FORECASTING PROJECT PROGRESS AND EARLY WARNING OF PROJECT OVERRUNS WITH PROBABILISTIC METHODS

A Dissertation

by

BYUNG CHEOL KIM

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 2007

Major Subject: Civil Engineering

FORECASTING PROJECT PROGRESS AND EARLY WARNING OF PROJECT OVERRUNS WITH PROBABILISTIC METHODS

A Dissertation

by

BYUNG CHEOL KIM

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Kenneth F. Reinschmidt
Committee Members,	Stuart Anderson
	David Trejo
	Daren B.H. Cline
Head of Department,	David V. Rosowsky

December 2007

Major Subject: Civil Engineering

ABSTRACT

Forecasting Project Progress and Early Warning of Project Overruns with Probabilistic Methods. (December 2007)

Byung Cheol Kim,

B.S., Seoul National University; M.S., Seoul National University

Chair of Advisory Committee: Dr. Kenneth F. Reinschmidt

Forecasting is a critical component of project management. Project managers must be able to make reliable predictions about the final duration and cost of projects starting from project inception. Such predictions need to be revised and compared with the project's objectives to obtain early warnings against potential problems. Therefore, the effectiveness of project controls relies on the capability of project managers to make reliable forecasts in a timely manner.

This dissertation focuses on forecasting project schedule progress with probabilistic methods. Currently available methods, for example, the critical path method (CPM) and earned value management (EVM) are deterministic and fail to account for the inherent uncertainty in forecasting and project performance.

The objective of this dissertation is to improve the predictive capabilities of project managers by developing probabilistic forecasting methods that integrate all relevant information and uncertainties into consistent forecasts in a mathematically sound procedure usable in practice. In this dissertation, two probabilistic methods, the

Kalman filter forecasting method (KFFM) and the Bayesian adaptive forecasting method (BAFM), were developed. The KFFM and the BAFM have the following advantages over the conventional methods: (1) They are probabilistic methods that provide prediction bounds on predictions; (2) They are integrative methods that make better use of the prior performance information available from standard construction management practices and theories; and (3) They provide a systematic way of incorporating measurement errors into forecasting.

The accuracy and early warning capacity of the KFFM and the BAFM were also evaluated and compared against the CPM and a state-of-the-art EVM schedule forecasting method. Major conclusions from this research are: (1) The state-of-the-art EVM schedule forecasting method can be used to obtain reliable warnings only after the project performance has stabilized; (2) The CPM is not capable of providing early warnings due to its retrospective nature; (3) The KFFM and the BAFM can and should be used to forecast progress and to obtain reliable early warnings of all projects; and (4) The early warning capacity of forecasting methods should be evaluated and compared in terms of the timeliness and reliability of warning in the context of formal early warning systems.

DEDICATION

To my family

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere gratitude to my advisor, Professor Kenneth Reinschmidt, for his continuous support, invaluable advice, and warm encouragement throughout the course of this research. With comprehensive knowledge and experience, he filled our discussions with incredibly insightful ideas and perspectives over various issues, which gradually opened my eyes to the importance of critical thinking balanced with the reality of having to make timely decisions. What I am most thankful for is his unbelievable patience in guiding me beyond academic bounds through a time of personal difficulties. I thank him sincerely for being a true mentor.

I also would like to extend my gratitude to my thesis committee members, Professor Stuart Anderson, Professor David Trejo, and Professor Daren Cline for many valuable comments and their time and effort in evaluating my dissertation. With their fresh views and useful suggestions on various issues, I was able to make a lot of improvement.

Last, but certainly not least, I want to thank my family for their love and encouragement. My parents, Yeon-kwon Kim and Hye-hun Jeong, have always been with me, giving me the strength to move forward and the wisdom of being thankful for everything and for each day. My wife, Seung-jin Lee, has been always there with her lovely smile and continuous encouragement. I am blessed to have a son and a daughter, Siyoung and Jayoung. Through the precious time with our children, my wife and I have a better appreciation of the true value of family.

In closing, I believe that the love and support from these people and those who I failed to mention should be shared with many more people in the future. In that sense, I feel like my true gratitude starts from here.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE Page

FIGURE Page

FIGURE Page

LIST OF TABLES

TABLE Page

CHAPTER I

INTRODUCTION

1.1 Project Management and the Role of Forecasting

Reliable forecasting is a critical component of project planning, controlling, and risk management. When at-completion project duration and cost are forecast before the start of a project, the process is carried out as a part of project planning and its results provide the baseline plan intended to complete the project on time and within budget. Once a project gets started, actual performance is monitored and analyzed to revise the estimates of the remaining work. The major purpose of execution phase forecasting is to obtain an early warning signal so that corrective or preventive actions may be taken in a timely manner. Such predictions need to be revised and compared with the scheduled completion time and the available budget. Therefore, the effectiveness of project controls relies on the capability of project managers to forecast final cost and completion time in a timely manner.

This research focuses on the schedule forecasting problem of on-going projects. Typically, three alternatives are available for project managers to update the original estimates, depending on the decision maker's perception of the relationship between past and future performance: (1) forecasting based on the original estimate; (2) forecasting based on a new estimate; and (3) forecasting based on the original estimate modified by

---

This dissertation follows the style and format of the *Journal of Construction Engineering and Management*.

past performance information (PMBOK® Guide 2004). The first two approaches, however, are valid only when any actual performance data observed from a project is considered irrelevant to the future performance of remaining jobs. In such cases, the remaining work is considered a separate project. The major focus of this dissertation is the third case in which project duration and cost at completion are updated using both the original estimate and actual performance data up to the time of forecasting. The third type of forecasting can is referred to as "projective forecasting" while the first two types are referred to as "estimative forecasting."

1.2 Motivations

This research is motivated by three observations regarding project management research in academia and practices in project management. The first motivation of this research is the presence of uncertainty in both future project performance and current performance measurement. Needs for forecasting some events arise only when there is uncertainty about the future. Furthermore, the level of uncertainty in forecasts may influence decisions about planning and controlling projects. Like all forecasts, forecasting the final outcomes of on-going projects is subject to uncertainties and prediction error. Effective forecasting methods for dealing with uncertainty are, of necessity, probabilistic. In addition, any prediction of project duration and cost should be accompanied by a level of confidence on the predicted values because those measures of uncertainty are also essential information for the project manager. The most common way of measuring the level of confidence that a decision maker put on predicted value is to estimate a prediction interval. Traditional approaches such as the Critical Path Method (CPM) and Earned Value Management (EVM) are deterministic and provide point forecasts. As a result, they do not provide information on the prediction bounds based on the likely accuracy of forecasts. The program evaluation and review technique (PERT) provides a semi-probabilistic evaluation of project duration. However, it has been criticized for systematic underestimation due to neglecting the influence of near-critical paths.

Before proceeding further, the term "prediction interval" needs to be defined precisely. In the literature on forecasting and decision making, prediction interval is often used interchangeably with confidence interval. In this dissertation, however, those two terms are carefully chosen according to the suggestion by Armstrong (2002), which is "The term confidence interval is usually applied to interval estimates for fixed but unknown parameters. In contrast, a prediction interval is an interval estimate for an (unknown) future value." A prediction interval consists of an upper and a lower limit at a prescribed probability, which are referred to as prediction bounds.

The second motivation of this research is the lack of reliable and consistent forecasting tools available to contractors, project managers, or program managers in public agencies, who would be the primary beneficiaries of this research. The traditional approaches do not provide forecasting methods that are consistently applicable to both schedule and cost predictions. For example, the fact that the standard EVM technique for forecasting the final cost at completion is not applicable to forecasting the project duration at completion has been criticized both by researchers (Short 1993; Vandevoorde and Vanhoucke 2006) and by practitioners (Leach 2005; Lipke 2003). As a result, project managers often resort to network-based schedule management techniques such as CPM for schedule prediction (Leach 2005). However, concurrent usage of EVM and CPM in the same project is likely to lead to inconsistent time and cost predictions because these approaches are based on inconsistent assumptions about the relationship between past performance and future performance. For example, a common cost forecasting technique in EVM is based on the assumption that future productivity (Cost Performance Index) will be the same as past productivity. In the typical critical path method, however, the durations for the remaining work activities are often fixed at the original estimates.

Another motivation of this research is the lack of a comprehensive and integrative forecasting framework which integrates all the information relevant to project performance predictions, such as detailed project plans, historical data, subjective knowledge from project managers' hands-on experiences, and measurement errors due to progress reporting rules or discrete reporting points. Exploiting the potential benefits of pre-construction planning information such as baseline project progress curves and probabilistic estimates of project duration and cost is critical to the accuracy of forecasting, especially during the early phase of project execution because of the lack of enough actual performance data to get statistically reliable predictions.

1.3 Problem Statement and Research Objectives

Aligning with the research motivations, the problem statement of this research is:

There is a need for reliable and consistent forecasting methods in construction project management. Currently available methods are mostly deterministic, overly simplified, or inconsistent in application and assumption, which make them unreliable or impractical. Advanced forecasting methods that integrate all relevant information and uncertainties into consistent project performance predictions in a mathematically sound way will enable project managers in the real world to make decisions in a more effective and efficient way.

The purpose of this dissertation is to improve the predictive capabilities of construction project managers by developing new probabilistic forecasting methods that integrate all relevant information and uncertainties into consistent schedule forecasts in a mathematically sound procedure usable in practice. To achieve this purpose, the research objectives are identified as follows:

1. **Probabilistic Forecasting Methodologies**: Improve project forecasting methods by explicitly identifying and accounting for uncertainties in project performance and errors in measurements, and by providing prediction bounds on the predicted values.

2. **Integrative Forecasting Methodologies**: Integrate all relevant information from different sources in a mathematically sound way. The methodologies to be developed should be based on information available from standard construction project management systems such as the work breakdown structure, the network schedule, CPM, or EVM.
3. **Consistent Forecasting Methodologies**: Develop forecasting methods that can be applied to both schedule and cost performance forecasting in a consistent way.

1.4 Research Hypothesis

To achieve the research objectives, two project performance forecasting methods are developed based on state-of-the-art theories and methods in decision making and forecasting. With the methods, the Kalman filter forecasting method and the Bayesian adaptive forecasting method, a series of parametric studies is carried out to test three research hypotheses which are developed to demonstrate the forecasting performance of new methods and to compare them with other methods currently used in the construction industry.

Hypothesis 1: The use of prior information, as used by Bayesian, along with actual performance data increases the quality of forecasting performance with regards to the accuracy, timeliness, and reliability of warning signals.

Hypothesis 2: The Kalman filter model and the Bayesian adaptive model outperform the conventional methods such as CPM and EVM with regards to the accuracy, timeliness, and reliability of warning signals.

Hypothesis 3: The relative performance of the Kalman filter and Bayesian adaptive forecasting varies depending on the types of information available at the time of forecasting. For example, the information about the baseline plan would be more useful in the Bayesian adaptive approach than in Kalman filter forecasting.

The results from the parametric study will provide useful information for developing some practical guidelines for potential users of the new methods.

1.5 Research Scope

This dissertation focuses on forecasting project schedule progress with probabilistic methods. From a literature review, it has been revealed that, despite a large literature on forecasts in various fields such as marketing, management, and engineering, very little research has addressed the forecasting issue in the context of construction project management. Especially, the literature on probabilistic project performance forecasting is very limited. Therefore, the research objectives and other issues addressed in previous sections require many original approaches in terms of methodology development, computer programming, data collecting, evaluating criteria development, and design of hypothesis tests. Some primary intellectual challenges in the research are identified and summarized in Table 1.1.

Table 1.1 Intellectual challenges in the research

Areas	Subjects	Challenges
Methodology Development	Kalman filter forecasting	❑ Incorporating the nonlinear baseline plan – S-curve – into forecasting algorithm. ❑ Developing algorithms for the cases in which the baseline plan is and is not available.
	Bayesian adaptive forecasting	❑ Developing S-curves library that covers, in combination or separately, a wide range of potential project progress curves. ❑ Developing easy and intuitive procedures for incorporating subjective information into the analysis in a systematic way.
Methodology Validation & Parametric Study	Artificial progress generation	❑ Developing a computer algorithm for random progress generation.
	Evaluation metrics	❑ Developing performance evaluation metrics for probabilistic forecasting methodologies.
	Design of hypothesis test	❑ Designing hypothesis test in a way that yields statistically significant results with minimum computation time.

1.6 Dissertation Overview

This dissertation is organized as follows.

- ❑ Chapter I addresses the outline of this research, including the motivations, problem statement, and the objectives and scopes of this research.
- ❑ Chapter II overviews the literature on project performance forecasting methods with an emphasis on practical issues. Three state-of-the-art methods in the construction industry are also reviewed.

- In Chapter III, the Kalman filter forecasting method is developed based on the general frameworks of project control and the Kalman filter. An example is presented to demonstrate the core properties of the Kalman filter forecasting method.
- In Chapter IV, the Bayesian adaptive forecasting method is developed based on Bayesian inference and S-curve models. Depending on the flexibility of S-curve models in fitting various progress curves of projects, two separate methods are developed: the multi-model Bayesian adaptive forecasting method and the BetaS-curve Bayesian adaptive forecasting method. Each method has its own unique properties and their performance is demonstrated with two numerical examples.
- An extensive parametric study is carried out in Chapter V. Using a large set of artificial projects, statistically meaningful results about the research hypotheses are obtained.
- Chapter VI addresses some practical issues regarding implementation of the new forecasting methods. Useful suggestions for project managers are made. In addition, step-by-step input guidelines for the new methods are presented.
- Chapter VII summarizes the contributions and major conclusion of the research and suggests some future research issues.

CHAPTER II

LITERATURE REVIEW

2.1 Introduction

The purpose of this chapter is to identify the important issues in developing a new forecasting method and the state-of-the-art methods in project performance forecasting in the context of decision making and construction management. It is often said that forecasting is an art rather than science. The literature review in this chapter focuses on identifying practical issues that must be taken into account in formulating new methods in order to make the methods more acceptable for potential users.

2.2 Forecasting Methods for Decision Making

Forecasting is an essential part of decision making under uncertainty. A need for forecasting arises *only* when there is uncertainty about the future and some aspects of the future can not be controlled (Armstrong 2002). If everything relevant to an event is certain and the future of the event is deterministically predicted or controlled based on what is known at the point of forecasting, any decisions about it can be made according to the decision maker's preference for expected outcomes. Otherwise, decisions should be made based on forecasts which account for the uncertainty about the future.

The limits of deterministic approaches and the need for probabilistic models in engineering and management decision making have been repeatedly addressed over the last four decades (Ang and Tang 1975; Barraza et al. 2004; Hertz 1979; Spooner 1974).

The use of probabilistic methods, however, is often avoided among practitioners, largely because of the lack of appropriate methods, the lack of sufficient reliable data, and the additional difficulty in dealing with uncertainty in a quantitative way during the decision making process.

Plenty of forecasting methods are available for engineering and management decision making and they can be characterized in several ways. Makridakis et al. (1982) classified forecasting methods into four groups: purely judgmental approaches, causal or explanatory methods, extrapolative (time series) methods, and any combination of the three. On the other hand, Al-Tabtabai and Diekmann (1992) classified forecasting models as econometric models, time-series models and judgmental models. With so many alternatives available, selecting the right method for a specific problem is itself a challenging decision. Some comprehensive overviews of the management forecasting methods and selecting the right method can be found in the literature (Armstrong 2002; Chambers et al. 1971). For example, Georgoff and Murdick (1986) evaluated 20 forecasting techniques according to 16 criteria to provide a guide for how to choose the best technique or combination of techniques.

Use of subjective information and judgment in forecasting is another important issue in forecasting literature. Some forecasting methods use objective data observed in a quantitative way, while others rely on subjective information. For example, the Delphi method makes forecasts by deriving a consensus from a group of experts through a sequence of questionnaires. For many forecasting methods, however, there is no clear-cut distinction between the forecasting methods that rely on objective data and other methods based on subjective information. Some forecasters believe that subjective information or the use of it is the second best alternative, which can be justified only when there are no objective data available (Chau 1995), while others suggest use of subjective information as a way of improving forecasting performance (Georgoff and Murdick 1986). For example, Al-Tabtabai and Diekmann (1992) conducted a study of applying social judgment theory to cost forecasting of a construction work package and asserted that both historical data and competent judgments based on experience and knowledge need to be included in a forecasting technique.

2.3 Evaluation Criteria for Project Performance Forecasting Methods

Evaluating forecasting methods, whether they are probabilistic or deterministic, is an essential challenge of all forecasters. Many approaches have been investigated over the last four decades since the seminal work by Bates and Granger (1969). A study (Carbone and Armstrong 1982) showed that accuracy was recognized as the most often used evaluation criterion for forecasting methods among both practitioners and researchers. Some empirical studies on accuracy of forecasting methods include extrapolation methods (Makridakis et al. 1982) and project performance forecasting methods (Teicholz 1993; Zwikael et al. 2000).

Makridakis et al. (1982) conducted an empirical study on the accuracy of extrapolation methods. Twenty four forecasting methods based upon time series analysis were applied to 1001 series and differences in performance were evaluated. In the study, accuracy, although not the only factor, was used as a single measure of overall forecasting performance. Teicholz (1993) described the characteristics of a desirable forecasting method as accuracy, unbiasedness, timeliness, and stability. In his study of forecasting final cost of construction projects, two forecasting methods – the moving average and the up-to-date average – were used to estimate unit cost, which was used to adjust the cost performance of future work. Instead of purely statistical measures of accuracy, such as absolute percentage error or mean square errors, Teicholz measured the average accuracy of project cost forecasting by the area enveloped by the actual final cost and forecast final costs from monthly measurements plotted against percent complete. In a similar way, timeliness was defined as the forecasting accuracy during the first 50% of the project duration. Zwikael et al. (2000) evaluated five forecasting models in EVM using three performance measures: the mean square error, the mean absolute deviation, and the mean absolute percentage error.

2.4 Conventional Methods for Project Performance Forecasting

In the project management community, the best practice for project performance forecasting is to use the earned value method for cost performance forecasting and the critical path method for schedule forecasting. In this section, basic features of those two methods are summarized and some limitations will be discussed. In addition, Monte Carlo simulation is addressed in the context of project cost and schedule forecasting.

2.4.1 Earned Value Management

The earned value method (EVM) integrates the project's scope, cost, and schedule by using a resource-loaded project schedule and provides a systematic way of measuring, analyzing, communicating, and controlling the actual performance of a project. EVM is endorsed by the Project Management Institute (PMI) (PMBOK® Guide 2004) and is in use by many public agencies (its use is mandated by NASA, DoD, DOE, and others) and by private owners and contractors to evaluate complex projects. It has the advantage of being universally applicable over a wide range of project types and sizes, because every project, no matter how large or complex, is represented by the three functions: the planned value (PV or the budgeted cost of work scheduled), the earned value (EV or the budgeted cost of work performed), and the actual cost (AC or the actual cost of work performed). Then both the schedule and cost performance of a project is measured in terms of four performance indicators:

- Cost Variance (CV) = EV – AC
- Cost Performance Index (CPI) = EV / AC
- Schedule Variance (SV) = EV – PV
- Schedule Performance Index (SPI) = EV / PV

The ultimate goal of EVM is to provide reliable early warning signals about the schedule and cost performance of a project (Fleming and Koppelman 2006). The standard EVM prediction for the project duration and cost at completion rests on the assumption that the cumulative performance index ($SPI_C(t) = EV(t)/PV(t)$ or $CPI_C(t) = EV(t)/AC(t)$) will represent the performance efficiency of the jobs in the future. Then, the estimate at completion (EAC) at time *t* is equal to the cost already spent (AC) plus the adjusted cost for the remaining work:

Reliable forecasting is a critical component of project planning, controlling

$$EAC(t) \cong AC(t) + \frac{[BAC - EV(t)]}{CPI_C(t)} = \frac{BAC}{CPI_C(t)} \tag{2.1}$$

where BAC is the Budget at Completion, or the original total budgeted cost. One estimate of duration at completion (EDAC) can also be calculated in a similar way.

$$EDAC(t) \cong t + \frac{[BAC - EV(t)]}{SPI_C(t)} \tag{2.2}$$

where the cumulative Schedule Performance Index (SPIC), which is defined as the ratio of the EV and the PV, represents the efficiency of schedule performance up to the time of forecasting.

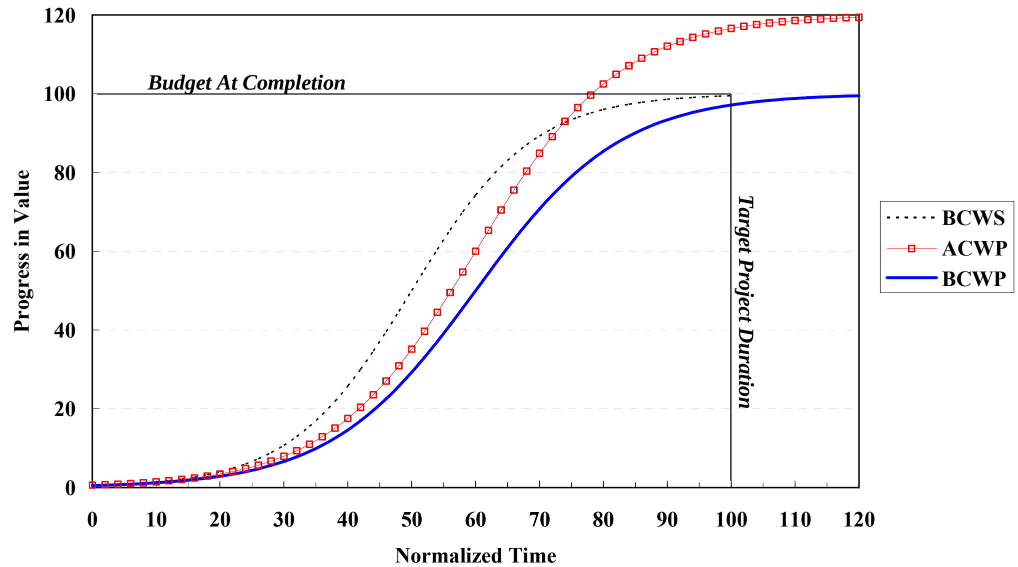
The cumulative CPI is known to stabilize once a project is twenty percent complete and "will not likely change by more than plus or minus 10% at the point of project completion" (Fleming and Koppelman 2006). Although the use of EVM forecasting formulas for cost performance has been supported widely, many modified versions of the standard formula have been suggested as summarized in Table 2.1 (Anbari 2003; Christensen 1993). Christensen (1993) conducted an extensive study of comparing various formulas for EAC forecasting, including linear and nonlinear regression models, and reported that the accuracy depends on the type of a system and the phase of a contract. It is apparent that different formulas should be applied to different situations and the accuracy of prediction will be assured only for appropriate matches.

Table 2.1 Extended formulas for the EAC. EAC=AC(*t*) +[BAC – EV(*t*)]/PF

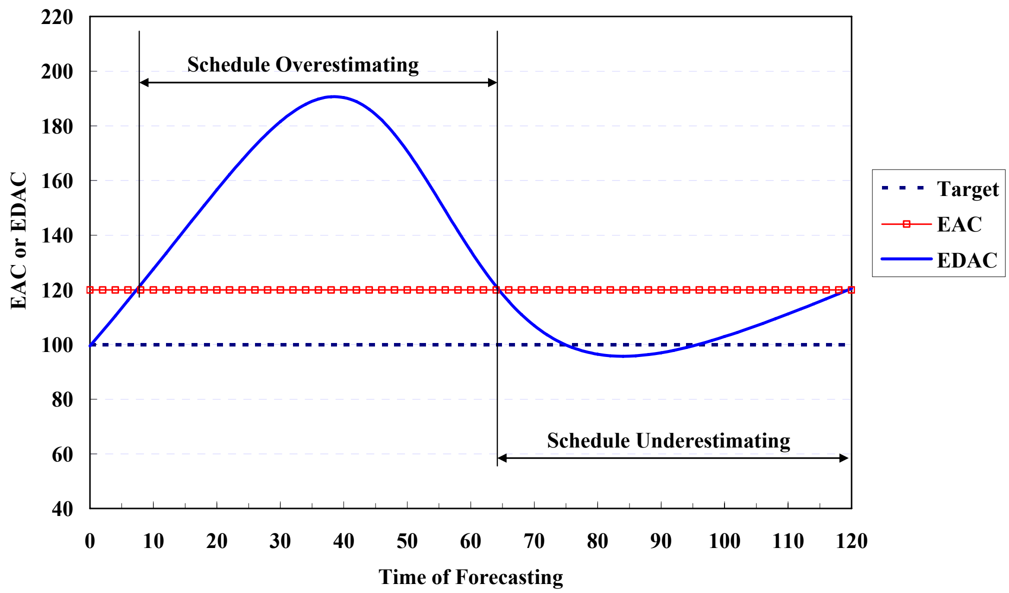
Types	Performance Factor	Description
Original	PF = 1	When past performance is not a good indicator of future performance.
Standard	PF = CPI_{cum}	The standard formula.
Composite Factor	PF = CPI_{cum}*SPI_{cum}	The product of CPI and SPI is called the critical ratio (Anbari 2003) or the schedule-cost index (Christensen 1993).
Moving average	PF = $CPI_m(t)$	Moving average of incremental CPI over the *m* latest time intervals. CPI_3, CPI_6, and CPI_{12} are often used.
Weighted	PF = w_1*CPI_{cum} + w_2*SPI_{cum}	Determination of the values of w_1 and w_2 is a judgment call of project managers. $0.5CPI_{cum}+0.5SPI_{cum}$, $0.75CPI_{cum}+0.25SPI_{cum}$, or $0.8CPI_{cum}+0.2SPI_{cum}$ can be used.
% Complete	PF = (%C)*CPI_{cum} + (1-%C)*SPI_{cum}	This is a modified version of the weighted method. Instead of using fixed weights, the weights change according to the percent complete of a project.

On the other hand, the schedule forecasting method using the cumulative PV and EV has been criticized for systematic distortion in results (Leach 2005; Lipke 2003; Short 1993; Sparrow 2005; Vandevoorde and Vanhoucke 2006). For example, Short (1993) demonstrated the erroneous behavior of Schedule Variance under two scenarios: when non-critical activities cause schedule variance and when certain activities are performed out of sequence from the baseline plan. Figure 2.1 shows the behavior of the EAC and the EDAC during project execution when a project undergoes 20% overruns in schedule and cost. While the EAC is accurately predicted from the beginning of the project, the EDAC overshoots the actual project duration, which is 120 time steps, as early as 10% of the planned project duration. Surprisingly enough, the EDAC at time 40 overestimates the project schedule delay by about 350% of the actual delay of 20 time steps. This simple example clearly demonstrates the limitations of the EVM forecasting approach.

To improve the capacity of EVM's schedule forecasting, several modified forecasting formulas have been suggested (Anbari 2003; Lipke 2003) as summarized in Table 2.2. Recently, Vandevoorde and Vanhoucke (2006) compared three different approaches for the EDAC. In the study, the authors demonstrated that the earned schedule method, originally proposed by Lipke (2003), was the only method that provides reliable forecasting results. Later, Vanhoucke and Vandevoorde (2006) reported similar results based on extensive simulation about the relative performance of the three formulas in Table 2.2 under different progress scenarios.



(a) EVM Progress Records under 20% Overruns



(b) EVM Forecasting

Figure 2.1 Distortions in EVM schedule forecasting

Table 2.2 Extended formulas for the EDAC

Types	EDAC(t)	Description
The planned value method	EDAC(t) = PD/PF PF = SPI or SCI	PD is total planned project duration.
The earned duration method	EDAC(t) = t + (PD-ED)/PF PF = SPI or SCI	The earned duration (ED) is the product of the actual duration and SPI, i.e. ED=AD*SPI
The earned schedule method	EDAC(t) = t + (PD-ES)/PF PF = SPI*	The earned schedule (ES) is the planned time to achieve the current EV. SPI* = ES/t.

Based on the previous study, the schedule forecasting formula based on the earned schedule is selected as the representative EVM schedule forecasting method in the following hypothesis tests. Given the three fundamental performance functions, the time performance index (TPI) is defined as

- ES(t) = Earned Schedule at time t, which is determined as the time based on the baseline plan to finish EV(t). Therefore, ES(t) is calculated from the equation,

$$PV\left(ES\left(t\right)\right)=EV\left(t\right) \quad \text{or} \quad ES\left(t\right)=PV\left[EV\left(t\right)\right]^{-1} \tag{2.3}$$

- TV(t) = ES(t) - t. TV stands for Time Variation.
- TPI(t) = ES(t)/t. TPI stands for Time Performance Index.

Then, the estimated duration at completion at time t is defined in terms of the planned duration at completion (PDAC) and TPI at time t.

$$EDAC\left(t\right)=\frac{PDAC}{TPI\left(t\right)} \tag{2.4}$$

2.4.2 CPM

The critical path method (CPM), which was developed by DuPont, Inc., in the 1950's, is the *de facto* standard schedule management tool in construction projects. Excellent introductions to CPM can be found in many sources (Meredith and Mantel 1995; Oberlender 2000). In a typical CPM, the total duration of an on-going project is determined based on the assumption that future tasks will proceed as planned regardless of past performance. For example, at the time of forecasting, actual completion dates of finished tasks and percent complete of on-going tasks are updated into a schedule network. Then, the starting and finishing dates of remaining tasks are deterministically calculated according to the precedence relations and the original estimates of individual task durations. The shortcomings inherent to classic CPM have been criticized since the early 1960s (Cottrell 1999; Lu and AbouRizk 2000). Recently, Galloway (2006a; 2006b) conducted an extensive survey on the use of CPM scheduling for construction projects and education in worldwide universities.

2.4.3 Monte Carlo Simulation

Monte Carlo simulation is probably the most common simulation technique in engineering and management. In dealing with a complex system, Monte Carlo methods represent uncertainty associated with some variables in the system with random numbers from the estimated probability distributions for those variables. A Monte Carlo simulation starts with drawing a set of random numbers from the distributions for the variables under consideration. Then a deterministic analysis is carried out to obtain a result based on the set of random values for variables. A Monte Carl simulation repeats these two steps over and over until some statistically significant results can be obtained.

In project management and decision making, Monte Carlo simulation, or simply simulation, has been used as a formal risk analysis technique for more than four decades. In a classic *Harvard Business Review* article, "Risk Analysis in Capital Investment," David Hertz (1964) showed the importance of risk assessment in business decision making with a Monte Carlo simulation. Van Slyke (1963) addressed typical problems in the program evaluation and review technique (PERT) with simulation. Because of the relative complexity in calculating total project duration from probabilistic estimates of component work packages, Monte Carlo simulation based on network schedules has been intensively investigated by several researchers (Finley and Fisher 1994; Hulett 1996; Lee 2005; Lu and AbouRizk 2000).

Recently, Barraza et al. (2005) conducted a study of probabilistic forecasting of project duration and cost using network-based simulation. In the study, the correlation between past and future performance is simplified by adjusting the parameters of probability distributions of future activities with the performance indices (for example, the Cost Performance Index as defined in the earned value method) of finished works. However, this method is not supported by any empirical or statistical evidence. Lee (2005) also presented a network-based simulation approach to compute the probability to complete a project in a specified time. In the cost estimating area, Monte Carlo methods were tested as a method for dealing with correlation between random variables (Chau 1995; Touran 1993; Touran and Wiser 1992).

CHAPTER III

KALMAN FILTER FORECASTING METHOD

3.1 Introduction

A new probabilistic forecasting method, the Kalman filter forecasting method (KFFM), is developed here for forecasting project progress and probability distribution on project duration at completion. The KFFM is formulated based on the general frameworks of project control and the Kalman filter.

Project control is a continuous process in which project performance and process are monitored and measured. The primary goal of project control is to identify adverse deviations from plan so that corrective or preventive actions can be taken in a timely manner (PMBOK® Guide 2004). The efficiency and effectiveness of project control rely on the viability of the plan and the accuracy of the measured performance and progress. However, currently available project planning tools and performance measuring methods in real construction projects are neither perfect nor free from errors or uncertainties. For example, project plans such as project schedule and cost baseline are developed based on assumptions about the availability of resources such as bulk material, labor, equipment, and key personnel in a timely manner and at reasonable costs. Furthermore, implementing even a good plan is subject to a degree of risk. On the other hand, performance measuring systems also include a degree of uncertainty because typical construction projects consist of a wide range of non-homogeneous activities with different physical properties and measurement units, which often overlap each other.

The Kalman filter provides a probabilistic framework that incorporates various kinds of prior information available under typical construction management standards and practices. Along with the actual performance data being generated by a project, prior knowledge of the project progress pattern and the measurement system are taken into account to forecast the future progress of a project.

This chapter is organized as follows. In Section 3.2, the concept and properties of the Kalman filter are reviewed and its application to the problem of forecasting project progress and final completion date at completion is addressed. Section 3.3 develops the general concepts into the Kalman filter forecasting method. Depending on the availability of the project baseline plan, different approaches are presented. In addition, a method of initializing the Kalman filter parameters with prior performance information is discussed. A numerical example is presented in Section 3.5.

3.2 The Kalman Filter and Its Application to Project Performance Forecasting

3.2.1 The Kalman Filter

The Kalman filter is a recursive algorithm used to estimate the true but hidden state of a dynamic system using noisy observations. Since the seminal paper by Rudolph E. Kalman about a solution to the discrete-data linear filtering problem (Kalman 1960), the Kalman filter has expanded its application areas from tracking algorithms for radar systems to image processing. In the context of civil engineering, the Kalman filter and the extended Kalman filter have been applied on structural identification problems, structural control, and forecasting (Awwad et al. 1994).

The Kalman filter algorithm provides a recursive learning cycle shown in Figure 3.1. Within the Kalman filter framework, the state of a dynamic system, or the knowledge of any system of interest, is represented by two sets of variables: the state variables and the error covariance variables. The error covariance represents the uncertainty in the estimates of the state variables. The states and covariance are updated through two stochastic linear models: the measurement model and the system model. The measurement model updates prior information using new observations and the system model predicts the future state of the system at the following time step.

It should be noted that the recursive learning cycle of the Kalman filter should be triggered with initial estimates of the states and the associated covariance. Kalman filter theory is too extensive to be covered here. Good introductions to the Kalman filter can be found in books (Brookner 1998; Welch and Bishop 2001; Zarchan and Musoff 2000). In the Kalman filter literature, different notations are currently in use, which may cause unnecessary confusion. This dissertation follows the notation used by Welch and Bishop (2001).

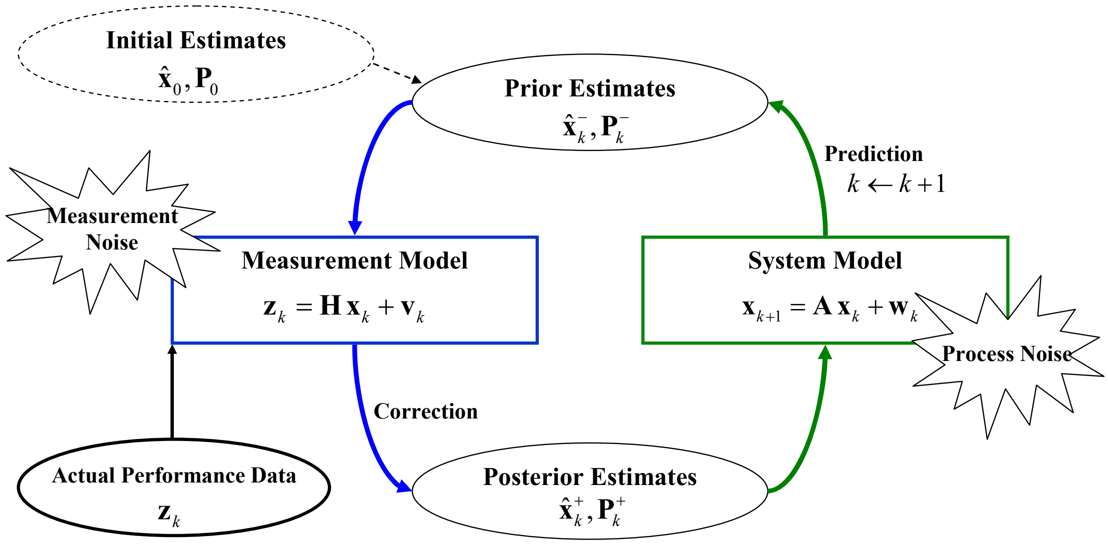


Figure 3.1 Recursive learning cycle of the Kalman filter

3.2.2 Application of the Kalman Filter to Project Performance Forecasting

Based on the general framework of the Kalman filter, the Kalman filter forecasting method has been derived by analogy with the problem of tracking satellites or missiles.

Figure 3.2 represents the block diagram of the Kalman filter forecasting method. Construction projects are executed under the guidance of a set of plans such as the work breakdown structure, the schedule, the budget, and the resource plan, which are collectively referred to as the baseline plan. A common way of representing the progress at the project level is cumulative progress curves, which are also called S-curves. The baseline progress curve, which is defined as the cumulative progress curve that is based on the baseline plans, specifies the amount of work to be done at some particular time, the budget allocated to complete the work, and time derivatives of these values. In the Kalman filter formulation for project performance forecasting, the cumulative progress of a project throughout the execution phase is assumed to evolve according to a dynamic state-space equation which determines the baseline progress curve of the project. The dynamic progress equation of a project represents our knowledge of the status of the project in the future and, inevitably, is subject to some degree of errors and uncertainties.

Once the project proceeds, actual performance, especially, the work done and the costs expended, is continuously monitored by performance measuring systems such as the earned value method and network schedule. The reported performance data are to be compared with the baseline plan and analyzed to obtain some indicators about how well the project is going. However, the true states of on-going projects are not observable due to measurement errors.

Using the prior performance information available, the actual performance data, and the inherent errors and uncertainties in execution and measurements, the Kalman filter is applied to obtain an optimal estimate of the project's true state that minimizes the mean squared error. After the effect of noise or random fluctuation in the observed data is eliminated, the Kalman filter proceeds by repeating the prediction process in the recursive learning cycle until the completion state is reached.

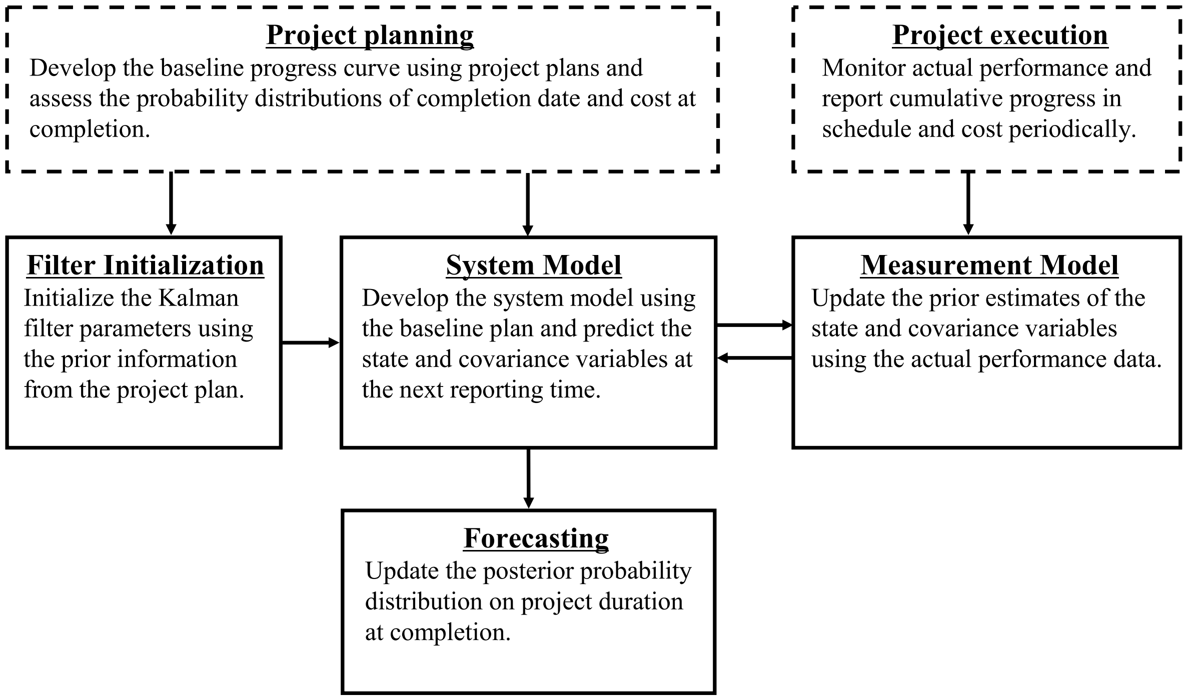


Figure 3.2 Application of the Kalman filter to project performance forecasting

3.3 Kalman Filter Forecasting Method

3.3.1 Formulation for the Case with a Baseline Plan

A project baseline is the approved time phased plan for a project, against which actual project performance is compared and deviations are measured in terms of cost, schedule, and other performance measures (PMBOK® Guide 2004). In the Kalman filter forecasting model, the baseline plan, or simply, the baseline of a project is represented by a cumulative progress curve that shows the amount of work to be done at a specific time throughout the execution of the project. With a baseline plan as a guide for execution, project controlling efforts focus on the deviation between the plan and the actual, instead of dealing with the plan and the actual separately.

Figure 3.3 shows the inputs required to apply the Kalman filter method to projects with baseline plans. Along with the actual performance data, the Kalman filter method requires prior information about the planned duration at completion (PDAC), the budget at completion (BAC), the baseline progress curve, and the probability distribution of the project duration *T*. The availability of PDAC and BAC is justified by the definition of a project, which is "… every project has a definite beginning and definite end. The end is reached when the project's objectives have been achieved" (PMBOK® Guide 2004). Here, the PDAC represents the definite end of a project and, together with the BAC, constitutes the project's objectives.

The need for the probability distribution of the project duration T arises from the probabilistic nature of the system model in the general Kalman filter formulation in Section 3.2. Because the future project status determined by a system model is not certain – if it is, the future can be foreseen accurately and performance measurements or control are not necessary – the uncertainty associated with the system model is taken into account with process noise. In the context of project performance forecasting, the process noise can be interpreted as the variation in the project performance due to the inherent uncertainty in the plan and in the execution of the plan.

For projects with a baseline plan, the Kalman filter forecasting method represents the dynamic progress of a project's performance in terms of the time or cost deviation between the baseline and the actually observed performance data. In this section, the formulation for schedule forecasting is derived.

Major components of the Kalman filter schedule forecasting method for projects with a baseline plan are shown in Table 3.1. The remaining part of this section addresses each component in detail.

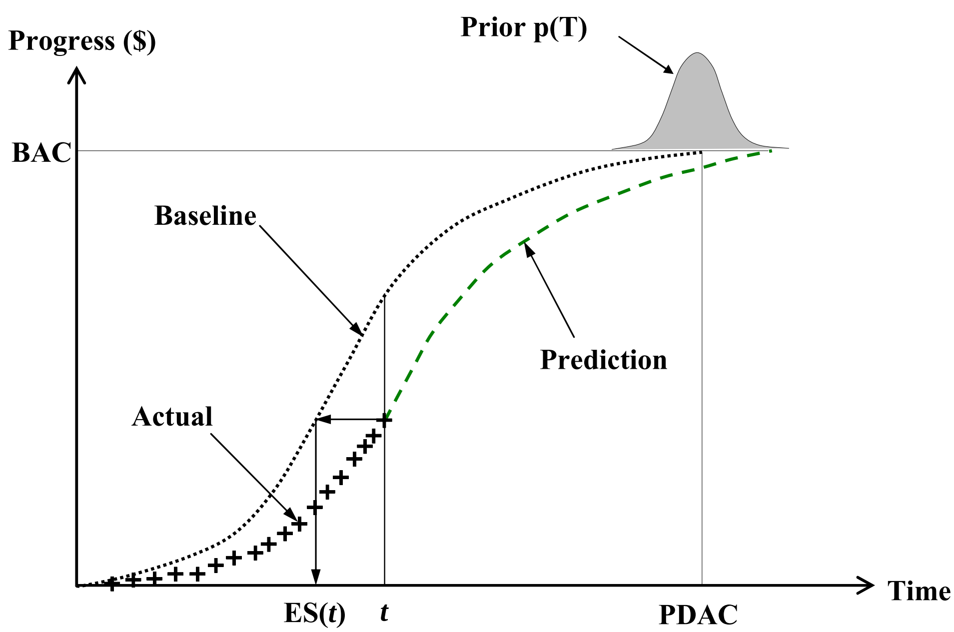


Figure 3.3 Kalman filter forecasting when a baseline is available

Table 3.1 The Kalman filter forecasting model

Components	Equations	Descriptions
State vector	$\mathbf{x}_k = \begin{Bmatrix} TV_k \\ \dfrac{dTV_k}{dt} \end{Bmatrix}$	TV_k is the time variation that is defined as the earned schedule minus the time of forecasting.
Dynamic system model	$\mathbf{x}_k = \mathbf{A}_k \cdot \mathbf{x}_{k-1} + \mathbf{w}_{k-1}$ $\mathbf{A}_k = \begin{bmatrix} 1 & \Delta T \\ 0 & 1 \end{bmatrix}$	$\mathbf{A_k}$ is the transition matrix and $\mathbf{w}_{k-1}$ is a vector of random process noise.
Measurement model	$\mathbf{z}_k = \mathbf{H}\,\mathbf{x}_k + \mathbf{v}_k$ $\mathbf{H} = \{1 \quad 0\}$	**H** is the observation matrix and $\mathbf{v}_k$ is a vector of random measurement noise.
Prediction process	$\hat{\mathbf{x}}_k^- = \mathbf{A}_k \hat{\mathbf{x}}_{k-1}^+$ $\mathbf{P}_k^- = \mathbf{A}_k \mathbf{P}_{k-1}^+ \mathbf{A}_k^{\ T} + \mathbf{Q}_{k-1}$	Before observing a new TV_k at time period *k*, the prior estimates of the state vector and the error covariance matrix are calculated.
Kalman Gain	$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}^T \left(\mathbf{H}\mathbf{P}_k^- \mathbf{H}^T + \mathbf{R}_k \right)^{-1}$	$\mathbf{K}_k$ is the Kalman Gain at time period *k*, which is determined in such a way that minimizes the posterior error covariance matrix.
Updating process	$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \mathbf{K}_k \left(\mathbf{z}_k - \mathbf{H}\hat{\mathbf{x}}_k^- \right)$ $\mathbf{P}_k^+ = \left[\mathbf{I} - \mathbf{K}_k \mathbf{H} \right] \mathbf{P}_k^-$	The posterior estimates of the state vector and the error covariance matrix are calculated using the Kalman Gain.

Definition of States

When a project starts with a baseline progress curve to meet the objectives of the project, actual progress during construction is compared with the baseline to identify deviations from the plan and to take corrective actions, if necessary. For example, the planned value (PV) and the earned value (EV) in the earned value method represent the amount of work planned to be done and the amount of work actually done, respectively, at a specific time. The Kalman filter forecasting method for schedule forecasting focuses on the time variation (TV), which is defined as the deviation between an actual reporting time and the planned time to complete the work actually done at the reporting time (Barraza et al. 2004). The planned time to achieve the work actually done is named as the earned schedule (ES) (Lipke 2003). Using the EVM terminology, the time variation is defined as follows.

$$\mathrm{TV}(t) = \mathrm{ES}(t) - t \tag{3.1}$$

By definition, the ES at time t is determined by

$$\mathrm{PV}\big(\mathrm{ES}(t)\big) = \mathrm{EV}(t) \tag{3.2}$$

Because the planned values of real projects are mostly calculated only at discrete reporting points, the earned schedule at a specific time can be approximated by an interpolation between two consecutive PVs that satisfy $\mathrm{EV}(t) \geq \mathrm{PV}(k)$ and $\mathrm{EV}(t) < \mathrm{PV}(k+1)$ (Vandevoorde and Vanhoucke 2006). Once the k is determined, the linear interpolation equation to calculate the earned schedule is

$$\mathrm{ES}(t) = k + \frac{\mathrm{EV}(t) - \mathrm{PV}(k)}{\mathrm{PV}(k+1) - \mathrm{PV}(k)} \tag{3.3}$$

In the KFFM, cumulative progress of a project is modeled as a system with two states that evolve over time: the time variation (TV) and its rate of change over a reporting period. Then, the state vector of the Kalman filter forecasting method is defined as

$$\mathbf{x}_k = \begin{Bmatrix} x_{k,1} = TV_k \\ x_{k,2} = \dfrac{dTV_k}{dt} \end{Bmatrix}, \qquad k = 0,1,... \tag{3.4}$$

Here, $x_{k,1}$ is the time variation at any time interval k. The second element in the state vector is the incremental TV, which is calculated as $x_{k,2} = x_{k,1} - x_{(k-1),1}$.

Transition Process

With the state vector, the transition process of the Kalman filter is defined as

$$\mathbf{x}_k = \mathbf{A}_k \cdot \mathbf{x}_{k-1} + \mathbf{w}_{k-1} \tag{3.5}$$

where $\mathbf{A}_k$ is the transition matrix and $\mathbf{w}_{k\text{-}1}$ is a vector of random process noise. The transition matrix $\mathbf{A}_k$ is obtained by assuming a constant rate $x_{k,2}$ between two consecutive observations and a linear approximation of $x_{k,1}$.

$$\begin{aligned} x_{k,1} &= x_{(k-1),1} + \Delta T\, x_{(k-1),2} \\ x_{k,2} &= x_{(k-1),2} \end{aligned} \tag{3.6}$$

where ΔT is the reporting period. Then,

$$\mathbf{A}_k = \begin{bmatrix} 1 & \Delta T \\ 0 & 1 \end{bmatrix} \tag{3.7}$$

The random process noise term represents the random error or uncertainty in the system and is assumed to have zero mean, white, Gaussian components with covariance matrix $\mathbf{Q}_k$ defined by

$$E\left[\mathbf{w}_i, \mathbf{w}_k^T\right] = \begin{cases} \mathbf{Q}_k & \text{for } i = k, \\ \mathbf{0} & \text{for } i \neq k, \end{cases} \tag{3.8}$$

The whiteness assumption of the process noise implies that the process noise is not correlated in time.

Measurement Process

The state components in (3.4) represent the status of a project. In real projects, however, the status of an ongoing project can be estimated only through a performance measurement process, which is defined as follows.

$$\mathbf{z}_k = \mathbf{H}\,\mathbf{x}_k + \mathbf{v}_k \tag{3.9}$$

where $\mathbf{z}_k$ and $\mathbf{H}$ are the measurement vector and the observation matrix, respectively. Since the time variation is the only measure of progress in the Kalman filter formulation, the dimension of the measurement vector $\mathbf{z}_k$ is one and the observation matrix $\mathbf{H}$ becomes

$$\mathbf{H} = \{1 \quad 0\} \tag{3.10}$$

The uncertainty in the measurement process is taken into account by adding a random noise term into the measurement process. The measurement noise vector $\mathbf{v}_k$ is assumed to have zero-mean, white, Gaussian components with covariance matrix $\mathbf{R}_k$ defined by

$$E\left[\mathbf{v}_i, \mathbf{v}_k^T\right] = \begin{cases} \mathbf{R}_k & \text{for } i = k, \\ \mathbf{0} & \text{for } i \neq k, \end{cases} \tag{3.11}$$

The measurement noise $\mathbf{v}_k$ is assumed to be uncorrelated with the process noise $\mathbf{w}_k$. Since the measurement vector has dimension one, $\mathbf{R}_k$ is a scalar. The size of the measurement noise depends on the accuracy of measuring tools or methods, and must be estimated from experience in other similar projects.

The State Estimate and the Error Covariance

In the Kalman filter approach, knowledge about the state of a system is represented by two quantities: the state estimate $\hat{\mathbf{x}}_k$ and the error covariance $\mathbf{P}_k$. At a given time period k, both quantities are further classified depending on whether the estimation is carried out before or after the new observation $\mathbf{z}_k$. Estimates *before* new observations are called the *prior* state estimates and the *prior* error covariance, and are denoted as $\hat{\mathbf{x}}_k^-$ and $\mathbf{P}_k^-$, respectively. Likewise, estimates *after* new observations are called the *posterior* state estimates and the *posterior* error covariance, and are denoted as $\hat{\mathbf{x}}_k^+$ and $\mathbf{P}_k^+$, respectively.

The error covariance is defined from the state error estimates.

$$\begin{aligned} \mathbf{P}_k^- &= E\left[\mathbf{e}_k^- \mathbf{e}_k^{-T}\right] \\ \mathbf{P}_k^+ &= E\left[\mathbf{e}_k^+ \mathbf{e}_k^{+T}\right] \end{aligned} \tag{3.12}$$

Here, the error vector $\mathbf{e}_k$ is defined as the difference of the true – but hidden – value of the system state $\mathbf{x}_k$ and an estimate $\hat{\mathbf{x}}_k$.

$$\begin{aligned} \mathbf{e}_k^- &= \mathbf{x}_k - \hat{\mathbf{x}}_k^- \\ \mathbf{e}_k^+ &= \mathbf{x}_k - \hat{\mathbf{x}}_k^+ \end{aligned} \tag{3.13}$$

Prediction Process: Prior State Estimate and Prior Error Covariance

Before observing new data, the prior state estimate at time period *k* is calculated based on the previous state estimate and the transition process.

$$\hat{\mathbf{x}}_k^- = \mathbf{A}_k \hat{\mathbf{x}}_{k-1}^+ \tag{3.14}$$

The prior error covariance is obtained from Equation (3.13) and (3.14).

$$\begin{aligned} \mathbf{e}_k^- \quad &= \mathbf{x}_k - \hat{\mathbf{x}}_k^- \\ &= \left\{ \mathbf{A}_k \mathbf{x}_{k-1} + \mathbf{w}_{k-1} \right\} - \mathbf{A}_k \hat{\mathbf{x}}_{k-1}^+ \\ &= \mathbf{A}_k \left\{ \mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1}^+ \right\} + \mathbf{w}_{k-1} \\ &= \mathbf{A}_k \mathbf{e}_{k-1}^+ + \mathbf{w}_{k-1} \end{aligned} \tag{3.15}$$

From the definition of the error covariance in Equation (3.12),

$$\begin{aligned} \mathbf{P}_k^- \quad &= E\left[\mathbf{e}_k^- \mathbf{e}_k^{-T} \right] \\ &= E\left[\left\{ \mathbf{A}_k \mathbf{e}_{k-1}^+ + \mathbf{w}_{k-1} \right\} \left\{ \mathbf{A}_k \mathbf{e}_{k-1}^+ + \mathbf{w}_{k-1} \right\}^T \right] \\ &= E\left[\mathbf{A}_k \mathbf{e}_{k-1}^+ \mathbf{e}_{k-1}^{+\ T} \mathbf{A}_k^{\ T} \right] + E\left[\mathbf{A}_k \mathbf{e}_{k-1}^+ \mathbf{w}_{k-1}^{\ T} \right] + \\ &\quad E\left[\mathbf{w}_{k-1} \mathbf{e}_{k-1}^{+\ T} \mathbf{A}_k^{\ T} \right] + E\left[\mathbf{w}_{k-1} \mathbf{w}_{k-1}^{\ T} \right] \end{aligned} \tag{3.16}$$

From the independence property between the error vector and the system noise vector, Equation (3.16) becomes

$$\begin{aligned} \mathbf{P}_k^- \quad &= E\left[\mathbf{A}_k \mathbf{e}_{k-1}^+ \mathbf{e}_{k-1}^{+\ T} \mathbf{A}_k^{\ T} \right] + E\left[\mathbf{w}_{k-1} \mathbf{w}_{k-1}^{\ T} \right] \\ &= \mathbf{A}_k \mathbf{P}_{k-1}^+ \mathbf{A}_k^{\ T} + \mathbf{Q}_{k-1} \end{aligned} \tag{3.17}$$

The Kalman Gain

In the Kalman filter, the posterior state estimate $\hat{\mathbf{x}}_k^+$ is determined as a linear combination of the prior state estimate $\hat{\mathbf{x}}_k^-$ and the weighted difference between the new actual measurement $\mathbf{z}_k$ and the predicted measurement $\mathbf{H}\,\hat{\mathbf{x}}_k^-$.

$$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \mathbf{K}_k\left(\mathbf{z}_k - \mathbf{H}\hat{\mathbf{x}}_k^-\right) \tag{3.18}$$

The critical issue is to determine the weight for the difference $\mathbf{z}_k - \mathbf{H}\,\hat{\mathbf{x}}_k^-$ which is called the measurement innovation. The weight matrix **K** is called the Kalman Gain and is determined in such a way that minimizes the posterior error covariance in Equation (3.12). One solution to the minimization problem is given by

$$\mathbf{K}_k = \mathbf{P}_k^-\mathbf{H}^T\left(\mathbf{H}\mathbf{P}_k^-\mathbf{H}^T + \mathbf{R}_k\right)^{-1} \tag{3.19}$$

It should be noted that, when the transition matrix **A** and the measurement matrix **H** are constant over time, the Kalman Gain $\mathbf{K}_k$ depends only on the initial error covariance $\mathbf{P}_0$ and the noise covariance matrixes $\mathbf{Q}_k$ and $\mathbf{R}_k$ regardless of the actual observation $\mathbf{z}_k$. For example, when the size of the measurement covariance matrix $\mathbf{R}_k$ increases the Kalman Gain decreases. This inverse relationship is intuitively correct because the effect of a new observation decreases as the accuracy of measurement decreases or as the uncertainty in measurement increases. Once the Kalman Gain is known, the posterior state estimate is obtained from Equation (3.18) and the posterior error covariance can be calculated by

$$\mathbf{P}_k^+ = \left[\mathbf{I} - \mathbf{K}_k\mathbf{H}\right]\mathbf{P}_k^- \tag{3.20}$$

3.3.2 Formulation for the Case without a Baseline Plan

In the Kalman filter formulation for projects with a baseline plan, a cumulative progress curve based on the plan serves as a target trajectory against which actual performance is compared. The resulting dynamic system model for project progress is defined in terms of the time deviation between the baseline and the actual performance reported after each reporting period. Since a prediction at a specific time point is made in terms of the relative performance against the baseline, the overall progress pattern of a baseline is automatically taken into account.

When there is no baseline to be used as a progress template, forecasts can be made based only on the actual performance data observed from the project itself. However, predictions that appear to be equally plausible in terms of their fit to the actual data can lead to a wide range of different results according to the assumptions about the relationship between the past and the future of a project. For example, Figure 3.4 conceptually depicts some typical predictions that can be made based on the same actual progress data. Even though it would be possible to select the "best" model using some mathematical methods, for example, the least squares method, huge differences in results according to the selection of one extrapolating function against others can be hardly justified. Obviously, better fit to the past data does not guarantee better forecasts. At the very least, it is possible to find some mathematical curve that fits pretty well both the actual data and the planned duration of a project (See the sigmoidal prediction in Figure 3.4). In other words, a prediction can be made and used to justify whatever status the project is in rather than to predict the plausible future and to use that information for better management of the remaining work.

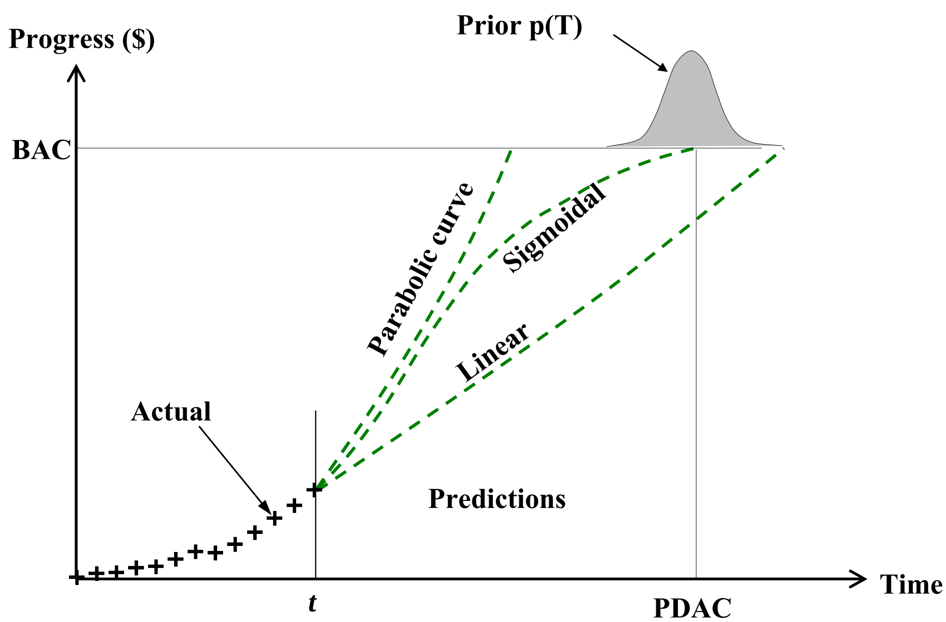


Figure 3.4 Predictions when a baseline is not available

The Kalman filter approach needs to address the same issue of selecting a plausible relationship between the past and the future. More specifically, the system model should be established in such a way that the next state of a project can be predicted based on the current state. In the absence of a baseline progress curve, alternative information should be used to establish a plausible progress template. For example, subjective judgments or experience from similar projects can be used to select an appropriate progress curve that reasonably describes the progress pattern of a given project.

3.3.3 Kalman Filter Forecasting

The application of the Kalman filter to forecasting can be regarded as an extrapolation of the system states. To get the state estimates at time $J > k$, the Kalman Gain is set to zero during the interim periods from $k + 1$ to $J - 1$. Then the Kalman updating routines for the state estimate become

$$\hat{\mathbf{x}}_J^+ = \hat{\mathbf{x}}_J^- = \mathbf{A}^{J-k}\hat{\mathbf{x}}_k^+ \tag{3.21}$$

In the case of the error covariance, the updating algorithm for forecasting becomes

$$\begin{aligned} \mathbf{P}_k^- &= \mathbf{A}_{k-1}\mathbf{P}_{k-1}^-\mathbf{A}_{k-1}^T + \mathbf{Q}_{k-1} \\ \mathbf{P}_k^+ &= \mathbf{P}_k^- \end{aligned} \tag{3.22}$$

3.3.4 Initialization of Kalman Filter

Because of the recursive nature of Kalman filter, initial values of state variables and error covariance must be determined before the first updating based on an actual performance report. In addition, the process noise covariance matrix and the measurement error matrix need to be estimated in advance. These processes are collectively referred to as the initialization of Kalman filter. The primary challenge in the initialization of Kalman filter is that it needs to be done in the absence of any objective observations from the project to be analyzed. The importance of initialization has been emphasized by many authors (Ansley and Kohn 1985; Casals et al. 2000; de Jong 1989). This section addresses the initialization process of the current Kalman filter forecasting method.

First of all, the initial state estimate and its error covariance are set to be zero. Unlike other time series forecasting problems, the project performance forecasting has or should have a clear starting point in terms of the physical work to be done, start time, and initial cost, which are likely to be set to be zero. In addition, it would be reasonable to assume that there is no uncertainty in the initial state of the project progress in terms of time and money.

The covariance matrix of process noise Q_k is modeled as a unit matrix multiplied by a scalar constant.

$$\mathbf{Q}_k = \begin{bmatrix} q & 0 \\ 0 & q \end{bmatrix} \tag{3.23}$$

In a covariance matrix of process noise, the diagonal terms represent the variances of each state variable and the off-diagonal terms represent the covariances between the state variables. In the case when only random errors of each state variable are considered, the off-diagonal terms are zero. The variances of state variables are assumed to be constant, not because of any rationale or empirical evidence supporting the assumption but because of the lack of information supporting an alternative assumption. As a result, the covariance matrix of process noise is modeled with a single variable, q, which represents the value of all diagonal terms.

In the Kalman filter forecasting method, the value of this variable is determined from the prior probability distribution – either from network-based simulation or subjective judgments – for the total project duration. Given the mean, μ_T, and variance, σ_T^2, of a project duration T, q is determined by the Kalman filter forecasting equations in (3.21) and (3.22), in such a way that the uncertainty in the system model is consistent with the prior estimate of the uncertainty in project performance. That is, at $k = \mu_T$, the dynamic system model reaches the completion of the project with the same uncertainty estimated before the beginning of a project.

$$\mathbf{P}_{\mu_T}^{+} = \mathbf{P}_{\mu_T}^{-} = \begin{bmatrix} \sigma_T^2 & 0 \\ 0 & \sigma_T^2 \end{bmatrix} \tag{3.24}$$

The measurement error covariance matrix R_k represents the accuracy of actual performance measurements and plays an important role in Kalman filter forecasting. For example, as the measurement error approaches zero, the Kalman Gain in Equation (3.19) increases, resulting in the posterior estimates of state variables being influenced more by actual measurements than by the prior estimates of state variables by the system model. Therefore, by setting an appropriate value for the measurement error, the sensitivity of the forecasts to the actual performance data is adjusted.

In the Kalman filter model for project performance forecasting, the progress of a project is measured with a single value, the cumulative progress of the project, and the measurement error term can be determined based on prior information from experience and judgment.

3.3.5 Calculation of the EDAC

The objective of progress forecasting is to estimate and update the estimated duration at completion (EDAC) at any time t_k. Since the state variables of the Kalman filter forecasting method represents the estimate of time variation and its rate of change over time, results from the method need to be converted into practical measures that can be used by project managers. This section addresses the problem of calculating the EDAC at a specific time t_k and its associated prediction bands.

The process of determining EDAC is illustrated in Figure 3.5. At any time t_k, the estimate of Time Variation $\widehat{TV}_k$ is obtained from the Kalman filter analysis as $\hat{x}_{k,1}$. Then, the expected earned value at that time of analysis $\widehat{EV}_k$ is determined by the baseline progress curve.

$$\begin{aligned} \widehat{EV}_k &= PV\left(\widehat{ES}_k\right) \\ \widehat{ES}_k &= t_k + \widehat{TV}_k = t_k + \hat{x}_k \end{aligned} \tag{3.25}$$

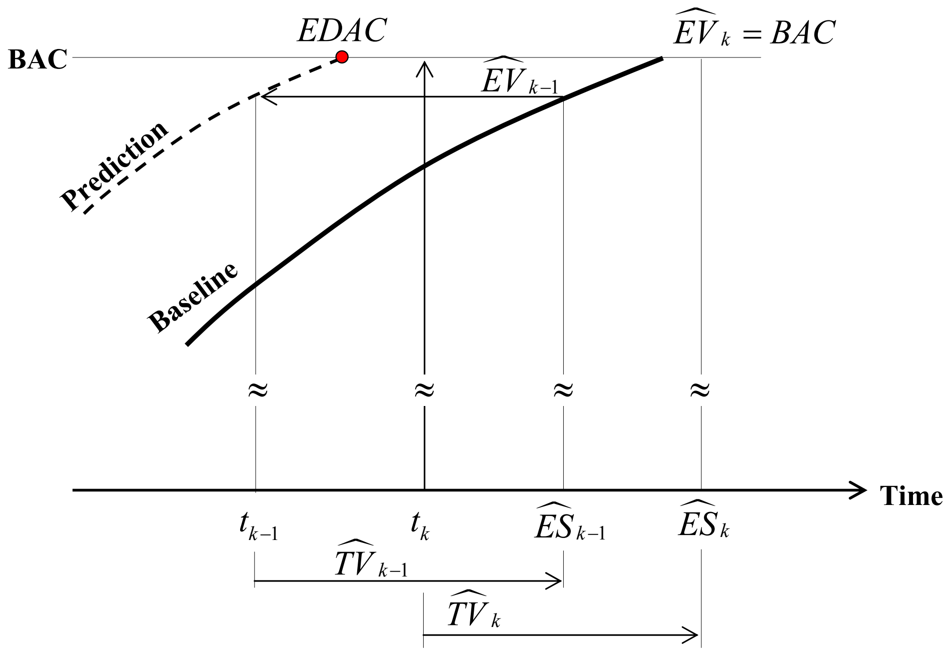
Then, the EDAC is determined as the time that satisfies

$$\widehat{EV}_k \geq BAC \tag{3.26}$$

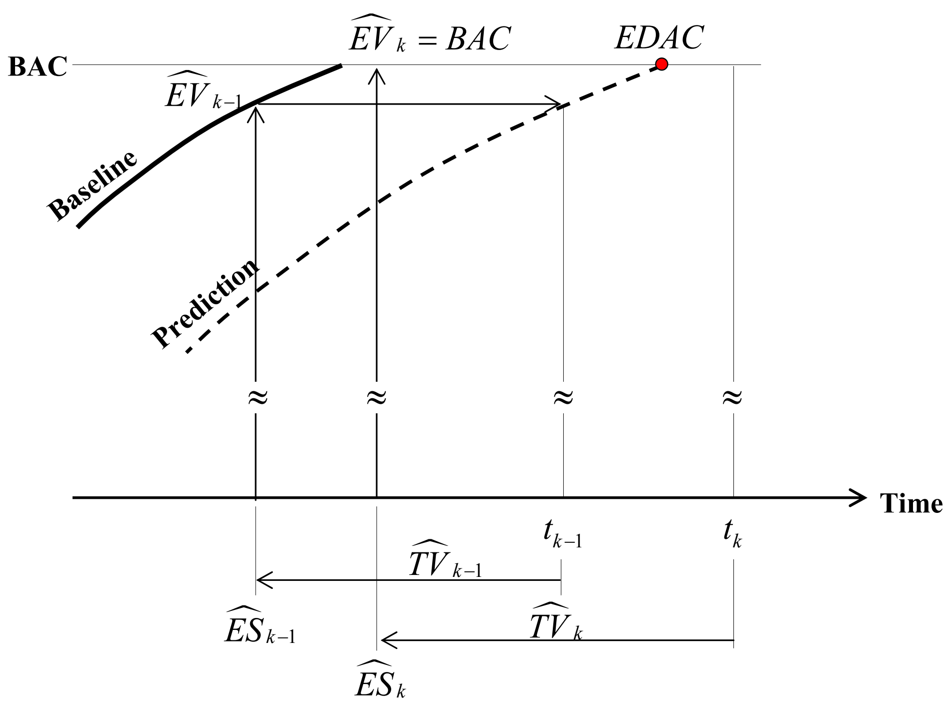
Once such t_k is found, the correct value of the EDAC must lie between t_{k-1} and t_k. In the applications in this dissertation, the exact value of EDAC is approximated as t_k.

$$\begin{aligned} EDAC &= t_k - \alpha\left(t_k - t_{k-1}\right), \qquad 0 \leq \alpha < 1 \\ &\cong t_k \end{aligned} \tag{3.27}$$

The prediction bounds of the EDAC can be obtained directly from the Kalman filter results in terms of the error covariance matrix $\mathbf{P}_k$.



(a) EDAC for a project ahead of schedule. $\widehat{TV}_k > 0$



(b) EDAC for a project behind schedule. $\widehat{TV}_k < 0$

Figure 3.5 Calculating EDAC from the Kalman filter forecasting output

3.4 Example 3.1

The Kalman filter forecasting method derived in the previous sections is used to predict and update future progress of an artificial project using randomly generated actual progress data. The purpose of this example is to demonstrate the probabilistic and adaptive nature of the KFFM.

First, a random activity network of a project with 100 activities and 144 precedence relations is generated using a random network generation method augmented with a redundancy elimination technique, which is discussed in Chapter V. Then, Monte Carlo simulation is carried out to determine the planned progress curve and the prior probability distribution of the completion date. The completion dates and corresponding progress curves for a set of random activity durations are calculated with the early start forward calculation in CPM. It is assumed that all activities in the project are homogeneous with the same stochastic properties in duration and cost. This homogeneous assumption is based on a scheduling practice in which activities in a project network are defined in such a way that they are small enough to be managed effectively, and not too small to cause additional burdens in management. The input data used for generating the progress curves are summarized in Table 3.2.

Table 3.2 Input data for random progress curve generation

Input parameters		Value
Number of activities		100
Number of effective precedence relations		144
Activity duration for planned progress	Mean Standard deviation	5 (weeks) 2 (weeks)
Activity duration for actual progress	Mean Standard deviation	5 (weeks) 2 (weeks)

The planned and the actual progress curves used in the following forecasts are shown in Figure 3.6. From a Monte Carlo simulation based on a random activity network and the assumed probability distributions of activity durations and costs, the prior distribution of the completion date is estimated to have a mean of 67 weeks and a standard deviation of 6.0 weeks. The planned progress curve in Figure 3.6 is determined by averaging stochastic progress curves from 5000 iterations over the progress dimension.

The cross marks represent a simulated progress curve with completion date at 79 weeks. The simulated curve is generated from the same network schedule for the planned progress curve. Therefore, the curve represents an "actual" progress of the project, which is determined at random due to the uncertainty in the assumed activity durations and costs. It should be noted that, although the complete actual progress is determined in advance, the actual performance data are assumed to be available only after each reporting period up to the time of forecasting.

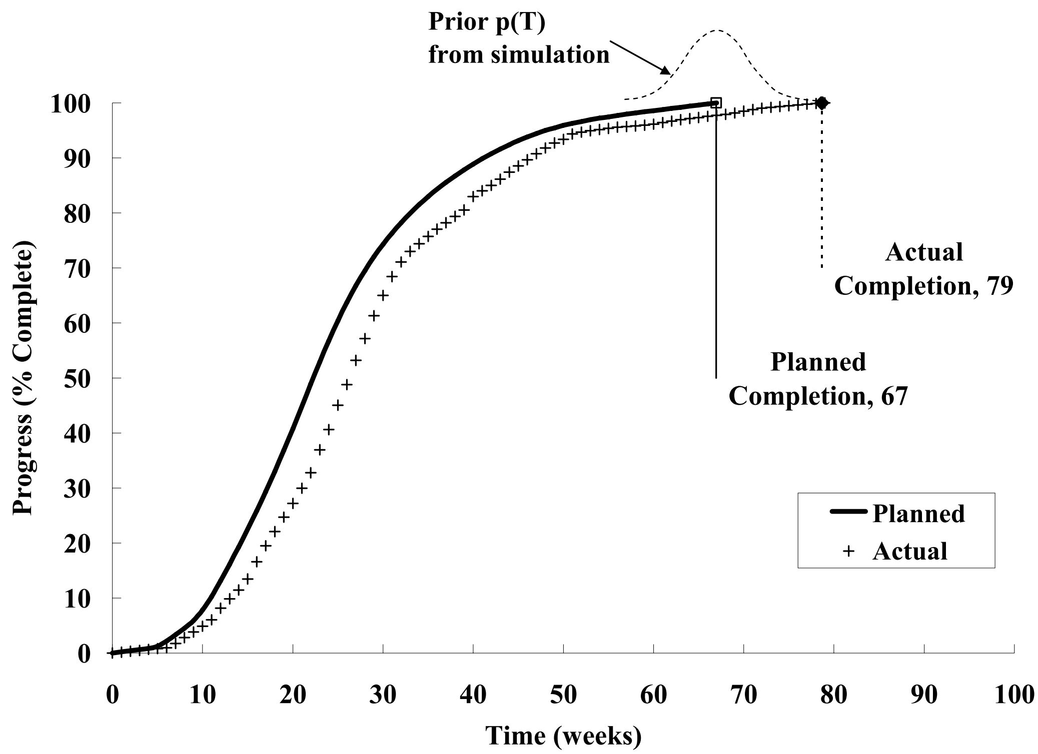
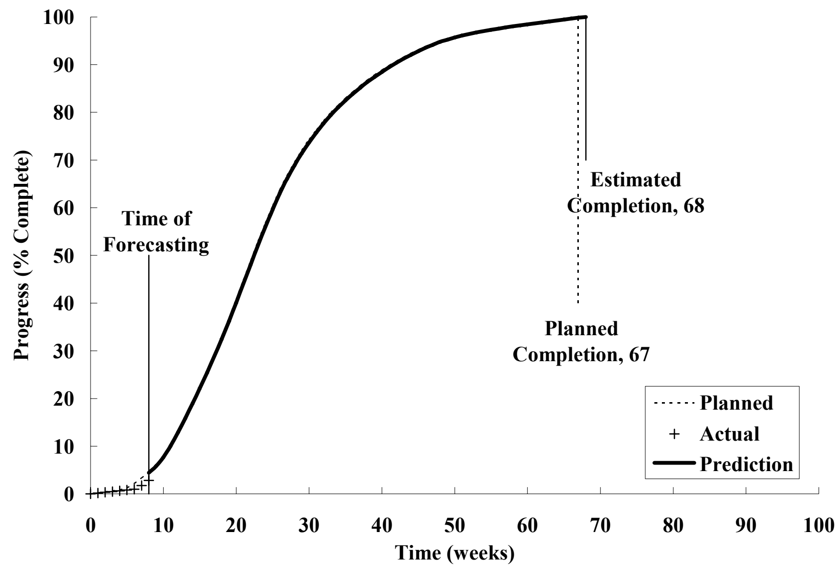


Figure 3.6 The planned progress and the simulated "actual" progress for Example 3.1 (Prior p(*T*) is drawn not to scale)

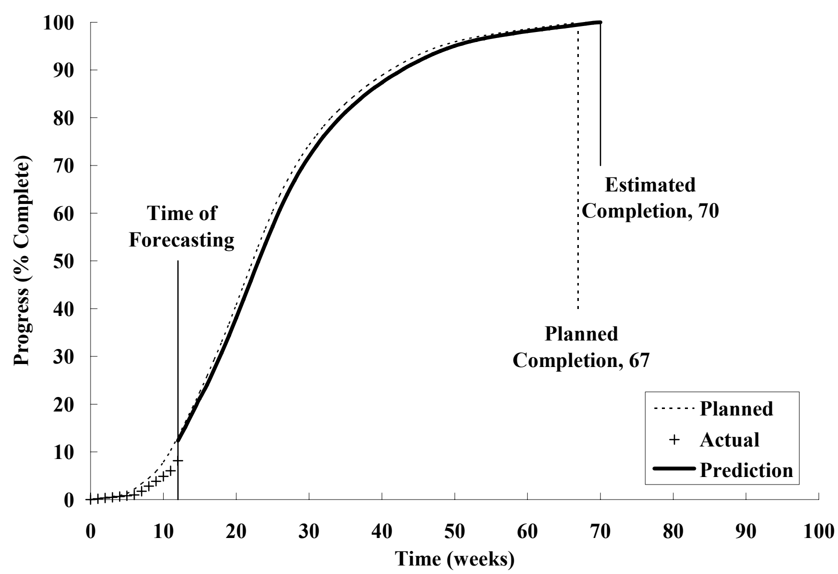
The first prediction is made after 8 weeks and the future progress is shown as a thick solid line in Figure 3.7(a). The graph shows that the forecast from the KFFM is very close to the plan in spite of some deviations in the actual progress from the plan. This is because the prediction is influenced more by the plan than it is by the small number of data points reported. This is attributed to two characteristics of the KFFM. First, the KFFM uses both the prior performance information – the planned progress – and the actual progress data. Second, the KFFM takes into account measurement errors in terms of the measurement noise term in Equation (3.11). The forecasts in Figure 3.7

are made with the variance of measurement error $R = 4$. It is also important to note that the prediction made at week 8 retains the sigmoid shape of the planned progress curve, even though only a few data points have been observed. That is, the KFFM predicts that the future progress will be S-shaped because the planned progress curve is S-shaped, not because the past data contain this information.

The other three predictions in Figure 3.7 (b), (c), and (d) are made at different time points and show the adaptive nature of the Kalman filter forecasting. The forecast would be repeated after every reporting period to incorporate the new data, but here the intermediate updates have been omitted. Comparing the forecasts at different points of project execution, it is easy to see that the forecast has moved away from the plan and closer to the actual data, due to the accumulated discrepancies between the actual reports and the planned value. When actual data are relatively few, for example at week 12, the Kalman filter prediction for the longer range forecasts is still influenced by the plan, and stays closer to it. However, this tendency diminishes as more actual data accrue. After 28 weeks, the actual data dominate the prediction and the predicted progress has moved away from the plan curve and very close to the unknown actual completion date.

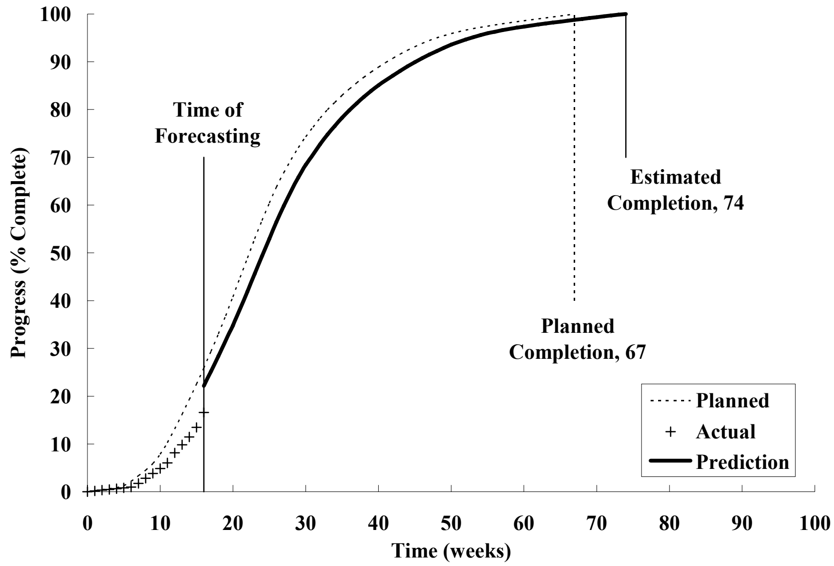


(a) Prediction at 8th week

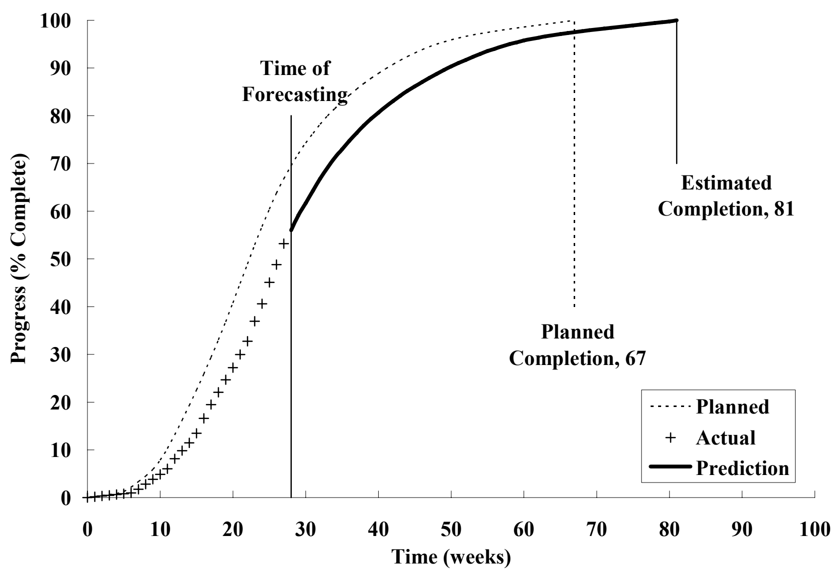


(b) Prediction at 12th week

Figure 3.7 Adaptive nature of the prediction by the KFFM



(c) Prediction at 16th week



(d) Prediction at 28th week

Figure 3.7 (Continued)

Future progress curves, as in Figure 3.7, make predictions by the KFFM more useful for project managers because they provide a visual representation of the plausible path up to the estimated completion date. However, the primary merit of the KFFM is its probabilistic nature. That is, the KFFM provides prediction bounds on predicted completion dates, which indicate the possible outcomes at a given confidence level. The history of EDAC over time is shown in Figure 3.8. The prediction bounds in the graphs are determined at 5% confidence level on both sides.

The prediction starts with the prior probability distribution of the project duration *T*, which is obtained from the simulation based on the project schedule. As more actual performance reports become available, the EDAC moves from the planned completion date toward the actual completion date. It should be noted that the width of prediction interval tends to narrow as more data become available. More importantly, after about 8 weeks, the actual completion date falls inside the prediction bounds and stays inside during most of the remaining period.

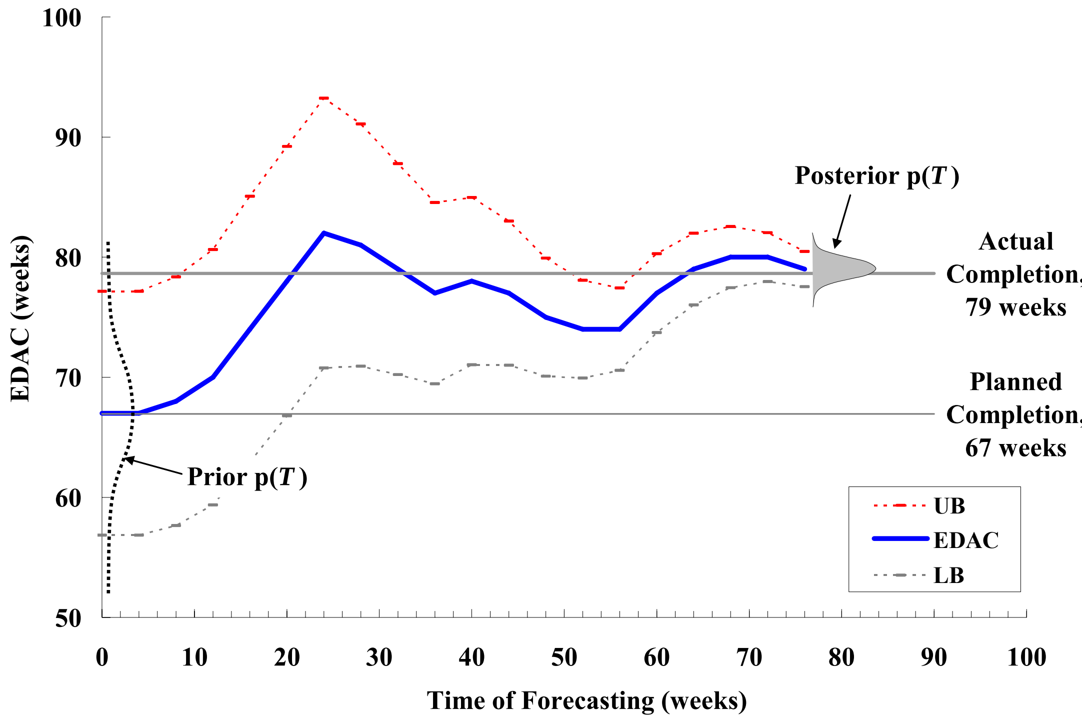
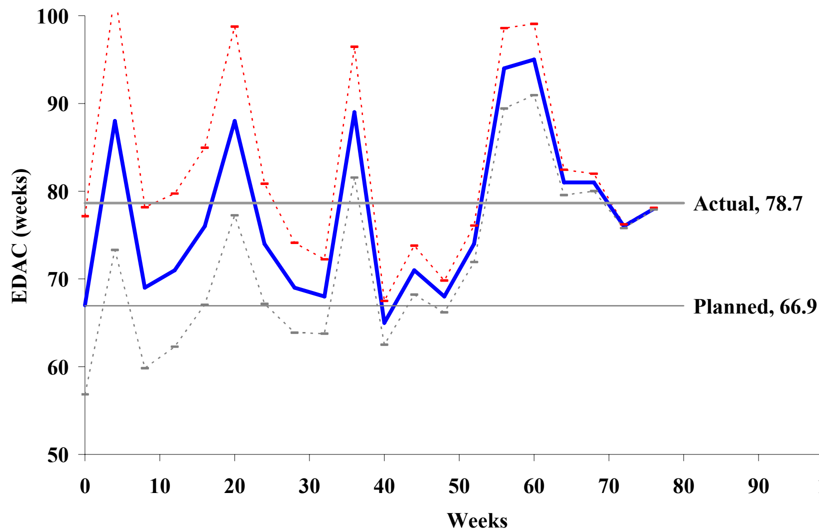


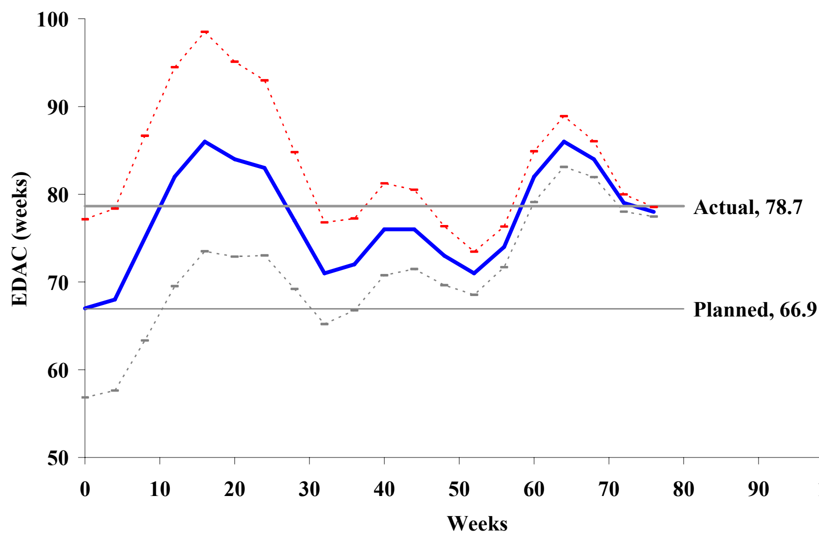
Figure 3.8 History of the EDAC with prediction bounds (The prior and posterior p(T) are drawn not to scale)

Another merit of the KFFM is its flexibility in adjusting the sensitivity of predictions to actual performance data according to the user's belief on the accuracy of measurements. The same project has been predicted with different values for the variance of measurement error R. The results are shown in Figure 3.9. When $R = 0$, which corresponds to a perfect measurement system that always gives the accurate state of the project without error, the predictions show large fluctuations over time. If $R = 0$, the Kalman Gain in Equation (3.18) becomes 1, which indicates that posterior state estimates are completely determined by the new observations. Then the predictions are made based only on the time variation and its rate of change, which are calculated based on the latest observation. As a result, predictions made with the Kalman filter forecasting equation (3.21) are more influenced by the specific states measured at the time of forecasting.

However, perfect measurement systems can hardly be realized in real world project management. The graphs in Figure 3.9 (b) show that even with a small measurement error, $R = 0.25$, the large fluctuation observed with $R = 0$ significantly disappears. The other two graphs with $R = 1$ and $R = 9$ reveal that increasing the measurement errors results in slower conversion of predictions to the actual completion date. With $R = 9$, one can get a smoother EDAC curve and, more importantly, the actual completion date lies inside the prediction bounds after 12 weeks throughout the completion point.

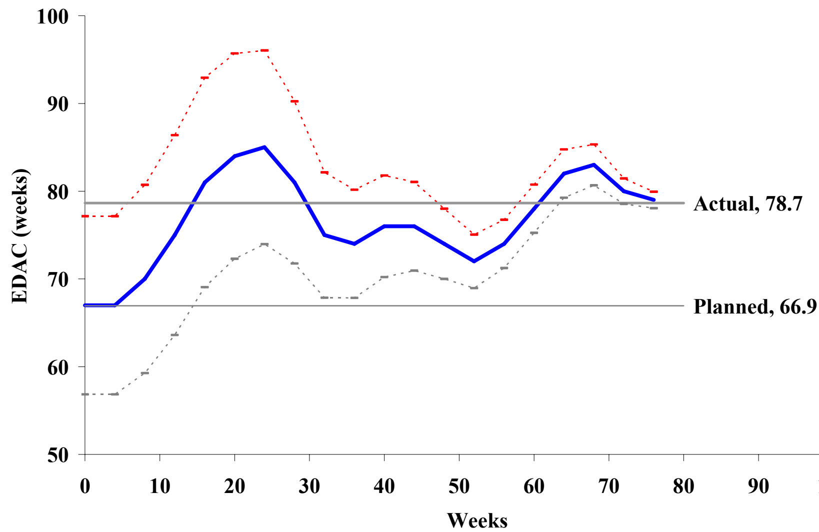


(a) $R = 0$

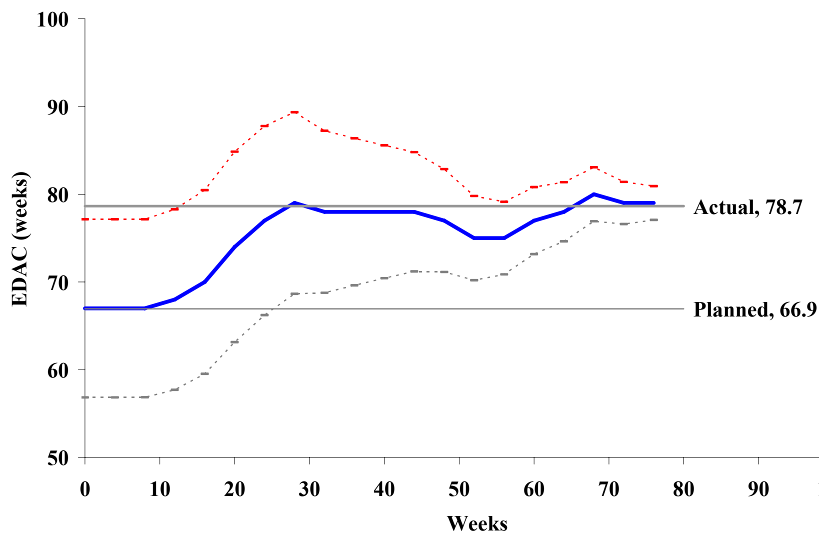


(b) $R = 0.25$

Figure 3.9 Influence of measurement errors on the prediction by the KFFM



(c) $R = 1$



(d) $R = 9$

Figure 3.9 Influence of measurement errors on the prediction by the KFFM (Continued)

3.5 Chapter Summary

The Kalman filter forecasting method is developed based on the general project control framework and the Kalman filter. The KFFM can be characterized with three attributes: (1) The KFFM is a probabilistic schedule forecasting method that provides prediction bounds on the predicted project duration; (2) The KFFM uses prior performance information in conjunction with the actual performance data; and (3) In the KFFM, measurement errors inherent in the real-world projects are explicitly taken into account, which allows the user to adjust the sensitivity of forecasts to the actual performance data according to their confidence on the accuracy of measurements.

In Section 3.2, the Kalman filter and its application to project performance forecasting problem were addressed. Based on the general frameworks of the Kalman filter and project controlling, the Kalman filter forecasting method was formulated in Section 3.3. In the KFFM, relevant prior performance information and uncertainties in typical construction projects are seamlessly integrated into consistent predictions. For example, a probabilistic estimate of project duration based on preconstruction plans is used to initialize the process noise matrix and a planned progress curve is used to supplement small samples of actual data, especially early in a project. With an example in Section 3.4, basic properties of the KFFM were demonstrated. Based on the formulation derived in this chapter, a thorough evaluation of the KFFM and a comprehensive comparison against other forecasting methods will be carried out in the following chapters.

CHAPTER IV

BAYESIAN ADAPTIVE FORECASTING METHOD

4.1 Introduction

A probabilistic method for project performance forecasting has been developed based on Bayesian inference and S-curve functions. S-curves, which are also called progress curves, represent cumulative progress of a project over the execution period. The Bayesian adaptive forecasting (BAF) method is a regression model that fits S-curves to cumulative progress curves of a project and updates the parameter estimates of the S-curves using a Bayesian inference approach. In the BAF method, prior performance information which is available before the inception of a project is effectively used to make reliable forecasts. The prior performance information may have various forms such as detailed project plans, historical data, and subjective judgment of project engineers. Bayesian inference provides a systematic framework for integrating the prior performance information with actual progress data reported during the execution. Effective use of the prior performance information in project performance forecasting overcomes a typical limitation of forecasting during the early phases of project execution when project managers suffer from the lack of enough actual performance data.

Every project is unique and can be characterized by a unique S-curve (Blyth and Kaka 2006; Kenley and Wilson 1986). This observation serves as the premise that every project proceeds following a characteristic cumulative progress curve. The underlying strategy of the BAF method is to identify the characteristic progress curve of a project using the prior performance information and to use that curve, which is named as the progress curve template, to forecast future progress of the project. The shape or pattern of individual progress curves for various projects can be characterized by some indicators such as the type, size, and location of a project. However, even for the same project, different progress curves can be constructed according to the strategies and competence of the organization. These factors collectively influence the shape of progress curves. A reasonable way, probably the most reliable way, of obtaining the progress curve template is to construct it from the detailed plans of the project itself. For example, the planned value (or the Budgeted Cost of Work Scheduled) distributed over time, which is used as the baseline of project performance analysis in the earned value method, can be directly used as a progress curve template for forecasting.

Once the progress curve template for a project is determined, it is approximated by fitting some S-curve models. If a model is found that fits the progress curve template reasonably well, it is used to forecast future performance of the project by updating the parameters of the model with actual performance data. This approach is valid under the assumption that an actual progress curve will also proceed following the same S-curve model that approximates the progress curve template. In real projects, if the actual progress deviates significantly from the plan, then something is amiss with the project.

This chapter is organized as follows. Section 4.2 describes S-curve models in construction management and a new S-curve model is presented. In Section 4.3, the general formulation of Bayesian adaptive forecasting is derived based on Bayesian inference. Practical solutions to computation of multi-dimensional integration are presented using a Monte Carlo integration technique. Based on the general framework of BAF, two different methods are developed in this research. Multi-model Bayesian adaptive forecasting (MBAF) method in Section 4.4 uses a group of fixed-shape S-curve models. Each model in the group proceeds separately using the same performance information. At the final stage, however, different predictions from individual models are combined with the Bayesian model averaging technique. Section 4.5 addresses BetaS-curve Bayesian adaptive forecasting (BBAF) method which uses the BetaS-curve derived in Section 4.2 as a single S-curve model. The BetaS-curve can approximate a wide range of progress curves. This approach is relatively simpler than MBAF and can be used to systematically quantify prior information from various sources of preconstruction information. In Section 4.6, a simulation-based test has been carried out to quantify the predictive nature of a progress curve technique.

4.2 S-curve Models

Experience has shown that typical cumulative progress curves of projects show S-like patterns, regardless of the unit of measurement, for example, cumulative costs, labor hours, or percentage of work (PMBOK® Guide 2004). In project management, S-curves, which are also called progress curves (Schexnayder and Mayo 2003), represent cumulative progress over time, which represents the amount of work done or to be done by a specific time throughout the execution period. An S-curve of a project connects the beginning point and the completion point of a project in terms of time and progress, typically, the budget. The completion point clearly indicates the objectives of the project in time and money and the curve connecting the points is the path to follow to finish the project in time and on budget. An S-curve for a project can be constructed from the resource-loaded project schedule for the project.

Many projects throughout different industries are actually planned and controlled by S-curves (Blyth and Kaka 2006; Miskawi 1989). In the construction industry, S-curves have been used as a graphical tool for measuring progress (Barraza et al. 2000; O'Brien and Plotnick 1999), as control limits based on early start and late start dates of activities (Schexnayder and Mayo 2003), and as a cash flow forecasting tool at relatively early stages of project life cycle (Fellows et al. 2002). For example, in cash flow forecasting, different standard S-curves are developed for different project groups based on a sample of historical projects (Blyth and Kaka 2006).

However, previous studies of the use of the S-curve as a quantitative performance management tool for on-going projects, not just visual displays, are very limited (Barraza et al. 2000; 2004; Cioffi 2005; Murmis 1997). Murmis (1997) generated a symmetric S-curve from a normal distribution and forced it to pass fixed points of the cumulative progress curve. Murmis applied the curve to detect problems in project performance. A more flexible S-curve was presented by Cioffi (2005). He modified a typical sigmoid curve used frequently in ecology by imposing two project progress constraints: the slope of the rise in the S-curve and the time at which half the total work has been completed. Barraza and his co-researchers tried to use a set of S-curves generated from a network-based simulation as a visual project control tool (2000) and, later, extended the concept to a probabilistic forecasting method by adjusting parameters of probability distributions of future activities with performance indices (the CPI and SPI in the earned value method) of finished activities (2004). Useful as they are, these previous works are not applicable to the current research of forecasting at-completion project duration and cost of ongoing projects, largely because of poor flexibility of the suggested S-curves and the lack of a mathematically sound forecasting framework based on actual performance information available at the time of forecasting.

The S-curve is a universal characteristic of all projects, regardless of the type, size, and complexity of a project. However the individual shapes for various projects vary according to the nature of the projects (Blyth and Kaka 2006; Miskawi 1989). Note that even linear projects with straight-line progress curves can be taken as S-curves (Table 4.2). Therefore, it is useful to develop a library of S-curve models that can be used to fit the individual progress curves. In the literature of forecasting, a lot of diverse S-curves can be found. For example, Meade and Islam (1998) identify 29 models for technological forecasting. In construction management, Skitmore (1992) compared four S-curve models for cash flow forecasting. Most of the models in the literature, however, have fixed shapes or very limited flexibility in representing various progress patterns such as front-end loaded, normal, and back-end loaded progress. And so, they can be categorized as fixed-shape S-curve models. The Multi-model BAF method uses such fixed-shape S-curve models while the BetaS-curve BAF method relies on a single more flexible model.

4.2.1 Fixed-shape S-curve Models

The accuracy and reliability of the BAF method depend heavily on how well a mathematical function or functions fit the progress curve template. A reasonable way of improving the quality of predictions from the BAF is to start with a large set of diverse S-curve functions and to screen them according to their goodness of fit to the progress curve template.

Among the S-curve models in the literature, five basic S-curve models are chosen to demonstrate the Multi-model BAF method and their formulas are shown in Table 4.1. The Pearl and Gompertz curves belong to the class of logistic functions that have many application areas in biology and economics. The Dual-Gompertz function is a modified form of the Gompertz to represent the progress of a project with back-end loading. It should be noted that these three logistic curves are defined inside the asymptotic limits $0 < w(t) < S$. In the application to project performance forecasting, an asymptotic approach to some specific value is problematic because every project should have finite beginning and completion states. Furthermore, while Function46 and Function50 have explicit parameters for the project duration (b) and the total work to be done (S), the logistic curves are defined over an infinite interval.

Table 4.1 Examples of simple S-curve models with two parameters

Name	Function
Pearl(t; a,b)	$w(t)=\dfrac{S}{1+a\exp(-bt)}$
Gompertz(t; a,b)	$w(t) = S\exp\{-a\exp[-bt]\}$
Dual-Gompertz(t; a,b)	$w(t) = S(1-\exp\{-a\exp[bt]\})$
Function46(t; n,b)	$w(t)=S\left[\dfrac{b-t}{b}\right]^{n+1}\left[\left(\dfrac{b-t}{b}\right)(n+1)-(n+2)\right]+S$
Function50(t; n,b)	$w(t) = S\left[\dfrac{(n+3)(n+2)(n+1)}{2b^2}\right]\left[\dfrac{b-t}{b}\right]^{n+1}\times \left[\dfrac{2b(b-t)}{n+2}-\dfrac{(b-t)^2}{n+3}-\dfrac{b^2}{n+1}\right]+S$

(Note: $w(t)$ represents the cumulative progress at time t and S represents the final state at completion)

In this research, a set of boundary conditions is assumed for the logistic curves to make them more suitable for project performance forecasting. Suppose that a project proceeds following the Pearl curve given as

$$w(t)=\frac{S}{1+a\exp(-bt)} \tag{4.1}$$

where w and S represent the cumulative progress at time t and the final state at completion, respectively. For a project with duration T, the two parameters a and b in the original formulas can be determined by a set of initial and completion conditions at t = 0 and t = T, respectively.

$$\begin{aligned} w(0) &= s_0 S \\ w(T) &= s_T S \end{aligned} \tag{4.2}$$

The two coefficients s_0 and s_T are the amount of work that define the beginning and completion, respectively, of the project. For example, s_0 = 0.01 and s_T = 0.99 represent the case in which a project starts with one percent of total work already done and is considered to be finished when 99 percent of the total work is done. Once the boundary conditions are given, the original parameters are calculated as

$$\therefore a = \frac{1}{s_0} - 1 \qquad \therefore b = \frac{\ln(a) - \ln\left(\frac{1-s_T}{s_T}\right)}{T} \tag{4.3}$$

With the same approach, the parameters of Gompertz and Dual-Gompertz functions in Table 4.1 are calculated for the same boundary conditions as follows.

$$\begin{aligned} &\text{Gompertz:} && \therefore a = -\ln(s_0) && \therefore b = \frac{\ln(a) - \ln\{-\ln(s_T)\}}{T} \\ &\text{Dual-Gompertz} && \therefore a = -\ln(1-s_0) && \therefore b = \frac{\ln\{-\ln(1-s_T)\} - \ln(a)}{T} \end{aligned} \tag{4.4}$$

It should be noted that the parameter *a* in Equation (4.3) and (4.4) is determined by the initial condition coefficient s_0 only, while the parameter *b* is determined by the project duration *T* as well as the boundary coefficients s_0 and s_T. Figure 4.1 shows the resulting curves generated with s_0= 0.01 and s_T= 0.99.

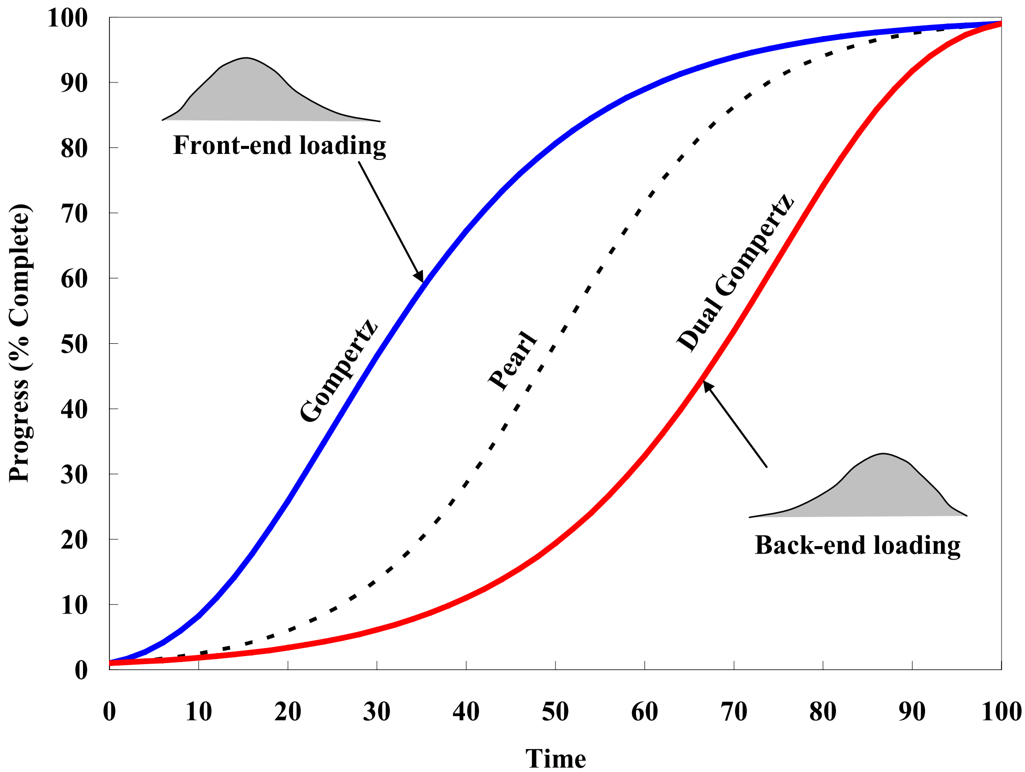
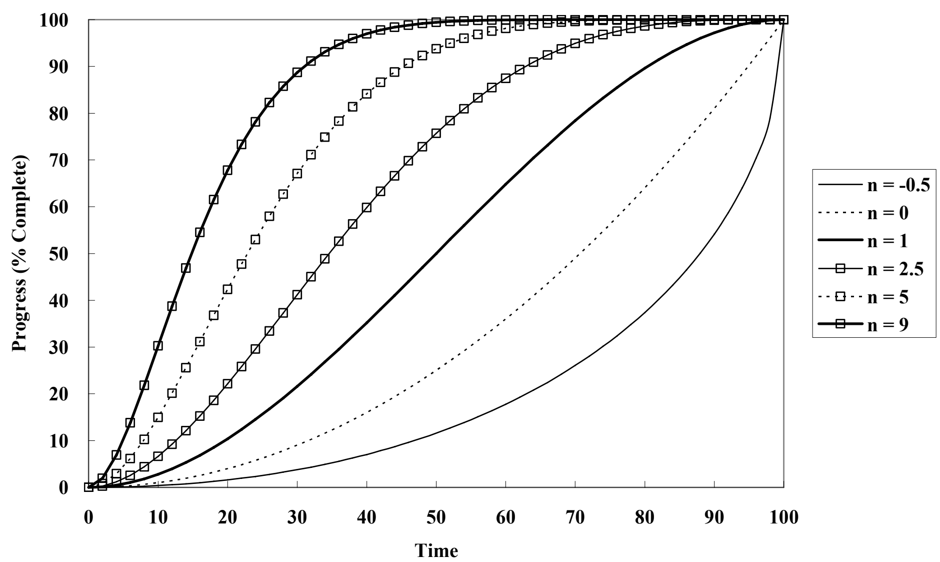
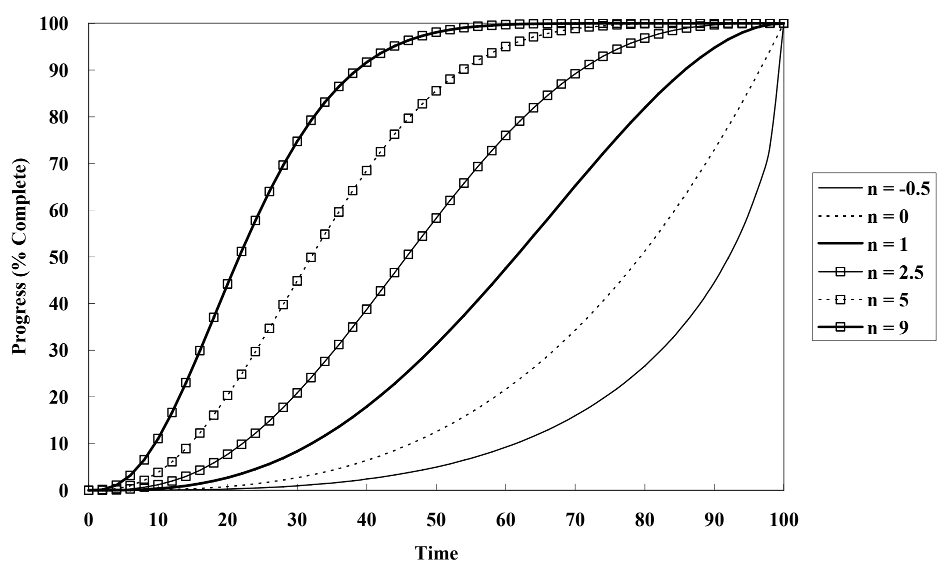


Figure 4.1 Examples of logistic curves (w(0)= 1, w(100)=99, T=100, and S=100)

"Function46" and "Function50" in Table 4.1 were developed by Reinschmidt (2006) and successfully used in the numerical example in this chapter. Recently, Function50 was applied in a probabilistic forecasting of the progress of nuclear power plants in the United States (Gardoni et al. 2007). These two functions have two advantages over the logistic curves above. Firstly, they can represent a range of different shapes by adjusting the first parameter n, which must have a value greater than -1. Figure 4.2 shows some examples of Function46 and Function50 with different shape factors. Secondly, the second parameter b directly represents the project duration. These two properties play important roles in their application in the Multi-model BAF. However, it should be noted that the flexibility of Function46 and Function50 in fitting various progress curves is relatively limited compared with the BetaS-curve model which has two shape parameters instead of one. Therefore, in this research, Function46 and Function50 are regarded as fixed-shape S-curves whose shape parameter must be fixed in advance based on the prior performance information.



(a) Function46



(b) Function50

Figure 4.2 Examples of Function46 and Function50

4.2.2 BetaS-curve Model

The BetaS-curve model is derived based on the beta distribution. The beta distribution provides a flexible S-curve function that can represent a wide range of cumulative progress patterns. The beta distribution has a long history of application in project management, especially in approximating subjective estimates and fitting curves to observed data. For example, the three-point estimates used in the program evaluation and review technique (PERT) are approximated by the beta distribution (Malcolm et al. 1959). AbouRizk and his co-researchers conducted a series of studies about fitting probability distributions to construction activity durations (AbouRizk and Halpin 1992; AbouRizk et al. 1991; 1994). In his study of fitting the beta distribution to construction duration data, a visual interactive procedure was tried (AbouRizk et al. 1991) and later more rigorous statistical methods – moment matching, maximum likelihood, and least-squares estimation – were tested against 80 construction data sets (AbouRizk et al. 1994).

In statistics, the beta distribution is a continuous probability distribution that is defined on the finite interval A to B with two shape parameters α and β. The probability density function (PDF) of a random variable X is:

$$f\left(x;\alpha,\beta\right)=\frac{\Gamma\left(\alpha+\beta\right)}{\Gamma\left(\alpha\right)\Gamma\left(\beta\right)}\frac{\left(x-A\right)^{\alpha-1}\left(B-x\right)^{\beta-1}}{\left(B-A\right)^{\alpha+\beta-1}},\quad \alpha>0,\beta>0,A\le x\le B \tag{4.5}$$

where $\Gamma\left(z\right)$ represents the Gamma function,

$$\Gamma\left(z\right)=\int_0^\infty t^{z-1}e^{-t}dt,\qquad \forall z>0 \tag{4.6}$$

The mean and variance of the random variable X are given as:

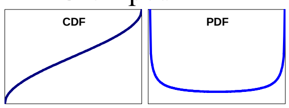
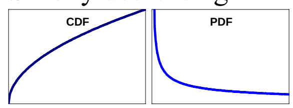
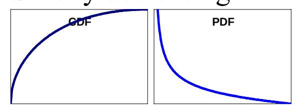
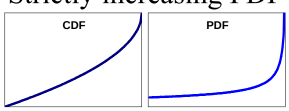
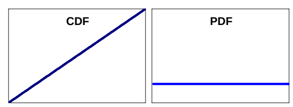
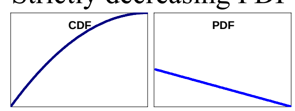
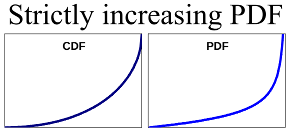
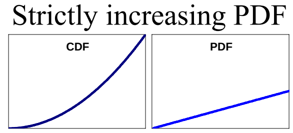
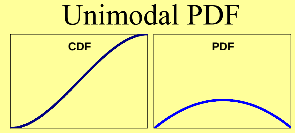
$$\mu_x = A + (B - A)\frac{\alpha}{\alpha + \beta}$$
$$\sigma_x^2 = \frac{(B - A)^2 \alpha\beta}{(\alpha + \beta + 1)(\alpha + \beta)^2} \qquad (4.7)$$

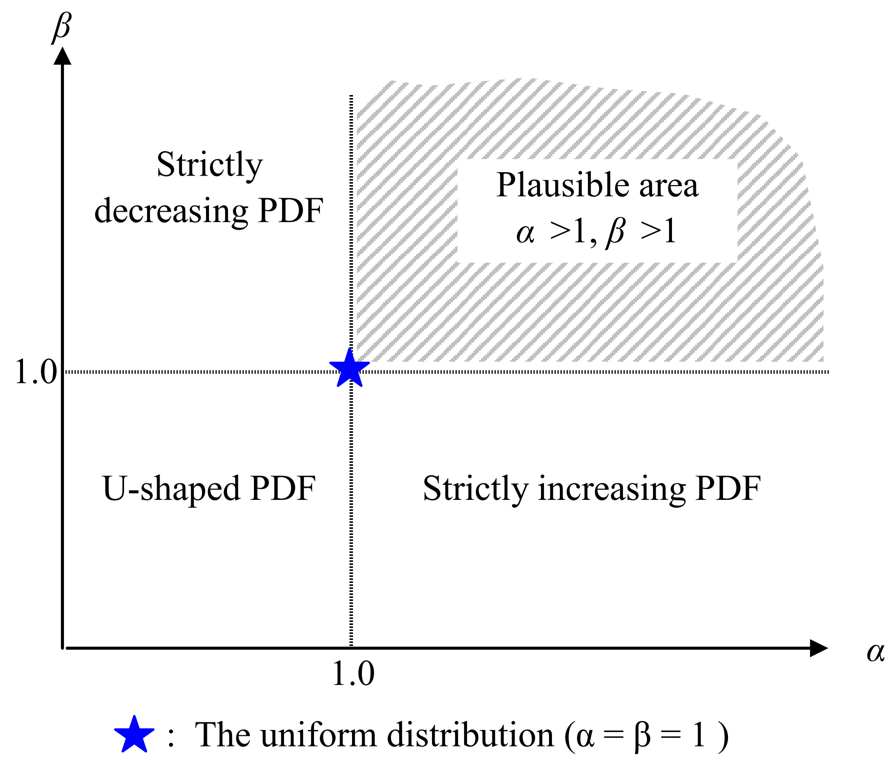
Like the Normal distribution, the cumulative distribution function of the beta distribution has no closed form expression.

The primary advantage of applying the beta distribution is the fact that the beta distribution can generate a wide range of shapes with only two parameters. For example, the beta distribution with $\alpha = 1.0$ and $\beta = 1.0$ is the uniform distribution. Table 4.2 shows some typical shapes of the beta distribution with different combinations of α and β.

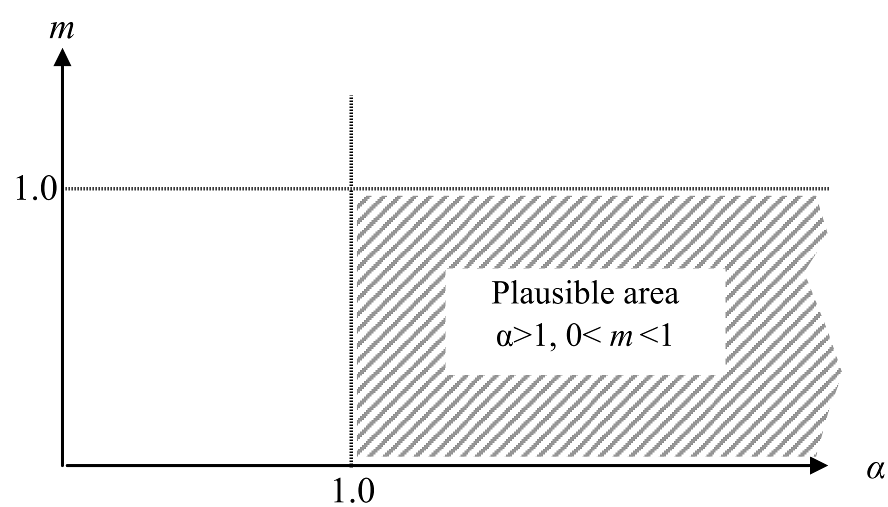
The BetaS-curve model is derived by imposing two constraints on the beta distribution. First, the model is defined over a finite interval [0, T], where T represents the project duration. Second, the PDF must have a unimodal shape which resembles the typical resource level distribution of projects during the execution period. The plausible range of the shape parameters that satisfy the unimodal condition is when $\alpha > 1.0$ and $\beta > 1.0$ shown in Figure 4.3 (a). More shapes of the beta distribution with $\alpha > 1$ and $\beta > 1$ are shown in Appendix A. Roughly speaking, bell-shaped curves similar to the Normal distribution are generated when $\alpha > 2$ and $\beta > 2$.

Table 4.2 Various shapes of the beta distribution

	$0 < \beta < 1$	$\beta = 1$	$\beta > 1$
$0<\alpha<1$	U-shaped PDF  (a) α=0.5, β=0.5	Strictly decreasing PDF  (b) α=0.5, β=1.0	Strictly decreasing PDF  (c) α=0.5, β=2.0
$\alpha = 1$	Strictly increasing PDF  (d) α=1.0, β=0.5	Uniform PDF  (e) α=1.0, β=1.0	Strictly decreasing PDF  (f) α=1.0, β=2.0
$\alpha >1$	Strictly increasing PDF  (g) α=2.0, β=0.5	Strictly increasing PDF  (h) α=2.0, β=1.0	Unimodal PDF  (i) α=2.0, β=2.0



(a) α-β plane



(b) α -m plane

Figure 4.3 Transformation of parameters and the plausible areas for project progress curves over α-β plane (a) and over α-m plane (b)

In the general definition of the beta distribution given in Equation (4.5), the location of the mode of the beta distribution on the interval A to B is given by the formula,

$$mode = \frac{B(\alpha-1)+A(\beta-1)}{(\alpha-1)+(\beta-1)} \tag{4.8}$$

From the first constraint A = 0 and B = T, the normalized mode (m) of the corresponding PDF is defined as:

$$m = \frac{mode}{T} = \frac{(\alpha-1)}{(\alpha-1)+(\beta-1)} \tag{4.9}$$

Note that when both α and β approach 1, $m = 0.5$.

Given α and m, the value of β is determined from Equation (4.9).

$$\beta = \frac{\alpha-1}{m} - (\alpha-2) \tag{4.10}$$

Then, the plausible area in the α-m plane is shown in Figure 4.3 (b). It should be noted that, in the context of construction project management, estimating the point in time when the maximum progress rate occurs is more meaningful than estimating a proper combination of α and β because it is much more intuitive.

Based on the properties of the beta distribution and the characteristics of project performance forecasting, the BetaS-curve model is defined as

$$\text{BetaS-curve}(t;\alpha,m,T):$$

$$\frac{dw(t)}{dt} = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{x^{\alpha-1}(T-x)^{\beta-1}}{T^{\alpha+\beta-1}}, \tag{4.11}$$

$$\text{where } \alpha > 1,\ 0 < m < 1,\ T > 0, \text{and } \beta = \frac{\alpha-1}{m} - (\alpha-2)$$

4.3 The General Formulation of Bayesian Adaptive Forecasting Method

4.3.1 Bayesian Inference

The heart of Bayesian inference is to update our knowledge in the light of new observations. A Bayesian approach also provides a systematic way of combining all pertinent information from various sources in terms of probability distributions. In the BAF, the parameters of an S-curve model are estimated and updated through Bayes' law whenever new actual performance data become available. If a project manager has an initial estimate of project progress (that is, a project plan) and if this progress curve is fitted to some known model with associated parameters (Θ), the belief in the individual model parameters can be updated with actual performance data (*D*) as the project proceeds. Bayes' law for this case can be written as:

$$P(\Theta \mid D) = \frac{P(D \mid \Theta) P(\Theta)}{P(D)} \tag{4.12}$$

where $P(\Theta)$ is the *prior distribution* reflecting the belief in parameters before observing new outcomes; $P(D \mid \Theta)$ is the conditional *probability* that the particular outcomes *D* would be observed, given the parameters Θ; $P(D)$ is the marginal distribution of the observables *D*; and $P(\Theta \mid D)$ is the *posterior distribution* of the parameters Θ given that the outcomes *D* were observed.

4.3.2 Bayesian Updating of Model Parameters

For an S-curve model with m parameters, the Bayesian updating process proceeds as follows. Let Θ denote the set of parameters $\{\theta_1, \theta_2, ..., \theta_m\}$. Then parameters are chosen independently so that the prior probability distribution of the parameter set is represented as

$$p(\Theta) = p(\theta_1)\, p(\theta_2) \cdots p(\theta_m) \tag{4.13}$$

Once a project gets started, actual progress is reported periodically and the data can be represented as a series of discrete values D.

$$D : (w_i, t_i),\; i = 1, ..., N \tag{4.14}$$

where w_i represents the cumulative progress reported at time t_i and N is the number of records up to the time of forecasting.

The likelihood of the data conditional on the parameters chosen is measured based on the errors between the actual times of performance reporting and the planned times determined by a specific S-curve model and the parameters, $T_M(w_i \mid \Theta)$. The goal is to seek an S-curve model and its associated parameters that make the errors normally distributed with zero mean and standard deviation σ. It is assumed that the random errors corresponding to different observations are uncorrelated. Then, the likelihood of the data conditional on the parameters can be calculated as the product of the likelihood of each observation.

$$
\begin{aligned}
p(D \mid \Theta) &= \prod_{i=1}^{N} p(t_i, w_i \mid \Theta) \\
&= \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2} \left(\frac{t_i - T_M(w_i \mid \Theta)}{\sigma} \right)^2 \right]
\end{aligned} \tag{4.15}
$$

It should be noted that the value of σ is determined by decision makers or forecasters to adjust the sensitivity of predictions to the actual data reported.

The marginal distribution of the observables D is determined from

$$
p(D) = \int p(D, \Theta) d\Theta \tag{4.16}
$$

where the joint probability distribution of data and parameters is constructed from Equation (4.13) and (4.15) as $p(D,\Theta) = p(D \mid \Theta)\, p(\Theta)$.

The ultimate goal of the Bayesian approach to prediction problems is to obtain a posterior marginal distribution of each model parameter conditional on the observed data. Using fundamental properties of conditional distributions, the posterior marginal distribution of parameter θ_i can be derived by integrating the joint parameter distribution conditional on the observed data, which is determined from equations (4.12) through (4.16), with respect to the remaining parameters $\Theta^{-i} = \{\theta_1, \cdots, \theta_{i-1}, \theta_{i+1}, \cdots, \theta_m\}$.

$$
p(\theta_i \mid D) = \int p(\Theta \mid D) d\Theta^{-i} \tag{4.17}
$$

4.3.3 Computation of Posterior Distributions using Monte Carlo Integration

Computing the posterior distributions derived above requires multifold integration over the parameters used in the analysis. In this research, a Monte Carlo integration technique has been successfully applied without resorting to more sophisticated methods such as importance sampling and Markov Chain Monte Carlo method.

Posterior marginal distributions of model parameters

The major results from the Bayesian updating process include posterior marginal distributions of the model parameters conditional on the observed data. Using the property of the conditional distribution, the posterior marginal distribution of a parameter θ_i is obtained by integrating the joint parameter distribution conditional on the observed data with respect to the remaining parameters Θ^{-i}.

$$\begin{aligned} p(\theta_i \mid D) &= \int p(\Theta \mid D) d\Theta^{-i} \\ &= \int \frac{p(\Theta, D)}{p(D)} d\Theta^{-i} \\ &= \int \frac{p(D \mid \Theta)\, p(\Theta)}{p(D)} d\Theta^{-i} \end{aligned} \tag{4.18}$$

Since random samples can be drawn from the prior distribution of θ_i, the integration in Equation (4.18) is approximately calculated using the Monte Carlo integration technique.

$$p(\hat{\theta}_i \mid D) = \frac{1}{p(D)} \frac{1}{N} \sum_{k=1}^{N} p(D \mid \theta_i, \Theta_k^{-i})\, p(\theta_i), \quad \Theta_k^{-i} \sim p(\Theta^{-i}) \tag{4.19}$$

Marginal likelihood of data

A typical challenge in the Bayesian approach to real problems is to calculate the marginal likelihood of data, $p(D)$, in Equation (4.19).

$$\begin{aligned} p(D) &= \int p(D,\Theta)\,d\Theta \\ &= \int p(D\,|\,\Theta)\,p(\Theta)\,d\Theta \end{aligned} \tag{4.20}$$

With Monte Carlo integration, a solution can be computed as

$$\begin{aligned} p(\widehat{D}) &= \int p(D\,|\,\Theta)\,p(\Theta)\,d\Theta \\ &\cong \frac{1}{N}\sum_{k=1}^{N} p(D\,|\,\Theta_k), \qquad \Theta_k \sim p(\Theta) \end{aligned} \tag{4.21}$$

Posterior estimates of model parameters

The future progress of a project can be estimated using the posterior estimates of model parameters. Once posterior marginal distributions of the model parameters are constructed, the expectation of each parameter conditional on the observed data is given as

$$\mathrm{E}[\theta_i\,|\,D] = \int \theta_i\, p(\theta_i\,|\,D)\,d\theta_i \tag{4.22}$$

If random samples from $p(\theta_i\,|\,D)$ can be generated, a numerical solution to Equation (4.22) can be easily calculated. However, the posterior marginal distribution of θ_i in the form of Equation (4.19) does not appear to be proper for random sampling techniques such as the Gibbs sampling. In this research, the Monte Carlo integration is again used.

$$
\begin{aligned}
\mathrm{E}\left[\theta_i \mid D\right] &= \int \theta_i p\left(\theta_i \mid D\right) d\theta_i \\
&= \int \left[\frac{\theta_i p\left(\theta_i \mid D\right)}{p\left(\theta_i\right)}\right] p\left(\theta_i\right) d\theta_i \\
&= \int \left[\frac{\theta_i p\left(D \mid \theta_i\right)}{p\left(D\right)}\right] p\left(\theta_i\right) d\theta_i \\
&= \frac{1}{p\left(D\right)} \int \left[\theta_i p\left(D \mid \Theta\right)\right] p\left(\Theta\right) d\Theta \\
&\cong \frac{1}{p\left(D\right)} \frac{1}{N} \sum_{k=1}^{N} \left[\theta_{i,k} p\left(D \mid \Theta_k\right)\right], \qquad \Theta_k \sim p\left(\Theta\right)
\end{aligned}
\tag{4.23}
$$

Prediction interval estimation

Primary outputs from the BAF method include the prediction interval around forecasts. The posterior probability density distributions of parameters can be generated from Equation (4.19). However, the prediction bounds corresponding to a predetermined acceptable level can not be directly obtained from the posterior distribution unless corresponding cumulative probability distributions are also available. To overcome this problem, it is assumed that the posterior distributions of parameters can be approximated as normal distributions. Then, the variance of the approximate normal distribution can be calculated from elementary statistics.

$$
Var\left[\theta_i \mid D\right] = E\left[\theta_i^2 \mid D\right] - \left\{E\left[\theta_i \mid D\right]\right\}^2 \tag{4.24}
$$

where the first term on the right side is calculated as

$$
\begin{aligned}
\mathrm{E}\left[\theta_i^2 \mid D\right] &= \frac{1}{p\left(D\right)} \int \left[\theta_i p\left(D \mid \Theta\right)\right] p\left(\Theta\right) d\Theta \\
&\cong \frac{1}{p\left(D\right)} \frac{1}{N} \sum_{k=1}^{N} \left[\left(\theta_{i,k}\right)^2 p\left(D \mid \Theta_k\right)\right], \quad \Theta_k \sim p\left(\Theta\right)
\end{aligned}
\tag{4.25}
$$

4.3.4 Use of Prior Performance Information

Along with its probabilistic nature, the use of prior performance information in conjunction with actual performance data characterizes the originality of the BAF method. The prior performance information is defined as all relevant performance information other than actual performance data, which is available even before the inception of a project. Table 4.3 summarizes some typical types of the prior performance information. The primary information that should be relied on for project performance forecasting is the past performance data observed in the project itself. Early in a project, however, project managers may suffer from a lack of sufficient actual performance data to make reliable forecasts, resulting in deferring any judgment about performance control at the risk of missing the opportune time to take appropriate corrective actions. Moreover, the value of even highly reliable forecasts progressively decreases as the project continues.

The Bayesian adaptive forecasting method provides a systematic way of making more reliable forecasts sooner, when they are more valuable to project management. In the general framework of Bayesian inference and forecasting, the BAF method combines all relevant information concerning future performance of an ongoing project, including documented project plans, historical data, subjective judgment, as well as the actual performance data being generated by the project itself.

Table 4.3 Information used in the Bayesian adaptive forecasting method

Types	Descriptions
Project activity network **G** = G(n, r)	A project activity network is represented as a directed acyclic network G(n, r), where n is the number of activities in the network and r is the number of Finish-to-Start relations among the activities.
Activity durations and costs **A** = {a_1, a_2, …, a_n}	Each element of the information set, **A**, has probabilistic estimates of activity duration and cost.
Historical data **H** = {h_1, h_2, …, h_m}	A historical data set might be represented as a series of discrete performance indicators throughout the lifecycle of past projects. m is the number of historical progress data sets.
Subjective information **S** = {c_1, c_2, …, c_s}	s is the number of subjective information sets that can be chosen by decision makers. A subjective information set may be represented in terms of numerical constraints for the model parameters.

In the BAF method for predicting schedule performance of a project, the prior performance information consists of two elements: the prior probability distribution of project duration and the progress curve template. Figure 4.4 shows these elements in a graphical way. First, the prior distribution of project duration represents the best probabilistic estimate of the project duration, which is made without observing actual performance data. When an activity network and probabilistic estimates of activity durations and costs are available for a project, well-known methods such as PERT (Malcolm et al. 1959) and a network-based CPM simulation (Lee 2005; van Slyke 1963) can be used to generate the prior distribution of project duration. In the absence of detailed project plans, a subjective probability estimate can be made based on a single value estimate of the project duration. For example, if the total project duration of a project is fixed as 60 months, someone with experience in similar projects might be able to estimate the possible range of the actual project duration as, for example, between 58 months and 64 months. Then one can further approximate a probability distribution over the range. The planned project duration may be chosen according to the level of risk accepted by the organization. Regardless of the methods used to generate it, a prior probability distribution of project duration is updated later with the actual performance data.

The second element of prior performance information, the progress curve template, represents the prior knowledge of the project manager and project engineers about the plausible progress pattern of the actual performance. As pointed out in Section 4.1, the fundamental premise of the Bayesian adaptive forecasting method is that there exist characteristic progress patterns for individual projects. The progress curve template of a project represents the characteristics of the project in terms of the cumulative progress over time. In the BAF method, the progress curve template is used to forecast future performance of the project. The overall reliability of the BAF relies on the degree to which actual performance data match the progress curve template.

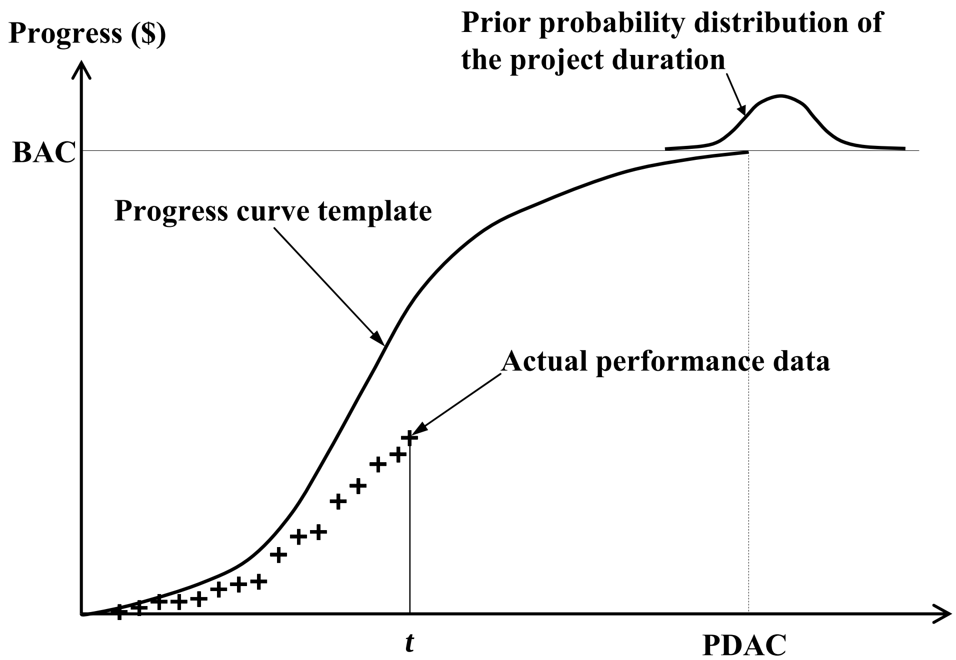


Figure 4.4 Two elements of the prior performance information and the actual performance data

It should be noted here that incorporating prior performance information into project performance forecasting does not necessarily mean that the quality of outcomes will be better than that of the forecasts based purely on actual performance data. Use of misinformation will delay the convergence of forecasts to the actual value. In a Bayesian approach, the quality of prior information can be measured in terms of the degree of bias and the precision of a prior probability distribution. The degree of bias of a prior probability distribution represents how far, on the average, the mean of the probability distribution is from the right value. On the other hand, the precision of a prior distribution is defined as the inverse of its variance. Ideally, the Bayesian approach works best with unbiased and precise prior distributions. However, when a prior distribution is precise but seriously biased, that will leave little chance for the information in the future to influence the final outcomes. Therefore, it is important to constantly evaluate the appropriateness of prior performance information in light of actual performance data.

4.4 Multi-model BAF

Based on the general formulation of the BAF in Section 4.3, two different methods – the Multi-model BAF method and the BetaS-curve BAF method – are developed and derived in this section and in Section 4.5, respectively. These two methods are distinguished by the use of the progress curve template in Figure 4.4 which is built based on prior performance information.

The outline of the Multi-model BAF method is shown in Figure 4.5. Once the progress curve template and the probability distribution of total duration for a project are developed from the prior performance information available, the Multi-model BAF method proceeds through five steps: (1) selecting models, (2) generating prior distributions of model parameters; (3) updating model parameters, (4) forecasting, and (5) combining forecasts. Among theses steps, the third step is carried out according to the general BAF framework in Section 4.3. The remaining part of this section discusses the other steps in detail.

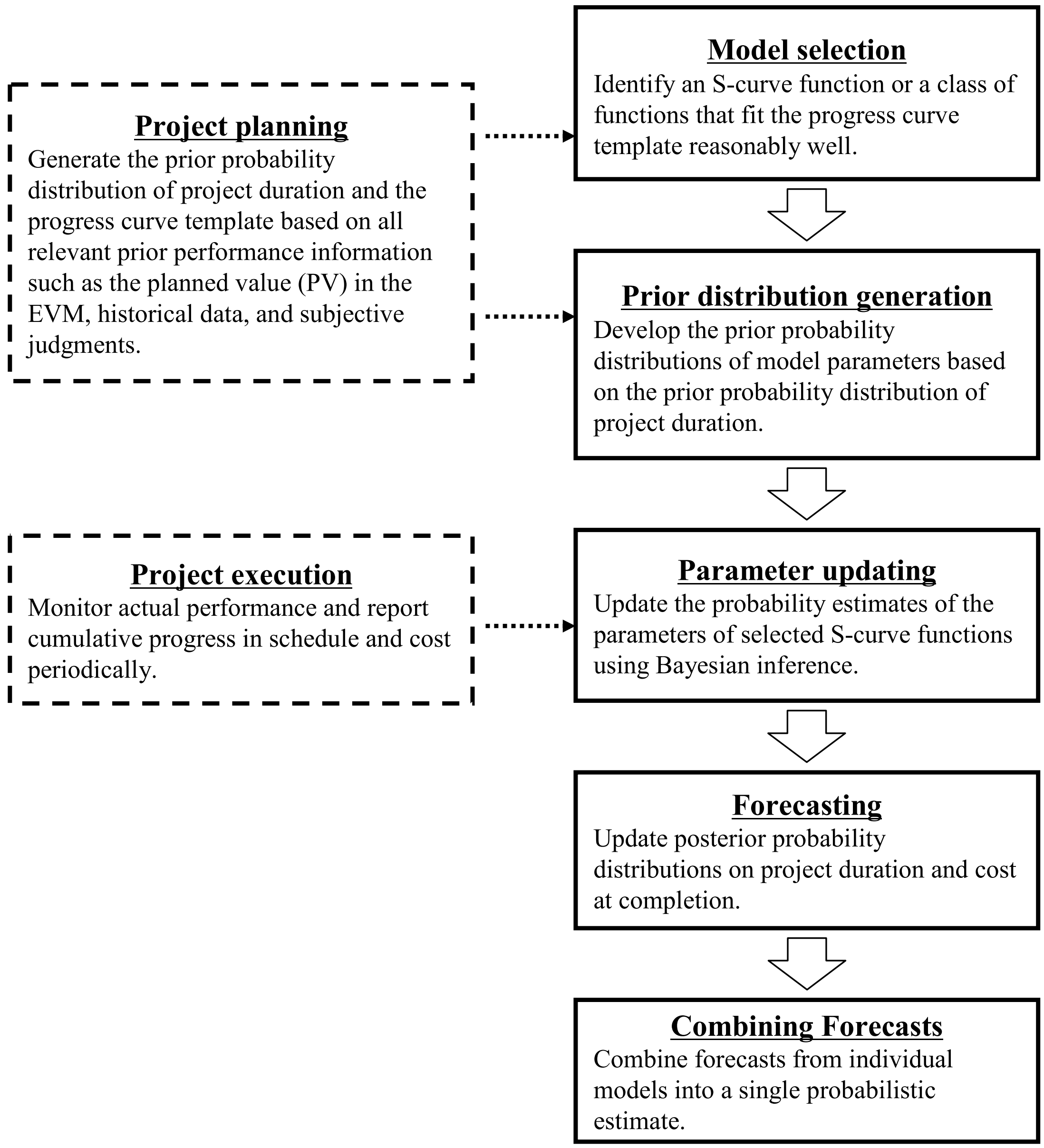


Figure 4.5 Outline of the multi-model Bayesian adaptive forecasting

4.4.1 Selecting Models

Once the two elements of prior performance information are developed for a project, they are separately incorporated into the Multi-model BAF method. The progress curve template is used to test potential S-curve models and to select models that fit the template reasonably well. This step is necessary because of the limitation of fixed-shape S-curves in representing various shapes of real project progress curves. As the name says, each S-curve model belonging to the fixed-shape category has its unique progress pattern which may or may not fit well to the progress curve template of a project. In the case of Function46 and Function50, the progress pattern can be adjusted by changing the first parameter n. However, the range of possible S-curve patterns generated by Function46 and Function50 is still limited as shown in Figure 4.2. In this research, the first parameter n of Function46 and Function50 is considered a constant that is determined in advance rather than a random variable which is updated with actual performance data.

It should be noted that the pool of potential S-curve models is expandable. That is, the user can add his or her own favorites. In the Multi-model BAF method, all available fixed-shape S-curve models are tested to fit the progress curve template. Then, the models with good fitting results are chosen as the candidate models whose parameters are updated with actual performance data. A simple way of evaluating and comparing the fitness of different S-curve models to the progress curve template is the method of least squares (Please refer to Appendix B for further information). Once proper models are selected, they are used to fit actual performance data and to generate posterior distributions of the model parameters from which the project duration at completion is predicted.

4.4.2 Generating Prior Distributions of Model Parameters

The other element of prior performance information, the prior probability distribution of project duration, needs first to be converted to the probability distribution of the corresponding parameters of individual models, which will be later updated as actual performance is reported periodically. In the case of Function46 and Function50, this transformation is unnecessary because parameter b itself represents the project duration. For the logistic functions (the Pearl, the Gompertz, and the Dual Gompertz functions) in Table 4.1, however, a simulation approach can be used to build the prior probability distribution of the slope parameter b from the prior distribution of project duration $p(T)$. For example, if the project duration is estimated to have a normal distribution with mean of 100 time intervals and standard deviation of 10 time intervals, the corresponding probabilistic distribution of the slope parameter can be obtained using Equation (4.3) and (4.4) in Section 4.2.1. A straightforward way of doing this is Monte Carlo simulation. Figure 4.6 shows the prior distributions of parameter b in the Pearl, the Gompertz, and the Dual Gompertz models under the assumption that s_0 is 0.01.

The simulation approach to generating prior distributions of the model parameters has the advantage that a prior distribution of the slope parameter from various forms of the project duration can be used as long as the duration is represented in a probabilistic way. For example, one can use a sequence of random durations from a

Monte Carlo simulation of a network schedule as input to generate corresponding random parameters. Considering that risk management is getting recognized as an essential part of project management, obtaining such data appears to have become easier than ever. Furthermore, commercial software packages such as Primavera® Project Planner and MS Project® provide similar functions for probabilistic estimation.

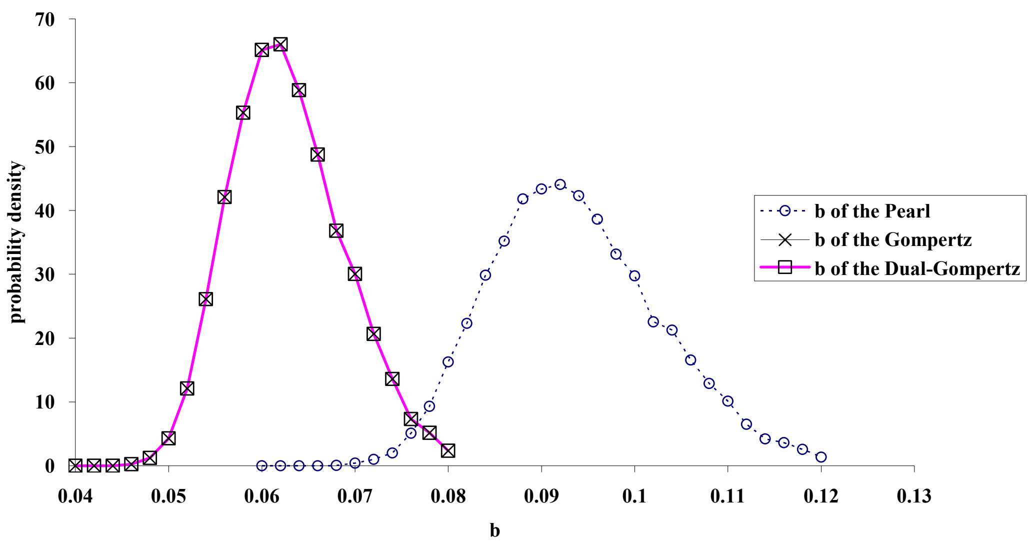


Figure 4.6 Prior distributions of the slope parameter of the Pearl, the Gompertz, and the Dual Gompertz functions for the prior distribution of project duration $N(100,10^2)$

4.4.3 Combining Predictions

In the Multi-model BAF, the limited flexibility of individual S-curve models is overcome by the Bayesian model averaging technique. The basic idea of the MBAF is to predict future progress of a project by using different models and combining predictions from the component models with weights according to the relative model likelihood.

Combining predictions has a long history of recognition as a viable solution to improve the quality of forecasting. Decision makers who need to predict any quantity of interest often get confused with a large range of forecasted values from various models which are based on different ideas and information. In one way, a decision maker might try to select the best among the plausible forecast models. However, if there are many independent forecasts that appear to be based on some reliable methodologies and information, it might be a waste of information to stick to the simplicity of using one method at the cost of discarding all the other sources of information. An axiomatic idea underlying combining forecasts is to maximize information usage by integrating all methods and relevant information (Bates and Granger 1969; Bunn 1975). It should be noted that if some methods are highly correlated, only one is needed. When original forecasts have complementary nature in terms of methodologies or/and exclusive information used, a reasonable decision maker would try to identify the best combination which yields a combined forecast better than any of the original forecast. Combining predictions is also useful when dealing with diverse estimates or advices from a plethora of human experts. One approach that can be used for combining diverse experts' forecasts is to apply a weight to each expert.

Since the paper by Bates and Granger (1969), the efficiency of linear combination of forecasts has been examined by many researchers (Bunn 1985; Newbold and Granger 1974; Winkler and Makridakis 1983). Most of them used variance of the error distribution as a measure of efficiency of combining methods, while some researchers try to apply the perspective of multi-criteria decision making process (de Menezes et al. 2000; Reeves and Lawrence 1991).

Bayesian model averaging (BMA) provides a coherent and intuitive framework for combining forecasts from the perspective of model uncertainty. Model uncertainty arises when a decision maker selects a model, presumably the best model, and makes an inference from data as if the selected model were the exact one that had generated the given data. Ignoring model uncertainty can lead to underestimation of the uncertainty about the quantities of interest and over-confident inferences (Hoeting et al. 1999; Raftery et al. 1997). Bayesian model averaging is an approach to overcome the problems due to ignored model uncertainty by combining all plausible models in a weighted average sense. The weights for individual models represent the adequacy of the model and can be determined as their posterior model probability. In addition, the burden of justifying the choice of a single predictive model can be avoided with the Bayesian model averaging technique (Hoeting et al. 1999).

In the Multi-model BAF method, the Bayesian model averaging technique is used to integrate different predictions from individual S-curve models into a single prediction. Once a project gets started, actual performance data, D, are periodically monitored and used to update the forecasts, for example the estimated duration at completion (EDAC), from individual models. Then the combined prediction can be calculated as

$$p(EDAC \mid D) = \sum_{k=1}^{m} p(EDAC \mid M_k, D)\, p(M_k \mid D) \tag{4.26}$$

where m is the number of models in the combination. This is an average of the posterior distributions from all component models, weighted by their posterior model probability. The posterior model probability for individual models is calculated as

$$\begin{aligned} p(M_k \mid D) &= \frac{p(D \mid M_k)\, p(M_k)}{p(D)} \\ &= \frac{p(D \mid M_k)\, p(M_k)}{\sum_{i=1}^{m} p(D \mid M_i)\, p(M_i)}, \qquad k = 1, \ldots, m \end{aligned} \tag{4.27}$$

where $p(D \mid M_k)$ is the likelihood of data D conditional on a model M_k and $p(M_k)$ is the prior model probability for M_k.

4.4.4 Example 4.1*

An artificial project has been analyzed to demonstrate the performance of the Multi-model BAF model. Table 4.4 shows the major inputs used in the generation of the artificial project data. From a Monte Carlo simulation, the mean and the standard

*Reprinted with permission from “An S-curve Bayesian model for forecasting probability distributions on project duration and cost at completion” by Kim and Reinschmidt (2007). Construction Management and Economics - the 25th Anniversary Conference, Copyright [2007] by Taylor & Francis http://www.informaworld.com.

deviation of the prior distribution of the completion date are estimated as 71 weeks and 6.6 weeks, respectively. Figure 4.7 shows the planned progress and the simulated "actual" progress used in the numerical example. The planned is determined by averaging stochastic progress curves from 5000 iterations over the progress dimension. The dotted line represents a simulated progress curve with completion date at week 71.

Table 4.4 Input data for random progress curve generation

Input parameters		Value
Number of activities		200
Number of effective precedence relations		279
Activity duration for planned progress	Mean	4 (weeks)
	Standard deviation	2 (weeks)
Activity duration for actual progress	Mean	4 (weeks)
	Standard deviation	2 (weeks)

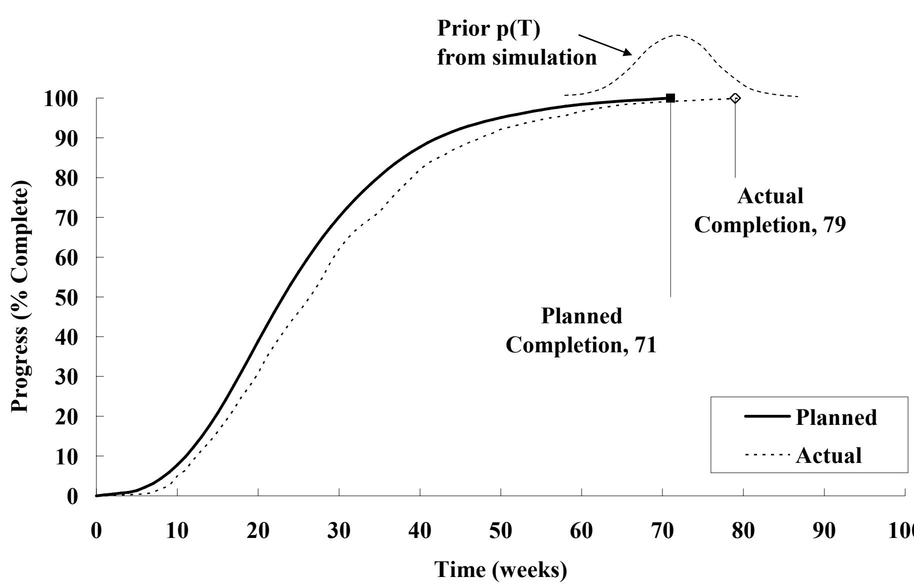


Figure 4.7 The planned progress and the "actual" progress for Example 4.1 (Prior p(*T*) is drawn not to scale)

The five fixed-shape S-curve models discussed earlier are fitted to the planned progress curve. The least squares method is used to measure the goodness of fit and four functions other than the Dual Gompertz turn out to fit the planned progress curve reasonably well. The fitted graphs are shown in Figure 4.8 and the corresponding prior estimates of the model parameters are shown in Table 4.5. It should be noted that the prior distribution of the completion date, which is obtained from the network-based schedule simulation, is also taken into account in terms of a random variable *T* in the formulas for slope parameter *b*. The results show that the Gompertz function best fits the current planned progress with the minimum sum of the squares of vertical deviations.

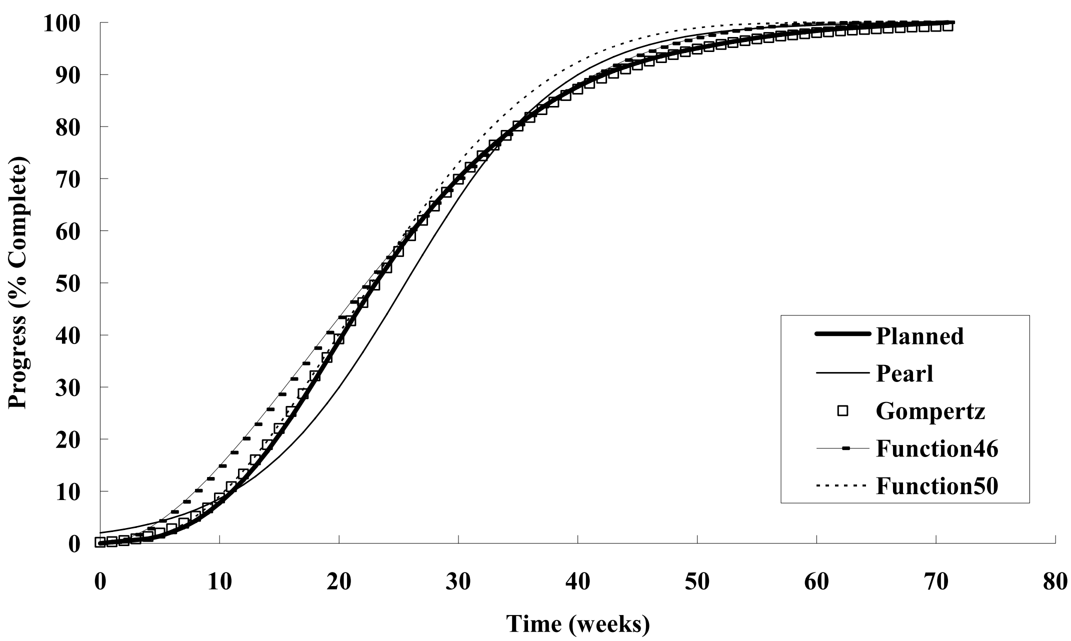


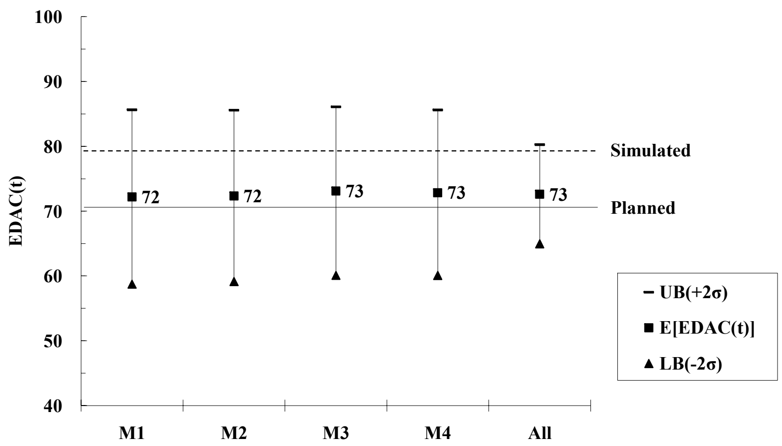
Figure 4.8 S-curve functions fitted to the planned progress

Table 4.5 Prior estimates of model parameters

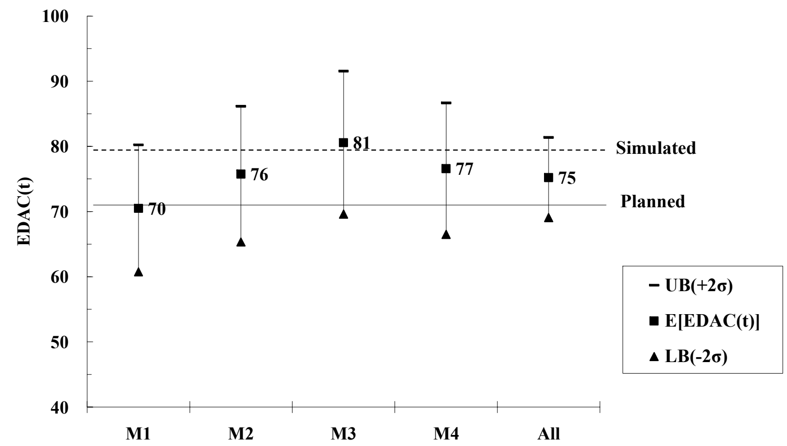
	Parameters for shift (a) or shape (n)	Slope parameter (b)	Sum of the squares of vertical deviations
Pearl	a = 49.0	[ln(49) – ln(1/49)]/T	1102
Gompertz	a = 6.377	[ln(6.337) – ln(0.00702)]/T	23
Function46	n = 3	T	765
Function50	n = 5	T	520

With the prior estimates of model parameters and weekly performance reports, EDAC is calculated after 10, 20, 30, and 40 weeks using the four S-curve functions. The results are shown in Figure 4.9. Different predictions from different models are combined with weights according to the relative reliability of the models, which is determined as the marginal probability of observing the actual performance data conditional on each model. The predictions at week 10 show that all models provide similar results that are close to the prior estimates of the completion date. This is attributed to the fact that the Bayesian adaptive model uses both prior information and actual progress data, and the prediction is influenced more by the plan than it is by the small number of data points reported up to this time.

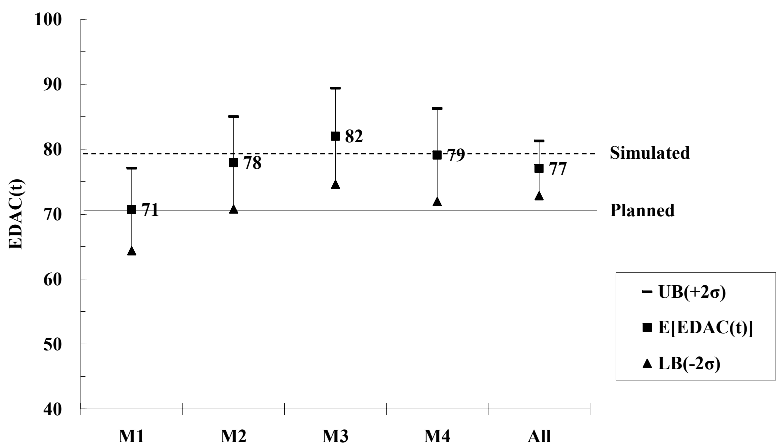
The other three graphs in Figure 4.9 show the adaptive nature of the BAF. Undoubtedly the forecast would be repeated after every reporting period to incorporate the new data, but here the intermediate updates have been omitted. Comparing the forecasts at different points of project execution, it is easy to see that each model responds to the actual performance data in a different way. After 20 weeks, predictions by Gompertz (M2), Function46 (M3), and Function50 (M4) move away from the plan and approach the actual completion date, due to the accumulated discrepancies between the actual reports and the planned progress curve. However, it should be noted that the planned completion date is still inside of the prediction bounds or control limits that are determined as two standard deviations above and below the expected EDAC. The prediction bounds (control limits) should be determined in advance by project managers as a level of risk accepted by the owner or organization. As more actual data accrue, the prediction bounds on the predicted values become narrower for all models. After 30 weeks, the combined prediction gets closer to the actual completion date, and more importantly, the prediction bounds indicate that the probability of completing the project within the planned completion date becomes lower than the predetermined acceptable level. Predictions after 40 weeks clearly show these patterns and the combined forecast provides more accurate prediction of the actual completion date than its component forecasts.



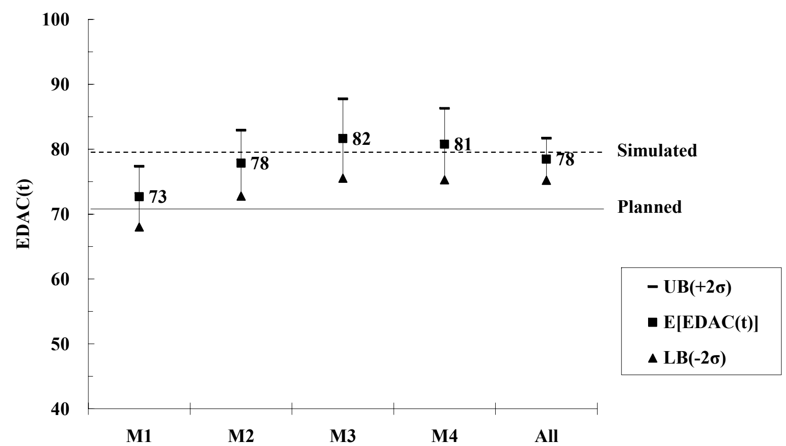
(a) Predictions after 10 weeks



(b) Predictions after 20 weeks



(c) Predictions after 30 weeks



(d) Predictions after 40 weeks

Figure 4.9 The estimated duration at completion at different times (M1-Pearl; M2-Gompertz; M3-Function46(*n*=3); M4-Function50(*n*=5))

4.5 BetaS-curve BAF

Instead of using multiple S-curve models, the BetaS-curve BAF (BBAF) method uses the BetaS-curve model in Section 4.2.2 to fit a wide range of progress curves for diverse projects. The flexibility of the BetaS-curve model is attributed to the property of having three free parameters rather than two. Along with the interval parameter T which directly represents the project duration, the BetaS-curve model has the two additional shape parameters α and m, which enables it to incorporate into forecasting prior performance information in various forms such as schedule network, activity estimates, historical data, and subjective judgment.

The outline of the BetaS-curve BAF method is shown in Figure 4.10. The method consists of three steps: (1) generating prior distributions of model parameters; (2) updating model parameters, and (3) forecasting. It should be noted that the BBAF method does not require the model selection and the combining forecasts steps in the Multi-model BAF method.

Once the prior performance information of a project is developed as discussed in Section 4.3.4, the information is used to generate the corresponding prior distributions of the BetaS-curve parameters. Depending on the types of information at hand, different approaches can be used. Specific methods proposed for different situations will be discussed below. It should be mentioned here that, regardless of the types of approach used, all their final results are represented as probability distributions of the corresponding BetaS-curve parameters.

The periodic performance data are used to revise prior beliefs on the model parameters through Bayesian inference. The process of monitoring and updating is cyclic and this adaptive property represents a key aspect of the Bayesian approach in dealing with sequential data. Updated distributions of model parameters are used to generate meaningful project information such as the estimated duration at completion (EDAC), associated prediction intervals, and the probability of meeting the planned project duration.

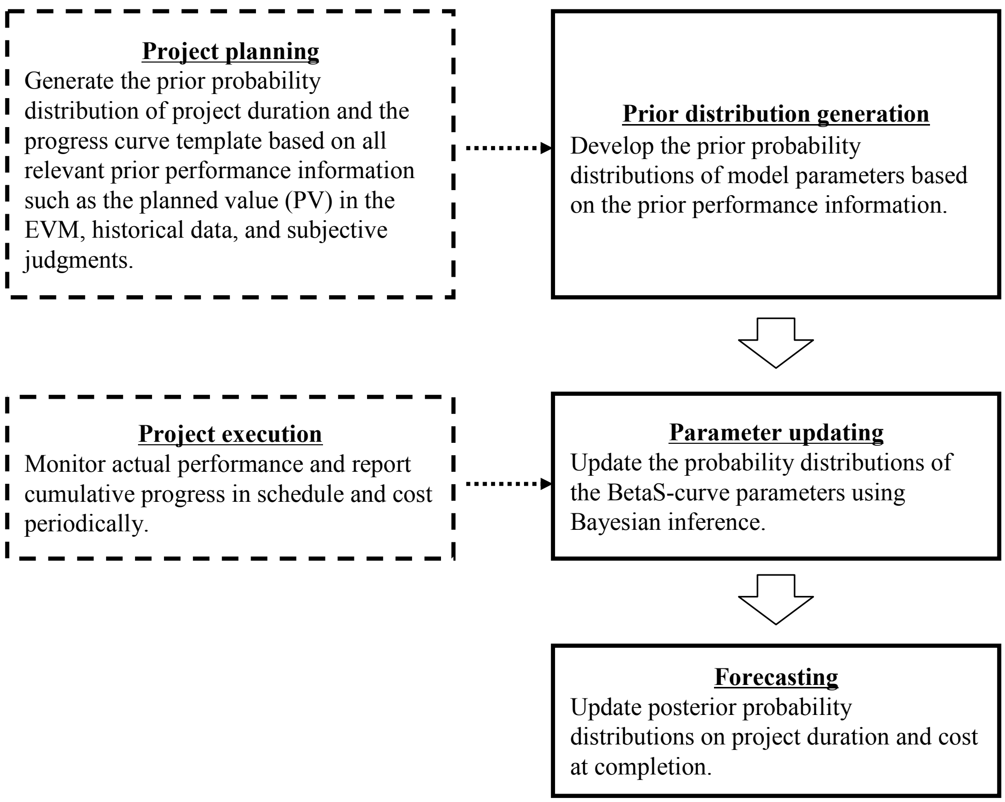


Figure 4.10 The outline of the BetaS-curve BAF method

4.5.1 Generating Prior Distributions of BetaS-curve Parameters

The BeatS-curve model has three paramters: α, m, and T. Just like Function46 and Function50, the BetaS-curve model has a parameter (T) that explicitly represents the project duration. As a result, the probability estimate of project duration based on prior performance information is directly used as the prior probability distribution of parameter T.

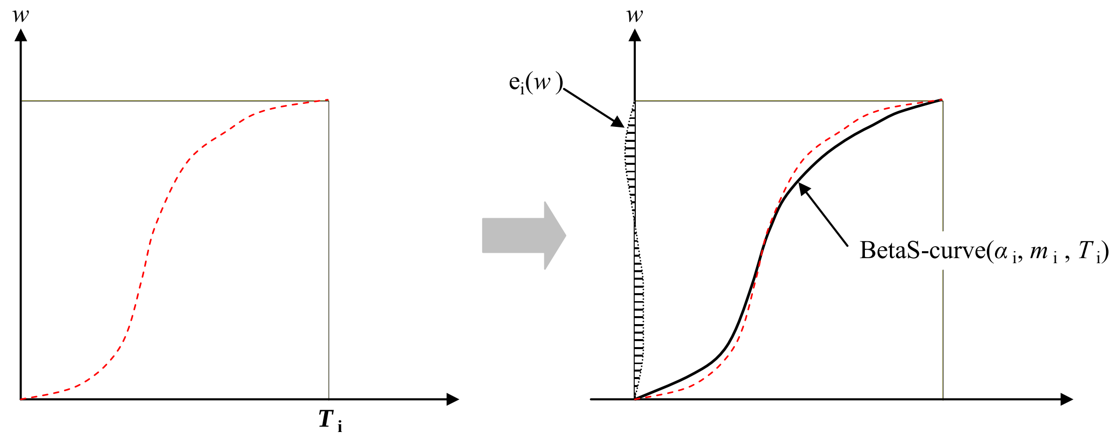
In order to construct the prior distributions of the other two parameters, α and m, different approaches can be applied, depending on the types of available information. In this dissertation, two typical situations are addressed.

- Prior distributions based on a stochastic network schedule $p(\alpha,m,T \mid \mathbf{G}, \mathbf{A})$;
- Prior distributions based on historical data or other subjective judgment $p(\alpha,m,T \mid \mathbf{H})$ or $p(\alpha,m,T \mid \mathbf{S})$.

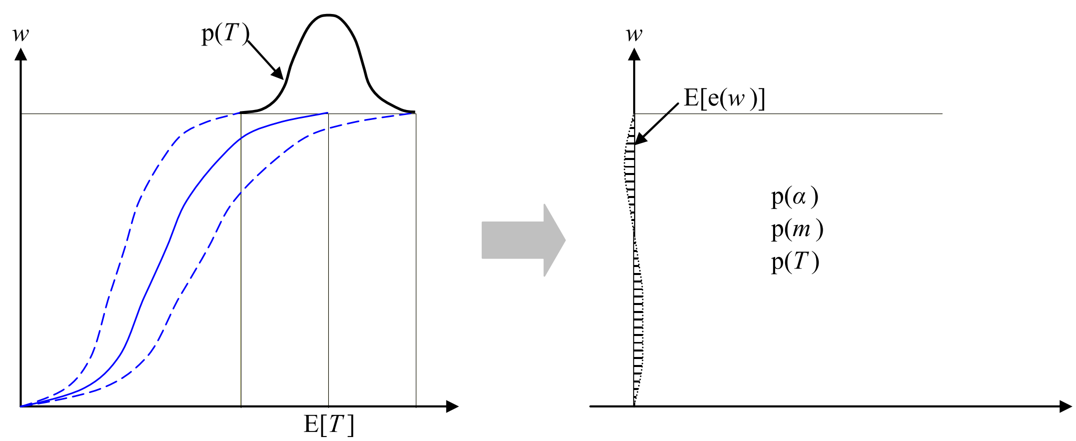
Prior distributions based on stochastic network schedule

A stochastic network schedule of a project is defined as a network schedule of the project in which activity durations and costs are estimated in a probabilistic way, for example, three-point estimates. Unlike typical, deterministic network schedules, the project duration based on a stochastic network schedule is calculated by a network-based Monte Carlo simulation and the results are represented as a probability distribution of project duration. Given a stochastic network schedule for a project, a large sample of potential progress curves, which are collectively called stochastic S-curves of the project (Barraza et al. 2000; 2004), can be generated using the simulation approach. Network-based schedule simulation has been used mostly as a method for quantifying schedule risk before the inception of a project (Lee 2005). In this research, a typical network-based schedule simulation is combined with a curve fitting technique to develop prior probability distributions of the shape parameters of BetaS-curve model. With the method, the stochastic nature of progress curves of a project can be quantified in a systematic way and represented as a set of prior probability distributions of the BetaS-curve model parameters.

The basic concepts and procedures of the method are shown in Figures 4.11 and 4.12, respectively. Given an activity network **G** and probabilistic estimates of activity durations and costs **A,** a group of plausible S-curves, or stochastic S-curves, of the project can be generated using a network-based schedule simulation. Each of the stochastic S-curves can be fitted with the BetaS-curve model. Along with the resulting best-fit parameters (α, m, T), the error distribution $e(w)$ is determined over the level of progress w, which represents the horizontal error between the S-curve generated in a simulation and the BetaS-curve approximation. If this fitting process is applied repeatedly to all S-curves generated, a set of marginal probability distributions of individual parameters – p(α), p(m), p(T) – can be obtained along with the correlation coefficients among them.



(a) Fitting a single progress curve



(b) Fitting a large sample of progress curves

Figure 4.11 Generating prior distributions for the BetaS-curve parameters from a large sample of progress curves (Error curves are drawn not to scale)

The overall error distribution from n fittings is measured in terms of the average fitting error (AFE), which is defined as the horizontal deviation, at specific level of progress w, between the average progress curve and the BetaS-curve determined by the expected values of the marginal probability distributions of parameters.

Given n random progress curves for a project, the resulting sequence of best-fit parameter values (α_i, m_i, T_i), for $i = 1,\ldots, n$, can be regarded as a set of random values drawn from the probability distributions, whatever they are, of the model parameters. The parameter values can directly be used in the computation of posterior distributions of interest. This method is referred to as the direct sequence method. An advantage of the direct sequence method is that dependences among model parameters, if any, are automatically taken into account. Otherwise, the prior distributions of model parameters can be approximated with reasonable probability distributions so that random numbers can be drawn for the Bayesian updating computation. Figure 4.12 depicts the procedures of developing prior distributions of BetaS-curve model parameters based on a stochastic network schedule.

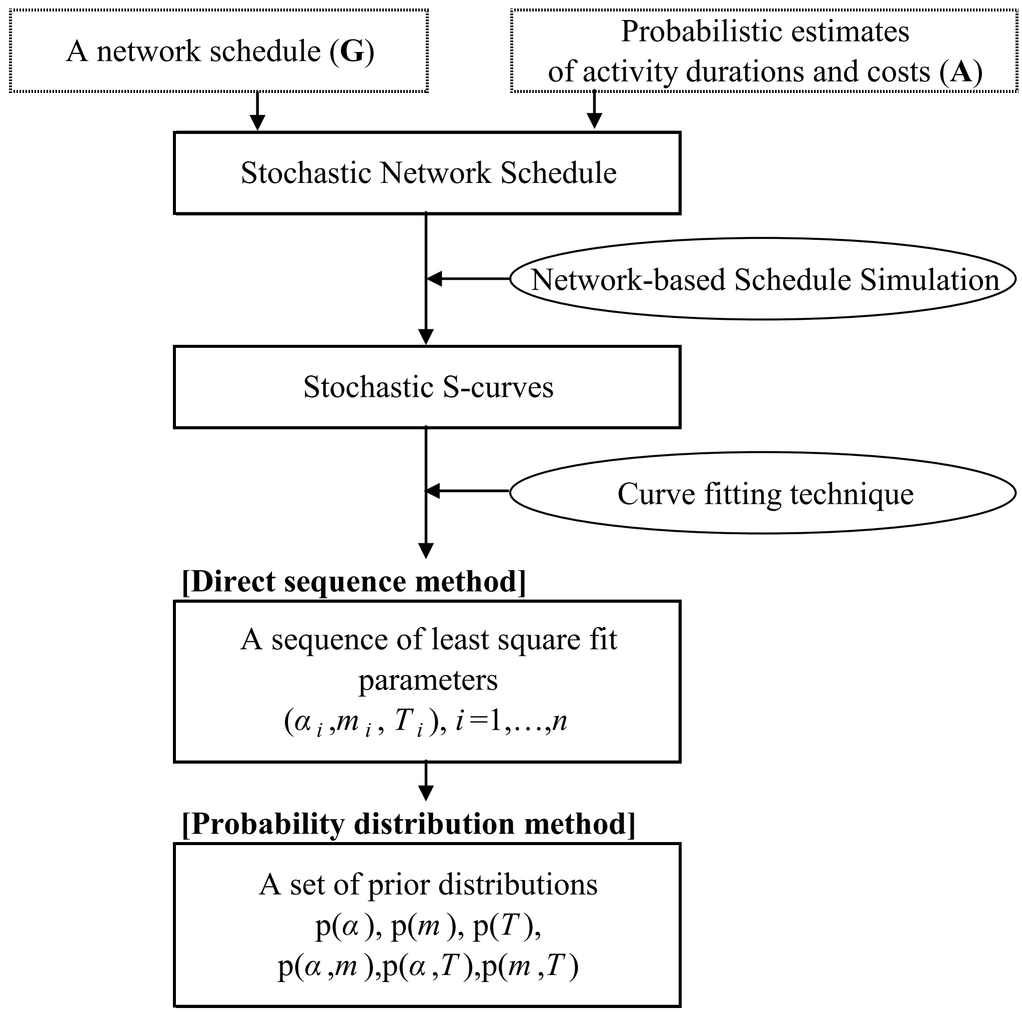


Figure 4.12 Procedures of developing prior distributions, given an activity network (G) and probabilistic estimates of activity durations and costs (A)

Prior distributions based on historical records or other subjective judgment

Historical data about cumulative progress curves from similar projects are valuable information for developing prior distributions of model parameters, especially in the absence of reliable project plans such as WBS, project schedule, and activity level estimates. When cumulative progress records of past projects are available, the BetaS-curve can be fitted to the data and the resulting best fit parameter values can be used in choosing reasonable probability distributions for model parameters. The best-fit parameters can be determined with the least squares method discussed in Appendix B.

This method is based on the premise that similar construction projects share some unique progress characteristics in terms of the time of the peak progress rate (or mode) and the overall progress pattern. For example, front-end loading projects and back-end loading projects would be distinguished by the mode parameter, m, of the BetaS-curve while linear and nonlinear projects would have different ranges of α parameter. When information directly related to a specific project is not available, subjective judgments about actual progress patterns can be applied.

4.5.2 Example 4.2

The BetaS-curve BAF is applied to an artificial project which is generated following the same procedure used in Example 3.1. In this example, four cases of prior distributions are used in predictions and the results are compared. The main purposes of the series of predictions for the same project are to demonstrate the merits and properties of BetaS-curve BAF model with different sets of prior distributions for the model parameters and to show how effectively the method quantifies prior performance information and how the forecasting performance of the BBAF is influenced by the use of different prior information. The major inputs are summarized in Table 4.6. The resulting baseline progress curve and the simulated "actual" progress curve to be used in the forecasting are shown in Figure 4.13.

It is assumed that the network schedule and the probabilistic estimates of activity durations are established by the project team. With this information, a set of stochastic progress curves can be generated. Then, the curve fitting technique in Section 4.5.1 is applied to individual progress curves to quantify the stochastic nature of the project progress in terms of the distributions of corresponding parameters. In this example, 500 random progress curves were generated and the resulting means and standard deviations of the BetaS-curve parameters are shown in Table 4.7 along with frequency diagrams in Figure 4.14. In the calculation of the posterior distributions of the parameters, normal approximation of the simulation result is used. The solid lines in Figure 4.14 show that normal distributions approximate the frequency histograms reasonably well.

Table 4.6 Input data for random progress curve generation

Input parameters		Value
Number of activities		200
Number of effective precedence relations		433
Activity duration for planned progress	Mean	4 (weeks)
	Standard deviation	1 (week)
Activity duration for actual progress	Mean	4 (weeks)
	Standard deviation	1 (week)

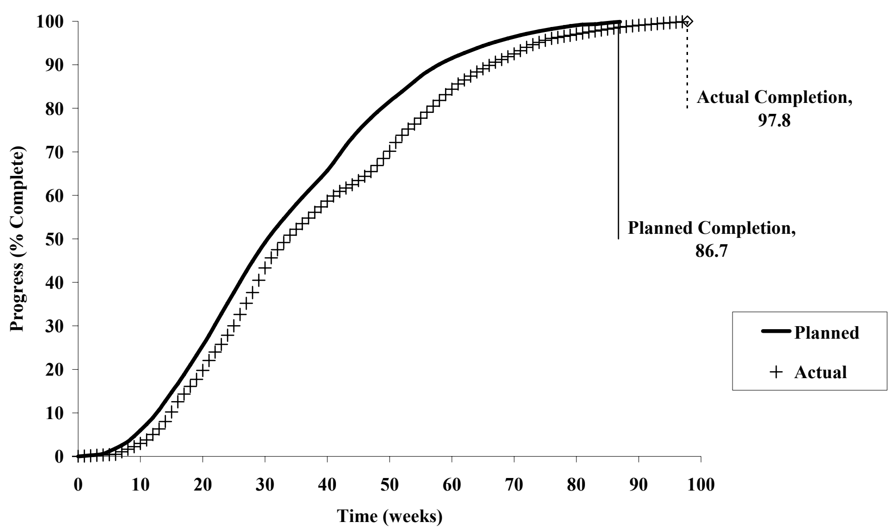
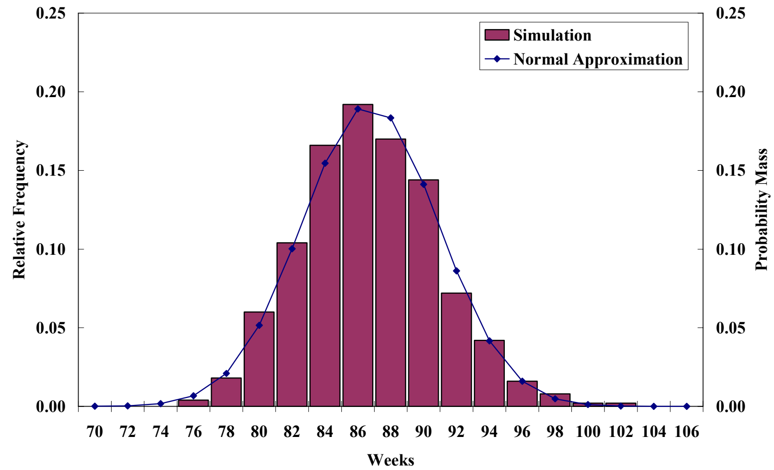


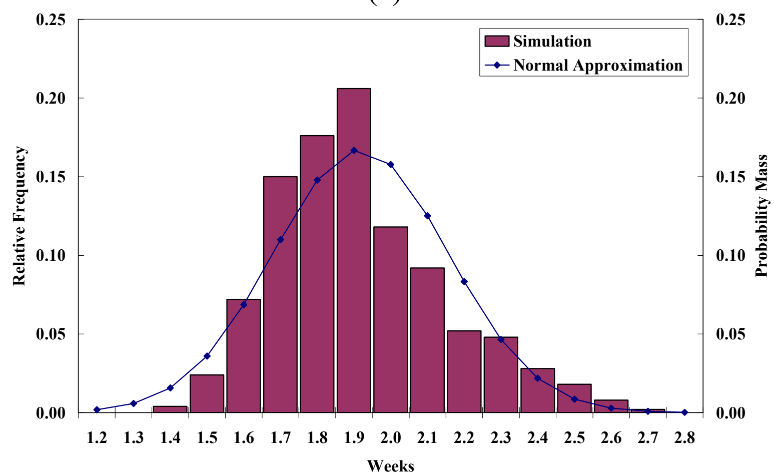
Figure 4.13 The planned progress and the "actual" progress for Example 4.2

Table 4.7 Statistical properties of the BetaS-curve parameters

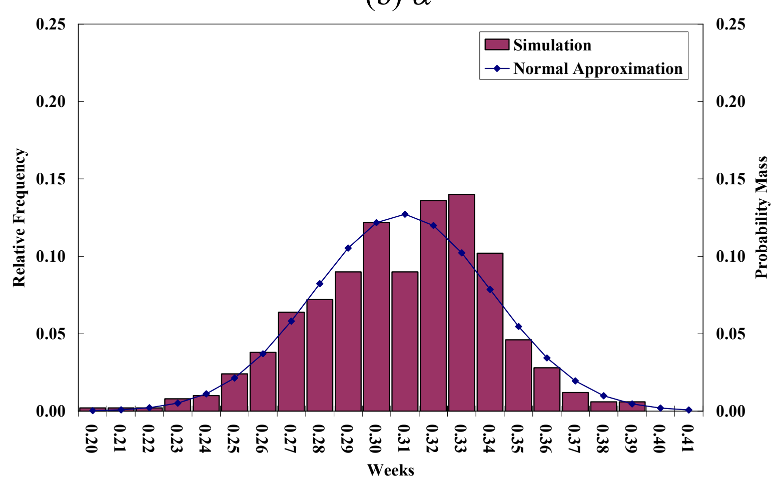
Parameters	Mean	Standard Deviation	Correlation Coefficients		
			T	α	m
T	86.7	4.15	1	-0.16	-0.18
α	1.91	0.23	(symmetric)	1	0.73
m	0.30	0.03	(symmetric)	(symmetric)	1



(a) T



(b) α



(c) m

Figure 4.14 Prior distributions of the BetaS-curve parameters

In this example, four cases of different prior distributions are applied to the prediction of the estimated duration at completion (EDAC). The four cases are summarized in Table 4.8. For the project duration parameter, *T*, two types of prior distributions are compared in the analysis: the informative prior distribution and the noninformative prior distribution. Figure 4.15 shows the two prior distributions along with the planned and the actual project durations. The informative prior is constructed from the network schedule simulation results in Table 4.8. On the contrary, the noninformative prior distribution is built as a uniform distribution over [0, 250]. Such noninformative priors can be used when a decision maker decides to make a prediction based only on actual performance data.

In the case of the shape parameters, α and m, two options are tested: fixed shape priors and probabilistic shape priors. Fixed priors for the shape parameters can be used when a project manager decides to make a prediction with a single, fixed progress curve template. In this example, the expected values from the simulation results are used as in Case B and Case D. Predictions made with fixed priors, however, ignore the uncertainty in the shape parameters under the assumption that the chosen values are the accurate shape parameters for the actual progress curve. When taking account of the uncertainty in the shape of actual progress curves, probabilistic priors should be used for the shape parameters.

Table 4.8 Four cases of prior distributions used in Example 4.2

Parameters	Informative prior for T		Noninformative prior for T	
	Case A	Case B	Case C	Case D
T	$N(86.7, 4.15^2)$	$N(86.7, 4.15^2)$	Uniform(0,250)	Uniform(0,250)
α	$N(1.92, 0.239^2)$	1.92	$N(1.92, 0.239^2)$	1.92
m	$N(0.30,0.0314^2)$	0.30	$N(0.30,0.0314^2)$	0.30

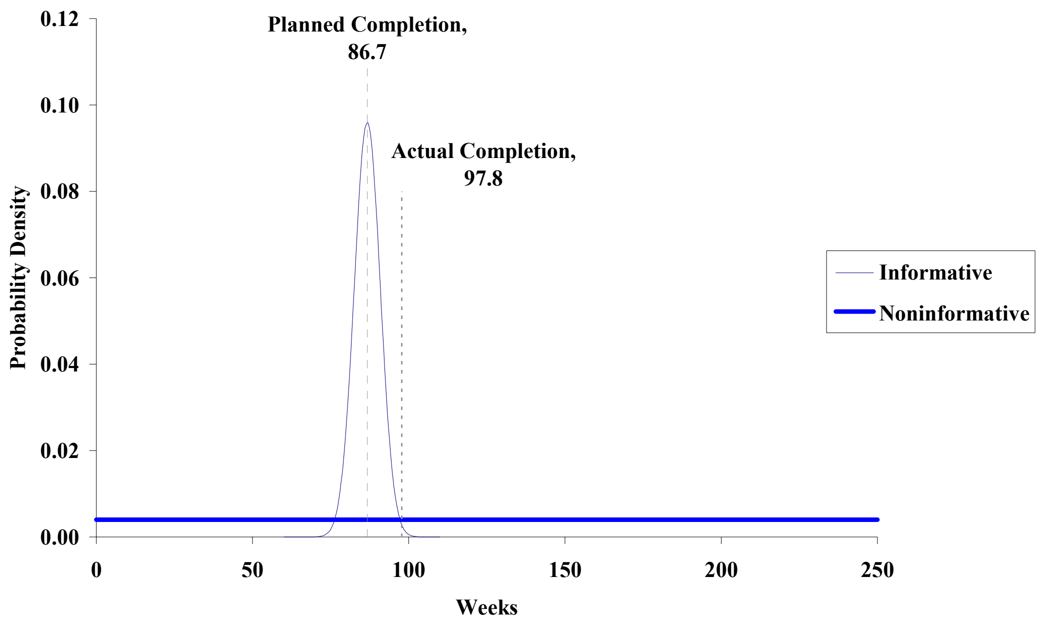
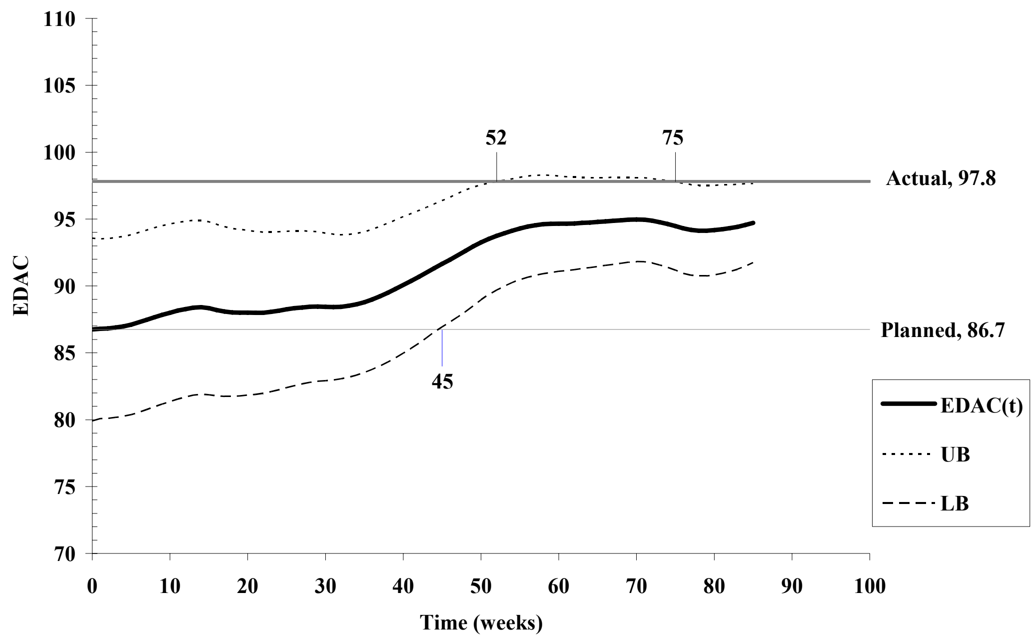


Figure 4.15 Two types of prior distributions for the project duration

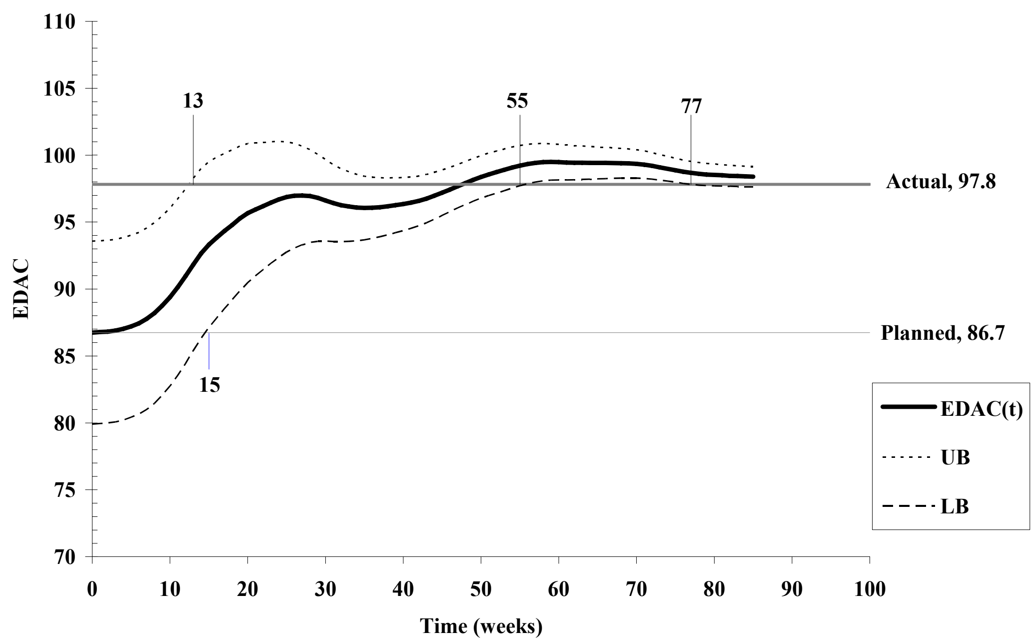
The BetaS-curve BAF model is repeatedly applied to the same progress data with the four prior cases and the time histories of the EDAC are shown in Figures 4.16 and 4.17. In the graphs, the thick solid line represents the mean of the posterior distribution of the EDAC over the forecasting time. The upper and lower bounds are determined at the 5 percent confidence level on each side. That is, the upper bound shows the duration that would not be exceeded with probability 0.95, while the lower bound is the duration that would be exceeded with probability 0.95. Therefore, an actual project duration has a 90% probability of lying between the bounds.

Forecasts made with the informative prior for *T* but different priors for the shape parameters are shown in Figure 4.16. The results reveal that the use of probabilistic priors for the shape parameters (Case A) results in a slow conversion of EDAC to the actual project duration. On the contrary, when fixed shape priors are used (Case B), the mean of the EDAC shifts quickly from the planned duration to the actual duration. The results indicate that the same actual data has stronger influence on the update of the prior distribution of the project duration when the true values of the other two parameters are assumed to be correctly known. However, it should be noted that using fixed shape priors does not guarantee fast convergence to the *correct* actual duration. Obviously, if biased estimates are used for the fixed parameters, predictions may approach the wrong conclusions.

The impact of using different shape priors is also found in the profiles of prediction intervals. Results in Figure 4.16 show that the width of prediction bounds narrows as more data are observed. However, the rate of narrowing is greater with Case B than Case A. This can be explained by the additional uncertainty included in the Case A priors in terms of probabilistic distributions for the shape parameters. In Case B, new information is used to update the distribution of project duration only. However, the same information is used in Case A to update all the three parameter distributions of a BetaS-curve model.



(a) Prior Case A



(b) Prior Case B

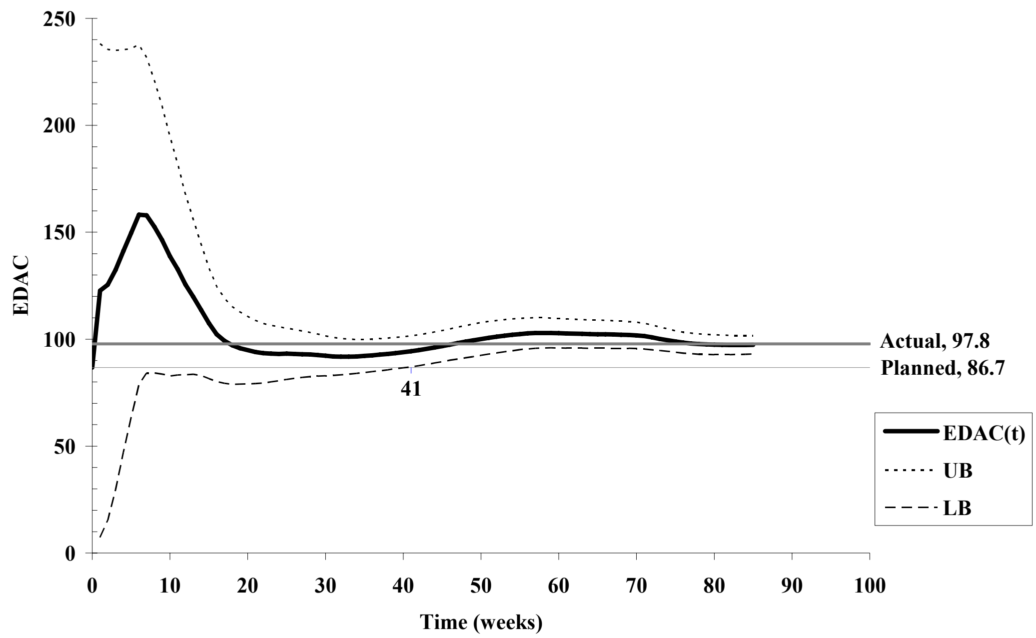
Figure 4.16 EDAC(*t*) with the informative prior distribution of project duration

The results when a noninformative prior distribution is used for the project duration are shown in Figure 4.17. The results show that the average EDAC quickly responds to the actual performance data even with probabilistic priors for the shape parameters (Case C). What makes the difference between Case A in Figure 4.16 and Case C in Figure 4.17 is the use of the noninformative prior instead of the distribution from the network simulation. As shown in Figure 4.15, for someone with a belief that the normal distribution constructed from a network schedule simulation represents the likely project duration, the actual duration in this example, belongs to a rather extreme case, which corresponds to a value that would not be exceeded with probability of 99%. As a result, it takes more data to adjust the prior belief. On the contrary, when a uniform distribution is used, all values in the range are assumed to be equally likely in the beginning and forecasts are made based on the new data being reported as the project proceeds. As a result, forecasts based on noninformative priors are more adaptive to the actual data than those based on informative priors.

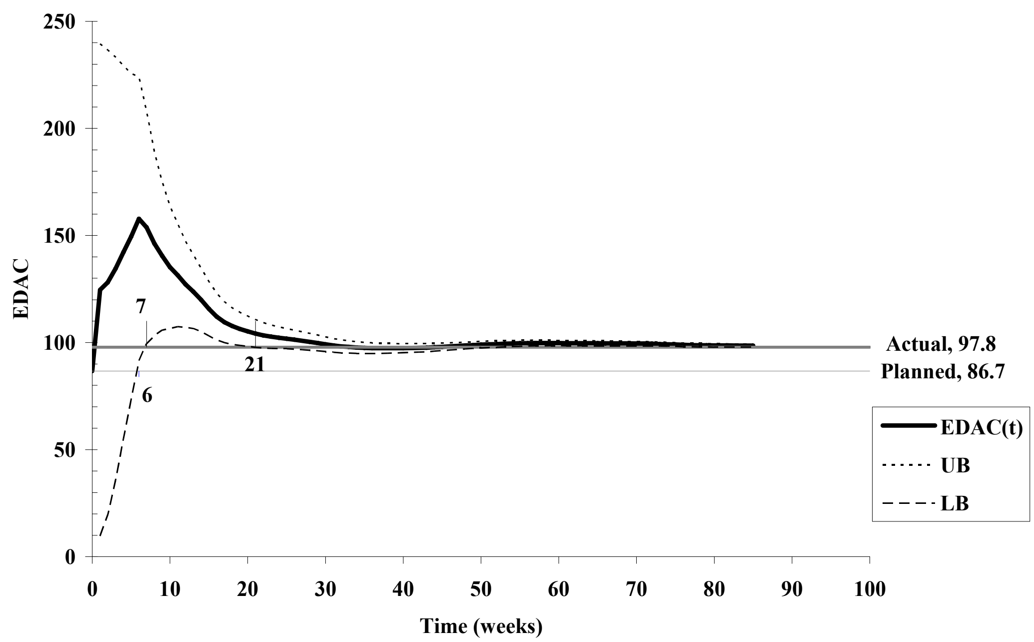
The overall forecasting performance of the four cases can be compared with two metrics: the length of time period during which the actual project duration lies between the prediction bounds and the time when the planned duration first falls outside of the prediction bounds. For example, forecasts with Case C priors include the actual project duration throughout the execution period while Case A starts to cover the actual duration at week 52. On the other hand, the earliest warning about schedule overrun can be obtained with Case D as early as sixth week.

The decision as to whether to use a noninformative prior or not should be made by the project manager. What the results in this example show is that, when the major concern of the project manager is some extreme cases, such as schedule or cost overruns, rather than normal outcomes, the noninformative prior can be useful because the results are determined entirely by the data.

With actual performance reports being generated after each reporting period, prior distributions of model parameters are repeatedly updated. Then a BetaS-curve with the mean of the posterior parameter distributions is extrapolated to the future. Figure 4.18 shows the updated project progress curves with the Case B prior set after 5, 10, 15, and 20 weeks. Note that the prediction progress curves in the graph are adjusted with the errors between the planned progress curve and its best-fit BetaS-curve. The graphs clearly show the adaptive nature of the method, which was observed in the example for the Kalman filter forecasting method (Section 3.4). At week 5, the prediction almost overlaps the plan. As more data are gathered and the discrepancies between the plan and the actual build up, however, predictions move away from the plan and closer to the actual data.

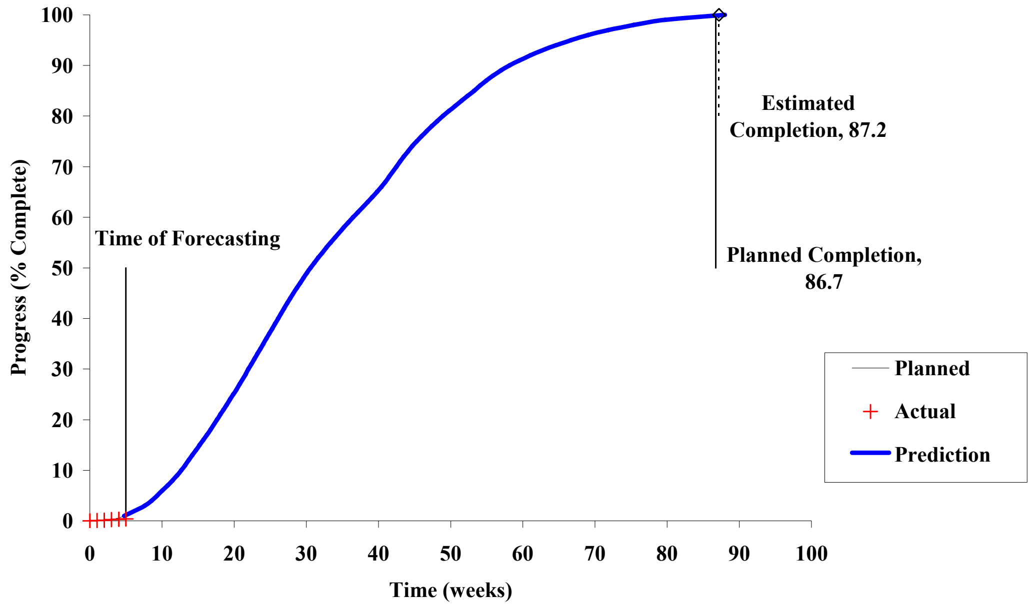


(a) Prior Case C

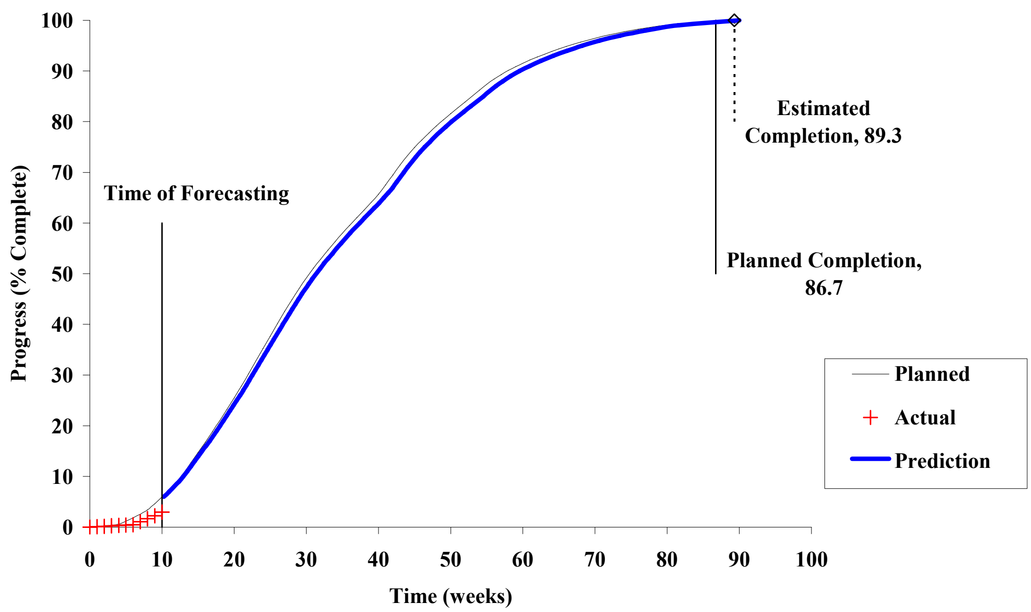


(b) Prior Case D

Figure 4.17 EDAC(t) with the noninformative prior distribution of project duration

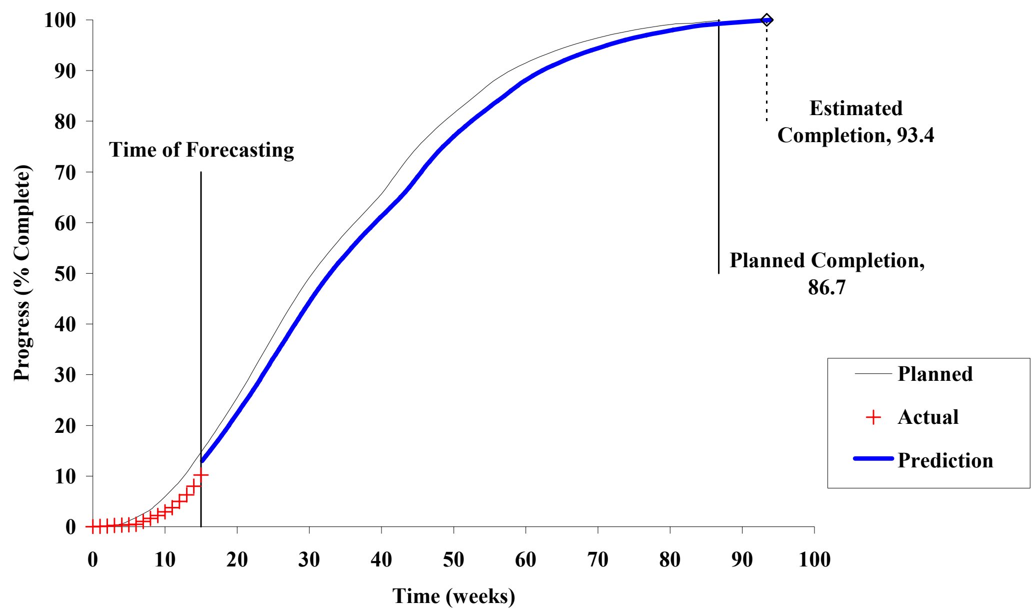


(a) Prediction at 5th week

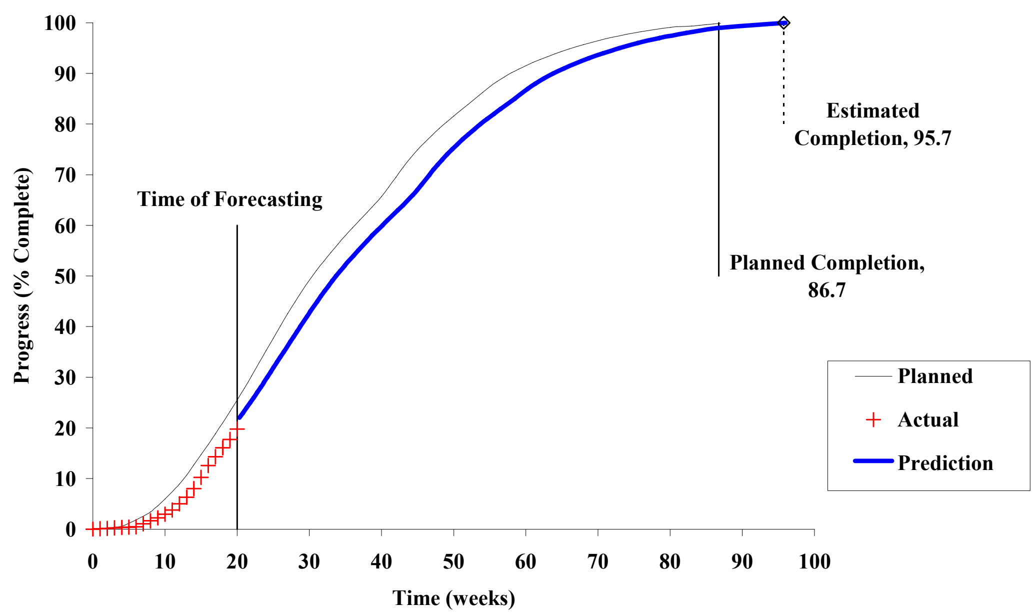


(b) Prediction at 10th week

Figure 4.18 Adaptive nature of the prediction by the BBAF



(c) Prediction at 15th week



(d) Prediction at 20th week

Figure 4.18 (Continued)

4.6 Predictive Power of a Progress Curve Template Based on Project Plans

A fundamental strategy of the BAF method is to identify some mathematical S-curve models for the progress curve template of a project using all relevant prior performance information and to update parameters of the selected models in light of actual performance data. Naturally, the reliability of the BAF method depends, in large part, on the degree to which an actual progress curve matches the progress curve template constructed before the inception of a project.

In this section, an empirical test is carried out to quantify the predictive power of a progress curve template of a project. The purpose of this test is to investigate the potential relation between the progress curve template based on prior performance information and the actual progress pattern and, if one exists, to assess the relation in a quantitative way.

Figure 4.19 shows the outline of the test. The test takes into account two types of uncertainty in project performance: the inter-project variation due to different schedule network structures and the activity level variation within individual projects. The inter-project variation is taken into account by using a large set of artificial projects. In total, 10,000 projects are generated by a random schedule network generation technique. Detailed information about the technique is provided in Chapter V. While all projects are assumed to consist of 200 activities, the number of precedence relations connecting those activities varies from project to project and Figure 4.20 shows the distribution of the resulting number of precedence relations in the test projects. On the other hand, the activity level variation within individual projects is simulated by modeling activity durations as random variables from a known normal distribution. Then, for each project, the progress curve template can be determined with the mean values of activity durations. On the other hand, an actual progress curve is randomly generated using a set of activity durations that are randomly drawn from the assumed probability distribution. It should be noted that the progress curve template of a project is determined by the characteristics of its network structure, while the actual progress curve is subject to the combined effect of the random network structures and the random activity durations.

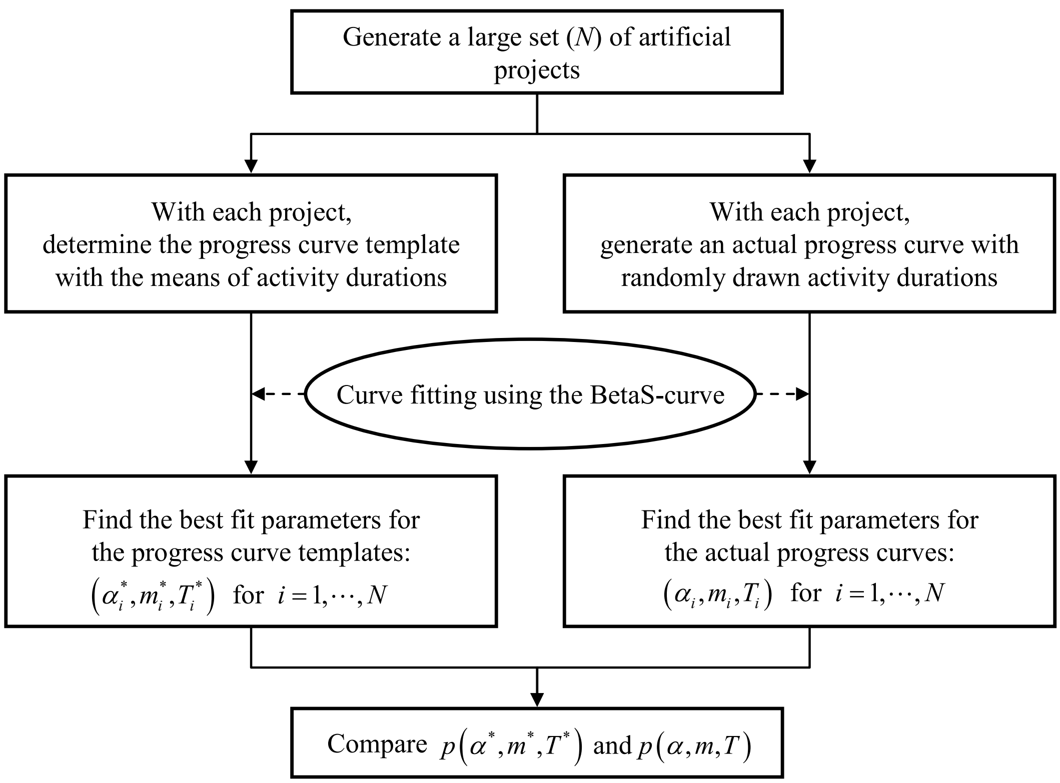


Figure 4.19 Outline of the test of predictive power of a progress curve template based on prior performance information

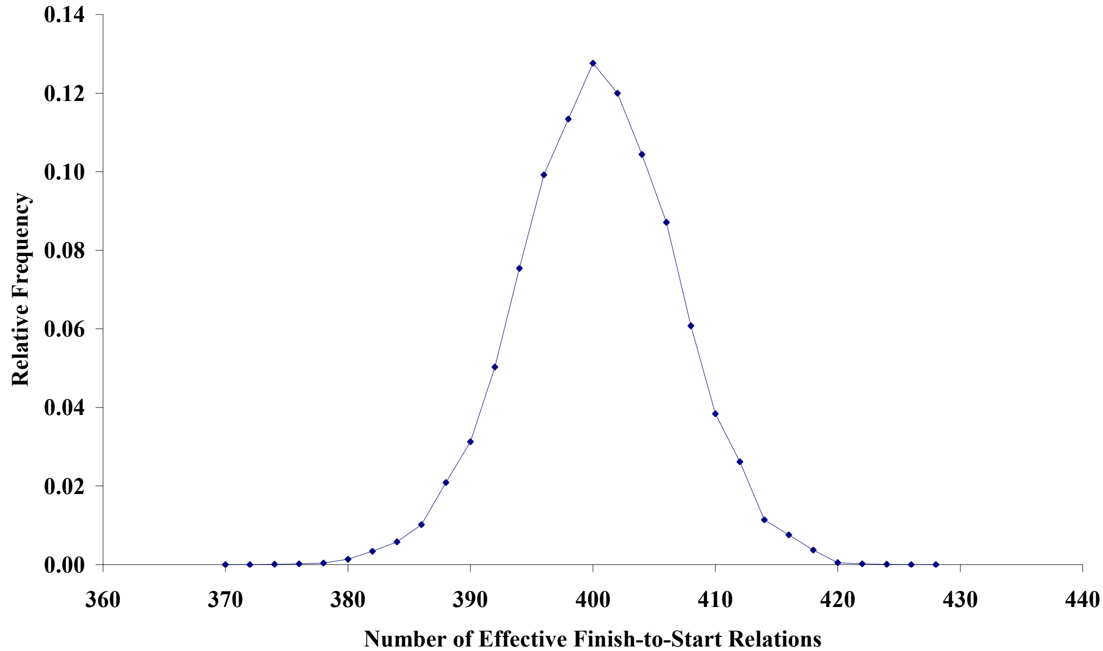
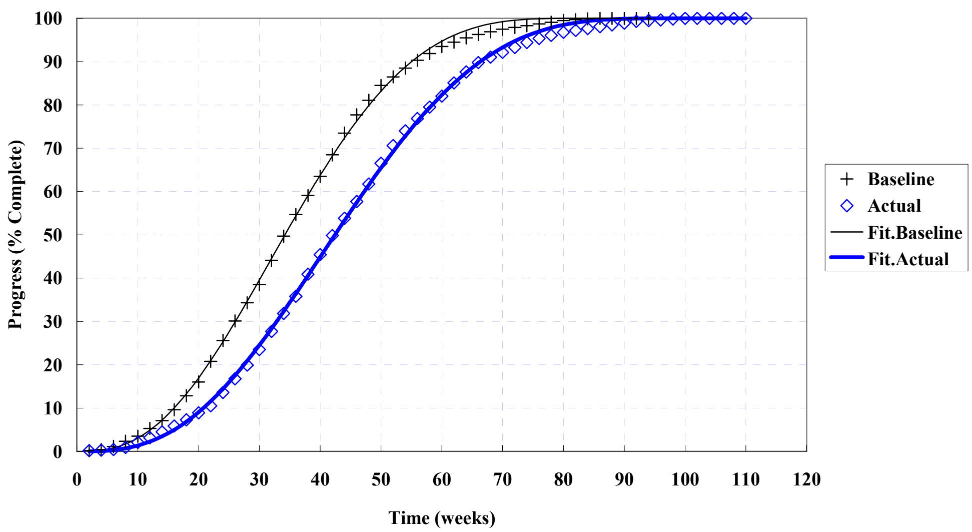
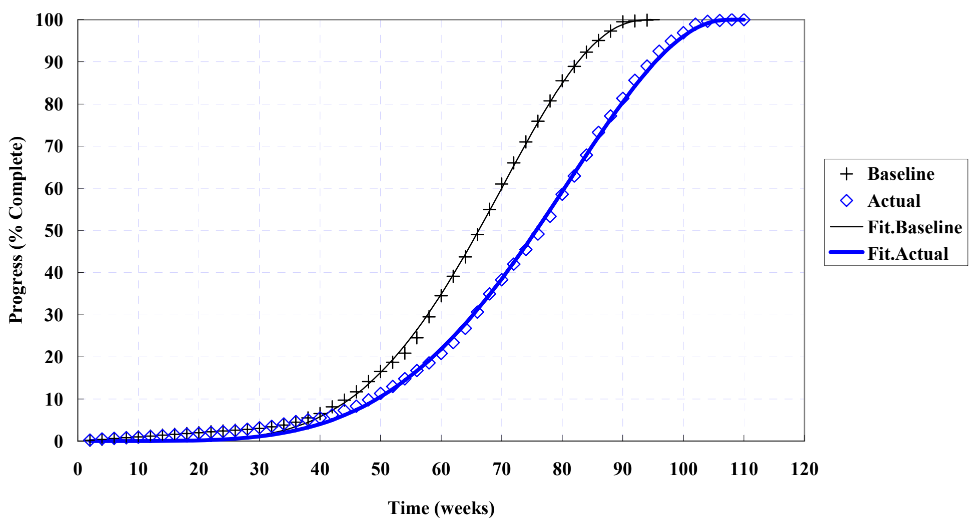


Figure 4.20 Distribution of the number of precedence relations of 10,000 test projects

Once a pair consisting of a progress curve template and an actual progress curve is determined, their stochastic nature can be measured with the curve fitting technique using the BetaS-curve model. In the test, two different execution options – the early start progress and the late start progress – are considered as shown in Figure 4.21.



(a) Early start progress



(b) Late start progress

Figure 4.21 Examples of the curve fitting technique to different execution options

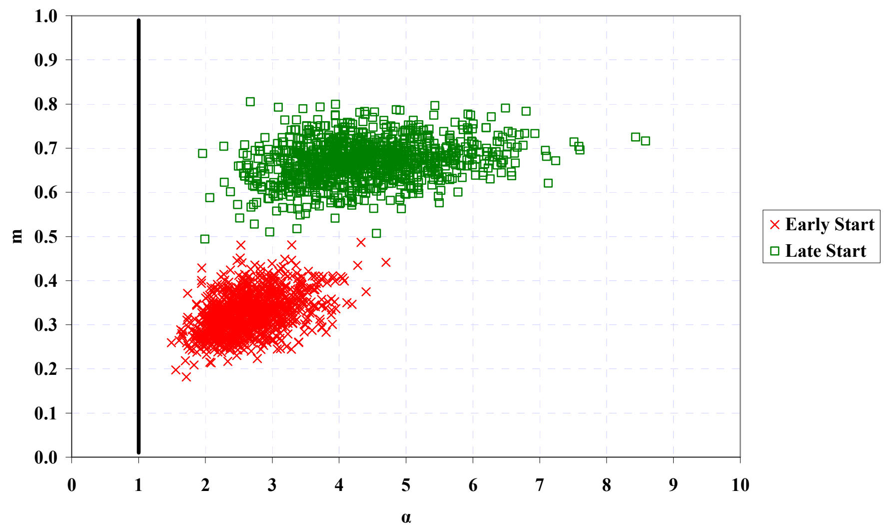
The statistical properties of the best fit parameters for the 10,000 projects are summarized in Table 4.9 and graphically shown in Figure 4.22. From the results, some conclusions can be drawn.

- Depending on the execution policy – the early start policy and the late start policy – the mode parameter m has different ranges. This result indicates that different progress patterns due to different resource loading types, such as front-end loading and back-end loading, can be characterized by different range estimates of parameter m.
- Regardless of the execution policy, strong correlations between the best fit shape parameters for the progress curve template and those for the actual progress curves are observed. All correlation coefficients between α and m are about 0.9 regardless of the execution policy. Since correlation coefficients represent the degree to which an actual progress curve matches its corresponding progress curve template, these results strongly support the rationale of using the progress curve template based on detailed project plans as a predictive model. Such a strong match of progress patterns is attributable to the fact that both progress curves – the baseline and the actual – are generated from the same activity network. In other words, the results indicate that, when the activities in a project are executed according to the scheduled order, the actual progress curve, in spite of the variations at the activity level, tends to follow a progress curve similar to the progress curve template based on the schedule network and probabilistic estimates of activity durations. The assumption about the execution order of activities is realistic because a network schedule is often

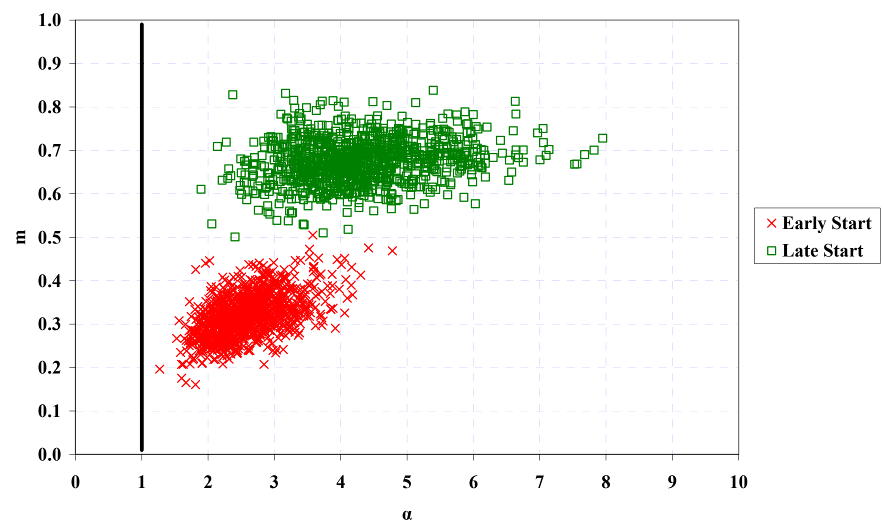
a part of legal contracts between owners and contractors. Once a project schedule is officially accepted and approved by the project management team, it is set as the schedule baseline and should be revised *only* in response to approved changes.

Table 4.9 Statistical properties of the best fit parameters for the progress curve templates and those for the actual progress curves

Execution policy	Model parameters	Progress curve template		Actual progress curve		Correlation coefficient	
		mean	std.dev.	mean	std.dev.		
Early Start	α	2.65	0.48	2.59	0.50	(α^*,α)	0.89
	m	0.33	0.04	0.32	0.05	(m^*,m)	0.88
	T	95.10	9.77	96	10.21	(T^*,T)	0.91
Late Start	α	4.34	0.92	4.30	0.93	(α^*,α)	0.90
	m	0.67	0.05	0.68	0.05	(m^*,m)	0.89
	T	95.23	9.72	97	10.13	(T^*,T)	0.91



(a) Baseline progress



(b) Actual progress

Figure 4.22 Scatter diagrams of the best fit parameters for the progress curve templates and the actual progress curves

4.7 Chapter Summary

A probabilistic framework for forecasting project progress and the probability distribution on project duration at completion has been developed based on Bayesian inference and S-curve models. The Bayesian adaptive forecasting method is a regression model that fits S-curves to cumulative progress curves of a project and updates the parameter estimates of the S-curves using a Bayesian inference approach.

This chapter started with a review of S-curve models in previous research. In Section 4.2, five useful S-curve models were introduced and the BetaS-curve model was proposed. Then, the general framework of the Bayesian adaptive forecasting was derived in Section 4.3. The framework can be characterized by three features of being probabilistic, integrative, and adaptive. The BAF method makes use of all relevant performance information available from standard construction management practices and theories. Information used in the BAF method can be grouped into two categories: the prior performance information and the actual performance data. The prior performance information includes all preproject (or preconstruction) information in various forms such as project plans, historical data, and subjective experience. From the prior performance information, the probability distribution of project duration and the progress curve template are constructed.

A fundamental idea of the BAF method is that every project proceeds following a characteristic progress pattern and that prior performance information can be used to identify the underlying pattern in advance before actual performance data reveal it. Because of the characteristic of projects of having a definite beginning, the data samples of actual performance early in a project is too small to make a reliable statistical inference about the overall progress pattern of the project. For example, when a project starts slowly, accelerates, and then tails off near closing – in fact, many projects actually follow this pattern – actual performance data at the outset do not reveal sufficient information about typical turning points in the future progress, such as when the project reaches its peak work rate or when the tailing off starts. In the BAF method, a progress curve template based on prior performance information is used to supplement the shortage of actual performance data by providing a holistic view of the overall progress pattern of a project. As more actual data are available, the contribution of the progress template becomes smaller. Mathematical S-curve models are used to formulate this fundamental concept within the general framework of Bayesian inference by quantifying the progress template of a project and adjusting the probability distributions of model parameters in light of new actual performance reports.

Depending on the types of S-curve models used in forecasting, two methods have been presented: the Multi-model BAF (Section 4.4) and the BetaS-curve BAF (Section 4.5). Both approaches are based on the general framework of BAF and proceed through basic three steps: (1) generating prior distributions of model parameters; (2) updating model parameters, and (3) forecasting. The Multi-model BAF however needs two additional steps because it starts with a group of fixed-shape S-curve models. Once the progress curve template for a project is developed, the goodness of fit of each model to the template is evaluated to select the models with reasonable fit. Then the general BAF framework is separately applied with each of the selected models. At the final stage, forecasts from different S-curve models are combined via the Bayesian model averaging technique. On the other hand, the BetaS-curve method uses the BetaS-curve model as a single mathematical function to approximate a wide range of progress curves. With two shape parameters – the nonlinearity parameter α and the mode parameter m –, the BeatS-curve model has a greater capability of representing extremes and variability in project progress patterns than the fixed-shape models used in the Multi-model BAF method.

In Section 4.6, the BetaS-curve model and the curve fitting technique used in the BetaS-curve BAF method were applied to an empirical test of the predictive power of a progress curve template based on project plans. With a large set of artificial projects, it has been shown that there exist strong correlations between the shape parameters of the progress curve templates and the actual progress curves when a project is executed according to the planned schedule network.

In addition to the probabilistic nature, another important merit of the Bayesian adaptive forecasting method is the use of prior performance information in conjunction with actual performance data. Prior performance information from detailed project plans, historical data, and subjective judgments of project managers is effectively combined with actual performance data to make reliable forecasts of future performance. As demonstrated in the two examples, proper use of prior performance information may lead to significant improvements in the quality of predictions early in the project. Conventional forecasting models based on CPM and EVM rely dominantly on either original project plans or actual performance data being observed during the execution. For example, a typical CPM makes little use of the actual performance information because only up-to-date performance, whether it is good or bad, is considered and the original estimates of remaining jobs are not adjusted according to the past performance. On the contrary, the fundamental concept of EVM forecasting formulas is to linearly extrapolate the efficiency of past performance to the future performance, regardless of the reliability or amount of past data available at the time of forecasting.

Another merit of the BAF methods is easy implementation. The BAF method can be incorporated into current construction management and control systems without additional burden of system changes or new data acquisition. Typical input data required by the BAFM are available in standard construction practices and standards. In addition, Monte Carlo integration is effectively used to perform the Bayesian updating calculation without relying on more sophisticated, time-consuming techniques such as importance sampling and Markov Chain Monte Carlo method.

CHAPTER V

PARAMETRIC STUDIES

5.1 Design of the Parametric Studies

5.1.1 Outline

Every project is unique in its objectives, the plan to achieve the objectives, and the actual progress guided by the plan. Naturally, the forecasting performance of a method varies from one project to another according to the specific situations of the project of interest. Even with the same performance information available, some projects may be harder to predict for some methods than for others due to the characteristics of individual methods. Therefore, it is challenging by itself to evaluate the forecasting performances of various methods and to compare them in an objective way.

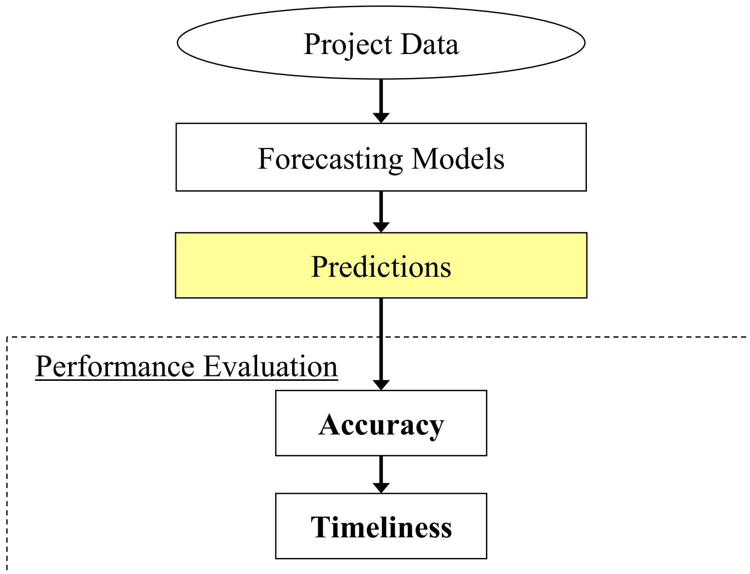
This section includes a series of parametric studies about the forecasting performance of the two methods – the Kalman filter forecasting method (KFFM) and the Bayesian adaptive forecasting method (BAFM) – which were presented in previous chapters. In addition, two conventional methods – the earned value method (EVM) and the critical path method (CPM) – are compared with the new methods. The purpose of the parametric studies is to evaluate the forecasting performance of the KFFM and BAFM and compare them with the conventional methods in a statistically meaningful way. The results from the study will serve as a guideline for potential users and help them build a better understanding of the new methods as well as the methods they have been relying on.

In spite of the crucial role of forecasting in successful project management, previous research about evaluating performances of different forecasting methods is very limited (Teicholz 1993; Vanhoucke and Vandevoorde 2006). Teicholz (1993) compares three forecasting methods for final cost and budget using the data from 121 real projects. Vanhoucke and Vandevoorde (2006) compares three earned value forecasting formulas for project completion date using a large set of simulation data. In spite of the differences in the forecasting targets and the data used, previous approaches commonly measure the accuracy of forecasts in terms of the average of the errors between the prediction and the actual over the entire project execution. Furthermore, the timeliness of a method is evaluated based on the average of the same errors over a specific period of time. For example, Teicholz (1993) defines timeliness as "the forecast accuracy during the first 50% of the project (measured by budget percent complete)" and Vanhoucke and Vandevoorde (2006) compares timeliness with the change of accuracy along the completion stage of projects.

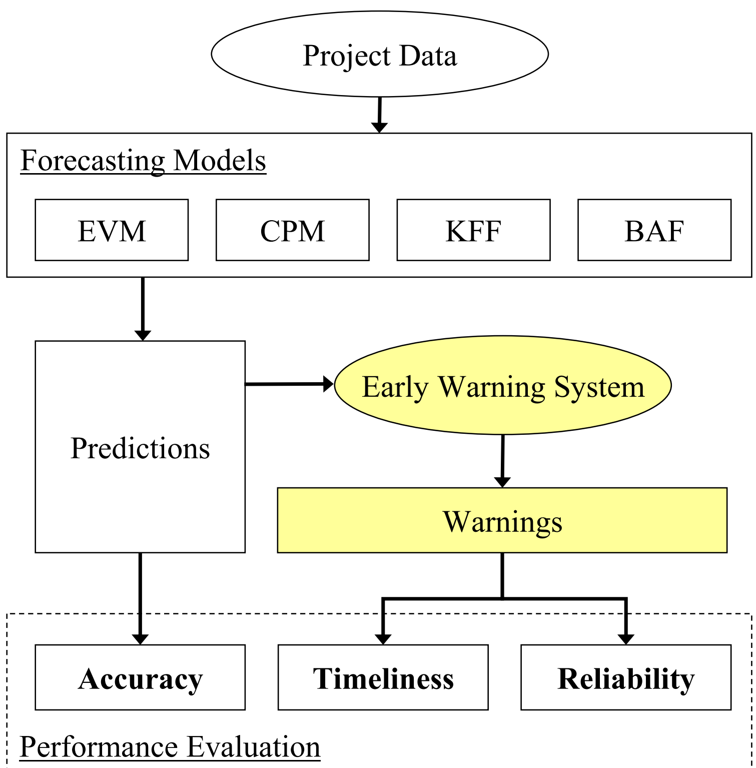
However, these approaches have two limitations. First, the accuracy measure in terms of the average error of prediction over the entire execution has little practical meaning for project managers. It would be more useful if accuracy of various methods is compared at specific points of time. For example, forecasting performance evaluation results must be able to answer a question such as "After 20 percent of project duration has passed, what is the expected accuracy of the predicted project duration from this method?" Another limitation in the previous approach is that the timeliness measure of a forecasting method is based on the belief that more accurate predictions lead to earlier warnings. Although, at least intuitively, the assumption appears reasonable, the timeliness of a forecasting method should be evaluated in terms of the timing of warning which is determined in the larger context of a decision making system.

In this dissertation, a new evaluation framework for project performance forecasting methods is proposed. The primary goal of the framework is to obtain *statistically* meaningful and *practically* informative results of various forecasting methods in terms of the accuracy of prediction, the timeliness of warning, and the reliability of warning. Figure 5.1 compares the new framework to the conventional one. The new approach is based on the common performance control practices used in the construction industry, artificial data from extensive simulation, and new evaluation criteria developed to take account of the characteristics of project performance forecasting. The timeliness and reliability of a method are evaluated based on the warnings from an early warning system, instead of average accuracy during a period of time.

This chapter is organized as follows. In Section 5.1, fundamental issues related to the parametric studies are reviewed. The limitations of using real projects are discussed and artificial project data generation is presented as an alternative method. A brief review of the conventional forecasting methods is also presented. Section 5.2 addresses the random progress generation technique used to develop a large set of diverse artificial project data. In Section 5.3, typical early warning systems in project management are reviewed and it is shown that in spite of individual characteristics they share common core elements. Section 5.4 reviews the evaluation criteria used in the



(a) Conventional approach



(b) New approach

Figure 5.1 Frameworks for evaluating project performance forecasting methods

following performance evaluation. In this research, the two most commonly used criteria (the mean percentage error and the mean absolute percentage error) in the forecasting literature are modified and several original criteria are proposed based on the characteristics of project performance forecasting and general probabilistic performance evaluation. Results from the tests of the three research hypotheses are summarized in Section 5.5, Section 5.6, and Section 5.7, respectively. Overall comparison between the four forecasting methods is presented in Section 5.8. Finally, the results are summarized in Section 5.9.

5.1.2 Data Collection

To achieve the goal of obtaining statistically meaningful results about the performance of various forecasting methods, a large set of project performance data is required. These data should be independent of any of the forecasting methods under comparison. Showing some real cases in which a forecasting method makes better forecasts than other methods does not prove the superiority of the method because it is almost always possible to find some other cases in which another method outperforms the others. A similar issue has been addressed by Hazelrigg (2003) in his study about validating engineering design alternative selection methods. The author argues that merely testing a decision method in various contexts does not validate the decision method because the results can be made by chance and the only viable method is to rely on *mathematical validation of the procedure*. Furthermore, in order to verify that a decision recommended by a decision method is indeed the best choice, all possible choices must be tested independently so that the results of each choice can be compared (Hazelrigg 2003).

The same argument by Hazelrigg (2003) holds true for evaluating and comparing different forecasting methods. In that sense, the use of real project data has several shortcomings because real data from past projects are limited in amount and, more importantly, affected by unknown managerial activities during the execution. First, for any historical project data to be used in performance evaluation, a complete set of the planned and the actual performance data should be available. Furthermore, to get a statistically meaningful outcome, a large enough number of complete project data is required, which can hardly be achieved in the real world. Second, given the actual progress data, any volitional interference – whether good or bad – by the project team or management should be identified and their effects on the actual progress should be quantified. Probability distributions of future performance predictions, as forecasts from the current state, are conditional on *no* control actions being taken in the future.

Another challenge in evaluating project performance forecasting methods arises from the fact that it is practically impossible to conduct empirical tests using real projects. First of all, a prediction of a project overrun is a self-negating prediction. That is, a warning about a future problem should be used to reduce its probability of coming true, resulting in negating the prediction. Furthermore, in order to compare forecasting methods with a real project, the same project should be executed independently from any of the forecasting methods under consideration or each forecasting method should be applied independently. However, either situation is not possible because forecasting is an essential function of management and the same project can not be repeated independently with each forecasting method.

In this research, a large set of artificial project data is used. Testing new ideas and methods under artificial or simulated environments is common and a generally accepted approach in the engineering and scientific research community. For example, new earthquake design codes for buildings or bridges may be tested on a huge shaking table that generates artificial earthquake vibrations. The primary idea of using artificial data is to evaluate or test the performance of forecasting models against a wider range of plausible situations instead of relatively small sets of real data (Vanhoucke and Vandevoorde 2006).

Artificial projects are generated with a common project simulation approach based on network schedule and their actual performance data are obtained through random executions that are independent of any forecasting methods. It should be noted that the fundamental concept of this approach is to evaluate the forecasting performance of various methods with a large set of diverse project data which are generated artificially under the control of critical factors such as the complexity of the activity network, the accuracy of preconstruction information (i.e., a probabilistic estimate of project duration and a baseline progress curve), and execution scenarios (i.e., ahead of schedule and behind schedule). Recently, Vanhoucke and Vandevoorde (2006) conducted a study comparing different EVM schedule forecasting formulas. In the study, they tried to draw general conclusions that hold under various situations, rather than to show specific forecast results based on a limited number of real project data. To achieve this goal, they generated 3100 random activity networks for projects with 30 activities.

5.1.3 Forecasting Methods

In this chapter, four forecasting methods are used separately or as a combination. They are the Kalman filter forecasting method, the Bayesian adaptive forecasting method, the EVM, and the CPM. Since the EVM and the CPM were reviewed in Chapter II and the KFFM and the BAFM have been discussed thoroughly in previous chapters, this section focuses on the assumptions made for CPM and EVM.

Critical path method

Critical path method has been reviewed in Chapter II. In this research, the project duration at completion is estimated based on two assumptions. First, it is assumed that all activities in a network schedule are successfully managed to start at the earliest possible time according to the early start policy. Second, it is assumed that actual completion dates of on-going activities that have been completed more than 50 percent as of the time of forecasting can be correctly estimated. For the other on-going activities with completion rate less than 50%, the original estimates are used for prediction.

Earned value method

Earned value method has been reviewed in Chapter II in detail. It should be noted that the same performance metric based on the earned schedule (ES) concept is used as the primary performance state variable in the Kalman filter method. Therefore, any differences in the predicted project durations at completion by the EVM and the KFFM are strictly the results of methodological differences rather than any differences in input information.

5.1.4 Selection of Decision Parameters

In the generation of the artificial project data, specific values of design parameters are selected according to common management practices in the construction industry and the results from previous research. This section summarizes major parameters used in the data generation, including the number of activities in a schedule network, the number of precedence relations in the network, the activity durations and costs, and the time horizon of forecasting.

The number of activities in individual project networks

The number of activities is an important element in project scheduling. Project scheduling is a process of identifying all activities required to complete the project and defining the interface between them to ensure that the activities are managed more effectively and in the right order. Typical project scheduling starts with decomposing the whole project into smaller components – activities or work packages – so that they can be estimated, planned, and managed more efficiently. In general, the size of a project is a primary factor that determines the number of activities in the project schedule. However, the appropriate number of activities in a schedule network depends on the level of scheduling detail which is determined by the planner according to many factors, such as the physical nature of a project, the competency of the organization, and the procurement plan specific to the project. Therefore, there are no systematic rules that determine the optimal number of activities for a project schedule.

In the 2002 Unified Facilities Guide Specification (UFGS) for Network Analysis Systems, appropriate ranges of the number of activities are recommended for construction projects with different sizes (Nassar and Hegab 2006). Most artificial projects used in the parametric studies are generated to have 200 activities. According to the UFGS, this number is recommended for projects in cost range of 1.0 to 5.0 million dollars.

The number of precedence relations in the network

Given the number of activities in a schedule network, the number of precedence relations among the activities is the most important factor that influences the complexity of the network structure. To take into account of potential effects of network complexity on the forecasting performance, the number of precedence relations in artificial projects is chosen to represent a reasonable range of complexity levels in construction projects.

A precedence relation is based on the interdependence of activities and shows the order or sequence in which the activities are planned to be executed. In a typical precedence network schedule, four basic relationships between activities are determined by how the completion (or start) of an activity restricts or restrains the start (or completion) of following activities. They are the finish-to-start (FS), finish-to-finish

(FF), start-to-start (SS), and start-to-finish (SF) relations. In the generation of artificial projects used in the hypothesis tests, only the FS relationship is used because it is most common in real project activities.

The number of precedence relations in a network is directly related to the complexity of a schedule network. Many different measures are proposed in the literature to characterize the complexity in activity networks. A well-known measure for activity-on-node networks is the coefficient of network complexity (CNC), which is defined as the number of arcs over the number of activities (Davies 1973; De Reyck and Herroelen 1996; Elmaghraby and Herroelen 1980; Kaimann 1974). Demeulemeester et al. (2003) argue that the CNC does not correctly measure the hardness of a project scheduling problem.

The complexity measure used in this research is based on the order of strength (OS) proposed by Mastor (1970). The OS is defined as the number of precedence relations divided by the maximum number of precedence relations, which is $N(N\text{-}1)/2$ for a network with N activities (Demeulemeester et al. 2003). Recently, Nassar and Hegab (2006) proposed a modified complexity measure based on the OS. The measure evaluates the complexity of a schedule network based on the range of possible non-redundant precedence relations. For a schedule network with N activities, the minimum number of non-redundant relations is (N -1) and the maximum number of non-redundant relations in a network with $N \geq 6$ is given by (Kolisch et al. 1995)

$$\text{Maximum number of precedence relations} = \begin{cases} \dfrac{N^2-4}{4} & \text{if } n \text{ is even} \\ \dfrac{N^2-5}{4} & \text{if } n \text{ is odd} \end{cases} \tag{5.1}$$

Then the network complexity (C_N) of a project with *N* activities connected with *x* non-redundant relations is defined as (Nassar and Hegab 2006)

$$C_N(x) = 100\frac{\log(x)-\log(\min)}{\log(\max)-\log(\min)} \tag{5.2}$$

where the min and the max are (*N*-1) and the maximum number of precedence relations calculated with Equation (5.1). In use of the C_N, the authors recommended a schedule network with $C_N < 30$ as an acceptable level of complexity for scheduling.

The durations and costs of activities

It is assumed that all activities in individual artificial projects are homogeneous in terms of the probabilistic properties of duration and cost. The durations and costs of activities are assumed to be normally distributed. Specific values of the mean and variance of activities are chosen in such a way that the overall project duration spans over the range from 70 to 90 weeks to ensure proper length of the time horizon of forecasting.

The time horizon of forecasting

The forecasting of project duration at completion is carried out up to the point of planned project duration. Forecasting is no longer meaningful when the planned completion date has already passed. In a practical point of view, when a project passes an original completion date without rescheduling, prediction of the actual completion date should be based on thorough investigation over the remaining activities instead of past performance data.

5.2 Random Progress Generation

5.2.1 Random Network Generation

An alternative to the data observed from real projects is artificial data generated by mathematical models. To meet the purpose of the research, an original random progress generating algorithm was developed and programmed in Visual Basic for Applications in Excel® 2003, which is named RanPRO (the random project progress generator). RanPRO is a program that generates a set of project progress curves – the planned progress and the actual progress – using a random network generation technique and the standard critical path analysis method (Figure 5.2). Although the need for random schedule network generation has been recognized and investigated by many researchers (Agrawal et al. 1996; Demeulemeester et al. 1993; Demeulemeester et al. 2003), the method presented in this research is an improved one because it takes the effect of redundant relations into account.

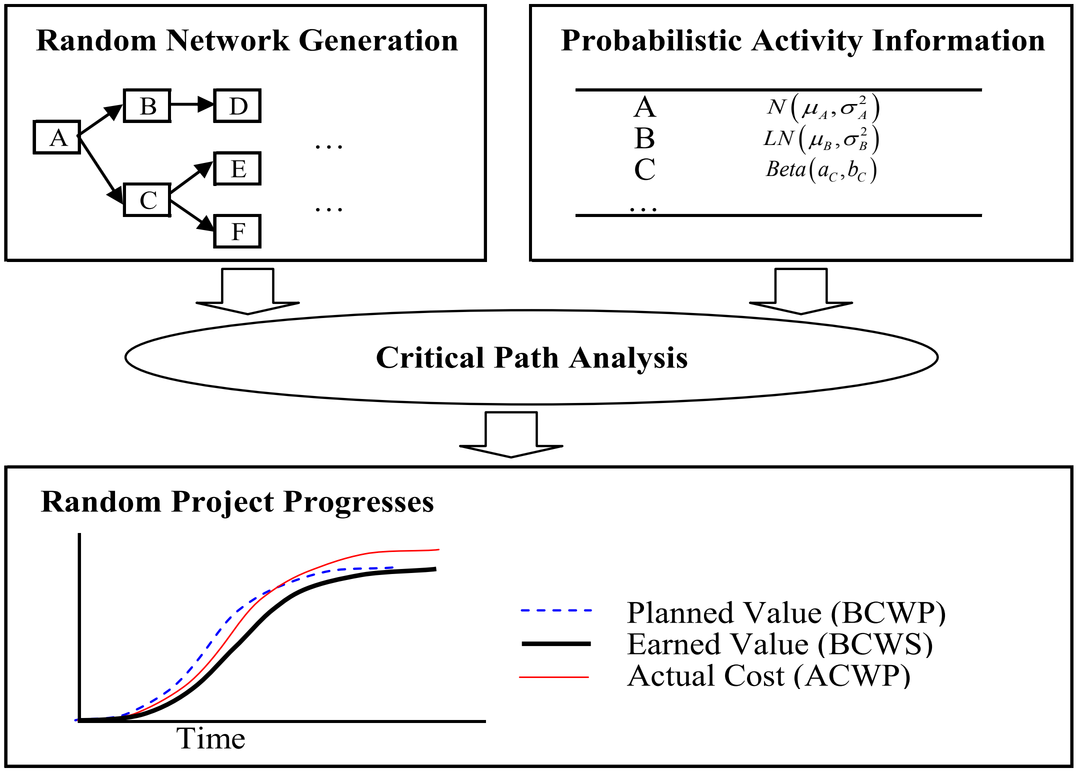


Figure 5.2 Random project progress generation

Demeulemeester et. al. (1993) presented a random activity network generation method that represents precedence relations between activities in a matrix form, which is named the precedence matrix. In the method, a schedule network should satisfy three schedule constraints.

Schedule Constraints

1. An arc always leads from a small number activity to a larger one.
2. There is one start and one end activity.
3. Each internal activity has at least one immediate preceding activity and at least one immediate following activity.

Under the schedule constraints, a schedule network with N-activities can have a maximum of $N(N\text{-}1)/2$ precedence relations. All the feasible relations can be represented in an N-by-N matrix, as shown in Figure 5.3 (a) for the case with ten activities. The number '1' in an element (i, j) of a precedence matrix indicates that Activity-i is linked to Activity-j by the Finish-to-Start relation. In other words, one needs to finish Activity-i to start Activity-j.

	1	2	3	4	5	6	7	8	9	10
1		1	1	1	1	1	1	1	1	1
2			1	1	1	1	1	1	1	1
3				1	1	1	1	1	1	1
4					1	1	1	1	1	1
5						1	1	1	1	1
6							1	1	1	1
7								1	1	1
8									1	1
9										1
10										

(a) Complete relations

	1	2	3	4	5	6	7	8	9	10
1		1	0	0	1	0	0	1	0	0
2			1	0	1	0	0	1	0	1
3				1	0	1	0	0	1	0
4					1	0	0	1	0	0
5						0	0	1	1	0
6							1	1	0	0
7								0	1	1
8									1	0
9										1
10										

(b) A random network with redundancy

Figure 5.3 A random network generated by the deletion method

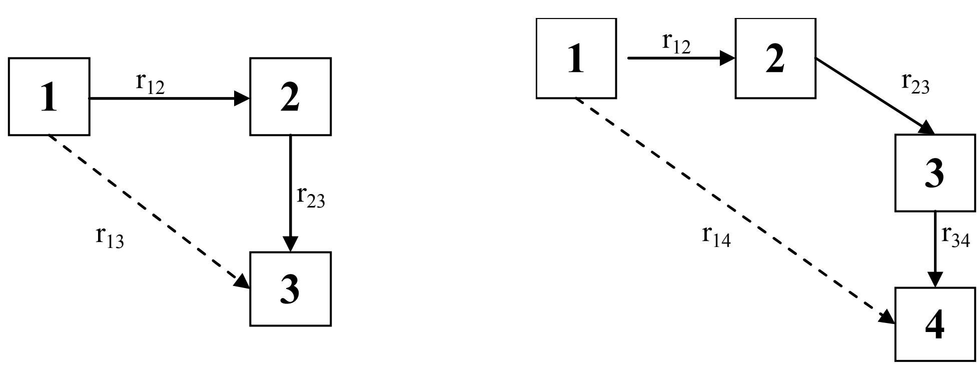
The deletion method by Demeulemeester et. al. (1993) starts with the complete precedence matrix and proceeds by randomly selecting a predetermined number of relations to eliminate. Before a relation has been eliminated from the precedence matrix, it must be checked whether eliminating the selected relation does violate any schedule constraints. The precedence matrix in Figure 5.3 (b) is an example that was generated to have 20 active relations.

5.2.2 Redundancy Elimination

Redundant relations

The *redundant relation* in an Activity-on-Node (AoN) schedule network is defined as a precedence relation that is dominated by other precedence relations and, as a result, elimination of it has no impact on the schedule performance of the network. Figure 5.4 shows two cases of redundancy in AoN schedule networks: the direct redundancy and the latent redundancy. The direct redundancy occurs when a relation between two activities is dominated by another activity that goes between the two with direct relations with both of them. The latent redundancy occurs when a relation between two activities is dominated by a combined effect of relations between two or more activities.

The deletion method does not account for the redundant links between activities. As a result, it may provide an unrealistic measure of network complexity. Furthermore, the presence of redundant links in a schedule network can lead to additional computational burden in the following modules.



(a) Direct redundancy (b) Latent redundancy

Figure 5.4 Redundant relations in AoN schedule network

Redundancy elimination algorithm

A simple algorithm is presented for the elimination of all redundant relations in a random network generated by the deletion method. The algorithm is based on a set of three order constraints and ensures a complete enumeration, in the right order, of all indirect relations between any two activities in the network. More specifically, all combinations of three activities – Activity-*i*, Activity-*j*, and Activity-*k* –, which are assumed to be $i < j < k$, are examined. Then, there are only four cases that make a relation between Activity-*i* and Activity-*k* redundant as shown in Table 5.1. From the four cases, three ordering constraints for complete enumeration of all combinations of two activities can be derived.

Order constraints (OC)

- [OC1] Relation between Activity-*i* and Activity-*j* should be examined before the relation between Activity-*i* and Activity-*k*. Therefore, the third activity should be arranged in ascending order.
- [OC2] Relation between Activity-*j* and Activity-*k* should be examined before the relation between Activity-*i* and Activity-*k*. Therefore, the first activity should be arranged in descending order.
- [OC3] There should be at least one activity between Activity-*i* and Activity-*k*.

Table 5.1 Four cases of a redundant relation

Activity-*i* and Activity-*j*	Activity-*j* and Activity-*k*	Activity-*i* and Activity-*k*
Precedence	Precedence	Redundant
Precedence	Redundant	Redundant
Redundant	Precedence	Redundant
Redundant	Redundant	Redundant

Activity-*i*	Activity-*j*	Activity-*k*
1	2	3
2	3	4
1	2	4
1	3	4
3	4	5
2	3	5
2	4	5
	···	
1	(*N*-1)	*N*

(a) Complete enumeration

```
For k = 3 To NNODE       'Third ACT-k
For i = (k - 2) To 1 Step -1   'First ACT-i
For j = (i + 1) To (k - 1)     'Second ACT-j
   If PM(i, j) * PM(j, k) <> 0 Then
      PM(i, k) = 2        'Redundant Relationship.
   End If
Next j
Next i
Next k
```

(b) Algorithm (Visual Basic)

Figure 5.5 An algorithm for redundant relation elimination

A computer code was written to detect the latent redundancy and shown in Figure 5.5. Figure 5.6 (a) shows the network after eliminating direct redundancy. Figure 5.6 (b) shows all the direct and latent redundancies in the same network along with the final precedence relations.

It should be noted that even though the sample network was originally generated to have 10 activities and 20 relations, eliminating redundant relations reduces the meaningful relations to 11 – about half the target. More importantly, Figure 5.6 (b) clearly shows that based on the final relations there are just five potential relations left, which are smaller than the difference, 20 – 11 = 9, to meet the target.

	1	2	3	4	5	6	7	8	9	10
1		1	0	0	R	0	0	R	0	0
2			1	0	1	0	0	R	0	1
3				1	0	1	0	0	1	0
4					1	0	0	R	0	0
5						0	0	1	R	0
6							1	1	0	0
7								0	1	R
8									1	0
9										1
10										

(a) The random network without direct redundancy

	1	2	3	4	5	6	7	8	9	10
1		1	R	R	R	R	R	R	R	R
2			1	R	R	R	R	R	R	R
3				1	R	1	R	R	R	R
4					1	0	0	R	R	R
5						0	0	1	R	R
6							1	1	R	R
7								0	1	R
8									1	R
9										1
10										

(b) The random network without redundancy

Figure 5.6 The effect of redundancy elimination in a random network

5.2.3 Project Progress Curve Generation

Once an activity network for a project is determined, durations of all activities in the project are combined into the network to get a project schedule. The start and finish dates of each activity can be calculated by the critical path analysis. Then the project progress curve is generated by imposing the cost estimate of each activity on the project schedule.

5.3 Early Warning Systems in Project Management

In evaluating forecasting methods, statistical significance is often confused with practical significance or reliability (Armstrong 2002). Statistical error measures such as the mean percentage error (MPE) and the mean absolute percentage error (MAPE) make evaluation procedures a straightforward process because they provide absolute values for the accuracy of forecasting methods. However, practical significance of forecasting methods should be evaluated in the context of a larger decision making system in which the methods are used as aids to provide unbiased predictions and reliable prediction bounds.

In project management, forecasting methods are used to provide reliable estimates of the degree of success in achieving the project objective. One should manage projects not by the current indicators but by the forecasts at completion. In a typical project control system, forecasts are evaluated and compared against the objectives in order to make a decision as to whether the current deviation is significant or not. An early warning system (EWS) provides a formal procedure for evaluating the significance of the deviation between the plan and the actual progress. Therefore, an early warning system is an essential part of project controlling process. With a reliable EWS, a project team is able to decide the timing when additional attention is required to detect some symptoms or early indicators of future problems.

Early warning systems can be characterized with two attributes: the warning metric and the warning criterion. In this section, two probabilistic early warning systems are reviewed according to the nature of warning metrics and warning criteria. The purpose of this section is to show that the early warning system used in the parametric studies in this chapter are, in fact, completely general and, therefore, the assessment of forecasting performance based on the warning systems can be considered a general one rather than case-specific.

EWS based on critical risk

Project objectives, especially the completion date and the budget of a project, should be realistic and achievable. Under a competitive business environment, however, decisions about the project duration and cost in a bid are often made based on a strategic balance between the probability of winning the bid and the expected profits from the project.

For example, an organization may win a project and start it with some accepted level of risk, P_0, which represents the probability of overrun. During the execution, the actual performance is monitored and measured to revise the probability of overrun. Let $P(t)$ denote the probability of overrun as evaluated at time t. Then a project team needs to determine whether the current level of overrun risk is negligible or it is the time for raising a warning flag. This decision must be made based on an objective criterion according to the objectives of the project and risk management strategies of the organization. For example, a warning signal can be transmitted when the current overrun probability exceeds the critical risk for the project, P_C. Figure 5.7 (a) shows an example of the EWS based on critical risk.

EWS based on Duration-at-Risk

Value-at-Risk (VaR), which was originally developed to quantify market risk, is defined as the percentile of gains or losses at a given level of confidence (Jorion 2000). In the construction industry, Ye and Tiong (2000) proposed a method for capital investment decision making based on the VaR concept. The method, the NPV-at-risk method, provides the net present value of a project at a given confidence level.

In a similar way, the Duration-at-Risk (DaR) for project schedule performance can be defined as follows. Given a probability distribution of project duration at time t, $T(t)$, the Duration-at-Risk, $T_p(t)$, of the project duration at a given probability p, is the $(1-p)^{th}$ percentile of the distribution. That is,

$$P\left[T(t) > T_p(t)\right] = p \tag{5.3}$$

Then, $T_p(t)$ represents the project duration that has a $(1 - p)$ probability of finishing earlier.

The primary merit of the DaR is that it provides an intuitive single-value indicator for a probability distribution of project duration. Once the probability p is chosen, the resulting Duration-at-Risk T_p is the single criterion that can be used to make decisions about performance control. For example, a warning can be made when the DaR at a specific time is greater than a predetermined control limit as shown in Figure 5.7 (b).

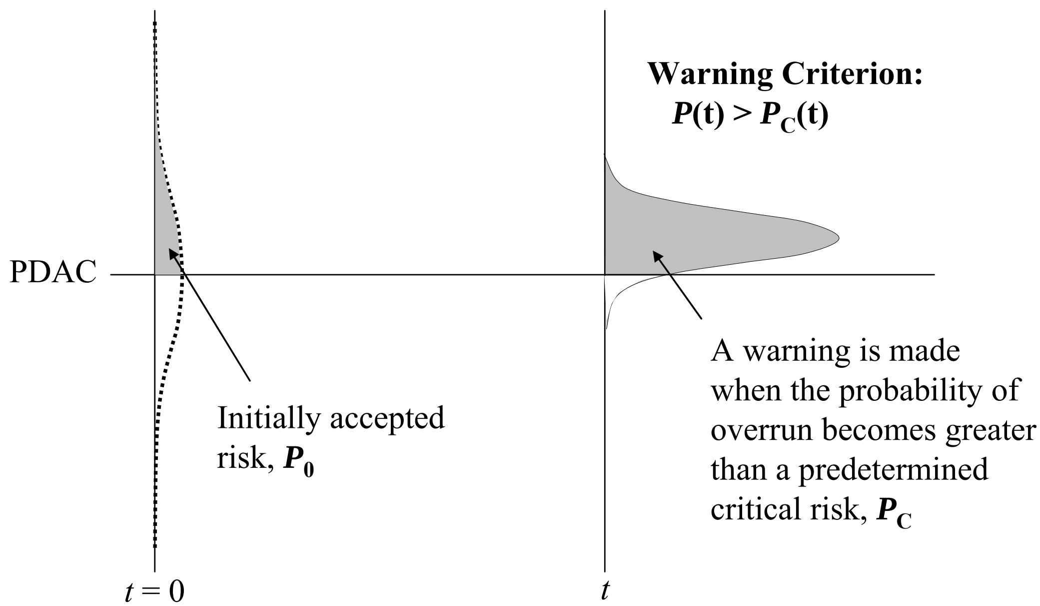
In practice, the control limit is determined at some value greater than the planned completion date to allow some buffer in making official warnings. For example, when the prior distribution of project duration is available, the planned duration at completion (PDAC) can be determined at some confidence level p, that is, PDAC = $\mathrm{DaR}_p(0)$. Then a warning is made at some time t when the $\mathrm{DaR}_p(t)$ calculated from the posterior distribution of project duration exceeds the control limit.

It should be noted that the safety margin between the control limit (CL) and the PDAC can be explicitly included in project objectives in terms of schedule contingency. Contingency is defined in many different ways. According to PMBOK®, contingency is "a provision in the project management plan to mitigate cost and/or schedule risk." Simply put, contingency is something added to the baseline estimate to account for uncertainties. Adding contingency as a separate item in project planning is a recommended practice in the project management community (Uyttewaal 2002). However, practical implementation of contingency may differ from organization to organization according to its culture, policy, and the attitude of the owner or top management against risk. For a project with schedule contingency as a separate item, the PDAC can be determined as

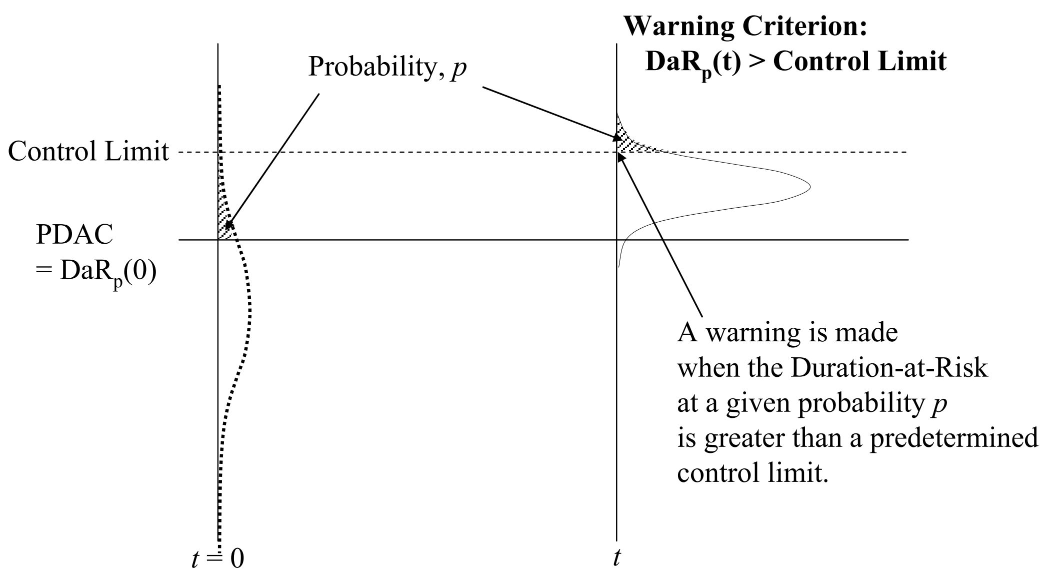
$$\left(PDAC\right) = \left(Baseline\,Estimate\right) + \left(Contingency\right) \tag{5.4}$$

In this situation, a schedule warning occurs when $DaR_p(t) > PDAC$.

In the parametric studies in this chapter, the EWS based on DaR is applied with confidence level p = 0.5. Then the PDAC is the mean of the prior probability distribution of project duration and the $DaR_{0.5}$ at time t is the mean of the posterior probability distribution of project duration. The control limit for warning can be determined from the prior distribution of project duration at a critical level $\alpha < 0.5$.



(a) EWS based on critical risk



(b) EWS based on DaR

Figure 5.7 Early warning systems

(Note: All probability distributions are drawn not to scale)

5.4 Evaluation Criteria for Project Forecasting Method

5.4.1 Accuracy, Timeliness, and Reliability

Establishing proper criteria is a critical step for evaluating and comparing forecasting methods. In the forecasting literature, accuracy is known as the most often used criterion among both practitioners and researchers (Carbone and Armstrong 1982). In the previous work on evaluating project performance forecasting methods (Vanhoucke and Vandevoorde 2006; Zwikael et al. 2000), accuracy of forecasts is assessed with well-known statistical error measures, such as the mean square error (MSE), the mean percentage error (MPE), and the mean absolute percentage error (MAPE). In this dissertation, conventional concepts of the MPE and the MAPE are modified to obtain results that are practically informative and statistically meaningful.

From a practical perspective, however, project managers may be more concerned about the timeliness and reliability of warnings than the accuracy in terms of statistical error measures. The ultimate goal of forecasting is to provide management with warning signals about the degree of success in achieving the project objectives (for example, the scope, the budget, and the completion date). Therefore, warning signals should be reliable and, more importantly, should be transmitted as early as possible. The timeliness of warning signals has been recognized as a desirable characteristic of project performance forecasting (Teicholz 1993) and has served as an important criterion for evaluating forecasting performance (Vanhoucke and Vandevoorde 2006). In this dissertation, two timeliness criteria – the overrun warning point and the probability of correct warning at different stages of completion – have been proposed and applied in the parametric studies. Both criteria are discussed in detail in the following sections. It should be noted here that the overrun warning point works only for probabilistic forecasting models, for example, the KFFM and the BAFM, because it relies on the prediction bounds on single-point forecasts. As a result, when deterministic methods such as CPM and EVM are involved in comparison, the probability of correct warning at different stages of completion is applied.

The reliability of warnings is also an important factor, especially for practitioners, that need to be considered in the evaluation of forecasting methods. Typical responses to an overrun warning may include the time consuming and disruptive process of identifying the root causes, evaluating the significance, and taking appropriate actions to put the project back on track. Even when the warning turns out to be false it may have repercussions. For example, the project team may still need to spend some time on preparing and disseminating additional performance analysis reports that justify the taking-no-action decision.

The overall performance of forecasting methods must be measured in terms of accuracy, timeliness, and reliability. An ideal case would be one in which all of the three elements are maximized simultaneously. However, in most situations, trade-offs among these factors may need to be made according to the managerial strategy or priority of the project. In the remaining parts of this section, the evaluation criteria used in the parametric studies are addressed in detail. Correct understanding of the criteria is essential for correct interpretation of results.

5.4.2 MAPE and MPE

The literature about the evaluation of project performance forecasting methods is very limited. Teicholz (1993) compared two models – the sliding moving average and the up-to-date average – for forecasting final cost and budget of construction projects in terms of the area which is enveloped by the forecast final costs over time and the actual final cost. Zwikael et al. (2000) evaluated five forecasting models in EVM using three performance measures: the mean square error (MSE), the mean absolute deviation (MAD), and the mean absolute percentage error (MAPE). Recently, Vanhoucke and Vandevoorde (2006) used the mean percentage error (MPE) and the MAPE in a study of comparing earned value metrics for schedule forecasting.

Most evaluation criteria used in the previous work can be characterized by the fact that an evaluation of a method or a comparison of methods is made based on the accuracy of forecasting, which is measured, fundamentally, in terms of the average deviation between the forecasts and the actual over a certain period of time. For example, when the estimated duration at completion (EDAC) is forecast over N periods (t, $t+N\Delta t$), the MAPE is defined as

$$MAPE\left(t, t+N\Delta t\right) = \frac{1}{\left(N+1\right)} \sum_{i=0}^{N} \frac{\left|ADAC - EDAC\left(t+i\Delta t\right)\right|}{ADAC} \times 100 \tag{5.5}$$

where ADAC is the actual duration at completion.

The MAPE is useful for comparing relative accuracy of the methods. However, an average error of forecasts over a period of time has little practical significance because it is difficult for decision makers, for example, project managers and owners, to interpret the result and use it to make a better decision. That is, the average error over a period of time does not answer the question of how accurate a prediction made at a specific time during the execution period is.

In this research, this limitation in the conventional measures for forecasting accuracy evaluation has been overcome by evaluating the percentage error (PE) and the absolute percentage error (APE) at a specific time *t*. These are defined as follows:

$$\begin{aligned} APE(t) &= \frac{\left|APDU - EDAC(t)\right|}{APDU} \times 100 \\ PE(t) &= \frac{APDU - EDAC(t)}{APDU} \times 100 \end{aligned} \tag{5.6}$$

Given a project *k*, the mean absolute percentage error and the mean percentage error are calculated from a set of forecasts for a large set of random executions.

$$\begin{aligned} MAPE_k(t) &= \frac{1}{N}\sum_{i=1}^{N} \frac{\left|APDU_{k,i} - EDAC_{k,i}(t)\right|}{APDU_{k,i}} \times 100 \\ MPE_k(t) &= \frac{1}{N}\sum_{i=1}^{N} \frac{APDU_{k,i} - EDAC_{k,i}(t)}{APDU_{k,i}} \times 100 \end{aligned} \tag{5.7}$$

where *k* and *i* are the indices for projects and random executions for individual projects, respectively; *N* is the number of executions for project *k*.

5.4.3 Overrun Warning Point

Although solutions to making decisions under uncertainty have been investigated as early as the 1960s (Hertz 1968), probabilistic results from management models often make decision makers confused and reluctant to rely on them, largely because of the lack of specific and intuitive rules for alternative selection and the chance of preference reversal under different criteria. A new performance measure is proposed in this research based on the characteristics of projects and probabilistic forecasting approaches. The new metric, the overrun warning point (OWP), serves as objective measure of forecasting performance, combining the accuracy of mean prediction, the associated uncertainty, and the risk-attitude of decision makers.

The overrun warning point is defined as the time when a probabilistic forecasting method generates a warning signal about project schedule or cost overrun according to a predetermined acceptable tolerance of decision makers. OWP is an integrative measure of the overall quality of a forecasting method because it is a function of the average value of prediction, its associated uncertainty, and the tolerance limits of decision makers. OWP is also a direct measure of timeliness of a forecasting method and represents the ability of the method to issue an early warning signal in a timely manner.

The graphs in Figure 5.8 illustrate how the OWP is determined from a series of probabilistic predictions of the EDAC of a hypothetical project undergoing roughly 28% schedule overrun. The mean and prediction bounds are estimated every 10 time units. The upper bound (UB) and lower bound (LB) are shown at 10% confidence level on both sides. From the result, one can expect the forecasting system to issue an overrun warning at around 34 time units. The overrun warning is the sign that the probability of schedule overrun exceeds the project's acceptable level which is specified with prediction bounds.

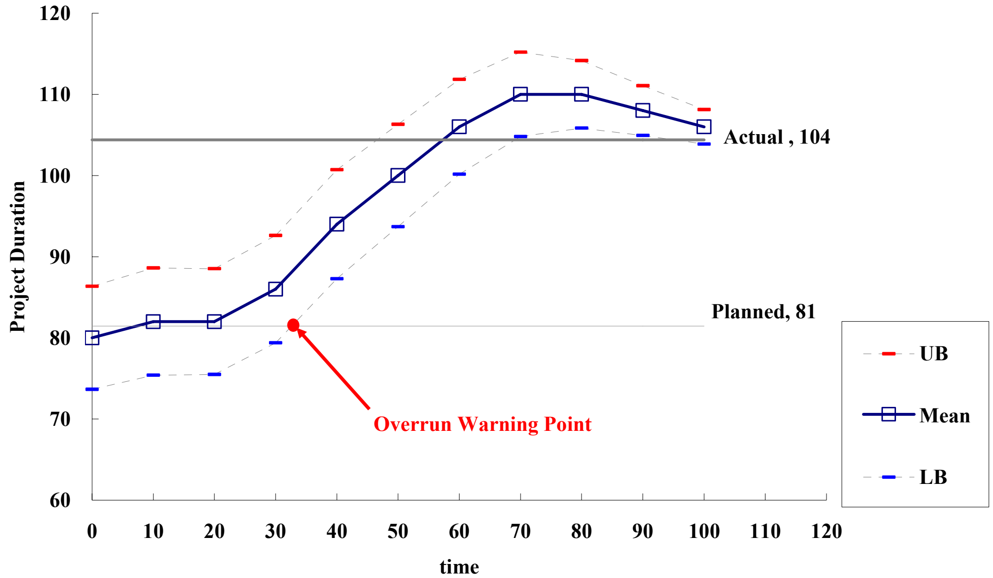


Figure 5.8 Example of the OWP

5.4.4 Probability of Warning at Different Stages of Execution

Probability of warning is an original criterion for the forecasting performance evaluation based on simulation. Given a project k, the probability of warning at time t, $PW_k(t)$, is defined as the probability that a forecasting method transmits a warning signal, whether it is correct or not, for a large set of potential executions under consideration.

$$PW_k(t) = \frac{\text{the number of forecasts that transmit overrun warning at time } t}{\text{the total number of forecasts at time } t} \tag{5.8}$$

The average of PW across many projects is named the mean probability of warning (MPW) and is defined as

$$MPW(t) = \frac{1}{m}\sum_{k=1}^{m} PW_k(t) \tag{5.9}$$

where m represents the number of projects used in calculation.

A set of executions used in the PW calculation is referred to as a target group of the PW. When a target group is chosen from a population with specific requirements, the PW represents a conditional probability of warning under the specific situation. For example, if a target group consists of executions that finish behind schedule, the corresponding $PW(t)$ represents the probability of detecting overrunning executions and issuing correct warnings at time t. On the contrary, for a group of executions finishing ahead of schedule, the $PW(t)$ represents the probability of transmitting false warnings even though the executions finish ahead of schedule.

With a large set of data from the simulation of artificial projects, the probability of warning is integrated with the scenario based evaluation of forecasting performance. That is, the PW at time t given an execution scenario represents the probability of getting a warning for the specific cases belonging to the scenario. The profiles of PW over the forecasting span may be different from each other depending on the scenarios.

5.5 Test of Hypothesis 1

5.5.1 Test Design and Data Generation

The first hypothesis of this research states as follows:

Hypothesis 1: The use of prior information, as used by Bayesian, along with actual performance data increases the quality of forecasting performance with regards to the accuracy, timeliness, and reliability of warning signals.

The first hypothesis is about the effect of incorporating prior performance information available during the preconstruction phase on the accuracy, timeliness, and reliability of warning signals. The BetaS-curve BAF method is chosen to represent the BAF approach because its flexibility in dealing with various types of prior performance information.

Figure 5.9 shows the structure of the artificial project data set used for the hypothesis tests. First, three project groups of ten artificial projects, 30 projects in total, are generated according to predetermined complexity levels for the schedule network. To take into account the potential influence of different network structures of the projects within the same project group, the ten projects in each project group are randomly generated. Projects in the same project group are meant to have the same network complexity level. However, the resulting complexity levels of the component projects vary to some extent, due to the randomness in the redundancy elimination technique – that is, the number of redundant precedence relations in a randomly generated network is not deterministic.

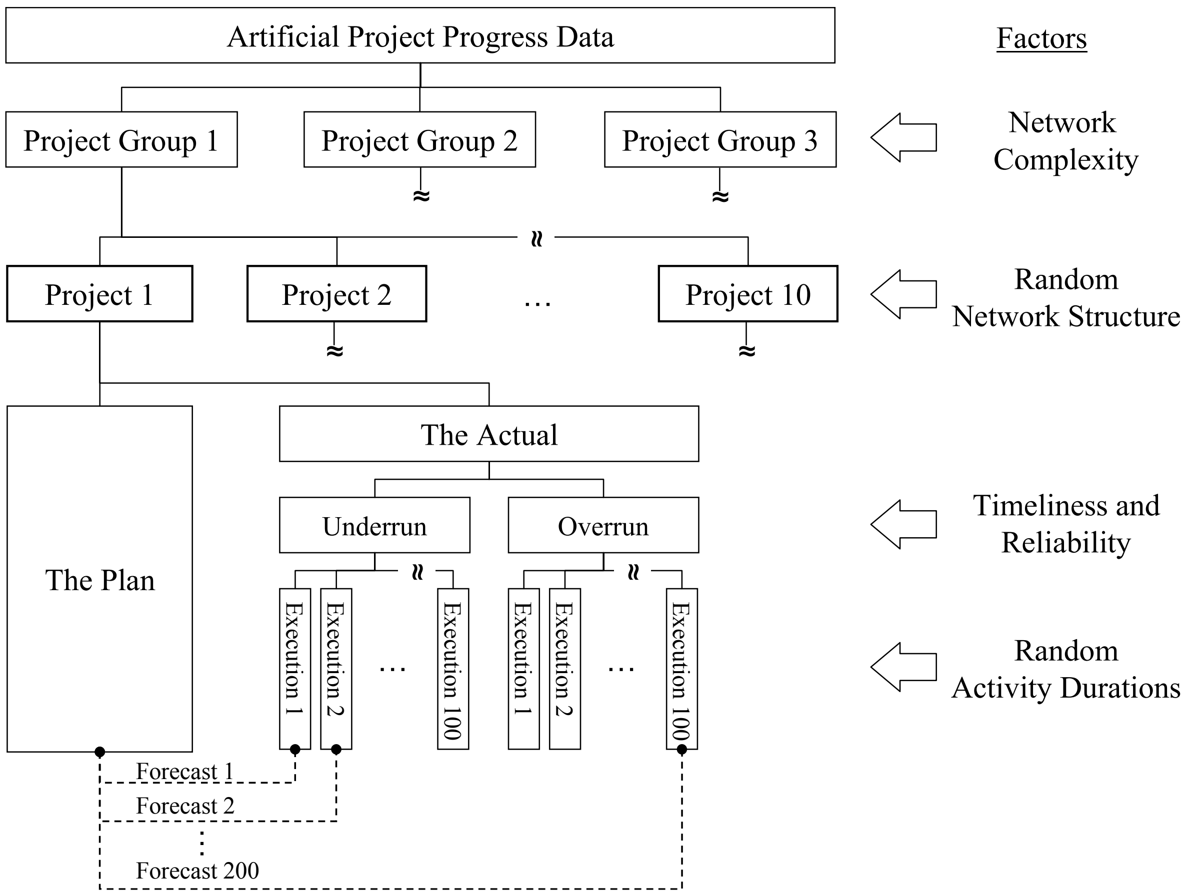


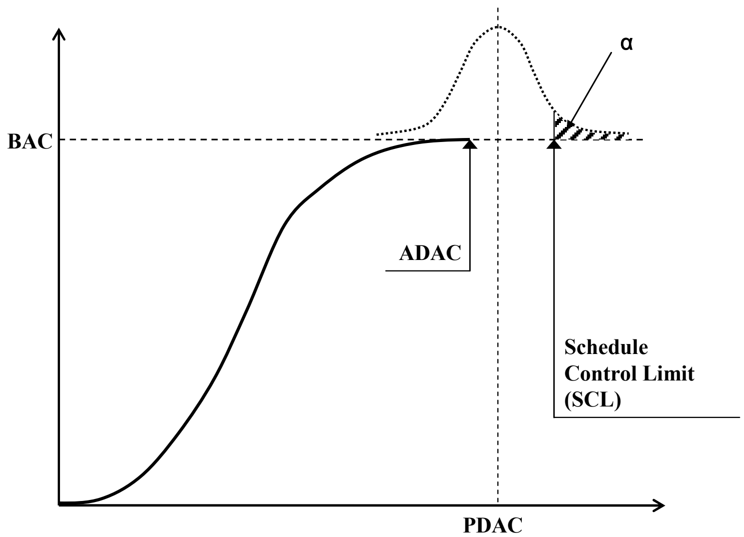
Figure 5.9 Structure of the artificial project progress data set

It is assumed that both activity network and probabilistic estimates of activity durations and costs are known to the project team in advance. This information is used in a network-based schedule simulation in order to generate a large set of random progress curves. From the simulation results, the planned duration at completion (PDAC), or the planned project duration, is chosen at the mean of the estimated probability distribution of the project duration. The planned progress curve corresponding to the PDAC is determined by averaging the random progress curves from the simulation over the progress dimension. The results in this section are obtained with 500 random progress curves.

Then, artificial project executions are randomly generated with the project network and random activity durations. Each of the resulting progress curves is considered the "actual" progress of the project and separately used in forecasts to update the estimated duration at completion (EDAC) after each reporting period. To take into account the potential influence of the stochastic nature of randomly chosen actual progress curves, each project is repeatedly analyzed against 100 random executions (Figure 5.9).

Two sets of random progress curves are generated according to two execution scenarios: the overrun scenario (OS) and the underrun scenario (US). The underrun scenario is the case when a simulated execution finishes earlier than the PDAC and the overrun scenario is the case when it finishes after a schedule control limit, which has to be determined by the project team in advance (Figure 5.10). The schedule control limit (SCL) is also used as a criterion for issuing warning signals. That is, a warning signal is transmitted when the EDAC becomes greater than the SCL. The SCL can be determined based on prior estimates of project duration and a critical risk level α.

According to the execution scenarios, two different sets of artificial project executions – the underrun group and the overrun group – are generated. The underrun group is used to evaluate the probability of false warning and the overrun group is used to measure the probability of correct warning.



Underrun Scenario:

ADAC ≤ PDAC

Overrun Scenario:

ADAC > SCL

Figure 5.10 Two execution scenarios in the hypothesis tests (Note: BAC – the budget at completion; PDAC – the planned duration at completion; ADAC – the actual duration at completion; α - the critical risk for the schedule control limit)

In sum, the total number of forecasts in the test of Hypothesis 1, in other words, the total number of pairs of planned and actual progress curves, is determined as the product of the number of project groups, the number of projects in each project group, the number of execution scenario, and the number of random executions in each scenario, which is 3x10x2x100 = 6000.

Recently Vanhoucke and Vandevoorde (2006) reported a study of comparing three earned value metrics. In the study, two scenarios – an 'ahead of schedule' scenario and a 'project delay' scenario – are simulated by assuming different probability distributions for the generation of random activity durations. More specifically, a triangular distribution with a long tail to the left is used to draw random activity durations for the ahead-of-schedule projects while another triangular distribution with a long tail to the right is used for the delayed projects. However, influencing the actual duration of each activity to ensure that the resulting execution finishes ahead of or behind schedule has some limitations. First, as the authors mentioned, "comparison between scenarios is of little value" and comparison should be made within each scenario. Another limitation is that it is hard to generalize the results. Forecasting performance of a method for specific conditions that are manipulated to meet a certain criterion can hardly represent the forecasting performance under general conditions in which the future depends purely upon the variations of individual activity durations.

In this research, the two execution scenarios are defined according to a typical early warning system and are generated based on the same probability distribution for activity durations. Therefore, it is possible to compare results across different scenarios.

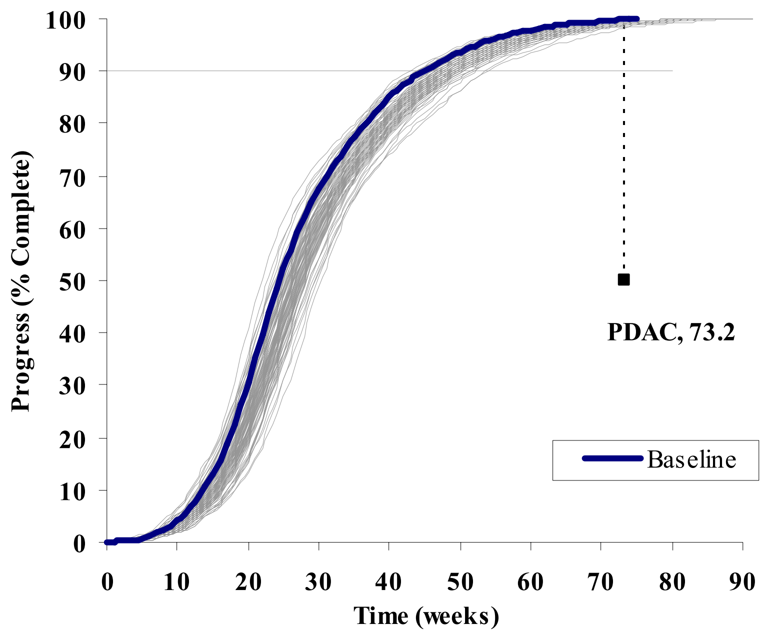
Furthermore, the schedule control limit that is used as the overrun criterion as well as the warning criterion can be adjusted according to the level of risk for performance control.

An example of random executions under different execution scenarios is shown in Figure 5.11. In the example, the SCL corresponding to $\alpha = 0.03$ is about 80 weeks. With the planned duration of 73 weeks, all random executions under the overrun scenario end up finishing with at least 10 percent schedule delay. It should be noted that a random execution belonging to the overrun scenario is not necessarily behind schedule throughout the execution. Figure 5.11 (a) shows that even at the 90% completion point there are some executions whose cumulative progress indicates the projects are ahead of schedule, even though they eventually result in at least 10 percent overrun. A similar pattern can be found in the random executions finishing ahead of schedule.

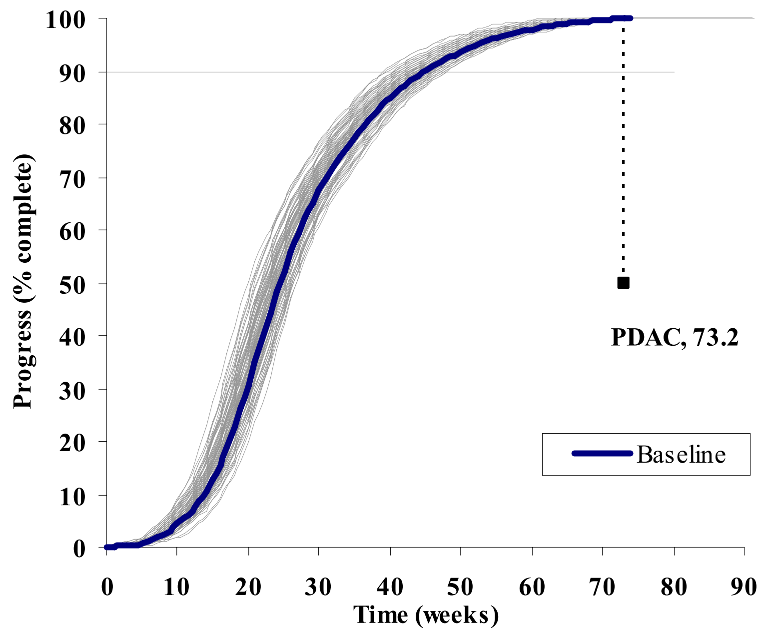
The forecasting performance is measured at different stages of completion. Six evaluation points are chosen as follows:

- 10%T: the time point corresponding to 10 percent of the PDAC
- 20%T: the time point corresponding to 20 percent of the PDAC
- 30%T: the time point corresponding to 30 percent of the PDAC
- 40%T: the time point corresponding to 40 percent of the PDAC
- 50%T: the time point corresponding to 50 percent of the PDAC

90%C: the time point when 90 percent of the project scope has been completed



(a) Overrun scenario



(b) Underrun scenario

Figure 5.11 The planned and random progress curves (100 for each scenario). The critical risk α used for the schedule control limit is 0.03

In order to investigate the influence of prior information on forecasting performance, four cases of prior performance information are chosen. Prior information to the BetaS-curve BAF method consists of the information about project duration and information about the progress curve template (Section 4.3.4). Regardless of the forms and sources of information, prior performance information must be represented in terms of the probability distributions of the corresponding model parameters, that is, T for the project duration and α and m for the shape of the progress curve.

Table 5.2 shows the four combinations of prior distributions used in this parametric study. For all cases, results from a network-based schedule simulation are used to determine the prior distributions of individual parameters. In Prior Case 3, the standard deviation is doubled to examine the potential effect. In the case of shape parameters, the first case (PC1) consists of the mean values of the prior distributions obtained from the simulation. With fixed shape parameters, actual performance data influence only the posterior distribution of project durations. On the contrary, the second (PC2) and the third cases (PC3) use the whole distributions from the simulation. As a result, a comparison between PC1 and PC2 shows the influence of using stochastic S-curves instead of a single, presumably best, S-curve. In Prior Case 4, the prior distributions of the shape parameters are assumed based on subjective judgment, instead of relying on simulation results. The purpose of PC4 is to demonstrate the effect of using relative vague prior information on the forecasting performance of the Bayesian adaptive forecasting model.

Table 5.2 Prior information cases to be compared in the test of Hypothesis 1

Prior Case (PC)	Types of the BetaS-curve parameters	
	Duration parameter (T)	Shape parameters (α and m)
PC1	$T_N \sim N\left(\mu_T, \sigma_T^2\right)$	$\alpha_{\mathrm{N}} \sim \mu_\alpha$, $m_{\mathrm{N}} \sim \mu_m$
PC2	$T_N \sim N\left(\mu_T, \sigma_T^2\right)$	$\alpha_{\mathrm{N}} \sim N\left(\mu_\alpha, \sigma_\alpha^2\right)$, $m_{\mathrm{N}} \sim N\left(\mu_m, \sigma_m^2\right)$
PC3	$T_N \sim N\left(\mu_T, \left(2\sigma_T\right)^2\right)$	$\alpha_{\mathrm{N}} \sim N\left(\mu_\alpha, \sigma_\alpha^2\right)$, $m_{\mathrm{N}} \sim N\left(\mu_m, \sigma_m^2\right)$
PC4	$T_N \sim N\left(\mu_T, \sigma_T^2\right)$	Subjective judgment $\alpha_{\mathrm{N}} \sim$ Beta(1.5, 1.5, 1, 9) $m_{\mathrm{N}} \sim$ Beta(1.5, 1.5, 0.1, 0.9)

5.5.2 Test Data

A large set of artificial project performance data is generated to resemble a wide range of real-life projects. Different levels of network complexity are taken into account with three project groups. Each project group consists of 10 random projects. With individual projects, 100 random executions are generated for the two execution scenarios. In total, 6000 sets of project performance data have been used to evaluate the effect of using different prior information in the BetaS-curve forecasting model.

The properties of the 30 projects in the test data are shown in Table 5.3. The average network complexity index (C_N) of each project group is 9.9, 15.5, and 19.6, respectively. The planned duration at completion ranges from 64 weeks to 91 weeks. It should be noted that the average time to reach the 90 percent completion point is about 60 percent of the planned duration. This means that, on average, it takes about 40 percent of the planned project duration to end the last ten percent of the project. This is typical of projects with front-end-loaded resource allocation. This pattern is attributed to the early start forward calculation method assumed in the generation of the cumulative progress curves. Although it is possible to generate random progress curves representing normal or back-end-loaded projects, the early start calculation in the CPM is chosen because it does not require additional assumptions other than that there is no resource constraint in executing a project.

Table 5.3 Properties of random project data for the test of Hypothesis 1

Project Group	Project #	NR	C_N	PDAC	90%C	S.C.L.	Ratio to PDAC 90%C	Ratio to PDAC S.C.L
PG 1	1	292	9.8	72.9	44.0	79.8	60.4%	110%
	2	292	9.8	73.2	44.9	80.4	61.4%	110%
	3	294	10.0	64.9	36.7	71.8	56.5%	111%
	4	292	9.8	65.6	41.8	71.5	63.7%	109%
	5	294	10.0	71.5	40.7	77.6	56.9%	109%
	6	293	9.9	84.6	49.5	92.7	58.5%	110%
	7	292	9.8	73.0	45.6	80.1	62.4%	110%
	8	294	10.0	76.0	38.2	83.9	50.2%	110%
	9	293	9.9	76.8	44.6	83.7	58.0%	109%
	10	293	9.9	70.0	44.8	76.5	64.0%	109%
	Avg.	293	9.9	72.8	43.1	79.8	59.2	110%
PG 2	1	368	15.7	79.5	49.9	86.3	62.7%	108%
	2	377	16.3	78.7	52.6	85.8	66.8%	109%
	3	359	15.1	91.9	55.3	101.0	60.2%	110%
	4	368	15.7	81.5	50.3	89.5	61.6%	110%
	5	367	15.6	74.0	48.8	80.4	66.0%	109%
	6	368	15.7	77.6	44.5	85.4	57.3%	110%
	7	362	15.3	77.9	47.1	85.3	60.4%	109%
	8	359	15.1	76.7	48.6	83.7	63.4%	109%
	9	360	15.1	88.6	51.9	97.5	58.6%	110%
	10	368	15.7	83.0	54.1	90.2	65.2%	109%
	Avg.	365	15.5	81.0	50.3	88.5	62.2%	109%
PG 3	1	424	19.3	77.5	46.5	85.1	60.0%	110%
	2	431	19.7	78.0	50.3	85.0	64.5%	109%
	3	435	20.0	74.7	44.5	81.4	59.5%	109%
	4	439	20.2	74.9	53.1	82.0	70.9%	109%
	5	437	20.1	80.6	53.0	88.4	65.8%	110%
	6	424	19.3	86.3	61.1	94.0	70.8%	109%
	7	429	19.6	80.4	51.6	88.4	64.2%	110%
	8	418	18.9	78.3	50.4	85.4	64.4%	109%
	9	428	19.6	76.1	47.1	83.6	61.8%	110%
	10	423	19.3	72.5	48.0	80.3	66.2%	111%
	Avg.	428	19.6	77.9	50.6	85.4	64.8%	110%

(Note: 1) NR denotes the number of nonredundant precedence relations in a network; 2) C_N is the complexity index; 3) PDAC is the planned duration at completion; 4) 90%C denotes the time at 90% completion; 5) S.C.L. denotes the schedule control limit, which is determined at $\alpha = 0.03$.)

5.5.3 Results Summary

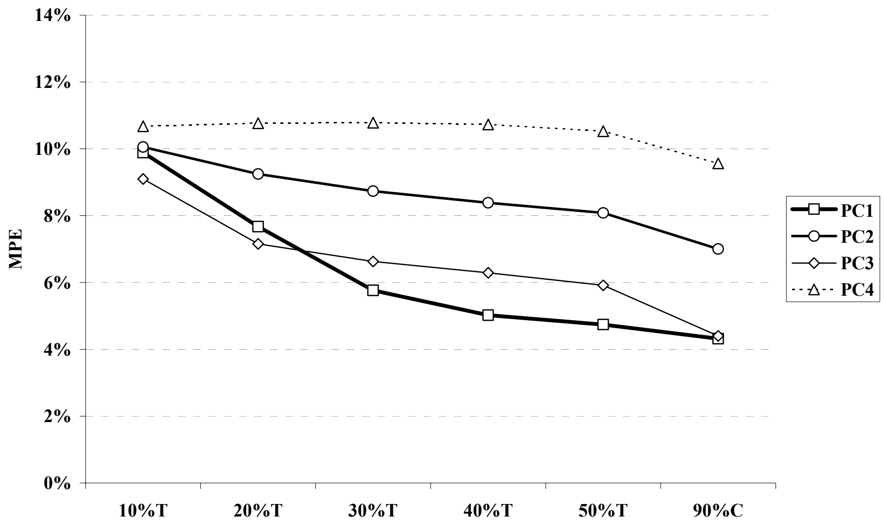
This section summarizes the results in terms of the accuracy of predictions and the timeliness and reliability of overrun warnings.

Accuracy

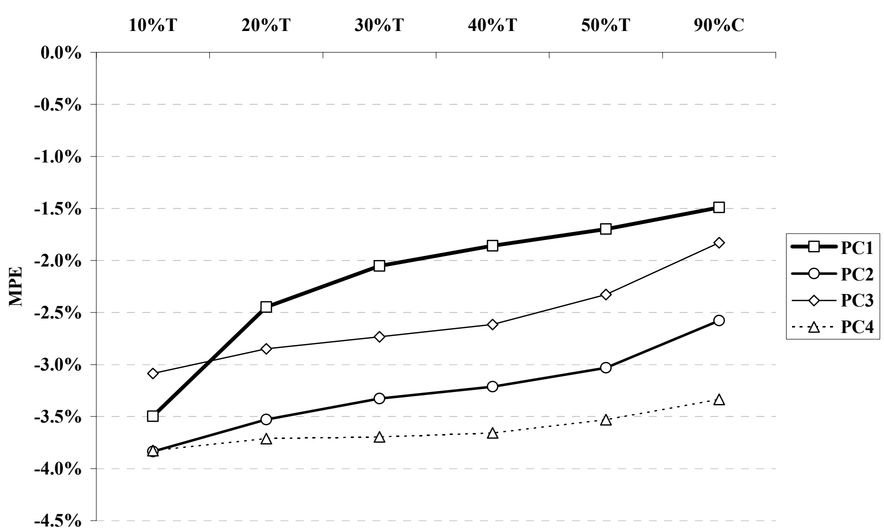
Figure 5.12 shows the overall MPE for the four prior cases, which are aggregated from the 30 projects in the test data. Each data point represents an average over 3000 forecasts at a specific time. The results can be summarized as follows.

- Predictions based on PC1 and PC3 outperform the other two cases at all six evaluation points, regardless of the execution scenarios. In more detail, PC3 provides more accurate predictions than PC1 at early evaluation points (at 10%T and 20%T under the overrun scenario and at 10%T under the underrun scenario).
- Predictions based on PC4 show the largest MPE at all evaluation points. More importantly, while all other three cases show improvement in accuracy as the time of forecasting increases from 10%T to 90%C, PC4 shows slight changes at 50%T and 90%C.

The 95 percent confidence interval on the MPE under the overrun scenario is shown in Figure 5.13. The confidence intervals indicate that the differences in the MPE values are statistically significant on the 5% level.

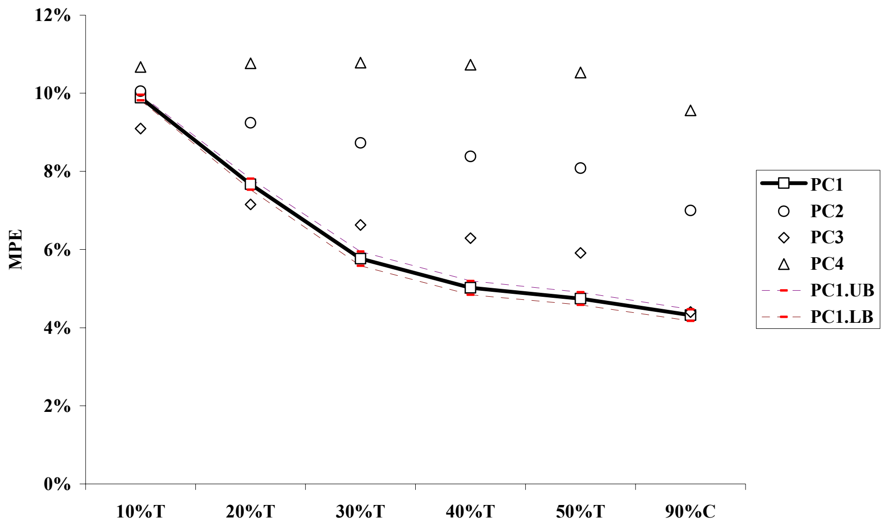


(a) Overrun scenario

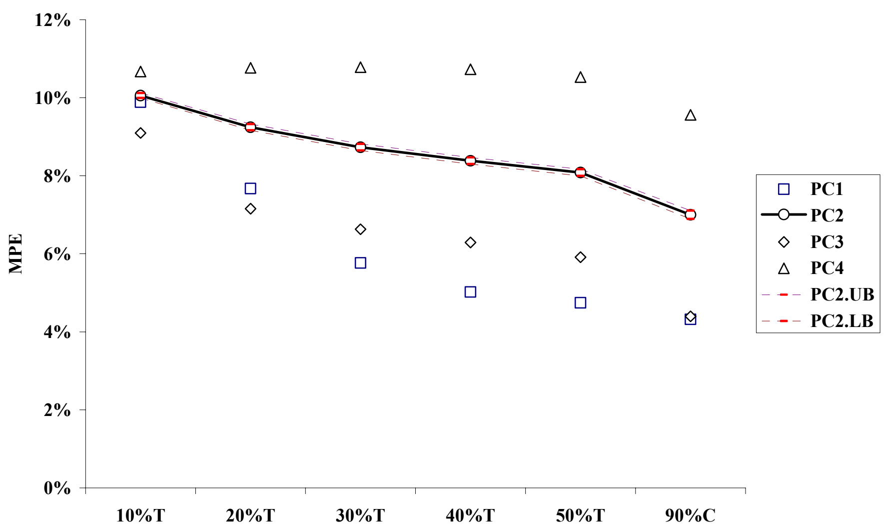


(b) Underrun scenario

Figure 5.12 MPE from 3000 forecasts

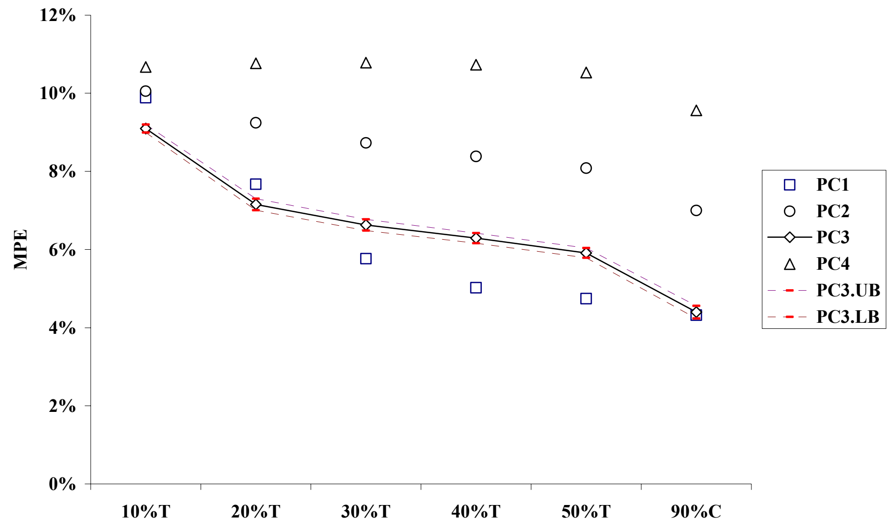


(a) PC1

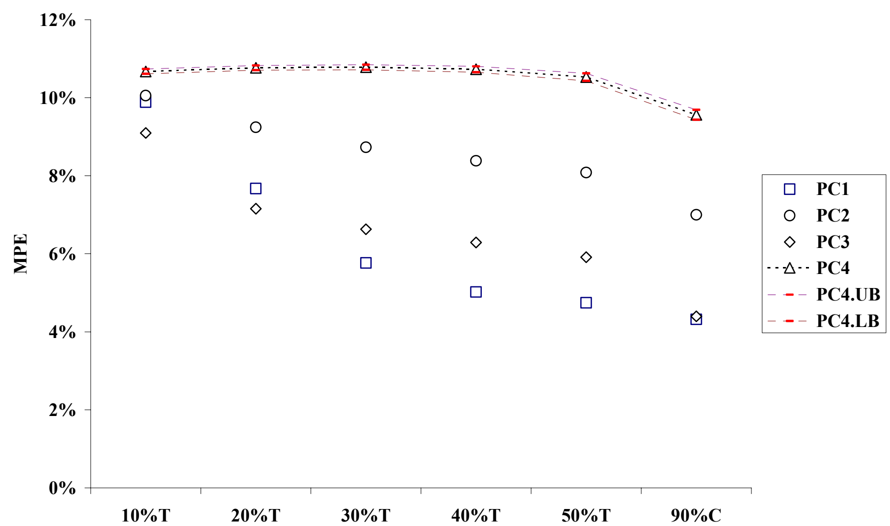


(b) PC2

Figure 5.13 Confidence intervals on MPE under the overrun scenario



(c) PC3



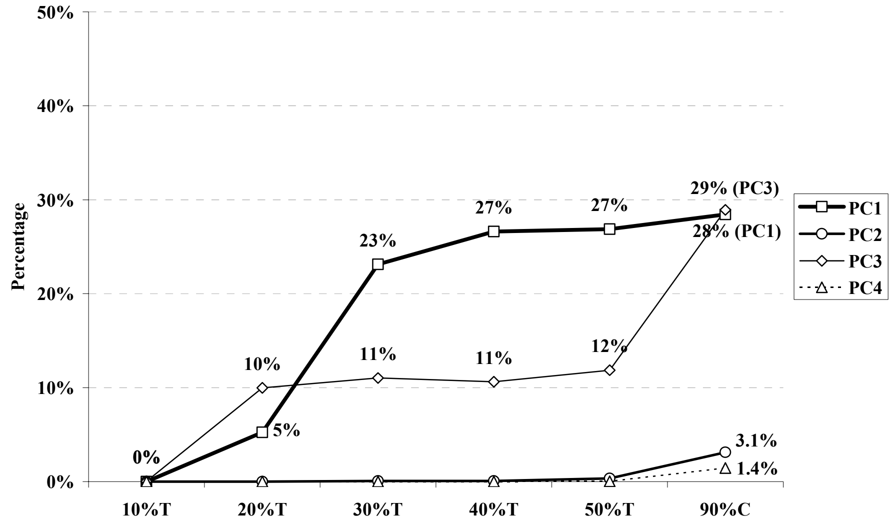
(d) PC4

Figure 5.13 (Continued)

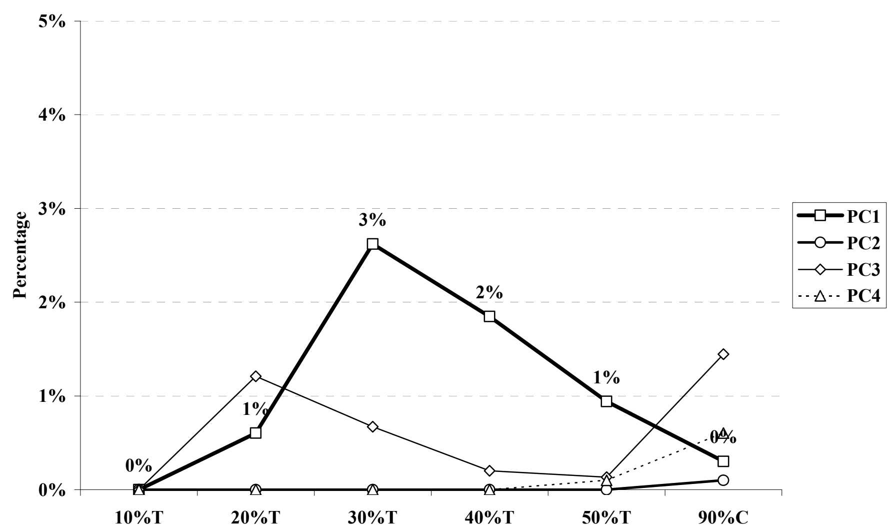
Timeliness and reliability of warning

The probability of warning from 3000 random executions under each execution scenario is shown in Figure 5.14. The results are summarized as follows.

- Forecasts based on PC1 and PC3 outperform the other two cases in their timeliness of warning. At 20%T, PC3 provides more correct warnings than the other cases. At 30%T, 40%T, and 50%T, the probability of correct warning by PC1 is greater than twice the probability by PC3. However, these differences disappear at the 90% completion point.
- The probability of warning under the underrun scenario is much smaller than the probability of warning under the overrun scenario. The PW of PC1 increases during the first three evaluation points, hitting the peak of 3% at 30%T. However, at the later evaluation points – at 40%T, 50%T and 90%C –, the PW decreases. This pattern confirms that the BetaS-curve BAF model provides, on average, more reliable warnings as a project proceeds.
- Early warning potentials of PC2 and PC4 are almost zero up to 50%T. Even at 90%C, the probability of correct warning is merely 3.1% for PC2 and 1.4% for PC4. It should be noted that the accuracy of predictions by PC2 shows a similar pattern to those by PC1 and PC3 under both execution scenarios. However, when it comes to the probability of (correct or false) warning, the results in Figure 5.14 (a) show that PC2 fails to detect overrunning projects in a timely manner. This shows the importance of evaluating project performance forecasting models in terms of the probability of warning as well as the accuracy measures.



(a) Overrun scenario



(b) Underrun scenario

Figure 5.14 Overall percentage of warning under different scenarios

5.6 Test of Hypothesis 2

5.6.1 Test Design and Data

The second hypothesis of this research states as follows:

Hypothesis 2: The Kalman filter model and the Bayesian adaptive model outperform the conventional methods such as CPM and EVM with regards to the accuracy, timeliness, and reliability of warning signals.

This hypothesis is about relative performance of different forecasting methods under diverse project situations with respect to execution scenarios, stages of completion, and complexity levels of activity network. The new forecasting methods – the KFF and the BAF – are compared with the CPM and the EVM.

General introductions to and major assumptions of the CPM and the EVM used in the test are discussed in Section 2.4 and Section 5.1.3. It should be recalled here that, except for the CPM, all other methods use the same form of actual performance data which are given in typical earned value performance measures, that is, the planned cumulative progress and the actual cumulative progress. The Bayesian adaptive forecasting model uses cumulative progress data over the time horizon. However, both the EVM and the KFF focus on the horizontal deviation of the actual progress from the plan, which is measured in terms of the earned schedule in Equation (2.3). Among the four prior information cases in Section 5.5, the first type of prior distribution is chosen in the BetaS-curve forecasting model.

The same procedure used in the test of Hypothesis 1 is applied to generate the test set of artificial project data for Hypothesis 2 (Figure 5.9). First, three project groups with different levels of network complexity are generated. Each project group consists of 10 projects with different properties in terms of the number of nonredundant precedence relations, the planned duration at completion, and the cumulative progress pattern of the baseline progress curve. The diversity at the project level is achieved by the random progress generation technique discussed in Section 5.2.

The properties of individual projects are summarized in Table 5.4. The mean project durations of the three project groups are 72, 77, and 86 weeks for PG1, PG2, and PG3, respectively. The ratio of the 90% completion point to the mean project duration is about 60%. The SCL is determined at $\alpha = 0.10$, which corresponds to about seven percent schedule overrun.

Given that the planned progress and the actual execution are determined, the four methods are individually applied to forecast the EDAC after each reporting period. Forecasting performance is evaluated at six points. However, the 50%T in the test of Hypothesis 1 is replaced with the 5%T in order to take a closer look at the forecasting performance during the early phase of execution.

Table 5.4 Properties of random project data for the test of Hypothesis 2

Project Group	Project #	NR	C_N	PDAC	90%C	S.C.L.	Ratio to PDAC	
							90%C	S.C.L
PG 1 (C_N = 10)	1	288	9.4	80.0	44.8	85.8	56.0%	107%
	2	292	9.8	72.0	39.4	77.2	54.7%	107%
	3	292	9.8	73.9	41.9	78.6	56.7%	106%
	4	295	10.1	73.8	43.2	78.7	58.5%	107%
	5	294	10.0	76.5	45.9	81.9	60.0%	107%
	6	293	9.9	68.3	41.6	73.7	61.0%	108%
	7	293	9.9	58.3	35.1	62.5	60.3%	107%
	8	298	10.3	64.9	36.4	69.1	56.2%	107%
	9	290	9.6	81.1	45.4	86.1	55.9%	106%
	10	292	9.8	71.2	43.6	75.7	61.3%	106%
	Avg.	292.7	9.8	72	41.7	76.9	58.1%	107%
PG 2 (C_N = 16)	1	369	15.8	80.3	47.9	85.6	59.6%	107%
	2	376	16.2	74.5	46.9	79.4	63.0%	107%
	3	356	14.8	83.6	42.8	88.8	51.3%	106%
	4	358	15.0	76.8	46.2	81.7	60.2%	106%
	5	364	15.4	78.4	50.0	83.1	63.8%	106%
	6	367	15.6	79.6	46.6	85.3	58.6%	107%
	7	365	15.5	68.1	38.8	72.9	57.0%	107%
	8	364	15.4	85.9	56.9	91.2	66.3%	106%
	9	364	15.4	73.6	39.8	79.0	54.1%	107%
	10	366	15.6	67.1	42.2	71.3	63.0%	106%
	Avg.	364.9	15.5	77	45.8	81.8	59.7%	107%
PG 3 (C_N = 19)	1	433	19.8	86.7	57.9	92.1	66.7%	106%
	2	417	18.9	85.1	51.6	90.6	60.7%	106%
	3	429	19.6	85.7	47.7	91.3	55.7%	106%
	4	422	19.2	86.0	63.0	91.6	73.3%	107%
	5	421	19.1	81.8	53.7	87.0	65.7%	106%
	6	430	19.7	80.2	51.2	84.5	63.9%	105%
	7	415	18.8	86.4	52.9	91.7	61.2%	106%
	8	426	19.4	84.0	50.7	90.0	60.4%	107%
	9	413	18.6	84.6	51.9	90.0	61.4%	106%
	10	430	19.7	91.9	57.1	97.9	62.1%	107%
	Avg.	423.6	19.3	86	53.8	90.6	63.1%	106%

(Note: 1) NR denotes the number of nonredundant precedence relations in a network; 2) C_N is the complexity index; 3) PDAC is the planned duration at completion; 4) 90%C denotes the time at 90% completion; 5) S.C.L. denotes the schedule control limit, which is determined at $\alpha = 0.1$.)

5.6.2 Accuracy of the EDAC

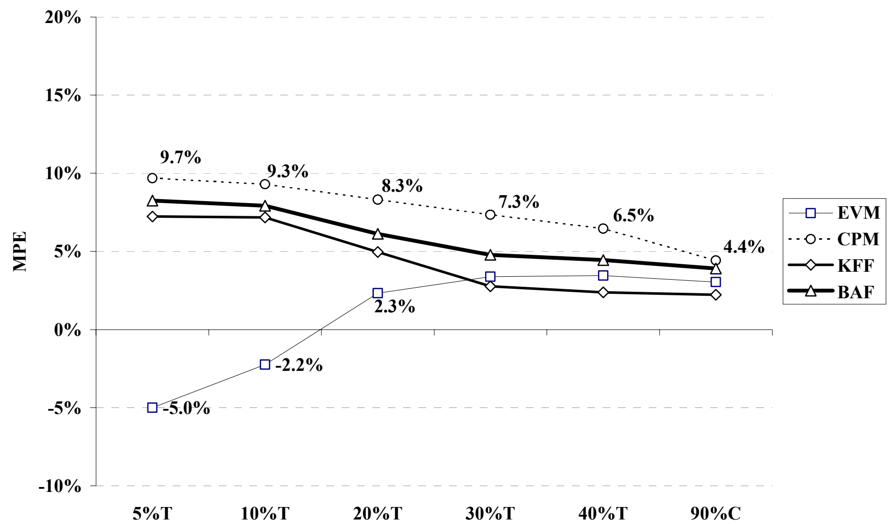
This section summarizes the results in terms of two accuracy metrics: the MPE and the MAPE. In the following section, the timeliness and reliability of warnings by different forecasting methods are evaluated in terms of the probability of warnings under the overrun and the underrun scenarios.

Accuracy - MPE

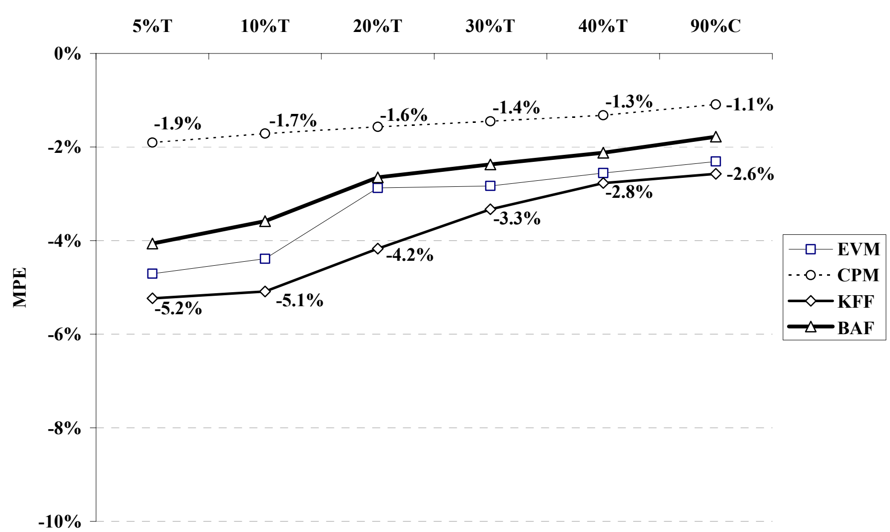
The accuracy of different methods is measured in terms of the MPE. The results are shown in Figure 5.15 at the six evaluation points. Each point in the graphs represents the average over 3,000 forecasts from 30 random projects. The results can be summarized as follows.

- Under the overrun scenario, the MPE of predictions by the CPM, the KFF, and the BAF models decreases with forecasting time, which means that the forecasting accuracy of the CPM, the KFF, and the BAF improves as a project is executed and progressed. For underruning executions, all four methods' accuracy improves over time.
- Among the CPM, the KFF, and the BAF methods, the order of forecasting accuracy is KFF > BAF > CPM under the overrun scenario. However, the order is reversed to CPM > BAF > KFF under the underrun scenario. These orders are consistent at all the six evaluation points.
- The MPE profile of the EVM for overrunning projects is different from the others. During the early stages, 5%T and 10%T, the MPE has negative values, which means that predictions by the EVM, on average, overestimate the project duration.

However, at the following evaluation points, the MPE are positive. This inconsistent pattern is attributed to the large variation in the EVM predictions, which results in errors with different signs.



(a) Overrun scenario



(b) Underrun scenario

Figure 5.15 MPE of the EDAC under different execution scenarios ($\alpha = 0.10$)

Figure 5.16 shows the variability in the observed percentage errors for each method. The result from the EVM shows that the EVM has a relatively large variation during the early evaluation points. It also shows that the small value of MPE is a result of averaging the large overestimating errors and the large underestimating errors during the early stages. Another noticeable pattern found in Figure 5.16 is that the variability of the percentage error samples from the KFF and the BAF is getting larger during the early stages and, later, slightly decreasing. Increasing uncertainty early in a project is counterintuitive because as the forecasting time passes more actual data are generated, which one might assume would decrease the uncertainty in forecasting errors. This pattern found in the KFF and the BAF can be explained by the use of prior information. Early in a project, both methods make forecasts based more on prior information than on a small number of actual data. As a result, forecasts made during early in a project tend to be close to the prior estimate of the project. That is, $EDAC(t) \cong PDAC$. Then, from the definition of the percentage error, the percentage error for *i*-th execution of project *k* is

$$\begin{aligned} PE_{k,i}(t) &= \frac{APDU_{k,i} - EDAC_{k,i}(t)}{APDU_{k,i}} \times 100 \\ &\cong \frac{APDU_{k,i} - PDAC_{k}}{APDU_{k,i}} \times 100 \end{aligned} \tag{6.10}$$

Therefore, when informative prior is used and the predictions are dominated by the prior rather than a small samples of actual data, the variability in percentage errors is dominated by the variability in actual project durations, not by the prediction uncertainty in EDAC(*t*). In other words, the Bayesian and Kalman priors used in the test reduce prediction uncertainty due to small sample data. Obviously, the pattern in Figure 5.16 (c) and (d) can be avoided when noninformative priors are used. For example, the forecasts in Figure 4.17 (page 112) show that when noninformative prior is used in the BAF method, large errors can be observed even early in a project.

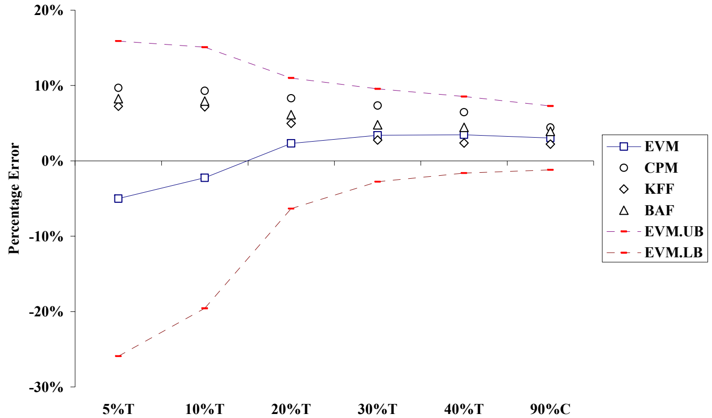
The 95 percent confidence interval in Figure 5.17 shows that the differences in the MPE values in Figure 5.15 (a) are statistically significant on the 5% level.

Accuracy –MAPE

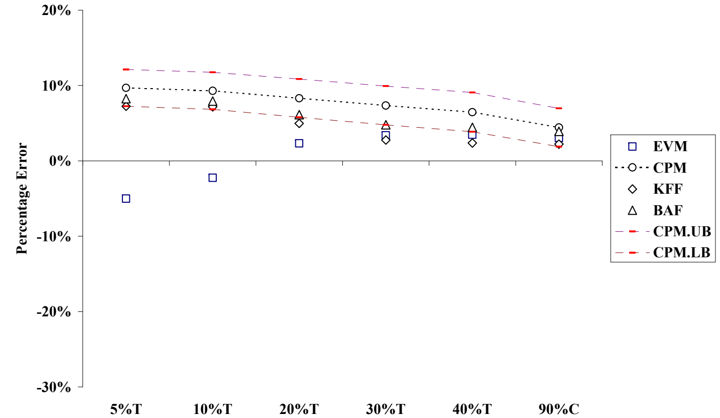
Another commonly used evaluation measure for forecasting methods is the mean absolute percentage error (MAPE). Since the absolute value of a percentage error is used, the MAPE avoids the offset effect in the EVM results in Figure 5.15. The results in Figure 5.18 can be summarized as follows.

- Among the CPM, the KFF, and the BAF, the same orders of forecasting accuracy measured with MPE in Figure 5.15 are observed with MAPE. That is, KFF > BAF > CPM under the overrun scenario and CPM > BAF > KFF under the underrun scenario. However, the BAF and the KFF coincide at 30%T under the overrun scenario.

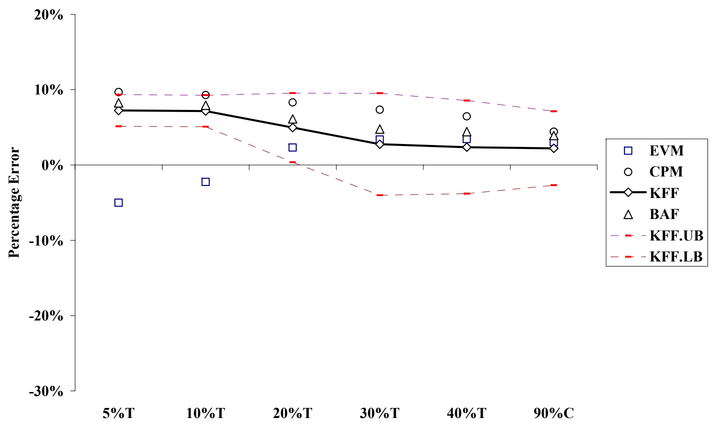
The EVM shows the largest errors at 5%T and 10%T for overrunning executions, and at 5%T, 10%T, and 20%T for underrunning executions. However, the accuracy improves and gets very close to the KFF and BAF at 30%T, 40%T, and 90%C.



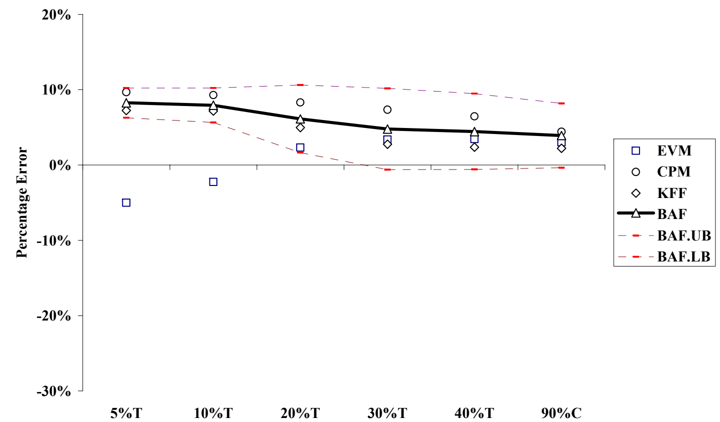
(a) EVM



(b) CPM

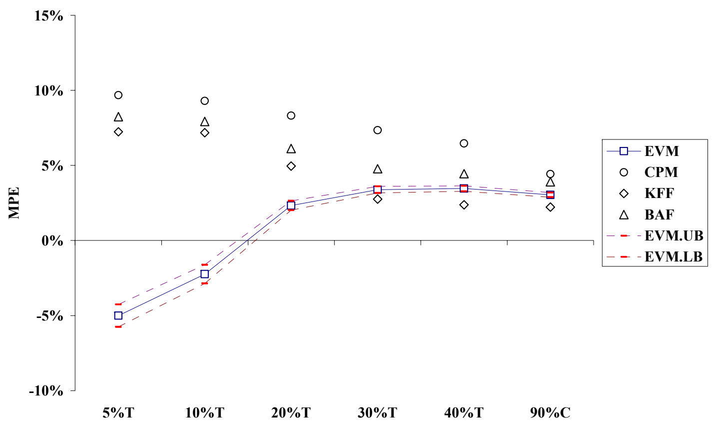


(c) KFF

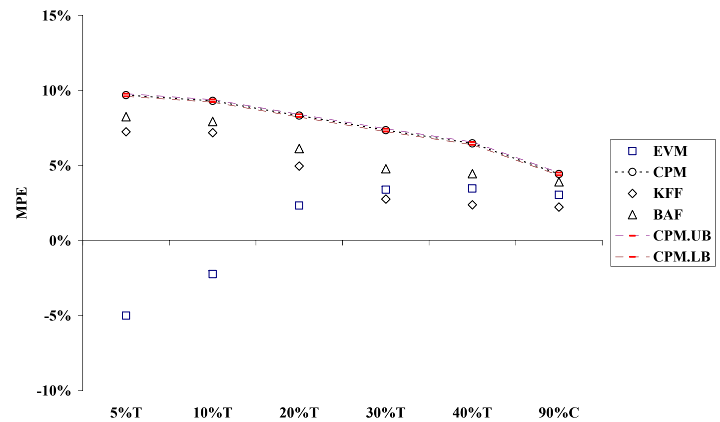


(d) BAF

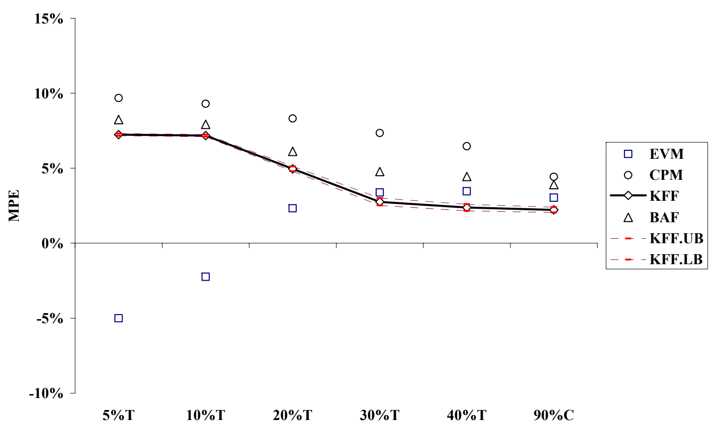
Figure 5.16 Variability in percentage errors under the overrun scenario
(The upper and lower bounds are determined at a standard deviation above and below the mean.)



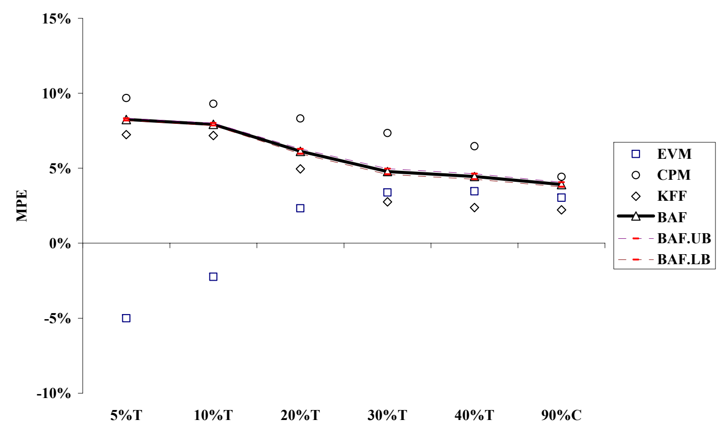
(a) EVM



(b) CPM

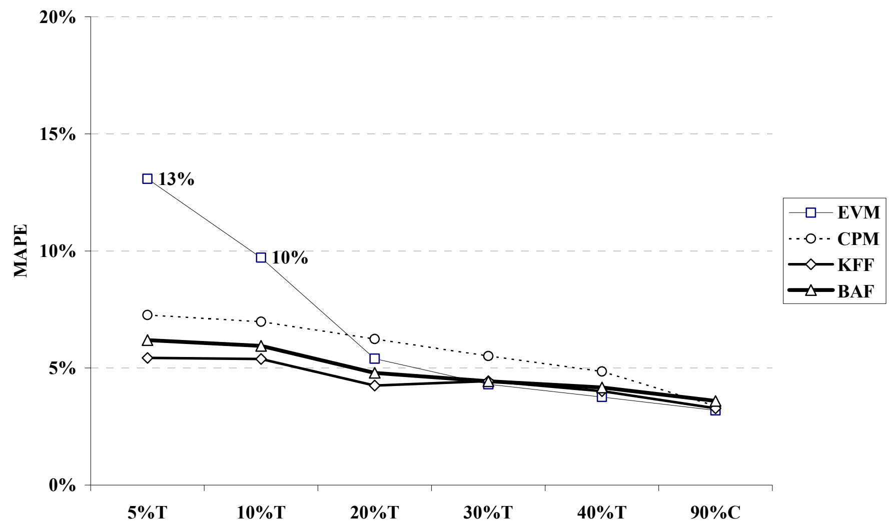


(c) KFF

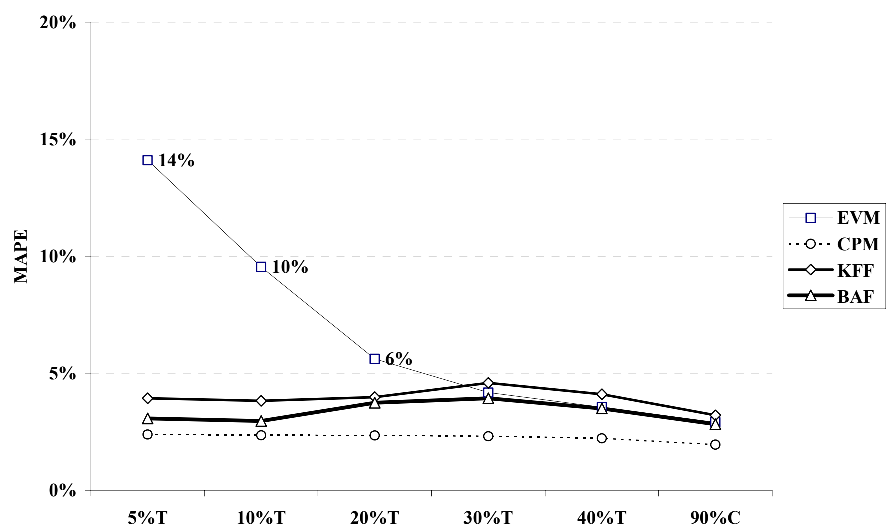


(d) BAF

Figure 5.17 Confidence intervals on MPE under the overrun scenario
(The upper and lower bounds are determined at a standard deviation above and below the mean.)



(a) Overrun scenario



(b) Underrun scenario

Figure 5.18 MAPE of the EDAC under different execution scenarios (α = 0.10)

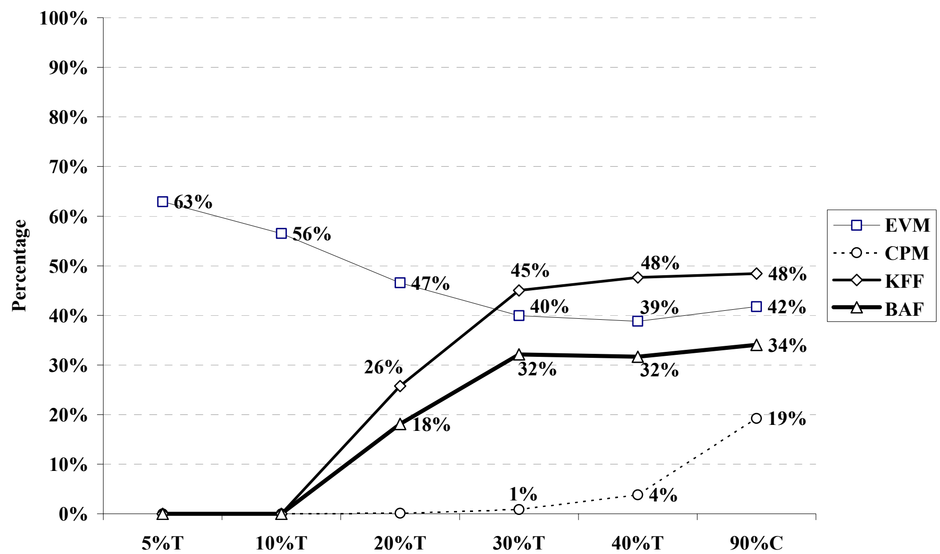
5.6.3 Timeliness and Reliability of Warning

Given a project, the probability of warning is calculated from 100 random executions. The probability of warning at specific times represents the probability that a method transmits a warning signal for a set of random executions in an execution scenario. The PW under the overrun scenario represents the probability of correctly detecting a schedule overrun at completion and, therefore, it can be used to measure the timeliness of a forecasting model to give an early warning against projects that belong to the overrun scenario. However, it is also important for a decision maker to understand the probability of false warning, that is, the probability that a forecasting method gives a warning against executions which, in fact, are going to finish ahead of the planned completion date. An abundance of false warnings eventually damages the reliability of the forecasting method. Furthermore, decision makers relying on such methods may end up ignoring warnings because of "warning fatigue", which is often observed among the people under frequent warnings such as security threats or hurricane evacuations.

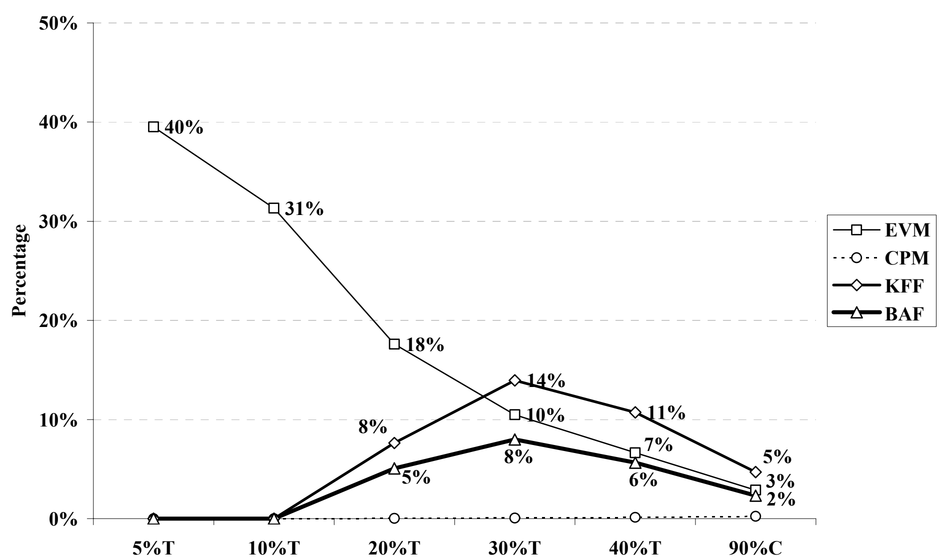
Therefore, the two graphs in Figure 5.19 show the two sides of the same coin. The results can be summarized as follows.

- **CPM**: The CPM forecasts are extremely conservative in giving warning signals. The risk of false warning against underruning projects is almost zero at all the six evaluation points (Figure 5.19 (b)). The chance of giving correct warnings for overrunning projects remains negligible up to 40%T. It increases slightly at 40%T to 4%. However, even at 90%C, it is merely 19%. These results can be explained by the retrospective nature of the CPM forecasting. Schedule forecasts by the CPM are based on what has already happened and the performance of remaining jobs in the future is not adjusted according to the past performance. That is, CPM assumes that the correlations between past activity durations and future activity durations are always zero. Therefore, it can be concluded that, when it comes to predicting the future, an overrun warning by the CPM is extremely reliable because is almost free from false warnings. But the cost of such reliability is that, its early warning capability is very poor and the chance of getting early warnings for overrunning projects is very low.

- **EVM**: In sharp contrast to the CPM, the EVM shows, at the first three evaluation points, higher probabilities of warning than the other methods, regardless of the execution scenarios. Under the overrun scenario, the chance of getting a correct warning from the EVM is over 50% even at 5%T and 10%T. However, the probability of correct warning decreases with forecasting time. This is counter-intuitive because the reliability of warning should increase as more actual information is acquired along the progress of a project. Nonetheless, the results in Figure 5.19 (a) indicate that the EVM gives more correct warnings early in the project than the other three methods. However, the results under the underrun scenario shed a different light on how the early warnings from the EVM should be interpreted. Figure 5.19 (b) reveals that the EVM tends to give a lot of false warnings during the early stages of a project. At 5%T and 10%T, the probability of a false warning is about 40% and 31%, respectively. Although the risk of false warning for projects decreases over time – at 40%T, it is 7% –, the high risk of false warning during the earlier stages may deter decision makers from counting on the EVM.



(a) Overrun scenario



(b) Underrun scenario

Figure 5.19 Probability of warning under different execution scenarios ($\alpha = 0.10$)

- **KFF**: The Kalman filter forecasting method shows relatively stable and reasonable warning patterns under both **execution** scenarios. Under the overrun scenario, the probability of correct warning remains negligible up to the 10%T, then it starts to grow to reach 26% at the 20%T and 45% at the 30%T. On the other hand, the probability of false warning by the KFF is again negligible up to the 10%T, increases to peak at 30%T, and starts to decline at the following evaluation points. The maximum value of the probability of false warning is about 14% at the 30%T. These results are reasonable because forecasts become more reliable with more actual performance data.
- **BAF:** The Bayesian adaptive forecasting method shows similar profiles of probability of warning over time to those of the KFF. However, the response of the BAF to actual performance data is slower that that of the KFF, which indicates that the BAF is less capable for giving early warnings against overrunning projects and more reliable in avoiding false warnings against underruning projects. It should be noted that the forecasting performance of the BAF is influenced by the use of different priors or different measurement errors. The same is also true for the KFF method. Predictions shown in Figure 3.9 (page 53), Figure 4.16 (page 109), and Figure 4.17 (page 112) show the wide range of possible predictions that can be made by the KFF and the BAF based on the same actual performance data. Therefore, comparison between the BAF and the KFF based on the results in Figure 5.19 can not be generalized as the overall performance of the methods.

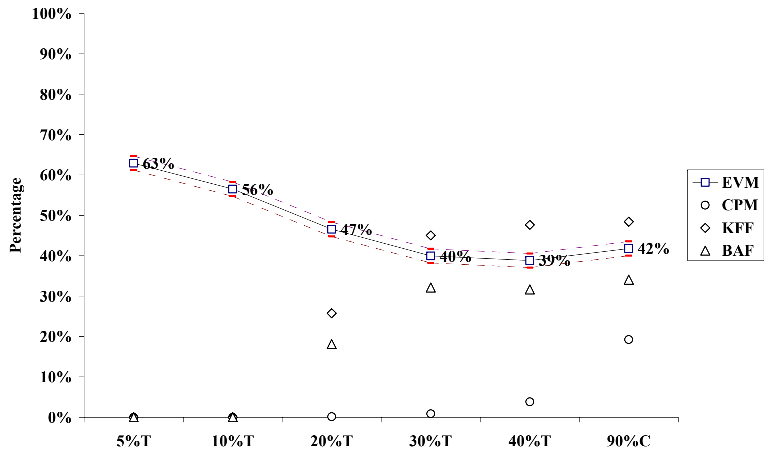
The 95 percent confidence interval on the PW in Figure 5.19 are shown in Figure 5.20 and Figure 5.21, which show that the differences in the PW are statistically significant on the 5% level.

5.6.4 Influence of the Network Complexity on Forecasting Performance

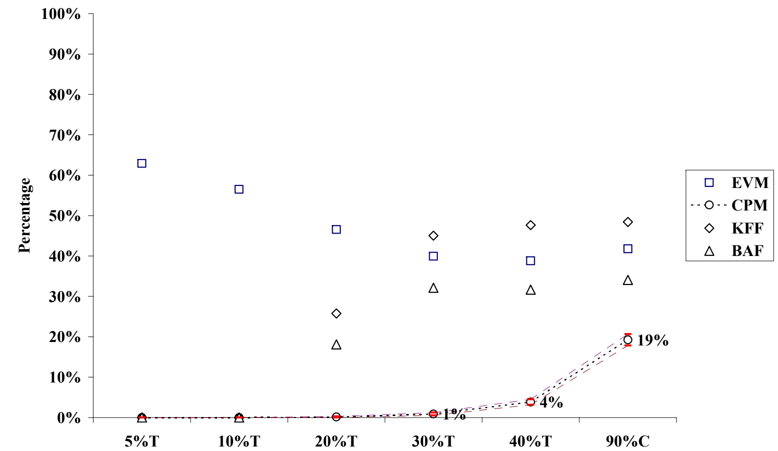
It is commonly taken for granted that as the network complexity of a project increases it gets more difficult to schedule and control the project because of the increasing interrelationships among network activities. This section investigates the influence of the network complexity, as defined in Section 5.1.4, on the forecasting performance of the EVM, the CPM, the KFF, and the BAF methods.

In addition to the results obtained from the three project groups with different levels of network complexity, a group of linear projects is analyzed as a reference case. Each linear project consists of 20 activities, all of which are on the critical path. Figure 5.22 shows the baseline and the first 100 executions out of 1000 random executions under each scenario.

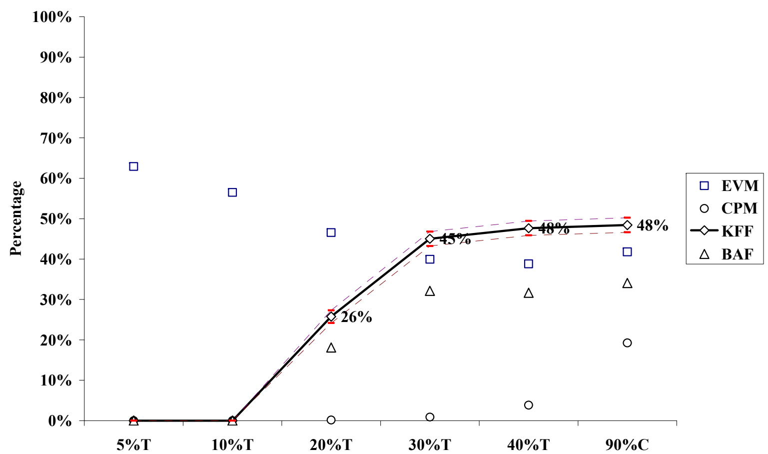
The major results in terms of MPE, MAPE, and PW are shown in Figure 5.23. From the graphs, one can see that the characteristic patterns of forecasting methods, which were discussed in Section 5.6.2 and Section 5.6.3 are also found with the linear projects.



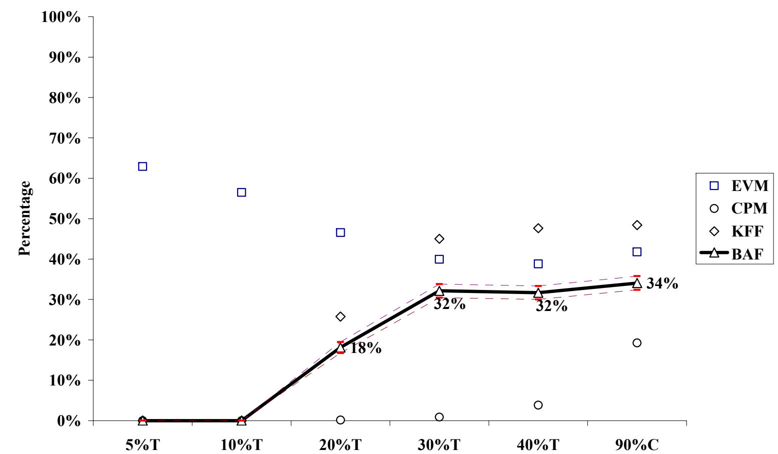
(a) EVM



(b) CPM

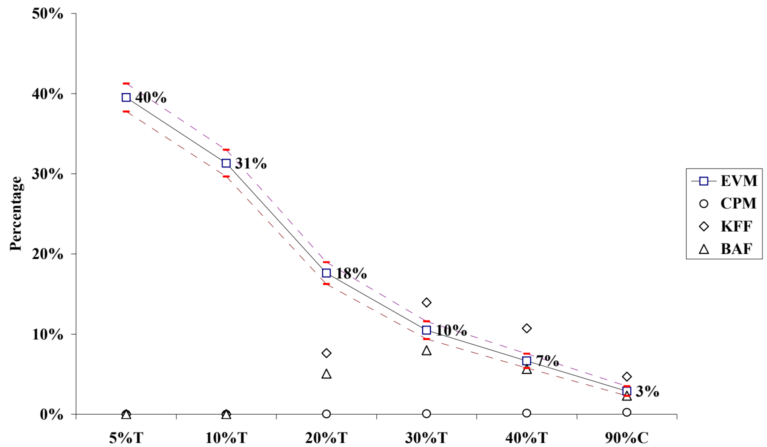


(c) KFF

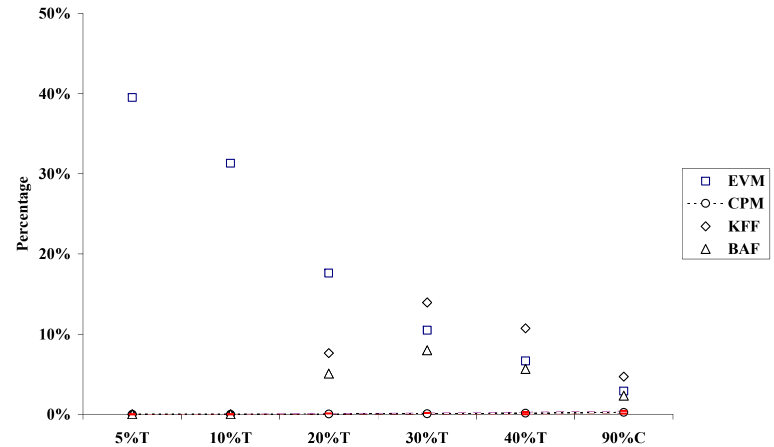


(d) BAF

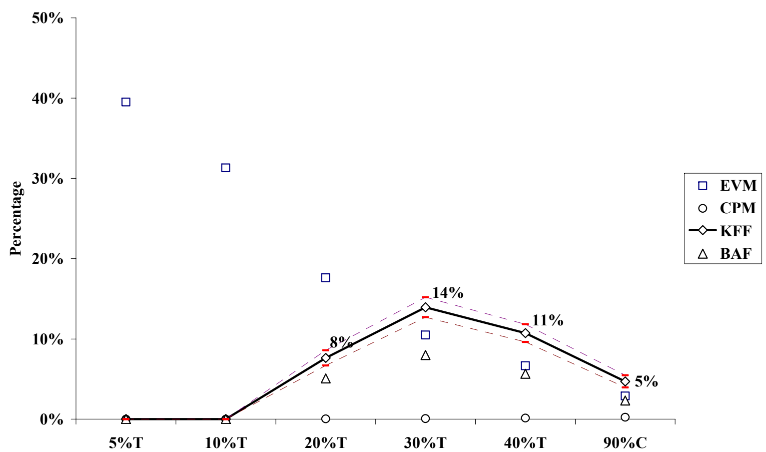
Figure 5.20 Confidence intervals on PW under the overrun scenario



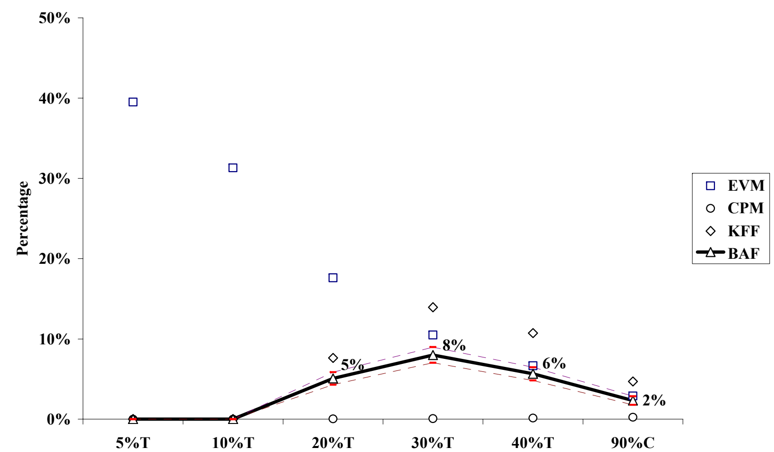
(a) EVM



(b) CPM

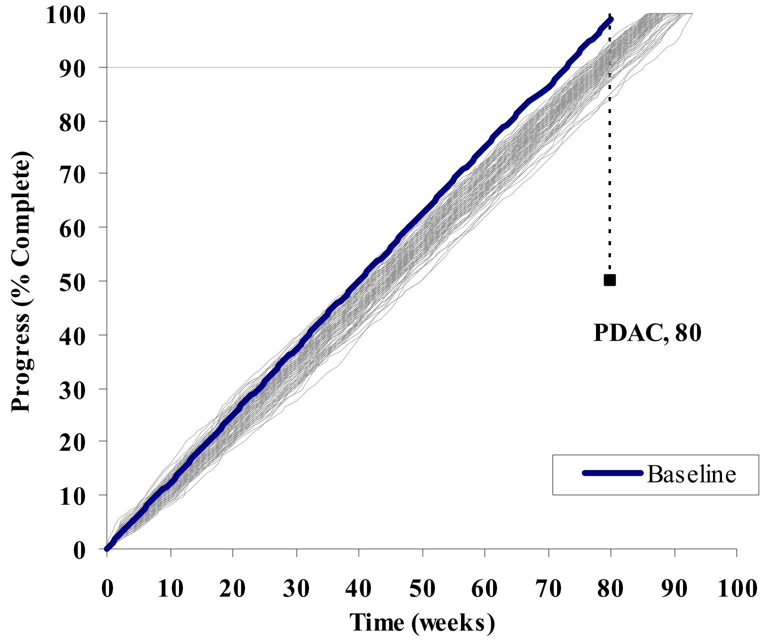


(c) KFF

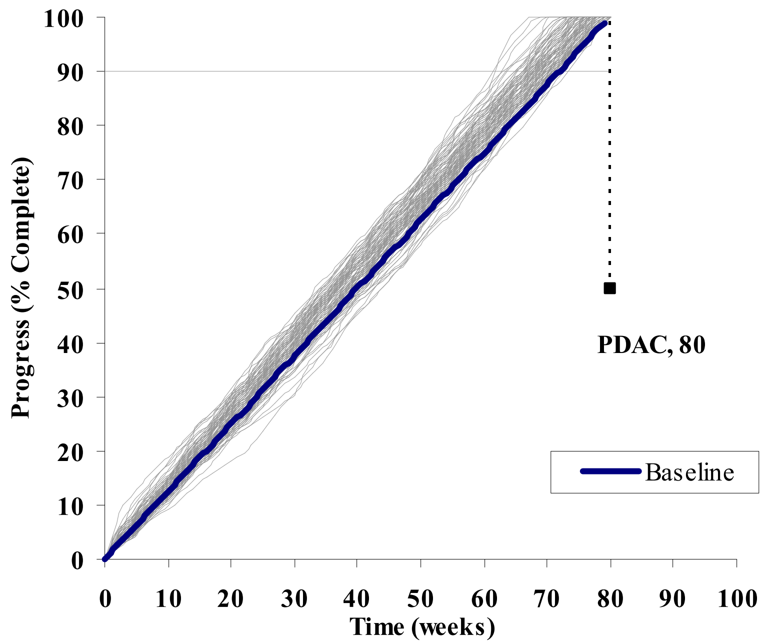


(d) BAF

Figure 5.21 Confidence intervals on PW under the underrun scenario

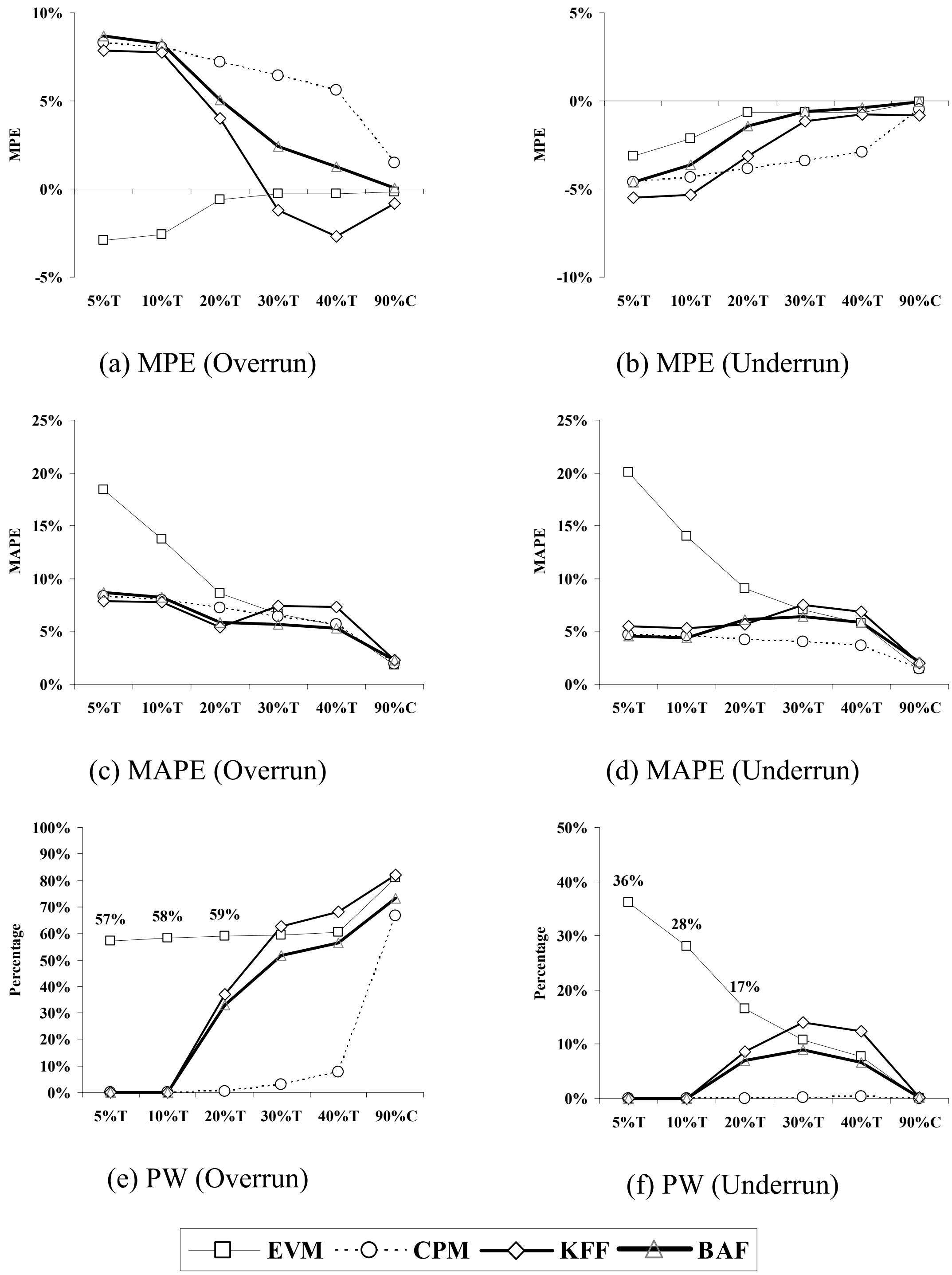


(a) Overrun scenario



(b) Underrun scenario

Figure 5.22 The baseline and random executions (first 100 executions for each scenario) of the linear project. The critical risk α used for the schedule control limit is 0.10



(a) MPE (Overrun) (b) MPE (Underrun)

(c) MAPE (Overrun) (d) MAPE (Underrun)

(e) PW (Overrun) (f) PW (Underrun)

Figure 5.23 Forecasting performances for linear projects

The forecasting performances of each method for projects with different network complexities are shown in Figure 5.24 through Figure 5.27 and the results can be summarized as follows.

- The EVM: In most evaluation points, the MPE tends to be smaller for a linear project than more complex projects. Especially, the MPE for a linear project under the overrun scenario is almost zero after the 20%T point (Figure 5.24 (a)). However, this does not necessarily mean that the EVM provides more accurate predictions for linear projects than more complex projects. As discussed in Section 5.6.2, the small MPE may result from the off-set between positive errors and negative errors. The results in Figure 5.24 (c) and (d) show that the MAPE for linear projects is not any better than those for other more complex projects. In the case of the probability of correct warning (Figure 5.24 (e)), the EVM shows higher PW for linear projects at 20%C, 30%C, 40%C, and 90%C. However, the influence of network complexity on the probability of false warning for projects finishing ahead of schedule is not obvious.

- The CPM: Under the overrunning scenario, the CPM shows better accuracy in both the MPE and the MAPE for linear projects. The improved accuracy leads to increase the probability of correct warning in Figure 5.25 (e). However, up to 40%T, the PW remains still less than 10 %. At 90%C, the probability of correct warning for linear projects is 67% which is about three times of the second highest PW for Project Group 3.
- The KFF: Forecasts by the KFF do not reveal any dominant pattern among different project groups in terms of MAPE and MPE. In the case of probability of correct warning (Figure 5.26 (e)), the KFF makes more correct warnings for linear projects than more complex projects. However, the graphs show that the probability of correct warning for Project Group 3 is higher than Project Group 1 and Project Group 2, which are less complex than Project Group 3.
- The BAF: Similar patterns observed with the KFF are also found in the results from the BAF. Again, no dominant pattern has been observed in forecasting performance of the BAF depending on the level of complexity of projects except for the probability of correct warning for linear projects.

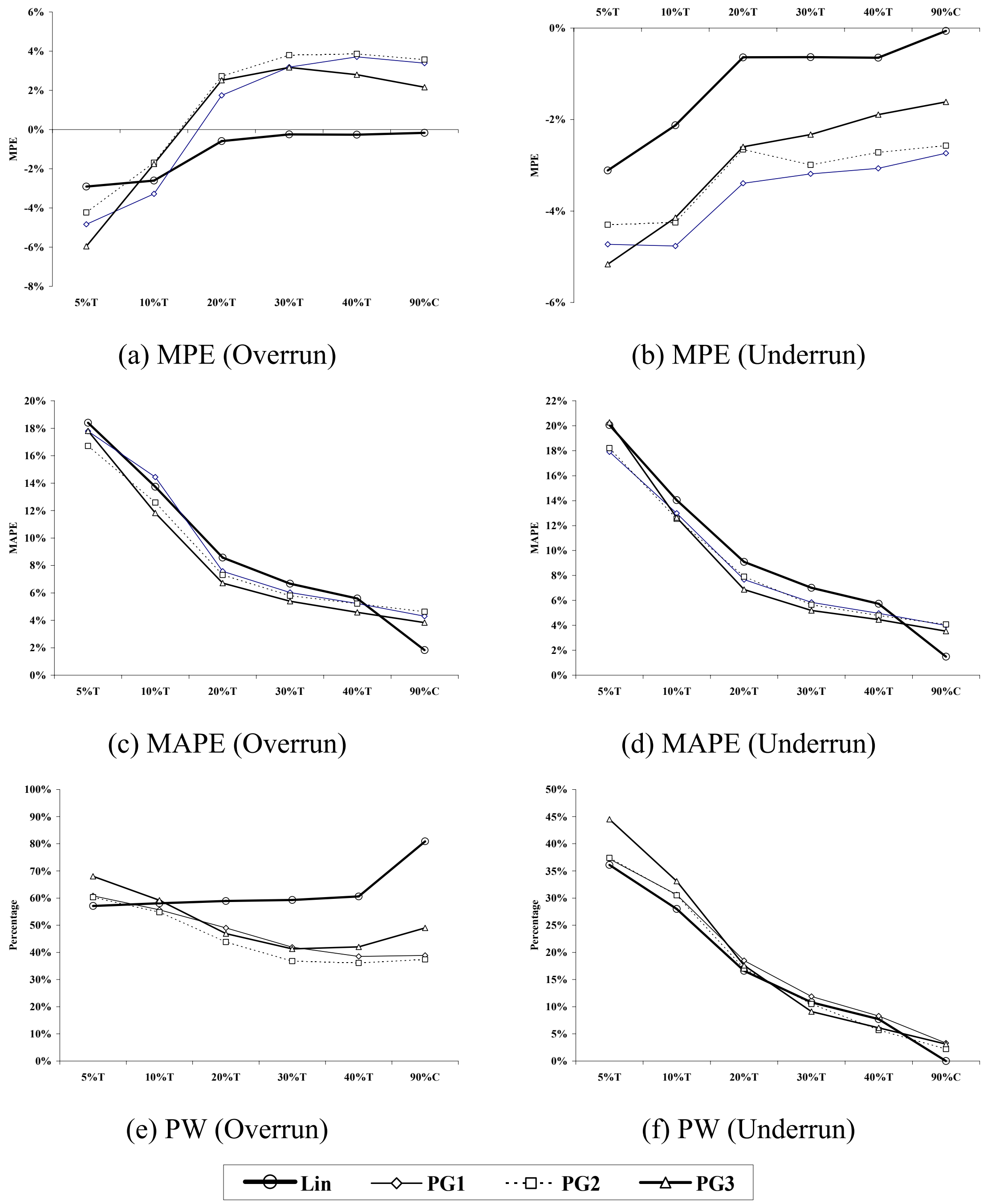


Figure 5.24 Forecasting performance of the EVM for projects with different network complexities (Note: PG denotes Project Group in Table 5.4)

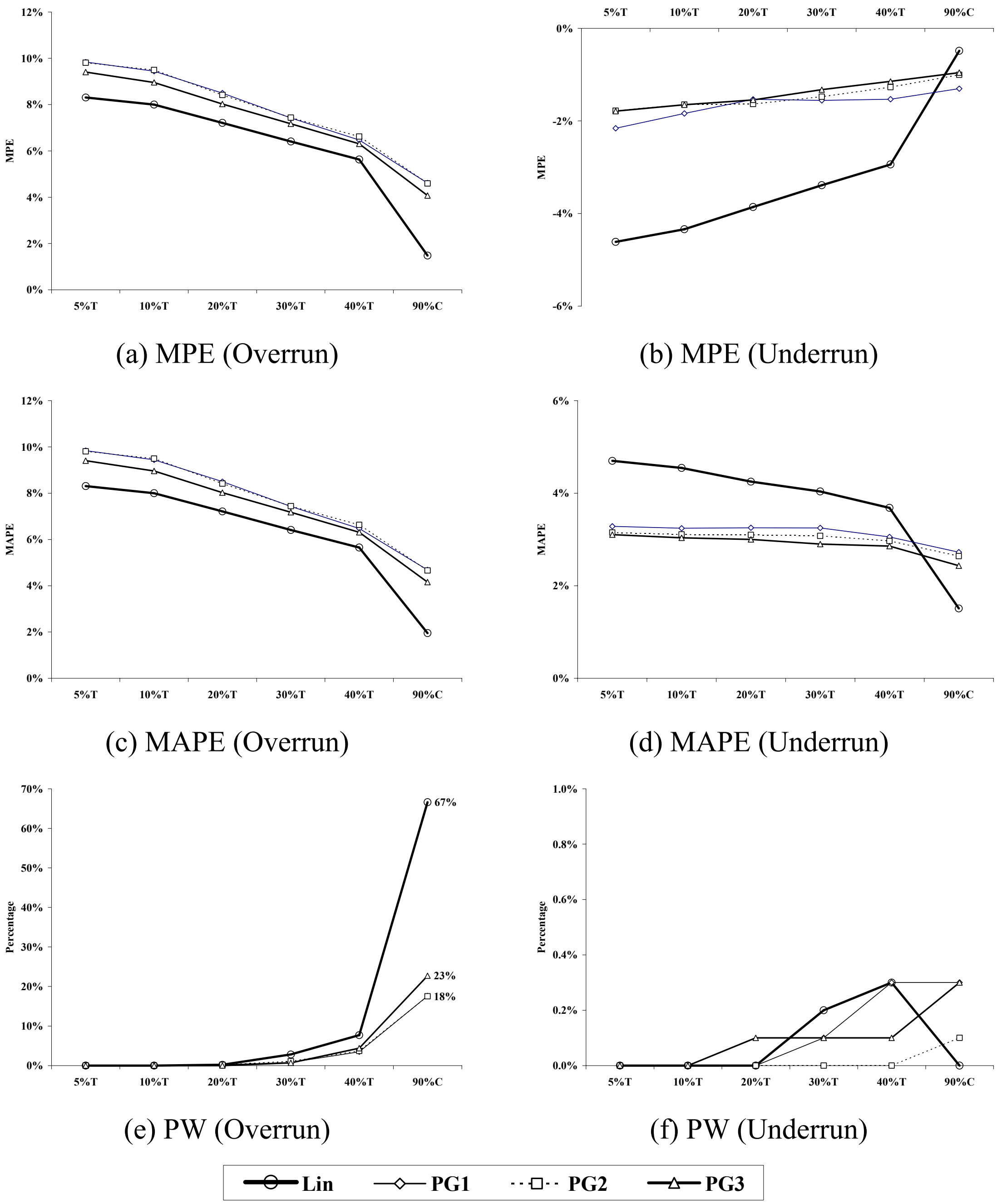
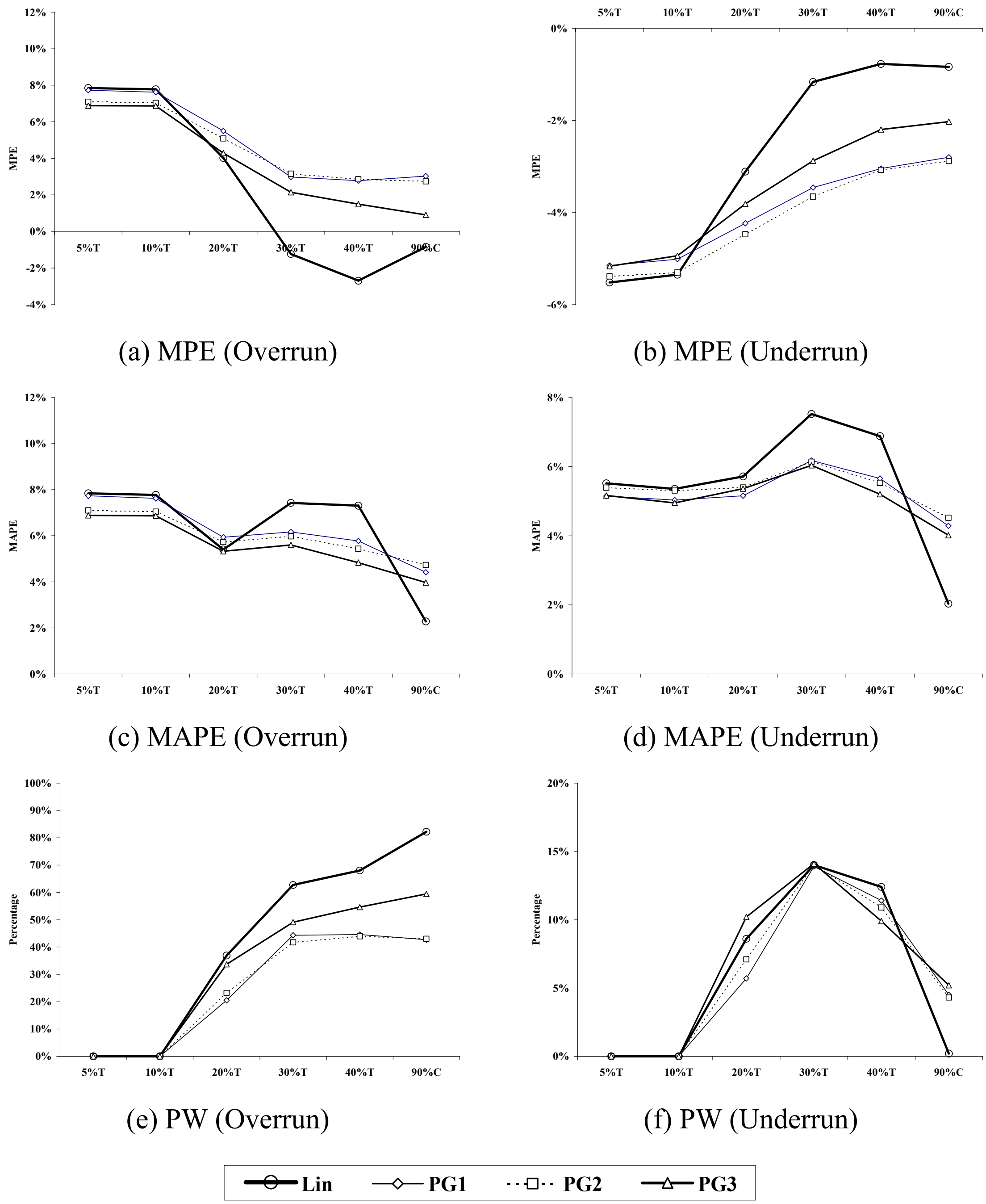


Figure 5.25 Forecasting performance of the CPM for projects with different network complexities



(a) MPE (Overrun) (b) MPE (Underrun)

(c) MAPE (Overrun) (d) MAPE (Underrun)

(e) PW (Overrun) (f) PW (Underrun)

Figure 5.26 Forecasting performance of the KFF for projects with different network complexities

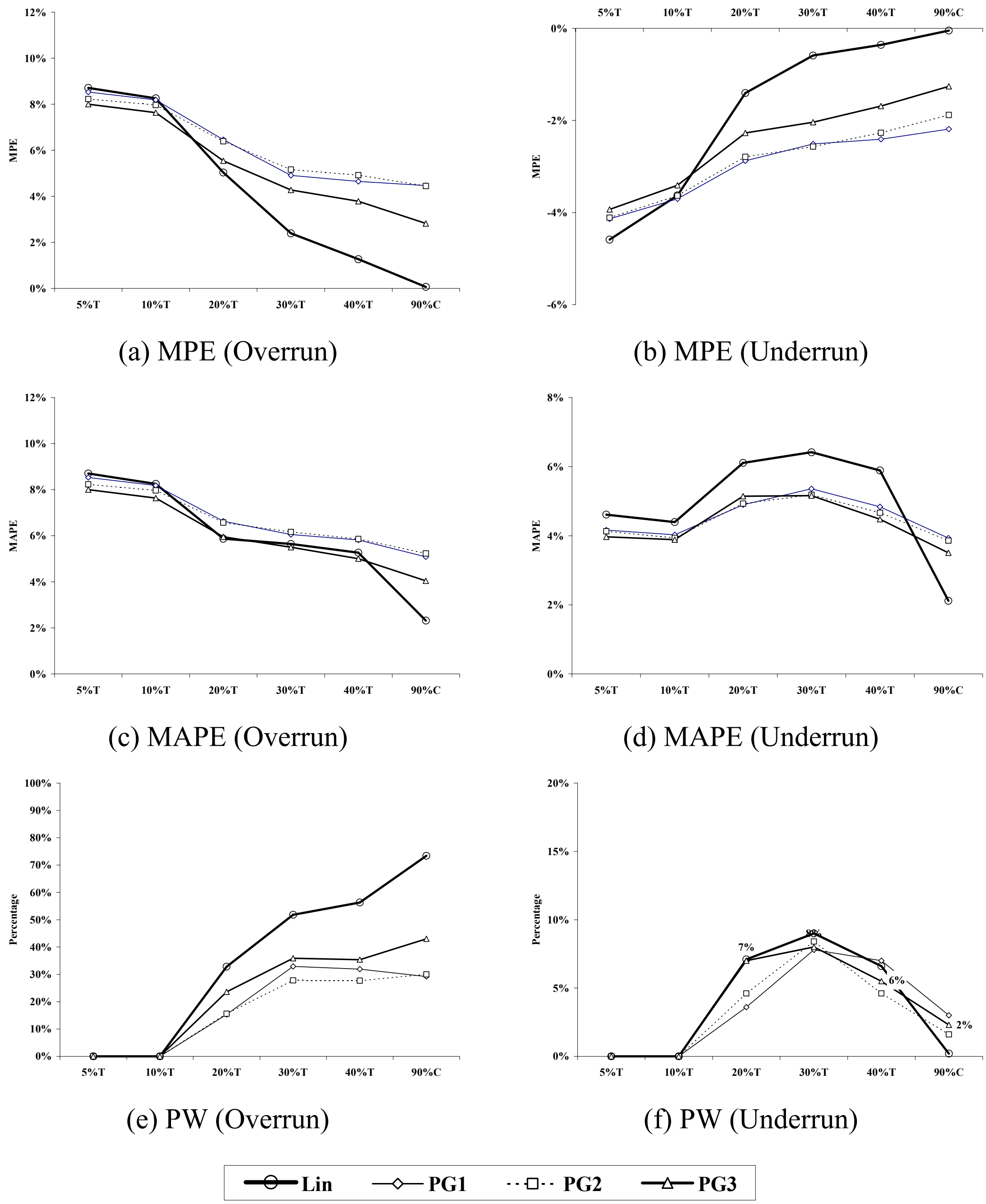


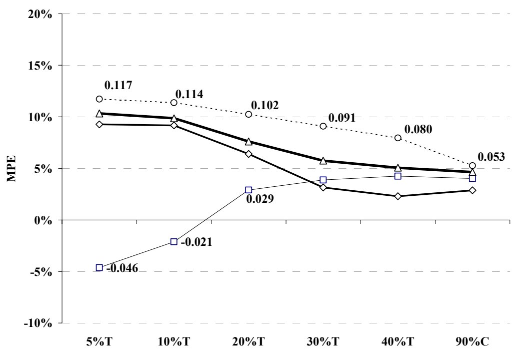
Figure 5.27 Forecasting performance of the BAF for projects with different network complexities

5.6.5 Effect of the Level of Critical Risk on Forecasting Performance

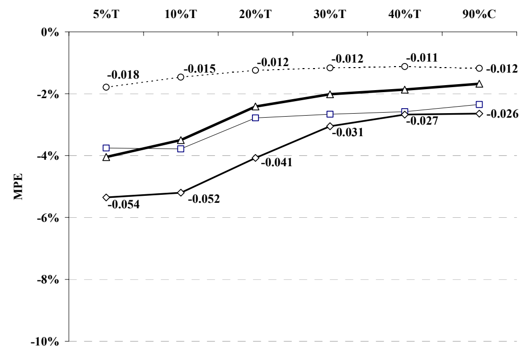
In the forecasts shown in the previous parts of this section, the critical risk level, α = 0.1, was chosen arbitrarily to determine the schedule control limit. Section 5.6.5 repeats the same analysis in sections 5.6.2 and 5.6.3 with a different value for α. The purpose of this test is to examine the potential influence of the critical risk level on the performance of forecasting methods.

To focus on more extreme cases – when project managers are more worried about severe delays than mild ones –, α = 0.03 is used to determine the schedule control limit. Major forecasting performance measures at different stages are shown in Figure 5.28.

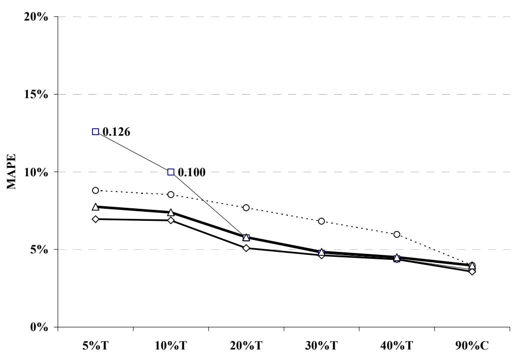
The results show that the change in the level of critical value, which must be determined by project managers according to their own attitudes toward risk, does not fundamentally change the forecasting performance patterns observed in previous sections.



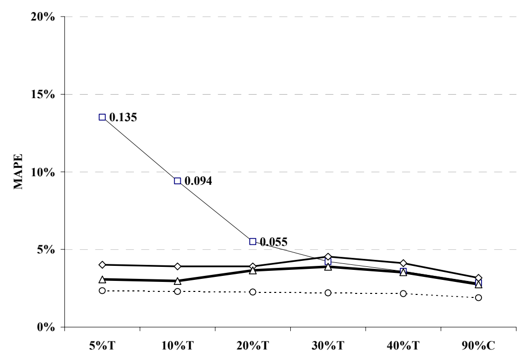
(a) MPE (Overrun)



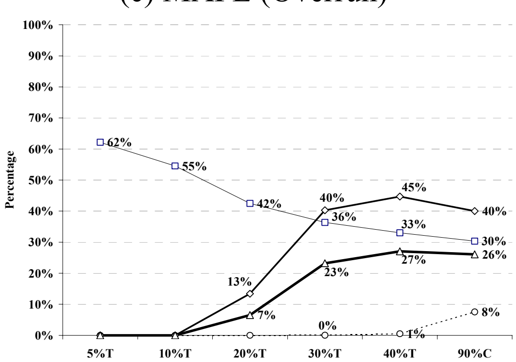
(b) MPE (Underrun)



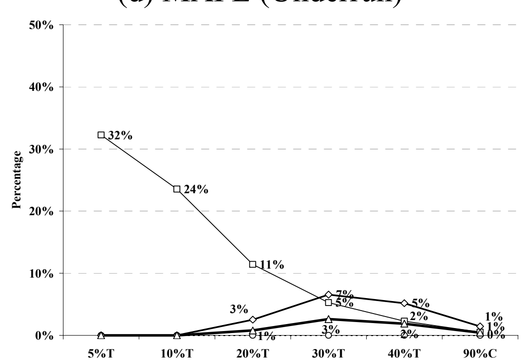
(c) MAPE (Overrun)



(d) MAPE (Underrun)



(e) PW (Overrun)



(f) PW (Underrun)

EVM CPM KFF BAF

Figure 5.28 Forecasting performances by different methods ($\alpha = 0.03$)

5.7 Test of Hypothesis 3

5.7.1 Test Design

The third hypothesis states as follows:

Hypothesis 3: The relative performance of the Kalman filter and Bayesian adaptive forecasting varies depending on the types of information available at the time of forecasting. For example, the information about the baseline plan is more useful in the Bayesian adaptive approach than in Kalman filter forecasting.

The Kalman filter forecasting method and the Bayesian adaptive forecasting method share several common attributes. First of all, both methods make use of prior performance information to supplement actual performance data. The major prior information required is the probability distribution of project duration and the planned cumulative progress curve. Both methods use actual performance data measured in terms of project level cumulative progress. In addition, both the KFFM and the BAFM yield probabilistic predictions with prediction bounds on single-point estimates.

However, from a methodological point of view, the two methods are formulated based on different premises. While the KFFM focuses on the dynamic nature of project performance, the BAFM relies on well-known statistical approaches such as regression methods, curve fitting techniques, and Bayesian inference. More specifically, the KFFM is based on the premises that the progress of a project can be represented by a set of dynamic equations with uncertainty and the true state of the project can only be measured with some inherent errors. The dynamic equations for a given project can be established with prior information from the project plan, historical data, or subjective judgment. In the BAFM, the same prior information is used to identify proper S-curve models and to construct the prior distributions of model parameters. Then, the probability distribution of each model parameter is revised with actual performance data through Bayesian inference. Therefore, even with the same actual performance data, specific predictions by the KFFM and the BAFM may differ from each other due to the differences in the use of information.

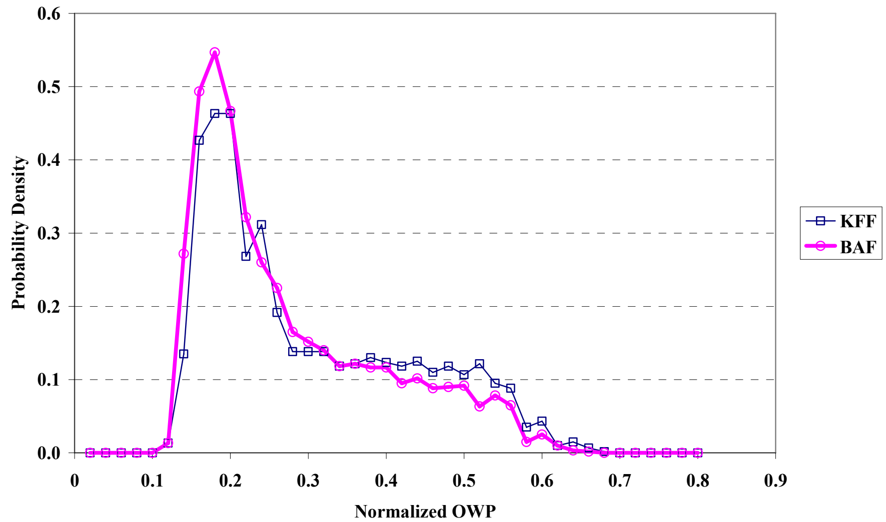
In this section, the relative performance of the KFFM and the BAFM is compared under two typical situations in real projects: when the baseline plan is available and when the baseline plan is not available.

5.7.2 The KFF and the BAF When the Baseline Is Available

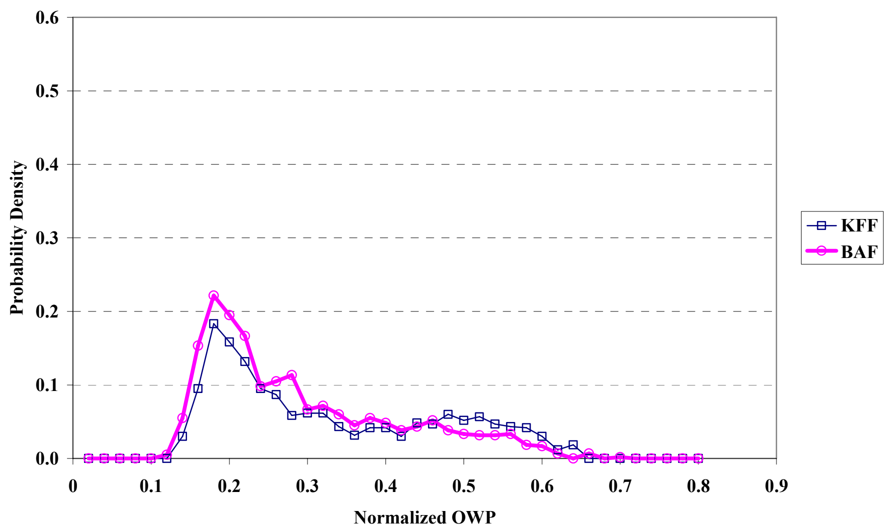
The forecasting performance of the KFFM and the BAFM when the baseline is available has already been discussed in Section 5.6. The results in Section 5.6, however, are limited to single-point predictions about the EDAC in order to make comparisons against deterministic methods – the CPM and the EVM – possible. Meaningful as they are, the results do not incorporate the probabilistic nature of both methods. In this section, a comparison between the KFFM and the BAFM is carried out with the overrun warning point (OWP) introduced in Section 5.4.3.

Using the data in Section 5.6, the overrun warning points from 3000 executions under each execution scenario have been calculated. The overrun warning points are determined at the 90% confidence level, which corresponds to a 90% probability of exceeding the planned duration at completion (PDAC). The probability distributions of the OWP from the KFFM and the BAFM are shown in Figure 5.29. Since each project has different PDAC, the normalized OWP – the OWP over the PDAC – is used in the results. The results indicate that both methods provide a similar pattern in issuing early warnings, regardless of the execution scenarios. During the first 10% of project duration, the probability of warning from both methods is almost zero. Thereafter the number of warnings increases and the probability of getting an overrun warning reaches a peak near 20% of the project duration. The probability distributions are highly skewed to the right. The most interesting thing is that the difference between the KFFM and the BAFM is very small.

The probability of getting an overrun warning at a specific time is also compared and shown in Figure 5.30. Under the overrun scenario, the probability of getting at least one overrun warning before a project reaches the 90% completion point is 72% and 69% for the BAFM and the KFFM, respectively. For underrunning projects, the probabilities are 13% and 10% for the BAFM and the KFFM, respectively. The results can be used effectively by project managers. For example, at 30% of project duration, the probability of getting a correct warning against the overrunning executions in the test is 49% by the BAFM. But there is still 8% chance of a false warning even when the project is underrunning.

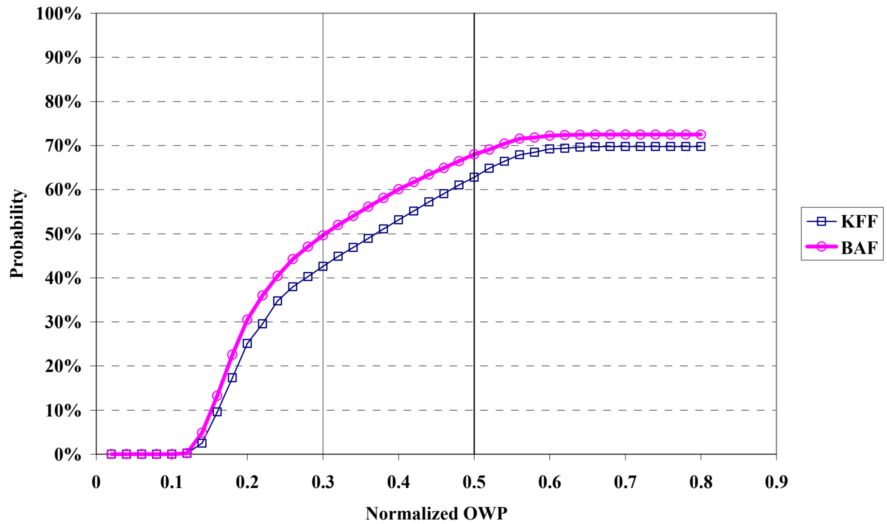


(a) Overrun

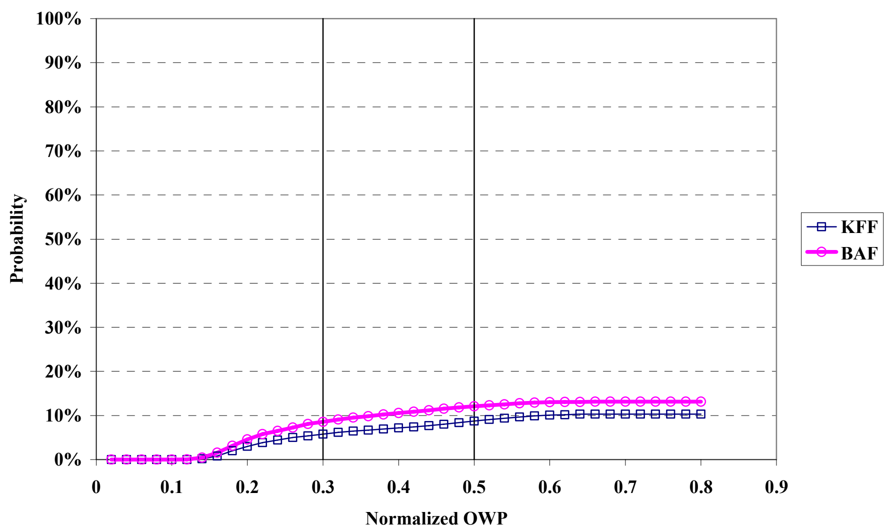


(b) Underrun

Figure 5.29 Probability distributions of the OWP from 3000 random executions



(a) Overrun



(b) Underrun

Figure 5.30 Probability of the OWP from 3000 random executions

5.7.3 The KFFM and the BAFM When the Baseline Is Not Available

The fundamental concept of the KFFM and the BAFM is to add additional value to the conventional project performance forecasting based only on actual performance data by making better use of prior performance information and by providing prediction bounds on predictions. Since prior performance information from various sources is represented by the baseline progress curve for a given project, the presence of a reliable baseline plan is crucial to elicit the best performance from the KFFM and the BAFM. At the very least, starting with a reliable plan to guide execution and control of a project is a must-do practice in project management.

However, even for a project in which such a baseline plan is not available, the new methods can still be useful. In the KFFM and the BAFM, the importance of a reliable baseline arises from the fact that it conveys the knowledge about the overall progress of a project, which is often not revealed until the closing phase of the project, by the actual performance data being generated by the project itself. Therefore, in the absence of the baseline based on specific plans for a project, other sources of prior information, such as historical data from similar projects and subjective judgment, may serve as alternatives. For example, Figure 5.31 shows the best fitting parameters of the BetaS-curve for the thirty artificial projects used in Section 5.6 (Table 5.4). The means of α and m from the sample are 2.55 and 0.3, respectively. Then, this information can be useful for another project that is believed to share some similarity in progress pattern with the 30 projects in the past.

From a methodological point of view, the BAF approach provides a project manager with more flexibility in approximating various forms of prior performance information. From the historical data in Figure 5.31, for instance, appropriate probability distributions of α and m can be used instead of the mean values.

In the case of the KFFM, options for using rather uncertain baseline information are limited. Fundamentally, the predictive power of the KFF approach lies in the accuracy of the system model which predicts the future state of a project based on the current state. The measurement model is used to correct the prediction according to the relative uncertainty included in the prediction and the measurement. In the absence of a baseline progress curve based on a detailed project plan, the system model should be established based on alternative sources such as historical data and subjective judgment.

5.8 Comparison of Forecasting Methods

From the results discussed in the previous parts of this chapter, a brief comparison of the four forecasting methods is shown in Table 5.5 for basic properties and Table 5.6 for forecasting performance.

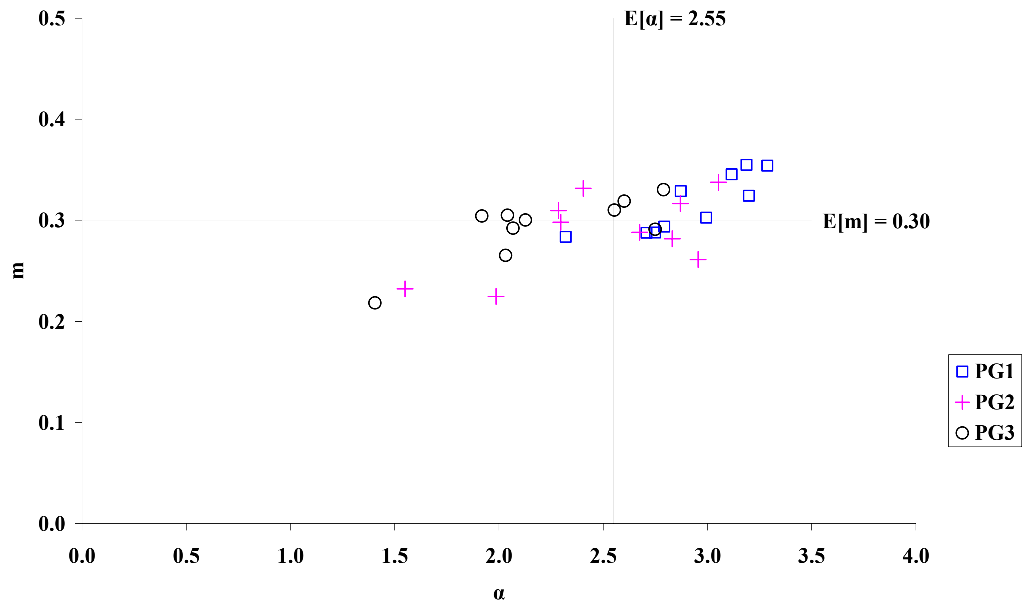


Figure 5.31 The best fitting parameters of the BetaS-curve for the projects used in Section 5.6

Table 5.5 Comparison of forecasting methods: Basic properties

Criteria	CPM	EVM	KFFM	BAFM
Major properties	- Activity level control - Retrospective	- Project level control - Linear extrapolation	- Project level control - Probabilistic prediction - Use of prior information	- Project level control - Probabilistic prediction - Use of prior information
Input requirements	- Activity network - Activity estimates	- The planned value - The earned value - The actual cost	- Prior probability distribution of project duration - The baseline curve - The actual progress	- Prior probability distribution of project duration - The baseline curve - The actual progress
Limitations	- Only for schedule	- For schedule and cost	- For schedule and cost	- For schedule and cost
Applicability	- Only for projects with a network schedule available	- Universally applicable to all projects of all types, sizes, and complexities.	- Universally applicable to all projects of all types, sizes, and complexities.	- Universally applicable to all projects of all types, sizes, and complexities.
Ease of implementation	- Activity level knowledge is required - Activity level updates - Commercial software	- Project level update - Simple formulas based on three variables	- Project level update - Simple formulas based on three variables - Basic knowledge about probabilistic forecasting	- Project level update - Simple formulas based on three variables - Basic knowledge about probabilistic forecasting
Ease of communication	- Difficulty increases with the number of activities	- Understanding of the three basic variables is required	- Understanding of EVM and probabilistic forecasting is required	- Understanding of EVM and probabilistic forecasting is required

Table 5.6 Comparison of forecasting methods: Forecasting performance

Criteria	CPM	EVM	KFFM	BAFM
Accuracy<1> (Start~20%T) [F.5.15-F.5.18]<2>	□□□□□ CPM updates prediction based only on actually observed performance.	□□□□□ EVM is best in MPE but worst in MAPE due to a large variability in predictions.	□□□□□ KFFM provides the most accurate predictions in both MPE and MAPE.	□□□□□ KFFM is slightly better than BAFM, but both are better than CPM in both MPE and MAPE.
Accuracy (30%T~90%C) [F.5.15-F.5.18]	□□□□□ CPM updates prediction based only on actually observed performance.	□□□□□ EVM gets stabilized and provides as accurate predictions as KFFM and BAFM in both MPE and MAPE.	□□□□□ KFFM provides more accurate predictions than early stages. Difference in KFFM and BAFM is negligible.	□□□□□ BAFM provides more accurate predictions than early stages. But difference in BAF and KFF is negligible.
Timeliness of warning [F.5.19-F.5.21]	□□□□□ CPM provides almost no warnings up to the 30%T and, even at the 90%C, the PW is mere 19%.	□□□□□ EVM provides early warnings but EVM can be applied only after the performance has stabilizes.	□□□□□ Up to the 10%T, no warnings come from the KFFM. After that point, the PW increases, reaching 45% at the 30%T.	□□□□□ After the 10%T, the PW starts to increase and reach 32% at the 30%T.
Reliability of warning [F.5.19-F.5.21]	□□□□□ The likelihood that CPM provides a false warning is almost zero.	□□□□□ The likelihood that EVM provides a false warning is about 40% at the 5%T.	□□□□□ The probability of false warning increases to 14% at the 30%T and starts to decrease.	□□□□□ The probability of false warning increases to 8% at the 30%T and starts to decrease.
Flexibility	□□□□□ Deterministic results.	□□□□□ Predictions can be adjusted with various performance factors that can be chosen by a user.	□□□□□ Sensitivity of prediction to the actual data can be adjusted by the measurement error term.	□□□□□ Sensitivity of prediction to the actual data can be adjusted by the likelihood variance term.

(Note <1> Accuracy is compared based only on the results under the overrun. <2> 'F' abbreviates 'Figure'.
<3> □□□□□(Very Poor); □□□□□(Poor); □□□□□(Medium); □□□□□(Good); □□□□□(Very Good))

5.9 Chapter Summary

In this chapter, a series of parametric studies has been carried out to evaluate, in a statistically meaningful way, the forecasting performances of the Kalman filter forecasting method in Chapter III and the Bayesian adaptive forecasting method in Chapter IV. In addition, a comparison of the new methods against two conventional methods – the earned value method and the critical path method – has been made.

To obtain statistically meaningful results, a new forecasting performance evaluation framework is proposed. The evaluation framework in this research can be characterized in two ways. First, it relies on a large set of artificial project data rather than limited real data selectively chosen from past projects. Several limitations in using real data for evaluation of project performance forecasting methods have been identified (Section 5.1.2) and the process of generating random project performance data is proposed based on a random network generation method from literature, which is improved with a redundancy elimination technique (Section 5.2). With the artificial project generation technique, a large set of project data can be generated to draw statistically meaningful conclusions. Furthermore, since artificial data are independent from the forecasting methods under comparison, it is possible to compare forecasting performance across different methods. Another characteristic of the evaluation framework is that the timeliness and reliability of a forecasting method are directly measured in terms of the warning instead of accuracy-based indicators. This was possible by incorporating an early warning system into the evaluation process. In Section 5.3, two typical early warning systems in project management have been reviewed. Furthermore, new concepts such as the overrun warning point and the probability of warning under different execution scenarios have been proposed and seamlessly integrated into the evaluation framework.

Using the new forecasting performance evaluation framework, the three research hypotheses have been tested and addressed in Section 5.6, Section 5.7, and Section 5.8, respectively. Based on the results from the hypothesis tests, basic properties and relative forecasting performance of the two new methods in this research are compared with the earned value method and the critical path method (Section 5.8).

In closing, it should be mentioned here that the comparison made in this chapter should not be considered *deterministic* – one method is better than the other with certainty – or *exclusive* – one method should be used and another method should be discarded. The dynamics and diversity of real project management still demand project managers to make their own decisions based on their best knowledge under specific conditions. Furthermore, better decisions often are made based on a combination of different approaches rather than a single model, presumably the best model, chosen for some reason. For example, according to the National Hurricane Center Forecast Verification Report 2006, consensus models that combine results from other hurricane forecasting models outperform individual models (Franklin 2007).

CHAPTER VI

IMPLEMENTATION OF THE KFFM AND THE BAFM

6.1 Introduction

Ease of implementation is an important factor for the selection of forecasting methods. Yokum and Armstrong (1995) conducted a study of criteria for selecting forecasting methods and examined the differences in selection criteria according to the various roles of forecasters, such as decision maker, practitioner, educator, and researcher. The study shows that, although the respondents ranked accuracy as the most important criterion, other criteria such as ease of implementation, ease of use, ease of interpretation and flexibility are also comparable in importance. Decision makers put significantly higher weights on flexibility and ease of implementation than did the other three groups of forecasters.

This chapter includes some recommendations and guidelines for the project manager and other potential users regarding the implementation of the KFFM and the BAFM in their projects. In Section 6.2, some recommendations for project managers and other potential users are suggested. In Section 6.3, practical guidelines for implementing the KFF and the BAF methods in real world projects are presented. The purpose of the guides is to help project managers and other potential users build better understanding of and insight into the new methods, so that they apply those methods to their projects efficiently.

6.2 Recommendations for Project Managers

6.2.1 General Attitudes of the Project Manager

In using the KFF and the BAF methods, the project manager should be aware of and try to practice the following recommendations about the general attitudes regarding project performance forecasting.

- Take part in the forecasting as an information provider: The project manager is not merely an end user of the forecasts provided to him or her. It is important for the project manager to take part in the planning and forecasting process actively, hopefully, from the very beginning. Unlike the CPM and the EVM, in which forecasts are made, in most parts, with limited, fixed data, the KFFM and the BAFM allow forecasters to utilize additional performance information from various sources such as project plans, historical data, and subjective judgment. One of the sources of relevant information is the knowledge and experience of the project manager.
- Acquaint yourself with the new methods: The more you understand a tool, the more effectively you can use its outcome. The project manager may not have to be an expert in Kalman filter, Bayesian inference and other theories or techniques used in the formulation of the method. However, understanding the basic features of the methods is important for the project manager to have more confidence in the forecasts, to make better decisions based on them, and to communicate them more effectively.

- Use forecasting methods as a tool for decision making: The purpose of project performance forecasting is not to foresee the *exact* final status of a project. The ultimate goal is to get reliable confirmation of success or an early warning about performance failure so that effective and timely actions can be made to *avoid* problems in the future. In that sense, successful forecasts for a project can not be verified by the project's result. Therefore, the best use of forecasting methods can only be realized by using them reasonably and within a rational decision making framework. For example, it is important to establish a formal early warning system in which forecasts, possibly from different methods, are used.
- Avoid exploiting the flexibility of methods to see what you want to see: The flexibility in the KFFM and the BAFM should be used in the way that makes the best use of all relevant, unbiased performance information available in a specific situation. It should not be used as a way of getting what the project manager wants to get.

6.2.2 Selecting the Right Method

To help the project manager decide which methods will be appropriate in a particular project situation, a table is developed based on the results in previous chapters. The forecasting methods selection table in Table 6.1 evaluates the four forecasting methods compared in Chapter V, which are the KFFM, the BAFM, the CPM, and the EVM, against four important criteria described below.

Table 6.1 Forecasting method selection table

Criteria	Options	Methods			
		KFFM	BAFM	EVM	CPM
Types of responsibility	Project manager	●	●	●	●
	Program manager	●	●	●	
Types of forecasts	Probabilistic forecasts	●	●		
	Single-point forecasts	●	●	●	●
Major concern of forecasts	Timeliness of warning	●	●	●	
	Reliability of warning	●	●		●
	Timeliness and reliability of warning	●	●		
Properties of the baseline progress curve	A single curve with properly fitted S-curve models	●	●	●	
	A single curve without S-curve models	●		●	
	A group of curves based on historical data or subjective judgments		●		

Criterion 1. Types of responsibility

While a project manager is in charge of a single project, a program manger needs to coordinate a group of related projects to achieve the program's objectives. For project managers, all the four methods in comparison may be found useful. However, for program managers, the CPM is not recommended because of two reasons. First, CPM requires an activity-level understanding of the project. Unlike project managers who may have a thorough understanding of the project at hand, program managers may not possess in-depth knowledge of the projects, for example, the precedence constraints between activities due to technology, practice, efficiency, and other factors such as, how many activities are on the critical path and what are the status of those critical activities. Second, even when the program manager is capable of handling the details of individual projects, it would be time-consuming. Tracking actual progress of multiple projects and making forecasts should be carried out periodically and dealing with large amount of very detailed information in CPM would impede the overall program management process.

Criterion 2. Types of forecasts

Uncertainty is an important factor in planning and making decisions about the future progress of a project. Only the KFFM and the BAFM provide prediction bounds that represent the range of possible outcomes at a given confidence level. As discussed repeatedly throughout this dissertation, probabilistic forecasts are to be preferred against deterministic forecasts. There are at least two major reasons. First, probabilistic forecasts (KFFM or BAFM) convey information about the accuracy of the forecasts that is not conveyed in deterministic (EVM or CPM) forecasts. The accuracy of the forecasts is represented as prediction intervals, from which the users may obtain estimates of risks associated with the forecasts and be able to answer some key questions such as "What is the risk or probability of exceeding the planned completion date?" and "What is the expected completion date?" Deterministic methods provide no variances of forecasts and no estimates of risk of overrun.

Another reason that probabilistic forecasts should be preferred is that, contrary to popular belief, single-point forecasts by deterministic methods do not represent expected values of the forecasts. For example, the basic earned value method for forecasting the estimated cost at completion (EAC) is given as

$$EAC = \frac{BAC}{CPI} \tag{6.1}$$

where BAC is the budget at completion and CPI is the cost performance index. Assume that the CPI in this equation is a random variable with mean μ_{CPI} and variance σ^2_{CPI}. Then the expected value of EAC can be approximated in terms of the Taylor series expansion up to the second order term (Ang and Tang 2006; Howard 1971).

$$\mathrm{E}\left[EAC\right] \cong \frac{BAC}{\mu_{CPI}} + \frac{BAC}{\left(\mu_{CPI}\right)^3}\sigma^2_{CPI} \tag{6.2}$$

Equation (6.2) shows that even if the expected value of CPI is used in Equation (6.1), the resulting value is not the expected value of EAC because the second term in Equation (6.2) is ignored, let alone the higher-order terms ignored in the Taylor series expansion. This example shows the case when a deterministic forecasting method gives a prediction, but one does not know what the prediction means.

Criterion 3. Major concern of forecasts

Getting reliable warnings, sooner than later, is the major concern of project performance forecasting. The parametric studies in Chapter V, however, reveal that no method under comparison outperforms others simultaneously on all criteria, such as the timeliness and the reliability of overrun warnings. Therefore, a trade-off between the timeliness and the reliability of warnings may be necessary in selecting adequate forecasting methods.

A simple way of doing the trade-off analysis is to compare the cost of errors in a specific project situation. In project performance forecasting, two types of warning errors can be defined according to the combination of forecasts and actual performance: the warning failure error and the false warning error. Table 6.2 shows the combination. The cost of false warning error can be defined as the monetary or non-monetary impact from a false warning on the performance of a project. For example, when a red flag is raised against a significant schedule overrun, the first thing for the project team to do is to authenticate the warning by investigating the root causes of the warning. Even when the warning turns out to be a false alarm, results from such investigations should be documented and reported. On the other hand, the cost of warning failure error is a kind of opportunity cost that might have been saved by taking additional preventive actions in a timely manner if there were a correct warning.

Table 6.2 Types of error in project performance forecasting

Forecasts	Actual results	
	Overrun	Underrun
Overrun Warning	Correct warning	False warning error
No Warning	Warning failure error	Correct no-warning

Obviously, it would not be easy to assess these two types of error costs and to compare them in individual projects. First of all, there is no accepted method that can be used systematically to make a quantitative evaluation. The results in Chapter V, in that sense, can be useful in selecting appropriate methods and interpreting results from different methods. For example, warnings from the EVM are very unreliable early in a project. The reliability of warning improves over time by collecting more data, but it would be difficult to assess the reliability of EVM for individual projects, because it is deterministic. Therefore, the EVM is not recommended for such projects in which a warning entails some formal investigation that might impede the proper execution of work or erode the credibility of the forecast. However, if the project manager's top priority is to detect every possible symptoms of an overrun, the EVM will be more useful than other methods. On the contrary, the CPM provides extremely reliable warnings but its early-warning potential is literally zero.

The KFFM and the BAFM provide viable alternatives to the EVM and the CPM. These methods should be used when a balance between the timeliness and reliability of warnings is required. Furthermore, it should be pointed out that the KFFM and the BAFM allow the user to incorporate subjective information in terms of the prior distribution of project duration and the measurement errors. This flexibility can be used to adjust the sensitivity of forecasts to the actual data, which eventually influences the reliability and timeliness of warnings.

Criterion 4. Properties of the baseline progress curve

A baseline progress curve is a primary input for KFFM, BAFM, and EVM. However, some general properties of the baseline curve may influence the selection of the proper method. For example, when historical data or subjective judgments are used to construct a group of possible progress curves, the BAF method can be applied effectively because of its flexibility in dealing with a range of different progress shapes. Once a single progress curve is used as the baseline, the BAF method works well when the baseline curve is reasonably approximated with some S-curve models.

6.3 Quick Implementation Guide of the KFFM and the BAFM

6.3.1 Basic Features

Some basic features of the KFFM and the BAFM are summarized below.

- The KFFM and the BAFM are probabilistic methods. First and foremost, the inputs – the prior performance information and the actual performance data – are represented in terms of probability distributions. In the KFFM, the prior probability distribution of project duration is used to determine the size of the process noise term and a single-point estimate of the actual performance at a specific time is combined with the measurement error term into a probability distribution. In the BAFM, the prior probability distribution of project duration is updated with the actual performance data through the Bayesian updating process.

- The KFFM and the BAFM can be used for any project, regardless of its size, type, and complexity. The actual performance data used in both methods are cumulative progress measures which may be constructed in various forms such as percentage of completion, labor hours, and cumulative costs. As a result, the new methods are free from the differences in sizes, types, and complexities in individual projects. EVM is a typical project management system that measures actual performance in terms of a cumulative progress curve.
- The KFFM and the BAFM use prior performance information to supplement the lack of predictive power of small samples of actual performance data early in a project. The best information to be used for forecasting is the actual performance data observed in the project itself. However, actual performance data early in a project are too few to reveal significant information about the future progress of a project. The new methods overcome this problem by using prior performance information based on project plans, historical data, or subjective judgment.
- The KFF method focuses on building more refined models of the dynamic nature of projects and on the uncertainty in measurements. In addition, forecasts by the KFF method are optimal in a sense that they minimize the sum of the squares of prediction errors.
- The reliability of the BAFM is influenced by the degree of match between the progress curve template and the actual progress curve. The degree of match between

these two curves represents the predictive power in the progress curve template. A practical way to improve the predictive power is to establish the progress curve template of a project based on a realistic project schedule.

6.3.2 Input Preparation for the KFFM

The KFFM forecasts using four input elements: the prior probability distribution of the project duration, the planned cumulative progress curve, the size of measurement error, and the actual cumulative performance data. Figure 6.1 (a) shows the KFF method input elements.

[KFF Input-1] The planned progress curve

The planned progress curve or the baseline represents the amount of work to be done at each specific time. The baseline of a project as shown in Figure 6.1 (a) can be generated from a value loaded project schedule. The value here is chosen as the quantity that best represents the overall progress of the project in the most efficient way. In earned value management, the budgeted cost of each activity is used as the measure of the activity's value. However, other quantities such as engineer hours, labor hours and percentage of work can be used as long as they represent the cumulative progress of the project. Once the values of individual activities in a project are evaluated, they are distributed over the time horizon according to the starting dates and the finishing dates in the plan.

Commercial programs such as Primavera Project Planner® and Microsoft Office Project® provide functions to automatically generate the planned progress curve. However, a simple bar chart can also be used effectively.

[KFF Input-2] The prior probability distribution of project duration

The prior probability distribution of the project duration is used to determine the uncertainty associated in the process model. That is, with a given level of uncertainty in the process model, the project duration forecast by the KFF method before the inception of a project should be consistent with the prior probability distribution of the project duration.

Project duration is assumed to be normally distributed by the central limit theorem and its mean and variance can be determined in many ways such as network-based CPM simulation, three-point estimates, and subjective estimates of the first and second moments. The simulation approach has been discussed in Chapter II. Many approaches have been proposed to approximate subjective distributions of random variables (Keefer and Bodily 1983; Moder and Rodgers 1968; Perry and Greig 1975). For example, range estimating or three-point estimating provides a simple but effective way of representing the user's uncertainty (Curran 1989; King et al. 1975; Oberlender 2000).

[KFF Input-3] The actual performance data

Actual performance should be constantly monitored and measured periodically. The actual cumulative progress at a specific time is estimated by aggregating activity level progress over each reporting period and summing all periodic progress in the past. Because of the periodical nature of the reporting cycle, the actual progress data are estimated in a discrete manner as shown in Figure 6.1 (a).

[KFF Input-4] The variance of measurement error

The last input element required to run the KFF method is the variance of measurement error, which represents the uncertainty in reported performance data. Estimating the measurement error is probably the trickiest aspect of the implementation of the KFF method. Measurement errors in project management arise from, roughly, two sources: the error in measuring time to report and the error in measuring performance at the reporting times. Most projects report progress according to some calendar-based periods such as weeks and months. However, the net work hours during each reporting period may vary from one period to another. For example, the number of days in a month varies from month to month. In addition, most performance data after each reporting period are subject to a certain amount of errors. For example, project level progress in any quantity of interest is estimated, in most cases, as a sum that aggregates various activities with diverse physical and managerial properties, whose progress is measured by individual task managers who may have different scales for the amount of work actually done. Currently, no viable methods for measuring these errors with actual data have been reported. In this dissertation, therefore, it is recommended to estimate the variance of measurement error by subjective judgments.

It should be noted that the measurement errors in the KFF formulation are estimated in terms of the time variation, that is, the deviation of the earned schedule from the forecasting time *t* (Figure 6.1 (a)). However, in most projects, project progress is measured at fixed times rather than the time at which a certain amount of work is accomplished. Naturally, project managers or other project engineers with experience would feel more comfortable with making range estimates over the potential errors in the amount of actual work done at a specific time. In such cases, a probability distribution of the measurement error in progress at time *t* should be converted to the probability distribution of earned schedule as shown in Figure 6.1 (a).

A simple but efficient way of doing this transformation is using a simulation approach. Given the baseline curve and the distribution of the measurement error in progress, a large set of random values can be drawn from the measurement error distribution and corresponding earned schedules can be calculated with Equation (2.3) in Section 2.4.1.

6.3.3 Input Preparation for the BAFM

In spite of different roles and interpretations within individual methodologies, conceptually the same elements discussed for the KAF method are also used in the BAF method. This section addresses the input elements focusing on the differences between the KFFM and the BAFM.

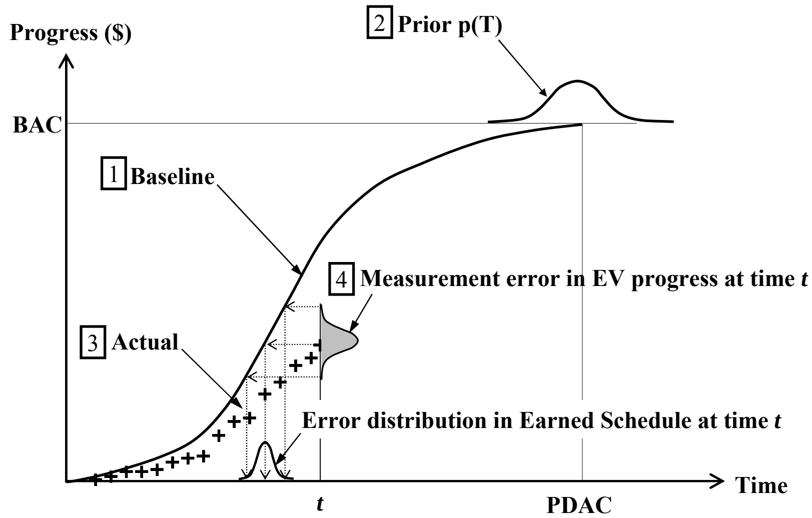
[BAF Input 1] The progress curve template

The same approach used for the KFFM can be applied to generate the planned progress curve. It should be noted, however, the BAF method provides a more robust framework than the KFF method for the development of the progress curve template. For example, the BetaS-curve model allows users to account for uncertainty in the progress curve template by imposing probabilistic shape parameters instead of fixed values.

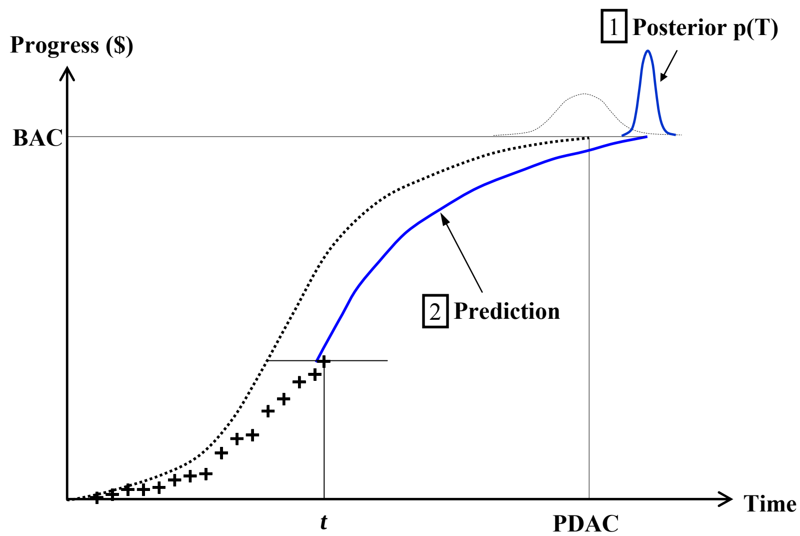
[BAF Input-2] The prior probability distribution of project duration

In the BAF method for schedule forecasting, the project duration is considered a random variable to be updated as a project progresses. Unlike the KFF method, the prior probability distribution of the project duration can be approximated with a non-normal distribution as long as random numbers can be drawn efficiently from the distribution. The prior probability distribution of project duration can be obtained in many ways using analytical tools or subjective judgment. The range estimating techniques for the KFFM can be used.

When selecting the prior probability distribution of project duration, special caution should be taken to ensure that the prior distribution includes a proper range of plausible outcomes. The BAF method *updates*, rather than independently estimates, the prior distribution of project duration in light of new actual performance data. Any new information is used to adjust prior distribution. Therefore, if the prior distribution is strongly biased from the actual project duration, using the prior information results in impeding the convergence to the right value, or even ending up a wrong one.



(a) Inputs



(b) Outputs

Figure 6.1 Inputs and outputs of the KFFM and the BAFM

[BAF Input-3] The actual performance data

The same data generated for the KFF method can also be used in the BAF method. Please refer to Section 6.3.2.

[BAF Input-4] The variance of measurement error

The measurement error in the BAF formulation for schedule forecasting is defined as the horizontal deviation between the actual progress curve and a progress curve that is generated by some S-curve model. In the BAF method, the standard deviation of measurement error is used for calculating the probability of observing the actual progress curve, conditional on some parameters (Please refer to Equation (5.15) in Section 5.3.2). In practice, the variance of measurement error should be determined based on subjective judgment for the same reasons addressed in the discussion of the KFF method. In addition, it should be noted that the variance can be assumed to be not constant.

6.3.4 Results Interpretation

The new methods provide forecasting results in two forms: the probability distribution of EDAC and the expected progress curve.

[Output-1] The posterior probability distribution of the project duration

The primary output from the KFFM and the BAFM for schedule forecasting is the posterior probability distribution of the project duration (Figure 6.1 (b)), which is represented in terms of the mean and the variance and may be represented by a normal distribution. Given these results, the project manager is able to make decisions based on the uncertainty rather than single-point forecasts. For example, the duration-at-risk in Section 5.3 may be used.

[Output-2] The future progress

An advantage of the KFFM and the BAFM over the EVM is that it provides an updated progress curve that leads to the expected project duration. With this curve, the project manager may gain better insight into the forecasts at completion.

6.4 Chapter Summary

Based on the results in previous chapters, a brief implementation guide for potential users of the KFF and the BAF methods was presented. Ease of implementation is an important factor that influences a decision maker's choice of forecasting methods. The objective of this chapter is to help potential users develop a better understanding of the core features of the new methods in this dissertation and apply them in their projects effectively.

This chapter addresses some practical issues. In Section 6.2, some useful recommendations for the project manager have been suggested about appropriate attitudes in forecasting. Then four criteria for selecting proper forecasting methods were discussed. In Section 6.3, quick implementation guides for the KFF method and the

BAF method have been presented, focusing on basic features, input preparation, and results interpretation.

CHAPTER VII

CONCLUSIONS

7.1 Contributions

Major contributions of this research to the construction industry and the project management community are: (1) Two probabilistic forecasting methods for project progress and early warning of schedule overruns were developed; (2) Project schedule forecasting methods – both conventional and the newly developed – were evaluated and compared; and (3) With the new forecasting method evaluation framework developed in this dissertation, independent comparisons between forecasting methods can be made with respect to the practical and statistical significance of forecasts.

(1) Two probabilistic forecasting methods for project progress and early warnings of schedule overruns were developed.

In this dissertation, two probabilistic methods for project schedule forecasting have been developed. The Kalman filter forecasting method (KFFM) and the Bayesian adaptive forecasting method (BAFM) have the following advantages over the critical path method (CPM) and the earned value method (EVM) based on the earned schedule metric:

1. <u>The KFFM and BAFM methods explicitly account for uncertainty in forecasts and provide prediction bounds on predictions</u>. The level of uncertainty or risk associated in a prediction is critical to efficient planning and decision making. Combined with the prescribed level of acceptable risk of the organization, the

prediction bounds can be used to determine the specific time at which current deviations from the original plan should be considered significant such that the original project targets are likely to be missed. With such warning signals, project managers and decision makers in top management will be able to take necessary actions in a timely manner.

2. <u>The KFFM and the BAFM make better use of the information available in typical construction project environments</u>. The most difficult problem in project performance forecasting is, probably, the lack of sufficient data to work with at the early stages of a project. In the KFFM and the BAFM, prior performance information in various forms such as detailed project plans, historical data, and subjective judgments are systematically integrated with actual performance data. For example, the baseline progress curve of a project is used to supplement the small samples of actual performance data, which is typical early in a project.

3. <u>The KFFM and the BAFM provide a systematic way of incorporating measurement errors into forecasts</u>. In project performance forecasting, one should make the best of actual performance data. However, measurement errors are inherent in project management and the degree of errors may differ from project to project. Therefore, it would be attractive for the users that the KFFM and the BAFM allow them to adjust the sensitivity of forecasts to actual performance data by adjusting some parameters according to their belief on the accuracy of measurements.

In spite of the methodological advantages addressed above, the KFFM and the BAFM can be efficiently implemented by any construction project, regardless of its type, size, and complexity. Furthermore, both methods rely on performance information which is available in standard practices and methodologies in the construction industry. Especially, the EVM provides an ideal management system in which the new methods can be integrated seamlessly. It should also be noted that the programs used in this dissertation are developed in Visual Basic for Application for Excel® and can be used in any computer with Excel®.

(2) Project schedule forecasting methods – both the conventional and the newly developed – were evaluated and compared.

Both the EVM and the CPM have been in use in the project management community for several decades. Surprisingly enough, no research has been done, to the best of the author's knowledge, about the comparison between them, especially regarding their early warning capacities. Some experts advocate the CPM because of its capacity of focusing on critical activities, while others prefer the EVM-based forecasting formulas, on the grounds of simplicity. In this dissertation, the advantages and challenges of the CPM and the EVM have been evaluated with respect to three vital criteria in schedule forecasts: the accuracy at different stages of a project, the early warning capacity, and the reliability of warnings. The results, which will be summarized in the next section, are intuitive and can be explained by some fundamental attributes of each method. But still, they would give many researchers and practitioners some fresh perspectives about those methods that they have used and are using now. Especially, the limitations in the CPM and the EVM, which were revealed in this dissertation, show the needs for better forecasting methods for project schedule management. At this point, the comparison between the KFFM and the BAFM versus the conventional methods would be useful both for practitioners who need something that can be used right away and for researchers who are interested in developing new forecasting methods.

(3) With the new forecasting method evaluation framework developed in this dissertation, independent comparisons between forecasting methods can be made with respect to practical and statistical significance of forecasts.

The new evaluation framework for project performance forecasting methods is based on artificial project data and a formal early warning system. With the random progress generation technique proposed in this dissertation, a large sample of independent projects can be generated so that statistically meaningful results can be obtained. It should be reminded that real project data can not be used in the comparison between new and conventional forecasting methods because real projects are likely to have been influenced by some of the conventional methods. Because artificial data are independent of any forecasting methods, whatever they are, it becomes possible to compare new forecasting methods against conventional ones that might have been employed in real projects in the past. Another salient aspect of the new evaluation framework can be found in the early warning system incorporated to simulate a warning process in typical project management. As a result, the early warning capacity of a forecasting method is directly evaluated in the situation for which the forecasts are used rather than being inferred from statistical error measures.

7.2 Conclusions

Major conclusions from this research are: (1) The state-of-the-art EVM schedule forecasting method can be used to obtain reliable warnings only after the project performance has stabilized; (2) The CPM is not capable of providing early warnings due to its retrospective nature; (3) The KFFM and the BAFM should be used to forecast progress and to obtain reliable early warnings of all projects; and (4) The early warning capacity of forecasting methods should be evaluated and compared in terms of the timeliness and reliability of warning in the context of formal early warning systems.

(1) The usual EVM schedule forecasting method can be used to obtain reliable warnings *only after* the project performance has stabilized.

In practice, EVM forecasting formulas for the estimate at completion (EAC), for example EAC = BAC/CPI_C, are recommended to be used only for projects that are at least 15 to 20 percent complete because of the inherent instability in the cumulative CPI measurements (Fleming and Koppelman 2006). This practice is supported by some empirical studies as addressed in Chapter II. On the other hand, predictive potentials of various schedule forecasting formulas based on EVM are still a contentious issue among professionals. Some EVM experts argue that earned value schedule data alone are not sufficient and should not be relied on to predict the final completion date for a project

(Fleming and Koppelman 2006), while others struggle to find proper schedule counterparts of the cost forecasting formulas in use (Lipke 2006; Vanhoucke and Vandevoorde 2006). For the former, the CPM often emerges as a rational solution to the schedule forecasting problem.

The results in this dissertation are in support of the potential benefits from a state-of-the-art EVM schedule forecasting method. Based on some recent findings in the literature, the earned schedule method is chosen as the state-of-the-art EVM schedule forecasting method. The parametric study in Chapter V reveals that early in projects the earned schedule method provides extremely erroneous forecasts. However, over a certain period, roughly 20 percent of the original project duration, the forecast error decreases rapidly and gets closer to the size of errors from the CPM, the KFFM, and the BAFM. Although the results in this dissertation were obtained from an artificial experiment, there is a close analogy between the schedule forecasting performance observed in this dissertation and the cost forecasting performance in the previous literature: both of them need some time to let the performance stabilize.

However, it should be noted that there is a serious limitation in the earned schedule method. That is, there is no way to detect the right time when the performance data of a project have stabilized within an acceptable level. Therefore, it can be concluded that the earned schedule method, which is known as the best schedule forecasting metric in the EVM, provides reliable early warnings only after the project performance has stabilized, but no one knows when it stabilizes.

(2) The CPM is not capable of providing early warning.

The CPM can not be used to predict schedule overruns *before* they actually happen. All that a CPM forecast for the project completion date says is how much the project is *currently* behind its planned schedule. In the CPM, there are no proper, accepted algorithms for systematically updating the original estimates of future jobs according to past performance data. This retrospective nature results in poor early warning capacity for detecting overrunning projects before they actually fall behind schedule. The results in Section 5.6 regarding the performance of the CPM reveal that the CPM has literally no predictive power and detects overrunning projects only when they already have fallen behind schedule.

(3) The KFFM and the BAFM should be used to forecast progress and to obtain reliable early warnings.

There are at least three compelling reasons why the KFFM and the BAFM should be applied on all project, especially the projects that employ the earned value management system. First, the KFFM and the BAFM provide prediction bounds on predictions but the CPM and the EVM do not. Prediction bounds indicate the range of possible outcomes at a given confidence level. In project management, the level of uncertainty or risk in forecasts can affect the planning and controlling decisions regarding future performance of a project.

Second, the KFFM and the BAFM share unique merits of the EVM but still provide stable forecasts from project inception. Just like the EVM, the KFFM and the

BAFM are universally applicable over a wide range of projects. However, the KFFM and the BAFM provide reliable forecasts from the outset of a project and do not require the stabilization period for the EVM forecasting methods. This characteristic can be attributed to the adaptive nature of the KFFM and the BAFM. The EVM forecasts show significant variability early in a project because the forecasts are made based only on small samples of actual data. On the contrary, the KFFM and the BAFM make forecasts using both the actual data and the prior performance information from project plans, historical data, and subjective judgments. As the project proceeds, the influence of the prior information decreases and forecasts become more influenced by the actual performance data. As a result, the problem of waiting some time – no one knows how long it is – for forecasts to stabilize can be avoided.

Third, the KFFM and the BAFM allow the users to adjust the sensitivity of forecasts to actual performance data according to their confidence on the accuracy of performance measurements. Measurement errors are almost inherent in project management because of human errors, inadequacy of rules for determining earned value, and variations in reporting intervals. As a result, the level of accuracy of actual performance data may differ from project to project. Therefore, the flexibility of the KFFM and the BAFM in dealing with measurement errors will be useful in real projects under diverse situations.

(4) The early warning capacity of forecasting methods should be evaluated and compared in terms of the timeliness and reliability of warning in the context of formal early warning systems.

In the context of project performance forecasting, there is a tendency to believe that more accurate forecasts will lead to earlier warnings. For example, Vanhoucke and Vandevoorde (2006) state that "we want to continue our research in order to improve the forecast accuracy of the EV metrics … in order to improve the early-warnings potential of the metrics." However, forecast accuracy based on statistical error measures should not be confused with the early warning capacity of forecasting methods. The results in this dissertation provide evidence that forecast accuracy of a method, useful as it is, is not an appropriate metric for the timeliness of warnings by the method. For example, according to the results in Section 5.6, the differences in accuracy between the CPM and other methods were not practically significant, but the chance of correct warnings from the CPM at the 40 percent project duration point was merely 4% against 32% from the BAFM, 39% from the EVM, and 48% from the KFFM. Statistical significance of forecasts should not be confused with practical significance which can be measured in the situations for which they are used. In this research, the early warning capacity of forecasting methods was evaluated in terms of their performance in an early warning system.

7.3 Further Study

In this research, two probabilistic forecasting methods, the KFFM and the BAFM, were developed and their applicability was demonstrated. In addition, an extensive parametric study has been carried out about the use of prior performance information and forecasting performance of the new methods against the EVM and the CPM. Based on the results established in this dissertation, some suggestions have been made about future study.

Application of the KFF and the BAF methods to cost performance forecasting

In this dissertation, application of the KFFM and the BAFM to cost performance forecasting is not covered. From a methodological point of view, as long as project progress is measured in some cumulative metric, the distinction for whether the metric represents schedule performance (i.e., the budgeted cost of work performed) or cost performance (i.e., the budgeted cost of actual cost) is irrelevant. However, it would be beneficial to investigate the cost forecasting counterparts of the KFFM and the BAFM formulated in this dissertation, and to compare them with other conventional methods such as the EVM and a Monte Carlo simulation approach.

Integration of schedule forecasting and cost forecasting

Schedule performance and cost performance are closely related and should be analyzed collectively. When it comes to forecasting, a schedule forecast should precede a cost forecast because, in general, project duration itself is an important factor of the project cost. The Kalman filter has a significant potential as an integrated framework because it provides an extremely flexible framework for combining new state variables.

REFERENCES

AbouRizk, S. M., and Halpin, D. W. (1992). "Statistical properties of construction duration data." *Journal of Construction Engineering and Management*, 118(3), 525-544.

AbouRizk, S. M., Halpin, D. W., and Wilson, J. R. (1991). "Visual interactive fitting of beta distributions." *Journal of Construction Engineering and Management*, 117(4), 589-605.

AbouRizk, S. M., Halpin, D. W., and Wilson, J. R. (1994). "Fitting Beta distributions based on sample data." *Journal of Construction Engineering and Management*, 120(2), 288-305.

Agrawal, M. K., Elmaghraby, S. E., and Herroelen, W. S. (1996). "DAGEN: A generator of testsets for project activity nets." *European Journal of Operational Research*, 90(2), 376-382.

Al-Tabtabai, H., and Diekmann, J. E. (1992). "Judgemental forecasting in construction projects." *Construction Management & Economics*, 10(1), 19.

Anbari, F. T. (2003). "Earned value method and extensions." *Project Management Journal*, 34(4), 12.

Ang, A. H., and Tang, W. H. (1975). *Probability concepts in engineering and design*, John Wiley & Sons, New York, NY.

Ang, A. H., and Tang, W. H. (2006). *Probability concepts in engineering:Emphasis on applications to civil and environmental engineering*, 2nd Ed., John Wiley & Sons, New York, NY

Ansley, C. F., and Kohn, R. (1985). "Estimation, filtering, and smoothing in state space models with incompletely specified initial conditions." *The Annals of Statistics*, 13(4), 1286-1316.

Armstrong, J. S., ed. (2002). *Principles of forecasting: A handbook of researchers and practitioners*, Kluwer Academic, Boston, MA.

Awwad, H. M., Valdes, J. B., and Restrepo, P. J. (1994). "Streamflow forecasting for Han River basin, Korea." *Journal of Water Resources Planning and Management*, 120(5), 651-673.

Barraza, G. A., Back, W. E., and Mata, F. (2000). "Probabilistic monitoring of project performance using SS-curves." *Journal of Construction Engineering and Management*, 126(2), 142-148.

Barraza, G. A., Back, W. E., and Mata, F. (2004). "Probabilistic forecasting of project performance using stochastic S curves." *Journal of Construction Engineering and Management*, 130(1), 25-32.

Bates, J. M., and Granger, C. W. J. (1969). "The combination of forecasts." *OR*, 20(4), 451-468.

Blyth, K., and Kaka, A. (2006). "A novel multiple linear regression model for forecasting S-curves." *Engineering, Construction and Architectural Management*, 13(1), 82.

Brookner, E. (1998). *Tracking and Kalman filtering made easy*, John Wiley & Sons, Inc, New York.

Bunn, D. W. (1975). "A Bayesian approach to the linear combination of forecasts." *Operational Research Quarterly*, 26(2, Part 1), 325-329.

Bunn, D. W. (1985). "Statistical efficiency in the linear combination of forecasts." *International Journal of Forecasting*, 1(2), 151-163.

Carbone, R., and Armstrong, J. S. (1982). "Evaluation of extrapolative forecasting methods: Results of a survey of academicians and practitioners." *Journal of Forecasting*, 1(2), 215-217.

Casals, J., Jerez, M., and Sotoca, S. (2000). "Exact smoothing for stationary and non-stationary time series." *International Journal of Forecasting*, 16(1), 59-69.

Chambers, J. C., Mullick, S. K., and Smith, D. D. (1971). "How to choose the right forecasting technique." *Harvard Business Review*, 49(4), 45-70.

Chau, K. W. (1995). "Monte Carlo simulation of construction costs using subjective data." *Construction Management & Economics*, 13(5), 369.

Christensen, D. S. (1993). "The estimate at completion problem: A review of three studies." *Project Management Journal*, 24(1), 37.

Cioffi, D. F. (2005). "A tool for managing projects: an analytic parameterization of the S-curve." *International Journal of Project Management*, 23(3), 215-222.

Cottrell, W. D. (1999). "Simplified program evaluation and review technique (PERT)." *Journal of Construction Engineering and Management*, 125(1), 16-22.

Curran, M. W. (1989). "Range estimating: Measuring uncertainty and reasoning with risk." *Cost Engineering*, 31(3), 18.

Davies, E. M. (1973). "An experimental investigation of resource allocation in multiactivity projects." *Operational Research Quarterly (1970-1977)*, 24(4), 587-591.

de Jong, P. (1989). "Smoothing and interpolation with the state-space model." *Journal of the American Statistical Association*, 84(406), 1085-1088.

de Menezes, L. M., W. Bunn, D., and Taylor, J. W. (2000). "Review of guidelines for the use of combined forecasts." *European Journal of Operational Research*, 120(1), 190-204.

De Reyck, B., and Herroelen, W. (1996). "On the use of the complexity index as a measure of complexity in activity networks." *European Journal of Operational Research*, 91(2), 347-366.

Demeulemeester, E., Dodin, B., and Herroelen, W. (1993). "A random activity network generator." *Operations Research*, 41(5), 972-980.

Demeulemeester, E., Vanhoucke, M., and Herroelen, W. (2003). "RanGen: A random network generator for activity-on-the-node networks." *Journal of Scheduling*, 6(1), 17-38.

Elmaghraby, S. E., and Herroelen, W. S. (1980). "On the measurement of complexity in activity networks." *European Journal of Operational Research*, 5(4), 223-234.

Fellows, R., Langford, D., Newcombe, R., and Urry, S. (2002). *Construction management in practice*, 2nd Ed., Blackwell Publishing, Malden, MA.

Finley, E. D., and Fisher, D. J. (1994). "Project scheduling risk assessment using Monte Carlo methods." *Cost Engineering*, 36(10), 24.

Fleming, Q. W., and Koppelman, J. M. (2006). *Earned value project management*, 3rd Ed., Project Management Institute, Inc., Newtown Square, PA.

Franklin, J. L. (2007). "2006 National hurricane center forecast verification report." National Weather Service, National Hurrican Center, Tropical Prediction Center, Miami, FL.

Galloway, P. D. (2006a). "Comparative study of university courses on critical-path method scheduling." *Journal of Construction Engineering and Management*, 132(7), 712-722.

Galloway, P. D. (2006b). "Survey of the construction industry relative to the use of CPM scheduling for construction projects." *Journal of Construction Engineering and Management*, 132(7), 697-711.

Gardoni, P., Reinschmidt, K. F., and Kumar, R. (2007). "A probabilistic framework for Bayesian adaptive forecasting of project progress." *Computer-Aided Civil and Infrastructure Engineering*, 22(3), 182-196.

Georgoff, D. M., and Murdick, R. G. (1986). "Manager's guide to forecasting." *Harvard Business Review*, 64(1), 110-120.

Hazelrigg, G. A. (2003). "Validation of engineering design alternative selection methods." *Engineering Optimization*, 35(2), 103.

Hertz, D. B. (1964). "Risk analysis in capital investment." *Harvard Business Review*, 42(1), 95-106.

Hertz, D. B. (1968). "Investment policies that pay off." *Harvard Business Review*, 46(1), 96.

Hertz, D. B. (1979). "Risk analysis in capital investment." *Harvard Business Review*, 57(5), 169-181.

Hoeting, J. A., Madigan, D., Raftery, A. E., and Volinsky, C. T. (1999). "Bayesian model averaging: A tutorial." *Statistical Science*, 14(4), 382-401.

Howard, R. A. (1971). "Proximal decision analysis." *Management Science*, 17(9, Theory Series), 507-541.

Hulett, D. T. (1996). "Schedule risk analysis simplified." *PM Network*, 1996(July), 23-30.

Jorion, P. (2001). *Value at risk: The new benchmark for managing financial risk*, 2nd Ed., McGraw-Hill, New York, NY.

Kaimann, R. A. (1974). "Coefficient of network complexity." *Management Science*, 21(2), 172-177.

Kalman, R. E. (1960). "A new approach to linear filtering and prediction problems." *Transactions of the ASME-Journal of Basic Engineering*, 82(Series D), 35-45.

Keefer, D. L., and Bodily, S. E. (1983). "Three-point approximations for continuous random variables." *Management Science*, 29(5), 595-609.

Kenley, R., and Wilson, O. D. (1986). "A construction project cash flow model - an idiographic approach." *Construction Management & Economics*, 4(3), 213.

King, E. P., Sampson, C. B., and Simms, L. L. (1975). "An alternative to Monte Carlo sampling in stochastic models." *Management Science*, 21(6), 649-657.

Kolisch, R., Sprecher, A., and Drexl, A. (1995). "Characterization and generation of a general class of resource-constrained project scheduling problems." *Management Science*, 41(10), 1693-1703.

Leach, L. (2005). "EVM & critical chain project management." *NASA Project Management Challenge 2005*, Houston, TX.

Lee, D.-E. (2005). "Probability of project completion using stochastic project scheduling simulation." *Journal of Construction Engineering and Management*, 131(3), 310-318.

Lipke, W. H. (2003). "Schedule is different." *The Measurable News*, 2003(March), 31-34.

Lipke, W. H. (2006). "Earned schedule leads to improved forecasting." *3rd International Conference on Project Management (ProMAC 2006)*. Sydney Australia.

Lu, M., and AbouRizk, S. M. (2000). "Simplified CPM/PERT simulation model." *Journal of Construction Engineering and Management*, 126(3), 219-226.

Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., Newton, J., Parzen, E., and Winkler, R. (1982). "The accuracy of extrapolation (time series) methods: Results of a forecasting competition." *Journal of Forecasting*, 1(2), 111-153.

Malcolm, D. G., Roseboom, J. H., Clark, C. E., and Fazar, W. (1959). "Application of a technique for research and development program evaluation." *Operations Research*, 7(5), 646-669.

Mastor, A. A. (1970). "An experimental investigation and comparative evaluation of production line balancing techniques." *Management Science*, 16(11), 728-746.

Meade, N., and Islam, T. (1998). "Technological forecasting - model selection, model stability, and combining models." *Management Science*, 44(8), 1115-1130.

Meredith, J. R., and Mantel, S. J. (1995). *Project management: A managerial approach*, John Wiley & Sons, Inc, Hoboken, NJ.

Miskawi, Z. (1989). "An S-curve equation for project control." *Construction Management & Economics*, 7(2), 115.

Moder, J. J., and Rodgers, E. G. (1968). "Judgment estimates of the moments of PERT type distributions." *Management Science*, 15(2, Application Series), B76-B83.

Murmis, G. M. (1997). "'S' curves for monitoring project progress." *Project Management Journal*, 28(3), 29-35.

Nassar, K. M., and Hegab, M. Y. (2006). "Developing a complexity measure for project schedules." *Journal of Construction Engineering and Management*, 132(6), 554-561.

Newbold, P., and Granger, C. W. J. (1974). "Experience with forecasting univariate time series and the combination of forecasts." *Journal of the Royal Statistical Society. Series A (General)*, 137(2), 131-165.

O'Brien, J. J., and Plotnick, F. L. (1999). *CPM in construction management*, 5th Ed., McGraw-Hill, New York, NY.

Oberlender, G. D. (2000). *Project management for engineering and construction*, 2nd Ed., McGraw-Hill, Boston, MA.

Perry, C., and Greig, I. D. (1975). "Estimating the mean and variance of subjective distributions in PERT and decision analysis." *Management Science*, 21(12, Application Series), 1477-1480.

Project Management Institute (PMI). (2004). *A guide to the project management body of knowledge*, 3rd Ed., Project Management Institute, Inc., Newtown Square, PA.

Raftery, A. E., Madigan, D., and Hoeting, J. A. (1997). "Bayesian model averaging for linear regression models." *Journal of the American Statistical Association*, 92(437), 179.

Reeves, G. R., and Lawrence, K. D. (1991). "Combining forecasts given different types of objectives." *European Journal of Operational Research*, 51(1), 65-72.

Reinschmidt, K. (2006). "Project risk managment: Class notes." Zachry Department of Civil Engineering, Texas A&M University, College Station, TX.

Schexnayder, C. J., and Mayo, R. (2003). *Construction management fundamentals*, McGraw-Hill, Boston, MA.

Short, J. W. (1993). "Using schedule variance as the only measure of schedule performance." *Cost Engineering*, 35(10), 35.

Skitmore, M. (1992). "Parameter prediction for cash flow forecasting models." *Construction Management & Economics*, 10(5), 397.

Sparrow, H. (2005). "Integrating schedule and EVM metrics." *NASA Project Management Challenge 2005*, Houston, TX.

Spooner, J. E. (1974). "Probabilistic estimating." *Journal of the Construction Division*, 100, 65-77.

Teicholz, P. (1993). "Forecasting final cost and budget of construction projects." *Journal of Computing in Civil Engineering*, 7(4), 511-529.

Touran, A. (1993). "Probabilistic cost estimating with subjective correlations." *Journal of Construction Engineering and Management*, 119(1), 58-71.

Touran, A., and Wiser, E. P. (1992). "Monte Carlo technique with correlated random variables." *Journal of Construction Engineering and Management*, 118(2), 258-272.

Uyttewaal, E. (2002). *Dynamic scheduling with Microsoft Project 2002*, J. Ross Publishing, Boca Raton, FL.

van Slyke, R. M. (1963). "Monte Carlo methods and the PERT problem." *Operations Research*, 11(5), 839-860.

Vandevoorde, S., and Vanhoucke, M. (2006). "A comparison of different project duration forecasting methods using earned value metrics." *International Journal of Project Management*, 24(4), 289-302.

Vanhoucke, M., and Vandevoorde, S. (2006). "A simulation and evaluation of earned value metrics to forecast the project duration." *Journal of the Operational Research Society*, advance online publication, 13 September 2006.

Welch, G., and Bishop, G. (2001). "An introduction to Kalman filter." University of North Carolina, Chapel Hill, NC.

Winkler, R. L., and Makridakis, S. (1983). "The combination of forecasts." *Journal of the Royal Statistical Society. Series A (General)*, 146(2), 150-157.

Ye, S., and Tiong, R. L. K. (2000). "NPV-at-risk method in infrastructure project investment evaluation." *Journal of Construction Engineering and Management*, 126(3), 227-233.

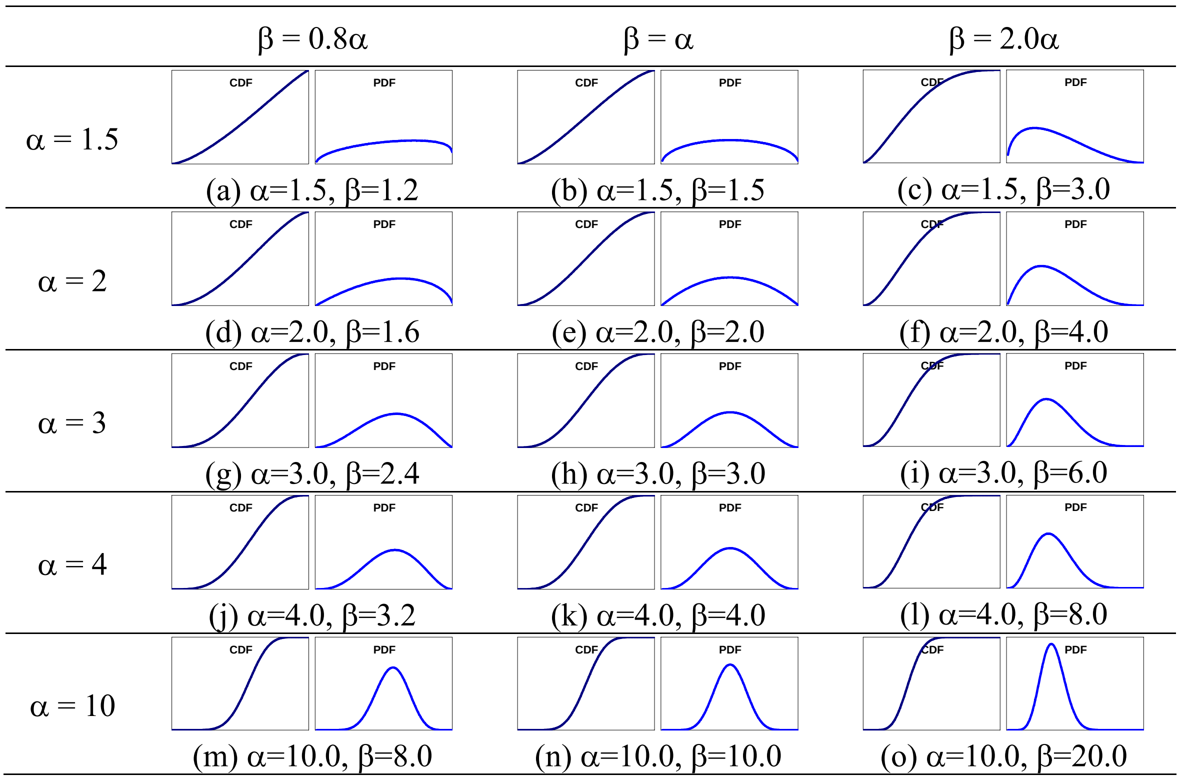
Yokum, J. T., and Armstrong, J. S. (1995). "Beyond accuracy: Comparison of criteria used to select forecasting methods." *International Journal of Forecasting*, 11(4), 591-597.

Zarchan, P., and Musoff, H. (2000). *Fundamentals of Kalman filtering: A practical approach (progress in astronautics and aeronautics)*, AIAA (American Institute of Aeronautics & Astonautics), Inc., Reston, VA.

Zwikael, O., Globerson, S., and Raz, T. (2000). "Evaluation of models for forecasting the final cost of a project." *Project Management Journal*, 31(1), 53.

APPENDIX A

More examples of the beta distribution with α > 1 and β > 1



APPENDIX B

The method of least squares

The method of least squares, which was proposed by the German scientist Karl Gauss (1777-1855), is a method for estimating the parameters of a regression model in a way that minimizes the sum of the squares of the (vertical, horizontal, or others – such as Generalized Least Squares) deviations of the observations from the fitted value.

The curve fitting technique used in this research utilizes the flexibility of the beta distribution in representing a wide range of shapes. The least squares estimates of the beta distribution parameters α, β, A and B should result in a curve that is, in one sense, a "best fit" to the progress curve at hand.

Suppose that a project progress curve is given in a discrete set of paired data points, (t_i, w_i) i = 1,2,…,N. Then the method of least squares can be applied to fit the beta distribution in two directions: vertical and horizontal. If the objective is to minimize the sum of the squares of the *vertical* deviations of the observed value from the fitted curve, the problem is defined as,

$$\begin{aligned} &\textit{Minimize} && \sum_{k=1}^{N}\left\{F_{beta}\left(t_i;\alpha,\beta,0,T\right)-\frac{w_i}{w_F}\right\}^2 \\ &\textit{Subject to}: && \alpha>1 \\ & && \beta>1 \\ & && T=known \end{aligned} \qquad \text{(B.1)}$$

where w_F is the progress level at completion and F_{beta} represents the cumulative distribution function (cdf) of the beta distribution.

In a similar way, if the objective is to minimize the sum of the squares of the *horizontal* deviations of the observed value from the fitted curve, the problem is defined as,

$$\begin{aligned} &\textit{Minimize} && \sum_{k=1}^{N}\left\{F_{beta}^{-1}\left(\frac{w_i}{w_F};\alpha,\beta,0,T\right)-t_i\right\}^2 \\ &\textit{Subject to}: && \alpha>1 \\ & && \beta>1 \\ & && T=known \end{aligned} \qquad \text{(B.2)}$$

where F_{beta}^{-1} represents the inverse of the cumulative distribution function for a specified beta distribution.

In this dissertation, the Solver function in the spreadsheet program Excel® is used to obtain the least squares estimates of the best-fit S-curve parameters.

VITA

Name: Byung Cheol Kim

Address: 185-3 Kyo-hyun dong, Chung-ju
Chung-chung Buk-do, South Korea.

Email Address: byungcheol.kim@gmail.com

Education:

2003 B.S., Civil Engineering, Seoul National University, Seoul, Korea
2005 M.S., Civil Engineering, Seoul National University, Seoul, Korea
2007 Ph.D., Civil Engineering, Texas A&M University, College Station, Texas

Experience:

2005 – 2007 Research assistant. Zachry Department of Civil Engineering, Texas A&M University
2000 – 2001 Project Controls Engineer, Office of International Projects, Samsung E&C Corporation, Seoul, Korea
1999 – 2000 Bridge Design Planner, Turnkey Project Office, Samsung E&C Corporation, Seoul, Korea
1996 – 1998 Bridge Construction Engineer, Youngjong Grand Bridge construction project, Samsung E&C Corporation, Inchon, Korea
1995 – 1995 Bridge Construction Engineer, Buk-bu Expressway construction project, Samsung E&C Corporation, Seoul, Korea

Publication:

Kim, Byung-Cheol (2003), *Introduction to Project Management*, Sewha publishing, Seoul, Korea.