



AMERICAN UNIVERSITY OF BEIRUT

LEARNING CURVES AND THE  
CONSTRUCTION INDUSTRY

by

DAOUD ARAM KIOMJIAN

A thesis

submitted in partial fulfillment of the requirements  
for the degree of Master of Engineering Management  
to the Department of Industrial and Engineering Management  
of the Faculty of Engineering and Architecture  
at the American University of Beirut

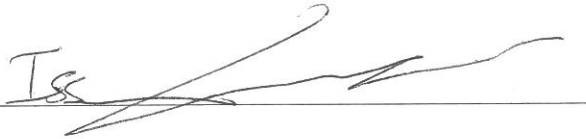
Beirut, Lebanon  
April 2016

AMERICAN UNIVERSITY OF BEIRUT

LEARNING CURVES AND THE  
CONSTRUCTION INDUSTRY

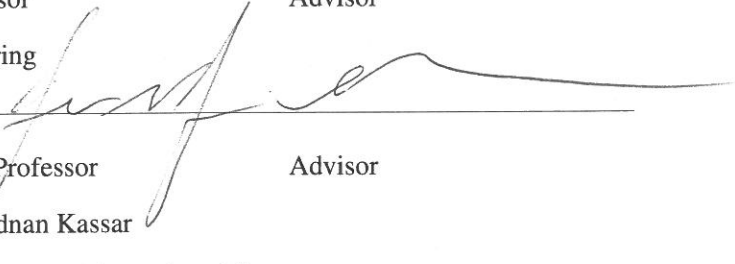
by  
DAOUD ARAM KIOMJIAN

Approved by:



Dr. Issam Srour, Associate Professor  
Civil and Environmental Engineering

Advisor



Dr. Faith Jordan Srour, Assistant Professor  
Lebanese American University-Adnan Kassar

Advisor

School of Business-Department of IT and Operations Management



Dr. Walid Nasr, Assistant Professor

Member of Committee

Department of Industrial and Engineering Management

Date of thesis defense: April 15, 2016

# AMERICAN UNIVERSITY OF BEIRUT

## THESIS, DISSERTATION, PROJECT RELEASE FORM

Student Name: Kiomyjian Daoud Aram  
Last First Middle

Master's Thesis       Master's Project       Doctoral Dissertation

I authorize the American University of Beirut to: (a) reproduce hard or electronic copies of my thesis, dissertation, or project; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes.

I authorize the American University of Beirut, **three years after the date of submitting my thesis, dissertation, or project**, to: (a) reproduce hard or electronic copies of it; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes.

Daoud 03/05/2016  
Signature Date

# Acknowledgements

My sincere gratitude goes to my advisors, Professor Issam M. Srour and Professor Faith Jordan Srour for their continuous support, motivation and patience during the course of this research.

My recognition and gratitude is also addressed for the University Research Board at the American University of Beirut for funding this research.

I would also like to thank, my thesis committee member, Professor Walid Nasr for his insightful comments and encouragement.

# An Abstract of the Thesis of

Daoud Aram Kiomjian for Master of Engineering  
Major: Engineering Management

Title: Learning Curves and the Construction Industry

Learning is expected to have a significant effect on the performance of construction crews, this performance is expected to improve with experience and repetition. This is particularly true for repetitive construction projects, where the worker repeats the same task multiple times throughout the course of the project. This performance improvement is numerically portrayed mathematically by Learning Curves. Researchers have developed numerous learning curve models that vary in complexity and purpose. However, learning curve theory is not yet very popular among industry practitioners. The main driver behind this shy popularity, is the lack of consensus in the literature regarding a learning curve model that best suites the construction industry.

In an attempt to rectify the above shortcoming, this study presents a new learning curve model. The presented model resembles the traditional Wright model by assuming an exponential form; however it employs recursion in order to place more emphasis on recent data. The research methods, used to develop this method and other aspects of this study include, a literature survey and case study analysis for a real life construction project and other case studies that were extracted from the literature. The developed model and the findings of the literature survey were used to develop a learning based automated scheduling tool. This tool displayed acceptable performance, when tested on case studies.

The applicability of learning curve theory was extended for another dimension of construction projects, which is quality. The results of a real life construction project from the MENA region, have revealed a correlation between learning and productivity, however the same was not observed for learning and quality. The findings of this paper indicate that the learning curve model to be used varies according to the project characteristics and location. The findings also indicate that the relationship between learning and quality is more complex than that between learning and productivity.

# Contents

<b>Acknowledgements</b>	<b>v</b>
<b>Abstract</b>	<b>vi</b>
<b>List of Figures</b>	<b>ix</b>
<b>List of Tables</b>	<b>x</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Learning Curves and Productivity in the Construction Industry:A Review and a Proposed Model</b>	<b>4</b>
2.1 Background and Literature Review . . . . .	4
2.1.1 The Learning Process . . . . .	5
2.1.2 Learning Curve Models . . . . .	7
2.1.3 Learning and Forgetting . . . . .	13
2.2 The Proposed Model . . . . .	14
2.3 Empirical Evaluation of Learning Models . . . . .	15
2.4 Summary and Future Research Directions . . . . .	18
<b>3 Learning Curves and Quality in Construction: A Case Study</b>	<b>21</b>
3.1 Introduction and Background . . . . .	21
3.2 Case Study:Learning in a Mega Project . . . . .	25
3.2.1 PCC Installation: The Productivity Learning curve . . . . .	27
3.2.2 Learning and Quality: Is it That Simple ? . . . . .	29
3.3 Conclusions and Discussion . . . . .	34
<b>4 The Integration of Learning and Construction Scheduling Tools</b>	<b>36</b>
4.1 Introduction and Background . . . . .	36
4.2 System Implementation . . . . .	40
4.2.1 Input Processing Block . . . . .	41
4.2.2 Learning Curve Calculations and Optimization . . . . .	43

4.2.3	Evaluation and Model Selection . . . . .	46
4.2.4	Scheduling and Output Display . . . . .	46
4.3	System Verification and Validation . . . . .	47
4.3.1	Model Robustness . . . . .	54
4.4	Conclusions and Future Research Directions . . . . .	59
<b>5</b>	<b>Conclusions</b>	<b>61</b>
	<b>Bibliography</b>	<b>65</b>



# List of Figures

2.1	Hypothetical Learning Curve based on Thomas et al. (1986) . . . . .	7
2.2	Learning Models . . . . .	8
3.1	Project Organizational Chart . . . . .	26
3.2	Cumulative Average PCC Production . . . . .	28
3.3	Cumulative Average Rejection Rates for Civil Materials . . . . .	30
3.4	Rejection Causes for Trench Excavations . . . . .	33
4.1	The Architecture of the Developed System . . . . .	41
4.2	Input Windows for the Tool . . . . .	43
4.3	Tool Output Window . . . . .	47
4.4	Message Box when a small data set is introduced) . . . . .	50
4.5	Case Study Data Represented in Cumulative Average (a) and Unit Data Formats (b) . . . . .	51
4.6	Errors Resulting from Introducing Anomalies to the Data of Unit 1 . . .	55
4.7	Errors Resulting From Introducing An Anomaly to the Data of Case Study 2 . . . . .	56
4.8	Errors Resulting From Introducing An Anomaly to the Data of Case Studies 3 and 4 . . . . .	57
4.9	Errors Resulting from Introducing Anomalies to the Data of Unit 5 . . .	58

# List of Tables

2.1	Parameters for models derived from optimization over the first half of each respective dataset. . . . .	18
2.2	Summary of fit and predictability of the models across the four cases.	19
3.1	Rejected Trench Excavation Inspections . . . . .	32
4.1	Summary of the Project Characteristics . . . . .	48
4.2	Summary of Forecast Testing Results . . . . .	49
4.3	Summary of Selected Learning Curve Models . . . . .	53

# Chapter 1

## Introduction

Labor productivity has a significant impact on the success and the overall efficiency of construction projects, where the labor associated costs could represent as much as 50% of the overall project cost (Kazaz et al., 2008). Labor costs also represent the costs that entail the highest uncertainty in construction projects. Labor costs are naturally an indicator of the worker's productivity, which is logically expected to vary significantly in an industry with a highly transient workforce. The construction industry is also a main contributor to national economies worldwide. In the U.S., this industry represents about 14% of the gross national product (Thieblot, 2002), in Qatar the industry adds \$7.7 Billion to the overall GDP. Whereas in Europe the industry is responsible for 7% of the employment (Proverbs et al., 1998). The considerable economic importance of the construction industry, coupled with effect of labor on the productivity of the industry, warrants the analysis of the factors pertaining to labor productivity and the development of methods to improve this productivity.

This study examines a specific attribute of labor productivity in the construction industry, namely worker learning and experience. Tasks performed by workers in con-

struction projects are typically repetitive in nature, and are thus expected to be performed multiple times throughout and across construction projects. Regardless, of the project nature a steel fixer will be tying rebars and carpenters will be working at assembling formwork at some point in the project. The findings of the literature suggest, that the performance of workers is expected to improve with repetition (Thomas et al., 1986). Worker learning impacts the time required by the worker to get integrated into the construction site and thus be familiar with the project, with his supervisors and the assigned tasks. This is particularly important for an industry with a transient workforce, where the workforce is dynamic across and projects and therefore the integration process will be repeated throughout the project lifetime. In order to properly apply the concepts of learning to productivity improvement and labor management they must be quantified and transformed into measurable quantities. This is done by using mathematical models called learning models, that are used to numerically measure learning and the improvement in productivity with experience or repetition.

Since their introduction into the construction industry in 1965, learning curve concepts have been applied throughout the various stages of construction projects starting with bidding, planning, design, construction and even claims. Learning curve concepts have also been applied to numerous construction activities such as rebar fixing (Jarkas, 2010), floating caisson construction (Panas and Pantouvakis, 2013), and prestressed concrete piles (Hinze and Olbina, 2009).

While there is abundant literature on the relationship between learning and productivity, the relationship between learning and quality has received less attention. Yet, other studies have focused on the effects of the learning curve on another dimension for the success of construction projects, which is quality of the built product. The quality of the worker's work, is expected to improve with his/her experience. With the significant costs associated with quality problems and bearing in mind the poor quality

culture in the construction industry, addressing the relationship between quality and learning in the construction industry could offer a significant contribution to this industry.

Despite the promising applications of the learning curve theory in the construction industry, its usage by industry practitioners is still rather limited. This can be attributed to the lack of consensus on a learning curve model. Multiple learning curve models exist in the literature and data originating from different countries have been fitted by different learning curve models. An agreement on a data representation format for usage with learning models is also absent from the literature. These factors had a negative impact on the attractiveness of the learning curve for construction management practitioners.

The learning attributes of the construction industry are to be carefully examined if the learning theory is to be used by industry practitioners in this industry. Hence, a few research questions arise: Is there a universal learning curve model that can be applied to all construction activities across different geographic locations? If not, what are the learning curve models to be used in the construction industry? Is there an actual correlation between quality and learning in the construction industry ?

The main goal of this research is to study the applicability of the learning curve in the construction industry, and consequently develop a scheduling tool that integrates the learning curve concept into common scheduling and planning software used in this industry. In particular the objectives are: (1) sketch the learning process in the construction industry in order to evaluate the applicability of various learning curve models and propose a new learning curve model,(2) study the relationship between learning and quality in the construction industry; and (3) propose an automated tool that allows the integration of learning curve concepts into common planning and scheduling tools used in the industry.

## **Chapter 2**

# **Learning Curves and Productivity in the Construction Industry: A Review and a Proposed Model**

### **2.1 Background and Literature Review**

The learning curve theory has found applications in the realms of multiple industries such as manufacturing (Argote and Epple, 1990): the automotive industry (Baloff, 1971), the chemical industry (Lieberman, 1984), photovoltaic production (Nemet, 2006), and semiconductors (Cook, 1991; Gruber, 1992, 1994, 1996, 1998; Chung, 2001). Recently, the learning curve has also found a place in the service sector, where the effect of learning appears in filing (Sturm, 1999), banking (Chambers and Johnston, 2000), and software installation (Saraswat and Gorgone, 1990). As observed in the fields above, the learning curve theory can be used to support resource management and planning. These benefits can be extended to the construction industry, where the effects of learning can be employed to improve the productivity of construction

sites (Rosenbaum et al., 2012).

The earliest reference to the learning curve theory in the construction industry is the 1965 Economic Commission for Europe report titled *Effect of Repetition on Building Operations and Processes on Site*. This seminal report extended the learning effect from the manufacturing industry into the construction industry, while noting that the usefulness of the learning effect in the construction industry is due to the repetitive nature of labor intensive tasks. After that report, the applicability of the learning curve theory has flourished in the construction industry, where it was applied across the various organizational levels, project phases, and site activities.

This chapter brings forth from the literature a series of learning curve models and evaluates them through the lens of the construction industry. This section examines the learning curve literature to explain the learning process. This section also presents an analysis of the applicability of the various learning curve models in the construction industry. We then build on the pros and cons of these models to generate a model that provides both good fit and predictability. The last section provides an empirical analysis of the various learning curve models to demonstrate their usefulness in the construction in terms in terms of predictive capability and fitness. This chapter concludes with a summary of the findings and recommends an agenda for future research.

### **2.1.1 The Learning Process**

If a new worker has freshly joined a site team, s/he will require a certain time window to perform at his/her full capacity. Learning in the construction industry is therefore a gradual process and is not instantaneous. The learning process is divided into two major stages: the operational learning stage and the routine acquiring stage

(Thomas et al., 1986; Gottlieb and Haugbølle, 2010). During the operational learning stage the worker gets introduced to his/her task and acquires the basic skills associated with it. For example, it is during this stage that an electrician knows that s/he should tie embedded electrical conduits to the rebars in order to fix them while pouring concrete and it is during this phase that s/he learns that styrofoam is placed inside electrical boxes in order to prevent concrete pouring. It is not until the second stage that the performance of the worker is boosted and s/he discovers shortcuts for performing the task. During this stage the electrician knows the optimal spacing of the metal ties and discovers shortcuts for tying them (e.g. faster and safer knots, the metal wire duty that simplifies the work etc... ). Yet these two stages are not entirely separate and they might be interfering to a certain extent; a worker might discover certain shortcuts to performing his task while being introducing to it.

Since the learning curve is a mathematical reflection of the learning process, it is natural for the learning curve to also be divided into two stages as it is shown in Figure 2.1. The first is the reduction in time due to learning, and the second is the “plateau” phase where no further improvement is observed. This is where the worker becomes an expert at the assigned task and reaches the minimum time possible to achieve that task. The learning curve has two points that are of particular interest: The *startup point* and the *standard production point*. The startup point, represents the time required to complete the first cycle. After this point the learning curve becomes steeper until it reaches the standard production point where maximum performance has been reached (Hinze and Olbina, 2009). It is the region located between these two points where maximum performance improvement happens. After the standard production point is reached, no further improvement is witnessed and performance reaches a steady state.



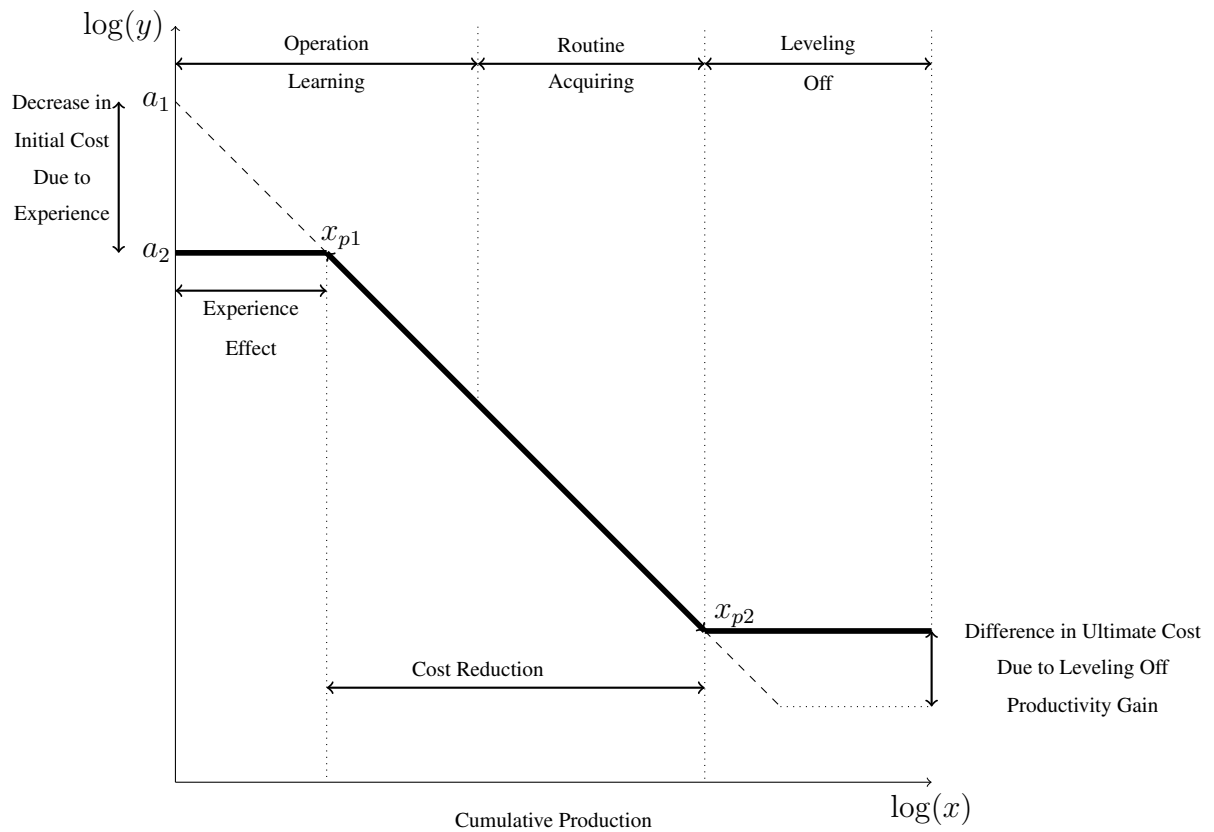


Figure 2.1: Hypothetical Learning Curve based on Thomas et al. (1986)

## 2.1.2 Learning Curve Models

Based on this two point framework for the learning process, multiple learning curve models were presented in the literature. Figure 2.2 highlights the main classes of learning models and presents the mathematical differences between them; the boxes in dark gray are the models reviewed in this study. This selection was based on our desire to highlight models of practical use within the field of construction – both multivariate and stochastic models, while useful, are not practical as they require significant data and expertise to calibrate properly.

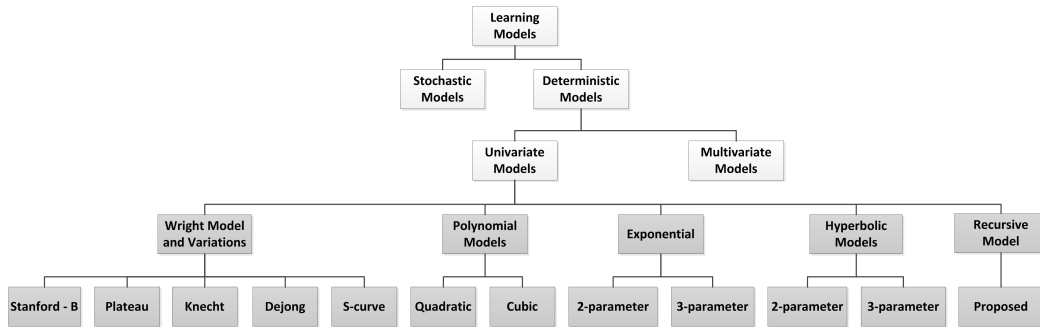


Figure 2.2: Learning Models

## The Wright Model

The oldest learning curve model is the *Wright Model*,  $y = Ax^{-n}$  (Anzanello and Fogliatto, 2011). In this model,  $y$  is the cumulative average, cost per unit,  $A$  is the cost of the first unit,  $x$  is the repetition or repetition cycle number, and  $n$ , a value between zero and one, is the slope of the logarithmic curve (Thomas et al., 1986). The learning rate,  $L$ , can be derived from the slope of the logarithmic form by using  $L = 2^{-n}$ . Therefore, the higher the slope the lower the learning rate (Thomas et al., 1986). Accordingly, when the learning rate is 100%, the parameter  $n$  which represents the slope of the curve would be zero and no further learning could occur.

This learning rate information is of particular interest to project managers and planners, since they can be used to benchmark performance. Learning rates for various trades, that can be used as a starting point or benchmark, are listed in Hijazi et al. (1992) and Gottlieb and Haugbølle (2010). The Wright Model remains the most widely used learning curve model due to its simplicity (Baloff, 1971; Globerson and Gold, 1997) and its ability to provide acceptable precision while having a simple mathematical structure (Vits and Gelders, 2002).

### **Stanford-B Model**

One of the pitfalls of the Wright model despite its popularity and applicability in the construction industry, is that it ignores the worker's previous experience. In order to overcome this shortcoming the United States' Department of Defense developed the *Stanford-B Model*,  $y = A(x + B)^{-n}$ , where  $B$  represents the number of experience units and shift the curve downwards (Badiru, 1992; Nembhard and Osothsilp, 2002). If  $B = 0$ , then this model reduces to the basic Wright Model (Badiru, 1992; Gottlieb and Haugbølle, 2010).

### **Plateau Model**

Both the Stanford and Wright model, assume that the perfect performance could be reached and tasks could be completed instantaneously or with zero time, however this is not realistic and violates the two point understanding of the learning process. The *Plateau Model*, proposed by Baloff (1971),  $y = C + Ax^{-n}$  solves this problem through the use of  $C$ , representing the steady state performance of the worker; all other terms are as introduced previously (Li and Rajagopalan, 1997; Anzanello and Fogliatto, 2011).

### **DeJong Model**

The main factor that leads to the asymptotic behaviour of the learning process after a significant number of cycles, is the use of technology and mechanization. It is therefore essential to consider mechanization in any learning curve model. For example, the performance of workers pulling cables manually is only limited by the speed of their hands. Whereas, when pulling using a motor, the threshold is enforced by the speed of the shaft of the motor. Therefore, when the operation is less mechanized and highly

manual the time or cost is more prone to decrease with learning (Kara and Kayis, 2005). The *DeJong Model*,  $y = A[M + (1 - M)x^{-n}]$ , developed in 1957 includes an incompressibility factor,  $M$ , that ranges from 0 to 1 (Hijazi et al., 1992). Thus, if the process is completely mechanized,  $M = 1$ , no improvement will arise with more repetitions (Badiru, 1992).

### **S-curve Model**

The *S-curve model*,  $y = A[M + (1 - M)(x + B)^n]$ , developed following World War Two combines both mechanization and experience (Badiru, 1992). In this model,  $M$  is the incompressibility or mechanization factor and  $B$  represents the acceptable units of previous experience. Zhang et al. (2014) developed a similar variation of the S-Curve Model which accommodates for the effects of experience, steady state labor productivity and mechanization. This model, termed the *Improved Learning Curve Model*, is represented by  $y = AM + C_0 + [A(1 - M) - C_0](x + B)^{-n}$ , where  $C_0$  is the standard time needed to complete the product under optimal conditions with perfect labor; all other terms are as before.

### **Other Wright Variants**

The literature features several other less cited variations of the Wright Model such as the Levy and Knecht Models. The former model is  $y = [1/\beta - (1/\beta - x^b/A)k^{-kx}]^{-1}$ , where  $\beta$  is a task defined coefficient and  $k$  is the performance of the labor in steady state (Levy, 1965). The latter model of Knecht (1974),  $y = (Ax^{b+1})/(b + 1)$ , was built for long production runs over which the learning rate changes. This model was never tried in the construction industry, however the construction industry has cases where similar items are constructed in large numbers, such as the installation of air-

field ground lighting at airports.

### **Polynomial Models**

Most planners and estimation engineers assume flat learning rates and calculate costs. Of course, learning is rarely so steady. As such, the Polynomial Models – quadratic and cubic – may exhibit better fit. Both the *Quadratic Model*,  $y = A + \beta_0x + \beta_1x^2$ , where  $A$  is the cost of the first unit,  $\beta_0$  is the initial slope and  $\beta_1$  is the quadratic factor (Everett and Farghal, 1994) and the *Cubic Model*,  $y = A + \beta_0x + \beta_1x^2 + \beta_2x^3$ , where  $\beta_2$  is the cubic factor, require one to assume or calculate the coefficients. Unlike the learning rate and the cost of the first unit, these parameters have no direct practical meaning and thus may be difficult to estimate or justify. Moreover, these models do not have any limiting parameters allowing for negative estimates of the costs. Therefore, they are unsuitable for using past data to extrapolate and predict future performance as is necessary when preparing bids in construction.

### **Exponential Models**

The exponential model was recommended by the Norwegian Building Research Institute as a means to improve predictive capabilities of the model (United Nations. Economic Commission for Europe. Committee on Housing and Planning, 1965). The most basic *Exponential Model* is that proposed by Knecht (1974),  $y = Ax^{-n}e^{cx}$ , where  $c$  is a constant. Just as the Wright model has variations, so too does the exponential model. The *Three-Parameter Exponential Model*,  $y = [k(1 - e^{-(x+p)/r})]^{-1}$ , includes  $k$  as a maximum performance parameter expressed as the number of units per operation time,  $p$  as the previous experience parameter expressed in units of time, and  $r$  as the learning rate given in units of time;  $x$  in this model is the number of units of operation

time. By taking the inverse on the right hand side,  $y$  becomes the cumulative average time to complete a unit after the passage of  $x$  units of time. The *Two-Parameter Exponential Model* is identical to the three-parameter model save for the exclusion of the term  $p$ . The value of the exponential models is their ability to estimate and forecast data over long production runs where the bound on learning is encountered. The limitation of the exponential models is that they are best applied to simple tasks such as fixing steel, manual excavation, cable pulling, installing wiring devices and backfilling (Mazur and Hastie, 1978).

### **Hyperbolic Models**

A final set of models, the hyperbolic models, was originally designed to capture the effect of learning within compound measures of performance (Wong et al., 2007). These models have since been adopted to capture the number of units that can be produced within  $x$  units of time (Anzanello and Fogliatto, 2011). In this paper, the models are manipulated further, by taking the inverse, to represent the cumulative average time required to produce a unit after  $x$  units of time have already been invested. As such the *Two-Parameter Hyperbolic Model*,  $y = [k(x/(x + r))]^{-1}$ , and the *Three-Parameter Hyperbolic Model*,  $y = [k((x + p)/(x + p + r))]^{-1}$ , allow for the inclusion of a maximum performance parameter,  $k$ , a measure of previous experience,  $p$ , and the learning rate,  $r$ . The hyperbolic model can be used for novel and complex tasks (Anzanello and Fogliatto, 2011). Moreover, by adjusting the learning rate appropriately, one can estimate costs after long breaks or model the performance of the workers during crash periods when they are fatigued (Uzumeri and Nembhard, 1998).

All of the models presented here assume that learning is measured as a cumulative average time to complete a task after a given number of repetitions or amount of time

(as in the exponential and hyperbolic models) and that a decrease in that cumulative average time depends on only one variable – the number of repetitions. However, task time reduction within the construction industry likely depends on multiple site related factors. Unlike manufacturing, the layout in the location where a specific task is performed can easily vary from cycle to cycle despite the repetitive nature of the task itself.

### 2.1.3 Learning and Forgetting

While a construction manager may successfully model learning while planning their project, the results may be suspect if the manager forgets to include the impact of holidays. It is quite common in the construction industry that workers' productivity drops after holidays. The main reason behind this is that the workers forget their skills and need time to regain them again (Jaber and Guiffrida, 2008). Skill level depends on practice, and thus the performance of workers will drop after interruptions or breaks. The amount of information will keep on decaying until there is nothing more to forget (Anzanello and Fogliatto, 2011). Based on theories that forgetting and learning happen in similar manners, Globerson and Gold (1997); Jaber and Guiffrida (2004); Bailey and McIntyre (1997) used log-linear models to represent the forgetting process and Nembhard and Osothsilp (2002) used a hyperbolic model to represent forgetting.

Combining both learning and forgetting Lam et al. (2001) integrated these opposing actions with the Line-of-Balance technique to develop a new forgetting model specifically for the construction industry:  $F' = F - (F - F(K^{1-n} - (K - 1)^{1-n}))((a * H + 1)e^{-aH})$ . In this model,  $F'$  is the time for the first unit after interruption (eg. installing the first pipe after the vacation),  $F$  is the time for the first unit before inter-

ruption,  $a$  is the forgetting coefficient,  $H$  is the interruption duration,  $n$  is the learning rate, and  $K$  is the number of units being completed by a specific gang.

One of the important benefits of forgetting models is that they help in developing accurate schedules that model the worker's performance after long breaks such as Christmas and Easter holidays or Adha and Futr holidays in the Middle East and North Africa region (Bailey and McIntyre, 1997). In fact, courts have ruled that damages caused by interruptions are acceptable reasons for claims (Lam et al., 2001; Hinze and Olbina, 2009). Forgetting models can also be used to determine lot sizes and decide on inventory levels (Salameh et al., 1993; Jaber and Bonney, 1996; Jaber et al., 2009; Alamri and Balkhi, 2007). It is also important to determine the time when complete forgetting occurs (Jaber and Kher, 2004).

By analyzing the impact of the forgetting effect the contractor can plan his/her activities and the delivery of the materials to site in order to minimize the undesirable results of forgetting (Lam et al., 2001). The contractor might even resort to Just-in-Time delivery in order to decrease the idle time caused by site congestion due to stocks of inventory stored on site (Lam et al., 2001). Site congestion is a factor that impacts labor productivity (Tucker, 1986 and Thomas, 1987) and thus accommodating for the forgetting effect might also help in increasing labor productivity. Finally, it is worth mentioning that the losses incurred due to the forgetting effect may offer an acceptable basis for construction claims in countries such as Hong-Kong (Lam et al., 2001).

## **2.2 The Proposed Model**

From this review of learning models, we can see that experience alone is not the primary driver behind the decrease in task time. Several external factors such as the



amount of mechanization and the prevalence of interruptions in the project also impact learning. Furthermore, to be effective as a planning tool, the model should be able to capture site conditions. One strategy to capture site-specific information in the model is through the use of recursion. The need for a recursive model was noted by Adler and Clark (1991) who recognized that experience or repetition alone was not enough to fully explain gains and losses in productivity among workers. By including a term relating the time required for the last or previous item to that of the current or next item, we can capture intrinsic changes in the learning process.

Given these considerations, we propose a recursive model represented by  $y_n = y_{n-1}2^{-(r_{n-1}+\varepsilon)} + b_0(y_{n-1}H)^{-a} + M$ , where  $y_n$  is the cumulative average cost or time required to produce the  $n$ th unit,  $y_{n-1}$  is the cumulative average cost of the previous unit or the last unit before an interruption (if applicable),  $r_{n-1}$  is the learning rate associated with the  $n - 1^{th}$  item,  $\varepsilon$  is an additive factor that updates the learning rate between repetitions and prevents the productivity from reaching infinity,  $b_0$  is a binary parameter that allows for the exclusion (0) or inclusion (1) of the experience gained prior to an interruption,  $H$  is the length of the interruption,  $a$  is a forgetting factor and,  $M$  captures the steady state performance, likely influenced by mechanization.

### 2.3 Empirical Evaluation of Learning Models

The majority of the existing literature focuses on analyzing the goodness of fit of models rather than their ability to predict future performance (Farghal and Everett, 1997). The most widely accepted metric for the goodness of fit is Pearson's coefficient of determination,  $R^2$  (Thomas et al., 1986). However, when using the learning curve in planning activities it is equally important to analyze a model's capability to predict

future performance. The works of Everett and Farghal (1997) and Farghal and Everett (1997) are among the major contributors in this domain. The method suggested by these works requires splitting the existing set of data exactly in the middle and then applying the fitted equation to predict the future values in the second set of data. The predictability of the model is then evaluated as an absolute percentage difference between the predicted and actual costs.

In order to study the capabilities of the models presented here, in terms of both fit and predictive capability, we use benchmark data from four different projects cited in the literature. The first project requires estimating the effect of learning in the construction of each floor in a 40 floor building in China (Zhang et al., 2014). The second project models learning across 19 cycles associated with the construction of tunnel formwork (Farghal and Everett, 1997). The third project studies learning across the cementing of 10 floors within a comparatively smaller housing project in Poland (United Nations. Economic Commission for Europe. Committee on Housing and Planning, 1965). The fourth project examines learning across 26 cycles associated with installing formwork for floors in an office building (Jarkas and Horner, 2011). These projects were selected given their representative nature in terms of size, type, and era of construction; with older projects representing cases with potentially less automation.

The model parameters could be obtained either by an optimization process that minimizes the least sum of squares or by using expert opinion. In this paper, the relevant parameters for all models were derived via an optimization process over the first half of the data. The specific models resulting from the optimization of the parameters are shown in Table 2.1. Only one representative model was studied from each of

the categories: the Wright model was selected from the Wright model and variations, the quadratic model was selected from the polynomial models, the basic Exponential Model from the exponential models, the Two-Parameter Hyperbolic Model from the hyperbolic models, and the proposed model as the only recursive model. The models excluded include those requiring the estimation of experience, mechanization or performance bounds as reasonable estimates for these terms were unknowable from the data provided in the case studies.

The results of both fitting the models to the first half of the data and using the fitted models to predict the second half of the data are summarized in Table 2.2. The fit is stated as the  $R^2$  for which a value closer to 1 represents a better fit; the predictability was measured using the Mean Absolute Percent Error (MAPE) for which a value closer to 0% indicates a better result. Table 2.2 also provides the p-value for a paired t-test between the model predicted cumulative average times per unit and the actual observed cumulative average times. In these tests, the null hypothesis is that the difference between the predicted and actual data is zero; a p-value greater than 0.05 supports this null hypothesis.

The results in Table 2.2 indicate that the proposed model exhibits a level of fit comparable to all other models. This is to be expected as the parameters were set with the goal of optimizing the fit across the first half of the data in all cases and for all models. The ability of the proposed model to predict the values in the second half of the data is superior in the first two cases, moderately worse in the third case, and comparable in the fourth case. Overall, however, when performing a paired t-test between the actual and the predicted values, the proposed model is superior in three of the four cases (Cases 1, 2, and 4). Case 3 presents an interesting finding as all models confirm the null hypothesis of no difference between the observed and predicted values. Furthermore, the significance of the results in Case 3 indicate that the Exponential model is superior

Table 2.1: Parameters for models derived from optimization over the first half of each respective dataset.

Model	Case 1, Parameters	Case 2, Parameters	Case 3, Parameters	Case 4, Parameters
<b>Wright</b> , $y = Ax^{-n}$	$A = 10.95$ , $n = 0.17$	$A = 26.74$ , $n = 0.39$	$A = 100.31$ , $n = 0.14$	$A = 221.91$ , $n = 0.0074$
<b>Quadratic</b> , $y = A + \beta_0x + \beta_1x^2$	$A = 10.71$ , $\beta_0 = -0.48$ , $\beta_1 = 0.01$	$A = 29.25$ , $\beta_0 = -4.35$ , $\beta_1 = 0.26$	$A = 113.67$ , $\beta_0 = -14.46$ , $\beta_1 = 1.62$	$A = 221.90$ , $\beta_0 = -0.52$ , $\beta_1 = 0.01$
<b>Exponential</b> , $y = Ax^{-n} \exp^{cx}$	$A = 10.99$ , $n = 0.17$ , $c = 0.001$	$A = 26.63$ , $n = 0.43$ , $c = 0.013$	$A = 99.10$ , $n = 0.18$ , $c = 0.016$	$A = 221.91$ , $n = 0.01$ , $c = 0.00$
<b>2-Parameter Hyperbolic</b> , $y = [k(x/x + r)]^{-1}$	$k = 0.15$ , $r = 7.89$	$k = 0.16$ , $r = 88.81$	$k = 0.01$ , $r = 33.84$	$k = 0.005$ , $r = 3.66$
<b>Proposed</b> , $y_n = y_{n-1}2^{-(r_{n-1}+\varepsilon)} + b_0(y_{n-1}H)^{-a} + M$	$y_0 = 10.97$ , $r_0 = 1.55$ , $\varepsilon = 0.15$ , $b_0 = 0$ , $M = 6.40$	$y_0 = 27.00$ , $r_0 = 1.05$ , $\varepsilon = 0.12$ , $b_0 = 0$ , $M = 8.74$	$y_0 = 100.63$ , $r_0 = 0.83$ , $\varepsilon = 0.00$ , $b_0 = 0$ , $M = 34.25$	$y_0 = 221.43$ , $r_0 = 0.12$ , $\varepsilon = 0.00$ , $b_0 = 0$ , $M = 16.94$

to the others. Not only does this confirm the findings in the article from which this case originates, but it also highlights the models' capabilities on small projects. Case 3 had only 10 data points – five used for fitting and five used for prediction. The real value of the proposed model is in application to large projects where the learning rate is likely to exhibit changes over time as in Case 1 with 40 floors.

## 2.4 Summary and Future Research Directions

Over the last four decades, the learning curve concept has gained popularity in the construction industry. One reason behind this increased popularity lies an industry trend towards cost control as a reaction to the steady increase in labor and construction material costs. As such contractors are developing schedules, checking progress and gathering data about site productivity on a regular basis. These data can be used to

Table 2.2: Summary of fit and predictability of the models across the four cases.

Fit, $R^2$	Wright	Quadratic	Exponential	2-Parameter Hyperbolic	Proposed
Case 1	0.998	0.954	0.998	0.898	0.999
Case 2	0.998	0.957	0.9995	0.999	0.997
Case 3	0.821	0.821	0.831	0.725	0.863
Case 4	0.578	0.651	0.578	0.323	0.652
Predictability, MAPE	Wright	Quadratic	Exponential	2-Parameter Hyperbolic	Proposed
Case 1	4.56%	58.44%	2.98%	7.53%	0.57%
Case 2	9.59%	146.4%	2.80%	4.47%	0.55%
Case 3	2.23%	36.54%	2.08%	3.48%	2.99%
Case 4	0.54%	0.41%	0.54%	0.94%	0.51%
Paired T-Test, p-value; H0: $\mu_{x_1-x_2} = 0$	Wright	Quadratic	Exponential	2-Parameter Hyperbolic	Proposed
Case 1, n = 40	0.00	0.00	0.00	0.00	0.87
Case 2, n = 19	0.00	0.01	0.00	0.00	0.88
Case 3, n = 10	0.14	0.06	0.31	0.13	0.15
Case 4, n = 26	0.12	0.13	0.12	0.00	0.97

develop models describing cost reductions over time due to learning.

Various mathematical models have been developed for modeling learning in different industries. This study presents a summary of the Learning Models from a mathematical and practical point of view within the construction industry. This review led to the proposal of a new model that accommodates both mechanization and forgetting. The proposed model is similar to the Wright model, but through recursion places more emphasis on the time consumed by the previous unit rather than the time used to construct the first unit. Our model demonstrates less than 1% error in predicting cumulative average unit construction times in three out of the four cases studied.

## Chapter 3

# Learning Curves and Quality in Construction: A Case Study

### 3.1 Introduction and Background

The wide majority of the learning curve literature focuses on the relationship between learning and productivity. However, there are other proxies that define the success of a product. A contractor might complete a project consisting of 10 housing units within a short duration due to the effects of learning, yet all these houses could suffer from significant quality problems and are not conforming to construction standards and regulations. Thus, the productivity of the contractor is practically nil, since none of his/her works are accepted. As such, observing the relationship between quality and learning is important. This is particularly true for an industry where the poor quality culture has led to devastating effects such as cost overruns, schedule delays, rework and property loss (Larsen et al., 2015). The quality of the built product in the construction industry generally depends on two factors: *quality of the used material*, and the *workmanship* of the labors. The first originates from suppliers and is associated with

factory controlled environments. The latter depends on the characteristics of the labor pool employed on the project and on the quality control and assurance policies by the contractor.

When viewed from the perspective of the type of knowledge acquired by the employee, learning can assume one of two forms: conceptual learning and operational learning (Lapr e et al., 2000). Conceptual learning is related to the *know-why* of tasks. For example, if the labor knows that if s/he is using a vibrator then the motion will cause the concrete to spread and will eventually yield a slab with a better quality (Lapr e et al., 2000). Operational learning on the other hand is related to the know how, which is in our case how to use the vibrator or the steel fixer's knowledge of how to tie steel (Lapr e et al., 2000). Lapr e et al. (2000) argues that both types of knowledge are necessary for the success of construction projects.

Learning in an institution can also occur at different levels, namely the individual level or the organizational level. When occurring at the individual level, knowledge is acquired by a single employee and his/her performance is expected to improve according to this knowledge. It is when an electrician learns how to use a spring in order to pull cables, that s/he becomes faster at their assigned tasks. Organizational learning on the other hand is when the institution as a whole acquires knowledge and uses it to improve performance. A contracting firm specialized in infrastructure projects and that executes such projects with superior performance due to its collective experience, could be a good example of such knowledge. Learning is believed to start at the individual level and is then expected to be transferred to the higher collective organizational level (Love et al., 2015). Improving the quality of work, necessitates perfecting the workmanship and materials and therefore requires initiatives at the level of the firm or the organization. We can thus infer that the transfer of knowledge from the



individual to the organizational level is vital for reducing quality problems such as rework. Unfortunately, the temporary nature of the construction industry inhibits the transfer of knowledge and the dissemination of information among the various parties (Love et al., 2015). This is the main driver behind using indirect methods for transferring knowledge in the construction industry.

In order to avoid quality problems during the construction phase of projects, designers tend to conduct constructability reviews (Jergeas and Put, 2001). These reviews are basically a transfer of construction knowledge to design practitioners, as designers employ contracting knowledge in order to assess the feasibility and practicality of their designs. A designer with previous contracting experience could review the design of his peers for example, and recommend replacing all the openings used to link the constructed facility by sleeves, since these are easier and cheaper to execute. Another indirect usage of organizational learning in the construction industry, is the early involvement of contractors, where the knowledge acquired by the contracting firm involved, is used to develop practical designs and concepts (Mosey, 2009). When a contractor is involved at the early stages of an infrastructure project, s/he could recommend unifying the sizes and sections of utility manholes, in order to simplify the site works and foster learning. The final form of knowledge transfer mechanisms which is used in the industry is the "lessons learned workshops" held at the end of construction projects. This form of knowledge transfer is the most direct form of embedding the individual knowledge at the collective organizational level. During these workshops organizations could compare the outcome of a specific project to previous projects, in terms of the execution methods employed and problems that occurred (Schindler and Eppler, 2003). During these workshops the various parties involved in the project share their knowledge, and then these members disseminate this knowledge to other employees when they participate in the execution of future projects. A site engineer

could learn new and simpler construction methods that have been acquired from sub-contractors, and then use them in future projects. When knowledge reaches the organizational level, it can be employed to avoid pitfalls and quality problems such as rework, accidents and high waste. Love et al. (2015) have shown that, when properly employed, organizational learning can be used to prevent rework and thus reduce quality problems.

After defining the relationship between various forms of learning and quality, it is now necessary to provide means for quantifying quality in the construction industry. Quality can be evaluated according to two criteria: quality of design and conformance quality (Meirovich, 2006). Quality of design measures the alignment between the design provided by the designer or architect and the client's intentions. Quality of conformance on the other hand measures the fitness between the built product and its intended design and specifications. Normally, during the construction phase of projects, the design would have been completed and approved by the client or their representative. As such, the majority of the quality problems occurring during construction belong to the conformance category. Quality of construction projects is usually quantified according to the approval rates of inspection requests submitted to the client or their representative. These rates provide insights regarding quality problems such as rework. For example, the work of a contractor with a 80% approval is superior in terms of quality when compared to the work of a contractor with a 70% rejection rate.

In this chapter we focus on the quality issues pertaining to the workmanship, and examine the relationship between quality and learning on one hand, versus that between learning and productivity on the other hand. Unfortunately, there is no universally accepted definition for defining or measuring quality in the construction industry. The first section introduces a case study from the GCC country and examines the relationships -if any- between quality, learning, and productivity in this project. This

chapter concludes with a section that discusses the findings of the case study and provides recommendations for future research.

The main purpose of this chapter is to assess the link between learning, quality and productivity. From an operational perspective, learning is typically reduced to task repetition, however the unit of analysis for learning and its nature varies across the literature.

## **3.2 Case Study: Learning in a Mega Project**

The case study examined in this chapter involves the construction of an airport in a GCC country that is expected to be one of the largest transportation hubs in the region. Contractors started mobilizing in early 2011, however the actual construction and earth works did not start until 2012. Due to its complex nature, and its tight schedule, an unorthodox delivery method was adopted. This project was delivered using the construction management, where all the major project players such as the client, construction manager (CM), the architect/engineer (A/E) and the Main Contractor (MC) are involved in the delivery process. Figure 3.1 portrays the project organizational chart and details the various contractual and communication relationships among the multiple project parties.

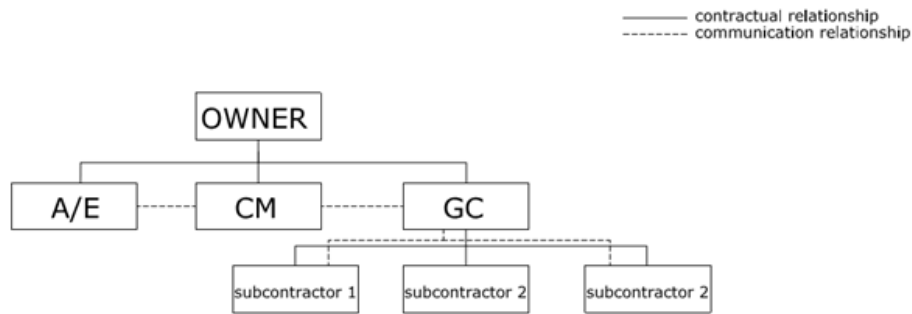


Figure 3.1: Project Organizational Chart

Due to the significant size of the project, it was divided into multiple contracts and the scope of these contracts was divided among multiple subcontractors. The nature of this project as an aviation project, dictated awarding all the aircraft related services, and special systems along with their associated construction works to contractors with previous experience in airport construction.

The preconstruction documentation for any site activity in this project is similar to regular projects, where the contractor submits for the CM's approval, the shop drawings and material qualification documents. For example, if a contractor is laying uPVC pipes in a certain zone A, all the shop drawings for this zone need to be approved before the site activities start. Moreover, the contractor should also submit the qualification of the uPVC pipe supplier before ordering the materials. Once the material is approved and ordered, the contractor again cannot proceed before the CM's representative inspects and approves the delivered material. When the site activity is completed, the CM inspects the works to check if they are conforming to standards and specifications. The CM's reply to any submitted request whether it is an inspection request or a review request can take one of the following grades:

- (I) Approved: The documents submitted or the works inspected are in conformance with the relevant standards and specifications, and are therefore approved by the

CM.

- (II) **Approved As Noted:** The documents submitted or the works inspected are approved with certain comments and require clarification or rectification by the contractor. Site works may proceed as noted.
- (III) **Revise and Resubmit:** The documents submitted or the works inspected are not approved and significant changes are required from the contractor in order to secure approval. Site works may not proceed.
- (IV) **Rejected:** The works and documents are in significant violation of the relevant specifications and standards, and are therefore rejected by the CM. These items cannot be resubmitted for reevaluation by the CM.

For the purpose of quality assessment in this study, the first two grades are counted as an approval, whereas the latter two are counted as a rejection.

### **3.2.1 PCC Installation: The Productivity Learning curve**

The concrete used in casting the final layer of the Aprons and parking stands is the portland cement concrete(PCC) type. PCC is casted in-situ in the form of panels, however prior to installing the panels, trenches are to be excavated, sub grade material must be laid and dowels are to be fixed. In order to ensure that the subcontractor responsible for PCC pouring is capable of producing good quality concrete, trial pours are held before mass production. The PCC pouring trials were held in the first quarter of 2012, where a total area of  $757.19m^2$  was completed and handed over to the CM. The actual site works for the PCC pouring activity did not start until July 2012 and continued until the early months of 2013.

The largest bulk of the work was completed during the period falling between July

and December 2012, during this period the overtime hours worked by the PCC crews were consistent. Due to the consistency in the working hours and the crew formation, the production volume can be taken as a proxy for productivity. PCC pouring is also a highly repetitive task and the only difference between two different PCC panels is in their coordinates. This fact coupled, with the crew consistency makes this case study acceptable for evaluating learning effects. During the first quarter of 2013, only the areas that were inaccessible due to site conditions were casted. Given the effect of these site conditions on productivity, only the period falling between July and December 2012 is evaluated in this study. The production data for this period is shown in Figure 3.2

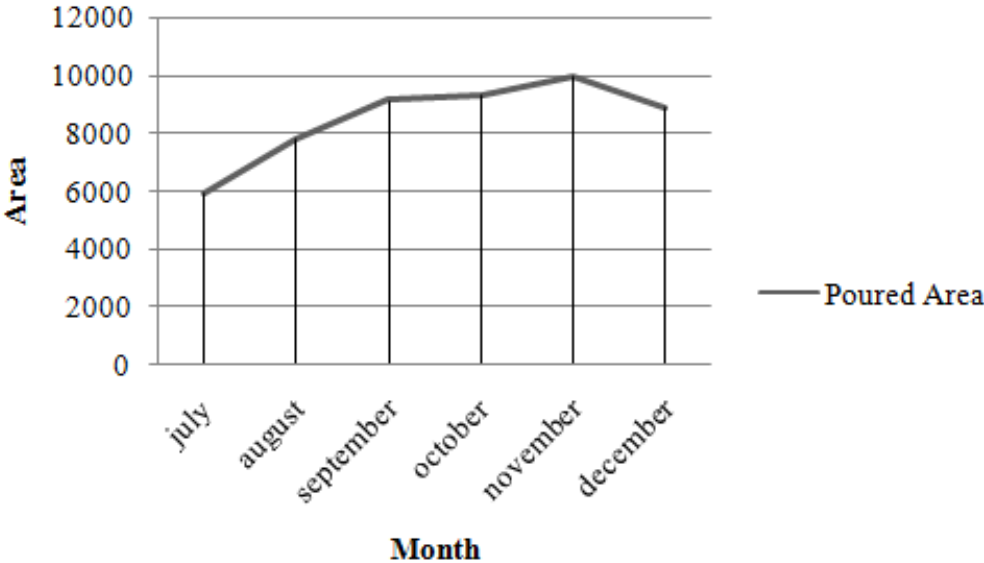


Figure 3.2: Cumulative Average PCC Production

As it can be seen from Figure 3.2, the productivity of the crew improved during the course of the study period, and the benefits of learning are clearly visible. In addition, the organization’s experience in installing PCC panels is also visible as the first month

alone witnessed a total area of 5,911.14m<sup>2</sup> being handed over to the CM. This provides an example of organizational learning as discussed in Section 3.1 above. The experience acquired by the crew in this project makes it logical for the contractor to transfer the team members to other projects, whose scope include the installation of PCC panels. Moreover, mixing the crew members with other members with less experience would also reflect good knowledge management, as this would allow for the transfer of experience and information among the members of the organization. These findings are in conformance with learning literature as productivity followed an increasing trend vis-a-vis learning in this product.

### **3.2.2 Learning and Quality: Is it That Simple ?**

The majority of the Learning literature in the construction industry focuses on the relationship between Learning and productivity. Unfortunately, and despite the prevalent quality problems in the construction industry, the amount of literature that looks into the relationship between learning and quality is minimal. Quality also has a significant impact on the progress of construction projects, since failing to secure the proper approvals, would prevent contractors from procuring materials and proceeding with site works. This would be the case in regular projects, with healthy quality assurance and control plans. However, this was not the case in the project under consideration as we discussed later in this section.

The first quality indicator to be observed, pertains to preconstruction documentation. As stated above, each subcontractor has to submit the necessary prequalification documents for the manufacturers of materials they intend to use at the project, in the form of a document called material submittal. In order to ensure alignment with previous section and since the PCC installation task falls under the Civil works trade,

only the material submittals related to civil works are considered. The materials under this category include grout, cement blocks, adhesives, anchors, bolts, nuts among many others. A significant portion of these materials is used in the PCC pouring activity, including the dowels, subgrade material and washed sand. The material submittal process started immediately after finalization of mobilization works in January 2012. Figure 3.3 below, displays the rejection rates for the civil material submittals for the period falling between January 2012 and May 2013. These rates were extracted from the logs, provided by the subcontractor executing the PCC panels. An interesting observation is that even after 18 months, more than half of the submittals were being rejected.

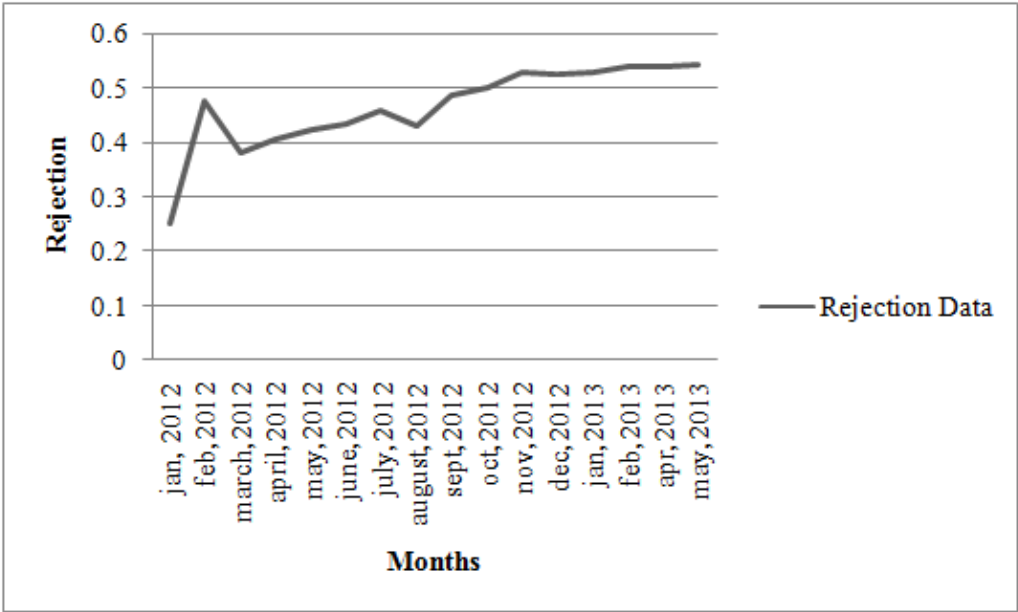


Figure 3.3: Cumulative Average Rejection Rates for Civil Materials

Although the literature states that quality improves with experience, the findings from Figure 3.3, contradict this theory. This indicates that, there are other factors that are coming into play and cancelling the positive effects of learning, which necessitates



a closer look into the review process adopted in this project. The material submittals were prepared by the contractors site teams in the GCC, however they were reviewed by the CM's technical team located in another country. This creates a significant communication problem, which inhibits the informal information exchange between the contractor and the CM. If the two teams, were located within acceptable geographic proximity, workshops could have been held between the members of the two organizations in order to assess the reasons behind these high rejection rates. Both the CM and the subcontractor, would have learned as one organization about the necessary measures to be taken in order to improve the rejection rates at the projects. These organizations would then learn know "why" the submittals are being rejected, instead of only knowing how to prepare and review a submittal.

Moreover, as the project progressed, the number of material submittals increased and the production volume spiked. This intensified the problem caused by the slow communication between the two parties located in two different countries. This would be a natural outcome, as the number of queries and clarifications requested by the CM to submittals is expected to grow with the number of these submittals. As such, increasing the production volume had a deterring effect on quality in this particular case. This indicates that, it is very important to take measures that correct or rectify quality problems that arise during the early stages of the project. Otherwise the negative effects of these problems are expected to grow with the increase in production volume. It is therefore necessary, to involve the contractors in the preparation of the quality plan for projects, so that they can understand the measures necessary to correct quality problems.

In order to ensure that the problem lies in the quality assurance and control measures adopted in the project and are not related to the activity, we decided to observe other activities. The simple site activity of excavating a trench was selected. The main

Table 3.1: Rejected Trench Excavation Inspections

Month	Total Submitted	Rejected	Reasons for Rejection		
			Incomplete Shop Drawings	Incomplete Site Works	Work Not Conforming to Quality Standars
Feb	1	1	1		
Mar	3	3	1	2	
April	2	2	2		
May	3	3	1		2
June	4	4	2	1	1
July	4	4	3	1	
Aug	1	1		1	
Sept	1	1	1		
Nov	13	13	12		1
Dec	28	27	26	1	

driver behind selecting this activity is its simplicity, which eliminates the effects of complexity on the quality of work. The activity inspected is related to the location and the leveling of the trench. Astonishingly, and among the 60 requests submitted during 2012 only one request was approved, and this did not occur until December. These rates were extracted from the subcontractor’s logs and are summarized in Table 3.1 and Figure 3.4 below. The significantly high rejection rates, requires to take a closer look at the causes of these rejections.

## Rejection Cause

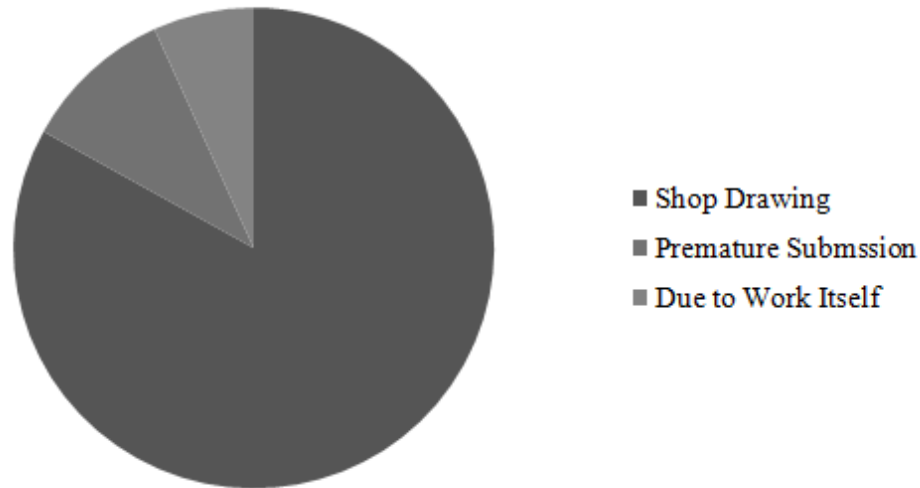


Figure 3.4: Rejection Causes for Trench Excavations

As it is shown in Figure 3.4, the major cause behind the rejections is the incomplete shop drawings. These shop drawings should have been approved during the preconstruction stages, however due to the poor implementation of quality measures, where site activities were allowed to proceed, before securing the necessary approvals. This also indicates that the reviewers of the CM were operating in silos since the whole CM's organization, did not observe the significant rejection rates and did take not any measures to reduce these high rejection rates. The second cause for rejection was the premature submission of inspection documents, where the contractor submits an inspection request before the site is ready or when it is not clean or tidy enough. This reflects a poor knowledge of the quality measures and of know how information on the contractor's side. Figure 3.4, also shows that only 7% of the rejections were actually caused by poor workmanship. This indicates that the rejection rates, are not reflective in any way of Learning as any possible benefits of learning were masked by the poor quality culture and measures at the project.

### **3.3 Conclusions and Discussion**

As it is clear from the results of the case study above, experience has played a role in improving productivity, yet it was not enough to have an impact on quality. On the contrary, we have seen that increasing the production volume could have a negative effect on quality, if it was coupled with poor communication and a poor quality culture. As such, we can notice that more than one factor comes into play when analyzing the relationship between quality and learning. Therefore any model, that attempts to represent quality variations during a construction project should be therefore multivariate in order to accommodate for the multiple factors affecting quality.

Poor communication and flawed quality procedures, could lead to significant problems throughout the project and could even hinder the benefits of learning or benefiting from using lean delivery methods such as the Construction Management delivery method. In addition, sharing knowledge among the various project parties is also necessary for rectifying quality problems and preventing them from propagating throughout the project life time. The common practice in the construction industry in the MENA region does not encourage such communications and therefore, it prevents having proper organizational learning. Studying the relationship between informal communication and organizational learning could be interesting topic for future research. The project was also divided into multiple scopes, awarded to different sub-contractors, this also intensified the communication problems and produced fluctuating quality patterns across the various project scopes. Weekly meetings should have been held or a lessons learned data base should have been established, in order to ensure proper transfer of information.

The quality of the site works was also assessed according to the findings of the resident engineer working at the CM organization. These findings do not represent a

real indication of the quality of work, since the inspector might reject the works due to subjective reasons or due to procedural complexities that are beyond the control of the site team. Accordingly, future research should focus on developing more objective methods for evaluating the quality of workmanship.

It is worth mentioning that this case study is under no circumstances representative of the construction industry in the MENA region. However, one of the most interesting findings of this case study is that learning is not always associated with quality improvements and in some cases where there is a poor quality culture, increasing the production volume could actually lead to more severe quality problems. Therefore, the major contribution of this work is that it portrays the complex relationship between quality and learning which necessitates developing a multivariate model that links quality to learning.

It is logical for future research to focus on the interplay between quality, learning and productivity. Specifically, it should focus on the type of learning that is typically associated with quality improvements. The methods for fostering such learning should also be questioned. Does it only require management intervention and training programs, or are certain policy changes required in terms of the quality programs applied ?

# Chapter 4

## The Integration of Learning and Construction Scheduling Tools

### 4.1 Introduction and Background

The success of any construction project relies heavily on three pillars: Time, Cost and Quality. Properly allocating resources and preparing accurate project schedules has a significant impact on these three pillars. Meeting the contractual schedule and quality requirements, entails using finances to acquire certain resources with specific productivity and a certain skill set. You can allocate ten skilled labors with experience to accomplish a certain task within two days but at double the cost of hiring ten unskilled labors who would need 15 days to complete the task with reduced quality. Preparing accurate schedules is therefore a complex task that requires meeting multiple objectives since it affects all the dimensions of project success. In addition, the tasks scheduled in the construction industry are related by complex interdependencies due to the fragmented nature of the industry. The multi objective nature of construction scheduling coupled with the complex interdependencies, might render traditional

scheduling tools such as CPM and LOB insufficient for preparing accurate schedules. The significant leaps in the fields of computation and optimization, has allowed automation to infiltrate the field of construction scheduling. The usage of a higher and more complex set of optimization models, such as Evolutionary Models, Artificial Neural Networks became more popular for scheduling and performance prediction purposes. For example, Georgy (2008), used genetic algorithm to allocate resources for repetitive construction projects. Long and Ohsato (2009), and Senouci and Al-Derham (2008) also used evolutionary algorithms to optimise schedules, while meeting time requirements and minimizing cost. Georgy et al. (2005) on the other hand, used artificial neural network for performance prediction in construction projects.

Although the studies referenced above had major contributions in improving the accuracy of construction schedules, their underlying computational methods require extensive calculations and significant mathematical knowledge. In a way these methods, represent a "black box" for industry practitioners, contractors and construction managers, who cannot analyze the implications of the output of these models. These models are also intended for use at the planning phase of the project. These models do not capture the actual performance data during the project. Construction schedules however, are highly dynamic documents that should be updated throughout the various project phases. Failing to do so, could jeopardize the success of the construction project.

Among the data that could be fed in into such tools , is the actual labor productivity which is expected to vary significantly throughout the project lifetime. It is therefore, natural to observe the main drivers behind this variability. The construction industry is known to be a labor intensive industry, where labor costs could amount to 50% of the total project cost (Kazaz et al., 2008), which indicates that the success of construction projects relies heavily on the characteristics of the labor pool used. Unfortunately,

labor is a resource that involves significant amount of uncertainty. This uncertainty could propagate and create a risk that affects the productivity of the project as a whole, which could have an impact on the management's capability of generating accurate schedules.

When planning repetitive projects, planners assume constant labor productivity, and that the rate of output produced by the workers will be uniform through the project lifetime (Al Sarraj, 1990). This assumption is not in line with the learning curve literature which states that labor productivity is expected to improve with repetition and experience. If we are using the traditional LOB method for the scheduling of trench excavation, and on the first month of work we noticed that  $1,000m^3$  were excavated, then we will assume that this rate will continue throughout the project lifetime. Therefore the scope that includes the excavation of  $10,000m^3$ , would be estimated to require ten working months. However, learning curve theory could tell us that it might require 8 months only. If the estimates were being used for bidding the company with the 8 months estimate would win the contract, given all the other bid items are similar. However, it is important to examine what models are the planners using to provide these estimates ?

The first attempt to integrate Learning Curve models, into performance prediction in the construction industry was the work of Everett and Farghal (1994). In their work, Everett and Farghal (1994) developed a methodology for using learning curves for the purpose of performance prediction, and they concluded that linear models were found to be the best predictors of future performance after examining a suite of learning curve models. The work of Everett and Farghal (1994) however, did not integrate this methodology with any common scheduling technique and it did not involve any automation. After this seminal work, many studies focused on integrating learning curve concepts into scheduling techniques. Arditi et al. (2001), integrated the log-linear



model with the Line-of-Balance techniques in order to determine the optimum crew sizes, and the start and finish times of activities. Ammar and Abdel-Maged (2012), used a similar philosophy, however, they assumed that work is continuous and target dates are met. Lutz et al. (1994) used the log-linear model for another purpose which is the simulation of learning development in construction projects. This work requires the user to have significant programming knowledge in order.

Although the studies mentioned above were all successful at integrating Learning curve concepts into construction scheduling, they were all based on a single learning curve model which is the log-linear model. The log-linear model is the most popular learning model due to its simplicity, and ease of applicability. Yet, this model might lead to erroneous forecasts under certain conditions (Hurley, 1996). The errors are caused by the tendency of the log-linear model to converge to zero as the number of repetitions is increased. In order to overcome this shortcoming, Zhang et al. (2014) proposed an alternative model, that was integrated in the Line-Of-Balance scheduling technique for the purpose of reducing crew sizes.

Despite the fact that Zhang et al. (2014) developed an improved learning curve model, the line of balance model they developed only integrated a single learning curve model. For a learning based scheduling technique to generate accurate schedules for a universal set of construction projects, a suite of learning curve models is to be considered, as different learning curve models are used for different project types.

In an effort to tackle these gaps in the literature, this paper utilizes non-linear optimization for the purpose of integrating learning curve concepts in construction scheduling. In this chapter we will present an optimization based scheduling tool that incorporates learning curve concepts. An array of learning curve models are embedded in the developed tool so that it can be used on multiple construction projects. This tool was developed using Macros in MS Excel, and it can therefore be used on per-

sonal computers which are normally available for site crews. This tool also allows the planner to generate construction schedules at various stages of the project. The first section provides an illustration of the tool's operation and the second section presents the test results for the various tests performed on the developed tool. The final section of the chapter provides conclusions and recommendations for future research.

## **4.2 System Implementation**

In order to automate the integration of learning curves into construction scheduling, a software tool was developed using Visual Basic Macro programming under MS Excel release 2007. The developed tool includes a graphical user interface, in order to facilitate its use by industry practitioners. The user can use the graphical user interface to import a schedule file, and select the number of cycles to be planned. The user can also allow the tool to select the best learning curve model, or select a preferred learning curve model. The system the operates through four functional blocks embedded in the tool as shown in Figure 4.1 :(1) Input Data Processing block, (2) LC Calculation and Optimization block, (3) Evaluation and Model Selection block, and (4) Scheduling and Output Display block.

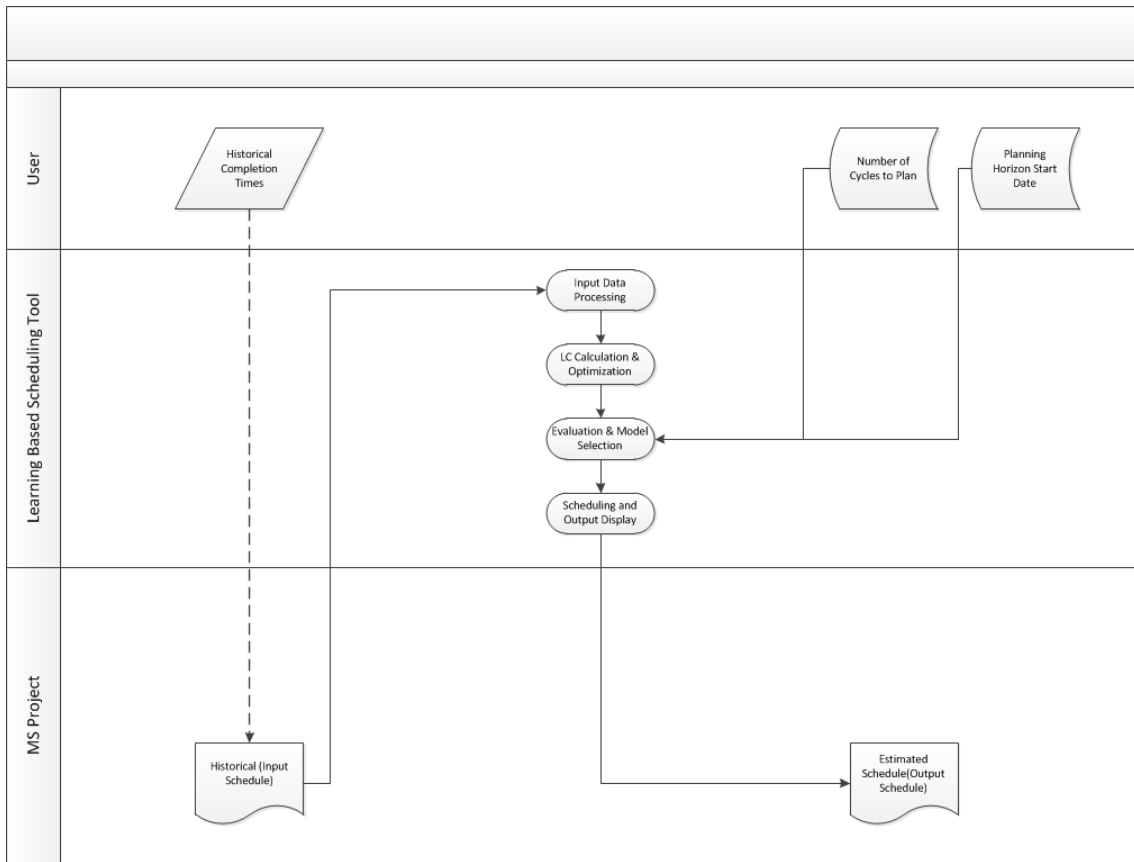
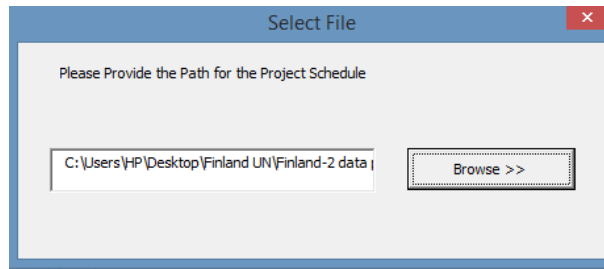


Figure 4.1: The Architecture of the Developed System

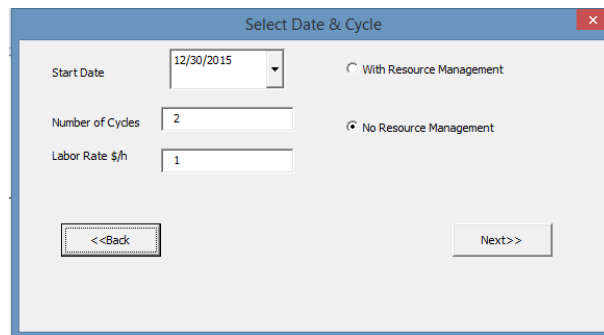
### 4.2.1 Input Processing Block

This functional block is responsible for importing the historical data from the schedule file specified by the user. The functional block accesses the schedule file and acquires the task names, their start dates, completion dates, durations and precedence relationships. The user can then either store the data in the acquired format or change it to the cumulative average format. Learning curve data is typically represented in one of two formats: (1) Cumulative Average Data, and (2) Unit Data. Using cumulative

average data works as a filter for the data by clearing the noise and emphasizing on the long term trends. The cumulative average is the mean completion time of all cycles up to and including the current one. Such data format, is recommended for preparing master schedules and thus, for planning over extended time horizons. Unit data on the other hand displays actual site progress and therefore emphasizes on short term site progress. Due to its emphasis on short term trends, this data representation format is recommended for day-to-day planning and for use by last planners. If the user decides to use the tool while the data is being represented in the unit format, then no further processing is required. However, if the cumulative average format is selected, then the cumulative average completion time for each cycle is calculated, after the actual completion times are exported. The input window used by the user to select the schedule file, and the window used for specifying the number of cycles to be planned are shown below in Figure 4.2.



(a)



(b)

Figure 4.2: Input Windows for the Tool

## 4.2.2 Learning Curve Calculations and Optimization

### Learning Curve Calculations

As stated previously in Chapter 2, multiple learning curves exist in the literature and these models vary in their nature and complexity. The model could be univariate or multivariate, and it could be deterministic or stochastic. The purpose of the models could even vary, as some models link productivity and learning, while others model the correlation between learning and productivity. The models included within this tool are: The Wright Model, The Exponential Model, The Hyperbolic Model and The Recursive Model.

The aforementioned models were selected according to their performance in the testing performed in Chapter 2 above, in addition to their applicability and suitability

for the construction industry. Although other models such as the multivariate models and stochastic models, have displayed superior fitness capabilities, yet they require significant computing power and the addition of certain add-ons, which might not be available for construction managers and planning engineers.

Once the learning models were selected, subroutines were developed to perform the necessary calculations. These subroutines start with initializing the parameters of each LC model, and calculating the associated completion times and fitness errors. Parameter initialization was done from within the code or was hard coded, since the optimization algorithm requires a starting or an initial solution. However, the parameters cannot be initialized to random values. If the initial solution is significantly distant from the optimal solution, the optimization algorithm might converge to a local optimum solution instead of the global optimum solution. As such, the parameters were initialized according to values that are recommended in the literature or according to a calibration process performed by the developers through running the optimization multiple times. In order to illustrate the parameter initialization process, let us consider that the parameters of the Wright Model ( $y = Ax^{-n}$ ), are being initialized. The initial cost parameter  $A$ , would be initialized to the completion time of the first cycle, which is inline with the definition of the parameter. The  $n$  parameter is initialized to 0.15 according to the findings of a report by the United Nations (United Nations. Economic Commission for Europe. Committee on Housing and Planning, 1965). This report suggests that the learning rate for construction activities is equal to 90%, since  $n = \log_2(\text{LearningRate})$ , initializing  $n$  to 0.15 would be reasonable. Once the learning curve estimates are obtained, point errors are then calculated and used to calculate the sum of squares for every model.

## Optimization

The obtained estimates for the model parameters are then optimized using a non-linear algorithm built in with MS Excel solver. The main driver behind selecting this solver is its commercial availability with MS Excel which would allow its usage by industry practitioners without any extensions or hardware upgrades. Like any other optimization tool, this one requires an objective function.

Naturally, the main goal of the developed tool is to forecast future completion times with minimal errors. This goal cannot be achieved directly, since future data is not yet available. Accordingly, the forecasting error will be minimized indirectly through adopting the methodology developed by Everett and Farghal (1994). The methodology is based on the assumption that the tool that provides the best fit will be the model providing the best forecasts. As such, the objective function becomes to minimize the difference between the estimates generated by the various learning curve models and the actual historical data. When represented in a mathematical form the objective function becomes:

$$\text{minimize } \sum_{i=1}^n (Estimate_{i,j} - Actual_i)^2$$

Where  $Estimate_{i,j}$  is the estimate for historical cycle  $i$  using Learning Curve Model  $j$  and that is calculated according to the mathematical equation of the model. A historical cycle is a previously completed construction cycle, e.g. floor, formwork installation and removal.  $Actual_i$  is the actual completion time for cycle  $i$ . As it is seen in the equation above, the tool aims at minimizing the Sum of Squares of the Errors and therefore improves the fitness of the various learning curve models.

### **4.2.3 Evaluation and Model Selection**

Once the optimized learning curve estimates are obtained, the model with the least sum of square error, or the model that best fits historical data is selected. This process is the final step towards achieving the goal of the system, which is generating accurate learning based schedules.

### **4.2.4 Scheduling and Output Display**

After selecting the best fitting learning curve model, it is used to generate completion times, finish and start dates of the future cycles. The number of future cycles to be planned is defined by the user within the input processing functional block. The user also feeds in the starting date of the planning horizon, which is considered to be the start date of the first cycle to be planned. While developing this functional block, we have assumed that there is no lag between two successive activities. For example, if two activities A and B are two successive activities then the completion date of activity A is considered to be the start date of activity B. Once only the necessary planning dates and completion time estimates are obtained, an MS Project File is created and the completion times are exported to it. The output window displayed by the tool, when planning is completed is shown in Figure 4.3.



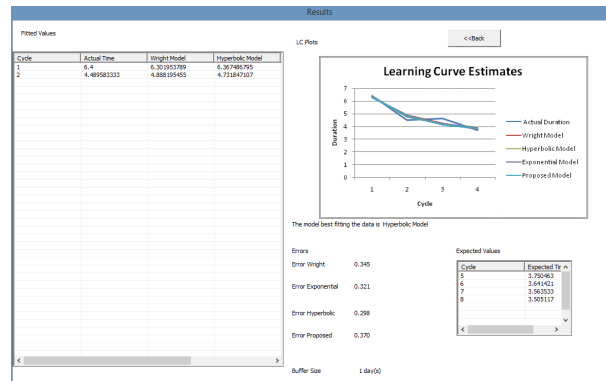


Figure 4.3: Tool Output Window

### 4.3 System Verification and Validation

In order to verify that the developed tool is meeting its intended functions, it was tested on five case studies. Table 4.1 provides a summary of the case studies used to test the tool. The number of points shown in the Table refers to the number of historical points or cycles already completed which is used to forecast the second set of the data (the future completion times). The scope of the first project includes the installation of 19 cycles of tunnel formwork in France. The second project involves the construction of eight identical housing units in Finland. The third and fourth projects evaluate learning for the construction of the first six stories of multistory buildings in Portugal. The final project involves the construction of 20 stories in a high rise Buildings in Hong Kong. As it is clear from Table 4.1 these projects originate from different geographic regions and are of different types, where some projects involve the constructs of highrise buildings while others are for tunneling works. These projects, are therefore representative of the construction industry in terms of both location and nature, and this is particularly important since previous research suggests that the best fitting learning curve model varies according to the project location and nature (Srouf

Table 4.1: Summary of the Project Characteristics

Case & Type	Source	Total Number of Cycles	Quarter	Half of the Data	Three quarters
Case 1 ,(Infrastructure)	(United Nations. Economic Commission for Europe. Committee on Housing and Planning, 1965)	19	5	10	14
Case 2, (Housing)	(United Nations. Economic Commission for Europe. Committee on Housing and Planning, 1965)	8	2	4	6
Case 3, (Residential)	(Couto and Teixeira, 2005)	6	2	3	4
Case 4, (Residential)	(Couto and Teixeira, 2005)	6	2	3	4
Case 5, (HighRise)	(Wong et al., 2007)	20	5	10	15

et al., 2015). For example, the exponential model was found to be the best fit for data originating from Norway (United Nations. Economic Commission for Europe. Committee on Housing and Planning, 1965), while the recursive model was found to be the best fit for high rise projects (Srouf et al., 2015).

The first test performed to evaluate the performance of the tool is adopted from the two-step procedure developed Everett and Farghal (1997). In order to start this procedure, the data is split into two smaller data subsets: the historical data, and future data. Historical data refers to the data subset that is used for extrapolation purposes and is therefore fed in as input for the tool. Future completion dates are the actual completion times of the rest of the data set and are compared to the estimates generated by the tool in order to evaluate its performance. Everett and Farghal (1997), split the data at its midpoint and used the first half as historical data. Such a process would only allow testing the tool at the middle of the project. In order to ensure that the tool can be used at different project stages, we decided to split the data at three points: the quarter data point, and in such a case the earlier 25% of the data is used to forecast the other 75%, the second boundary point was set at the midpoint of the data, and the final boundary point was set at the point representing three quarters of the data.

The accuracy of predicting future completion times is evaluated by observing the Mean Absolute Percent Error(MAPE), which is defined as the average of the absolute values of the differences between the forecasts and the actual completion times,

Table 4.2: Summary of Forecast Testing Results

Case	Cumulative Average-Error(%)			Unit-Error(%)		
	25%	50%	75%	25%	50%	75%
1	5.76	2.8	0.79	11.77	4.57	0.4
2	2.59	4.76	0.95	43.86	7.59	8.54
3	5.74	3.82	2.21	26.43	5.63	17.65
4	33.91	4.81	1.29	43.67	14.9	1.21
5	42.4	2.11	5.59	50.68	9.11	1.12

divided by the actual completion time. Although the sum of squares was used to evaluate the fitness of the models, the process developed by Everett and Farghal (1997), upon which the tool was developed, uses the MAPE for evaluating the accuracy of predictions. With this methodology, an error that is closer to 0%, indicates a better quality of predictions. The results of this testing procedure are shown in Table 4.2.

As it is shown in Table 4.2, the tool displays a better performance when data is represented in the cumulative average format, as it displayed superior performance with the cumulative average format for 12 out of the 15 tests. For example, when using the first half of the data as the historical data set for case 3, the MAPE when using the cumulative average format was 4.8%, while it amounted to 14.9% for the same case when representing data in the unit format. These results are expected since the cumulative average representation tends to smooth out the data, which places a higher emphasis on long term trends, and therefore trends that have existed in the historical data set are expected to continue in the future data set.

The results have also shown that the tool is more sensitive to the size of the data set when it is being used in the unit data format. For all of the five cases the error associated with using 25% of the data as a historical data and in the unit format, has exceeded 10%. This is also logical since, using unit data masks trends and introduces noise to the data. Accordingly, and in order to avoid misleading the users of the tool, we found that it would be appropriate to alert the user, if s/he is using a limited amount of data points represented in the unit data format. This alert is done via message box that is

prompted to the user, when the number of historical data points is less than 2. The Message box is shown in Figure 4.4 below:

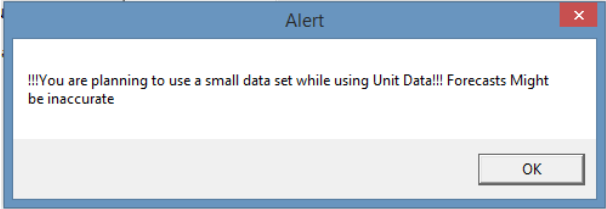
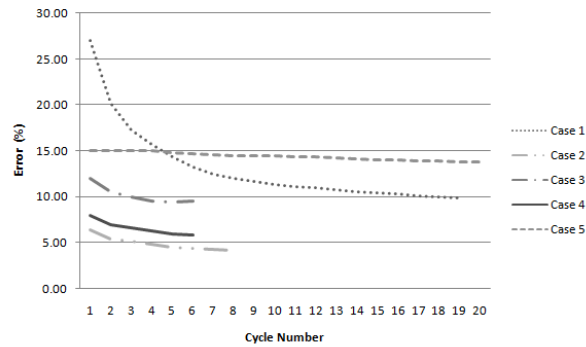
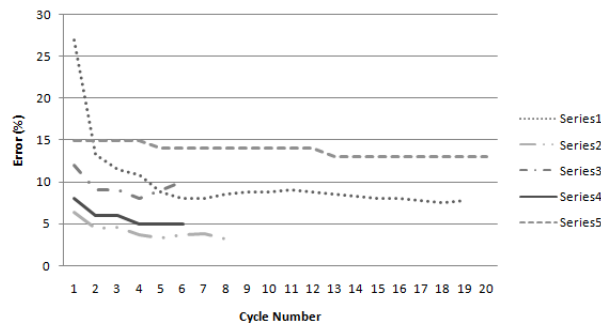


Figure 4.4: Message Box when a small data set is introduced)

In addition to being sensitive to the size of the data set, the size of the error generated could also vary according to the characteristics of the data itself. For example, when examining the results obtained for cases 4 and 5, the resulting errors were high when 25% of the data was used even when the data was represented in the cumulative average format. When observing the data itself in Figure 4.5, we can see that the trends existing in the first part quarter of the data does not persist in later parts of the data. Therefore, the tool is not recommended for use in projects where the construction methods or management methods are expected to change through the course of the project.



(a)



(b)

Figure 4.5: Case Study Data Represented in Cumulative Average (a) and Unit Data Formats (b)

It is also important to observe the learning to observe the learning curve models selected by the tool for the various cases and for the various dataset sizes. This is particularly important since it helps us ensure that the tool is not converging to a single learning curve model. Table 4.3 summarizes the various learning curve models selected by the tool for the various runs. As we can see in the table, each learning curve was selected at least once. We can also observe that the exponential model was the most widely selected model when the data is represented in the cumulative average format, when the unit data format is used, the recursive model was the most widely selected model. This is again a logical outcome, since the recursive model is expected

to be capable of handling noisy data.

Table 4.3: Summary of Selected Learning Curve Models

Case	Cumulative Average Data Representation				Unit Data Representation			
	Wright Model	Hyperbolic Model	Exponential Model	Recursive Model	Wright Model	Hyperbolic Model	Exponential Model	Recursive Model
<b>Case 1</b>								
Using 25% of the Data			x					x
Using 50 % of the Data			x					x
Using 75% of the Data			x					x
<b>Case 2</b>								
Using 25% of the Data				x				x
Using 50 % of the Data			x			x		
Using 75% of the Data	x						x	
<b>Case 3</b>								
Using 25% of the Data			x		x			
Using 50 % of the Data			x					x
Using 75% of the Data			x			x		
<b>Case 4</b>								
Using 25% of the Data				x				x
Using 50 % of the Data		x						x
Using 75% of the Data			x			x		
<b>Case 5</b>								
Using 25% of the Data				x				x
Using 50 % of the Data				x				x
Using 75% of the Data				x				x

### **4.3.1 Model Robustness**

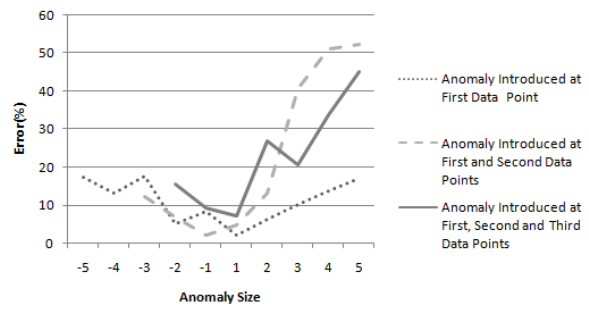
Noisy and scattered data are very common in the construction industry, and therefore there is a possibility that the historical data fed in as input into the tool could contain multiple anomalies or outliers. These outliers could be caused by faulty reporting, or by the significant uncertainty existing in the construction industry. Uncertainty is considered to be a fact of life in the construction industry and is expected to be more prevalent at the early stages of the project, when managers and site engineers and not yet very familiar with the project and the contract. Rejections and rework are expected to be more common during this stage of the project, due to improper documentation or site access problems. Therefore, it is necessary to ensure that the tool can handle these anomalies and verify its robustness when provided with noisy input data. Since outliers are common at the early stages of the project, we decided to introduce anomalies to the data at the first 3 data points. With these anomalies introduced we can perturb the data of the five case studies, and thus simulate a real case of faulty data reporting. In order to be consistent with the testing procedure and introduce anomalies in a uniform fashion, the outliers were introduced as multiples of the standard deviation of the data. Multiples of the standard deviation ranging between  $-5$  and  $+5$  were added to perturb the data for the purpose of robustness testing.

#### **Case Study 1**

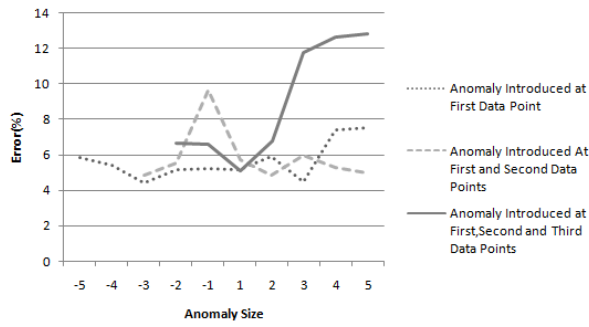
When anomalies were introduced to the first, second and third data points of the first case study, a correlation could be observed between the size of the anomaly and the resulting error in the cumulative average case. This is expected, since larger anomalies are expected to introduce higher perturbations to the data which would increase the



size of the error. The tool also displayed higher robustness when the error was only introduced to the first data point, instead of the first, second and third data points. This is also logical since fewer anomalies are expected to cause fewer errors. When representing the data in the unit format, there was no significant correlation between the size of the error and the size of the anomaly except when an anomaly was introduced to the first three data points at the same time. This was caused by the original profile of the data itself, as the first data point was already significantly distant from the mean by more than three standard deviations and thus represented an anomaly in itself. In this particular case, and surprisingly the tool performed better when the data was represented in unit format, as the original data itself contained a significant amount of noise and introducing errors did not have a significant effect on the original profile of the data. The results of the robustness testing are summarized in Figure 4.6.



(a)



(b)

Figure 4.6: Errors Resulting from Introducing Anomalies to the Data of Unit 1

## Case Study 2

Similarly, to the first case study a correlation was observed between the size of the anomaly and the size of the resulting error when the cumulative average format was used, but no visible correlation existed when using the unit data format. However, unlike the first case the tool performed better when data was represented in the cumulative average format. This can be explained by observing the size of the data set, as the data set contained only 8 data points and was not very scattered and introducing an anomaly would therefore alter the original data profile and introduce significant noise. The robustness test result for this case study are shown in Figure 4.7.

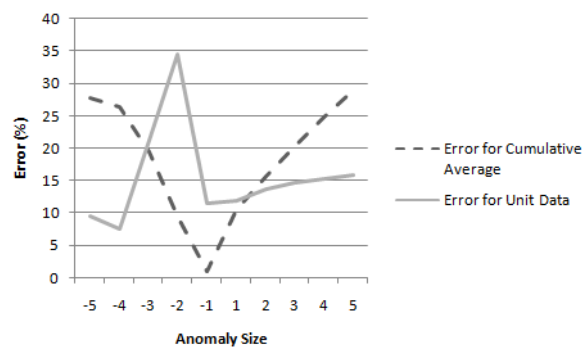


Figure 4.7: Errors Resulting From Introducing An Anomaly to the Data of Case Study 2

## Cases 3 and 4

Again the tool displayed superior robustness, when the cumulative average format was used and this could be attributed to the fact that cumulative cleans out the data and thus eliminates the effect of anomalies. For these two particular cases, the tool's performance was also superior when negative multiples of the standard deviation were introduced to the data, as compared to positive multiples. Adding, a positive multiple

to the early points of the data would create a false trend which would not persist in the later stages of the data, and would therefore lead to forecasting errors. Figure 4.8 summarizes the robustness testing results for these two case studies.

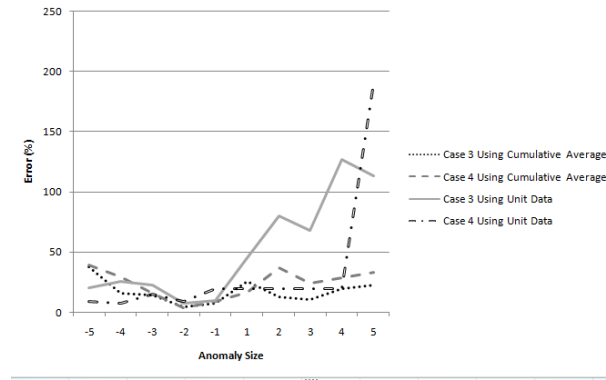
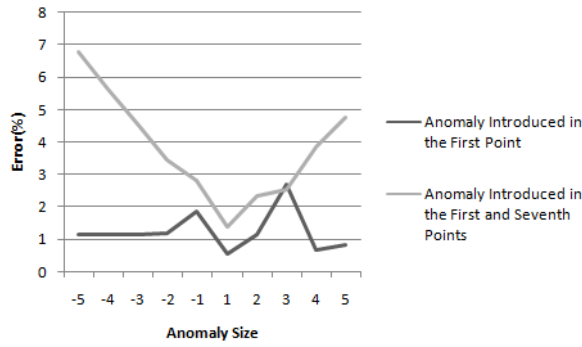


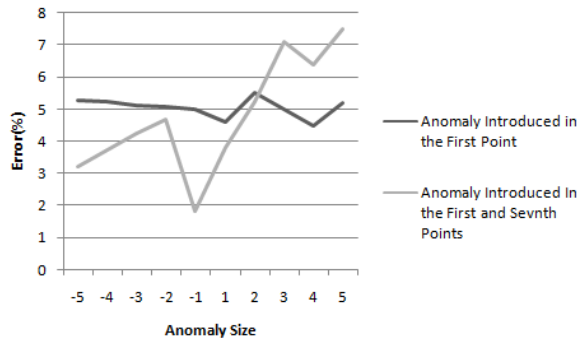
Figure 4.8: Errors Resulting From Introducing An Anomaly to the Data of Case Studies 3 and 4

### Case 5

Unlike the previous case studies, the performance of the tool was comparable when data is represented in the cumulative average and unit data formats. This result can be attributed to the size of the data set which is larger than the other cases and that is also less less scattered than the other cases. The size of the data and its scatter can also be used to explain the correlation between the size of the error and the size of the anomaly for both the cumulative average and unit data formats. Figure 4.9 illustrates the robustness testing results for this case study.



(a)



(b)

Figure 4.9: Errors Resulting from Introducing Anomalies to the Data of Unit 5

The testing results above indicate that the tool had an acceptable performance in terms of robustness, and on average the tool was more robust when the cumulative average format is used. We would recommend that the user, uses his/her expert judgement in order to decide on eliminating outlier data points, so that the accuracy of the predictions is improved.

## 4.4 Conclusions and Future Research Directions

In conclusion, the developed LC based scheduling tool was able to generate construction schedules with acceptable forecasting errors. This tool extends previous research by integrating learning curve concepts into automated construction scheduling. The tool includes a suite of learning curve models, which increases its applicability to a wide umbrella of construction projects, as different projects are fitted by different learning curve models. The tool was also built on the assumption that the learning curve model that best fits historical data, will be the learning curve model that generates the best forecasts. The tool can also be used at both the micro and macro project levels, since it allows the user to select one of data representation formats: cumulative average format and unit data format. These first format can be used at the macro level and thus used to generate master schedules, while the latter is recommended for weekly or daily schedules.

The tool was also developed with a graphical user interface, and is therefore practical and can be used by site engineers and construction managers. The inputs files required by the tool and the tool's executable version are also compatible with common off the shelf softwares, which helps in facilitating its usage by industry practitioners. When tested on project data extracted from the literature, and originating from different geographic areas, the tool was able to generate fairly accurate forecasts, even when the size of the historical data set was small. The tool's robustness was tested by introducing anomalies into the data and the tool displayed poor robustness for small construction projects and acceptable robustness for the larger ones.

Since this tool is used for generating schedules, we believe that the output can be utilized for resource allocating purposes. Despite the fact that no resource allocation functionalities, however planning engineers can use expert judgement and basic

heuristics for allocating resources for smaller projects. Yet, we still recommend that future research looks into the integration of resource allocation with LC based scheduling software. Moreover, the tool was based the assumption that the model best fitting historical data will be the model yielding the best forecasts. This assumption is built on the premise that the worker's performance depends solely on experience and repetition. Yet, recent research have shown that social connections and the management style have serious impact on the performance of employees. Therefore, we also recommend future research observes the possibility of developing a multivariate learning curve model, in order to account for these numerous factors and try and integrate it into the developed tool.

# Chapter 5

## Conclusions

This study investigated, using exhaustive literature surveys and case study research, the Learning Curve theory and its applicability in the construction industry. The relevance of the learning curve was evaluated for various success dimensions of construction projects such as quality and productivity. The r-squared coefficient was used to verify the applicability and fitness of various learning curve models to construction data for four case studies extracted from the literature and originating from various countries. The Mean Absolute Percent error on the other hand was used to evaluate the capability of various models to accurately predict completion times. The learning process in the construction industry was found to be divided into two stages, and is expected to reach a plateau. According to the literature survey conducted, the shape of the learning process was also found to be significantly affected by the worker's previous experience and by the amount of mechanization involved. No consensus also seemed to exist in the literature regarding the best learning curve model and as such, a recursive learning model was proposed, in an attempt to provide a learning curve model with superior performance.

Most importantly, the results of case study analysis, indicate that the learning curve profile varies according to the project location and characteristics. Therefore, the learning curve model that best fits data originating from Europe will not necessarily be the Learning curve model best describing productivity data for workers in East Asia. The learning curve model best describing housing projects is also not necessarily the learning curve model that provides the best sketch of the productivity profile for linear projects such as pipeline and infrastructure projects. Therefore, and despite the fact the paired t-test results indicated that the recursive model provides better predictive capability in 3 out of 4 case studies, we cannot say that a single learning curve model is the most universally superior model in terms of fitness and predictive capability. However, we can comfortably state that, the results of the case studies indicate that a certain suite of learning curve models such as the Wright Model, the Exponential Model, the Hyperbolic Model and the Recursive Model tend to provide better results when compared to other learning curve models. The findings also indicated that using an inappropriate learning curve model, could lead to significant forecasting errors as such models tend to provide negative completion estimates when the number of repetitions is significantly increased.

After establishing the relationship between learning and productivity in the construction industry, the connection between learning and quality was to be observed, especially with the endemic quality problems in the construction industry. A case study from the GCC countries was evaluated, and the findings reestablished the positive correlation between learning and productivity. However and in contrast to the relevant literature, no relationship existed between quality and learning for this particular case. This indicated that quality is a more complex project success dimension, and others factors such as the management style, project type, and employee skill are to be observed when studying the relationship between quality and learning.



The verified connection between learning and productivity was then used to develop a learning based scheduling software, that allows construction planners to generate construction schedules while taking into consideration the learning effects. The tool contains a suite of learning curve models, and can be used in multiple options and at multiple project stages. The findings of a comprehensive testing procedure conducted on the tool, indicated that it can provide estimates with acceptable accuracy, even when minimal data about the project is available. The tool was also found to be robust as it was able to handle data containing outliers.

The major contribution of this study, is that it eliminated the notion of a superior learning curve model, as the best learning curve model could vary according to multiple factors. Therefore, any study examining the learning curve in the construction industry should focus on an array of relevant learning curves and not on a single learning curve model. The study also concluded that both the cumulative and unit data representation format are fairly acceptable for the construction industry, however the selection of the format should vary according to the purpose of the use of the learning theory. The study has also contributed to the literature a new recursive learning curve model that is capable of handling outliers and noisy data. The main advantage of this model is that it accommodates for mechanization, interruption and places more emphasis on recent data points. The benefit of this model is that it provides the planner with an acceptable flexibility, as s/he can decided to include or exclude certain factors such as mechanization depending on the project characteristics. Consequently, the model is malleable and can be modified according to the project requirements.

Other benefits of this study include its findings on the complex nature between learning and quality. This complex relationship, cannot be reduced to a simple log-linear or univariate learning curve model. The findings also indicate that certain conditions such as poor communication and quality culture, could even mask the benefits

of learning. Therefore, management should take proper care when developing quality procedures for projects. Another outcome of this study, was a software that integrated LC into planning and thus, allows planners to utilize learning curve concepts without having any deep or significant learning curve model.

Nevertheless, there are several limitations in this study. The study only focused on deterministic learning curve models and no effort was done to evaluate stochastic learning curve models. The study also did not properly examine the status of the application of the learning curve concepts in the actual industry. Industry practitioners could have had insights that would have benefited the research team or even affected the course of the research. The research team also did not examine the various types of learning such as individual learning, organizational learning, autonomous learning, stochastic learning and the effects of each on productivity.

Future research should therefore observe if different types of learning are modeled by a different set of mathematical classes. In addition this research should also accommodate for randomness and include stochastic models in the developed tool, in order to factor in uncertainty which is a fact of life in the construction industry. The validity of this research could also be enhanced by using the LC concepts for the purpose of resource allocation and generating optimum crew mixes that would maximize the flow of knowledge throughout the organization.

# Bibliography

Adler, P. S. and Clark, K. B. (1991). Behind the learning curve: A sketch of the learning process. *Management Science*, 37(3):267–281.

Al Sarraj, Z. M. (1990). Formal development of line-of-balance technique. *Journal of Construction Engineering and Management*, 116(4):689–704.

Alamri, A. A. and Balkhi, Z. T. (2007). The effects of learning and forgetting on the optimal production lot size for deteriorating items with time varying demand and deterioration rates. *International Journal of Production Economics*, 107(1):125–138.

Ammar, M. and Abdel-Maged, A. (2012). Scheduling of repetitive projects with learning development effect. *Advances in Civil Engineering and Building Materials*, page 341.

Anzanello, M. J. and Fogliatto, F. S. (2011). Learning curve models and applications: Literature review and research directions. *International Journal of Industrial Ergonomics*, 41(5):573–583.

Arditi, D., Tokdemir, O. B., and Suh, K. (2001). Effect of learning on line-of-balance scheduling. *International Journal of Project Management*, 19(5):265–277.

- Argote, L. and Epple, D. (1990). Learning curves in manufacturing. *Science*, 247:920–924.
- Badiru, A. B. (1992). Computational survey of univariate and multivariate learning curve models. *Engineering Management, IEEE Transactions on*, 39(2):176–188.
- Bailey, C. D. and McIntyre, E. V. (1997). The relation between fit and prediction for alternative forms of learning curves and relearning curves. *IIE transactions*, 29(6):487–495.
- Baloff, N. (1971). Extension of the learning curve—some empirical results. *Operational Research Quarterly*, pages 329–340.
- Chambers, S. and Johnston, R. (2000). Experience curves in services: macro and micro level approaches. *International Journal of Operations & Production Management*, 20:842–859.
- Chung, S. (2001). The learning curve and the yield factor: the case of korea's semiconductor industry. *Applied Economics*, 33:473–483.
- Cook, J. E. (1991). A competitive model of the japanese firm. *Journal of Policy Modeling*, 13:93–114.
- Couto, J. and Teixeira, J. (2005). Using linear model for learning curve effect on highrise floor construction. *Construction Management and Economics*, 23(4):355–364.
- Everett, J. G. and Farghal, S. (1994). Learning curve predictors for construction field operations. *Journal of construction engineering and management*, 120(3):603–616.

- Everett, J. G. and Farghal, S. H. (1997). Data representation for predicting performance with learning curves. *Journal of construction engineering and management*, 123(1):46–52.
- Farghal, S. H. and Everett, J. G. (1997). Learning curves: Accuracy in predicting future performance. *Journal of construction engineering and management*, 123(1):41–45.
- Georgy, M. E. (2008). Evolutionary resource scheduler for linear projects. *Automation in Construction*, 17(5):573–583.
- Georgy, M. E., Chang, L.-M., and Zhang, L. (2005). Prediction of engineering performance: a neurofuzzy approach. *Journal of Construction Engineering and Management*, 131(5):548–557.
- Globerson, S. and Gold, D. (1997). Statistical attributes of the power learning curve model. *International journal of production research*, 35(3):699–711.
- Gottlieb, S. C. and Haugbølle, K. (2010). The repetition effect in building and construction works: A literature review. Technical report, SBI forlag.
- Gruber, H. (1992). The learning curve in the production of semiconductor memory chips. *Applied Economics*, 24:885–894.
- Gruber, H. (1994). The yield factor and the learning curve in semiconductor production. *Applied Economics*, 26:837–843.
- Gruber, H. (1996). Trade policy and learning by doing: the case of semiconductors. *Research policy*, 25:723–739.
- Gruber, H. (1998). Learning by doing and spillovers: further evidence for the semiconductor industry. *Review of Industrial Organization*, 13:697–711.

- Hijazi, A. M., AbouRizk, S. M., and Halpin, D. W. (1992). Modeling and simulating learning development in construction. *Journal of construction engineering and management*, 118(4):685–700.
- Hinze, J. and Olbina, S. (2009). Empirical analysis of the learning curve principle in prestressed concrete piles. *Journal of Construction Engineering and Management*, 135(5):425–431.
- Hurley, W. (1996). When are we going to change the learning curve lecture? *Computers & operations research*, 23(5):509–511.
- Jaber, M., Bonney, M., and Moualek, I. (2009). Lot sizing with learning, forgetting and entropy cost. *International Journal of Production Economics*, 118(1):19–25.
- Jaber, M. Y. and Bonney, M. (1996). Production breaks and the learning curve: the forgetting phenomenon. *Applied Mathematical Modelling*, 20(2):162–169.
- Jaber, M. Y. and Guiffrida, A. L. (2004). Learning curves for processes generating defects requiring reworks. *European Journal of Operational Research*, 159:663–672.
- Jaber, M. Y. and Guiffrida, A. L. (2008). Learning curves for imperfect production processes with reworks and process restoration interruptions. *European journal of operational research*, 189(1):93–104.
- Jaber, M. Y. and Kher, H. V. (2004). Variant versus invariant time to total forgetting: the learn–forget curve model revisited. *Computers & Industrial Engineering*, 46(4):697–705.
- Jarkas, A. and Horner, M. (2011). Revisiting the applicability of learning curve the-

- ory to formwork labour productivity. *Construction Management and Economics*, 29(5):483–493.
- Jarkas, A. M. (2010). Critical investigation into the applicability of the learning curve theory to rebar fixing labor productivity. *Journal of Construction Engineering and Management*, 136(12):1279–1288.
- Jergeas, G. and Put, J. V. d. (2001). Benefits of constructability on construction projects. *Journal of Construction Engineering and Management*, 127(4):281–290.
- Kara, S. and Kayis, B. (2005). The effect of the learning process in concurrent engineering projects. *Concurrent Engineering*, 13(3):209–217.
- Kazaz, A., Manisali, E., and Ulubeyli, S. (2008). Effect of basic motivational factors on construction workforce productivity in turkey. *Journal of Civil Engineering and Management*, 14(2):95–106.
- Knecht, G. (1974). Costing, technological growth and generalized learning curves. *Operational Research Quarterly*, pages 487–491.
- Lam, K. C., Lee, D., and Hu, T. (2001). Understanding the effect of the learning-forgetting phenomenon to duration of projects construction. *International Journal of Project Management*, 19:411–414.
- Lapré, M. A., Mukherjee, A. S., and Van Wassenhove, L. N. (2000). Behind the learning curve: Linking learning activities to waste reduction. *Management Science*, 46(5):597–611.
- Larsen, J. K., Shen, G. Q., Lindhard, S. M., and Brunoe, T. D. (2015). Factors affecting schedule delay, cost overrun, and quality level in public construction projects. *Journal of Management in Engineering*, 32(1):04015032.

- Levy, F. K. (1965). Adaptation in the production process. *Management Science*, 11(6):B-136.
- Li, G. and Rajagopalan, S. (1997). The impact of quality on learning. *Journal of Operations Management*, 15(3):181-191.
- Lieberman, M. B. (1984). The learning curve and pricing in the chemical processing industries. *The RAND Journal of Economics*, 15:213-228.
- Long, L. D. and Ohsato, A. (2009). A genetic algorithm-based method for scheduling repetitive construction projects. *Automation in Construction*, 18(4):499-511.
- Love, P. E., Ackermann, F., Teo, P., and Morrison, J. (2015). From individual to collective learning: A conceptual learning framework for enacting rework prevention. *Journal of Construction Engineering and Management*, 141(11):05015009.
- Lutz, J. D., Halpin, D. W., and Wilson, J. R. (1994). Simulation of learning development in repetitive construction. *Journal of Construction Engineering and Management*, 120(4):753-773.
- Mazur, J. E. and Hastie, R. (1978). Learning as accumulation: A reexamination of the learning curve. *Psychological Bulletin*, 85(6):1256.
- Meirovich, G. (2006). Quality of design and quality of conformance: Contingency and synergistic approaches. *Total Quality Management & Business Excellence*, 17(2):205-219.
- Mosey, D. (2009). *Early contractor involvement in building procurement: contracts, partnering and project management*. John Wiley & Sons.



- Nembhard, D. A. and Osothsilp, N. (2002). Task complexity effects on between-individual learning/forgetting variability. *International Journal of Industrial Ergonomics*, 29(5):297–306.
- Nemet, G. F. (2006). Beyond the learning curve: factors influencing cost reductions in photovoltaics. *Energy policy*, 34:3218–3232.
- Panas, A. and Pantouvakis, J. P. (2013). Simulation-based and statistical analysis of the learning effect in floating caisson construction operations. *Journal of Construction Engineering and Management*, 140(1).
- Proverbs, D., Holt, G. D., and Olomolaiye, P. (1998). Factors impacting construction project duration: a comparison between france, germany and the uk. *Building and environment*, 34(2):197–204.
- Rosenbaum, S., Toledo, M., and Gonzalez, V. (2012). Green-lean approach for assessing environmental and production waste in construction. In *Proceedings for the 20th Annual Conference of the IGLC, San Diego, USA*.
- Salameh, M. K., Abdul-Malak, M.-A. U., and Jaber, M. Y. (1993). Mathematical modelling of the effect of human learning in the finite production inventory model. *Applied mathematical modelling*, 17(11):613–615.
- Saraswat, S. P. and Gorgone, J. T. (1990). Organizational learning curve in software installation: an empirical investigation. *Information & management*, 19:53–59.
- Schindler, M. and Eppler, M. J. (2003). Harvesting project knowledge: a review of project learning methods and success factors. *International journal of project management*, 21(3):219–228.

in performance prediction. *Journal of construction engineering and management*, 133(6):474–482.

Zhang, L., Zou, X., and Kan, Z. (2014). Improved strategy for resource allocation in repetitive projects considering the learning effect. *Journal of Construction Engineering and Management*, 140(11).

- Senouci, A. and Al-Derham, H. R. (2008). Genetic algorithm-based multi-objective model for scheduling of linear construction projects. *Advances in Engineering Software*, 39(12):1023–1028.
- Srour, J. F., Kiomjian, D., and Srour, I. M. (2015). Learning curves in construction: A critical review and new model. *Journal of Construction Engineering and Management*, page 06015004.
- Sturm, R. (1999). Cost and quality trends under managed care: is there a learning curve in behavioral health carve-out plans? *Journal of Health Economics*, 18:593–604.
- Thieblot, A. J. (2002). Technology and labor relations in the construction industry. *Journal of Labor Research*, 23(4):559–573.
- Thomas, H. R., Mathews, C. T., and Ward, J. G. (1986). Learning curve models of construction productivity. *Journal of Construction Engineering and Management*, 112:245–258.
- United Nations. Economic Commission for Europe. Committee on Housing, B. and Planning (1965). *Effect of Repetition on Building Operations and Processes on Site: Report of an Enquiry*. Number 438 in Document (United Nations). UN.
- Uzumeri, M. and Nembhard, D. (1998). A population of learners: A new way to measure organizational learning. *Journal of Operations Management*, 16(5):515–528.
- Vits, J. and Gelders, L. (2002). Performance improvement theory. *International journal of production economics*, 77(3):285–298.
- Wong, P. S., On Cheung, S., and Hardcastle, C. (2007). Embodying learning effect