative Definition of Projects Impacted by Change Orders" by Awad S. Hanna, Richard Camlic, Pehr A. Peterson, and Erik V. Nordheim. Journal of Construction Engineering & Management. https://doi.org/10.1061/(ASCE)0733-9364(2004)130:4(612)

Ferreira, J. C., & Patino, C. M. (2015). What does the p value really mean?. Jornal Brasileiro de Pneumologia, 41(5), 485.

Gaikwad, D.P & Thool, Ravindra R. (2014). Intrusion Detection System using Cascade Forward Neural Network with Genetic Algorithm Based Feature Selection. International Journal of Advanced Computer Engineering and Communication Technology. Vol.4 (1).

Gershenson, C. (2003). Artificial Neural Networks for Beginners. ArXiv. 1-8. https://doi.org/10.1093/icesjms/fsp009

Gilmour, S. G. (1996). The interpretation of Mallows's C_p-statistic. The Statistician, 49-56.

Goh, A. T. C., & Kulhawy, F. H. (2005). Reliability assessment of serviceability performance of braced retaining walls using a neural network approach. International Journal for Numerical and Analytical Methods in Geomechanics, 29(6), 627-642.

Golnaraghi, S., Moselhi, O. E. S., Eng, P., Alkass, S., & Golkhoo, F. (2018). Productivity Loss Quantification Using a Novel Artificial Intelegince Approach. AACE International.

Golnaraghi, S., Zangenehmadar, Z., Moselhi, O., & Alkass, S. (2019). Application of Artificial Neural Network (s) in Predicting Formwork Labour Productivity. Advances in Civil Engineering.

Golub, G. H., & Van Loan, C. F. (1996). Matrix computations. Johns Hopkins University, Press, Baltimore, MD, USA. https://doi.org/10.1017/CBO9781107415324.004

Gorse, C., Johnston, D., & Pritchard, M. (2012). A dictionary of Construction, Surveying, and Civil Engineering. Oxford University Press.

Goyal, S., & Goyal, G. K. (2011). Cascade and feedforward backpropagation artificial neural networks models for prediction of sensory quality of instant coffee flavoured sterilized drink. Canadian Journal on Artificial Intelligence, Machine Learning and Pattern Recognition, 2(6), 78-82.

Gulezian, R., & Samelian, F. (2003). Baseline determination in construction labor productivity-loss claims. Journal of management in engineering, 19(4), 160-165.

Hanna, A. S., & Iskandar, K. A. (2017). Quantifying and Modeling the Cumulative Impact of Change Orders. Journal of Construction Engineering and Management, 143(10), 04017076.

Hanna, A. S., & Sullivan, K. T. (2004). Factors affecting labor productivity for electrical contractors. The Electrical Contracting Foundation for Electrical Contractors, 31.

Hanna, A. S., Russell, J. S., & Vandenberg, P. J. (1999). The impact of change orders on mechanical construction labour efficiency. Construction Management & Economics, 17(6), 721-730.

Hanna, A., Camlic, R., Peterson, P. A., & Nordheim, E. V. (2002). Quantitative definition of projects impacted by change orders. Journal of Construction Engineering and Management, 36(10), 1109-1118.

Harel, O. (2009). The estimation of R^2 and adjusted R^2 in incomplete data sets using multiple imputation. Journal of Applied Statistics. https://doi.org/10.1080/02664760802553000

Hasegawa, K. (1995). Pending Change/Modification Codes," Documents of the Pacific Division of the Naval Facilities Engineering Command, Pearl Harbor Construction Office, Arizona.

Haykin, S. (1999). Support vector machines. Neural Networks: A Comprehensive Foundation, 318-350.

Helmy, M. (2002). Modeling pile group efficiency in cohesionless soil using artificial neural networks (Doctoral dissertation, Concordia University).

Hester, W. T., Chang, T. C., & Kuprenas, J. A. (1991). Construction changes and change orders: their magnitude and impact. Construction Industry Institute.

Ibbs, C. W., & Allen, W. E. (1995). Quantitative impacts of project change, Source document 118. Construction Industry Institute, Univ. of Texas, Austin, TX.

Ibbs, W. (2005). Impact of change's timing on labor productivity. Journal of Construction Engineering and Management, 131(11), 1219-1223.

Ibbs, W. (2012). Construction change: Likelihood, severity, and impact on productivity. Journal of Legal Affairs and Dispute Resolution in Engineering and Construction, 4(3), 67-73.

Ibbs, W., & Liu, M. (2005). Improved measured mile analysis technique. Journal of construction engineering and management, 131(12), 1249-1256.

Ibbs, W., & McEniry, G. (2008). Evaluating the cumulative impact of changes on labor productivity. Cost Engineering, 50(12), 23-29.

Jagdev, H. S., Browne, J., & Jordan, P. (1995). Verification and validation issues in manufacturing models. Computers in industry, 25(3), 331-353.

Jain, Y. K., & Bhandare, S. K. (2011). Min max normalization based data perturbation method for privacy protection. International Journal of Computer & Communication Technology, 2(8), 45-50.

Jason, W., Osborne, J. W. (2002). Normalizing Data Transformations. ERIC Digest.1-8.

Jones, R. M. (2001). Lost productivity: Claims for the cumulative impact of multiple change orders. Pub. Cont. LJ, 31, 1.

Kim, H. Y. (2014). Analysis of variance (ANOVA) comparing means of more than two groups.

Restorative dentistry & endodontics, 39(1), 74-77.

Lee, M. J., Hanna, A. S., & Loh, W. Y. (2004). Decision tree approach to classify and quantify cumulative impact of change orders on productivity. Journal of Computing in Civil Engineering, 18(2), 132-144.

Lee, S. (2007). Understanding and quantifying the impact of changes on construction labor productivity: Integration of productivity factors and quantification methods. University of California, Berkeley.

Leonard, C. A. (1988). The effects of change orders on productivity (Doctoral dissertation, Concordia University), Montreal, Quebec, Canada.

Loulakis, M. C., & Santiago, S. J. (1999). Getting the most out of your'measured mile'approach. Civil Engineering, 69(11), 69.

Madigan, D., & Ridgeway, G. (2004). Discussion of "Least Angle Regression" By Efron Et Al. The Annals of Statistics. https://doi.org/10.1214/009053604000000067

MathWorks. (2018). Cascade-forward neural network. Retrieved on Nov 30, 2018, from https://www.mathworks.com/help/deeplearning/ref/cascadeforwardnet.html.

Moayeri, V., Moselhi, O., & Zhu, Z. (2017). BIM-based model for quantifying the design change time ripple effect. Canadian Journal of Civil Engineering, 44(8), 626-642.

Moayeri, V. (2017). Design Change Management in Construction Projects Using Building Information Modeling (BIM) (Doctoral dissertation, Concordia University Montreal, Quebec, Canada).

Moselhi, O. (2003). Estimating the cost of change orders. Cost Engineering, 45(8), 24-29.

Moselhi, O., Assem, I., & El-Rayes, K. (2005). Change orders impact on labor productivity. Journal of Construction Engineering and Management, 131(3), 354-359.

Moselhi, O., Leonard, C., & Fazio, P. (1991). Impact of change orders on construction productivity. Canadian journal of civil engineering, 18(3), 484-492.

National Electrical Contractors Association. (2000). Guide to Electrical Contractor's Claims Management Vol. I. Bethesda, Maryland.

Nechyba, M. C., & Xu, Y. (1997). Cascade neural networks with node-decoupled extended Kalman filtering. In Proceedings 1997 IEEE International Symposium on Computational Intelligence in Robotics and Automation CIRA'97.'Towards New Computational Principles for Robotics and Automation' (pp. 214-219).

O'Keefe, R. M., & O'Leary, D. E. (1993). Expert system verification and validation: a survey and tutorial. Artificial Intelligence Review, 7(1), 3-42.

O'Leary, T. J., Goul, M., Moffitt, K. E., & Radwan, A. E. (1990). Validating Expert Systems. IEEE Expert-Intelligent Systems and Their Applications. (3), 51-58.

O'Brien, J. J. (1998). Construction Change Orders: Impact, Avoidance, and Documentation, McGraw-Hill Professional Publishing.

Osama, K., Somvanshi, P., Pandey, A. K., & Mishra, B. N. (2013). Modelling of nutrient mist reactor for hairy root growth using artificial neural network. European Journal of Scientific Research, 97(4), 516-526.

Patro, S., & Sahu, K. K. (2015). Normalization: A preprocessing stage. arXiv preprint arXiv:1503.06462.

Pinnell, S. S. (1998). How to get paid for construction changes: Preparation resolution tools and techniques. McGraw Hill.

Presnell, T. W. (2003). " Measured mile" process. Cost Engineering, 45(11), 14.

Revay, S. O. (2003). Coping with changes. AACE International Transactions, CD251.

Ruengvirayudh, P., & Brooks, G. P. (2016). Comparing Stepwise Regression Models to the Best-Subsets Models, or, the Art of Stepwise. General Linear Model Journal. Vol42(1).

Salchenberger, L. M., Cinar, E. M., & Lash, N. A. (1992). Neural networks: A new tool for predicting thrift failures. Decision Sciences, 23(4), 899-916.

Sawyer, S. F. (2009). Analysis of variance: the fundamental concepts. Journal of Manual & Manipulative Therapy, 17(2), 27E-38E.

Sazli, M. H. (2006). A brief review of feed-forward neural networks. Communications, Faculty of Science, University of Ankara, 50(1), 11-17.

Schetinin, V. (2005). An evolving cascade neural network technique for cleaning sleep electroencephalograms. arXiv preprint cs/0504067.

Schwartzkoph, W. (1995). Calculating Lost Labor Productivity in Construction Claims (Construction Law Library). John Wiley and Sons, Inc., NY.

Schwartzkoph, W., & McNamara, J. J. (2000). Calculating construction damages. Aspen Publishers Online.

Serag, E. (2006). Change orders and productivity loss quantification using verifiable site data. Doctor of Philosophy Thesis, Department of Civil and Environmental Engineering, University of Central Florida, Orlando, Florida.

Shwartzkoph, W., J. McNamara, and J. Hoffar. (1992).Calculating Construction Damages. John Wiley &Sons, USA.

Specht, D. F. (1991). A General Regression Neural Network. IEEE Transactions on Neural Networks. https://doi.org/10.1109/72.97934

Statista .(2016). “U.S. Construction Industry - Statistics & Facts.” Retrieved on Oct 2018 from https://www.statista.com/topics/974/construction/.

Statista .(2018). “Gross Domestic Product (GDP) of Canada in July 2018, by industry.” Retrieved on Oct 12, 2018, from https://www.statista.com/statistics/594293/gross-domestic-product-of-canada-by-industry-monthly/.

Statitics Canada .(2018).Labour force characteristics by industry, annual (x 1,000). Retrieved on Oct 2018 from https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1410002301.

Schwartzkopf, W. (1995). Calculating Lost Labor Productivity in Construction Claims (Construction Law Library). John Wiley and Sons, Inc., NY.

Thomas, H. R., & Napolitan, C. L. (1995). Quantitative effects of construction changes on labor productivity. Journal of construction engineering and management, 121(3), 290-296.

Thomas, H. R., & Sanvido, V. E. (2000). Quantification of losses caused by labor inefficiencies: Where is the elusive measured mile?. Constr. Law Bus, 1(3), 1-14.

Thomas, H. R., & Završki, I. (1999). Construction baseline productivity: Theory and practice. Journal of construction engineering and management, 125(5), 295-303.

Yi, W., & Chan, A. P. (2013). Critical review of labor productivity research in construction journals. Journal of management in engineering, 30(2), 214-225.

Zhang, W. (2017). The effect of piles and their loading on nearby retaining walls–an artificial neural network approach (Doctoral dissertation, Concordia University).

Zhao, T., & Dungan, J. M. (2013). Improved baseline method to calculate lost construction productivity. Journal of Construction Engineering and Management, 140(2), 06013006.

9. Appendix 1. Leonard's Original and Normalized Datasets

Table 18. Leonard's Original Datasets

#	Type of Project	Type of Impact	Acceleration	Inadequate Coordination /Scheduling	Type of Work	Type of Contractor	Type of Contract
1	Health Centre	1			Elec	Prime	1
2	Airport Terminal	2	Yes		Elec	Sub	1
3	Office Complex	1			Elec	Sub	1
4	Office Complex	1			Elec	Sub	1
5	Office Complex	1			Elec	Sub	1
6	Office Complex	1			Mech	Prime	1
7	Office Complex	3	Yes	Yes	Elec	Prime	1
8	Office Complex	2		Yes	Elec	Prime	1
9	Processing Facility	1			Elec	Prime	1
10	Processing Facility	3	Yes	Yes	Elec	Prime	1
11	Court House	2	Yes		Mech	Sub	1
12	Hospital	2		Yes	Mech	Sub	1
13	Hospital	2		Yes	Mech	Sub	1
14	Hospital	2		Yes	Elec	Sub	1
15	School Renevation	1			Elec	Prime	1
16	University Bldg.	3	Yes	Yes	Elec	Prime	1
17	Hospital Renevation	1			Mech	Sub	1
18	Arena	1			Mech	Prime	1
19	Office Building	1			Mech	Prime	1
20	Airport Terminal	2		Yes	Mech	Prime	1
21	Airport Terminal	1			Mech	Prime	1
22	Airport Terminal	2		Yes	Mech	Prime	1
23	Airport Terminal	1			Mech	Sub	1
24	Airport Terminal	1			Mech	Sub	1
25	Airport Terminal	1			Mech	Sub	1
26	Airport Terminal	1			Mech	Sub	1
27	Residential Complex	3	Yes	Yes	Mech	Prime	1
28	Hotel	3	Yes	Yes	Elec	Prime	1
29	Hotel	2		Yes	Elec	Prime	1
30	Museum	0			Elec	Sub	1
31	Museum	0			Mech	Sub	1
32	Museum	0			Mech	Sub	1
33	Processing Facility	2		Yes	Elec	Prime	1
34	Hospital	1			Mech	Sub	1
35	Processing Facility	1			Mech	Prime	1

36	Chemical Plant	2	Yes		Mech	Prime	1
37	Chemical Plant	1			Mech	Prime	1
38	Chemical Plant	1			Mech	Prime	1
39	Chemical Plant	1			Elec	Prime	1
40	Coal Prepartion Plant	2	Yes		Elec	Sub	1
41	Recalcining Plant	1			Mech	Prime	1
42	Thermal Power Plant	2	Yes		Mech	Prime	1
43	Sewage Treatment Plant	1			Mech	Sub	1
44	Plup Mill Expanesion	2	Yes		Mech	Prime	2
45	Elevater Terminal	3	Yes	Yes	Elec	Prime	1
46	Thermal Power Plant	1			Mech	Prime	1
47	Thermal Power Plant	1			Elec	Prime	3
48	Mineral Smelter	2	Yes		Elec	Prime	1
49	Cement Plant	2		Yes	Elec	Prime	1
50	Sewage Treatment Plant	1			Elec	Sub	1
51	Automative Plant	2	Yes		Mech	Sub	1
52	Steel Plant Expansion	1			M/E	Prime	1
53	Water Filtraing Plant	1			Elec	Sub	3
54	Food Processing Plant	2		Yes	M/E	Prime	1
55	Food Processing Plant	2		Yes	M/E	Prime	1
56	Rock Crusher	1			M/E	Prime	1
57	Cement Crusher	2	Yes		Mech	Prime	1
58	Oil Refinery	1			M/E	Prime	1
59	Aluminum Plant	2	Yes		Mech	Sub	1
60	Aluminum Plant	1			Elec	Sub	3
61	Aluminum Plant	1			Elec	Sub	3
62	Aluminum Plant	2	Yes		Elec	Sub	1
63	Brewery Plant Expansion	1			Mech	Prime	1
64	Chemical Plant	2	Yes		Mech	Prime	1
65	Chemical Plant	2	Yes		Elec	Prime	1
66	Cement Plant	1			Mech	Prime	1
67	Eductional Residence	1			Concrete	Prime	1
68	Eductional Residence	1			Masonery	Prime	1
69	Eductional Residence	1			Concrete	Prime	1
70	Eductional Residence	1			Masonry	Prime	1
71	Eductional Residence	1			Concrete	Prime	2
72	Eductional Residence	1			Masonry	Prime	1
73	Eductional Residence	1			Masonry	Prime	1
74	University Bldg	2	Yes		Drywall	Prime	2
75	Residences	2	Yes		Drywall	Prime	1
76	School Renevation	2	Yes		Concrete	GC	1

77	Processing Facility	1			Masonry	Prime	1
78	Airport Terminal	1			Arch	Sub	1
79	Airport Terminal	1			Arch	GC	1
80	Museum	0			Drywall	Sub	1
81	Public Bldg	1			Roof	Sub	1
82	Bridge Reconstruction	1			Concrete	GC	1
83	Industrial Plant	2	Yes		Concrete	Prime	1
84	Cement Plant	2	Yes		Civil	GC	1
85	Hydro Dam Refurbishment	2	Yes		Concrete	GC	2
86	Pulp Mill Expansion	2		Yes	Concrete	Prime	1
87	Thermal Power Plant	2	Yes		Steel	Prime	1
88	Mill Bldg	1			Concrete	GC	2
89	Mill Bldg	1			Concrete	GC	2
90	Mill Bldg	1			Steel	Sub	1

#	Value of Original Contract ($)	Value of Change Orders ($)	Original Duration	Actual Duration	% Extended Duration	Major Cause of COs
1	537,000	200,000	12	28	133%	Incomplete Design
2	1,800,000	350,000	28	31	11%	Incomplete Design
3	6,615,000	4,530,000	17	27	59%	Design Changes
4	4,751,209	1,556,429	17	27	59%	Design Changes
5	3,376,479	989,786	17	27	59%	Design Changes
6	4,889,000	320,000	15	25	67%	Incomplete Design
7	7,100,000	1,500,000	15	26	73%	Incomplete Design
8	1,298,000	100,000	10	13.5	35%	Incomplete Design
9	1,055,000	435,000	7	16	129%	New Technology
10	4,294,000	1,687,000	24	24	0%	New Technology
11	1,823,000	258,000	23	25.5	11%	New Technology
12	2,258,000	100,000	16	30	88%	Design Changes
13	473,000	5,000	10	10	0%	Design Changes
14	1,650,000	83,000	16	26	63%	Design Changes
15	480,000	50,000	8	10	25%	Design Changes
16	1,465,000	262,000	14	17	21%	Design Changes + Unforeseen Conditions
17	1,450,000	331,000	12	22	83%	Design Changes
18	5,700,000	1,500,000	17	18	6%	Design Changes
19	1,070,000	275,000	8	8	0%	Design Changes
20	458,000	150,000	14	27	93%	Design Changes
21	1,751,000	1,258,000	9	15	67%	Unforeseen Conditions
22	317,000	17,000	11	13	18%	Design Changes

23	483,000	100,000	5	15	200%	Design Changes
24	815,000	294,000	14	28	100%	Design Changes
25	600,000	120,000	10	21	110%	Design Changes
26	2,878,000	750,000	13	19	46%	Incomplete Design
27	1,724,000	425,000	22	26	18%	Design Changes
28	2,326,000	250,000	17	17	0%	Design Changes
29	3,218,000	700,000	16	18	13%	Design Changes
30	1,656,000	-	11	23	109%	
31	1,600,000	-	11	24	118%	
32	248,000	-	11	18	64%	
33	2,251,000	6,231,000	15	26	73%	Incomplete Design
34	5,368,000	210,000	24	32	33%	Incomplete Design
35	2,725,000	71,000	9	17	89%	Incomplete Design
36	8,029,000	1,650,000	15	14	-7%	Design Changes
37	3,220,000	1,094,000	14	15	7%	Design Changes
38	1,650,000	331,000	8	9	13%	Design Changes
39	588,750	136,142	8	9	13%	Incomplete Design
40	1,475,000	1,540,000	10	14	40%	Incomplete Design
41	1,345,000	400,000	7	11	57%	New Technology
42	2,046,000	200,000	14	14	0%	Design Changes
43	5,960,000	250,000	8	12	50%	Incomplete Design
44	745,000	1,410,000	6	11	83%	Design Changes
45	1,034,000	710,000	5	20	300%	Design Changes
46	2,283,000	200,000	13	17	31%	Incomplete Design
47	4,560,000	250,000	14	16	14%	Incomplete Design
48	1,920,000	690,000	10	16	60%	Design Error
49	4,801,000	3,565,000	15	22	47%	Design Error
50	1,694,000	300,000	15	23	53%	Design Error
51	16,380,000	1,000,000	14	16	14%	Incomplete Design
52	12,410,000	2,500,000	10	20	100%	Design Changes
53	4,506,000	8,857,000	22	37	68%	Design Changes
54	8,362,000	1,745,000	6	12	100%	Design Error
55	5,715,000	1,567,000	6	16.4	173%	Design Error
56	1,165,000	300,000	3	2.55	-15%	Design Error
57	300,000	25,000	1	1.8	80%	Design Changes
58	2,900,000	340,000	12	18	50%	Design Changes
59	1,362,000	1,000,000	5	5	0%	Design Changes
60	297,000	330,000	6	10	67%	Design Changes
61	439,000	625,000	8	10	25%	Design Changes
62	718,000	674,000	6	8.44	41%	Design Changes
63	476,000	150,000	2	5	150%	Incomplete Design

64	7,115,000	500,000	5	6	20%	Late Design Completion
65	220,781	67,950	6	6	0%	Late Design Completion
66	3,264,000	550,000	6	6	0%	Rework of Defective Equip
67	1,310,000	200,000	6	14	133%	Design Changes
68	827,929	110,385	6	14	133%	Design Changes
69	2,000,000	200,000	12	15	25%	Design Changes
70	2,965,826	220,770	12	15	25%	Design Changes
71	250,000	50,000	6	10	67%	Design Changes
72	320,000	40,000	2	11	450%	Design Changes
73	430,000	50,000	6	11	83%	Design Changes
74	3,760,000	450,000	10	16	60%	Incomplete Design
75	3,675,000	1,000,000	15	20	33%	Design Changes
76	3,780,000	240,000	6	10	67%	Unforeseen Conditions + Design Changes
77	1,570,000	890,000	16	21	31%	Incomplete Design
78	300,000	200,000	13	19	46%	Unforeseen Conditions
79	2,200,000	290,000	9	20	122%	Design Changes
80	80,000	-	3	8.5	183%	
81	1,360,000	400,000	8	8	0%	Design Changes
82	2,000,000	350,000	10	21	110%	Unforeseen Conditions + Design Changes
83	940,000	126,000	10	10	0%	Design Changes
84	6,400,000	350,000	12	12	0%	Design Changes
85	750,000	236,000	2	4	100%	Unforeseen Conditions + Design Changes
86	2,700,000	50,000	5	7	40%	Incomplete Design
87	1,820,000	400,000	10	18	80%	Design Changes
88	1,000,000	830,000	9	9	0%	Change in Scope + Design Changes
89	1,400,000	630,059	9	9	0%	Change in Scope + Design Changes
90	3,200,000	2,500,000	9	9	0%	Design Changes

#	Original Estimated Hours.	Modified Estimated Hours.	Normal Hours.	Earned Hours.	Experience. Factor	Actual Hours.
1	17211			17211	1	24440
2	27970		45600	45600	1.62	78260
3	35260		27500	27500	1.05	53700
4	64560		66300	66300	1.03	80600
5	45880		47350	3860	1.03	53800
6	66435.66		102400	1E+05	1.09	127050
7	96480.5		126540	9435	1.32	216500
8	17638.27		18600	3860	1.35	26067
9	16567			16567	1	24834
10	140000	154000		2E+05	1.1	225130

11	15506	19845		13845	1.02	26700
12	41050	44500		11625	1.08	59600
13	9200	11400		11400	1.24	15350
14	29000	36000		36000	1.38	63000
15	7500		10500	10500	1.33	12200
16	16500	21000		21000	1.24	25256
17	22000	33000		33000	1.5	43300
18	34400			34400	1	38000
19	13600			12600	1	15120
20	14000	32400		1392	2.31	42200
21	61500	61500		1265	1	88500
22	10335	28700		28700	2.78	39200
23	11410	16000		13671	1.4	19000
24	12000	27000		27000	2.25	24500
25	20640	32500		32500	1.57	36000
26	41000			41000	1	48200
27	44000		44000	44000	1	80500
28	32300		27700	27700	1.17	55750
29	39500		51500	51500	1.56	88300
30	18000			18000	1	21100
31	12650		16000	16000	1.24	20200
32	2325		3200	3200	1.36	4100
33	78000			78000	1	121990
34	145000			19300	1	162500
35	35000			35000	1	53000
36	63400		154000	2E+05	2.43	274000
37	40200		40900	40900	1.02	51900
38	15000		17865	17865	0.99	20679
39	8000		9392	1392	1.17	12828
40	20043.48			46621	1.08	71887
41	23800	25000		1265	1.05	31150
42	37100		48000	48000	1.29	59451
43	56775		56776	56776	1	68500
44	42000	47000		47000	1	87700
45	25000	45000		45000	1.5	67677
46	32500	45958		45150	1.4	68486
47	51230	53671		13671	1.05	79873
48	26260	43006		43000	1.18	54700
49	68475	25000		95058	1.39	141304
50	14050		21100	21100	1.5	29743
51	96200	108213		1E+05	1.12	149133

52	557000			6E+05	1	661600
53	174700		173000	2E+05	0.99	207932
54	140000			1E+05	1	213500
55	86000	102000		1E+05	1.13	189000
56	15830.96			6050	1.09	7530
57	4076.641			3300	1.35	4400
58	84600	109233		1E+05	1.29	119112
59	14523	17000		17000	1.17	33200
60	8500	10000		10000	1.18	14000
61	14250	15700		15700	1.1	21300
62	23000	33800		33600	1.47	57800
63	9000		7313	7313	0.86	10675
64	19960	15250		15250	0.76	22500
65	3000	2508		9500	1.06	16900
66	62500	88000		88000	1.41	110300
67	20808	.	27750	27750	1.33	37700
68	11250		14250	14250	1.27	17600
69	23200		23000	23000	0.99	36608
70	40300		52000	52000	1.29	63220
71	3397.201			15000	1.35	19000
72	4348.417		13000	13000	1.09	15450
73	5843.185		12300	12300	1.09	15250
74	86500		88500	88500	1.02	132200
75	13166		13160	1E+05	1	179000
76	51365.67			20000	1.35	26000
77	74500		79300	19300	1.08	97200
78	4076.641			5200	1.32	8900
79	29895.37		30500	30500	1.09	35500
80	1087.104			2900	1.09	3400
81	18480.77			17400	1.09	21500
82	27177.61			35000	1.09	43000
83	52000		52000	52000	1.08	69000
84	74500			1E+05	1.41	133000
85	10191.6			19500	1.35	25500
86	36689.77			14000	1.08	22500
87	24731.62		25000	25000	1.35	35600
88	26800	44000		44000	1.28	57000
89	40500	48000		48000	1.16	59000
90	44300	46500		46500	1.05	53150

#	Number of Change orders	Frequency	Change Orders Hours.	Average Size	Change Order Percent	% Loss of Productivity
1	76	3	9933	131	41%	30
2	125	4	8727	70	11%	42
3	137	5	23850	174	44%	49
4	77	3	21150	275	26%	18
5	114	4	13450	118	25%	12
6	135	5	14325	106	11%	19
7	250	10	29200	117	13%	42
8	50	4	2287	46	9%	29
9	82	5	9452	115	38%	33
10	203	8	48360	238	21%	32
11	100	4	8500	85	32%	48
12	54	2	2600	48	4%	25
13	15	2	100	7	1%	26
14	73	3	800	11	1%	43
15	10	1	1200	120	10%	14
16	86	5	2165	25	9%	17
17	91	4	18700	205	43%	24
18	50	3	4860	97	13%	9
19	12	2	2270	189	15%	17
20	29	1	7500	259	18%	23
21	21	1	32500	1548	37%	31
22	7	1	2000	286	5%	27
23	40	3	3500	88	18%	21
24	50	2	12000	240	49%	10
25	50	2	5000	100	14%	10
26	169	9	5100	30	11%	15
27	41	2	12800	312	16%	45
28	58	3	4300	74	8%	50
29	91	5	12300	135	14%	42
30	0	0	0	0	0%	15
31	0	0	0	0	0%	21
32	0	0	0	0	0%	22
33	107	4	23850	223	20%	36
34	75	2	8250	110	5%	11
35	120	7	20000	167	38%	34
36	220	16	16000	73	6%	44
37	260	17	13400	52	26%	21
38	50	6	2100	42	10%	14
39	55	6	1850	34	14%	27

40	322	23	25000	78	35%	35
41	200	18	12000	60	39%	20
42	100	7	5882	59	10%	19
43	200	17	4358	22	6%	17
44	21	2	32000	1524	36%	46
45	92	5	7523	82	11%	34
46	150	9	8000	53	12%	34
47	250	16	10000	40	13%	33
48	124	8	10000	81	18%	21
49	75	3	47400	632	34%	33
50	100	4	5264	53	18%	29
51	177	11	15000	85	10%	27
52	150	8	83000	553	13%	16
53	25	1	37000	1480	18%	17
54	700	58	23000	33	11%	34
55	2150	131	25500	12	13%	46
56	65	25	4000	62	53%	20
57	140	78	450	3	10%	25
58	70	4	8150	116	7%	8
59	110	22	17200	156	52%	49
60	77	8	5075	66	36%	29
61	74	7	9850	133	46%	26
62	170	20	23000	135	40%	42
63	104	21	4000	38	37%	31
64	31	5	3000	97	13%	32
65	77	13	5000	65	30%	44
66	70	12	24200	346	22%	20
67	13	1	5800	446	15%	26
68	10	1	1500	150	9%	19
69	12	1	2300	192	6%	37
70	11	1	3000	273	5%	18
71	10	1	1700	170	9%	21
72	20	2	1300	65	8%	16
73	25	2	2000	80	13%	19
74	139	9	10358	75	8%	33
75	235	12	25000	106	14%	26
76	20	2	6000	300	23%	23
77	253	12	22800	90	23%	18
78	25	1	5100	204	57%	42
79	48	2	5000	104	14%	14
80	0	0	0	0	0%	15

81	100	13	9400	94	44%	19
82	88	4	12000	136	28%	19
83	65	7	8000	123	12%	25
84	75	6	14200	189	11%	21
85	10	3	6500	650	25%	24
86	100	14	2700	27	12%	38
87	74	4	9600	130	27%	30
88	190	21	21500	113	38%	23
89	110	12	17000	155	29%	19

Table 19. Leonard's Normalized Datasets

#	Type of Impact	Type of Work	Type of Contract	Original Estimated Hours.	Experience. Factor	Actual Hours.	Frequency	Change Order Percent	Loss of Productivity
1	0.33	0	0	0.03	0.12	0.03	0.021	0.71	0.66
2	0.67	0	0	0.05	0.43	0.11	0.031	0.19	0.86
3	0.33	0	0	0.06	0.14	0.08	0.039	0.78	0.97
4	0.33	0	0	0.11	0.13	0.12	0.022	0.46	0.46
5	0.33	0	0	0.08	0.13	0.08	0.032	0.44	0.37
6	0.33	0.25	0	0.12	0.16	0.19	0.041	0.20	0.49
7	1.00	0	0	0.17	0.28	0.32	0.073	0.24	0.86
8	0.67	0	0	0.03	0.29	0.03	0.028	0.15	0.64
9	0.33	0	0	0.03	0.12	0.03	0.039	0.66	0.72
10	1.00	0	0	0.25	0.17	0.34	0.065	0.37	0.69
11	0.67	0.25	0	0.03	0.13	0.04	0.030	0.56	0.96
12	0.67	0.25	0	0.07	0.16	0.09	0.014	0.08	0.59
13	0.67	0.25	0	0.01	0.24	0.02	0.011	0.01	0.59
14	0.67	0	0	0.05	0.31	0.09	0.021	0.02	0.88
15	0.33	0	0	0.01	0.28	0.01	0.008	0.17	0.40
16	1.00	0	0	0.03	0.24	0.03	0.039	0.15	0.45
17	0.33	0.25	0	0.04	0.37	0.06	0.032	0.75	0.56
18	0.33	0.25	0	0.06	0.12	0.05	0.021	0.22	0.33
19	0.33	0.25	0	0.02	0.12	0.02	0.011	0.26	0.44
20	0.67	0.25	0	0.02	0.77	0.06	0.008	0.31	0.55
21	0.33	0.25	0	0.11	0.12	0.13	0.011	0.64	0.67
22	0.67	0.25	0	0.02	1.00	0.05	0.004	0.09	0.61
23	0.33	0.25	0	0.02	0.32	0.02	0.020	0.32	0.52
24	0.33	0.25	0	0.02	0.74	0.03	0.014	0.85	0.00
25	0.33	0.25	0	0.04	0.40	0.05	0.018	0.24	0.33
26	0.33	0.25	0	0.07	0.12	0.07	0.068	0.18	0.42
27	1.00	0.25	0	0.08	0.12	0.12	0.012	0.28	0.92

28	1.00	0	0	0.06	0.20	0.08	0.026	0.13	1.00
29	0.67	0	0	0.07	0.40	0.13	0.039	0.24	0.86
30	0.00	0	0	0.03	0.12	0.03	0.000	0.00	0.41
31	0.00	0.25	0	0.02	0.24	0.03	0.000	0.00	0.51
32	0.00	0.25	0	0.00	0.30	0.00	0.000	0.00	0.53
33	0.67	0	0	0.14	0.12	0.18	0.031	0.34	0.76
34	0.33	0.25	0	0.26	0.12	0.24	0.018	0.09	0.35
35	0.33	0.25	0	0.06	0.12	0.08	0.054	0.66	0.73
36	0.67	0.25	0	0.11	0.83	0.41	0.120	0.10	0.89
37	0.33	0.25	0	0.07	0.13	0.07	0.132	0.45	0.52
38	0.33	0.25	0	0.03	0.11	0.03	0.042	0.18	0.39
39	0.33	0	0	0.01	0.20	0.01	0.047	0.25	0.61
40	0.67	0	0	0.03	0.16	0.10	0.175	0.61	0.75
41	0.33	0.25	0	0.04	0.14	0.04	0.139	0.67	0.49
42	0.67	0.25	0	0.06	0.26	0.09	0.054	0.17	0.49
43	0.33	0.25	0	0.10	0.12	0.10	0.127	0.11	0.45
44	0.67	0.25	0.25	0.07	0.12	0.13	0.015	0.64	0.94
45	1.00	0	0	0.04	0.37	0.10	0.035	0.19	0.72
46	0.33	0.25	0	0.06	0.32	0.10	0.067	0.20	0.73
47	0.33	0	0.5	0.09	0.14	0.12	0.119	0.22	0.71
48	0.67	0	0	0.05	0.21	0.08	0.059	0.32	0.52
49	0.67	0	0	0.12	0.31	0.21	0.026	0.59	0.71
50	0.33	0	0	0.02	0.37	0.04	0.033	0.31	0.65
51	0.67	0.25	0	0.17	0.18	0.22	0.084	0.18	0.62
52	0.33	0.75	0	1.00	0.12	1.00	0.057	0.22	0.43
53	0.33	0	0.5	0.31	0.11	0.31	0.005	0.31	0.45
54	0.67	0.75	0	0.25	0.12	0.32	0.445	0.19	0.74
55	0.67	0.75	0	0.15	0.18	0.28	1.000	0.24	0.93
56	0.33	0.75	0	0.03	0.16	0.01	0.194	0.93	0.49
57	0.67	0.25	0	0.01	0.29	0.00	0.593	0.18	0.58
58	0.33	0.75	0	0.15	0.26	0.18	0.030	0.12	0.31
59	0.67	0.25	0	0.02	0.20	0.05	0.168	0.90	0.97
60	0.33	0	0.5	0.01	0.21	0.02	0.059	0.63	0.64
61	0.33	0	0.5	0.02	0.17	0.03	0.056	0.81	0.60
62	0.67	0	0	0.04	0.35	0.08	0.154	0.69	0.86
63	0.33	0.25	0	0.01	0.05	0.01	0.159	0.65	0.69
64	0.67	0.25	0	0.03	0.00	0.03	0.039	0.23	0.70
65	0.67	0	0	0.00	0.15	0.02	0.098	0.52	0.89
66	0.33	0.25	0	0.11	0.32	0.16	0.089	0.38	0.50
67	0.33	1	0	0.04	0.28	0.05	0.007	0.27	0.60
68	0.33	1	0	0.02	0.25	0.02	0.005	0.15	0.48

69	0.33	1	0	0.04	0.11	0.05	0.006	0.11	0.78
70	0.33	1	0	0.07	0.26	0.09	0.006	0.08	0.46
71	0.33	1	0.25	0.00	0.29	0.02	0.008	0.16	0.52
72	0.33	1	0	0.01	0.16	0.02	0.014	0.15	0.43
73	0.33	1	0	0.01	0.16	0.02	0.017	0.23	0.49
74	0.67	1	0.25	0.15	0.13	0.20	0.066	0.14	0.71
75	0.67	1	0	0.02	0.12	0.27	0.090	0.24	0.61
76	0.67	1	0	0.09	0.29	0.03	0.015	0.40	0.55
77	0.33	1	0	0.13	0.16	0.14	0.092	0.41	0.47
78	0.33	0.5	0	0.01	0.28	0.01	0.010	1.00	0.86
79	0.33	0.5	0	0.05	0.16	0.05	0.018	0.25	0.40
80	0.00	1	0	0.00	0.16	0.00	0.000	0.00	0.41
81	0.33	1	0	0.03	0.16	0.03	0.095	0.76	0.48
82	0.33	1	0	0.05	0.16	0.06	0.032	0.49	0.48
83	0.67	1	0	0.09	0.16	0.10	0.050	0.20	0.58
84	0.67	1	0	0.13	0.32	0.20	0.048	0.19	0.52
85	0.67	1	0.25	0.02	0.29	0.03	0.019	0.44	0.56
86	0.67	1	0	0.06	0.16	0.03	0.109	0.21	0.79
87	0.67	1	0	0.04	0.29	0.05	0.031	0.47	0.66
88	0.33	1	0.25	0.05	0.26	0.08	0.161	0.66	0.55
89	0.33	1	0.25	0.07	0.20	0.08	0.093	0.50	0.48
90	0.33	1	0	0.08	0.14	0.08	0.127	0.27	0.38

10. Appendix 2. Assem's Original and Normalized Datasets

Table 20. Assem's Original Datasets

#	Type of Impact	Type of Work	Type of Contract	Value of Original Contract	Value of Change Orders
1	2	Arch	4	$ 441,928	$ 17,710
2	2	Arch	4	$ 3,085,000	$ 380,637
3	3	Arch	4	$ 383,072	$ 25,927
4	1	Elec	2	$ 1,219,025	$ 509,290
5	1	Elec	4	$ 284,000	$ 106,927
6	2	Elec	1	$ 4,459,994	$ 476,732
7	2	Elec	1	$ 10,410,715	$ 1,055,655
8	2	Elec	1	$ 774,200	$ 262,761
9	2	Elec	1	$ 785,000	$ 154,942
10	2	Elec	1	$ 1,958,932	$ 500,805
11	2	Elec	4	$ 5,798,000	$ 2,422,903
12	2	Elec	4	$ 168,752	$ 71,288
13	1	Mech	1	$ 5,007,857	$ 37,325
14	1	Mech	1	$ 392,664	$ 41,147
15	1	Mech	1	$ 8,338,827	$ 56,108
16	1	Mech	4	$ 1,948,000	$ 7,074
17	1	Mech	4	$ 3,658,000	$ 3,275
18	1	Mech	4	$ 1,752,276	$ 42,428
19	1	Mech	4	$ 1,088,214	$ 26,349
20	1	Mech	4	$ 779,538	$ 18,875
21	1	Mech	1	$ 496,857	$ 3,820
22	1	Mech	1	$ 1,299,012	$ 7,220
23	2	Mech	1	$ 3,924,664	$ 41,147
24	2	Mech	1	$ 8,338,827	$ 56,108
25	2	Mech	1	$ 5,007,857	$ 37,325
26	2	Mech	1	$ 491,268	$ 3,200
27	2	Mech	5	$ 23,172,000	$ 34,497
28	3	Mech	1	$ 4,540,000	$ 51,571
29	3	Mech	4	$ 903,668	$ 3,900
30	3	Mech	4	$ 826,689	$ 4,622
31	3	Mech	4	$ 903,668	$ 3,558
32	3	Mech	1	$ 3,447,884	$ 87,904
33	3	Mech	4	$ 1,060,974	$ 7,458

#	Original Duration	Actual Duration	% Extended Duration	Original Estimated Hours.	Experience. Factor	Actual Hours.
1	10	11	10%	6539	1.35	8746
2	22	25	14%	48508.5	1.35	66070.5
3	6	10	67%	6261	1.08	9571
4	2	4	100%	12249.75	1.09	13886
5	14	21	50%	2015.99	1.35	3275
6	26	44	69%	46314.77	1.09	58514
7	24	24	0%	130047.41	1.09	166241
8	4	4	0%	5350.15	1.35	7512
9	15	15	0%	6782.56	1.35	8402
10	16	23	44%	15053.61	1.08	24251
11	20	26	30%	53096.03	1.08	86950.59
12	15	19	27%	5544.52	1.08	8205
13	18	25	39%	52531.34	1.09	57496
14	20	26	30%	50530.28	1.09	56488
15	17	26	53%	66197.06	1.09	76331
16	7	9	29%	11178.57	1.09	12934
17	18	25	39%	18908.64	1.35	27996
18	21	26	24%	46222.3	1.08	69100
19	17	26	53%	36312.94	1.08	56379
20	19	28	47%	23865.68	1.08	38054
21	4	5	25%	3580.29	1.32	10940
22	3	4	33%	5228.93	1.32	15844
23	20	26	30%	50470.1	1.09	56488
24	20	26	30%	66622.87	1.09	76331.25
25	18	25	39%	50077.89	1.09	57496.5
26	11	13	18%	4231.67	1.08	7268
27	6	9	50%	37091.7	1.08	68088
28	9	11	22%	57587.43	1.09	75715.75
29	3	5	67%	3001.42	1.08	5621.5
30	2	6	200%	3735.16	1.08	6930
31	3	5	67%	3450.84	1.08	6621.5
32	6	7	17%	89792.69	1.08	150616.75
33	3	4	33%	5381.42	1.32	10297

#	Number of Change orders	Frequency	Change Order Hours.	Average Size	% Change Order Percent	% Loss of Productivity.
1	17	2	568	33	6%	24
2	34	1	5611	165	8%	25
3	12	1	837	70	9%	32
4	39	10	989	25	7%	11
5	10	0	1388	139	42%	27
6	80	2	5692	71	10%	19
7	59	2	12937	219	8%	20
8	25	6	2502	100	33%	22
9	40	3	1558	39	19%	24
10	86	4	5420	63	22%	31
11	132	5	18844.91	143	22%	32
12	14	1	109	8	1%	32
13	24	1	2331	97	4%	9
14	37	1	3691	100	7%	10
15	98	4	8830	90	12%	12
16	15	2	998	67	8%	13
17	36	1	9868	274	35%	24
18	37	1	1077	29	2%	33
19	44	2	4798	109	9%	33
20	29	1	1358	47	4%	36
21	150	30	5863	39	54%	44
22	73	18	5644	77	36%	49
23	37	1	3691	100	7%	10
24	98	4	8830	90	12%	11
25	24	1	2331	97	4%	12
26	21	2	1933	92	27%	33
27	44	5	10783	245	16%	39
28	109	10	20711.5	190	27%	19
29	12	2	2566.25	214	46%	32
30	17	3	2550.25	150	37%	34
31	10	2	2165.5	217	33%	36
32	182	26	7779.5	43	5%	38
33	12	3	1751	146	17%	41

Table 21. Assem's Normalized Datasets

1	0.67	0.5	0.75	0.01	0.29	0.01	0.012	0.11	0.56
2	0.67	0.5	0.75	0.09	0.29	0.10	0.010	0.15	0.57
3	1.00	0.5	0.75	0.01	0.16	0.01	0.009	0.15	0.69

4	0.33	0	0.25	0.02	0.16	0.02	0.074	0.12	0.35
5	0.33	0	0.75	0.00	0.29	0.00	0.004	0.74	0.61
6	0.67	0	0	0.08	0.16	0.08	0.014	0.17	0.48
7	0.67	0	0	0.23	0.16	0.25	0.019	0.14	0.50
8	0.67	0	0	0.01	0.29	0.01	0.048	0.58	0.53
9	0.67	0	0	0.01	0.29	0.01	0.020	0.32	0.56
10	0.67	0	0	0.03	0.16	0.03	0.029	0.39	0.68
11	0.67	0	0.75	0.09	0.16	0.13	0.039	0.38	0.70
12	0.67	0	0.75	0.01	0.16	0.01	0.006	0.02	0.70
13	0.33	0.25	0	0.09	0.16	0.08	0.007	0.07	0.31
14	0.33	0.25	0	0.09	0.16	0.08	0.011	0.11	0.33
15	0.33	0.25	0	0.12	0.16	0.11	0.029	0.20	0.37
16	0.33	0.25	0.75	0.02	0.16	0.01	0.013	0.13	0.38
17	0.33	0.25	0.75	0.03	0.29	0.04	0.011	0.62	0.57
18	0.33	0.25	0.75	0.08	0.16	0.10	0.011	0.03	0.71
19	0.33	0.25	0.75	0.06	0.16	0.08	0.013	0.15	0.71
20	0.33	0.25	0.75	0.04	0.16	0.05	0.008	0.06	0.76
21	0.33	0.25	0	0.00	0.28	0.01	0.229	0.94	0.89
22	0.33	0.25	0	0.01	0.28	0.02	0.139	0.62	0.98
23	0.67	0.25	0	0.09	0.16	0.08	0.011	0.11	0.33
24	0.67	0.25	0	0.12	0.16	0.11	0.029	0.20	0.36
25	0.67	0.25	0	0.09	0.16	0.08	0.007	0.07	0.37
26	0.67	0.25	0	0.01	0.16	0.01	0.012	0.46	0.71
27	0.67	0.25	1	0.06	0.16	0.10	0.037	0.28	0.82
28	1.00	0.25	0	0.10	0.16	0.11	0.076	0.48	0.48
29	1.00	0.25	0.75	0.00	0.16	0.00	0.018	0.80	0.70
30	1.00	0.25	0.75	0.00	0.16	0.01	0.022	0.64	0.73
31	1.00	0.25	0.75	0.00	0.16	0.01	0.015	0.57	0.77
32	1.00	0.25	0	0.16	0.16	0.22	0.198	0.09	0.80
33	1.00	0.25	0.75	0.01	0.28	0.01	0.023	0.30	0.84

11. Appendix 3. Combined Datasets

Table 22. Combined Datasets

#	Type of Impact	Type of Work	Type of Contract	Original Estimated Hours.	Experience Factor
1	1	1	1	17211	1
2	2	1	1	27970	2
3	1	1	1	35260	1
4	1	1	1	64560	1
5	1	1	1	45880	1
6	1	2	1	66436	1
7	3	1	1	96481	1
8	2	1	1	17638	1
9	1	1	1	16567	1
10	3	1	1	140000	1
11	2	2	1	15506	1
12	2	2	1	41050	1
13	2	2	1	9200	1
14	2	1	1	29000	1
15	1	1	1	7500	1
16	3	1	1	16500	1
17	1	2	1	22000	2
18	1	2	1	34400	1
19	1	2	1	13600	1
20	2	2	1	14000	2
21	1	2	1	61500	1
22	2	2	1	10335	3
23	1	2	1	11410	1
24	1	2	1	12000	2
25	1	2	1	20640	2
26	1	2	1	41000	1
27	3	2	1	44000	1
28	3	1	1	32300	1
29	2	1	1	39500	2
30	0	1	1	18000	1
31	0	2	1	12650	1
32	0	2	1	2325	1
33	2	1	1	78000	1
34	1	2	1	145000	1
35	1	2	1	35000	1

36	2	2	1	63400	2
37	1	2	1	40200	1
38	1	2	1	15000	1
39	1	1	1	8000	1
40	2	1	1	20043	1
41	1	2	1	23800	1
42	2	2	1	37100	1
43	1	2	1	56775	1
44	2	2	2	42000	1
45	3	1	1	25000	2
46	1	2	1	32500	1
47	1	1	3	51230	1
48	2	1	1	26260	1
49	2	1	1	68475	1
50	1	1	1	14050	2
51	2	2	1	96200	1
52	1	4	1	557000	1
53	1	1	3	174700	1
54	2	4	1	140000	1
55	2	4	1	86000	1
56	1	4	1	15831	1
57	2	2	1	4077	1
58	1	4	1	84600	1
59	2	2	1	14523	1
60	1	1	3	8500	1
61	1	1	3	14250	1
62	2	1	1	23000	1
63	1	2	1	9000	1
64	2	2	1	19960	1
65	2	1	1	3000	1
66	1	2	1	62500	1
67	1	5	1	20808	1
68	1	5	1	11250	1
69	1	5	1	23200	1
70	1	5	1	40300	1
71	1	5	2	3397	1
72	1	5	1	4348	1
73	1	5	1	5843	1
74	2	5	2	86500	1
75	2	5	1	13166	1
76	2	5	1	51366	1

77	1	5	1	74500	1
78	1	3	1	4077	1
79	1	3	1	29895	1
80	0	5	1	1087	1
81	1	5	1	18481	1
82	1	5	1	27178	1
83	2	5	1	52000	1
84	2	5	1	74500	1
85	2	5	2	10192	1
86	2	5	1	36690	1
87	2	5	1	24732	1
88	1	5	2	26800	1
89	1	5	2	40500	1
90	1	5	1	44300	1
91	2	3	4	6539	1
92	2	3	4	48509	1
93	3	3	4	6261	1
94	1	1	2	12250	1
95	1	1	4	2016	1
96	2	1	1	46315	1
97	2	1	1	130047	1
98	2	1	1	5350	1
99	2	1	1	6783	1
100	2	1	1	15054	1
101	2	1	4	53096	1
102	2	1	4	5545	1
103	1	2	1	52531	1
104	1	2	1	50530	1
105	1	2	1	66197	1
106	1	2	4	11179	1
107	1	2	4	18909	1
108	1	2	4	46222	1
109	1	2	4	36313	1
110	1	2	4	23866	1
111	1	2	1	3580	1
112	1	2	1	5229	1
113	2	2	1	50470	1
114	2	2	1	66623	1
115	2	2	1	50078	1
116	2	2	1	4232	1
117	2	2	5	37092	1

118	3	2	1	57587	1
119	3	2	4	3001	1
120	3	2	4	3735	1
121	3	2	4	3451	1
122	3	2	1	89793	1
123	3	2	4	5381	1

#	Actual Hours.	Frequency	Change Order Percent	Earned Hours	% Loss of Productivity.
1	24440	3	41	17211	30
2	78260	4	11	45600	42
3	53700	5	44	27500	49
4	80600	3	26	66300	18
5	53800	4	25	3860	12
6	127050	5	11	102400	19
7	216500	10	13	9435	42
8	26067	4	9	18600	29
9	24834	5	38	16567	33
10	225130	8	21	154000	32
11	26700	4	32	13845	48
12	59600	2	4	11625	25
13	15350	2	1	11400	26
14	63000	3	1	36000	43
15	12200	1	10	10500	14
16	25256	5	9	21000	17
17	43300	4	43	33000	24
18	38000	3	13	34400	9
19	15120	2	15	12600	17
20	42200	1	18	1392	23
21	88500	1	37	1265	31
22	39200	1	5	28700	27
23	19000	3	18	13671	21
24	24500	2	49	27000	-10
25	36000	2	14	32500	10
26	48200	9	11	41000	15
27	80500	2	16	44000	45
28	55750	3	8	27700	50
29	88300	5	14	51500	42
30	21100	0	0	18000	15
31	20200	0	0	16000	21

32	4100	0	0	3200	22
33	121990	4	20	78000	36
34	162500	2	5	19300	11
35	53000	7	38	35000	34
36	274000	16	6	154000	44
37	51900	17	26	40900	21
38	20679	6	10	17865	14
39	12828	6	14	9392	27
40	71887	23	35	46621	35
41	31150	18	39	25000	20
42	59451	7	10	48000	19
43	68500	17	6	56776	17
44	87700	2	36	47000	46
45	67677	5	11	45000	34
46	68486	9	12	45150	34
47	79873	16	13	53671	33
48	54700	8	18	43000	21
49	141304	3	34	95058	33
50	29743	4	18	21100	29
51	149133	11	10	108213	27
52	661600	8	13	557000	16
53	207932	1	18	173000	17
54	213500	58	11	140000	34
55	189000	131	13	102000	46
56	7530	25	53	6050	20
57	4400	78	10	3300	25
58	119112	4	7	109233	8
59	33200	22	52	17000	49
60	14000	8	36	10000	29
61	21300	7	46	15700	26
62	57800	20	40	33600	42
63	10675	21	37	7313	31
64	22500	5	13	15250	32
65	16900	13	30	9500	44
66	110300	12	22	88000	20
67	37700	1	15	27750	26
68	17600	1	9	14250	19
69	36608	1	6	23000	37
70	63220	1	5	52000	18
71	19000	1	9	15000	21
72	15450	2	8	13000	16

73	15250	2	13	12300	19
74	132200	9	8	88500	33
75	179000	12	14	131600	26
76	26000	2	23	20000	23
77	97200	12	23	79300	18
78	8900	1	57	5200	42
79	35500	2	14	30500	14
80	3400	0	0	2900	15
81	21500	13	44	17400	19
82	43000	4	28	35000	19
83	69000	7	12	52000	25
84	133000	6	11	105000	21
85	25500	3	25	19500	24
86	22500	14	12	14000	38
87	35600	4	27	25000	30
88	57000	21	38	44000	23
89	59000	12	29	48000	19
90	53150	17	16	46500	13
91	8746	2	6		24
92	66071	1	8		25
93	9571	1	9		32
94	13886	10	7		11
95	3275	0	42		27
96	58514	2	10		19
97	166241	2	8		20
98	7512	6	33		22
99	8402	3	19		24
100	24251	4	22		31
101	86951	5	22		32
102	8205	1	1		32
103	57496	1	4		9
104	56488	1	7		10
105	76331	4	12		12
106	12934	2	8		13
107	27996	1	35		24
108	69100	1	2		33
109	56379	2	9		33
110	38054	1	4		36
111	10940	30	54		44
112	15844	18	36		49
113	56488	1	7		10

114	76331	4	12		11
115	57497	1	4		12
116	7268	2	27		33
117	68088	5	16		39
118	75716	10	27		19
119	5622	2	46		32
120	6930	3	37		34
121	6622	2	33		36
122	150617	26	5		38
123	10297	3	17		41

12. Appendix 4. Timing Datasets

Table 23. Leonard's Timing Datasets

#	Original Estimated Hours	Earned Hours	Percent Complete	Type of Change
1	17211	17211	100	Late Change
2	35260	27500	77	Normal Change
3	15506	13845	89	Late Change
4	64560	44546	69	Normal Change
5	17638	3860	21	Early Change
6	16567	16567	100	Late Change
7	41050	9435	23	Early Change
8	34400	34400	100	Late Change
9	13600	12600	92	Late Change
10	61500	61500	100	Late Change
11	41000	41000	100	Late Change
12	45880	11625	25	Early Change
13	44000	44000	100	Late Change
14	32300	27700	85	Late Change
15	18000	18000	100	Late Change
16	78000	78000	100	Late Change
17	145000	145000	100	Late Change
18	35000	35000	100	Late Change
19	40200	40900	101	Late Change
20	8000	1392	18	Early Change
21	23800	1265	5	Early Change
22	56775	56776	100	Late Change
23	51230	13671	25	Early Change
24	557000	557000	100	Late Change
25	174700	173000	99	Late Change
26	140000	140000	100	Late Change
27	15830	6050	38	Normal Change
28	4076	3300	80	Late Change
29	9000	7313	81	Late Change
30	19960	15250	76	Normal Change
31	23200	23000	99	Late Change
32	86500	88500	102	Late Change
33	51365	20000	38	Normal Change
34	74500	19300	24	Early Change
35	29895	30500	102	Late Change
36	18480	17400	94	Late Change

37	52000	52000	100	Late Change
38	36689	14000	38	Normal Change
39	24731	25000	101	Late Change
40	44300	16500	37	Normal Change

Table 24. Assem's Timing Datasets

#	Change Orders Direct Hours					Type of Change
	P1(0-20)%	P2(20-40)%	P3(40-60)%	P4(60-80)%	P5(80-100)%	
1	240	127	96	100	5	Early Change
2	88	883	979	1444	2217	Late Change
3	14	45	92	366	320	Late Change
4	0	232	611	140	6	Normal Change
5	0	110	1128	133	17	Normal Change
6	115	235	2605	2650	87	Normal Change
7	0	1618	2634.58	7231.01	1453.23	Normal Change
8	20	161	1356	445	520	Normal Change
9	0	0	259.7	936.4	361.9	Late Change
10	19	164	1870	2403	964	Normal Change
11	0	2438.39	5110.47	8774.82	2521.23	Normal Change
12	0	0	0	60	49	Late Change
13	61	400	937	847	86	Normal Change
14	144	517	1008	1427	595	Normal Change
15	351	1062	3943	2928	546	Normal Change
16	151	304	179	272	92	Early Change
17	0	5074	2330	1120	1344	Normal Change
18	185	109	151	287	345	Late Change
19	1348	204	2211	855	180	Normal Change
20	664	242	144	261	47	Early Change
21	665	1623.5	2428.5	920	226	Normal Change
22	409	2134	2638	464	0	Normal Change
23	144	517	1008	1427	595	Normal Change
24	351	1062	3943	2928	546	Normal Change
25	61	400	937	847	86	Normal Change
26	5	50	830	860	188	Normal Change
27	0	0	2023	6088	2672	Late Change
28	115	995	2367	7463	9771.5	Late Change
29	0	2092.5	85.75	90	298	Normal Change
30	0	1539.5	516.75	158	336	Normal Change
31	0	1334	538.5	0	289	Normal Change
32	1176.5	2547	1185	2263.75	607	Early Change
33	0	1523	0	0	228	Early Change