


1996

Construction labor productivity modeling with neural networks and regression analysis

Rifat Sonmez
Iowa State University

Follow this and additional works at: <https://lib.dr.iastate.edu/rtd>

 Part of the [Civil Engineering Commons](#), and the [Industrial Engineering Commons](#)

Recommended Citation

Sonmez, Rifat, "Construction labor productivity modeling with neural networks and regression analysis " (1996). *Retrospective Theses and Dissertations*. 11491.
<https://lib.dr.iastate.edu/rtd/11491>

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

Construction labor productivity modeling with neural networks and regression analysis

by

Rifat Sonmez

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Department: Civil and Construction Engineering
Major: Civil Engineering (Construction Engineering and Management)

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1996

UMI Number: 9635357

**UMI Microform 9635357
Copyright 1996, by UMI Company. All rights reserved.**

**This microform edition is protected against unauthorized
copying under Title 17, United States Code.**

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

TABLE OF CONTENTS

| | |
|---|--------|
| LIST OF FIGURES | iv |
| LIST OF TABLES | vi |
| ACKNOWLEDGEMENTS | vii |
| CHAPTER 1. CONSTRUCTION LABOR PRODUCTIVITY MODELING | 1 |
| 1.1 Introduction | 1 |
| 1.2 Construction Labor Productivity Definitions | 4 |
| 1.3 Adjustment Factors | 5 |
| 1.4 Work Sampling | 6 |
| 1.5 Weather Models | 7 |
| 1.6 Repetition Studies | 12 |
| 1.7 Scheduled Overtime Studies | 13 |
| 1.8 Expectancy Theory Model | 16 |
| 1.9 Action Response Model | 17 |
| 1.10 Expert Systems | 19 |
| 1.11 Factor Model | 20 |
| 1.12 Neural Networks | 23 |
| 1.13 Construction Labor Productivity Modeling with Neural Networks and Regression Analysis | 25 |
| CHAPTER 2. CONSTRUCTION LABOR PRODUCTIVITY MODELING WITH NEURAL NETWORKS AND REGRESSION ANALYSIS | 27 |
| 2.1 Research Objective | 27 |
| 2.2 Description of the Data | 27 |
| 2.3 Limitations of the Data | 32 |
| 2.4 Modeling Methodology | 33 |
| 2.5 Data Identification | 33 |
| 2.6 Regression Analysis | 36 |
| 2.6.1 Regression Models for Concrete Pouring | 46 |
| 2.6.2 Regression Models for Formwork | 48 |
| 2.6.3 Regression Models for Concrete Finishing | 50 |
| 2.6.4 Regression Models for Granular Fill | 52 |

| | |
|--|----|
| 2.7 Neural Network Modeling | 53 |
| 2.7.1 Feedforward Neural Networks and Back Propagation | 53 |
| 2.7.2 Learning Rate and Momentum Factor | 56 |
| 2.7.3 Number of Hidden Units and Pruning | 57 |
| 2.7.4 Direct Connection Between Input and Output Units | 58 |
| 2.7.5 Adding Noise to Training Data | 58 |
| 2.7.6 Neural Network Models for Concrete Pouring | 59 |
| 2.7.7 Neural Network Models for Formwork, Concrete Finishing and Granular Fill | 62 |
| CHAPTER 3. RESULTS | 63 |
| 3.1 Model Comparison | 63 |
| 3.2 Sensitivity Analysis of the Models | 67 |
| CHAPTER 4. CONCLUSIONS | 76 |
| REFERENCES | 79 |
| APPENDIX CODE LISTING | 86 |

LIST OF FIGURES

| | |
|---|----|
| Figure 1.1. Factors Influencing Construction Labor Productivity | 2 |
| Figure 1.2. Effects of Temperature and Relative Humidity on Productivity (NECA) | 10 |
| Figure 1.3. Effects of Temperature and Relative Humidity on Productivity (Grimm-Wagner) | 10 |
| Figure 1.4. Effects of Temperature and Relative Humidity on Productivity (Koehn-Brown) | 11 |
| Figure 1.5. Effects of Temperature and Relative Humidity on Productivity (Thomas-Yiakoumis) | 11 |
| Figure 1.6. Effect of Repetition on Productivity (Thomas, Mathews, and Ward) | 14 |
| Figure 1.7. Effect of Repetition on Productivity (Baker Ridge Tunnel) | 14 |
| Figure 1.8. Effect of Scheduled Overtime on Productivity (50 Hours/Week) | 15 |
| Figure 1.9. Effect of Scheduled Overtime on Productivity (60 Hours/Week) | 15 |
| Figure 1.10. Action Response Model | 18 |
| Figure 1.11. Effects of Temperature and Relative Humidity on Productivity (Factor Model) | 22 |
| Figure 1.12. Effect of Crew Size on Productivity (Factor Model) | 22 |
| Figure 2.1. Flow Chart of the Modeling Methodology | 34 |
| Figure 2.2. Plot of Quantity versus PR for Pour Concrete | 37 |
| Figure 2.3. Plot of Number of Workers versus PR for Pour Concrete | 38 |
| Figure 2.4. Plot of Temperature versus PR for Pour Concrete | 39 |
| Figure 2.5. Plot of Quantity versus PR for Formwork | 40 |
| Figure 2.6. Plot of Number of Workers versus PR for Formwork | 41 |
| Figure 2.7. Plot of Temperature versus PR for Formwork | 42 |
| Figure 2.8. Plot of Quantity versus PR for Trowel Finish Slab | 43 |
| Figure 2.9. Plot of Number of Workers versus PR for Trowel Finish Slab | 43 |

| | |
|---|----|
| Figure 2.10. Plot of Temperature versus PR for Trowel Finish Slab | 43 |
| Figure 2.11. Plot of Quantity versus PR for Fine Granular Fill | 44 |
| Figure 2.12. Plot of Number of Workers versus PR for Fine Granular Fill | 44 |
| Figure 2.13. Plot of Temperature versus PR for Fine Granular Fill | 44 |
| Figure 2.14. Feedforward Neural Network Model with One Hidden Layer | 54 |
| Figure 2.15. Network Error for NP | 61 |
| Figure 3.1. Sensitivity Analysis of NP and RP for Quantity | 68 |
| Figure 3.2. Sensitivity Analysis of NP and RP for Number of Workers | 69 |
| Figure 3.3. Sensitivity Analysis of NP and RP for Temperature | 70 |
| Figure 3.4. Sensitivity Analysis of NF and RF for Quantity | 71 |
| Figure 3.5. Sensitivity Analysis of NF and RF for Number of Workers | 71 |
| Figure 3.6. Sensitivity Analysis of NT and RT for Quantity | 72 |
| Figure 3.7. Sensitivity Analysis of NT and RT for Temperature | 72 |
| Figure 3.8. Sensitivity Analysis of NG and RG for Quantity | 73 |
| Figure 3.9. Sensitivity Analysis of NG and RG for Number of Workers | 73 |
| Figure 3.10. Sensitivity Analysis of NG and RG for Precipitation | 73 |

LIST OF TABLES

| | |
|---|----|
| Table 1.1. Construction Labor Productivity Studies | 3 |
| Table 1.2. Neural Network Models Compared with Regression Models | 24 |
| Table 2.1. Structure of the Labor Database of the Contractor | 28 |
| Table 2.2. Structure of the Quantity Database of the Contractor | 28 |
| Table 2.3. Combined Worhour, Quantity, and Equipment Data | 29 |
| Table 2.4. Calculation of Weekly Average Temperature and Humidity Values | 30 |
| Table 2.6. Labor Productivity Factors | 32 |
| Table 2.7. Production Rate Statistics for Different Job Types | 35 |
| Table 2.8. Regression Statistics for Model RP-1 | 47 |
| Table 2.9. Steps through Model RP-1 to Model RP | 47 |
| Table 2.10. Regression Statistics for Model RF-1 | 49 |
| Table 2.11. Steps through Model RF-1 to Model RF | 49 |
| Table 2.12. Regression Statistics for Model RT-1 | 50 |
| Table 2.13. Steps through Model RT-1 to Model RT | 51 |
| Table 2.14. Regression Statistics for Model RG-1 | 52 |
| Table 2.15. Steps through Model RG-1 to Model RG | 53 |
| Table 2.16. Comparison of Neural Network Models for Concrete Pouring | 60 |
| Table 2.17. Neural Network Models for Formwork, Concrete Finishing, and Granular Fill | 62 |
| Table 3.1. Comparison of the Models | 66 |

ACKNOWLEDGEMENTS

I would especially thank to Dr. Jim Rowings for his support through out the course of my program. I would also like to thank all of the other members of my program of study committe; Dr. Mark Federle, Dr. Vasant Honavar, Dr William Meeker, Dr Charles Jahren and, Dr. John Even. They have brought many insightful and useful suggestions to help improve this dissertation. I sincerely thank Doug Powell for sharing his expertise, and for helping collect the data used in this study.

CHAPTER 1. CONSTRUCTION LABOR PRODUCTIVITY MODELING

1.1 Introduction

Construction labor productivity variations are results of several factors. These factors can be grouped into three main categories : (1) Management related factors including project team, management control, methods and equipment, materials and tools availability, crew composition, work sequence, scheduled overtime, congestion (Figure 1.1), (2) project related factors including specifications, design features, crew size, repetition, site conditions, temperature, humidity, precipitation, and (3) labor related factors including incentives, morale, fatigue, unionized labor, quality of craftsmanship, absenteeism, and turnover (Borcherding and Alarcon 1991; Neil and Knack 1984; Dalliva 1954).

Construction productivity models explain productivity variations by the factors included in the model. These models are needed for construction planning, estimating, and scheduling. In planning, productivity models of controllable factors (such as crew size or scheduled overtime) are needed for maximizing labor productivity to achieve lower labor cost and shorter project duration. In estimating, productivity models are used to predict labor costs; and finally in scheduling, productivity models are needed to forecast activity durations.

Although productivity modeling is an important part of construction planning, estimating, and scheduling, models developed so far are limited in explaining the variations of productivity. Most of these models included a single factor while neglecting the variations caused by other factors. Furthermore the models were based on a limited amount of data. Table 1.1 includes a list of construction labor productivity studies. Each of these studies and their limitations are discussed in detail in Sections 1.3 to 1.11.

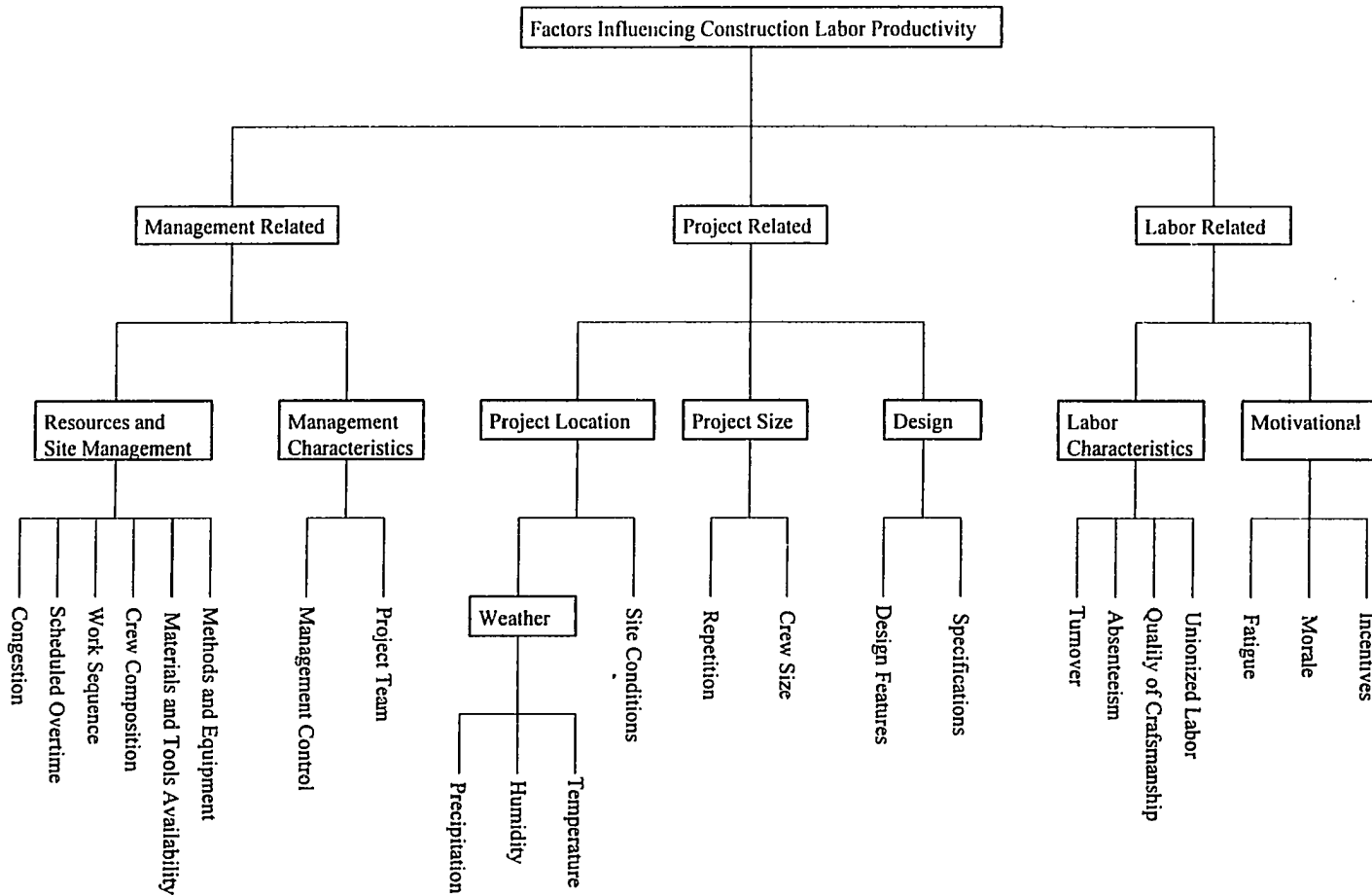


Figure 1.1 Factors Influencing Construction Labor Productivity

Table 1.1 Construction Labor Productivity Studies

| Category | Source(s) |
|-------------------------|---|
| Adjustment Factors | Dallavia (1954), Neil (1982), Neil and Knack (1984) |
| Work Sampling | Thomas et al. (1984), Liou and Borcharding (1986), Thomas (1991) |
| Weather | Clapp (1966), "The effect" (1974), Grimm and Wagner (1974), Koehn and Brown (1985), Thomas and Yiakoumis (1987) |
| Repetition | "Effects" (1965), McClure et al. (1980a and 1980b) Ward and Thomas (1984), Thomas et al. (1986), Oglesby et al (1987) |
| Scheduled Overtime | "Scheduled" (1980), "The Effects" (1988), Thomas (1992) "Effects" (1994) |
| Expectancy Theory Model | Maloney (1981 and 1986), Maloney and McFillen (1985, 1986 and 1987) |
| Action Response Model | Halligan et al. (1994) |
| Expert Systems | Hendrickson and Martinelli (1987), Christian and Hachey (1995) |
| Factor Model | Thomas et al. (1992), Sanders and Thomas (1993), Thomas and Sakarcan (1994) |
| Neural Networks | Moselhi et al. (1991) |

Regression analysis has been the common tool used in construction productivity studies, but in recent years neural networks have been a successful alternative to regression analysis for problems similar to construction labor productivity modeling. However the potential capabilities of neural networks for construction labor productivity modeling have not been examined. The focus of this

study was to develop a methodology for modeling construction labor productivity of different tasks with multiple factors. Use of neural networks was explored as a part of the overall modeling methodology. The methodology was used to develop multivariate productivity models for concrete pouring, formwork, concrete finishing, and granular fill. Results of the models were compared with the results of the models developed by other present construction labor productivity modeling methodologies.

1.2 Construction Labor Productivity Definitions

In the construction industry the meaning of the term productivity varies with its application to different areas. The term productivity usually refers to the output produced per unit input. The overall measure of the productivity can be defined by the total factor productivity (TFP). TFP is used by the several government agencies including the Department of Commerce, and is defined by:

$$TFP = \frac{\text{Total Output}}{\text{Labor} + \text{Material} + \text{Equipment} + \text{Energy} + \text{Capital}} \quad (1.1)$$

TFP is an economic model, in which inputs and outputs are measured in terms of dollars. TFP is not very useful for contractors, as it can be highly inaccurate if applied to a specific project because of difficulties in predicting the various inputs. At the project site contractors are usually interested in labor productivity (Thomas et.al. 1990). There is no standard definition of labor productivity, it can be defined in one of the following ways (Thomas and Mathews 1985):

$$\text{Labor Productivity} = \frac{\text{Output}}{\text{Labor Cost}} \quad (1.2)$$

or;

$$\text{Labor Productivity} = \frac{\text{Output}}{\text{Work Hour}} \quad (1.3)$$

Other terms such as efficiency may be used synonymously with the term productivity. Equation 1.3 is usually referred, to as, the production rate. Sometimes inverse of equation 1.3 is used by the contractors:

$$\text{Labor Productivity} = \frac{\text{Work Hour}}{\text{Output}} \quad (1.4)$$

The outputs in equations 1.2, 1.3 and 1.4 are usually measured in appropriate units for various kinds of product outputs for the tasks. Typical units are square feet, cubic yards, and tons.

In this study, equation 1.3 was used to calculate productivity because: (1) The equation fits into classical definition of productivity where productivity is the ratio of output of a production process to the corresponding input (Martin 1991; Skills 1992) and, (2) a higher productivity value defined by the equation indicates a higher output/input ratio.

1.3 Adjustment Factors

Several attempts have been made to determine effects of various factors on construction labor productivity. The early studies (Dalliva 1954; Neil 1982; Neil and Knack 1984) recommended use of adjustment factors to quantify the effects of these factors. Adjustment factors are generally values between zero and one, representing the effects of factors at various levels. The effect of a factor is reflected by multiplying the average productivity rate with the adjustment factor of the corresponding factor level.

Dalliva (1954) recommended adjustment factors for eight different groups including: general economy, amount of work, labor, supervision, job conditions, weather, equipment, and delay. The adjustment factors were given in terms of ranges for each of the three (low, average, high) factor

levels. For example, Dalliva suggested that for a high amount of precipitation, average productivity should be multiplied by a factor between 0.25 and 0.55. However, Dalliva neither defined a quantified range for the high amount of precipitation, nor suggested a quantified procedure to select the adjustment factor between 0.25 and 0.55.

Neil (1982), and Neil and Knack (1984) proposed adjustment factors for seventeen factors. These adjustment factors were also given in terms of ranges. Very little is known about how these adjustment factors were determined.

A set of adjustment factors commonly accepted by the construction industry does not exist. Adjustment factors recommended by different sources or experts may be different for the same factor levels. Adjustment factors method also does not define a quantified methodology for productivity prediction. Although this method has several limitations, several contractors use a similar approach for productivity prediction because of its simplicity (Sonmez 1992). The estimators quite often predict labor productivity by adjusting the average productivity for project conditions on the basis of their experience and judgment.

1.4 Work Sampling

Work sampling is a technique that measures the percent of time craftsmen spend in various categories of tasks, such as direct work, transporting materials, or waiting (Thomas 1991). Research in this area has been ongoing since the early 1980's. Thomas, Guevara, and Gustenhoven (1984) compiled data from construction of a nuclear power plant. The study focused on a ten worker pipefitter crew working in the containment building during a ten week period. The ratio of earned to actual workhours was used as the measure of the labor productivity. The data indicated a strong correlation between productivity and direct work. The conclusion was that work sampling can be used as an accurate estimator of construction productivity, provided that the direct work is narrowly defined.

In another work sampling study, Liou and Borcharding (1986) collected 45 data points from eleven nuclear power projects, and four fossil fuel power projects. The ratio of workhours to units completed was used as the measure of the productivity. Liou and Borcharding concluded that work sampling data had a significant relationship with the productivity. Thomas (1991) focused on the main assumption of the work sampling productivity models. The main assumption of the work sampling models was that the percent of direct work was related to labor productivity. The database of the study was obtained from seven sources including Thomas, Guevara, and Gustenhoven (1984), and Liou and Borcharding (1986). The database consisted of 288 data points compiled from 48 projects. Regression models of labor productivity where direct work was used as the independent variable were developed. The models had a very low coefficient of determination, indicating an insignificant relation between direct work and productivity. Thomas suggested that work sampling studies show how busy the crafts are, and cannot be used to predict labor productivity. Thomas (1990) also argued that the early studies by Thomas, Guevara, and Gustenhoven (1984), and Liou and Borcharding (1986) had overstated the productivity prediction capabilities of the work sampling models.

1.5 Weather Models

Several studies have been carried out to model the effects of weather on construction labor productivity. One of the early studies was performed by Clapp (1966). Clapp studied five housing projects in the United Kingdom and classified manhour losses due to weather into five categories: (1) Bad weather time in which the craftsman can not work but is paid; (2) reduced productivity in which less output is obtained with the same labor input; (3) rework because of damage or low quality workmanship resulting from frost, ice, wind or rain; (4) high absenteeism; (5) reduced working schedule. Clapp reported that considerable manhours were lost in all categories except for the second one.

The National Electrical Contractors Association (NECA) conducted a controlled experiment to quantify the effects of humidity and temperature on labor productivity ("The Effect" 1974). It was concluded that productivity varies as a function of temperature and humidity; however, the study had several limitations. The crew of the experiment consisted of only two electricians, and the task was limited to installation of electrical boxes and duplex outlets. Furthermore, there was only a single observation for each of the temperature and humidity levels because the study time was limited to six days. Finally factors other than temperature and humidity were not considered in the study although they might have significantly influenced the observations. Thomas and Yiakoumis (1987) argued that improved productivity resulting from familiarization with highly repetitious work could have affected the results for the NECA study. Plots of productivity versus temperature for different humidity levels were presented as the findings of the NECA study. Models explaining productivity variations due to temperature and humidity were not presented.

Grimm and Wagner (1974) studied mason productivity over a period of nine months during the construction of 283 test walls under regulated conditions. The Grimm and Wagner study did not include factors other than temperature and humidity like the NECA study. It was argued by Thomas and Yiakoumis (1987) that the exclusion of the repetition effect in the Grimm and Wagner study may have led to an overstatement of productivity losses due to weather. Work by Grimm and Wagner included contour plots of productivity versus temperature and humidity for the data but did not include productivity models.

Koehn and Brown (1985) developed two regression models to quantify productivity variations due to temperature and humidity. The data used to develop the models were obtained from a number of sources including NECA, and the Grimm-Wagner studies. The data collection methods of the five sources used in the study were not same; furthermore, some of the resources included very limited information about the way in which productivity was measured. The data set of the study included productivity data of seven different tasks or crafts. These were grouped as manual excavation,

erection, masonry, electrical, carpentry, laborer, and equipment excavation. All of the data from different tasks were combined to develop the models; however, terms presenting variations due to different tasks were not included in the models. The same limitation was also present in the model developed by Thomas and Yiakoumis (1987). This model was based on the data compiled from three building projects. The data included performance ratios (actual/expected productivity) of steel, masonry, and formwork tasks. Temperature and humidity were used as the independent variables to explain variations in the performance ratios. The effect of repetition was considered during calculations of the performance factors.

The data of the NECA and Grimm-Wagner studies, and the predictions of Koehn-Brown and Thomas-Yiakouimis models were plotted through Figures 1.2 to 1.5. In all of the plots at low temperature levels, productivity increases as the temperature increases, and at high temperature levels, productivity decreases as the temperature increases. But there are a lot of differences among the relative humidity effects of the four studies. The NECA plot indicates that increase in humidity rates decrease productivity only at high temperature rates, though the Grimm-Wagner plot suggests that productivity declines when relative humidity deviates from 60 % at all temperature levels. The Koehn-Brown model however, indicates that productivity does not improve when relative humidity increases from 35% to 60 %, unlike the Grimm-Wagner study. The model suggests that high humidity rates at low temperature levels result in a decreased productivity. The Thomas and Yiakouimis model, on the other hand, suggests a unique relation between temperature and relative humidity. The plot of the model indicates a very significant effect of relative humidity beyond 80 %. This effect results in a zero productivity beyond 85 % at all temperature levels.

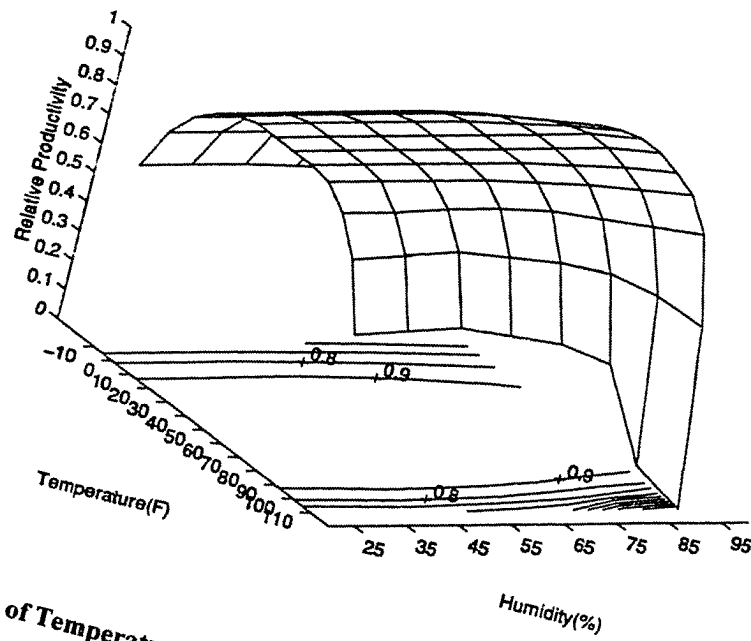


Figure 1.2 Effects of Temperature and Relative Humidity on Productivity (NECA)

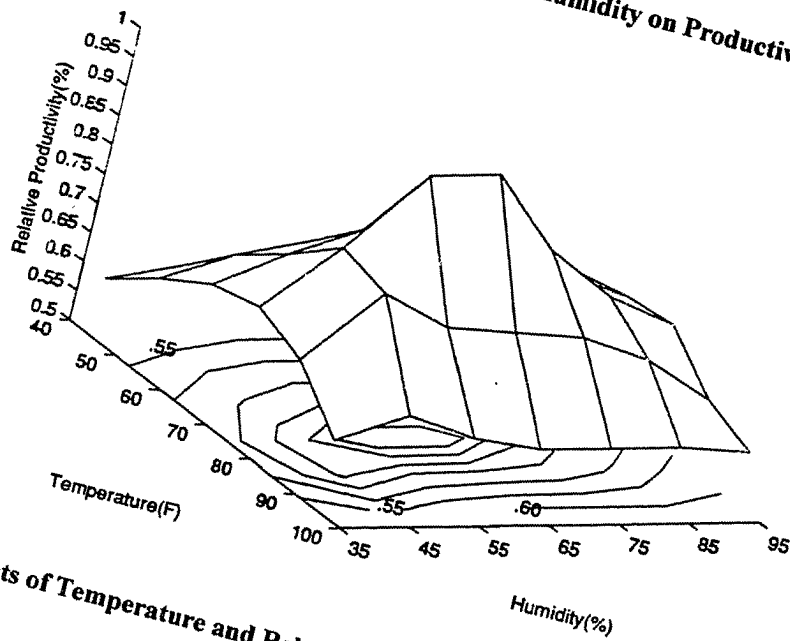


Figure 1.3 Effects of Temperature and Relative Humidity on Productivity (Grimm-Wagner)

Figure 1.5 Effects of Temperature and Relative Humidity on Productivity (Thomas-Yiakoumis)

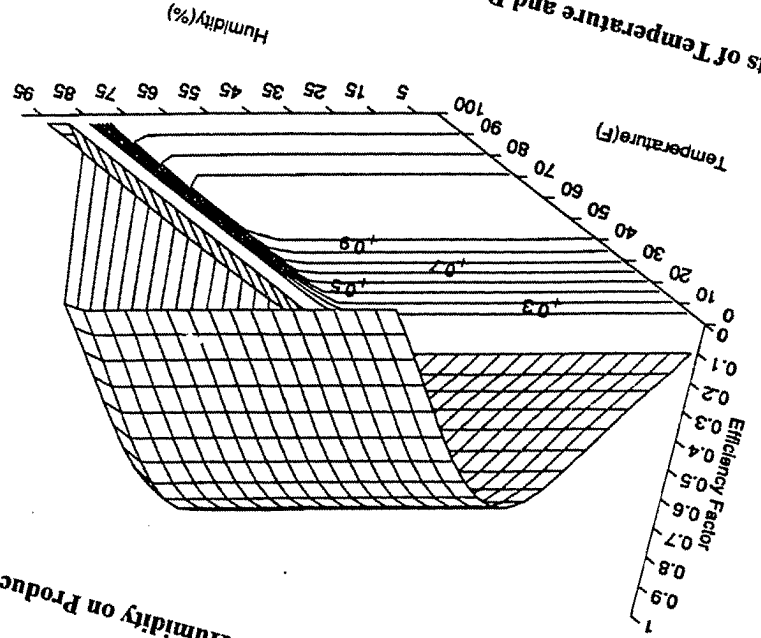
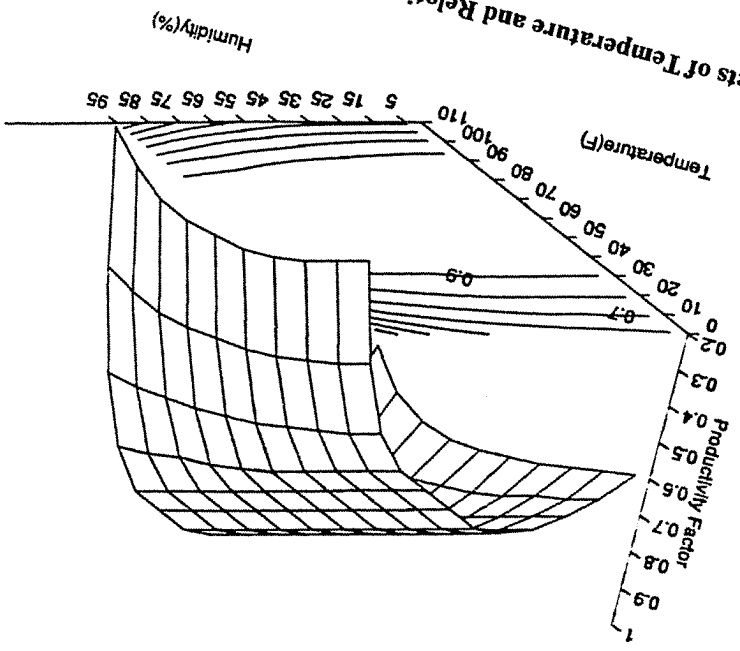


Figure 1.4 Effects of Temperature and Relative Humidity on Productivity (Kohn-Brown)



1.6 Repetition Studies

It is expected that productivity will improve with continuous repetition of a task as the crew becomes more familiar and skilled with the task. Repetition may also lead to improved equipment, crew, and material management, and development of more efficient techniques. The effect of repetition on construction labor productivity was modeled in several studies. These models are usually referred as learning curves.

Learning curves for construction tasks were developed in an early United Nations study ("Effects" 1965). In this study a number of reports from various European Countries were reviewed. The learning curve models developed by the Norwegian Building Research Institute and data of 44 residential building activities were also included in the findings of the study. The data reported in the United Nations study along with, data of a building project (Ward and Thomas 1984) and a bridge project (McClure et al. 1980a, 1980b) were used by Thomas, Mathews and Ward (1986) to compare different learning curve alternatives for construction tasks. Thomas, Mathews and Ward used cumulative average productivity, instead of unit productivity used by the Norwegian Building Research Institute.

The use of cumulative average productivity for the installation time indicated a relation between the installation time and the cumulative precast concrete plank number for the building project, although no relation was identified by Thomas, Mathews, and Ward from the initial plot of unit installation time versus cumulative plank number. Thomas, Mathews, and Ward concluded that the use of cumulative average productivity instead of unit productivity was superior for learning curve modeling. Their conclusions were based on the coefficient of determination values calculated for the two learning models. The author of this study, however, believes that the use of cumulative average productivity may lead to an overstatement of the effect of repetition on productivity. This overstatement is mainly due to the very few observations at the initial stages. As in the building project presented, most of the installation times were between one and five minutes at the initial

stage. But there were two observations at the seventeen minute range. These two observations caused a dramatic increase in the cumulative average productivity value at the initial stage, until their effect in the cumulative productivity value were balanced by enough observations. The decrease in the cumulative productivity curve given in Figure 1.6 is mainly because of this balancing. If the installation time decreased with repetition this would have been also observed in the plot of unit installation time versus cumulative plank number.

Productivity improvements due to repetition was observed at the Baker Ridge Highway Tunnel (Oglesby et al. 1989). A plot of number of shifts required to excavate each drift versus its drift number is given in Figure 1.7. Each tunnel drift was 1330 feet long. The progress was not always smooth because of the effects of factors other than repetition.

1.7 Scheduled Overtime Studies

Scheduled overtime is another factor that is believed to cause variations in labor productivity. Overtime was reported to have significant influence on labor productivity by Proctor & Gamble during their Green Bay operation ("Scheduled" 1980). Figures 1.8 and 1.9 illustrate the reported effect. The productivity observations were the average of the productivity rates for a week and the initial week of the project was taken as a baseline.

The Construction Industry Institute (CII) sponsored a three-year study of overtime in 1984 ("The Effects" 1988) which included data of seven projects. The overtime influences on productivity were not consistent. This study concluded that overtime does not necessarily cause a decrease in the productivity. In another CII overtime study data of electrical and mechanical crews from four projects were compiled (Thomas and Raynar 1994). Only three crews in the study worked an overtime schedule for at least four weeks or more. There was also inconsistency in the overtime schedule. For example, a crew may have worked five days in one week, six days the next week, and then return back to a five day work schedule. The plots of average productivity values of the crews

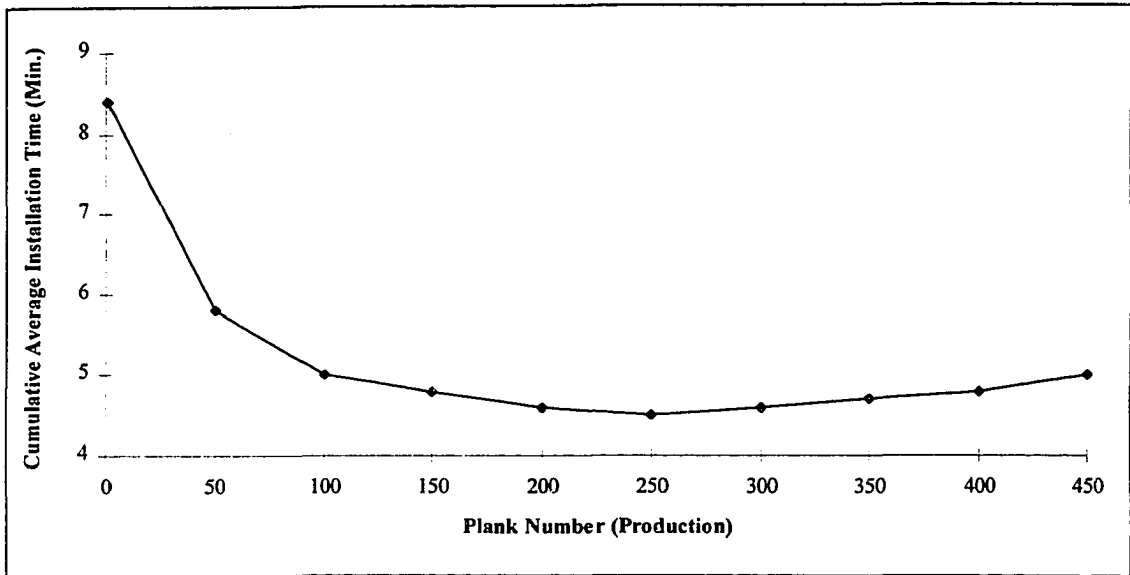


Figure 1.6 Effect of Repetition On Productivity (Thomas, Mathews and Ward)

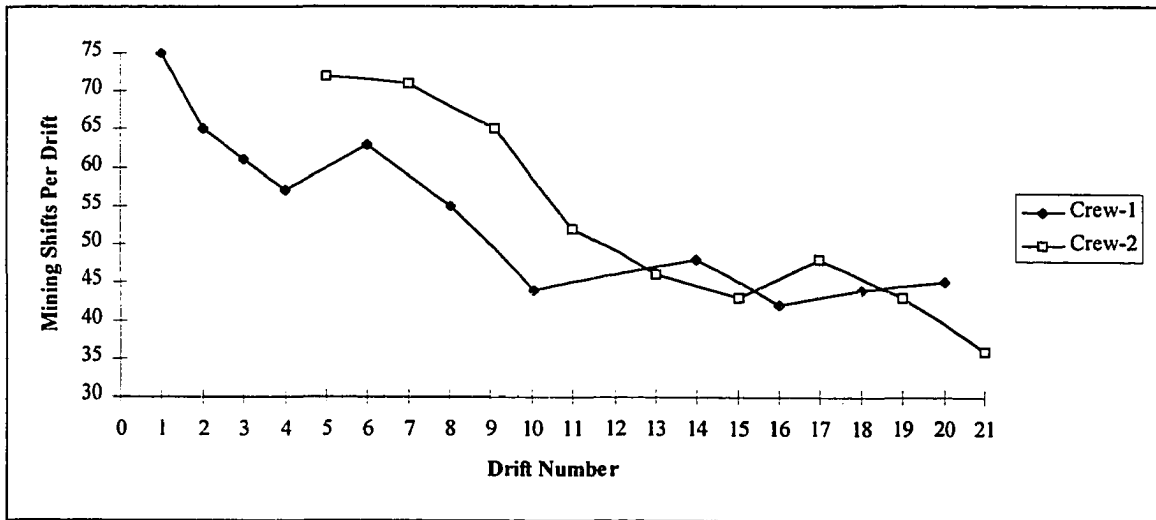


Figure 1.7 Effect of Repetition On Productivity (Baker Ridge Tunnel)

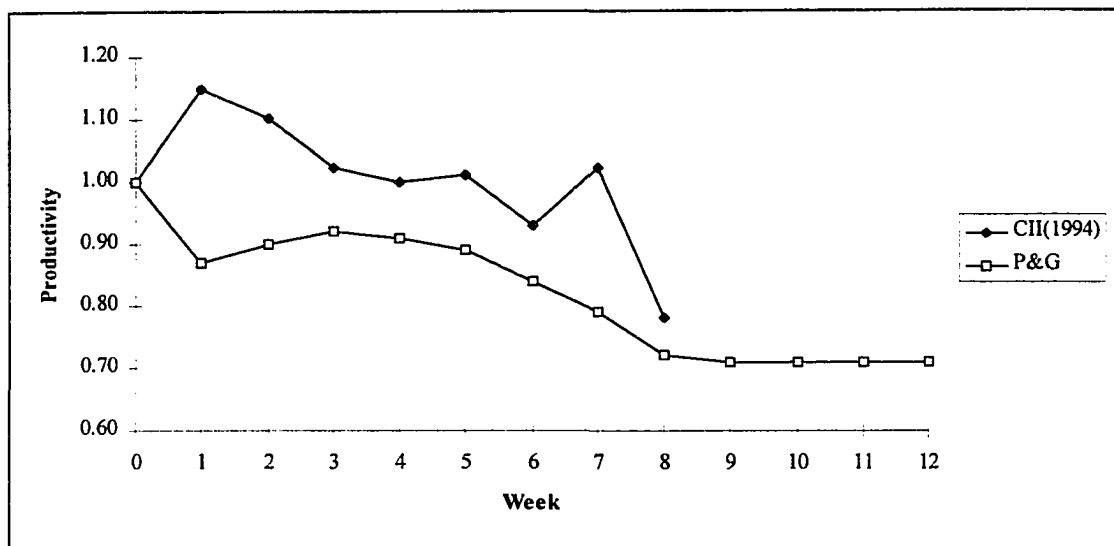


Figure 1.8 Effect of Scheduled Overtime On Productivity (50 Hours/Week)

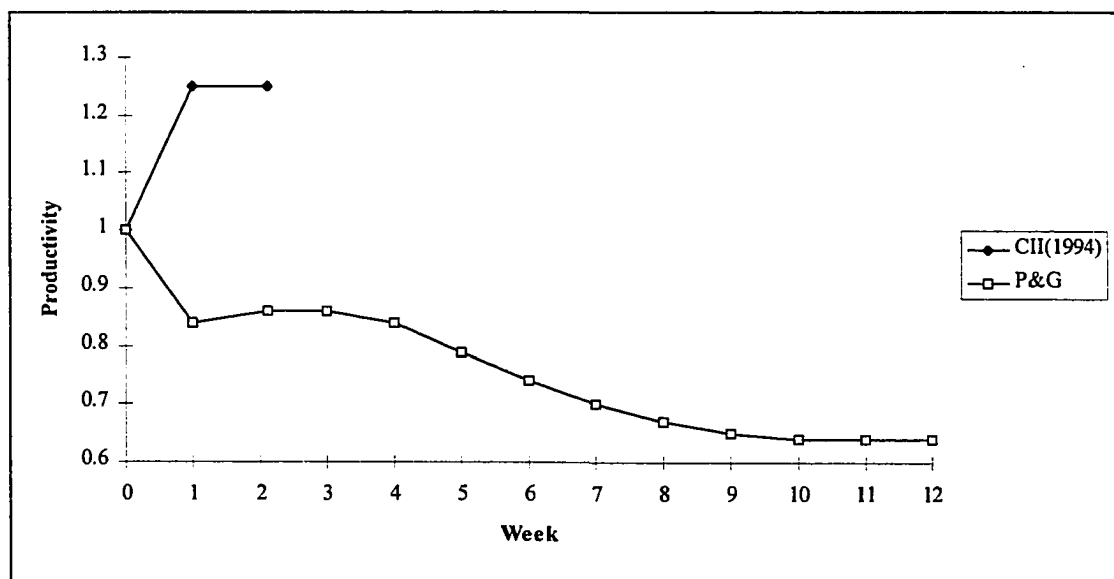


Figure 1.9 Effect of Scheduled Overtime On Productivity (60 Hours/Week)

for each week of overtime are given in Figures 1.8 and 1.9. The initial week (week 0) was taken as the baseline so the productivity values could be compared with the values of the Proctor & Gamble study. The productivity value for week seven of 1994 CII study was obtained from a single crew observation, whereas productivity values for the first four weeks were obtained from averages of at least four crew observations (Figure 1.8). In the findings of the 1994 CII study Thomas and Raynar concluded that the average loss due to productivity was within the range of 15 %. However, they also mentioned that the productivity losses due to overtime were not automatic but could range from 0 % to approximately 25 % for crews (projects) where there were no factors influencing productivity.

The methods used in the mentioned overtime studies were limited to plots of data and calculation of means. Statistical methods such as regression or variance analysis were not used; therefore, the statistical significance of the difference between regular time productivity versus overtime productivity for the studies was not available.

1.8 Expectancy Theory Model

A variety of motivational models were studied to understand construction motivation. The first motivational model validated for construction activities was the expectancy theory model (Maloney and McFillen 1985).

Expectancy theory explains variations in the performance by the effort that a worker is willing to exert on a task. Effort is related to the incentives and can be increased or decreased by job conditions, management actions, and rewards. Expectancy theory suggests that if the worker has adequate knowledge and skills, proper direction is given by the management, and constraints are removed, then the performance will be high. Although expectancy theory was initially developed as a theory of individual performance, it can also be applied at the crew level.

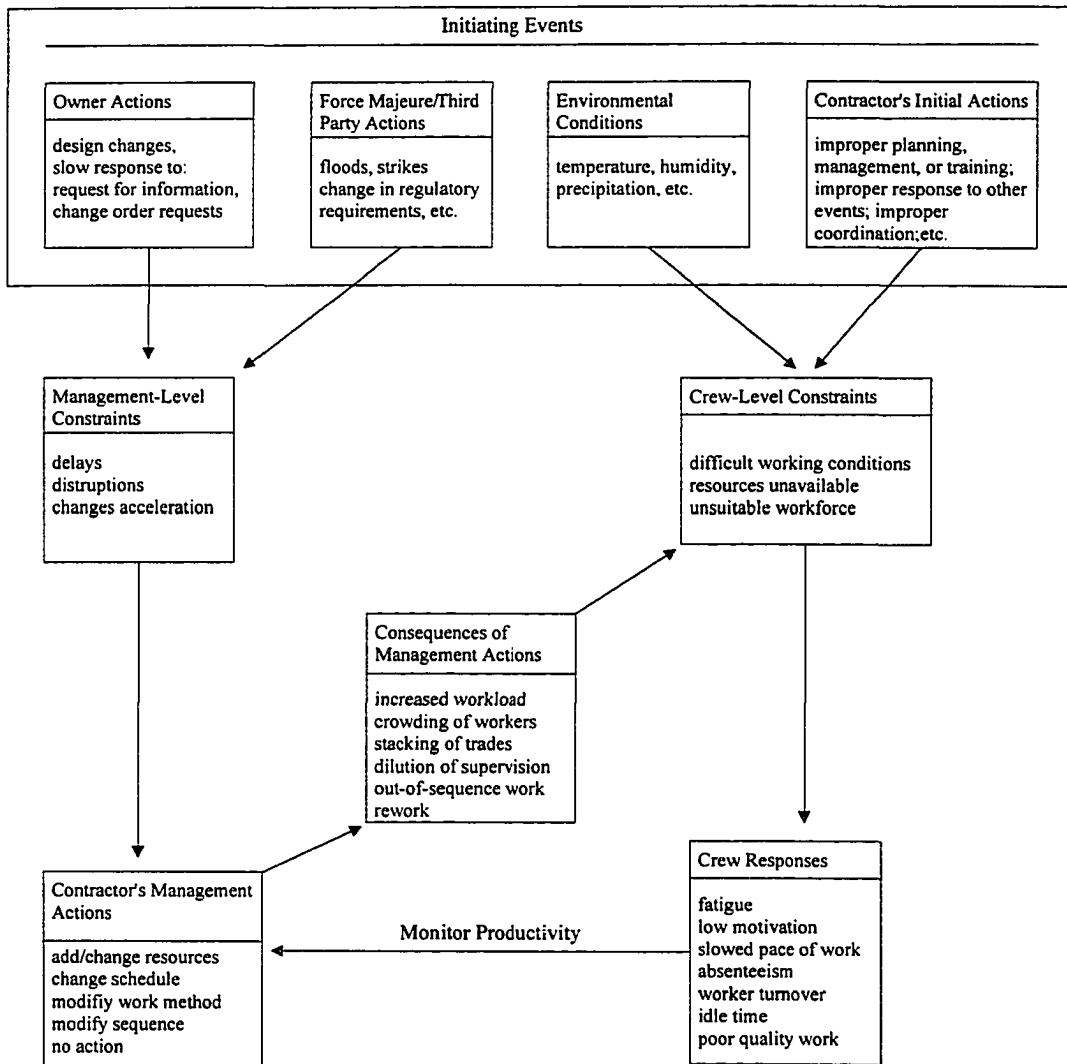
The expectancy theory model was validated by Maloney and McFillen (1985) through a survey. The survey included 703 responses from unionized construction workers of different crafts. The Maloney and McFillen model was not validated through direct productivity measurement, but only self-reported measures of productivity were collected. Therefore the model is not adequate to quantify the effects of factors or to predict productivity (as defined by equation 1.3).

1.9 Action Response Model

The action response model graphically depicts how a variety of factors may interact to cause a loss of productivity (Halligan et al. 1994). The action response model has six components; initiating events, management-level constraints, crew-level constraints, contractor's management actions, consequences of management actions, and crew responses (Figure 1.10). Initiating events, which include owner actions, force majeure and third party actions, environmental conditions, and contractor's initial actions may ultimately lead to reduced productivity. Owner actions and force majeure/third party actions may result in delays, disruptions, changes, or acceleration to the project. The contractor is usually made aware of these results through a directive or a change order.

Difficult working conditions, unavailability of resources, and unsuitable work force are typically caused either by the contractor's actions or by environmental conditions. The contractor may not become aware of these crew-level constraints unless labor productivity is measured. A variety of management actions can be taken to eliminate loss of productivity. However a choice of an inappropriate action may add additional constraints to the project.

Depending on the specifics of the job and contractor's management actions a variety of consequences are possible (Figure 1.10). These consequences may result in difficult working conditions, unavailability of resources, or unsuitable work force; each influence crew responses and may cause loss of productivity.



Note: Arrows indicate "may lead to" relationship

Figure 1.10 Action Response Model

The action response model was not validated by any technique; only case studies for the model were given. The model cannot be used to quantify the effects of the productivity factors or to predict productivity; however, it may be used to determine cause and effect relations in loss of productivity to go along with more rigorous statistical analysis in the future.

1.10 Expert Systems

The use of expert systems for construction productivity modeling was explored by two studies. The first expert system of construction labor productivity, called "MASON", was developed by Hendrickson, Martinelli, and Rehak (1987); and the second expert system was developed by Christian and Hachey (1995). MASON was used for activity duration and productivity prediction for masonry construction. The reasoning used by MASON was developed through interviews with a professional mason and a supporting labor. Size, type and location of the job, temperature, precipitation level, size of crew, type of labor, and material being used were the factors that were included in MASON to modify productivity. MASON was not validated by any technique such as cross validation or closeness of fit comparison with the factor model. MASON was limited to knowledge of a professional mason and a supporting laborer. Multiple experts were not consulted to reach a consistent model of labor productivity for masonry construction.

A more recent expert system was developed by Christian and Hachey (1995) to predict production rates for concrete placement. Three sources were used to determine the expert rules. These sources were heuristic knowledge, published knowledge, and field knowledge. Field knowledge consisted of observations obtained from eleven projects by video recording and stopwatch studies. The total observation time at each site varied between 68 and 263 minutes. The work sampling variation technique was used to analyze variations in the productivity data. The conclusion was that waiting time delay was a very significant cause for reduced productivity.

The expert system developed by Christian and Hachey had several limitations. First of all, the data sources used to develop the expert system had significant variations and inconsistencies. Very little is known about how these variations and inconsistencies were analyzed to determine the expert rules. Second, the work sampling technique was used to analyze the field productivity data. The limitations of work sampling for productivity modeling was discussed in Section 1.4. Third, the field data was very limited; the longest total observation period for a project was 263 minutes, whereas the shortest was 68 minutes. The total observation periods were very short to get a good sample that includes possible variations in productivity due to different factors. Finally, this expert system was not validated by any technique.

The rules of the two expert systems discussed consisted of adjustment factors for different factor levels. In MASON, the knowledge of a mason and a laborer was used to determine these adjustment factors, whereas, in the second expert system three different sources were used. Once these adjustment factors are determined, productivity prediction can also be done manually by multiplying normal productivity with the adjustment factors for the given factor levels, as discussed in Section 1.3.

Construction productivity modeling requires quantification of previous experiences. This quantification can be done by mapping nonlinear, noisy, productivity data. However expert systems lack the ability of mapping noisy data and generalizing solutions (Wassermann 1989; Zahedi 1991).

1.11 Factor Model

The factor model is a multivariate regression approach for modeling of construction labor productivity. Quantification of effects of the factors involves the statistical analysis of labor productivity. Data from thirteen projects were compiled by Sanders and Thomas (1993) to study factors affecting masonry productivity. However, all of the data compiled were not used to develop the model. The database including 465 samples were divided into two parts: Non disrupted or

normal working days and, disrupted or abnormal working days. Abnormal conditions were usually results of disruptions such as congestion, lack of materials, and bad weather. Only normal working days consisting of 286 samples were used to develop the factor model. The following model was suggested:

$$PDP = \alpha + \beta_1 + \beta_2 + \beta_3 + \omega + \theta + \lambda_1c + \lambda_2c^2 + \lambda_3c^3 \quad (1.4)$$

where PDP is the predicted daily productivity; α is a constant term representing standard conditions; β_1 is the work type coefficient; β_2 is the physical element coefficient; β_3 is the design detail coefficient; ω is the construction method coefficient; θ is the weather zone coefficient; λ_1 , λ_2 , and λ_3 are the corresponding coefficients for crew size terms; and c is the crew size.

The factor model given explained 41 % of the total variability of the non disrupted productivity data. Plots of factor model for temperature, humidity and crew size factors are given in Figures 1.11 and 1.12. To obtain Figure 1.11 temperature and humidity levels were varied, while rest of the factors were kept constant at their standard condition levels. The standard condition levels were defined by Sanders, and Thomas (1993). Figure 1.12 was obtained in a similar way, but this time crew size level was varied while the rest of the factors were kept constant at their standard condition levels.

The factor model suggests that productivity declines as temperature increases. This contradicts all of the previously discussed weather models (Section 1.5) where productivity improves as the temperature increases at low temperature levels. The contradiction may be because of the removal of the disrupted working days due to abnormal weather conditions. It is possible that these data points included low productivity values due to cold weather disruptions.

The factor model is limited to binary relations, except for the crew size factor. The model was compared with percent complete method for closeness of fit, however, the model has not been validated by a technique such as cross validation; therefore the predictive accuracy of the model is questionable.

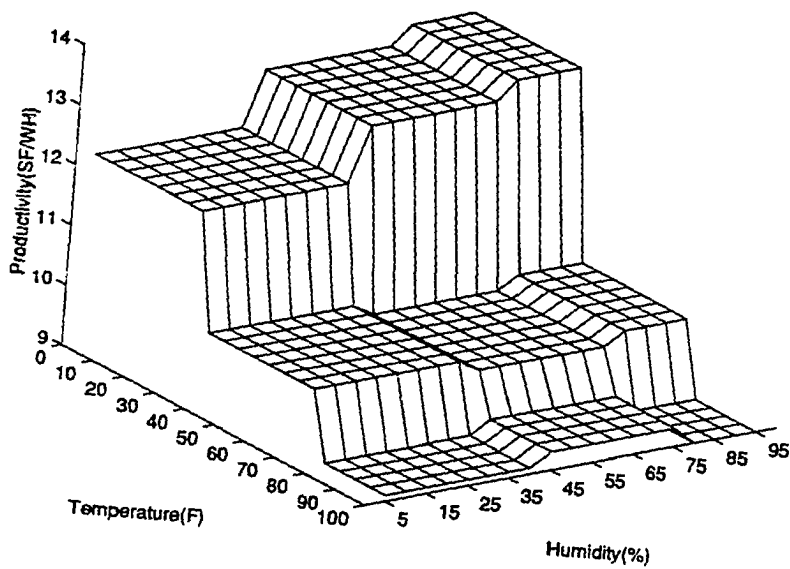


Figure 1.11 Effects of Temperature and Relative Humidity on Productivity (Factor Model)

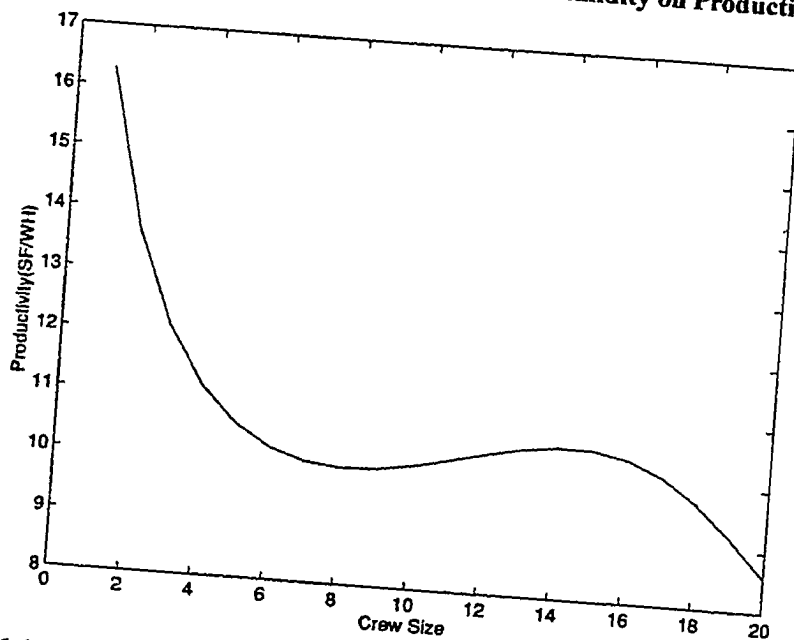


Figure 1.12 Effect of Crew Size on Productivity (Factor Model)

1.12 Neural Networks

Neural network models are algorithms for cognitive tasks, such as learning and optimization, that are based on concepts derived from the research into the nature of the brain (Muller and Reinhardt, 1990). Neural networks have the capability of learning from a number of input patterns (representing different problem encounters) and their associated output patterns (representing the conclusions and decisions). During the process called training, the network generalizes the knowledge, and becomes capable of providing solutions to the new problems even if only incomplete or noisy data are available. Once a network is trained using an adequately representative training set, it can be used to classify or to predict the output of the modeled system for a given input pattern. One of the attractive properties of such networks is their capacity for tolerating moderate amounts of noise, and variations in the input.

Neural networks provide a variety of powerful tools for optimization, function approximation, pattern classification, and modeling. Neural network models have been developed and used as an alternative to regression analysis since the back propagation algorithm was proposed. Table 1.2 includes some of the studies in which both neural networks and regression analysis were used to model a specific problem. Different comparison methods were used in the studies. In studies 1,3,5 and 6 only closeness of fit (See Section 2.6) was used to compare neural networks with the regression models. However a good closeness fit for a neural network model does not necessarily guarantee a good prediction performance, therefore generalization (prediction) performance of these neural networks is not known.

The regression model for concrete strength prediction (Williams et.al. 1992) was slightly better than the neural network model. In another study by the same author, it was concluded that neural networks cannot accurately predict the variations of construction cost indexes because the regression model was more accurate than the presented neural network model (Williams 1994). However the study had two main limitations. In the regression model only, one variable was used, but in neural

network nine variables were used to predict the variations of the construction cost index. Data identification was not conducted to determine which one of these nine variables had a significance influence on the construction cost index variations. It is quite possible that some of the nine variables used in the neural network model did not have a significant effect on the cost index, which may have resulted in overtraining. The second limitation of the study was twenty hidden units used

Table 1.2 Neural Network Models Compared with Regression Models

| No | Application | Comparison Method | Source |
|----|-----------------------|---|---------------------------|
| 1 | Stock Market | Closeness of Fit | Kimoto et al. (1990) |
| 2 | Electrical Load | Prediction Performance | Srinivasan et al (1991) |
| 3 | Industrial Production | Closeness of Fit | Niu et al. (1991) |
| 4 | Concrete Strength | Prediction Performance | Williams et al. (1992) |
| 5 | Pump Cost | Closeness of Fit | McKim (1993) |
| 6 | Bankruptcy | Closeness of Fit | Fletcher and Goss (1993) |
| 7 | Stock Ranking | Closeness of Fit & Prediction Performance | Refenes et al. (1994) |
| 8 | River Flow | Closeness of Fit & Prediction Performance | Karunanithi et al. (1994) |
| 9 | Cost Index | Prediction Performance | Williams (1994) |
| 10 | Soil Correlations | Closeness of Fit & Prediction Performance | Goh (1995) |

in the neural network model with only 215 training samples. Use of twenty hidden units, with nine input, and two output units requires estimation of 242 model parameters. 215 training samples may not be sufficient to estimate 242 model parameters. The two limitations; use of insignificant input variables and use of too many hidden units with few training facts, might have caused poor predictive accuracy for the neural network model.

In the studies 2,7, and 8 neural network models (Table 1.2) were reported to be more accurate than the regression models. The neural network models were compared with the traditional regression models that have been commonly used for the problems 2 and 8, but little is known about the class of the regression model used in study 7, in which the neural network model was reported to be much better than the regression model for both closeness of fit, and prediction performance.

In a recent study, Goh (1995) demonstrated the potential of neural networks to capture nonlinear interaction between various soil variables. Goh compared the neural networks with the previously used regression models, and concluded that neural network models were able to produce reasonably accurate predictions.

Neural networks with their modeling capabilities appear to be a powerful tool for construction labor productivity modeling, as was originally pointed out by Moselhi, Hegazy, and Fazio (1991) in an article about possible neural network applications in construction engineering and management. Since then, however, no studies have been published about use of neural networks for construction labor productivity modeling.

1.13 Construction Labor Productivity Modeling with Neural Networks and Regression

Analysis

Model fitting for construction labor productivity data requires quantification of the effects of factors on labor productivity and quantification of the interactions among the factors. This task of

identifying a mapping function from the independent variables to the dependent variables is analogous to that performed by some of the neural network models such as backpropagation. In statistics, regression analysis is the most common method to explore this relationship. The advantage of regression models lies in their generally, more parsimonious use of free parameters than the neural networks. Regression models require the user to decide a-priori on the class of relationships (linear, quadratic etc.) to be used in modeling. In the common use of neural network models, on the other hand, apart from the choice of a neural network architecture (which constrains the class of the models or the functions that can be learned), the user need not exert much effort to decide about the class of relationships, and can let the training algorithm do the work. However, it must be pointed out that many of the neural network approaches to model fitting are closely related to their statistical counterparts. A pragmatic approach, therefore, is to use a mix of tools and techniques drawn from both neural networks and statistical approaches for complex real world applications such as construction productivity modeling. This was the focus of this study.

CHAPTER 2. CONSTRUCTION LABOR PRODUCTIVITY MODELING WITH REGRESSION ANALYSIS AND NEURAL NETWORKS

2.1 Research Objective

The main goal of this study was to develop a methodology for modeling construction labor productivity of different tasks, to improve the present factor modeling methodology in terms of closeness of fit and prediction performance.

2.2 Description of the Data

The data of this study were compiled from eight projects of a building contractor, during 1992-1994 time frame. The projects were all located in Iowa and had a range between one to sixty three million dollars. Five projects were located in Des Moines, one in Ankeny, one in Newton, and one in Johnson. The data was compiled from the main frame database of the contractor and then transferred to a comma delimited ASCII file format. The database of the contractor did not directly include productivity data; however, labor, quantity and equipment data were kept in three separate databases for other purposes. The labor database (Table 2.1) included data of total weekly regular and overtime workhours that each employee spent on a particular task. Employee data was recorded daily, but cumulative weekly data for a task was put in to the labor database. The craft of the workers were also included in the labor database.

The second database; the quantity database included data of actual quantities completed for various tasks (Table 2.2). The quantity database was updated weekly, at the same day, when the labor database was updated. The task numbers of the quantity database were identical to the tasks numbers for the labor database.

The third database; the equipment database had data of equipment for all the projects in progress. The data on the concrete pump was obtained from the equipment database.

Table 2.1. Structure of the Labor Database of the Contractor

| Employee | Craft | Project No | Task No | Task Description | Type of Wokhours | Week Ending | Total Hours |
|------------|-----------|------------|---------|-----------------------|------------------|-------------|-------------|
| Employee-1 | Laborer | 1 | 033110 | Place Column Concrete | Regular | 930421 | 9 |
| Employee-2 | Laborer | 1 | 033110 | Place Column Concrete | Regular | 930421 | 7 |
| Employee-3 | Laborer | 1 | 033110 | Place Column Concrete | Regular | 930421 | 6.5 |
| Employee-4 | Carpenter | 1 | 033110 | Place Column Concrete | Regular | 930421 | 1 |

Table 2.2. Structure of the Quantity Database of the Contractor

| Project No | Week Ending | Task No | Task Description | Quantity |
|------------|-------------|---------|-----------------------|----------|
| 1 | 930421 | 033110 | Place Column Concrete | 32 |
| 1 | 930428 | 033110 | Place Column Concrete | 16 |
| 1 | 930505 | 033110 | Place Column Concrete | 19 |
| 1 | 930512 | 033110 | Place Column Concrete | 13 |
| 1 | 930519 | 033110 | Place Column Concrete | 16 |
| 1 | 930609 | 033110 | Place Column Concrete | 3.5 |

Data of the independent variable (Production Rate) and dependent variables excluding weather variables were compiled from the labor, quantity and equipment databases. Cumulative regular and overtime workhours, total number of workers, total number of laborer of each task were calculated for every week of a project, if there was a task in progress (Table 2.3). Production Rate (PR) was calculated by dividing total quantities completed by the total workhours spent for a task. Percentage overtime was obtained by dividing the total overtime hours by the total workhours. Percentage laborer was calculated in a similar way by dividing the total hours spent by the laborer craft only to the total work hours. Cumulative quantities were calculated by summing the quantities from the first week of construction until the week of interest.

The database of the contractor included several tasks, but only for a limited number of tasks the quantities were updated weekly. For some of the tasks that had weekly quantity updates the total number of data points obtained from eight projects were very few. There were four tasks which had weekly data and had sufficient amount of data points. These tasks were concrete pouring (P), formwork (F), concrete finishing (T), and granular fill (G). The data for the concrete pouring task included three job types; column (C), slab on grade (S), and walls over eight feet (W). The data for

Table 2.3. Combined Workhour, Quantity, and Equipment Data

| Week Ending | 930421 | 930428 | 930505 | 930512 |
|------------------------|--------|--------|--------|--------|
| Quantity(Cy) | 32 | 16 | 19 | 13 |
| Manhours(Hr) | 23.5 | 14 | 15 | 12 |
| Production Rate(Cy/Hr) | 1.36 | 1.14 | 1.27 | 1.08 |
| Number of Worker | 4 | 4 | 4 | 4 |
| Number of Laborer | 3 | 3 | 4 | 4 |
| Overtime Hours | 0 | 0 | 0 | 0 |
| Concrete Pump | 0 | 0 | 0 | 0 |
| Job Type | C | C | C | C |
| Project | 1 | 1 | 1 | 1 |

the formwork task also included three job types; column, grade beam (B), and walls over eight feet. Data for concrete finishing and granular fill tasks included only one job type. The only job type for concrete finishing was trowel finish slab, and for granular fill it was fine granular fill.

Temperature, percent relative humidity, and precipitation data for Des Moines were obtained from the local climatological publications of the national climatic data center (Table 2.4). The data consisted of observations of temperature in degrees Fahrenheit (F), and percentage relative humidity for three hour intervals, and cumulative amounts of precipitation in inches for one hour intervals.

Table 2.4. Calculation of Weekly Average Temperature and Precipitation Values

| Day | | 6:01-9:00am | 9:01-12:00am | 12:01-3:00pm | 3:01-6:00pm |
|--------|-----------------------------|-------------|--------------|--------------|-------------|
| 930415 | Temperature(F) | 35 | 37 | 36 | 35 |
| | Relative Humidity(%) | 92 | 89 | 93 | 92 |
| | Temperature(F) | 41 | 50 | 55 | 55 |
| 930416 | Relative Humidity(%) | 68 | 46 | 30 | 30 |
| | Temperature(F) | 53 | 58 | 64 | 59 |
| | Relative Humidity(%) | 96 | 97 | 58 | 65 |
| 930419 | Temperature(F) | 40 | 48 | 51 | 48 |
| | Relative Humidity(%) | 65 | 48 | 41 | 42 |
| | Temperature(F) | 46 | 56 | 57 | 55 |
| 930420 | Relative Humidity(%) | 52 | 25 | 26 | 29 |
| | Week Ending Temperature(F) | Average | | | 48.95 |
| | 930421 Relative Humidity(%) | Average | | | 59.2 |

Table 2.6. Labor Productivity Factors

| Group | Factors | Concrete Pouring | Formwork | Concrete Finishing | Granular Fill |
|----------------------------|-------------------------------|------------------|----------|--------------------|---------------|
| A. Job Complexity | 1. Quantities completed (q) | X | X | X | X |
| | 2. Job type | X | X | | |
| B. Crew Size & Composition | 3. Number of workers (n) | X | X | X | X |
| | 4. % Laborer (l) | X | X | X | X |
| C. Repetition | 5. Cumulative quantities (cq) | X | X | X | X |
| D. Weather | 6. Temperature (t) | X | X | X | X |
| | 7. Humidity (h) | X | X | X | X |
| | 8. Precipitation (p) | X | X | X | X |
| E. Equipment | 9. Concrete pump (u) | X | | | |
| F. Motivation & Fatigue | 10. % Overtime (o) | X | X | X | X |

2.3 Limitations of the Data

The data of this study, that were compiled from eight projects covering a two-year time frame included four tasks and ten productivity factors. However the data had limitations. The data were compiled weekly, daily variations in the production rate and factors were not included. The number of workers factor was the total number of workers participated in a task during a week. The number of workers was related to crew size, but it might not be actually equal to the daily crew size.

The weather data was limited to the weekly data for Des Moines, although three projects were not located at Des Moines. However, the locations of these three projects were all in Iowa, usually with similar weather conditions. The factors included in this study were limited to the ones given in Table 2.6.

2.4 Modeling Methodology

The modeling methodology for this study consisted of four stages: Data identification, regression analysis, neural network modeling, and model comparison (Figure 2.1). In the data identification stage average production rates for different job types were calculated. Plots of factors versus production rate for each job types of tasks were also included in the data identification stage. The purpose of the data identification was to identify the factors that might have an effect on production rate. Results of previous construction labor productivity studies were also used in combination with the data identification stage to define the initial regression model that has all the factors (of Table 2.6) that might have a significant effect on the production rate for a task. Next, the factors that did not significantly improve the regression model were dropped from the model at the regression analysis stage. The factors that were used in the final regression model were used to develop the neural network models. Several neural network models that have different characteristics were developed to improve prediction performance of the neural networks, at the neural network modeling stage. Parsimonious models were considered for the regression, and neural network models. A parsimonious model fits the data adequately, without using any unnecessary parameters. The principle of parsimony is important because in practice parsimonious models generally produce better forecasts (Pankartz 1983). Finally, at the model comparison stage, the results of regression and neural networks were compared with the results of the productivity models available in literature. Data identification, regression analysis, neural network modeling, and model comparison stages were discussed in detail in sections 2.5 to 3.1.

2.5 Data Identification

The first part of the data identification consisted of determination of minimum, maximum, average, and standard deviation values of production rate for different job types of the tasks (Table 2.7). The minimum, maximum, and standard deviation values indicated that there was a high

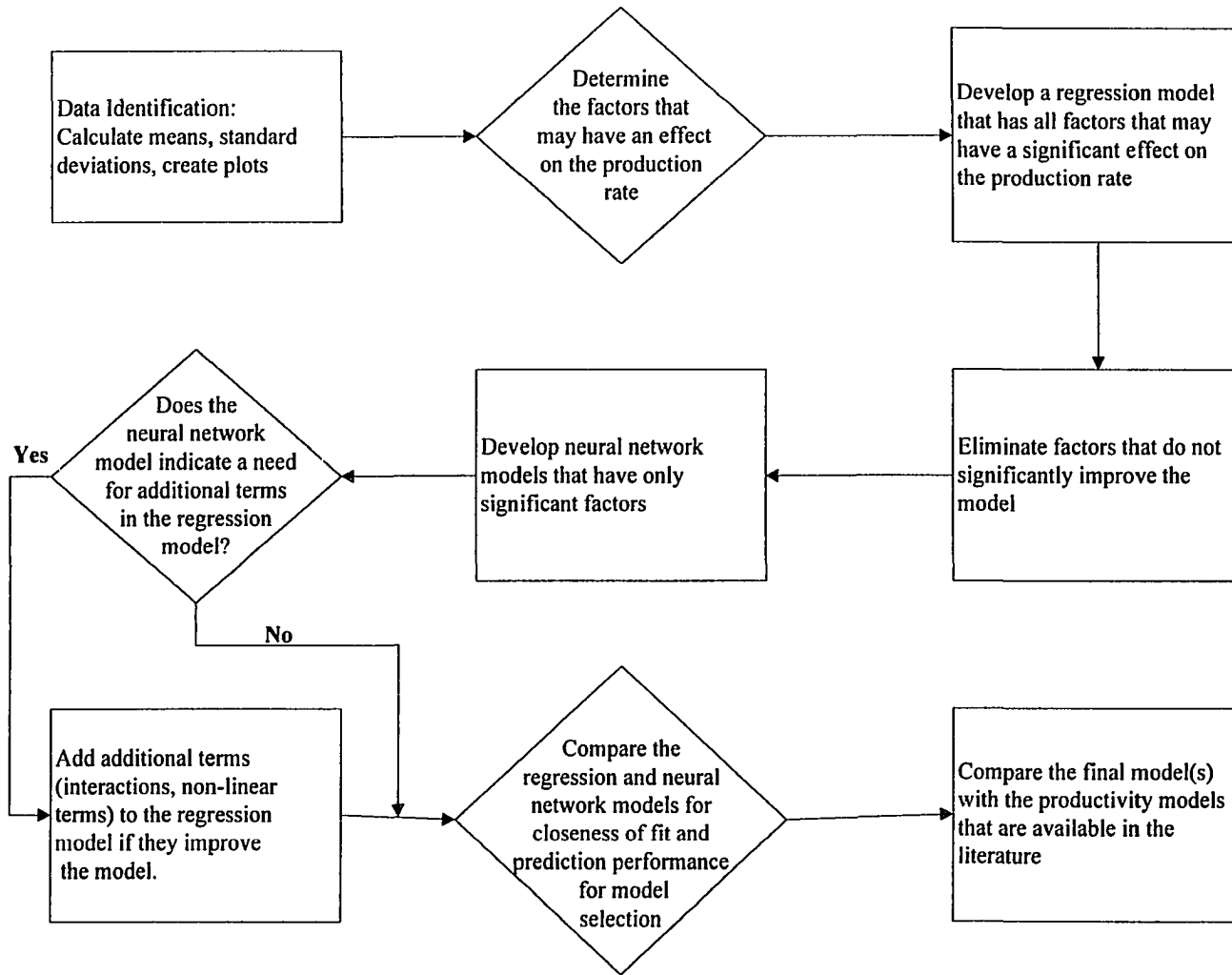


Figure 2.1. Flow Chart of Modeling Methodology

Table 2.7 Production Rate Statistics for Different Job Types

| Task | Job Type | Min. | Max. | Mean | Std.* | Ndp** | Unit PR |
|-----------------------|---------------|-------|--------|--------|--------|-------|---------|
| Concrete Pouring(P) | Column(C) | 0.42 | 3.50 | 1.28 | 0.65 | 21 | cy/hr |
| | Slab(S) | 0.22 | 11.41 | 2.45 | 2.83 | 24 | |
| | Slab Pump(SP) | 0.67 | 11.69 | 3.48 | 1.90 | 34 | |
| | Wall(W) | 0.58 | 10.38 | 4.09 | 2.16 | 33 | |
| Formwork(F) | Column | 1.93 | 15.15 | 8.70 | 3.69 | 21 | sf/hr |
| | Grade Beam(B) | 0.86 | 29.46 | 12.27 | 8.74 | 20 | |
| | Wall | 1.83 | 18.83 | 8.71 | 3.55 | 35 | |
| Concrete Finishing(T) | Trowel Finish | 24.38 | 406.25 | 197.17 | 114.20 | 46 | sf/hr |
| Granular Fill (G) | Fine Granular | 13.11 | 405.40 | 140.30 | 94.91 | 33 | sf/hr |

* Standard deviation

** Number of data points

amount of variation in the production rate values. The mean and standard deviations of production rate for different job types of concrete pouring and formwork indicated that job type may be a factor that has an effect on the production rate.

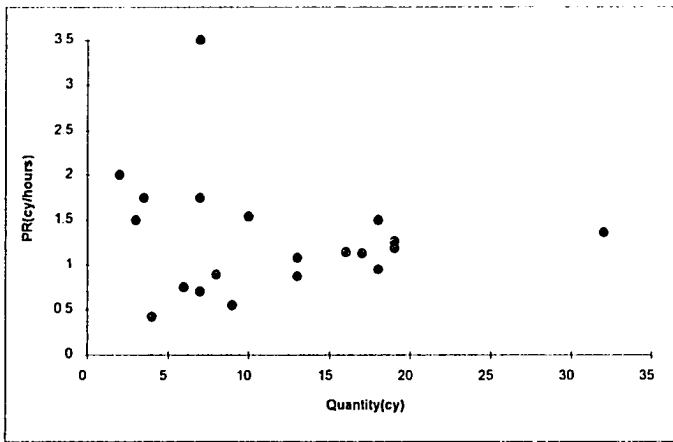
Data of production rate versus all of the factors were plotted to identify the factors that may influence the production rate. Plots of quantity, number of workers, and temperature versus production rate for all of the four tasks were made. Plots of these factors were given, as these factors were included in the initial regression models of all of the tasks. Plots were made for each job type of the concrete pouring and formwork tasks because of two reasons. The first reason was the differences in the levels of production rates between the job types. The second reason was the

possible differences among the effects of the factors on production rate for different job types. The plots were given in Figures 2.2 to 2.13.

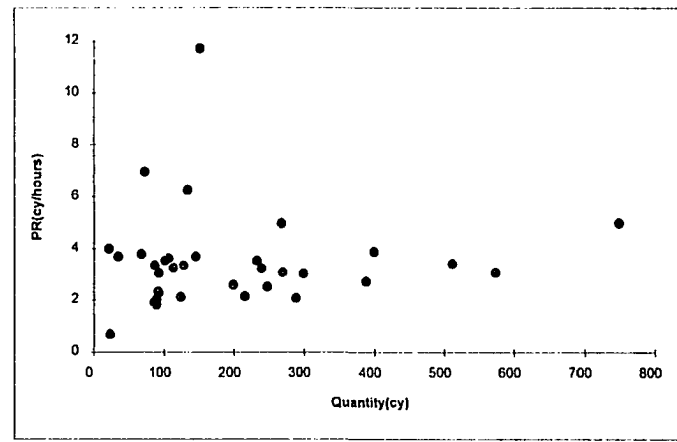
The effects of quantity on production rate for all the four tasks were easily identified from the quantity versus production rate plots; production rate increased when the amount of weekly quantities increased. Number of workers versus production rate plots indicated a possible decrease in the production rate as the number of workers increased, especially when the number of workers were more than five. However this relation was not very significant for the formwork task. The temperature versus production rate plots indicated an increase in production rate due to an increase in temperature for concrete pouring, and concrete finishing tasks.

2.6 Regression Analysis

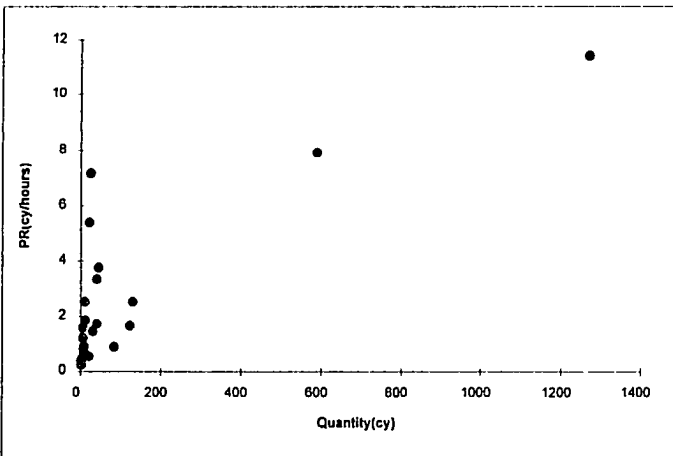
The factors that were identified at the data identification stage were used to develop the initial regression models. The initial regression models developed for each of the four tasks were pure linear regression models that included all of the factors that might have an effect on the production rate. The factors were the independent variables, and production rate was the dependent variable in the regression models. Quantity and number of workers were included in all of models as independent variables, because possible influences of the two factors were identified for all of the tasks from the plots. Temperature was also included as independent variable in the initial regression models of the all four tasks, because previous research suggested influence of temperature on labor productivity (Section 1.5). However, only for concrete pouring and concrete finishing, influences of temperature on production rate were identified at the data identification stage. The factors other than quantity, number of workers, and temperature were included in the initial regression models of the tasks if the data identification stage suggested possible effects of the factors on production rate.



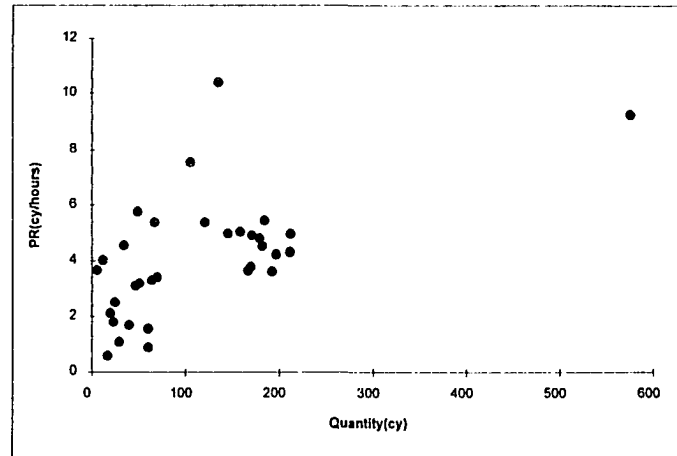
(a)



(b)

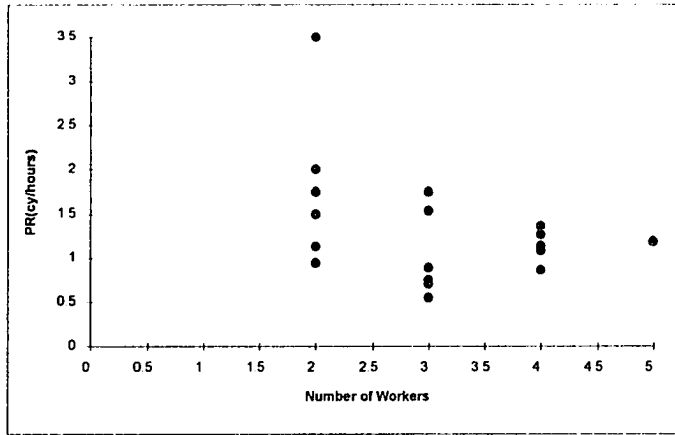


(c)

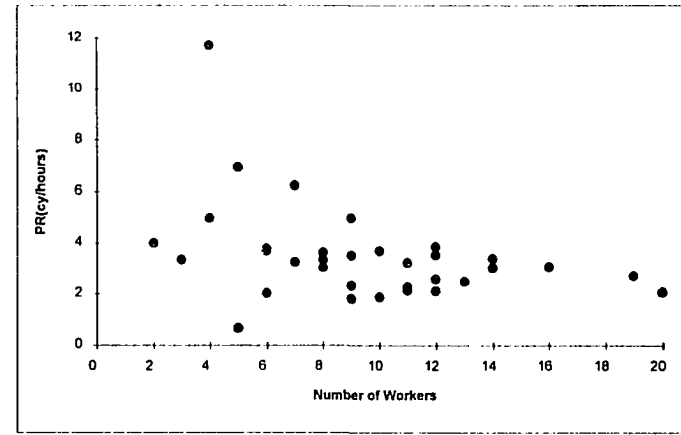


(d)

Figure 2.2. Plot of Quantity versus PR for Pour (a) Column, (b) Slab (Pump), (c) Slab, (d) Wall



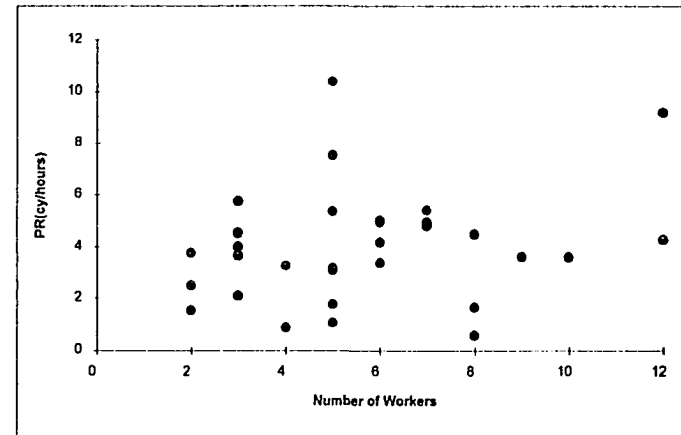
(a)



(b)

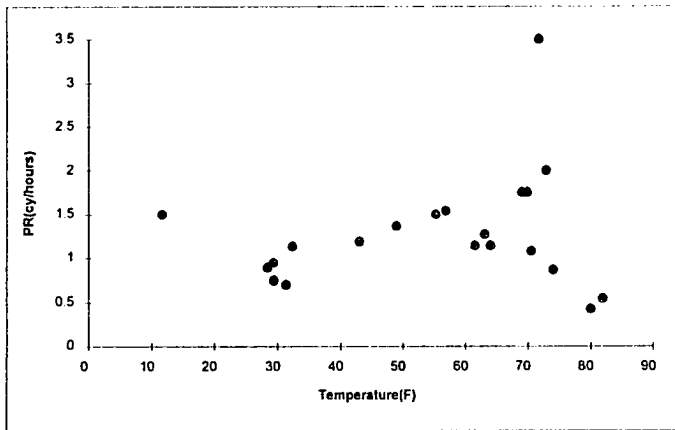


(c)

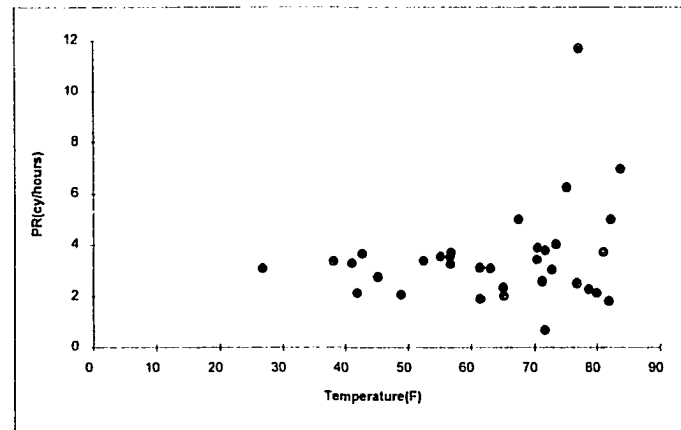


(d)

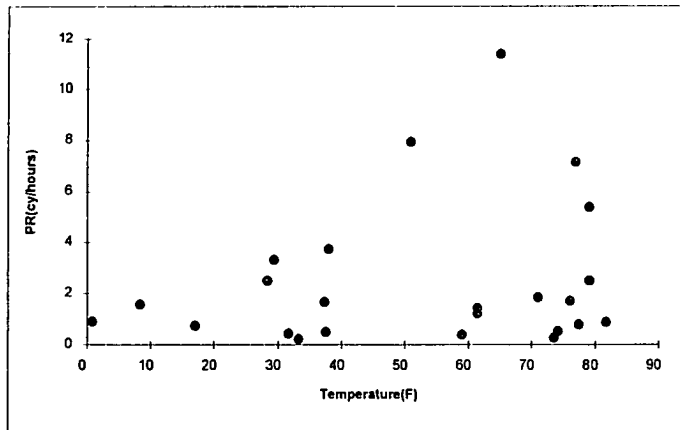
Figure 2.3. Plot of Number of Workers versus PR for Pour (a) Column, (b) Slab (Pump), (c) Slab, (d) Wall



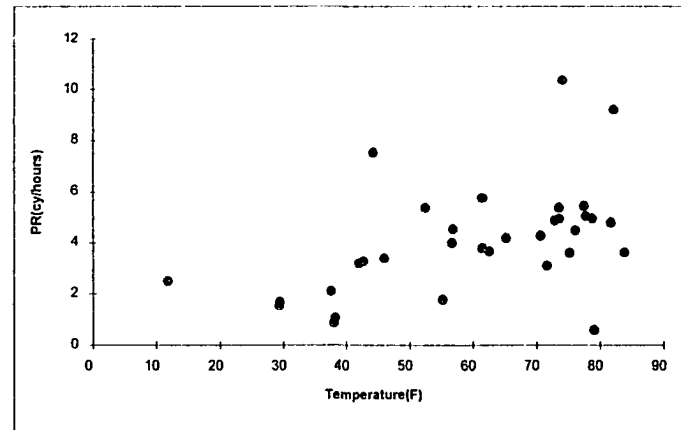
(a)



(b)

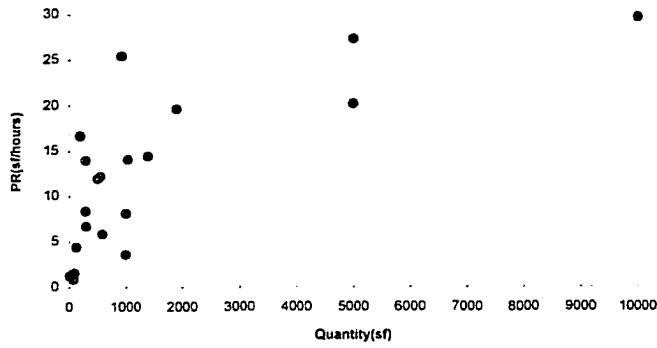


(c)

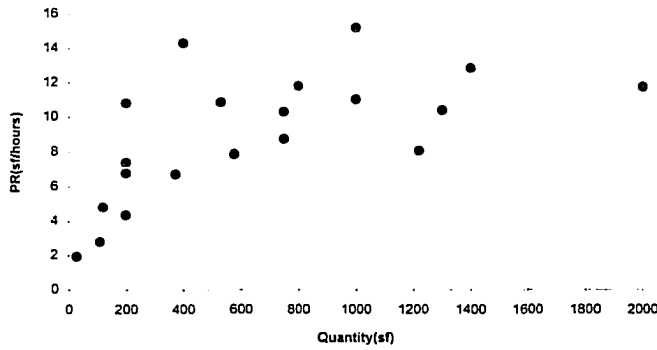


(d)

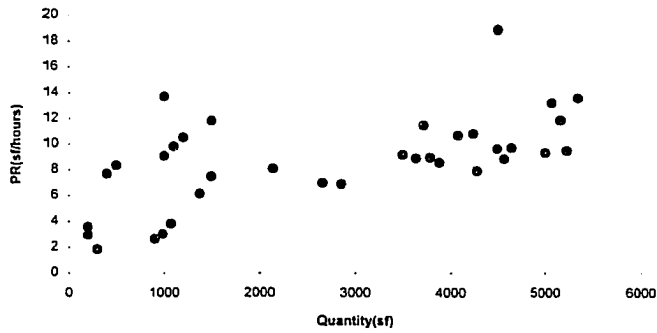
Figure 2.4. Plot of Temperature versus PR for Pour (a) Column, (b) Slab (Pump), (c) Slab, (d) Wall



(a)

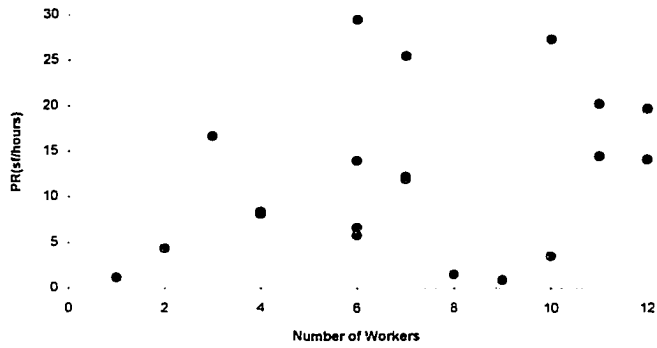


(b)

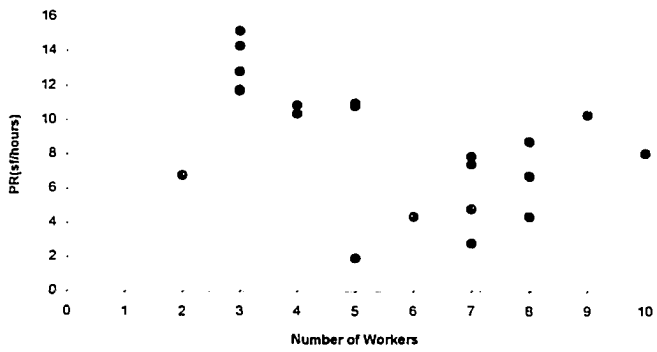


(c)

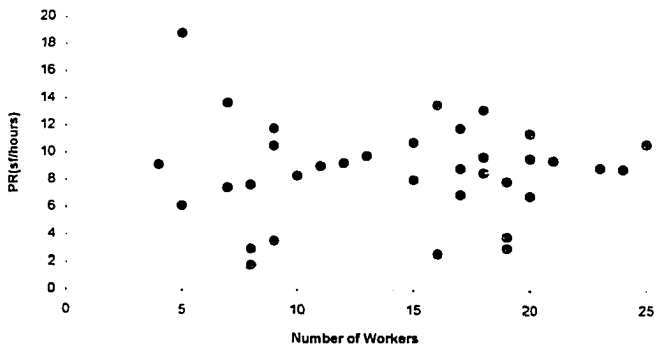
Figure 2.5. Plot of Quantity versus PR for Form (a) Grade Beam, (b) Column, (c) Wall



(a)

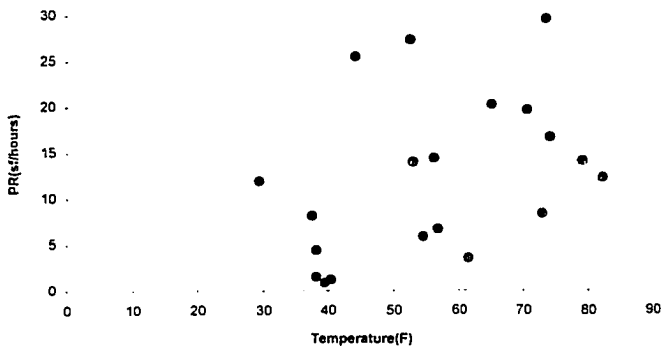


(b)

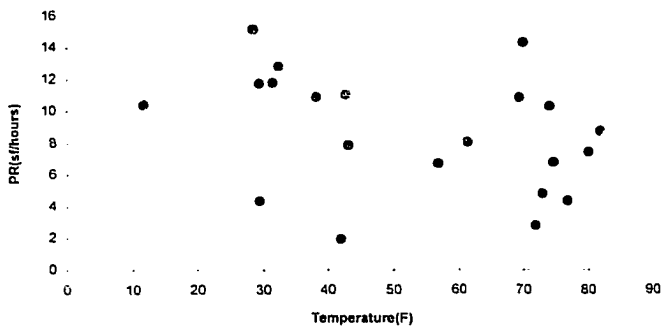


(c)

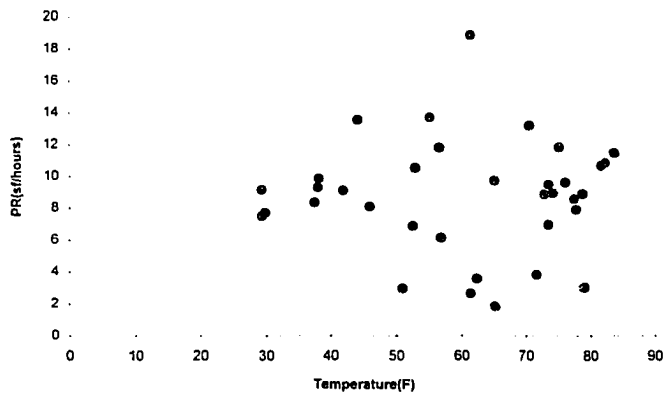
Figure 2.6. Plot of Number of Workers versus PR for Form (a) Grade Beam, (b) Column, (c) Wall



(a)



(b)



(c)

Figure 2.7. Plot of Temperature versus PR for Form (a) Grade Beam, (b) Column, (c) Wall

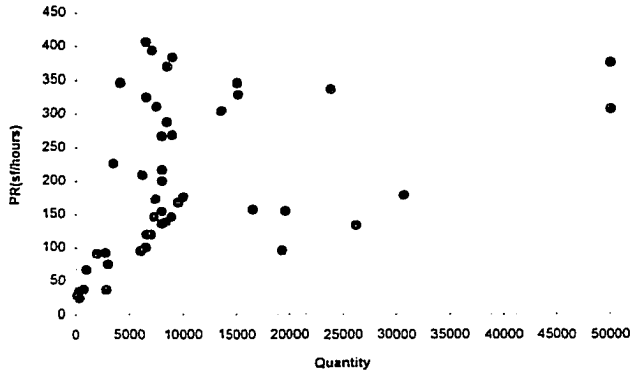


Figure 2.8 Plot of Quantity versus PR for Trowel Finish Slab



Figure 2.9 Plot of # of Workers versus PR for Trowel Finish Slab

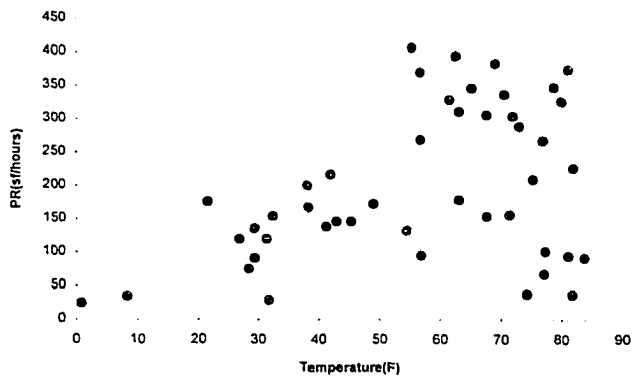


Figure 2.10 Plot of Temperature versus PR for Trowel Finish Slab

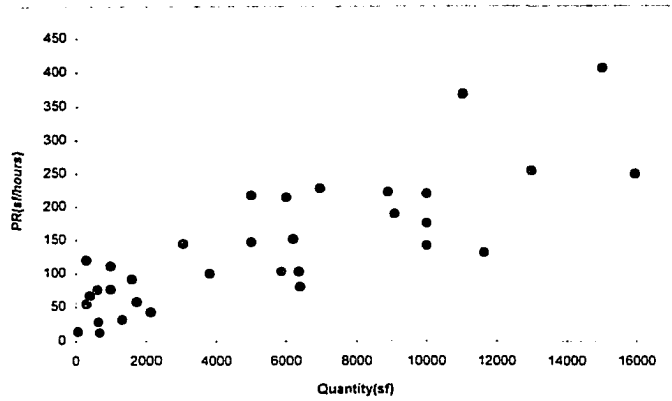


Figure 2.11 Plot of Quantity versus PR for Fine Granular Fill

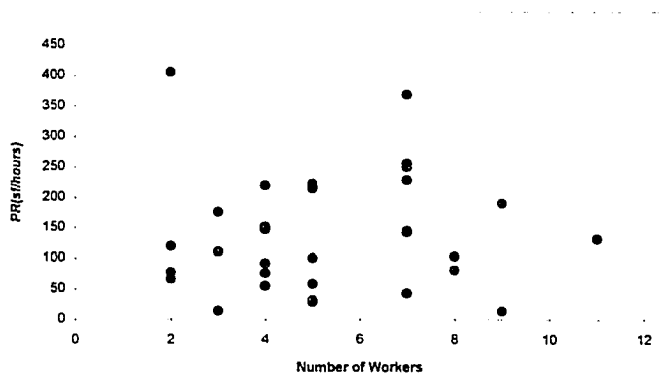


Figure 2.12 Plot of # of Workers versus PR for Fine Granular Fill

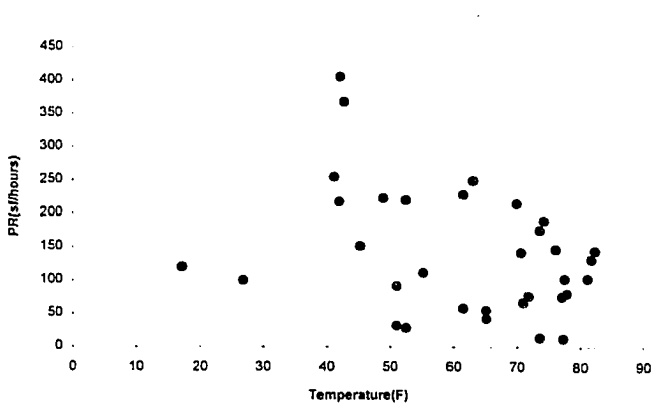


Figure 2.13 Plot of Temperature versus PR for Fine Granular Fill

Once the factors that may have possible effects on production rate for each of the tasks were identified, initial regression models for the tasks were developed. Next, the factors that did not significantly improve the model were dropped from the model. This was decided by dropping one factor at a time. P value of the factor in the regression model, comparison of closeness of fit, and prediction performance of the models with and without the factor were the criteria used to determine the significance of the factor in the model. Two error measures were used to compare closeness of fit and prediction performance of the models: Mean squared error (MSE), and mean absolute percent error (MAPE). MSE and MAPE were calculated by:

$$MSE = \frac{1}{n} \sum_{i=1}^n (Actual_i - Predicted_i)^2 \quad (2.1)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|Actual_i - Predicted_i|}{Actual_i} \times 100 \quad (2.2)$$

where n is the total number of data points for the model. Using MSE and MAPE together gives a better picture of closeness of fit and prediction performance because there was a high variation in the production rate values.

A procedure based on the cross validation technique was developed to compare prediction performance of the models. The procedure can be summarized in the following steps: (1) A project was selected as the test sample and a new data set was formed. The new data set included data of all of the remaining projects but not the data of the project that was selected as the test sample. (2) Model parameters for the model which is being evaluated for prediction performance were calculated with the new data set. (3) The model with the new parameters were used to predict the production rate values of the project which was selected as the test sample. Squared error and absolute percent error values of the model predictions for test sample were calculated. (4) All of the projects were selected as the test sample, one at a time, and steps 1-3 were repeated for each of the test samples.

(5) MSE and MAPE values were calculated by averaging the squared error, and absolute percent error values of the all test samples to compare prediction performance of the model being evaluated.

The procedure described to compare prediction performance requires determination of parameters for several models. However if only one data set was selected as the test sample, the chance of prediction performance of a model being better than prediction performance of another model due to randomness would increase, especially with the limited amount of data used in the study.

2.6.1 Regression Models for Concrete Pouring

Quantity, number of workers, temperature, overtime, and job type were the factors that were identified which might have effects on the production rate of concrete pouring. The initial model for concrete pouring was in the following form:

$$PR(P) = \alpha + \beta_1q + \beta_2n + \beta_3t + \beta_4o + \beta_5s + \beta_6u + \beta_7w \quad (2.3)$$

where; PR (P) was the predicted production rate for concrete pouring in cy/hours; α was the regression constant; β_1 to β_7 were regression coefficients for the factors; q was the weekly quantities in cubic yards; n was the number of workers; t was the temperature in Fahrenheit degrees; o=1 if there was overtime, o=0 otherwise; s=1 if job type was slab without pump, s=0 otherwise; u=1 if job type was slab, and concrete pump was present, u=0 otherwise; w=1 if job type was wall, w=0 if job type was column or slab. Regression statistics of model RP-1 are given in Table 2.8.

The overtime(o) term which had a P-value of 0.266 was dropped from the initial model to obtain model RP-2 (Table 2.9). The prediction performance of model RP-2 was better than the model RP-1. Next the slab (s) term which had P-value of 0.103 in model RP-2 was eliminated from the model RP-2 to obtain model RP-3. The closeness of fit, and prediction performance of RP-3 was similar to RP-2, but RP-3 was preferred over RP-2 because it was more parsimonious. Elimination of the

Table 2.8 Regression Statistics for Model RP-1

| | | | | | | |
|--------------|--------------|----------------|--------|---------|----------------|-----------|
| R Square | .530 | | | | | |
| Observations | 112 | | | | | |
| ANOVA | | | | | | |
| | df | SS | MS | F | Significance F | |
| Regression | 7 | 305.86 | 43.69 | 16.73 | 1.26E-14 | |
| Residual | 104 | 271.63 | 2.61 | | | |
| Total | 111 | 577.49 | | | | |
| | Coefficients | Standard Error | t Stat | P-value | Lower 95% | Upper 95% |
| Intercept | .508 | .600 | .847 | .3991 | -.682 | 1.698 |
| q | .008 | .001 | 7.583 | .0001 | .006 | .010 |
| n | -.200 | .064 | -3.125 | .0021 | -.327 | -.073 |
| t | .024 | .008 | 2.897 | .0051 | .007 | .040 |
| o | -.520 | .465 | -1.118 | .2661 | -1.441 | .402 |
| s | .851 | .494 | 1.723 | .0881 | -.129 | 1.832 |
| u | 2.173 | .578 | 3.762 | .0001 | 1.028 | 3.319 |
| w | 2.757 | .537 | 5.138 | .0001 | 1.693 | 3.821 |

Table 2.9. Steps through Model RP-1 to Model RP

| Step | Model | R ² | Closeness of Fit | | Prediction Performance | | Operation | Term | P-value |
|------|-------|----------------|------------------|------|------------------------|------|-----------|------|---------|
| | | | MSE | MAPE | MSE | MAPE | | | |
| i | RP-1 | .530 | 2.42 | 69.7 | 4.27 | 83.6 | drop | o | .266 |
| ii | RP-2 | .524 | 2.45 | 68.8 | 4.03 | 80.7 | drop | s | .103 |
| iii | RP-3 | .512 | 2.52 | 66.0 | 4.13 | 77.0 | drop | t | .009 |
| iv | RP-4 | .479 | 2.68 | 65.1 | 4.19 | 84.0 | add | t | |
| v | RP-3 | .512 | 2.52 | 66.0 | 4.13 | 77.0 | add | q*u | |
| vi | RP | .569 | 2.22 | 61.3 | 3.96 | 71.7 | | | |

temperature term from model RP-3 resulted in a worse prediction performance, and closeness of fit, so model RP-3 was preferred over model RP-4.

The sensitivity analysis of the neural network model for concrete pouring indicated possible interaction between quantity and concrete pump (u) terms. The rate of increase in production rate due to increase in the level of quantity was observed to be different for concrete pouring with pump than the rate of increase for concrete pouring without pump. The q*u interaction term improved both the closeness of fit, and prediction performance of the regression model RP-3. The model designated RP was selected as the final regression model.

2.6.2 Regression Models for Formwork

Quantity, number of workers, temperature, precipitation, and job type were the factors that were identified, which might have influenced the production rate of the formwork task. The first regression model of formwork (RF-1) was in the following form:

$$PR(F) = \alpha + \beta_1 q + \beta_2 n + \beta_3 t + \beta_4 p + \beta_5 b + \beta_6 w \quad (2.5)$$

where PR(F) was the predicted production rate for formwork in sf/hours; b=1 if job type was grade beam, b=0 otherwise; α was the regression constant; q,n,t,p,w were same as the variables defined by equation 2.4. The regression statistics of model RF-1 are given in Table 2.10.

The temperature term which had the highest P-value in model RF-1 was dropped from model RF-1 to obtain model RF-2 (Table 2.11). Because model RF-2 had better prediction performance than model RF-1 with a similar closeness of fit, elimination of terms were continued with wall term. Model RF-3 were similar to model RF-2 in terms of prediction performance, and closeness of fit, but was preferred because it was more parsimonious than model RF-2. Elimination of the precipitation term from Model RF-3 improved prediction performance with an insignificant loss of closeness of fit. Next the grade beam term was dropped from model RF-4 to obtain model RF-5. Model RF-5

was preferred over Model RF-4, because it had a similar prediction performance and closeness of fit with fewer parameters. Finally model RF-6 was obtained by dropping the number of workers term from model RF-5. Model RF-5 was selected as the final regression model for formwork (RF) because, model RF-6 had a poor prediction performance, with a significant loss of closeness of fit compared to model RF-5.

2.6.3 Regression Models for Concrete Finishing

The initial regression model for concrete finishing (RT-1) included quantity, number of workers, temperature and precipitation as the independent variables, and was in the following form:

$$PR(T) = \alpha + \beta_1q + \beta_2n + \beta_3t + \beta_4p \quad (2.5)$$

where; PT(T) was the predicted production rate for concrete finishing in sf/hours. The percentage of variation explained by the initial regression model (R^2) for concrete finishing (Table 2.12) was less

Table 2.12. Regression Statistics for Model RT-1

| | | | | | | |
|--------------|--------------|----------------|----------|---------|----------------|-----------|
| R Square | .308 | | | | | |
| Observations | 46 | | | | | |
| ANOVA | | | | | | |
| | df | SS | MS | F | Significance F | |
| Regression | 4 | 180876 | 45218.99 | 4.57 | 0.003838 | |
| Residual | 41 | 406011 | 9902.71 | | | |
| Total | 45 | 586887 | | | | |
| | Coefficients | Standard Error | t Stat | P-value | Lower 95% | Upper 95% |
| Intercept | 224.266 | 132.734 | 1.690 | .099 | -43.797 | 492.329 |
| q | .004 | .002 | 2.683 | .010 | .001 | .007 |
| n | -10.910 | 6.147 | -1.775 | .083 | -23.324 | 1.504 |
| t | 1.507 | .841 | 1.792 | .081 | -.192 | 3.206 |
| p | -1.399 | 1.449 | -.965 | .340 | -4.325 | 1.528 |

than percentage of variation explained by the initial regression models for concrete pouring, and formwork.

Terms that did not improve the initial regression model of concrete finishing task were eliminated one at a time (Table 2.13). Elimination of precipitation, and number of workers terms improved prediction performance, without reducing closeness of fit significantly (RT-3). However elimination of the temperature term resulted in a model that had a poor prediction performance, and a poor closeness of fit; therefore, temperature term was kept in the model. Sensitivity analysis of the neural network model for concrete finishing and low R^2 value of the model suggested that only a linear term for temperature may not be sufficient. The polynomial term t^2 was added to the model RT-3 to obtain next model; RT-4. But model RT-4 predicted negative production rate values at certain levels of temperature. The problem was eliminated by dropping the regression constant α from model RT-4. The final regression model; model RT had a good prediction performance, and closeness of fit compared to other regression models, and was selected as the adequate regression model for concrete finishing.

Table 2.13. Steps through Model RT-1 to Model RT

| Step | Model | R^2 | Closeness of Fit | | Prediction Performance | | Operation | Term | P-value |
|------|-------|-------|------------------|------|------------------------|-------|-----------|----------|---------|
| | | | MSE | MAPE | MSE | MAPE | | | |
| i | RT-1 | .308 | 8826 | 80.9 | 28350 | 161.0 | drop | p | .340 |
| ii | RT-2 | .292 | 9026 | 77.8 | 23493 | 135.8 | drop | n | .112 |
| iii | RT-3 | .248 | 9593 | 72.5 | 20304 | 118.6 | drop | t | .019 |
| iv | RP-4 | .168 | 10609 | 82.6 | 21391 | 127.0 | add | t | |
| v | RF-3 | .248 | 9593 | 72.5 | 20304 | 118.6 | add | t^2 | |
| vi | RT-5 | .307 | | | | | drop | α | .510 |
| vii | RT | .301 | 8920 | 69.5 | 18831 | 107.6 | | | |

2.6.4 Regression Models for Granular Fill

Quantity, number of workers, temperature and precipitation variables were included in the initial regression model for granular fill (RG-1). The model was in the following form:

$$PR(G) = \alpha + \beta_1q + \beta_2n + \beta_3t + \beta_4p \quad (2.6)$$

where: PT(6) was the predicted production rate for granular fill in sf/hours. Regression statistics of model 2.6 are given in Table 2.14. The temperature term was the first term dropped from the model. Model RG-2, the model without temperature term was similar to model RG-1 in terms of predictive performance, and closeness of fit (Table 2.15). RG-2 was preferred over RG-1 because it was more parsimonious. Elimination of precipitation term resulted in model RG-3 which had poor prediction performance, and closeness of fit compared to model RG-2. Model RG-2 was selected as the final regression model for granular fill (RG).

Table 2.14. Regression Statistics for Model RG-1

| | | | | | | |
|--------------|--------------|----------------|----------|---------|----------------|-----------|
| R Square | .797 | | | | | |
| Observations | 33 | | | | | |
| ANOVA | | | | | | |
| | df | SS | MS | F | Significance F | |
| Regression | 4 | 229742 | 57435.49 | 27.49 | 2.45E-09 | |
| Residual | 28 | 58503 | 2089.38 | | | |
| Total | 32 | 288245 | | | | |
| | Coefficients | Standard Error | t Stat | P-value | Lower 95% | Upper 95% |
| Intercept | 158.295 | 33.047 | 4.790 | .000 | 90.601 | 225.988 |
| q | .020 | .002 | 9.074 | .000 | .015 | .024 |
| n | -11.901 | 4.568 | -2.605 | .015 | -21.259 | -2.544 |
| p | -.799 | .577 | -1.384 | .177 | -1.981 | .384 |
| t | -56.171 | 30.945 | -1.815 | .080 | -119.558 | 7.217 |

Table 2.15. Steps through Model RG-1 to Model RG

| Step | Model | Closeness of Fit | | | Prediction Performance | | Operation | Term | P-value |
|------|----------|------------------|------|------|------------------------|------|-----------|------|---------|
| | | R ² | MSE | MAPE | MSE | MAPE | | | |
| i | RG-1 | .797 | 1773 | 33.6 | 2648 | 51.6 | drop | t | .177 |
| ii | RG-2 | .783 | 1894 | 33.6 | 2881 | 49.6 | drop | n | .024 |
| iii | RG-3 | .741 | 2263 | 41.1 | 3692 | 76.2 | add | n | |
| iv | RG-2(RG) | .783 | 1894 | 33.6 | 2881 | 49.6 | | | |

2.7 Neural Network Modeling

Neural network models used in this study consisted of backpropagation neural network models which have the capability of identifying a mapping function from the independent productivity variables (factors) to the dependent variable (production rate), and have been commonly used in the applications similar to construction productivity modeling (Sections 1.2 and 1.3). In the neural network models developed for the different tasks, only the factors that were determined to influence production rate significantly were included as input variables. These factors were defined through data identification and regression analysis stages. The architecture, and training algorithms of these models are discussed in Sections 2.7.1 to 2.7.5.

2.7.1 Feedforward Neural Networks and Back Propagation

Artificial neural networks are typically composed of interconnected units which serve as model neurons (Hinton, 1992). A multilayer feedforward neural network consists of a set of units (neurons) that are logically arranged into two or more layers (Figure 2.14). There is an input layer and an output layer, each containing at least one unit. Between input and output layers there are

and an output layer, each containing at least one unit. Between input and output layers there are usually one or more "hidden" layers. The term "feedforward" means that information flows in one direction only. The inputs to units in each layer come exclusively from the outputs of the units in previous layers, and outputs from these neurons pass exclusively to neurons in following layers.

Each connection between the input layer and a hidden unit has an associated weight W_{ij} . The net signal I_j to an individual hidden unit is expressed as the sum of connections between the input layer units, and that particular hidden unit plus the connection value W_{gj} from a bias node. This relationship may be expressed as:

$$I_j = \sum W_{ij}O_i + W_{gj},$$

where O_i is the signal produced by the input unit i . The signal from the hidden layer is processed

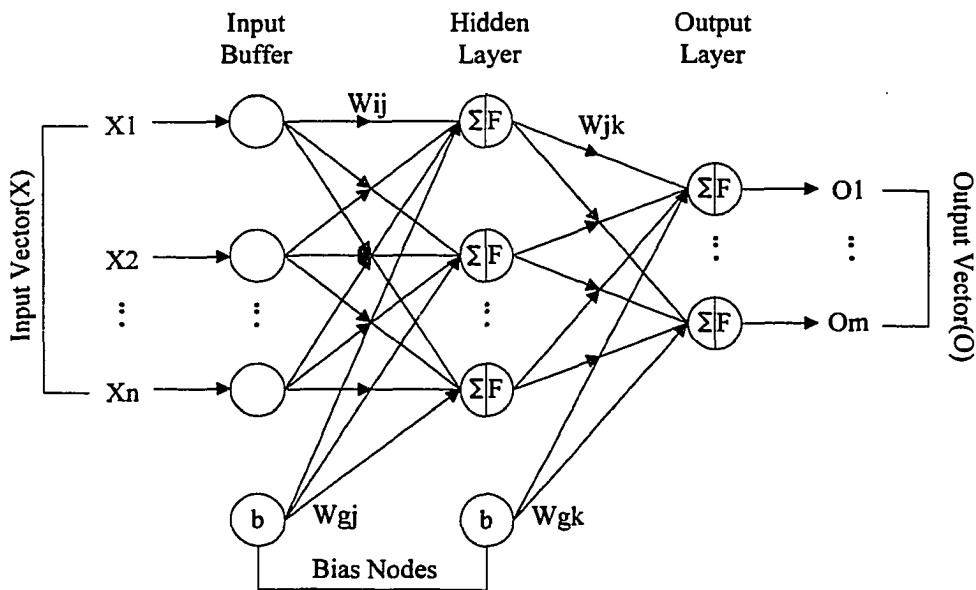


Figure 2.14 A Feedforward Neural Network Model with One Hidden Layer

with an activation function. In Rumelhart, Hinton and Williams (1986) network logistic function was used, which processed according to:

$$O_j = \frac{1}{1 + \exp(-I_j)}$$

The net signal to an output unit I_k is the sum of all connections between the hidden layer units and the respective output node, expressed as:

$$I_k = \sum W_{jk} O_j + W_{gk},$$

where W_{gk} represents a single connection weight from a bias unit. The net signal is again processed by logistic function to produce the final output value O_k , where:

$$O_k = \frac{1}{1 + \exp(-I_k)}$$

At the output layer the net signal O_k (estimated dependent variable) is compared to the actual value of the dependent variable, T_k , to produce an error signal. Rumelhart et al. (1986) used the "delta rule" to minimize the network error, and defined the process of weight adjustment by:

$$\Delta W_{jk}(n+1) = \eta \delta_{pk} O_{pj} + \alpha \Delta W_{jk}(n),$$

where η is the learning rate, and α is the momentum factor. The learning rate allows control on the magnitude of changes in weights. The momentum factor determines the effect of past weight changes on the current direction of movement in the weight space, and proportions the amount of the last weight change to be added into the new weight change.

The error signal δ is back-propagated to the connection weights between the hidden and output layers is defined as the difference between the target value T_{pk} for a particular input pattern p and the neural network's feed-forward calculations of the signal from the output layer O_k as:

$$\delta_{pk} = (T_{pk} - O_{pk}) O_{pk} (1 - O_{pk})$$

Then the connection weights between the input and hidden layers are changed by:

$$\delta_{pj} = O_{pj} (1 - O_{pj}) \sum_k \delta_{pk} W_{jk}$$

The data feed-forward and error back-propagation process (training) is continued until the desired accuracy or a certain number of iterations is reached.

The task in training is to determine a unique set of network weights (W 's) that enables the network to produce outputs (O 's) that match the set of target outputs (T 's), pertaining P training examples, when fed only with the respective inputs (X 's). When the desired mapping is achieved, the network is said to be knowledgeable about all P examples.

2.7.2 Learning Rate and Momentum Factor

One of the limitations of the backpropagation algorithm is the training speed. Several techniques have been proposed to overcome this. The simplest approach is to use a large learning rate coefficient η , however, this might result in high oscillations that may cause the algorithm to miss the global minimum of the network error. BrainMaker Professional Version 3.1, the neural network software that was used for this study includes a heuristic learning rate option which incorporates simple heuristics to dynamically adjust η during training. The heuristic learning rate option starts training with a large η of 1.0 and reduces η by a factor of 0.5, when the network error starts to fluctuate. Another limitation of the backpropagation algorithm is that it may be trapped in local minimum rather than the global minimum for the error function. But with heuristic learning rate algorithm large η can skip the local minimum to another point on the error surface, while small η can be used to go deeper in the valley of the global minimum.

Use of a momentum factor α is another procedure for improving the training speed of the backpropagation algorithm. The momentum factor is usually set to 0.9, adjusting the momentum factor has not been found to improve prediction performance (Pao 1989; BrainMaker 1993). The heuristic learning rate, and a momentum factor of 0.9 were used during training of all of the neural network models developed in this study. Training was stopped when the average squared error for the network converged, except for two of the neural network models described in Section 2.7.6.

2.7.3 Number of Hidden Units and Pruning

Number of hidden units is another parameter that needs to be decided before training can start. If too many hidden units are used the neural network may have a poor prediction (generalization) performance, however, if too few hidden units are used the neural network model may not have enough parameters to identify the mapping function. Defining the number of hidden units is highly problem dependent. Kolmogorov's Mapping Neural Network Existence Theorem states that any continuous function can be implemented with one hidden layer network structure using $2n+1$ hidden units; where n represents the number of input units (Kolmogorov 1957; Nielsen 1989). Use of $2n+1$ hidden units was also recommended by Caudill (1991).

$2n+1$ units were used for the initial neural network models of all the four tasks. However $2n+1$ units might be too many for certain tasks, which could result in a poor network prediction performance. One approach to decrease the number of parameters used in the neural network model is to decrease the number of hidden units. A similar approach is to remove the network connections which are not significantly contributing to the neural network model. The removal of hidden units or connections to decrease the number of parameters used in the neural network model is called pruning.

Le Cun (1989) discussed the importance of reducing the number of free parameters in a neural network to increase its likelihood of correct generalization without reducing the size of the network. Sietsma and Dow (1991) have studied generalization capabilities of different network configurations. Their experiments showed that networks (with input patterns that were not corrupted by noise) with fewer hidden units led to better generalization than the networks with more hidden units.

The number of hidden units for the second neural network models of each of the four tasks were reduced to $n+1$ from $2n+1$. Prediction performance for the second set of neural network models were compared with the prediction performance of the first set of neural network models to

determine if there was improvement. The number of hidden units were reduced further for the third model if there was a possibility of improvement in the prediction performance.

2.7.4 Direct Connection Between Input and Output Units

Direct connections between input and output units can be made in a neural network, in addition to the connections between input and hidden units, and hidden and output units. If a linear transfer function is used in the output units, the final neural network model can represent the dependent variables as linear and non-linear combinations of the independent variables. This may improve prediction performance of the neural network.

BrainMaker, the neural network software used for this study does not directly allow direct connections between input and output units. A program in Visual Basic was developed and was linked with BrainMaker to be able make direct connections between input and output units. The code for the program developed and used is given in the Appendix.

Direct connections with input and output units were experimented for concrete pouring with the use of the program developed (Section 2.7.6). Because the prediction performance of the neural network model with direct connections was not better than the prediction performance of the neural network model without direct connections for concrete pouring, neural networks with direct connections were not experimented for the tasks other than concrete pouring.

2.7.5 Adding Noise to Training Data

Adding random noise to training data is another technique that may improve prediction performance of the neural network model. Siestma and Dow (1991) experimented adding random noise to the input data and, concluded for their data set that adding noise improved the ability of a network to recognize representatives of the classes which were not in the training set. But

Lawrence (1993) argued that adding noise to a very noisy data set such as financial forecasting data may worsen the prediction performance of the neural network.

Adding noise to the training data was experimented with concrete pouring data. The closeness of fit and prediction performance of the neural network model trained, without noise added data, was better than the closeness of fit and prediction performance of the neural network trained, with noise added data, for concrete pouring. Adding noise to data of other tasks were not experimented because several characteristics of these data sets including level of variation were similar to the characteristics of concrete pouring data.

2.7.6 Neural Network Models for Concrete Pouring

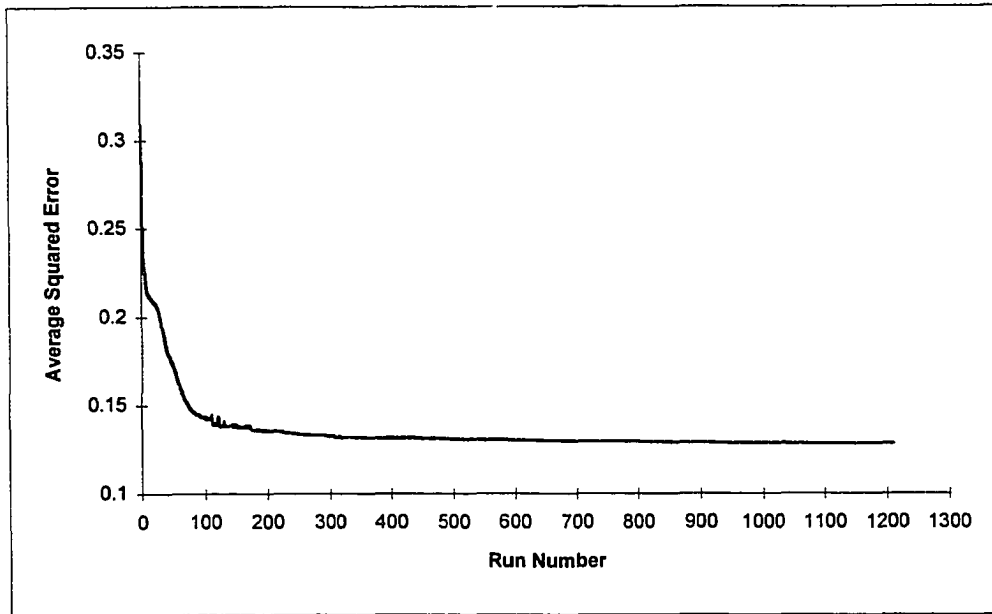
The neural network models for concrete pouring included six input variables that were included in the final regression model for the task. These variables were quantity (q), number of workers (n), temperature (t), and the variables associated with the job type; column and slab (cs), slab pump (u), and wall(w). The initial neural network model for concrete pouring; model NP-1, had 13 ($2n+1$) hidden units. The measures for comparison of closeness of fit and prediction performance were calculated as described in section 2.6 (Table 2.16). Next the number of hidden units were reduced to 7 ($n+1$) to obtain model NP-2. It was observed that the loss in closeness of fit was not very significant when the number of hidden units were reduced to 7; so the number of units were further reduced to 4 ($n/2+1$). The neural network model with 4 hidden units (NP-3) had a better prediction performance than the neural network with 7 hidden units. A random noise with gaussian distribution (mean zero, standard deviation 0.05) was added to the input data of the next neural network NP-4 during training. The average squared error of the model NP-4 did not converge as smoothly as the previous neural networks, because of the noise disruptions to the input data.

Table 2.16. Comparison of Neural Network Models for Concrete Pouring

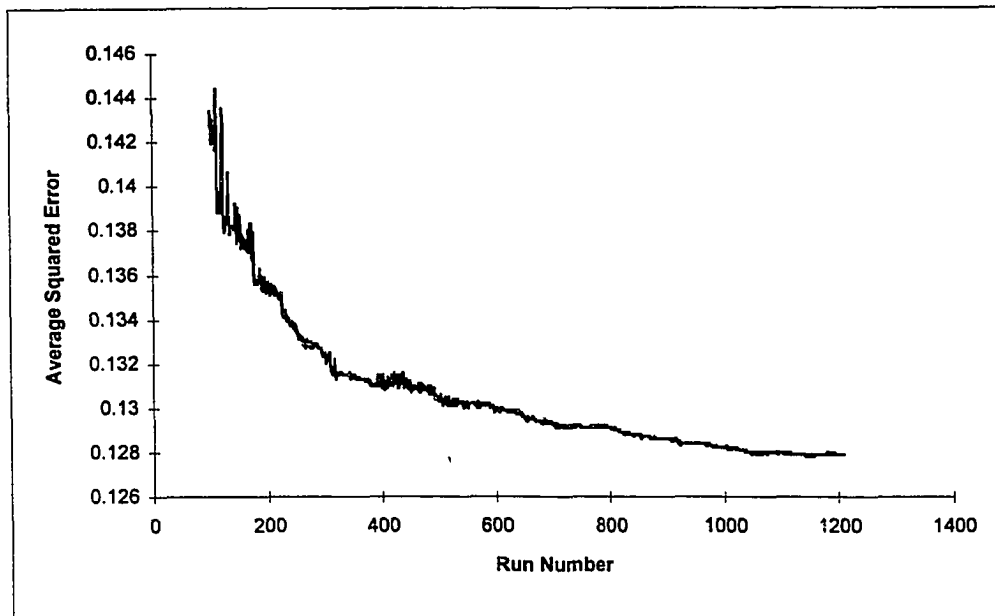
| Model | Closeness of Fit | | | Prediction Performance | | Hidden Units | Noise | Number of Runs |
|----------|------------------|------|------|------------------------|------|--------------|-------|----------------|
| | R ² | MSE | MAPE | MSE | MAPE | | | |
| NP-1 | .656 | 1.82 | 69.7 | 4.83 | 89.1 | 13 | none | 1291 |
| NP-2 | .591 | 2.11 | 72.2 | | | 7 | none | 1106 |
| NP-3(NP) | .587 | 2.12 | 70.5 | 3.67 | 76.8 | 4 | none | 1209 |
| NP-4 | .521 | 2.27 | 73.7 | 4.34 | 85.3 | 4 | 0.05 | 1000 |
| NP-5 | .506 | 2.55 | 61.2 | 4.54 | 75.6 | 4 | none | 100 |

Training of NP-4 was stopped at run number 1000, when a reasonable convergence was achieved. Prediction performance of NP-4 was worse than the prediction performance of NP-3; adding noise to input data did not improve generalization.

The final neural network for concrete pouring; NP-5, had direct connections between input and output units, in addition to the connections between input and hidden, and hidden and output units. The neural network had four hidden units like NP-3 and NP-4. Training for NP-5 was very slow because of the Visual Basic code added to BrainMaker to be able to make the direct connections. Executing BrainMaker commands from Visual Basic resulted in several Windows screen updates, which slowed the training. Due to slow training the number of runs were limited to 100. However 100 runs may be sufficient because most of the learning for models NP 1-4 (Figure 2.15a) took place in the first 100 runs. Direct connection between input and output units did not improve the prediction performance of model NP-3. Model NP-3 was selected as the final neural network model for concrete pouring.



(a)



(b)

Figure 2.15. Network Error of NP for (a) Runs 1-1209, (b) Runs 100-1209

2.7.2 Neural Network Models for Formwork, Concrete Finishing and Granular Fill

Two different neural network models were trained for each of the formwork, concrete finishing and granular fill tasks. The initial neural network models had $2n+1$ hidden units, and the second neural networks had $n+1$ hidden units. The input variables for the models were the same variables that were used in the final regression models of the tasks. The neural network for formwork with three hidden units (NF-2) had better prediction performance than the one with five hidden units (NF-1). NF-2, therefore was selected as the final neural network model for concrete pouring (Table 2.17). However, for concrete finishing, and granular fill tasks neural network models with $2n+1$ hidden units (NT-1, NG-1) had better prediction performance than the neural network models with $n+1$ hidden units. The neural network model with five hidden units for concrete finishing and, the neural network model with seven hidden units for granular fill were selected as the final neural network models.

Table 2.17. Neural Network Models for Formwork, Concrete Finishing, and Granular Fill

| Model | R ² | Closeness of Fit | | Prediction Performance | | Input Variables | Hidden Units | Number of Runs |
|----------|----------------|------------------|------|------------------------|-------|-----------------|--------------|----------------|
| | | MSE | MAPE | MSE | MAPE | | | |
| NF-1 | .464 | 16.51 | 62.2 | 25.37 | 76.8 | q,n | 5 | 1146 |
| NF-2(NF) | .428 | 17.63 | 65.6 | 24.84 | 71.6 | q,n | 3 | 1149 |
| NT-1(NT) | .545 | 5801 | 55.4 | 16796 | 102.1 | q,t | 5 | 1625 |
| NT-2 | .483 | 6614 | 59.4 | 21534 | 116.9 | q,t | 3 | 1538 |
| NG-1(NG) | .809 | 1674 | 33.9 | 3755 | 72.5 | q,n,p | 7 | 896 |
| NG-2 | .769 | 2027 | 42.9 | 3975 | 76.9 | q,n,p | 4 | 983 |

CHAPTER 3. RESULTS

3.1 Model Comparison

The final regression models and neural network models developed in Chapter-2 were compared with the average production rate, and factor models in the model comparison stage. The average production rate model is a practical approach to predict production rate which has been used by several contractors (Sonmez, 1992). The average production rate method is similar to the adjustments factors method described in Section 1.3, however, in average production rate method, only job type factor is used for adjustments. The average production rate model suggests use of average production rate values for different job types to predict the production rate. The average production rate model of concrete pouring (AP) that was used for comparison had the following form:

$$PR(P) = j_1c + j_2s + j_3w \quad (3.1)$$

where; $PR(P)$ was the predicted production rate for concrete pouring in cy/hours; $c=1$ if the job type was column, $c=0$ otherwise; $s=1$ if job type was slab, $s=0$ otherwise; $w=1$ if job type was wall, $w=0$ otherwise; j_1 was the average production rate for column job type; j_2 was the average production rate for slab job type; j_3 was the average production rate for wall job type.

The average production rate model for formwork (AF) was similar to model AP, and was in the following form:

$$PR(F) = j_1c + j_2b + j_3w \quad (3.2)$$

where; $PR(F)$ was the predicted production rate for formwork in sf/hours; $c=1$ if the job type was column, $c=0$ otherwise; $b=1$ if job type was grade beam, $b=0$ otherwise; $w=1$ if job type was wall, $w=0$ otherwise; j_1 was the average production rate for column job type; j_2 was the average production rate for grade beam job type; j_3 was the average production rate for wall job type.

Average production rate models for concrete finishing and granular fill included only one job types, because the data of the tasks were limited to one job type. The average production rate model for concrete finishing (AT) had the following form:

$$PR(T) = j \quad (3.3)$$

where; PR(T) was the predicted production rate for concrete finishing in sf/hours, and j was the average production rate for concrete finishing. The average production rate model for granular fill (AG) had the following form:

$$PR(G) = j \quad (3.4)$$

where; PR(G) was the predicted production rate for granular fill in sf/hours, and j was the average production rate for granular fill.

Factor models that were similar to the model 1.4 were the next set of models that were compared with the final neural network and regression models. Although model 1.4 was developed for masonry construction, it was suggested that same methodology could be used to model other labor intensive tasks (Sanders and Thomas, 1993). Data of temperature, humidity, number of workers, work (pour) method, and job type factors were available from the factors included in model 1.4. The factor model developed for concrete pouring (FP) was in the following form:

$$PR(P) = \alpha + \sum_{i=1}^{i=2} \beta_i j_i + \omega u + \sum_{i=1}^{i=8} \theta_i (th)_i + \lambda n \quad (3.5)$$

where; PR(P) was the predicted production rate for concrete pouring in cy/hours; α was the regression constant; $j_1=1$ if job type was slab, $j_1=0$ if job type was column or wall; $j_2=1$ if job type was wall, $j_2=0$ if job type was column or slab; $u=1$ if concrete pump was present, $u=0$ otherwise; $th_1=1$ if temperature(t) < 40 (F) and relative humidity (h) < 45 (%), $th_1=0$ otherwise; $th_2=1$ if $40 \leq t \leq 80$ and $h < 45$, $th_2=0$ otherwise; $th_3=1$ if $t > 80$ and $h < 45$, $th_3=0$ otherwise; $th_4=1$ if $t < 40$ and $45 \leq h \leq 80$; $th_4=0$ otherwise; $th_5=1$ if $t > 80$ and $45 \leq h \leq 80$, $th_5=0$ otherwise; $th_6=1$ if $t < 40$ and $h > 80$, $th_6=0$ otherwise; $th_7=1$ if $40 \leq t \leq 80$ and $h > 80$, $th_7=0$ otherwise; $th_8=1$ if $t > 80$ and $h > 80$,

$th_8=0$ otherwise (Sanders and Thomas, 1993); n is the number of workers; $\beta_i, \omega, \theta_i, \lambda$ were the regression coefficients.

The factor model for formwork (FF) was similar to model FP and was in the following the form:

$$PR(F) = \alpha + \sum_{i=1}^{i=2} \beta_i j_i + \sum_{i=1}^{i=8} \theta_i (th)_i + \lambda n \quad (3.6)$$

where: $PR(F)$ was the predicted production rate for formwork; $j_1=1$ if job type was grade beam, $j_1=0$ if job type was column or wall; $j_2=1$ if job type was wall, $j_2=0$ if job type was column or grade beam. rest of the terms were same as the equation 3.5. Factor models for concrete finishing (FT) and granular fill (FG) were similar to previous models and were in the following form:

$$PR(T) = \alpha + \sum_{i=1}^{i=8} \theta_i (th)_i + \lambda n \quad (3.7)$$

$$PR(G) = \alpha + \sum_{i=1}^{i=8} \theta_i (th)_i + \lambda n \quad (3.8)$$

where $PR(T)$ was the predicted production rate for concrete finishing in $sf/hours$; $PR(G)$ was the predicted production rate for granular fill in $sf/hours$.

Data of tasks did not include all of the nine temperature humidity levels defined. The terms associated with the temperature and humidity levels which were not included in the data were dropped from the factor models.

The closeness of fit and prediction performance values for the average production rate and factor models were calculated by the procedure described in Section 2.6. The final regression and neural network models developed for concrete pouring, formwork, and granular fill were superior to the factor, and average production rate models developed for the tasks, in terms of closeness of fit and prediction performance (Table 3.1). The regression model (RP) and neural network model (NP) for

Table 3.1. Comparison of the Models

| Task | Model | R ² | Closeness of Fit | | Prediction Performance | |
|--------------------|-------|----------------|------------------|-------|------------------------|-------|
| | | | MSE | MAPE | MSE | MAPE |
| Concrete Pouring | AP | .175 | 4.26 | 100.4 | 5.52 | 105.5 |
| | FP | .281 | 3.71 | 88.2 | 5.53 | 110.6 |
| | RP* | .569 | 2.22 | 61.3 | 3.96 | 71.7 |
| | NP | .587 | 2.12 | 70.5 | 3.67 | 76.8 |
| Formwork | AF | .080 | 28.34 | 91.6 | 71.65 | 110.6 |
| | FF | .117 | 27.20 | 89.7 | 76.85 | 150.3 |
| | RF* | .427 | 17.64 | 61.8 | 21.04 | 68.1 |
| | NF | .428 | 17.63 | 65.6 | 24.84 | 71.6 |
| Concrete Finishing | AT | .000 | 12758 | 102.0 | 16820 | 118.3 |
| | FT | .320 | 8675 | 80.1 | 17827 | 126.3 |
| | RT | .301 | 8920 | 69.5 | 18831 | 107.6 |
| | NT* | .545 | 5801 | 55.4 | 16796 | 102.1 |
| Granular Fill | AG | .000 | 8735 | 125.7 | 13917 | 180.6 |
| | FG | .016 | 8593 | 128.2 | 20500 | 219.6 |
| | RG* | .783 | 1894 | 33.6 | 2881 | 49.6 |
| | NG | .769 | 2027 | 42.9 | 3975 | 76.9 |

* was selected as the final model.

concrete pouring were very close in terms of closeness of fit and prediction performance, but regression model was selected as the final model because it was more parsimonious. Regression model for formwork (RF) had a slightly better prediction performance than the neural network model (NF). Concrete finishing was the only task among the four tasks in which the average production rate and factor models were close to the regression and neural network models in terms of closeness of fit, and prediction performance. However neural network model (NT) which had the best closeness of fit, and prediction performance among the four models was selected as the final model for concrete finishing. The regression model for granular fill had the best closeness of fit and

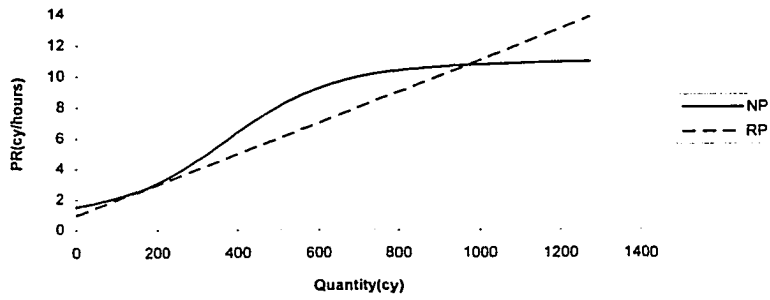
prediction performance among the four models, and was selected as the final model. Average production rate and factor models for granular fill had a very poor closeness of fit, and prediction performance compared to the regression and neural network models.

3.2 Sensitivity Analysis of the Models

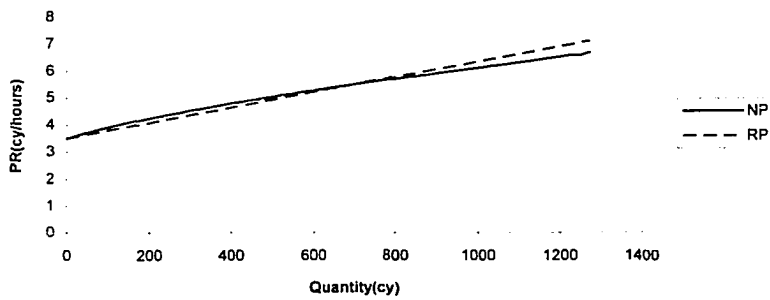
At the final stage of model comparison a sensitivity analysis was performed, to compare how the effect of a factor was quantified by the final regression and neural network models. To perform sensitivity analysis first the minimum, maximum and the average values of the factors for each of the tasks were determined. The values of each factor were then varied one at a time, while holding the values of the other factors at their mean value. The values were varied between minimum and the maximum values of the factor, and generally 25 values were used. For each value production rates were calculated by the regression and neural network models. The plots of the sensitivity analysis are given in Figures 3.1 to 3.10.

The sensitivity analysis of the factors for concrete pouring were performed for three different job types that were included in the final regression (RP), and neural network models (NP). Sensitivity analysis of the neural network model indicated that the rate of increase in the production rate of concrete pouring due to increase in quantity, with concrete pump, was not same as the rate of increase without concrete pump (Figures 3.1a-c). This relation was not included in the regression model RG-3, which was being considered for concrete pouring. An interaction term $q*u$ was added to the model RP-3 to obtain model RP. The interaction term improved model RP-3 in terms of closeness of fit and prediction performance (Table 2.9).

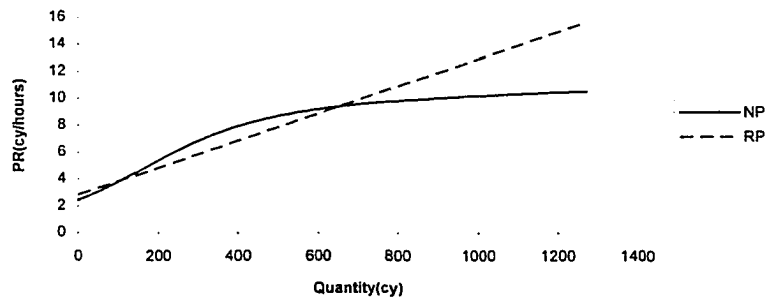
Increase in quantity also resulted in an increase in production rate for formwork, concrete finishing and granular fill tasks. However the rates of increase for different tasks were not same. The rate of increase for concrete finishing was less than the rate of increase for the other tasks.



(a)

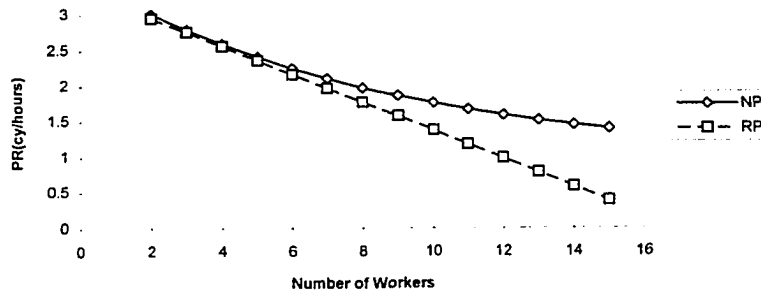


(b)

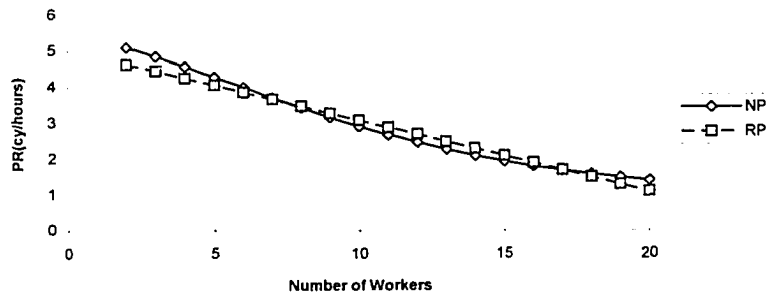


(c)

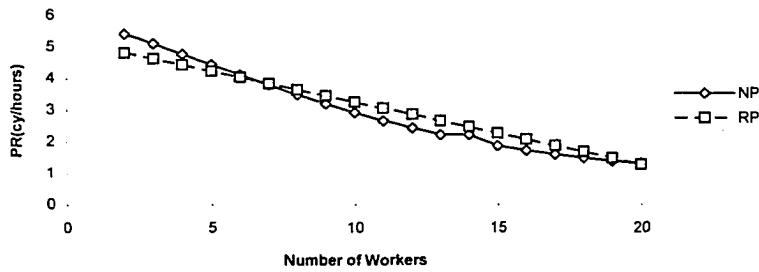
Figure 3.1. Sensitivity Analysis of NP and RP for Quantity (a) Column and Slab, (b) Slab Pump, (c) Wall



(a)

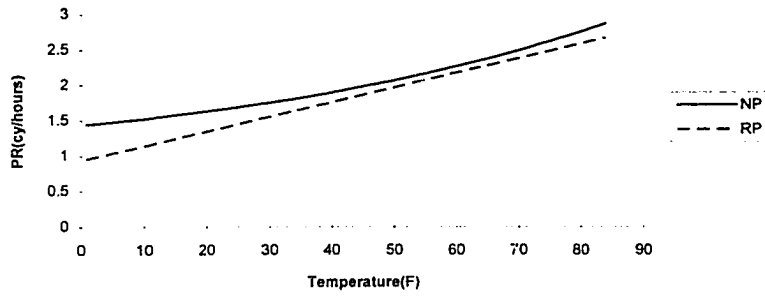


(b)

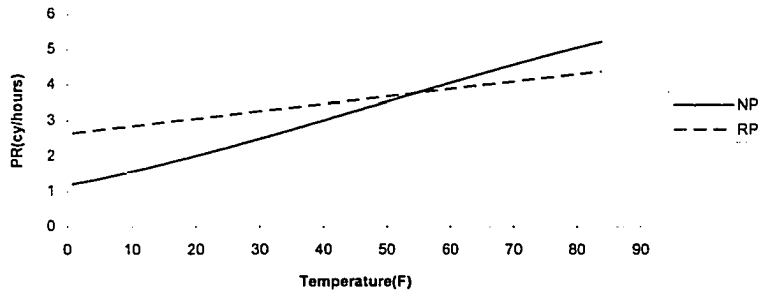


(c)

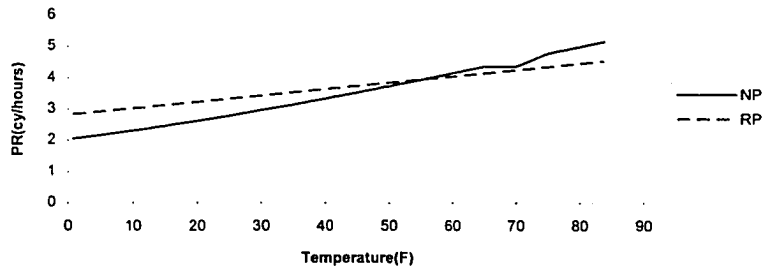
Figure 3.2. Sensitivity Analysis of NP and RP for # Workers (a) Column and Slab, (b) Slab Pump, (c) Wall



(a)



(b)



(c)

Figure 3.3. Sensitivity Analysis of NP and RP for Temperature (a) Column and Slab (b) Slab Pump, (c) Wall

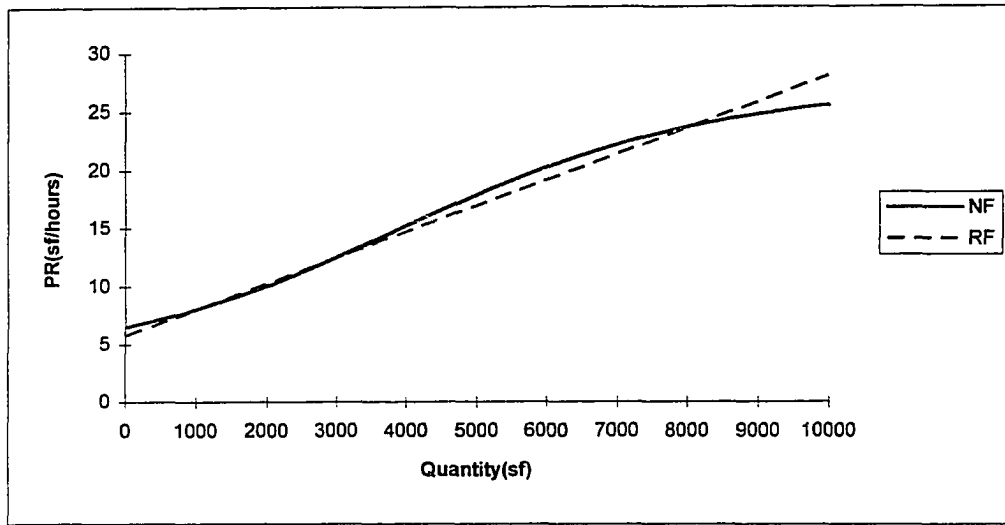


Figure 3.4 Sensitivity Analysis of NF and RF for Quantity

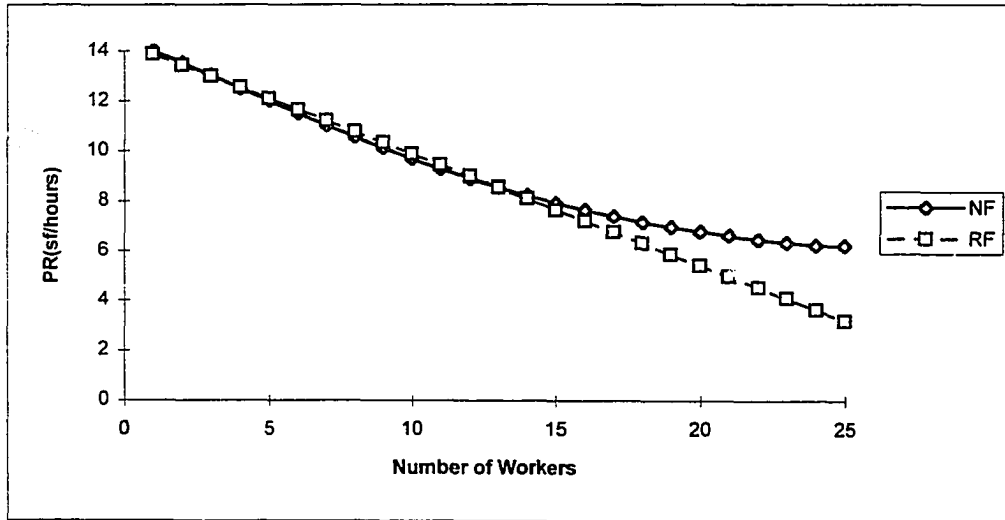


Figure 3.5 Sensitivity Analysis of NF and RF for Number of Workers

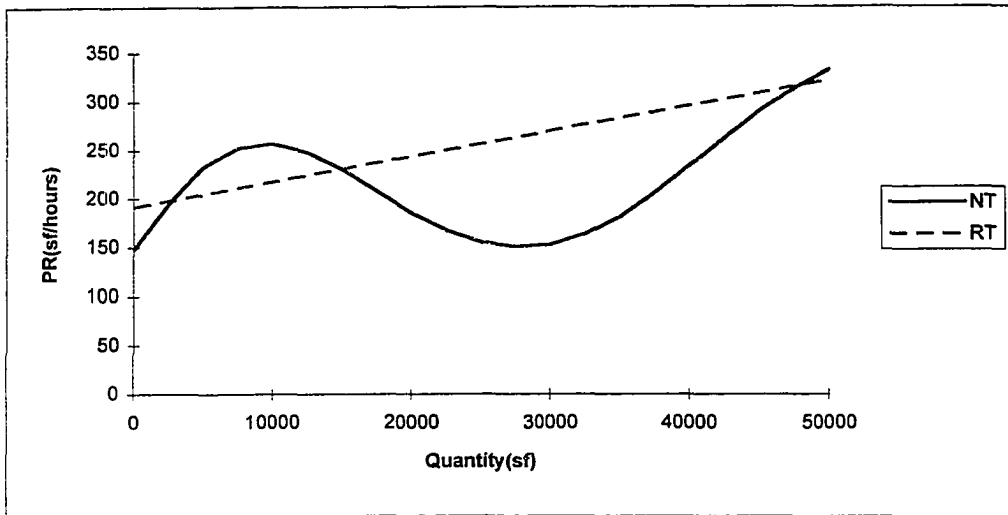


Figure 3.6 Sensitivity Analysis of NT and RT for Quantity

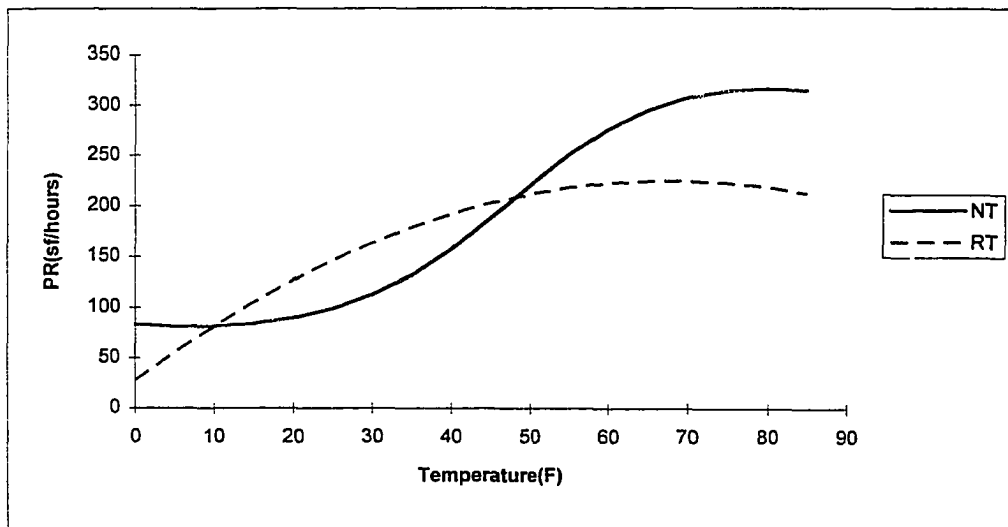


Figure 3.7 Sensitivity Analysis of NT and RT for Temperature

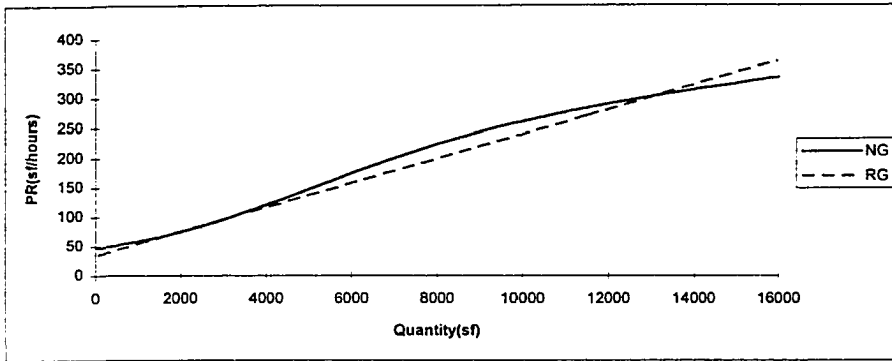


Figure 3.8 Sensitivity Analysis of NG and RG for Quantity

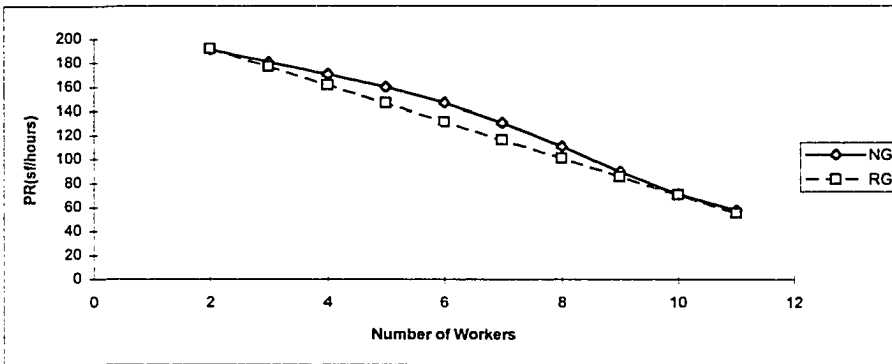


Figure 3.9 Sensitivity Analysis of NG and RG for Number of Workers

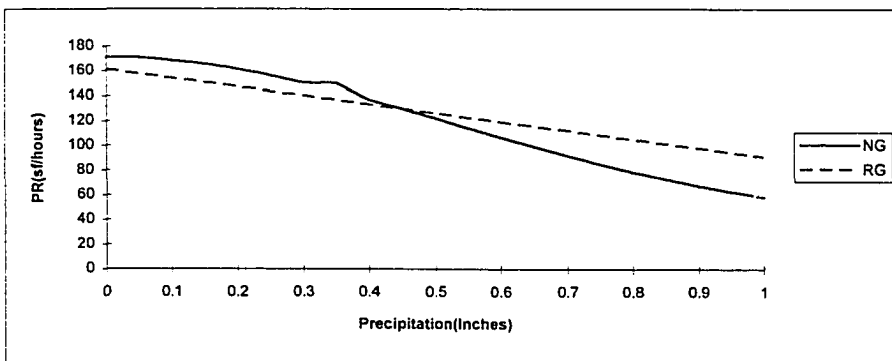


Figure 3.10 Sensitivity Analysis of NG and RG for Precipitation

The increase in production rate due to increase in the level quantity could be related to work complexity and repetition. Jobs with bulk quantities (higher quantity levels) are generally less complex than the jobs with fewer quantities, which may result in a higher production rate. Jobs with fewer quantities may also require a higher equipment, and material preparation time per unit quantity, than the jobs with bulk quantities. The crew may as well, become more familiar with the task as the amount of quantities completed for the task increases.

Number of workers was the next factor that sensitivity analysis was performed. Production rate decreased as the number of workers increased for concrete pouring, formwork, and granular fill tasks. However the rates of decrease in production rate due to increase in the number of workers were not same for the tasks. The decrease in production rate due to increase in number of workers could be related to turnover rate, and overcrowding. The description of the number of workers factor was given in Section 2.2. A large value of crew size will mean a large value for the number of workers, but opposite of the relation is not always true. A large value of number of workers may also indicate a high turnover rate.

Crew size was included in the factor model 1.4 which was developed by Sanders and Thomas. The model also suggested that productivity decreased as crew size increased. However, the rate of decrease in productivity due to number of workers that was suggested by the regression and neural network models was higher than the rate of decrease suggested by the factor model 1.4 due to crew size. This may be because the turnover rate, which is also believed to be a negative productivity factor was also reflected in the number of workers factor.

Temperature was another factor that was included in the regression, and neural network models of concrete pouring, and concrete finishing. The sensitivity analysis for temperature indicated that productivity improved as temperature increased. But there were limited data points for the regions where temperature was less than 20 °F, or more than 80 °F. The relation between temperature and productivity suggested by the final regression, and neural network models generally agreed with the

relation suggested by the previous weather models given in Section 1.5. But the suggested decrease in productivity at the high temperature levels was not observed because there were very few data points at the high temperature levels (80-110 °F). The rate of increase in productivity due to temperature for concrete finishing was higher than the rate of increase for concrete pouring. The rate of increase for concrete finishing was also higher than the rate suggested by the NECA, Grimm-Wagner, and Koehn-Brown models, but less than the rate suggested by the Thomas-Yiakouimis model.

Humidity, another weather related factor was not identified as a factor significantly influencing productivity of the four tasks studied. The precipitation factor, however, improved the model for the granular fill. The regression and neural network models of granular fill suggested that production rate decreased as the amount of precipitation increased.

CHAPTER 4. CONCLUSIONS

Research conducted to date on construction labor productivity usually has focused on the effect of a single factor while neglecting the effects of the other factors. Factor model was the only model that focused on quantification of the effect of multiple factors. It was suggested that the factor modeling methodology which was developed for masonry construction could also be used for other tasks. However the factor modeling methodology had several limitations.

The modeling methodology defined in this study for productivity modeling of labor intensive construction tasks, is an improvement over the factor modeling methodology in the way it addresses the following three issues. The first issue is that the methodology includes a stage that factor modeling methodology lacks, in which the factors influencing productivity are identified. The effects of factors on productivity for different tasks may not be same. It is important to include only the factors that contribute to the model of a task to achieve good prediction performance. It is also important not to be limited to the factors included in the factor model. There may be several factors that are not included in the factor model, but may have potential to improve the productivity model for a task. All of the factors that data are available should be considered for the data identification stage. The second issue of improvement is the methodology presented suggests use of parsimonious models which is not considered in the factor modeling methodology. Factor modeling methodology suggests use of eight binary terms for temperature and humidity, instead of few continuous variables. The principle of parsimony is important because, in practice parsimonious models generally produce better forecasts. Third issue of improvement is that the modeling methodology suggested includes a procedure to compare prediction performance. In factor modeling methodology only closeness of fit is used to compare different models. However a good closeness of fit does not always result in a

good prediction performance. Use of closeness of fit as the only performance measure may lead to models that fit data better, than the models that forecast better.

The modeling methodology suggested, factor modeling methodology and the average production rate method were used to model four different tasks. The models developed by the modeling procedure suggested in this study was better than the models developed by the average production rate method, and the factor modeling methodology, in terms of closeness of fit and prediction performance. The prediction performances of the models developed for concrete pouring, formwork, and granular fill, by the methodology presented in this study were significantly better than the prediction performances of the models developed by the factor modeling methodology. This significant difference verifies the importance of the three issues discussed for modeling.

This study included data, and models of multiple factors for multiple tasks. Factor model was developed for masonry construction only. The regression and neural network models developed in this study suggested that the effects of factors on productivity and, the rate of effects may not be same for different tasks. Therefore, this study suggests that use of one common productivity model for different tasks is not sufficient. Productivity models of the different tasks should be studied individually, although there may be similarities between the effects of the some factors on different labor intensive tasks.

Use of neural networks for construction labor productivity was also explored in this study. Neural networks with their mapping capabilities helped the overall modeling process. Neural networks have shown potential to identify the effects of the factors, especially when interactions and non linear relations were present.

The final models given could be improved by including other factors, for which data were not available in this study. The modeling methodology described could also be used for labor intensive tasks, other than concrete pouring, formwork, concrete finishing and granular fill. Project related factors, such as project type, design features, project team could also be studied by the described

methodology. However, this study requires data from several different types of projects. A more extended study may also include international factors. An international construction productivity study may help to establish international labor productivity norms.

Productivity models which are capable of explaining variations due to several factors will improve accuracy of labor cost estimates, and activity duration forecasts. Better understanding of construction productivity will also lead to more realistic expectations and better planning decisions.

REFERENCES

- Blank, T. B., and Brown, S.T. (1993). "Nonlinear multivariate mapping of chemical data using feed- forward neural networks." *Analytical Chemistry*, 65(21), 3081-3089.
- Borcherding, J.D., and Alarcon, L.F. (1991). "Quantitative effects on construction productivity." *The Construction Lawyer*, 11(1), 1, 36-48.
- BrainMaker Professional User's Guide and Reference Manual* (1993). California Scientific Software, Nevada City, Calif.
- Caudill, M. (1991). "Neural network training tips and techniques." *AI Expert*, 56-61, Jan.
- Caudill, M. (1992). "The view from now." *AI Expert*, 24-31, Jun.
- Chao, L., and Skibniewski, M.J. (1994). "Estimating construction productivity: Neural-network-based approach." *J. Computing in Civ. Engrg.*, ASCE, 8(2), 234-251.
- Cheng, B., and Titterington, D.M. (1994). "Neural networks: A review from a statistical perspective." *Statistical Sciences*, 4(1), 2-54.
- Christian, J., Hachey, D. (1995). "Effects of delay times on production rates in construction." *J. Constr. Engrg. and Mgmt.*, ASCE 121(1), 20-26.
- Churchland, P.S., and Sejnowski T.J. (1992). *The Computational Brain*. The MIT Press, Cambridge, Mass.
- Clapp, M.A. (1966). "The effect of adverse weather conditions on productivity on five building sites." *Construction Series, No. 21*, Building Research Establishment, Watford, England.
- Collins, J.M., and Clark, M.R. (1993). "An application of the theory of neural computation to the prediction of workspace behavior: An illustration and assessment of network analysis." *Personnel Psychology*, 46, 503-524
- Dalliva, L. (1954). *Estimating Production and Construction Costs*. Dalliva Co., Houston, Tex.
- Draper, N.R., and Smith, H. (1981). *Applied Regression Analysis*. John Wiley & Sons, New York, N.Y.
- Dutta, S., and Shenakar, S. (1988). "Bond rating: A non-conservative application of neural networks." *Proc., IEEE Int. Conf. on Neural Networks*, Vol 2, 445-450.
- "Effect of repetition on building operations and processes on site." (1965). *United Nations Pub. No. ST/ECE/HOU/14*, United Nations Committee on Housing, Building, and Planning, New York, N.Y.

- "Effect of scheduled overtime on construction projects." (1974). *A Construction Industry Cost Effectiveness Task Force Rep.*, Business Roundtable, New York, N.Y.
- "The Effects of Scheduled Overtime and Shift Schedule on Construction Craft Productivity." (1988). *Source Document 43*, Construction Industry Institute, Austin, Tex.
- "The effects of temperature on productivity." (1974). *Rep No. 5072*, National Electrical Contractors Association (NECA), Washington, D.C.
- Elman J. L., and Zipser D. (1988). "Learning the hidden structure of speech." *J. The Acoustical Soc. of America*, 83(4), 1615-1626.
- Fletcher D., and Goss E. (1993). "Forecasting with neural networks." *Information and Management*, 24, 159-167.
- Freeman J.A. (1994). *Simulating Neural Networks with Mathematica*. Addison-Wesley Publishing Company.
- Goh, A.T.C. (1995). "Modeling soil correlations using neural networks." *J. Computing in Civ. Engrg.*, ASCE, 9(4)
- Grimm, C.T., and Wagner, N.K. (1974). "Weather effects on mason productivity." *J. Constr. Engrg. and Mgmt.*, ASCE 100(3), 319-335.
- Halligan, D.W., Demsetz, L.A., Brown, J.D., Pace, C.B. (1994). "Action-response models and loss of productivity in construction." *J. Constr. Engrg. and Mgmt.*, ASCE 120(1), 47-64.
- Handa, V.K., and Abdalla, O. (1989). "Forecasting productivity by work sampling." *Constr. Mgmt. Econ.*, 7, 19-28.
- Hegazy T., Fazio P., and Moselhi O. (1994). "Developing practical neural network applications using back propagation." *Microcomputers in Civ. Engrg.*, 9(2), 145-159.
- Hendrickson, C., Martinelli, D., and Rehak, D. (1987). "Hierarchical rule-based activity duration estimation." *J. Constr. Engrg. and Mgmt.*, ASCE 113(2), 288-301.
- Hinton, G.E. (1992). "How neural networks learn from experience." *Scientific American*, 145-151, Sep.
- Hornik, K., Stinchcombe, M., and White, H. (1989). "Multilayer feedforward networks are universal approximators." *Neural Networks*, 2(5), 359-366.
- Hwang, J.N., Li H., Maechler, M., Martin, D., and Schimert, J. (1992). "Projection pursuit learning networks for regression." *Engineering Applications of Artificial Intelligence*, 5(3), 193-204.

- Jonson, R.A., and Dean, W.W. (1992). *Applied Multivariate Statistical Analysis*. Prentice Hall, New Jersey.
- Karunanithi, N., Grenney, W.J., Whitley, D., and Bovee, K. (1994). "Neural Networks for River Flow Prediction." *J. Computing in Civ. Engrg.*, ASCE, 8(2), 201-220.
- Kimoto T., Asakawa K., Yoda, M., and Takeoka, M. (1990). "Stock market prediction system with modular neural networks." *IJCNN Int. Joint Conf. on Neural Networks*, Vol. I, San Diego, Calif, 1-7.
- Koehn, E., and Brown, G. (1985). "Climatic effects on construction." *J. Constr. Engrg. and Mgmt.*, ASCE 111(2), 129-137.
- Kolmogorov, A. (1957). "On the presentation of continuous functions of many variables by superposition of continuous functions of one variable and addition." *Doki Akad Nauk*, 144, 953-956.
- Lawrence, J. (1993). *Introduction to Neural Networks*. California Scientific Software Press, Nevada City, Calif.
- Local Climatological Data: Monthly Summary*. (1992-1994). National Climatic Data Center, Asheville, N.C.
- Le Chun, Y. (1989). "Generalization and network design strategies." *Connectionism in Perspective*. Elsevier Science Publishers, New York, N.Y., 143-55.
- Lehr, M., and Widrow, B. (1993). "Commercial and industrial applications of neural networks." *Proc., 5th Workshop on Neural Networks*, San Francisco, Calif., 165-170.
- Leshno, M., Lin, V.Y., Pinkus, A., and Schocken, S. (1993). "Multilayer feedforward networks with a nonpolynomial activation function can approximate any function." *Neural Networks*, Vol. 6, 861-867.
- Liou, F.W., and Borcharding, J.D. (1986). "Work sampling can predict unit rate productivity." *J. Constr. Engrg. and Mgmt.*, ASCE 112(1), 90-103.
- Maloney, W.F. (1981). "Motivation in construction: A review." *J. Constr. Div.*, ASCE, 107(4), 641-647.
- Maloney, W.F. and McFillen, J. (1985). "Valence of and satisfaction with job outcomes." *J. Constr. Engrg. and Mgmt.*, ASCE 111(1), 53-73.
- Maloney, W.F. and McFillen, J. (1986). "Motivational implications of construction work." *J. Constr. Engrg. and Mgmt.*, ASCE 112(1), 137-151.
- Maloney, W.F. (1986). "Understanding motivation." *J. Mgmt. in Engrg.*, ASCE 2(4), 231-245.

- Maloney, W.F., and McFillen, J. (1987). "Motivational impact of work crews." *J. Constr. Engrg. and Mgmt.*, ASCE 113(2), 208-221
- Martin, C. J. (1991). *Labor Productivity Control*. Praeger Publishers, New York, N.Y.
- Masters, T. (1993). *Practical neural network recipes in C++*. Academic Press, San Diego, Calif.
- McClure, R.M., Willenbrock, J.H., and Henderson, J.D. (1980a). "Fabrication of segments for an experimental segmental bridge.", *Rep. to Pennsylvania Department of Transportation*. 134.
- McClure, R.M., Thomas, H.R., and Henderson, J.E. (1980b). "Construction of an experimental segmental bridge." *Rep. to Pennsylvania Department of Transportation*. 317.
- McKim, R.A. (1993), "Neural network applications to cost engineering." *Cost Engineering*, AACE., 35(7), 31-35.
- "Measuring productivity in construction." (1982). A-1 Report of the Construction Industry Cost Effectiveness Project. *The Business Roundtable*, New York, N.Y.
- Moselhi O., Hegazy, T., and Fazio, P. (1991). "Neural networks as tools in construction." *J. Constr. Engrg. and Mgmt.*, ASCE 117(4), 606-625.
- Muller, B., and Reinhardt, J. (1990). *Physics of Neural Networks: An Introduction*. Springer-Verlag, Berlin, Germany.
- Neil, J.M. (1982). *Construction Cost Estimating for Project Control*. Prentice-Hall Inc., Englewood Cliffs, N.J.
- Neil, J.M., and Knack, L.E. (1984). "Predicting Productivity." *AACE Trans., Paper H-3*, American Association of Cost Engineers (AACE), Morgantown, W. Va.
- Nielsen, R.H. (1989). *Neurocomputing*. Addison-Wesley Co., New York, N.Y.
- Niu, A., Wan, D.Li, L., Wee, W.G., and Carrier, C.W. (1991). "A comparative study of model fitting by using neural network and regression." *Applications of Artificial Neural Networks(II)* SPIE Vol. 1469, 495-505.
- O'Connor, L.V. (1969). "Overcoming the problems of construction scheduling on large central station boilers." *Foster Wheeler Corporation, American Power Conf.* Chicago, Ill.
- Oglesby, C., Parker, W.H., and Howell, A.G. (1989). *Productivity Improvement in Construction*. McGraw Hill, New York, N.Y.
- "Overtime and Productivity in Electrical Construction." (1969). *Rep. No. 5050*, 2nd Ed., National Electrical Contractors Association (NECA), Washington, D.C.

- Ott, L. (1988). *An Introduction to Statistical Methods and Data Analysis*. 3rd Ed., PWS-KENT Publishing Company, Boston, Mass.
- Page, G.F., Gomm, J.B., and Williams D. (1993). *Application of Neural Networks to Modelling and Control*. Chapman & Hall, London, England.
- Pankratz, A. (1983). *Forecasting With Univariate Box-Jenkins Models*. John Willey & Sons, New York, N.Y.
- Pao, Y. H. (1989). *Adaptive Pattern Recognition and Neural Networks*. Addison-Wesley, Reading, Mass.
- Rakes T. (1991) "Research in the neural network/expert system area." *Decision Line*, 22, 11-14.
- Refenes, A.P., Zapranis, A., and Francis, G. (1994). "Stock performance modeling using neural networks: A comparative study with regression models." *Neural Networks*, 7(2), 375-388.
- Ripley, B.D. (1994). "Neural networks and related methods for classification." *Journal of the Royal Statistical Society (B)*, 56(3), 409-456.
- Rochester, J.D. (1990). "New business uses for computing." *I/S Analyzer*, 28, 1-16.
- Rumelhart, D.E., Hinton, G.E., and Williams, R.J. (1986). "Learning internal representations by error propagation." *Parallel Distributed Processing*, Vol. 1, MIT Press, Cambridge, Mass.
- Sanders, S.R., and Thomas, R.H. (1991). "Factors affecting masonry-labor productivity." *J. Constr. Engrg. and Mgmt.*, ASCE 117(4), 626-644.
- Sanders, S.R., and Thomas, R.H. (1993). "Masonry productivity forecasting model." *J. Constr. Engrg. and Mgmt.*, ASCE 119(1), 163-179.
- "Scheduled Overtime Effect on Construction Projects." *Rep. No. C-2*, Business Roundtable, New York, N.Y.
- Sharda, R. (1994). "Neural networks for the MS/OR analyst: An application bibliography." *Interfaces*, 24(2), 116-130.
- Siestma, J. (1990). "The effect of pruning a back propagation network." *First Australian Conference on Neural Networks (ACNN'90)*, Sydney, Australia.
- Sietsma, J., and Dow, R.J. (1991). "Creating artificial neural networks that generalize." *Neural Networks*. Vol. 4, 67-79.
- Skills and Knowledge of Cost Engineering*. (1992). 3rd ed., AACE International, Morgantown, W.Va.

- Sonmez, R. (1992). "A data-based activity duration forecast system for construction projects." MS. thesis, Iowa State Univ., Ames, Ia.
- Srinivasan, D., Liew, A.C., and Chen, John S.P. "A novel approach to electrical load forecasting based on a neural network." *IEEE International Conference on Neural Networks*. Vol. 2, Singapore, 1172-1177
- Swank, J. (1989). "An exploratory study of the effects of crew size on productivity in masonry construction," MS. thesis, The Pennsylvania State Univ., Univ. Park, Pa.
- Thomas, H.R. (1991). "Labor productivity and work sampling: The bottom line." *J. Constr. Engrg. and Mgmt.*, ASCE 117(3), 423-444.
- Thomas, H.R. (1992). "Effects of scheduled overtime on labor productivity." *J. Constr. Engrg. and Mgmt.*, ASCE 118(1), 60-76.
- Thomas, H.R., Guevara, J.M., and Gustenhoven, C.T. (1984). "Improving productivity estimates by work sampling." *J. Constr. Engrg. and Mgmt.*, ASCE 110(2), 178-188.
- Thomas, H.R., Mathews C.T. (1985). *An Analysis of the Methods of Measuring Construction Productivity*. The Constr. Industry Inst., Austin, Tex.
- Thomas, H.R., Mathews, C.T., and Ward, J.G. (1986). "Learning curve models of construction productivity." *J. Constr. Engrg. and Mgmt.*, ASCE 112 (2), 245-258.
- Thomas, H.R., and Yiakouimis, I. (1987). "Factor model of construction productivity." *J. Constr. Engrg. and Mgmt.*, ASCE 113(4), 623- 639.
- Thomas, H.R., Maloney, W.F., Horner, M.W., Smith, G.R., Handa, V.K., and Sanders, S.R. (1990). "Modeling construction labor productivity." *J. Constr. Engrg. and Mgmt.*, ASCE 116 (4), 705-726.
- Thomas, H.R., Sanders, S.R., and Bilal, S. (1992). "Comparison of labor productivity." *J. Constr. Engrg. and Mgmt.*, ASCE 118(4), 635-650.
- Thomas, H.R., Raynar, K.A., (1994). "Effects of overtime on labor productivity: A quantitative analysis." *Source Document 98*, The Pennsylvania State Univ., Univ. Park, Pa.
- Thomas, H.R., and Sakarcan, A. (1994). "Forecasting labor productivity using factor model." *J. Constr. Engrg. and Mgmt.*, ASCE 120(1), 228-239.
- Thomas, H.R., and Napolitan, C.L. (1995). "Quantitative effects of construction changes on labor productivity." *J. Constr. Engrg. and Mgmt.*, ASCE 121(3), 290-298.

- Ward, J.G., Thomas, H.R. (1984). "A validation of learning curve models available to the construction industry." *Construction Management Research Series Rep. No. 20*, The Pennsylvania State Univ., Department of Civ. Engrg., Univ. Park, Pa.
- Wassermann, P.D. (1989). *Neural Computing: Theory and Practice*. Van Nostrand Reinhold, New York, N.Y.
- White, H. (1990) "Connectionist nonparametric regression: Multilayer feedforward neural networks can learn arbitrary mappings." *Neural Networks*, Vol. 3, 535-549.
- Williams, T.P, Khajuria, A., Balaguru, P. (1992). "Neural network for predicting concrete strength." *Proc., 8th Conf., Computing in Civ. Engrg.*, ASCE, Dallas, Tex.
- Williams, T.P. (1994). "Predicting changes in construction cost indexes using neural networks." *J. Constr. Engrg. and Mgmt.*, ASCE 120(2), 306-320.
- Zahedi, F. (1991). "An introduction to neural networks and a comparison with expert systems." *Interfaces*, 21(2), 25-38.

APPENDIX
CODE LISTING

*** This program makes direct connection between input and output units, in addition to the connections between input and hidden, and hidden and output units, for BrainMaker Professional.

*** Define Variables

```
Dim N, NV, C, FC, V, LNNT, HU, WL, WTS, FACTN, FFACTN, X, LF, LCWT, WT, NRUN, CO,
FF, TF, CT
NV = 6
NRUN = 100
ReDim FCT(1000, NV) As Single
ReDim OUT(1000, NV) As Single
ReDim NET(1000) As String
ReDim NWTS(NV)
LF = Chr(13)
```

*** Read Fact File

```
Open "c:\vb\cf.fct" For Input As #1 ' Open file
Line Input #1, N
Line Input #1, N
C = 1
Do While Not (EOF(1))
    If EOF(1) GoTo 10
        Input #1, FCT(C, 1), FCT(C, 2), FCT(C, 3), FCT(C, 4), FCT(C, 5), FCT(C, 6)
    If EOF(1) GoTo 10
        Input #1, OUT(C, 1)
    C = C + 1
    If EOF(1) GoTo 10
        Input #1, N
Loop
10 Close #1
FC = C - 1
C = 1
```

*** Create Necessary Fact Files

```
Do While C <= FC
    FF = CStr(C) & ".fct"
    TF = "C:\VB\C.FCT"
    FileCopy TF, FF
    Open FF For Append As #5
    Seek #5, 20
    Print #5, LF
    Print #5, FCT(C, 1), FCT(C, 2), FCT(C, 3), FCT(C, 4), FCT(C, 5), FCT(C, 6)
    Print #5, OUT(C, 1)
    Close #5
    C = C + 1
Loop
```

*** Make Direct Connections Between Input and Output Units

CO = 1

Do While CO <= NRUN

 C = 1

 Do While C <= FC

 V = 1

 Open "c:\vb\c.net" For Input As #3

 Do While Not (EOF(3))

 Input #3, NET(V)

 WTS = Left(NET(V), 7)

 V = V + 1

 LNNT = V - 1

 Loop

 Close #3

 V = 1

 Do While V <= LNNT

 WTS = Left(NET(V), 7)

 If WTS = "weights" Then

 WL = V

 V = LNNT + 1

 End If

 V = V + 1

 Loop

 V = 1

 Do While V <= NV

 FACTN = FCT(C, V)

 If FACTN = 0 Then FACTN = .0001

 If FACTN = 1 Then FACTN = .9999

 FFACTN = Log(1 / FACTN - 1)

 If FFACTN = 0 Then FFACTN = .0001

 NWTS(V) = 1 / FFACTN

 TFCT = 1 / (1 + Exp(1 / NWTS(V)))

 NET(WL + 1) = "#.##### #.##### #.##### #.##### #.##### #.##### " &
 Left(CStr(NWTS(V)),6)

 V = V + 1

 WL = WL + 1

 Loop

NET(5) = "filename trainfacts C:\VB\" & CStr(C) & ".FCT"

Open "c:\vb\c.net" For Output As #4

V = 1

```

Do While V <= LNNT
  Print #4, NET(V)
  V = V + 1
Loop

Close
AppActive = "Brianmaker Professional"
SendKeys "%F", True
SendKeys "R", True
SendKeys "c.net", True
SendKeys "~", True
SendKeys "%O", True
SendKeys "T", True
SendKeys "%F", True
SendKeys "S", True
SendKeys "~", True
SendKeys "Y", True
If CT = 5 Then
  SendKeys "%F", True
  SendKeys "X", True
End If
V = 1
Do While V <= LNNT
  NET(V) = ""
  V = V + 1
Loop
C = C + 1
Loop
CO = CO + 1
Loop

End

```