

A Study on Improving Forecasting Accuracy for Expenditures of Residential Building Projects through Selecting Similar Cases

이 준 성*
Yi, June-Seong

Abstract

Dynamic and fragmented characteristics are two of the most significant factors that distinguish the construction industry from other industries. Previous forecasting techniques have failed to solve the problems derived from the above characteristics, and do not provide considerable support. This paper deals with providing a more precise forecasting by applying Case-based Reasoning (CBR). The newly developed model in this study enables project managers to forecast monthly expenditures with less time and effort by retrieving and referring only projects of a similar nature, while filtering out irrelevant cases included in database. For the purpose of accurate forecasting, the choice of the numbers of referring projects was investigated. It is concluded that selecting similar projects at 5~6 % out of the whole database will produce a more precise forecasting.

The new forecasting model, which suggests the predicted values based on previous projects, is more than just a forecasting methodology; it provides a bridge that enables current data collection techniques to be used within the context of the accumulated information. This will eventually help all the participants in the construction industry to build up the knowledge derived from invaluable experience.

Keywords : Expenditure Forecasting, CBR (Case Based Reasoning), Residential Building

1. Introduction

Accurate information about the expenditures enables the managers to capture the problems as early as possible and to correct them with less difficulty. Along with planning incomes, forecasting expenditures is indispensable for ensuring a high level of credit reliability for a company. A cash flow statement indicates the movement of cash-in (income) and cash-out (expenditure) of a company over a certain period of time. The accuracy of the forecast depends on the precision of the prediction of the periodical income and expenditure.

As market competition becomes more intense, the importance of a comprehensive assessment of financial status has become crucial

for both the success of the project and the survival of the company. Despite its importance, cash flow planning, which is a part of the financial management, has not received enough attention from the construction industry. The reason for this is the ever-changing environment of construction projects, and consequently, the large amount of work involved in compilation and updating of such a plan.

With the application of statistical models, databases, and machine learning, researchers began to develop a set of new tools and techniques to analyze very large databases. However, most computer programs have been designed for processing the current independent project. Data from historic projects are invaluable asset for all the participants involved. If all this information were systematically structured to allow easy access to planners, it would enhance the certainties, as projects progress.

Construction managers are supposed to make decisions based on their experience and newly generated data as projects progress. In

* 일반회원, 위스콘신대 건설관리학과, 공학박사

many cases, there is little or no useful historical information available at the time the initial forecast is required. Thus, the early forecast must be based largely on subjective considerations. As more information becomes available, managers must modify their subjective estimates in terms of the actual data. However, the existing model mostly depends only on historic data without considering the current trends. The need for measuring the change over time for the current project is demanded in a rapidly fluctuating situation.

Most of the experienced knowledge extracted from a project, becomes useless without the interpretation by those people who were directly involved in the project, because they are the only ones who experienced and controlled all the tasks during the process. The obstacles in developing an appropriate model, which represents the nature of the construction industry, are derived from the following reasons: 1) a project-based system in the construction industry prevents sharing information and learning from historic data, 2) the lack of a standardized tool for analyzing the project progress exacerbates the efficiency in inheriting knowledge between projects, and 3) ignorance of the work-package structure as a project control tool misleads managers in capturing the problem.

Existing forecast models have weaknesses in capturing the significant variance derived from the set-off between elements, and work-packages. The Case-Based Reasoning (CBR) is regarded as the proper tool to resolve the problems listed above. It uses the knowledge represented in specific cases to solve a new problem based on the similarities between the new problem and the available cases.

The purpose of this research is to develop the dynamic forecasting model for the periodic expenditure over construction periods, which provides the basic information for the financial status of the project. To overcome the above-mentioned deficiencies and help managers make appropriate decision, the following factors are taken into consideration: 1) forecast based on information derived from respective work-packages, 2) diagnosis of sources of excessive variances, 3) generation of a standardized distribution function for each work-package, 4) providing the standardized S-curve for monthly expenditure throughout all construction periods, and 5) suggestion of the number of similar projects selected at each phase as projects progress for forecast with more accuracy. Another objective is to extract as much knowledge as possible from similar projects, which saves significant time and effort and increase the efficiency in

managing projects.

The collected data to build the systems were limited to residential building projects with lump sum contracts and a design-bid-build delivery type. The research was also limited to projects larger than \$9,000,000.

Some assumptions had to be made to maintain the problem at a manageable level. First, no cost indices compensating the difference in time frame were applied in comparing expenditure data between the projects in the data set. Though the researcher established the time laps between 1995 and 2003 as the horizon for this research, any modification based on the different fluctuations in material, wages, and equipment costs were not taken into account. All the values represented in terms of proportion against the costs at completion are able to standardize the data over the different time frame based on the assumption that the identical inflation rate were applied to the prices among the categories: e.g., material cost, wages, and equipment rental costs, etc. Secondly, It was assumed that all monthly cost data would be obtained on the last day of the month for which they would be accumulated up to that point. Lastly, the possible difference derived from various construction areas was disregarded in this research. To obtain a more reliable result, the distinction according to the location - the urban area and suburban area - should be reflected in developing the model.

2. Literature Reviews

2.1 S-Curve Model

The S-curve has been regarded as an effective tool for financial planning, monitoring and controlling projects. It is also recognized as a powerful tool for producing a forecast of future financial commitments.

2.1.1 Development of the S-Curve

The history of the S-curve dates back to the 19th century. Gompertz (1825) derived an asymmetrical logistic curve for graduating mortality data, which was also used for the population growth model. Since then several major advances have been undertaken by a number of researchers. More varied studies including logistic, log-logistic, normal integral, and log-normal have been conducted. Extensive application of the growth curves has been conducted in the business area. Preece and Baines (1978) examined the four different models for describing human growth. Meade (1984) applied the growth curve to generate forecasts of market

development.

2.1.2 The Application in the Construction Projects

A number of studies have been performed in budgeting, forecasting and monitoring construction projects' expenditure by utilizing mathematical models.

Drake (1978) used a mathematical model for predicting cost and duration at an early stage. Hudson (1978) developed the DHSS model to forecast the expenditure pattern of hospital projects. Bromilow (1979) adapted the exponential equation for the mathematical expression. Cook and Jepson (1979) applied a parabolic equation for predicting the expenditure pattern. Peer (1982) used regression techniques to produce a cash flow forecasting prior to construction. Bemy and Howes (1982) suggested the integrated model of polynomial and exponential curve for expenditure forecast. Singh and Woon (1984) exploited the features existing in the cash flow, and concluded that there is correlation between the project type and its expenditure pattern. Pattern (1982) emphasized the importance of cash flow information as a management tool. He presented a dynamic model for forecasting characteristics of a typical construction project. The model used the historical data as the basis of its information. Singh and Lokanathan (1992) expand the ability of a model to consider the interest charges and the internal rate of return for a given project based on the maximum negative cash flow.

2.1.3 The S-curve as a Forecasting Tool

The S-curve method is a commonly used tool for predicting the expenditure patterns of construction projects. This graph is generated by plotting the cumulative expenditure against time, of which axis are usually represented in percentile form. Since the incurred expenditure over the construction duration assumes a growth pattern, it forms the S-shaped curve. The use of the S-curve is based on its ability to forecast an expenditure pattern at an early stage, enabling the project manager to have better understanding of the project's future situation.

2.2 Cash Flow Forecasting

The importance of cash flow forecasting cannot be overemphasized. Cash flow forecasting provide valuable early warning systems which enable managers to recognize problems and solve them.

2.2.1 Detailed Approach for Cash Flow Forecasting

In calculating cash flow, a variety of information, which includes a number of items, resources, time intervals, payment periods, payment delay, amounts of retention, and the project cost breakdown structure, is required.

The calculation process demands tedious tasks, which necessitated the use of computers. Peterman (1972) developed a computer model to forecast contractors' cash flow based on contract bar charts and unit price information. Allsop (1980) linked the unit costs from estimating program to obtain cash flow. Jamieson (1980) simplified this process by the application of a cost curve using an average rate of markup.

2.2.2 Value Curves Approach for Cash Flow Forecasting

More efforts was devoted to develop the simplified model for cash flow forecasting. Bronmilow and Henderson (1974) used four building projects to establish their value S-curve. Ashley and Teicholz (1977) developed a model base on the value curve to assist in the analysis of cash flow throughout the whole project life. Oliver (1984) analyzed projects collected from three construction companies and concluded that projects are individually unique and pursued respective value curves based on historical data. Singh and Woon (1984) developed S-curves for high-rise commercial and residential buildings. Bemy and Howes (1982) modified a standard model to reflect the specific form of individual projects. Kenley and Wilson (1986) developed the value S-curve assuming that there should be errors caused by discrepancies from the average trend line rather than random error. Consequently, they claimed that each project has an individual line of central tendency. This concept was continuously explored and developed to claim that individual building projects with different sizes and types have their own characteristic cash flow.

2.2.3 Expenditure Forecasting

Forecasting the expenditure of a construction project has a critical role in the survival of company. A forecast of expenditure in construction is made based on the expected cost amount and construction period. As the project progresses, actual expenditure is compared with the forecast providing quick detection of any unfavorable trend. Based on actual expenditure to date and the expected final cost, the expenditure pattern of the current project is decided. Subsequent actual expenditure is compared with the new

forecast, which would be revised again and continuously updated. The incurred cumulative expenditure over time can be represented by a graph which takes the form of an S-curve. The traditional methods of forecasting the expenditure pattern demand huge information about the project as well as the laborious work and times. However, with the significant role of computers, a number of models have been developed, which are based on the use of a mathematical expression. The S-curves method has been widely utilized for the financial management of construction projects in predicting the future financial commitments including planning, monitoring and controlling.

2.3 Case-based Reasoning (CBR)

Case-based reasoning (CBR) is a major paradigm in automated reasoning and machine learning. In CBR, the system solves a new problem by noticing its similarity to one or several previously solved problems and by adapting their known solutions instead of working out a solution from scratch.

2.3.1 CBR Process

The process to describe CBR can be explained as followings:

- Retrieve the most similar case(s).
- Reuse the case(s) to attempt to solve the problem.
- Revise the proposed solution if necessary
- Retain the new solution as a part of a new case.

According to these steps, Aamodt and Plaza (1994) describe a CBR as a cyclic process comprising the 4 REOS; Retrieve, Reuse, Revise and Retain.

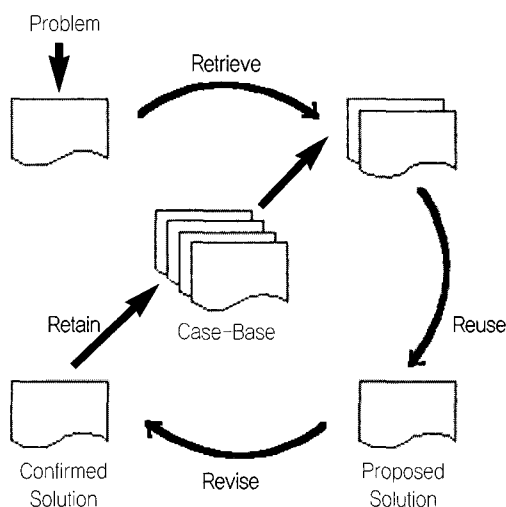


Figure 1. Case-based Reasoning Cycle

2.3.2 Practical Application

The past two decades have seen a number of trials of applying the CBR algorithm to the real world, especially in areas such as law, medicine, and strategic planning, where a huge amount of data has been accumulated. In the following sections, some of the most significant CBR systems developed from 1980 onwards are summarized especially in two categories: forecasting and the construction industry.

1) Examples in Forecasting

Stotler (1992) applies a CBR system for cost and sales prediction under uncertainty, which leads to an increase in business profit. He tried to retrieve the similar example in the past, and use that knowledge for business planning, e.g., staffing, sale volume, financing, etc. Lee et al. (1995) describes a CBR system that develops forecasts for cash flow accounts. From the proposed system in his research, fuzzy integrals are used to calculate the synthetic evaluations of similarities between cases instead of the usual weighted mean.

2) Examples in Construction

Domeshek and Kolodner (1992) developed ARCHIE, which is a CBR aid for the conceptual design of buildings. This model emphasizes case presentation and decomposition of a large case for easy understanding by the user. Its main goal was to capture and disseminated lessons learned from design experience so that future designers could avoid repeating similar mistakes. Maher and Zhang (1993) developed CADSYN which has cases decomposed into subparts during the representation of cases. It was realized through more flexible and efficient reuse of cases, wherein irrelevant information is pruned from consideration for each given design problem. Flemming (1994) proposed the SEED model with a goal to support the tasks at early phases of building design. SEED provides a systematic computational support for the rapid generation of designs with respect to recurring building types. Similar solutions in past projects are stored and retrieved as a prototypical designs. Dzeng (1995) applied CBR in developing a model that could generate a construction schedule for a new project based on the description of a new project and by using previous cases. It was devised to generate schedule for a new project by reusing schedules of cases that are similar to it. Soibelman (1999) developed a model that aim to assist engineers in the conceptual phase of the structural design of tall buildings by providing them with organized and reliable information. The objective is to provide designers with adapted past design

solutions with the help of a multi-reasoning mechanism creating a support system to enhance creativity, engineering knowledge and experience of designers.

3. Data

3.1 Data Characteristics and Analysis

The data required for the model were the monthly actual cost as well as the project summary. The structure and elements of the original data were based on those of the company that provided the data. Data from 88 completed residential building projects were analyzed in developing the model. The collected data could be divided into two main categories: project summary and monthly cost data.

3.2 Project Profiles

All the data were collected from the residential building projects that were constructed between 1993 and 2002. Table 1 reports descriptive statistics on each category of the project summary. The table depicts the mean, median, and standard deviation (StDev) for the following variables: duration, floor area, stories below ground, storey above ground, the number of buildings, the number of total housing units, and final costs at completion. The sample size corresponds the number of projects in the database. Statistical analyses were accomplished using Minitab for Windows v.13, and Microsoft EXCEL 2000.

Table 1. Projects Characteristics

Variable	N	Mean	Median	TrMean	StDev
Duration	88	29,386	29,000	29,288	4,290
Floor Area	88	907270	638573	790516	897288
No. Base	88	1,6364	2,0000	1,6000	0,6640
No. Floor	88	20,307	20,000	20,263	4,499
No. Bldg	88	7,091	6,000	6,538	5,364
No. Unit	88	591,9	452,5	563,9	401,6
Cost	88	41829	34154	39568	27557

The mean construction period for the sample residential building projects was 29.4 months. The mean value of the floor area was 907,270 square feet; the minimum value was 180,975 square feet and the maximum value was 6,995,796 square feet. The mean number of basement stories was 1.6. The mean number of building stories was 20.3. The mean value for the total units of the projects was 592 units. The mean value for the number of buildings was 7.1. The mean cost was \$41,829,100.

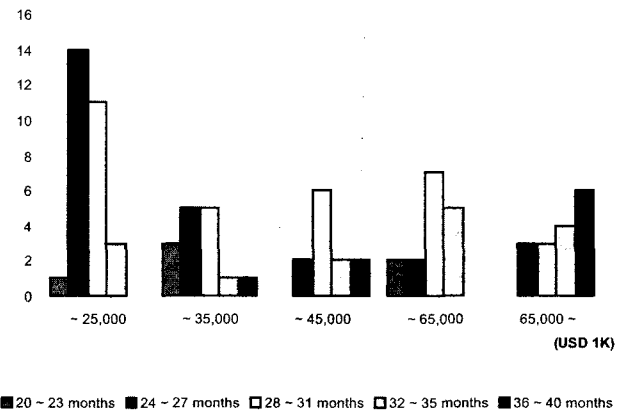


Figure 2. Projects Characteristics (cost & duration)

All the monthly-based cost data have been transformed into percentile value against the total cost for the purpose of standardization. Continued data collection using the methodology developed will serve to enable a more precise forecasting in the future. The primary focus of this research is to predict monthly expenditure based on the cost data at the work-package level from past projects. To match the current project to the similar ones, it was necessary to develop a process to standardize the variables and project durations.

3.3 Developing the Standard Curve

For this research, a quantitative study design is sufficient for the purpose of the proposed study. Quantitative methodologies manipulate variables and controls to predict monthly expenditure as projects proceed. A wide range of statistical techniques was utilized to predict dependent variables regarding expenditure in residential building projects. Establishing the cumulative curve as well as a

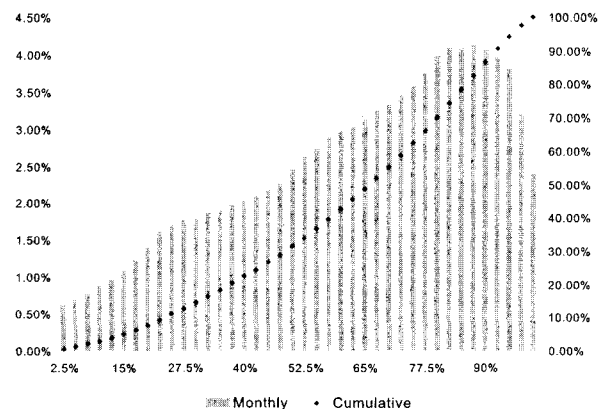


Figure 3. Progress Distribution and Cumulative Curve

progress distribution is very important. A significantly large sample size is required to obtain an accurate curve. Progress at each monitoring point is calculated and assembled into the standardized curve. The curve obtained from the sample might be rough. Some curves might be inconsistent with the overall curve shape. Nevertheless the discrepancy between the individual projects' S-curves and standardized one, it provides a base for controlling and planning at early stage. Progress distributions over time and cumulative curves are developed based on averaged data from the 88 dataset.

4. Development of the Forecasting Model

The purpose of developing the new model was to generate the precise prediction of periodic expenditure as a project proceeds by choosing similar projects as references for forecasting instead of all cases in a database. To analyze a large number of collections for this research, appropriate selection of computer tools were required, which would enable automation of the repetitive tasks with efficiency.

4.1 Theoretical Framework

The primary research concern will be focused on finding the efficient way of matching the similar projects in the past to the current one. The development of the forecasting model follows four steps: 1) data transformation; 2) development of standardized

expenditure pattern curves for respective projects in the database; 3) case matching; and 4) the development of an expenditure pattern curve for the current project in progress based on data derived from similar projects in the past.

The first step of the process is to transform monthly cost data into percentiles from dollar amounts. Thereafter, another conversion is performed to synchronize the different time frame into the same one with the current project. The second step is to develop the standardized expenditure pattern curves for individual projects. The case matching is applied in the third step for examining the similar projects compared with the current project in progress. The output of matching process is derived from the sum of squares of differences (SSD) between matched cases. At this stage, all the projects in the database are arrayed in the order that a project with smaller SSD coming first. Finally, forecasted expenditure curves are generated in the fourth step. For its validation, the results are compared with those of the existing method. The appropriate number (or proportion) for selecting projects from the database will be determined at this stage.

4.2 Model Development

In the model development, the first step is to define the case structure. The cases here will represent individual projects of which components are project features and respective monthly expenditure patterns from the 20-work package level. The format of input data should hold the consistency complying with the current accounting system by which projects are controlled.

The data analysis tool consists of four modules: 1) data entry; 2) data standardization; 3) statistical analysis and matching similar cases; and 4) forecast generation. These four modules comprise the new forecasting model that is developed in this research. More detailed descriptions for each model components are depicted below.

4.2.1 Data Entry

The first module, data entry, is aimed at avoiding extreme modification from the existing system. It is designed to provide an easy means by which an end-user can input data into the system. The first step of the program input request is to designate the construction duration in months according to that of the current project.

4.2.2 Data Standardization

The data standardization module performs the standardization of the time variable for the horizontal axis and the percentage complete

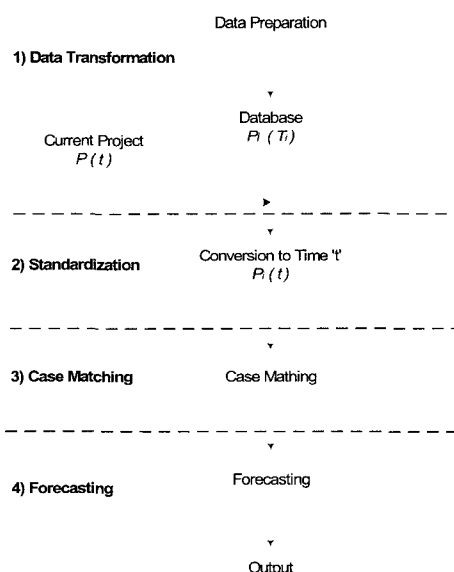


Figure 4. Theoretical Framework

for the vertical axis. The need to generate the data of a consistent number of data points for each project is crucial for the success of the newly developed forecasting model. The number of data points being generated is determined by the duration of the current project in progress. When the duration of the currently monitored project is 30 months, all the data points of previous projects in the database will be converted into 30 splits by applying either interpolation or extrapolation. From the expenditure curve for each work package, the assumption, that costs were expended in a linear fashion between the two adjacent points, is acceptable. The number of splits of all projects in the database is synchronized with the time frame of the project in progress.

4.2.3 Case Matching

The main performance at this stage is to compare the current project to previous ones in the database, and retrieve similar cases for a more precise forecasting. The degree of similarity between the current projects and ones in the database is determined by the sum of squares of difference (SSD). The new model reads the complete set of projects in the database and convert them into the identical time frame with the project in progress. The results of SSD can be obtained from 20-work package analysis. This process provides the available data a more accurate forecast by referring to only similar projects based on expenditure values at each work package level. The formula (1) explains how to calculate SSD of the 20- work package level analysis. Afterward, a calculation process takes place, and finally forecasts the remaining time points in advance based on the values derived from the similar projects. The appropriate number for selecting projects will be further addressed in the model validation part.

$$SSD = \sum_{i=1}^I \sum_{j=1}^n (W_{fij} - W_{dij})^2 \quad (1)$$

Where,

W_f: Currently forecasted project

W_d: Projects in database (j=1, 2, ..., M), M: number of projects in database

i: Work package (i=1, 2, ..., I), For 20- work package level analysis, the value of I will be 20.

j: month (j=1, 2, ..., n), n: forecast executed month

4.2.4 Forecasting

The program then calculates the average values derived from the

selected projects, and fills out the remaining blank data points with the average values of those matched projects. The model validation module enables users to modify the number of referring projects by their own decisions. The quantity for the ordinate axis is the value as a percent of the total expended. The horizontal axis measures time as a percent from the beginning of the construction.

4.3 Model Validation

The purpose of this section is to show how well the newly developed forecasting methodology predicts values for monthly expenditure. The above analysis is conducted at three different forecasting periods of times, which are 35%, 50% and 70% completion of the project. To prove the effectiveness and accuracy of the new methodology, the results will be compared with those of a conventional methodology that refers to the average of the whole database.

4.3.1 Measure of Forecasting Performance

The completed project was selected and evaluated for the purpose of model validation. Two separate modules are required to evaluate the model performance by % Error over the remaining construction duration and thereafter the forecasting point. One is to produce the forecast value at ith month (X_{F_i}), which is averaged from the selected projects. It is calculated by employing the formula (2) below:

(n: number of selected projects as references)

$$X_{F_i} = \frac{\sum_{j=1}^n X_{F_{ij}}}{n} \quad (2)$$

The other module is to generate % Error at the forecasting time, ith month. It is calculated by employing the formula (3) below:

$$\%Error \text{ at } i^{th} \text{ month} = \frac{|X_{F_i} - X_{A_i}|}{X_{A_i}} \times 100 \quad (3)$$

(X_{A_i}: actual data at ith month,

X_{F_i}: forecasted value at ith month from model)

Above two modules are integrated into one and provide overall model performance from the forecast conducted month (k) to the end of the construction period (N). It is calculated by employing the

formula (4) below:

(N: construction duration, k: forecasting occurred month)

$$\%Error = \frac{\sum_{i=k}^N \left| \frac{\sum_{j=1}^n \frac{X_{Fij}}{n} - X_{Ai}}{X_{Ai}} \right|}{(N - k)} \quad (4)$$

4.3.2 Evaluation on Forecasting Performance

The issue addressed in this section is to figure out the appropriate number of projects selection that provides the most accurate forecasting for monthly expenditure. The forecasting performances at three different phases (35%, 50%, and 70% completion of the projects) are measured and analyzed, as the number of projects selected increases. The project in progress is matched to the similar projects according to the results of the sum of squares of difference (SSD).

The summary of % Error results from testing the case project is given in Figure 5, as the number of projects selection increases 1 to 88. In its matching process, all the projects in database are ranked by the SSD values from 20-work package level. Regarding the forecasting evaluation from 35% completion, it can be seen that the % Error decreases as the number of selected projects increases until the adequate model is reached. Therefore, the adequate model for this dataset is to choose 12 projects as base reference for forecasting residential building projects. When the user selects 12 similar projects out of an 88 project dataset, the new model provides the prediction of monthly expenditure with a significant value of 0.76% Error. If all 88 projects are averaged as in traditional method, the value of 5.82% Error will be obtained.

It is encouraging that both results at other forecasting points (50% and 70%) also represent significant improvements in forecasting the monthly expenditure.

When prediction was conducted at the 50% completion time, the new model provided a forecast with an accuracy of 0.69% Error, based on the averaged curve of 19 projects. The % Error would be 2.69%, if all 88 projects were averaged in generating the standardized curve. Similarly, the accuracy was improved from 2.67% Error up to 0.78% Error by choosing 8 projects instead of 88 projects.

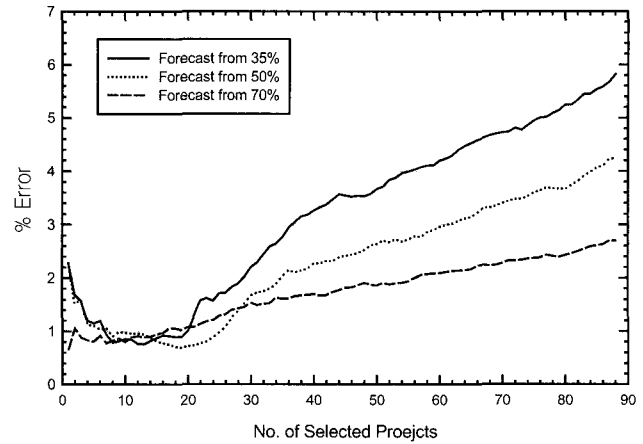


Figure 5. %Error Summary at Three Different Phases

5. Conclusion

This dissertation has taken an in-depth and focused look at one central issue: enhancing the forecast accuracy for monthly expenditure of residential building project based on the work-package level. This issue was addressed through the use of the Case-Based Reasoning concept by matching and retrieving similar projects from past projects. Result of the analyses indicated that the forecasting performance could be significantly improved by applying the new model, which is developed in this study.

The developed model produces forecasting with an increase in accuracy of 1.73 % Error at the point of 35 % progress, which would have been 4.39 % in the traditional methodology. Similar improvements are shown in from the point of 50 % and 70 % progress, which are 1.21 from 3.17 % Error, and 0.69 from 1.70 % Error respectively.

The primary contribution of this research is to increase the forecasting accuracy for expenditure pattern at each phase. The application of the new model is useful for forecasting the expenditure of residential building projects as they progress. Besides saving time, expenses, and the efforts of the project managers, there are several contributions both to the academic area and the construction industry. These include extending the forecasting horizon, applicability to different monitoring points, and enhancing the knowledge build-up and data exchange by referring to the similar projects in the past. The practical contributions are outlined below.

- Extending the Forecasting Horizon
- Deeper Insights into the Status of Project

- Generation of Expenditure Pattern Curves of Each Work Package
- Efficiency Improvements in Knowledge Build-up and Data Exchange

The new forecasting model developed in this research, which suggests the predicted values based on previous projects, has meaning beyond just a forecasting methodology. It provides a bridge that enables current data collection techniques to be used within the context of the accumulated information. This will eventually help all the participants in the construction industry to build up the knowledge derived from invaluable experience.

REFERENCES

- Aamodt, A. and E. Plaza (1994). "Case-Based Reasoning: Foundational Issues, Methodological Variations and System Approaches." *AI Communications* 7: 39-59.
- Allsop, P. (1980). *Cash Flow and Resource Aggregation from Estimators Data*. Construction Management, Loughborough University of Technology.
- Ashley, D. B. and P. Teicholz (1977). "Pre-Estimate Cash Flow Analysis." *Journal of Construction Division* 10(3): 111-113.
- Bemy, J. and R. Howes (1982). "Project Management Control Using Real Time Budgeting and Forecasting Models." *Construction Papers* 2: 19-40.
- Clark, F. D. (1985). "Labor Productivity and Manpower Forecasting." *Transaction, AACE: A.1.1-A.1.7*.
- Cook, B. and W. B. Jepson (1979). "Financial Control for Construction Firms." London, MacMillan.
- Gompertz, B. (1825). "On the Nature of the Function Express of the Law of Human Mortality." *Philosophical Transactions of the Royal Society XXXVI*: 513-585..
- Hudson, E. W. (1978). "DHSS Expenditure Forecasting Method." *Chartered Surveyor - Building and Quantity Surveying Quarterly*: 42-54.
- Koldner, K. (1992). "An Introduction to Case-Based Reasoning." *Artificial Intelligence Review* 6: 3-34.
- Makridakis, S. and S. C. Wheelwright (1985). *Forecasting Methods for Management*. New York, John Wiley & Sons.
- O'Leary, T. and S. N. Tucker (1996). *Techniques and Tools for Project Cash Flow Prediction in the Australian Construction Industry*. Proceeding of 40th Annual Meeting of AACE, Vancouver, Canada.
- Peer, Shlomo (1982). "Application of Cost-flow Forecasting Models." *Journal of the Construction Division, Proceedings of ASCE, 8(CO2)*: 226-232.
- Russell, J. S., Lawrence, Edward J. Jaselskis, et al. (1988). "Continuous Assessment of Project Performance." *Journal of Construction Engineering and Management(March)*: 64-71.
- Singh, S. and P. W. Woon (1984). *Cash Flow Trends for High-Rise Building Projects*. Proceedings of the 4th International Symposium on Organization and Management of Construction, University of Waterloo, Canada.
- Soibelman, L., and Kim, H. J. (2001). *Generating Construction Knowledge with Knowledge Discovery in Database*. Eighth International Conference (ICCCBE - VIII), Palo Alto, CA, ASCE.
- Stottler, R. H. (1992). "Case-Based Reasoning for Bid Preparation" *AI Expert (March)*: 44-49.
- Teicholz, P. (1993). "Forecasting Final Cost and Budget of Construction Projects." *Journal of Computing in Civil Engineering* 7(4): 511-529.
- Thomas, H. R., and Kramer, D.F (1987). *The Manual of Construction Productivity Measurement and Performance Evaluation*, Pennsylvania State University.
- Watson, I. (1997). *Applying Case-Based Reasoning: Techniques for Enterprise Systems*. San Francisco, Morgan Kaufmann Publishers, Inc.