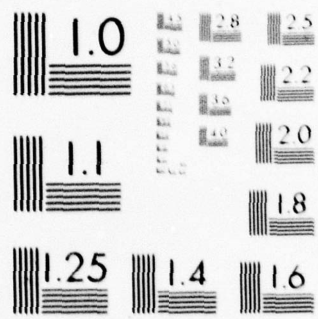
AD-A056 502 AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OHIO SCH--ETC F/G 5/1
AN EVALUATION OF A BAYESIAN APPROACH TO COMPUTE ESTIMATES-AT-CO--ETC(U)
DEC 77 R A HAYES
UNCLASSIFIED AFIT/GSM/77D-21

1 OF 1
AD
A056 502

END
DATE
FILMED
8 -78
DDC



MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

LEVEL II

UNITED STATES AIR FORCE

AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT/GSM/77D-21

(1) LEVEL II

AD A056502

DDC
RECEIVED
JUL 21 1978
B

AD No.
DDC FILE COPY

AN EVALUATION OF A BAYESIAN APPROACH
TO COMPUTE ESTIMATES-AT-COMPLETION
FOR WEAPON SYSTEM PROGRAMS

THESIS

AFIT/GSM/77D-21 Richard A. Hayes
Capt USAF

78 07 07 023

14

AFIT/GSM/77D-21

6

AN EVALUATION OF A BAYESIAN APPROACH

TO COMPUTE ESTIMATES-AT-COMPLETION

FOR WEAPON SYSTEM PROGRAMS.

9

THESIS

Master's thesis,

Presented to the Faculty of the School of Engineering

of the Air Force Institute of Technology

Air University

in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

by

10 Richard A./Hayes

Capt USAF

Graduate Systems Management

12 93p.

11 December 1977

Approved for public release; distribution unlimited.

012 225

acH

PREFACE

The objective of this paper was to research more accurate techniques of estimating the cost-at-completion on weapon system programs. After working in a program office in the Air Force as a program manager, I recognized the need for this type of information in managing weapon system programs. The reader should not construe the conclusions of this paper that there is any one best method all program managers should use. Even if one method of estimating is more accurate in the majority of programs over other methods, information can still be derived from those less "accurate" methods. The program manager or program analyst should estimate the final cost-at-completion using several techniques, and then analyze the differences between the estimates.

I would like to extend my appreciation to my advisor, Lt Col Adrian Harrell, and my reader, Major Charles McNichols, for their effort in guiding this research and struggling with the final manuscript. I would also like to thank Dr. N. Keith Womer for his technical advice on the research.

LIST OF FIGURES

Figure		Page
1	Work Breakdown Structure	12
2	Cost Variances	15
3	Cost Performance Report	17
4	Normal Probability Distribution	36
5	Variance Calculation	38
6	Sample Standard Deviation Calculations	40
7	Comparison of Methods: Program One	48
8	Comparison of Methods: Program Two	49
9	Comparison of Methods: Program Three	51
10	Comparison of Methods: Program Four	52
11	Comparison of Methods: Program Five	53
12	Summary of Data Analysis (entire program).	55
13	Summary of Data Analysis (first 12 months)	56
14	Summary of Data Analysis (first 6 months)	57
15	Variance Sensitivity Analysis	60
16	Changing Prior Distribution Mean	62
17	Using Non-Linear Regression to Determine Sample Distribution Mean	64

ABSTRACT

The Bayesian model developed by the author to predict costs-at-completion on weapon system programs is an extension of research done by M. Zaki El-Sabban. The model assumes cost is a random variable and is normally distributed. Budgeted costs are used to develop the prior probability distribution. Actual cost information is used for the Bayesian updating of the probability distribution. The mean of the updated probability distribution is the new estimated cost-at-completion for the program. The model was compared with a non-linear regression model and a linear extrapolation model on five weapon system programs. On three of the programs, the non-linear regression model estimated the final cost the greater percentage of the time. On the remaining two programs, the Bayesian model estimated the final cost the greater percentage of the time. The Bayesian model demonstrated several advantages over previous models: use at the beginning of the program, inclusion of subjective information, and giving weight to future program budgets.

CONTENTS

	Page
Preface	ii
List of Figures	iii
Abstract	iv
I. Introduction	1
Bayesian Approach	4
Objectives	5
Research Questions	5
Methodology	6
Limitations of Scope	7
Organization	7
II. Theory of C/SCSC	9
C/SCSC History	9
Objectives of C/SCSC	11
Requirements of C/SCSC	13
Cost Performance Report	16
Analysis of Data	18
Past Performance Factor	19
Regression Analysis	21
Time-series Analysis	23
Bayesian Approach	23
Summary	24
III. The Bayesian Model	26
Bayes' Theorem	26
El-Sabban's Approach	29
New Approach	31
Variance Calculation	34
Use of Subjective Data	41
Summary	41
IV. Data Analysis and Results	43
Data Collection Methodology	43
Data	44
Bayesian Method Modifications	45
Comparison of Methods	47
Analysis of Bayesian Method	58
Modification of the Bayesian Model	63
Summary	64

	Page
V. Conclusions and Recommendations	67
Research Quostions	67
Conclusions	69
Recommendations for Further Research	70
Bibliography	71
Appendix A	74
Appendix B	81

I. INTRODUCTION

In the past twenty years, cost estimating and predicting have become more and more critical in the weapon systems acquisition cycle of the Department of Defense. The Defense budget as a percentage of the federal budget and the nation's gross national product has declined, yet the weapon systems bought are more expensive and technologically more advanced. Even the weapon system acquisition process itself has become more complex with the incorporation of ideas such as reliability, maintainability, life-cycle costing, and compatibility between the services. This paradox of increasing complexity and costs coupled with a decreasing budget has caused a strong burden to be placed on management information systems to generate accurate data in a timely fashion to allow for the most competent decision-making. Particular emphasis has been placed on the cost information system in an attempt to determine as soon as possible if a cost overrun is developing so management action can be taken to minimize the effect of the overrun.

The Department of Defense started in 1961 to develop a management information system that would accurately report to the project offices the cost and schedule situation of the program accurately and in a timely fashion. The first attempt resulted in an adaptation of a technique used for scheduling, the Program Evaluation Review Technique (PERT), called PERT-COST. However, PERT-COST enjoyed only limited success and was phased out in the late 1960s (Acquino, 1977, p. 565).

By 1967, a new system was adapted by the Defense Department to replace PERT-COST. This system was called Cost/Schedule Control System Criteria (C/SCSC). There were two key elements in C/SCSC. First, it was a set of criteria that the contractor's system had to meet, not a new system imposed on the contractor. Second, it encompassed the idea of earned value. The concept of earned value more accurately determines whether the program is following the budget or not. More will be said about earned value and the details of C/SCSC in Chapter II.

C/SCSC has been in use for ten years and is now a mature system. Almost all major Defense contractors, including Defense installations such as Army arsenals, have been qualified under C/SCSC (Baumgartner, 1974, p. 33). An extensive literature search uncovered the following areas that have been identified as problem areas:

1. Accuracy of the data
2. Timeliness of the data
3. Contractor acceptance of the system
4. Contractor manipulation of the baseline budget
5. Rebudgeting of open work packages
6. More accurate estimating techniques.

The most complete work on C/SCSC history, how it works, and problems of the system (done by surveying both contractor and Government program managers) was done by Lieutenant Colonel Leonard Marrella (1973) in his dissertation, The Effect of the Cost Schedule Control Systems Criteria on Contractor Planning and Control.

The output of C/SCSC is presented in the cost performance report delivered monthly to the program office. The data in the monthly cost performance report is a summary of actual and budgeted costs of the program to date. The report not only gives costs for the total program, but also for the specific elements of the program. The data in the monthly cost performance report is a summary of actual and budgeted costs of the program to date. The report not only gives costs for the total program, but also for the specific elements of the program. The data presented in the cost performance report is used by managers for two primary purposes: 1) to determine problem areas in the program through analysis of the reported variances and, 2) to develop an estimate-at-completion (EAC) for the program through trend analysis.

This research problem centers on the latter area, developing an accurate estimate-at-completion at the earliest possible stage in the program. This gives the manager useful information for his docision making. Three basic methods have already been developed:

1. Past performance factor
2. Time-series analysis
3. Regression analysis.

The past performance factor method has many variations and has been used since the inception of the C/SCSC program (Holeman, 1974, p. A-1). A performance factor is generated from past historical data and applied to the budget-at-completion estimate to genarate the estimate-at-completion.

The time-series analysis approach developed by John Sincavage (AMSAV-PK) is being used by the B-1 Systems Project Office and the Army helicopter projects (Holeman, 1974, p. A-2). One major disadvantage with this method is that it requires a significant amount of data inputs (up to 50 data points, which translates into 50 months of data).

A non-linear regression model pioneered by Arthur Karsch of ASD/AC is just now being used by the Air Force on a widespread basis (Karsch, 1974, p. 19). To overcome the deficiency of small samples, Karsch developed a constrained model that uses parameters from historical data at first and modifies these parameters with data from the present program as more data is accumulated (Karsch, 1974, p. 19).

Bayesian Approach

A fourth approach was proposed by M. Zaki El-Sabban in 1973 when he was with the Army Aviation Systems Command (El-Sabban, 1973, p. 3). The major assumption in the use of this method is that the cost at any point in time along the path of a program follows a specific probability distribution. Although this method was published in 1973, this writer cannot find any follow-up research, nor actual use of this methodology for computing estimates-at-completion. Furthermore, in El-Sabban's paper only one month's example was given as a demonstration of the method. And the outcome was not compared to other techniques to determine its real benefit. Finally, El-Sabban never explained how the variance, a critical part of the estimate-at-completion calculation, should be derived.

Objectives

If the Bayesian method for generating estimates-at-completion can be shown to be an effective approach in forecasting final costs of weapon systems programs, it can be of benefit to the program manager. Managers have a continuing need for timely and accurate information. Presumably, the more perfect the information the analyst provides to the manager, the better the manager's decisions will be. The Bayesian method, if shown to be at least as accurate as the methods now in use, offers several advantages:

1. Easily closed formulas are used
2. No extrapolation is involved
3. Inclusion of subjective information is allowed
4. *It can be used early in the program.*

This research extends the ideas presented by El-Sabban, and attempts to determine if the Bayesian approach has advantages over methods presently in use. The research will first determine if El-Sabban's application of Bayesian statistics is correct. Second, it will try to determine methods for calculating variances on the assumed cost distributions, a critical step left out by El-Sabban. Finally, the research will determine how the Bayesian method compares with the presently used mathods to generate estimates-at-completion.

Research Questions

In order to satisfy the objectives of this research, three research questions were formulated.

The first research question concerned the theory itself:

1. Was the application of the Bayesian theory used by El-Sabban in his method correct?

The second research question concerned the application of the theory:

2. Can the variances of the assumed specific distributions be calculated with enough accuracy that a useful estimate-at-completion can be generated?

The final research question concerns the comparison of the Bayesian method with other methods now in use:

3. Is the Bayesian method more accurate in generating estimates than the presently used methods?

<u>Methodology</u>

This research is an effort to explore new methods to utilize data generated from a cost information system that is in compliance with C/SCSC. The objective is to determine if El-Sabban's applications of the Bayesian approach for estimates-at-completion is correct, and to determine how it compares with two methods presently being used by the Air Force, the past performance factor and non-linear regression methods. The extensive data requirements of time series analysis which limit its use and comparisons with this approach were not made. An extensive literature search was conducted to determine the state of the art of C/SCSC. Telephone interviews with the Comptroller's office in the Air Force and Army were made to obtain the latest information concerning methods presently used in the Department of Defense for generating estimates-at-completion.

A literature review of Bayesian statistics was performed. This literature review, along with discussions with Professors N.K. Womer and Charles McNichols, of the Air Force Institute of Technology's System Management Deaprtment, provided the basis for the determination of variance calculations.

Data used for the comparisons of the different methods was taken from completed and on-going programs at the Aeronautical Systems Division, Wright-Patterson Air Force Base, Ohio.

Limitations of Scope

There are several inherent limitations to a study of this kind. The first is that data from only a small number of programs will be used to compare the different methods. Since each weapon system program is unique, this limitation will bias the results to some degree. The second limitation is caused by assuming that the specific distribution of the cost data is normal. It is beyond the scope of this research to determine whether cost data follow either a normal distribution of some other distribution (e.g. beta distribution). If the Bayesian method proves useful, further research into the actual distribution of the cost data would be worthwhile. The final limitation is that only data from the cost performance reports will be directly used in developing the estimate-at-completion.

Organization

This research paper is divided into five chapters. The first chapter contains introductory material on C/SCSC and its present state of the art, defines the objective of the research, and presents the research questions to be answered.

Chapter II examines cost/schedule control systems criteria. The history and objective of C/SCSC are discussed along with an in-depth discussion on the different methods used to generate estimates-at-completion. This chapter provides the necessary background needed to understand C/SCSC.

Chapter III examines the theory used by El-Sabban in his application of Bayesian statistics. Also, the approach developed to determine the variances of the assumed distribution is discussed in this chapter.

Chapter IV includes the application of the Bayesian approach to actual data, and a comparison with past methods is made to determine if the Bayesian approach is a viable method of generating estimates-at-completion.

In Chapter V, the conclusions of the research are presented along with recommendations on the use of the Bayesian approach. In addition, suggestions for additional study are included.

II. THEORY OF C/SCSC

Before the actual Bayesian method used to estimate costs is discussed, an understanding of C/SCSC and the resultant information generated is needed. This chapter will first give a brief history of the development of C/SCSC, and then outline the basic requirements of a cost information system such as C/SCSC. An explanation of the type of information generated by the cost system will be given. Finally, the present methods of taking this information and generating cost estimates-at-completion will be discussed.

C/SCSC History

To control a program, the manager must anticipate deviations and changes, understand their impact on the program, and take the necessary action to prevent cost overruns and schedule slippages. The sooner the manager knows a problem is developing, the wider the options are for dealing with the problem. Therefore, a necessary prerequisite for good management is a responsive and accurate cost information system. The Department of Defense has made several attempts at developing cost information systems that could be used to measure contractor performance to effectively manage the Defense weapon acquisition programs.

The first attempt by the Department of Defense to develop a reliable system came in 1960 when the Polaris Program Management Office developed a system called Program Evaluation Review Technique (PERT). This system was externally imposed on the contractor. At first, PERT was concerned with just scheduling; the basic form of PERT focussed on finding the longest time-consuming path through a network of paths. Through the use of PERT the manager could actually calculate a probability of completion (Chase, 1977, p. 555). PERT was later expanded to include cost information and was called PERT-COST. However, PERT-COST did not achieve its objectives, as evidenced by increasing overruns on major weapon systems in the Sixties (e.g. F-14, F-111, C-5 programs).

There were several drawbacks to the PERT-COST system. The major drawback of PERT-COST was that it was externally imposed in a very rigid fashion on the contractor. Most contractors did not replace their system with PERT-COST just to satisfy government requirements. What developed was two systems. The contractor used its own system for decision-making and developed a *second pseudo-information system* to satisfy PERT-COST requirements (Fox, 1971, p. 413). This meant that the information the government program office received was not the data used by the contractor to make decisions.

In an effort to correct the drawbacks of PERT-COST, the Defense Department implemented the Cost/Schedule Control Systems Criteria approach in 1967. This approach specifies the criteria, or requirements, that a Defense contractor's system must meet in order to contract with the Department of Defense if the contract exceeds certain cost thresholds and is not a fixed price contract.

Objectives of C/SCSC

The Cost/Schedule Control Systems Criteria approach is defined and described in detail in the Department of Defense Instruction (DODI) 7000.2. It is further detailed in the following Defense Department pamphlets: AFSCP 173-5, AFLCP 173-5, AMCP 37-5, NAVMAT P5240, AMCP 715-10, NAVMAT P5243, AFSCP/AFLCP 173-6, DSAH 8315.1, and DSAAP 7641.6.

As stated in DODI 7000.2, C/SCSC has two objectives: to insure that Defense contractors use effective management control systems, and to insure these systems provide data which accurately indicates work progress while properly relating cost, schedule and technical performance. Further, the data must be timely, traceable, and must supply government managers with a practical level of summarization (Kemps, 1974, p. 4). Although the Defense Department would like to apply C/SCSC on all contract efforts, it realizes that on most of the smaller contracts C/SCSC is not cost effective (it costs more to implement than it would save). Therefore, the Defense Department has established thresholds below which C/SCSC is not mandatory. C/SCSC is applicable to research and development contracts over $50 million and production contracts over $200 million. In the Air Force, these thresholds have been lowered to $25 million and $100 million respectively. The Air Force has also developed a much less formal and detailed report called Cost Schedule Status Report that applies to contracts under the threshold but above $2 million. C/SCSC can also be applied to programs of national urgency and selected subcontracts that do not fall under the previous categories. C/SCSC is not applied to any firm-fixed price contracts.

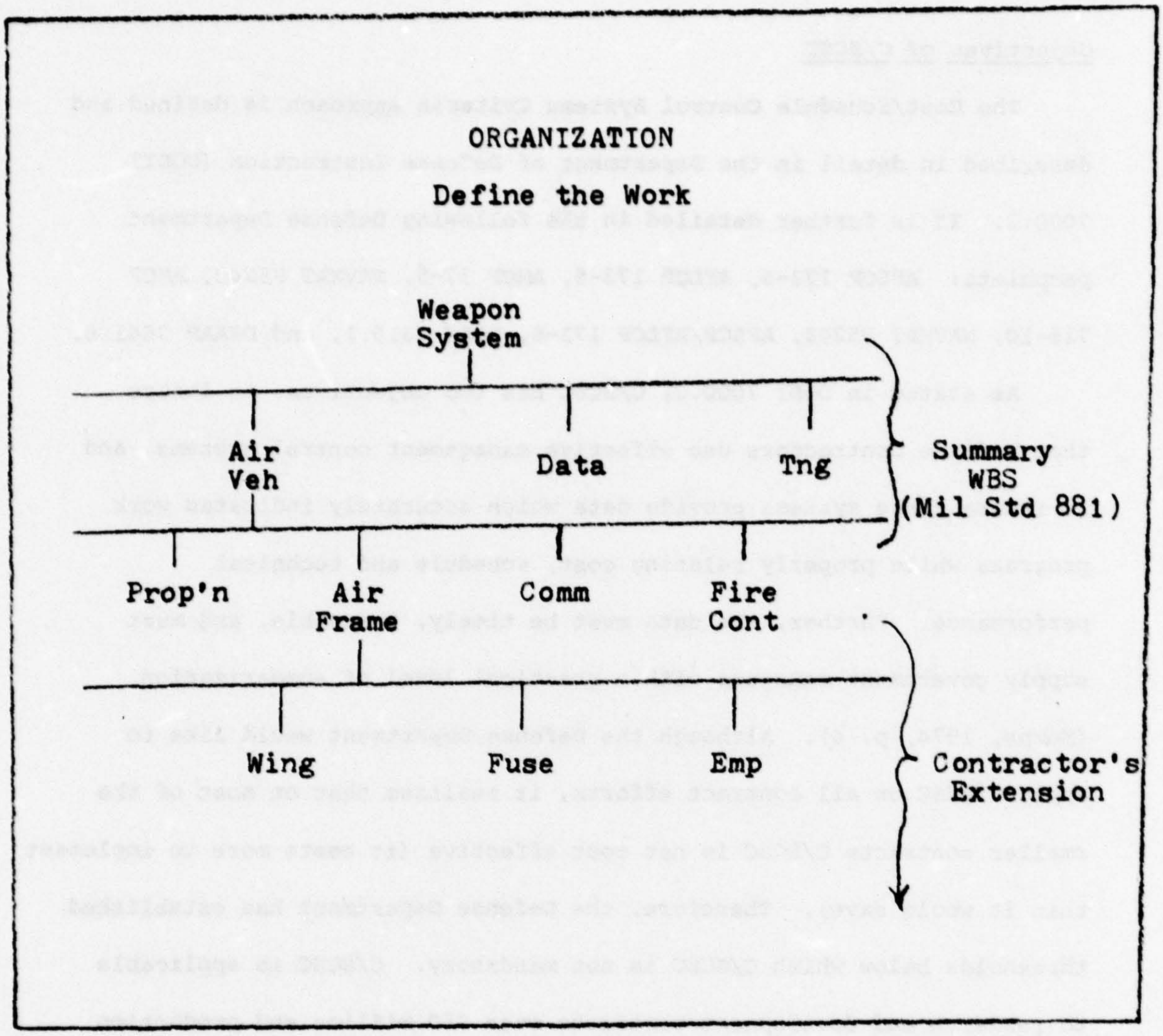


FIGURE 1. WORK BREAKDOWN STRUCTURE (Kamp, 1972)

Requirements of C/SCSC

To avoid the imposition of a new system and to minimize changes to the existing contractor's cost system, the Defense Department does not define a system to be used in 7000.2, but lists instead 35 major criteria that the existing contractor's system must meet in order to be validated, thus meeting contractual requirements. These 35 criteria are divided into five major categories: organization, planning and budgeting, accounting, analysis, and revisions.

Organization: C/SCSC requires that all work be defined and that a specific organization be assigned responsibility for its accomplishment. The primary vehicle used to organize the work is the program work breakdown structure. The objective is to subdivide the effort into manageable units of work. All work must be organized into short, clearly defined work packages that form the basis of the work breakdown structure. All costs must be accumulated from the bottom up as directly as possible.

Planning and budgeting: All work must be planned and budgeted before accomplishment with a separate budget assigned to each unit of work. Work is planned ahead to the level of detail required as the contract progresses. Budgets must be identifiable by cost elements such as labor, material, and other direct costs. Except for accounting adjustments, no retroactive budget changes are allowed for either completed work or for packages currently open. The fact that budgets are assigned to individual elements of work and retroactive changes cannot be made provides the basis for the analysis of the contractor's performance.

Accounting: The accounting system of the contractor must be able to record actual direct costs to the cost accounts. The system must be able to sum both direct and indirect costs from the level at which they are first recorded to the contract level. In general, any accounting procedure which is acceptable to the Defense Contract Audit Agency satisfies the requirements of C/SCSC.

Analysis: The system must be capable of comparing actual cost for the work performed to the budgeted cost of work performed. The critical feature of the contractor's system is to be able to not only define the actual cost for the work performed and the budgeted cost for the work scheduled, but also the budgeted cost for the work performed. This requirement is necessary to determine the actual cost variance of the program, and not bias the cost variance with schedule variances. This biasing is what happened formerly with systems that simply determined the actual cost for the work performed and compared it to the budgeted costs for the work scheduled to get a cost variance. The budgeted cost for the work performed is determined by adding up the budgets of those work packages which have been completed and an estimated amount of the budget finished in the work packages not completed. By having very short work packages, the distortion due to the subjectiveness of the estimated budget in work packages not finished will be less significant. Figure 2 further explains the determination of the cost variance.

Revisions: The contractor's system is required to incorporate contractual changes in a timely manner and to be able to reconcile original budgets with current performance measurement budgets by showing changes to authorized work or internal replanning actions. Any changes to the performance baseline must also be closely controlled and documented.

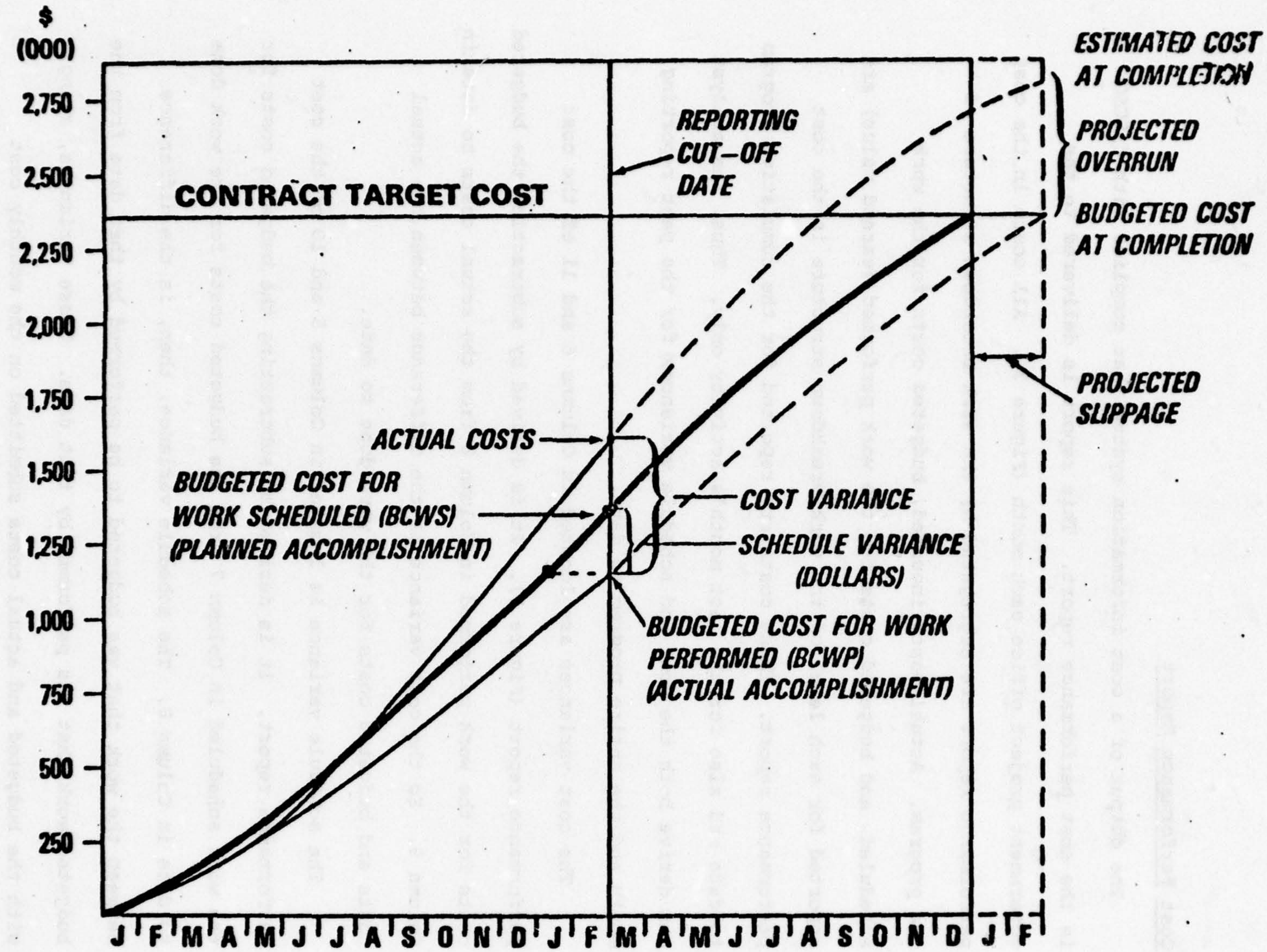


FIGURE 2. COST VARIANCES (Kemp, 1972)

Cost Performance Report

The output of a cost information system that complies with C/SCSC is the cost performance report. This report is delivered to the government project office each month (Figure 3). All costs in the cost performance report are segregated by the work breakdown structure of the program. Actual costs incurred, budgeted costs for the work scheduled, and budgeted costs for the work performed (earned value) are reported for each level of the work breakdown structure in the cost performance report. These costs are reported for the cumulative program to date and also for the last month's activity only. Thus, the analyst can derive both the cost and schedule variances for the past reporting month and the entire program to date.

The cost variances are located in Columns 6 and 11 of the cost performance report (Figure 3). It is derived by subtracting the budgeted costs for the work performed in Column 8 from the actual costs to date in Column 9. So the cost variance is the difference between the actual costs and budgeted costs for the work done to date.

The schedule variance is located in Columns 5 and 10 of the cost performance report. It is derived by subtracting the budgeted costs for the work scheduled in Column 7 from the budgeted costs for the work done to date in Column 8. The schedule variance, then, is the difference between the work that was budgeted to be performed by that date from the budgeted work that was performed by that date. These variances, along with the budgeted and actual costs submitted on the monthly cost performance report, will be used by the program analyst to predict future costs. The formulas are as follows:

COST PERFORMANCE REPORT -	Contract Number: A00019-69-C-0157	Program Name/Number: Boomerang/MGM-69A	Definitized Contract Target Cost	Report Period: 12-1-70 to 12-31-70	Signature, Title, & Date: J.S. Browning [illegible] 71	5 of 21

	Current Period: Budgeted Cost: Work Sched	Current Period: Budgeted Cost: Work Perform	Current Period: Actual Cost Work Perform	Current Period: Variance: Sched	Current Period: Variance: Cost	Cumulative to Date: Budgeted Cost: Work Sched	Cumulative to Date: Budgeted Cost: Work Perform	Cumulative to Date: Actual Cost Work Perform	Cumulative to Date: Variance: Sched	Cumulative to Date: Variance: Cost	At Completion: Budgeted	At Completion: Latest Revised Estimate	At Completion: Variance
Eng	6351	6173	6664	(178)	(491)	37249	33493	40455	(3756)	(6962)	108798	109536	(738)
Tooling	12	11	13	(1)	(2)	88	76	83	(12)	(7)	4257	4257	0
Q. C.	15	15	16	0	(1)	162	162	170	0	(8)	876	876	0
	526	524	532	(2)	(8)	3005	2997	3119	(8)	(122)	27463	27463	0
Equip.	231	230	232	(1)	(2)	1874	1870	1901	(4)	(31)	13729	13729	0
Material	12	12	12	0	0	94	94	95	0	(1)	686	686	0
Subcont.	3933	3784	3678	(149)	106	25080	22117	26005	(2963)	(3888)	63418	64018	(600)
Tng	72	75	80	3	(5)	274	271	285	(3)	(14)	2386	2386	0
	630	642	624	12	18	7570	7380	7470	(190)	(90)	18336	18336	0
ProjMgt	136	147	140	11	7	886	911	911	25	0	6699	6699	0
Data											10538	0	10538
Subtotal	11918	11613	11991	(305)	(378)	76282	69371	80494	(6911)	(11123)	257186	247986	9200
G & A									-				
Total													

FIGURE 3. COST PERFORMANCE REPORT (AMETA, 1974)

Cost Variance = Budgeted costs work performed - Actual cost work performed

Schedule Variance = Budgeted costs work performed - Budgeted costs for work scheduled

Analysis of Data

There are two basic types of analysis done with the data on the cost performance report: variance analysis and trend analysis. Variance analysis is done as the data is being received and entails no real computations. When the data is being compiled and an overrun or behind schedule is being reported (negative variance on the cost performance report), the manager or analyst goes to the next lower level in the work breakdown structure to determine what area(s) in that next lower level is causing the negative variance at the summary level. The analyst continues this procedure to the lowest level necessary to determine the cause of the negative variance so appropriate action can be taken to correct the situation. This technique is simple, but effective. It helps the manager determine where he should spend his time to correct the program problems (that show up on the report as variances).

The second type of analysis is trend analysis. Trend analysis takes on many forms, and focuses on predicting or forecasting future problems to prevent negative variances as early as possible. The various techniques in trend analysis use the data on the cost performance report to forecast costs at program completion (estimates-at-completion) for each of the summary elements in the work breakdown structure. The simplest method for determining the estimate-at-completion suggested by AFSCP 173-3 is:

Estimate at Completion	=	Actual Costs to Date	+	Budgeted Costs at Completion	-	Budgeted Costs to Date

This method does not attempt to evaluate future contractor cost performance; rather, it assumes the budget values will remain valid (Busse, 1977, p. 17). Besides this simple addition of the future budget to actual costs, there are three basic methods that have already been developed to forecast future costs in an attempt to determine problems at the earliest point in the program:

1. Past performance factor
2. Regression analysis
3. Time-series analysis.

This thesis proposes the use of a fourth method, the Bayesian approach.

Past Performance Factor

This technique of estimating the final cost uses a performance factor that is applied to the budgeted cost at completion to give an estimate-at-completion. This performance factor is derived from historical data, usually obtained from the cost performance report. There are several methods for computing the performance factors (or indices as they are sometimes called). The different methods amount to taking different information from the report to derive a performance factor. The theory behind this type of estimating is that past trends may be linearly extrapolated to the end of the program to forecast the final outcome. This method assumes that the same percentage of overrun (or underrun) will continue throughout the program. It assumes that an eventual overrun (underrun) at the end of the program depends completely on the current status reported without regard to its proximity to the end of the program.

One method of using a past performance factor is described in Status, Trends and Projections (AMETA, 1974, p. 17). This method uses the cumulative cost variance to generate an efficiency factor.

The percentage of budgeted costs for the work performed to the actual costs is a measure of cost efficiency with which work has been accomplished. An example from Figure 3 can illustrate this point:

$$\frac{\text{Budgeted costs (Col. 8)}}{\text{Actual costs (Col. 9)}} = \frac{69371}{80494} = 0.86$$

This means that for each budget dollar spent, 0.86 value was received. The question has been structured such that 1.0 is par and above 1.0 is better than par. To arrive at a final estimate (EAC), the final budgeted costs are divided by this efficiency factor:

$$\text{EAC} = \frac{\text{Final budget}}{\text{Efficiency factor}} = \frac{247986}{.86} = 288{,}356$$

Therefore, this method of using a past performance factor is predicting a $40360 overrun (288,356 - 247986 = 40360, as shown in Figure 3).

These factors can be developed using current monthly, cumulative, moving average, or weighted moving average cost data from the cost performance report. This method of determining estimates-at-completion is the most widely used by Defense Department program offices (Holeman, 1975, p. A-1). The reason for its popularity is because of its simplicity of computation. Electronics Systems Division has developed a standard computer program which takes the monthly cost performance report data and automatically computes estimates for each work breakdown structure level submitted in the report. The estimate-at-completion is computed six ways using several past performance factors including current month cost performance, cumulative cost performance, a moving average cost performance, and a combination of the cost and schedule performance factors. The computer cost information output also gives estimates using two other methods: historical performance factor derived from past contracts with same contractor and a non-linear regression technique discussed later.

Another more subjective performance factor was derived by Major J.B. Holeman (Holeman, 1974, p. 23). His performance factor, called predicted performance factor (PPF), replaced the normal cost performance factor discussed in the last paragraph $\left(\frac{\text{Actual costs}}{\text{Budget costs}}\right)$, and was constructed based on five different factors. The five factors that make up the new performance factor are inflation, overhead increase, unexpected technical problems, contract changes, and schedule variations. After the factor was calculated, Holeman used the following formula:

$$\text{EAC} = \text{Actual costs} + (\text{Final budget} - \text{present budget costs}) \cdot (\text{PPF})$$

<u>Regression Analysis</u>

Regression analysis is the estimation of a line or curve through existing plotted data (cost versus time). There are three basic types of regression analysis used to develop a trend and derive an estimate-at-completion. The first method is the "eyeball" technique where a line or curve is drawn free hand through the data. Although this may not give bad results, it is unlikely that two people will get the same exact result and it does not allow for the application of formal statistical procedures (AMETA, 1974, p. 36).

The second method is a regression technique known as the method of least squares. This technique calculates a line, from which the net of the squares of the vertical deviation is a minimum. The linear equation is

$y = mx + b$, where m is the slope of the line and b is the y-intercept when $x = 0$ (AMETA, 1974, p. 36). Although this technique can be done by hand, it is easiest done by computer.

The third regression technique used is the curvilinear least squares regression analysis developed by Arthur Karsch, ASD/ACC. He found that most samples observed did not show a straight line relationship in the long term, deciding that a curvilinear form was more realistic (Karsch, 1974, p. 13). He used the relationship:

$$Y = b_1 \cdot X^{b_2}$$

The factors b_1 and b_2 represent non-dimensional growth characteristics. Karsch developed both an unconstrained approach where b_1 and b_2 were determined only with the present program's historical data, and a constrained method where b_1 and b_2 are first derived from data available from completed programs. This constrained approach is an attempt to overcome the deficiency of small samples early in the program (Karsch, 1974, p. 19). This curvilinear regression analysis is being added to the Electronics System Division computer program as another method for deriving estimates-at-completion on weapons system programs (Karsch, 1977).

D. Busse has used Karsch's research to develop another technique for developing estimates-at-completion. Although his technique is not regression analysis, it used the curvilinear expression $b_1X^{b_2}$ that Karsch used. Busse takes information from latest monthly cost performance reports to arrive at his new estimate-at-completion:

$$EAC = Z\ (\text{Budget at completion})^{e}$$

where

$$e = \frac{\text{change actual costs}}{\text{actual cost to date}} \Bigg/ \frac{\text{change budgeted cost}}{\text{budgeted cost to date}}$$

$$Z = \frac{\text{actual costs to date}}{(\text{budgeted costs to date})^{e}}$$

The foundation of this method is based on the dynamic relationship between budgeted and actual costs and the sensitivity between the changes in those two variables (Busse, 1977, p. 23). This model is still in the development stage and is not yet being used operationally.

<u>Time-Series Analysis</u>

Another approach to developing estimates is using time-series analysis. This method was first used for developing estimates by the Army Aviation Systems Command on the UTTAS helicopter program, and later used in the Air Force by the B-1's Los Angeles office and the Munitions System Program Office at Eglin Air Force Base (Hayes, 1977). It is a form of least squares regression analysis that takes into account cyclical and seasonal variations. The big drawback is the amount of data needed -- up to fifty months of data. Time-series analysis is more applicable to continuous production programs than to research and development programs. Research and development programs do not necessarily follow seasonal or cyclical trends and do not last long enough to take advantage of the positive characteristics of time-series analysis.

<u>Bayesian Approach</u>

This thesis examines the viability of a Bayesian approach proposed by M. Zaki El-Sabban of the Systems Analysis Office at the Army Aviation Systems Command. Bayesian analysis represents an attempt to incorporate all relevant information directly into the process of making inferences about a state of nature. It uses Bayes' theorem to continually update (using each monthly cost performance report) an existing distribution developed from prior information (budgeted costs developed from historical information). The writer can find no further application of this approach than the single month example used by El-Sabban in his paper. There are several technical errors in El-Sabban's paper, and little guidance was given for the derivation of the two subjective variances needed to apply the Bayesian formula to derive an estimate-at-completion.

This research deals directly with the Bayesian approach and its usefulness to program offices. The next chapter considers the Bayesian approach in detail. Then in Chapter IV, results using the Bayesian approach will be compared to the cumulative cost variance past performance factor method and the non-linear regression analysis method. This comparison is to determine if the Bayesian method has any significant advantages over presently-used methods. The non-linear regression method was chosen since it is the most sophisticated type of regression analysis developed to date for cost analysis. The cost variance past performance factor method was chosen because it is the most widely used method used. The time-series method will not be used in the comparison because it is not widely used and has limited application.

Summary

Chapter II outlines criteria established by the Department of Defense that its contractors' accounting system must follow. The criteria was established in an effort to assure that cost data from these systems is valid. The Cost/Schedule Control Systems Criteria (C/SCSC) evolved after several previous systems were tried and failed. C/SCSC does not impose a new accounting system on the contractor, instead it requires the contractor's system to meet 35 major criteria established by the Department of Defense. These criteria cover five areas of an accounting system: organization, planning and budgeting, accounting, analysis, and revisions. The output of this accounting system should be accurate cost data that can be used by managers to determine the status of their programs. There are two types of analysis that can be done using the cost data, variance and trend analysis. This paper will concentrate on the validation of a new form of trend analysis using Bayesian statistics.

III. THE BAYESIAN MODEL

This chapter deals with the theory of Bayesian estimation. A brief review of Bayes' theorem, concentrating on its interpretation, is given. The derivation of the posterior probability assuming a normal probability distribution is then presented. It is shown how El-Sabban integrated cost data into Bayesian statistics to determine estimates of the final cost of a program. Finally, shortcomings in El-Sabban's approach are suggested, along with the author's approaches to correct these shortcomings.

Bayes' Theorem

Probability theory deals with uncertainty, and since managers deal frequently with uncertainty, probability theory is an important tool in the decision-making process. Probability theory is especially relevant in forecasting final costs on weapon system acquisition programs, due to the uncertainty of those programs (i.e., the process can lead to more than one outcome). The basis of the Bayesian approach is to assume that cost is a random variable and use probability theory to assign a useful measure to the likelihood that a certain actual cost will occur.

It will be assumed that frequencies can be assigned to the different possible outcomes of actual cost, and that these frequencies will have three properties:

1. The relative frequency of any event will be greater than or equal to zero.

2. The relative frequency assigned to the sample space will be equal to one.

3. If a given event can be represented as the union of two or more mutually exclusive events, the relative frequency of the given event represents the sum of the relative frequencies of the mutually exclusive events that comprise it.

With the above three properties, the function that maps the frequencies to the possible outcomes is called a probability distribution. The value of the distribution of any particular event is the probability of the event (Dyckman, 1969, p. 40). Probability distributions can be described by certain parameters. An example is the normal distribution which can be described by the two parameters, the mean and the variance.

An outcome, θ, is conditional when its probability of occurrence is dependent on another outcome, F. Therefore, if one of the events has occurred, the manager can take advantage of this information to improve his knowledge about the probabilities of the remaining event. The conditional probability of an event θ is defined as:

$$P(\theta|F) = \frac{P(\theta \cap F)}{P(F)}, \quad P(F) = 0$$

with $P(\theta \cap F) = P(F|\theta) \cdot P(\theta)$

$P(F)$ can be described by the theorem on total probability as:

$$P(F) = P(F|\theta_1) \cdot P(\theta_1) + P(F|\theta_2) \cdot P(\theta_2) + \ldots$$

$$P(F) = \sum P(F|\theta_i) \cdot (\theta_i)$$

Bayes' theorem then can be stated for the case of a continuous random variable as follows:

$$P(\theta|F) = \frac{P(F|\theta) \cdot P(\theta)}{\int P(F|\theta) \cdot P(\theta)\, d\theta}$$

Interpretation: $P(\theta)$ is called the prior probability since it is the probability that applies prior to the addition of the new evidence, F. This prior distribution reflects the amount of knowledge of θ (θ is a state of nature) prior to the experiment, F. If no information is known prior to the experiment, the prior probability distribution would be a uniform distribution where all outcomes are equally likely (White, 1971, p. 23). A prior distribution is usually obtained by fitting a distribution to historical data. In determining estimates-at-completion, the data used to determine the prior distribution is the budgeted costs, usually derived from historical data. For reasons to be discussed later, a normal distribution will be assumed for the cost data, with the budgeted cost assumed to be the mean of the prior distribution. Budgeted costs are used since the only information a manager has before a program is started is the budgeted cost for that program. $P(F)$ is the probability of the test observation of the test observation occurring (in the case of this study the actual reported cost value).

$P(F|\theta_i)$ is called the likelihood function. It is the probability or likelihood of observing the result or test observation, F, given the state of nature, θ_i. In other words, the likelihood function measures how likely event F will occur if the true state of nature is θ_i.

The result of Bayes' theorem, $P(\theta_i|F)$, is called the revised or posterior distribution, since it is the probability of θ_i after event F has occurred and been analyzed. It is the updating of $P(\theta_i)$ after more information from event F has been obtained. Thus, Bayes' theorem provides the framework to allow the redetermination of $P(\theta_i)$ based on additional information. This redetermination is important to the manager since decisions dependent on the knowledge of the probability of θ_i are better when based on additional information (White, 1971, p. 18). This is the method managers use informally in their minds to revise the probability of a certain state of nature based on additional information about that state of nature. This revision of judgment as additional information becomes available is thus handled in a formal explicit manner through the use of conditional probability and Bayes' theorem (Dyckman, 1969, p. 75).

Probability distributions can be described in terms of their parameters. A common example used before is the normal distribution, which is described by the two parameters, mean and variance. In Bayesian estimation, the parameter is looked upon as a random variable which has a prior distribution reflecting either the strength of one's belief about the possible values it can assume, or collateral information. Bayes' theorem can be used to combine prior information about a parameter with direct sample evidence to obtain the posterior distribution of the parameter. It will be shown later that using the parameters of the normal distribution, point estimates can be obtained by using this form of Bayesian updating of the distribution's parameters (Freund, 1971, p. 280).

<u>El-Sabban's Approach</u>

As shown in the last section, through Bayes' theorem a probability of a certain state of nature may be updated with the addition of new test data. M. Zaki El-Sabban proffered the idea that cost can be conceived to be a random variable that follows a specific distribution. By assuming that cost is a random variable with a certain distribution, a prior probability can be determined. This is done in the case of cost data with historical cost information. By using Bayes' theorem, the parameters of this probability distribution can be updated with new cost information, as the program office receives it, to form a posterior probability distribution. This updated posterior distribution will lead to an updated estimate-at-completion.

El-Sabban assumed a normal distribution for the cost data at any point along the path of the project. He stated that although this assumption may not be true, it may be considered a fair approximation (El-Sabban, 1973, p. 3). In the use of Bayes' theorem to derive an updated estimate, it is very convenient mathematically to assume a normal distribution. If the prior and sample distribution follow a normal distribution, then the posterior distribution will also be normally distributed. Determining the best distribution is a project by itself, and is determined by the author to be beyond the scope of this paper. Therefore, as in El-Sabban's appraoch, the normal distribution will be assumed primarily for ease of mathematical computations.

El-Sabbans' approach used the actual costs of the program (obtained from the cost performance report) as the prior distribution, and updated this prior distribution with the budgeted costs to derive a posterior probability distribution. He took the actual cost for the work performed (ACWP), budgeted cost for the work performed (BCWP), and the budgeted cost at completion (BAC), from the cost performance report. Letting u_o = BAC, X = ACWP, C = BAC BCWP, and cost be the random variable that at some point, o, along the project is normally distributed with a mean of u and a variance of o^2, then the likelihood function is:

$$P(\mu_o|\mu) = \frac{1}{\sqrt{2\pi}\,\sigma_o} \exp\left[\frac{-1}{2\sigma_o^2} (\mu_o - \mu)^2\right]$$

The prior probability is:

$$P(\mu) = \frac{1}{\sqrt{2\pi}\,C\sigma_a} \exp\left[\frac{-1}{2c^2\sigma_a^2} (\mu - c\bar{x})^2\right]$$

El-Sabban then derived the posterior distribution using the formula:

$$P(\mu|\mu_o) \propto P(\mu_o|\mu) \cdot P(\mu) \qquad \text{(El-Sabban, 1973, p. 4)}$$

New Approach

El-Sabban's approach has a major flaw in it. The situation as he describes is just the opposite of the actual situation. The prior probability should be developed from the budgeted costs for the program, not the actual costs, as El-Sabban theorized. This prior probability should then be updated with sample information; in this case the sample information is the actual cost data received each month in the cost performance report. Therefore, assuming a normal distribution for the prior and the sample probability distributions, the notation and assumptions are as follows:

Actual cost at point "a" $= \bar{x}_a$

Cost at point 'o' $\quad N(\mu,\sigma^2)$

Budgeted cost at completion $\quad N(\mu_b,\sigma_o^2)$

Again, assuming the prior probability is derived from the budgeted costs which are normally distributed, the prior probability distribution function is:

$$P(\mu_b) = \frac{1}{\sqrt{2\pi}\,\sigma_o} \exp\left[\frac{-1}{2\sigma_o^2} (\mu - \mu_b)^2\right]$$

The likelihood function is the probability of the actual cost at point "a", x_a, given some cost that is normally distributed with a mean of u:

$$P(\bar{x}_a|u) = \frac{1}{\sqrt{2\pi}\,\sigma_a} \exp\left[\frac{-1}{2\,\sigma_a^2} \cdot (\bar{x} - \mu)^2\right]$$

It is necessary to convert the sample evidence, the actual cost data at point "a", to an estimate of the cost at completion. This data will then be used to update the prior estimate. At point "a" in the program, the percentage of work completed is the amount of work performed divided by the total amount of work to be performed. The amount of work performed at point "a" is BCWP at point "a", which can be found in the cost performance report. The total amount budgeted for the project is the BAC, also taken from the cost performance report. The simpliest method of estimating the final project cost using the actual cost at point "a" ($\bar{x}_a$), is by multiplying $\bar{x}_a$ by the reciprocal of the percentage of work completed to date:

$$\text{New final estimate from sample} = \frac{\text{BAC}}{\text{BCWP}} \cdot \bar{x}_a$$

Since $\bar{x}_a$ is the actual cost at point "a", there is no variance (Actual cost $\sim$ $N(\bar{x}_a,0)$). However, a variance does appear in the extrapolation of the actual cost at point "a" to an estimated cost of completion. Letting c = BAC $\div$ BCWP, and $\mu_a = c\bar{x}_a$, the specific distribution of the new estimate-at-completion derived from the sample (actual costs) data is:

$$u \sim N(\mu_a,\ c_a^2\sigma_a^2)$$

To incorporate this sample data with the prior probability distribution into an eventual posterior probability, the likelihood function has to be changed to:

$$P(c\bar{x}_a \mid \mu) = \frac{1}{\sqrt{2\pi}\, c\sigma_a} \exp\left[\frac{-1}{2c^2\sigma_a^2}(c\bar{x}_a - \mu)^2\right]$$

$$= P(\mu_a \mid \mu)$$

Using the relationship used by El-Sabban to derive the posterior, $P(\mu \mid \mu_o) \propto P(\mu_o \mid \mu) \cdot P(\mu)$, and two pages of algebraic manipulation that can be found on pages 508-509 of Dyckman (1969) <u>Management Decision Making Under Uncertainty</u>, the posterior distribution looks like:

$$P(\mu \mid \mu_a) = \frac{1}{\sqrt{2\pi}\sqrt{\frac{\sigma_o^2\sigma_a^2}{\sigma_o^2+\sigma_a^2}}} \exp\left[-\left(\mu - \frac{\sigma_a^2\mu_o + \sigma_o^2 c^2 \bar{x}}{\sigma_o^2 + \sigma_a^2} \div \frac{2\sigma_o^2\sigma_a^2}{\sigma_o^2+\sigma_a^2}\right)\right]$$

This is the posterior probability density function, and it is normal with a mean:

$$\mu = \frac{(\sigma_a^2 c^2)\mu_o + c\bar{x}_a\sigma_o^2}{\sigma_o^2 + c^2\sigma_a^2}$$

And the variance is:

$$\sigma^2 = \frac{\sigma_o^2 \cdot \sigma_a^2 c^2}{\sigma_o^2 + \sigma_a^2 c^2}$$

The resulting posterior probability density function shows that when assuming the prior and sample density functions are normally distributed, the posterior is also normally distributed. This is extremely valuable since the posterior probability becomes the prior probability when the next monthly cost performance report (the next test sample) is received. Since the prior is again a normal distribution, the same formula can be used to determine the next posterior probability.

It should be noted that the posterior probabilities are just weighted averages of the prior mean and sample mean, with the variances of each mean determining the actual weights of each mean. In the specific case of cost data, the estimated cost is a weight of the budgeted costs and actual costs. The weights are the reciprocals of the respective variances, σ_o^2 and $c^2\sigma_a^2$. By letting $(c\sigma_a)^{-2} = h_o$, then:

$$\mu = \frac{c\mu_a h_a + \mu_o h_o}{h_a + h_o}, \text{ and the variance } = \frac{1}{h_a + h_o}$$

The new variables h_a and h_o are precision parameters, with the precision parameter of the posterior mean being the inverse sum of the sample and prior precision parameters. The constant, "c", is the same proportionality constant used before to estimate the cost at completion from the actual cost at point "a" in the program and is equal to BAC $\div$ BCWP.

Variance Calculation

El-Sabban's Approach: The major problem that has surfaced in the attempt to use Bayesian analysis has been the determination of the variances, both for the prior probability and the sample distributions. El-Sabban gave little guidance on how to determine the variances in his research effort. He stated that fair estimates may be obtained by one of two methods: 1) subjective estimates through personnel who are knowledgeable about the particular contract; 2) using the cost variances reported in the earlier cost performance reports, which "are indicative of the dispersion, e.g., assuming they loosely follow a normal distribution, then calculating σ in the usual manner." (El-Sabban, 1973, p. 7). El-Sabban did not give any further explanation on how either of his two methods were to be used. In using the first method, it is assumed that the analyst is knowledgeable enough about the program to estimate the variances. The second method uses variances reported on the cost performance report, which is incorrect. The variance given in the cost performance report is an accountant's variance which is not related to the variance statistic σ^2 of the normal distribution. The accountant's variance, if used, will give the exact opposite result to that desired. The posterior mean is weighted between the prior mean (budgeted costs) and the sample mean (actual costs). If the accountant's variance is used as the variance of the distribution, the larger the variance on the actual costs, the more weight will be placed on the budgeted cost (due to the inverse relationship of the precision parameters). However, the large accountant's variance means the budget is not being followed very closely and more weight should be placed on the sample mean than the budgeted mean. Therefore, although the accountant's variance on the cost performance report may indicate dispersion, it is not related to the distribution variance and cannot be used in Bayes' formula.

In the one example El-Sabban provided, he used percentages of the mean, $\sigma_a = 0.1\mu_a$ and $\sigma_o = 0.05\mu_o$ (Sabban, 1973, p. 8). El-Sabban gave no explanation as to why he used percentages of the mean to determine the variance, nor why he used different percentages for each mean (i.e., why did not he use either 0.1 or 0.05 for each). This point is critical, because the difference between the sizes of the two variances has much more to do with the weight of each mean to the new mean of the posterior distribution mean than the size of the variance as compared to the mean of the distribution. So El-Sabban has given very little insight into the development of the variance, a critical input to the calculation of a new estimate-at-completion.

<u>New approach for the prior distribution variance</u>: The situation of using one observation from a distribution with an unknown variance to

update the prior probability estimate is unique and not covered in the literature. For ease of explanation, the derivation of variances will be split. The derivation of the prior probability variance will be discussed first, followed by the discussion on the variance of the sample distribution.

If the general shape of the probability distribution to be assigned is known, its parameters can be developed from probability statements derived from historical data on the distribution. For example, this can be done with the normal distribution through the use of the following formula:

$$(\overline{x} - \mu) \div \sigma(\overline{x}) = Z$$

An example of the way this formula is used can be shown by assuming the mean is 50, and the statement can be made that the probability of the actual cost being between 40 and 60 is 50%

$(P(40 < \overline{x} < 60 : \mu = 50) \leq .50)$, then:

$$(60 - 50) \div \sigma(\overline{x}) = .67$$

$$\sigma(\overline{x}) = 0.15$$

This example can be better visualized through the diagram:

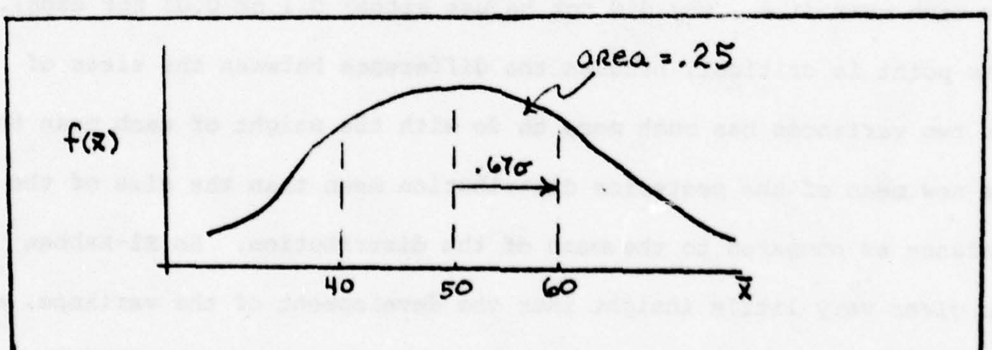


Figure 4. Normal Probability Distribution

Therefore, to determine the variance of the prior probability distribution in the case of cost data where a normal distribution is being assumed, historical data is needed concerning the accuracy of past budgeted costs. The judgement of the analyst should be used as to what set of historical facts is to be used for the basis of the probability statements. Some examples of data that can be used are: Department of Defense contracts in general, type of weapon system only, and past performance of the specific contractor. An example would be the data gathered by Peck and Scherer on twelve major weapon systems as reported by Anthony Babiarz in his thesis on weapon system cost growth (Babiarz, 1975, p. 1). Peck and Scherer found the standard deviation to be 170%, with a mean of 220% of the estimated costs. This study was done in the late 1950s.

Another study in the 1960s on 22 programs showed a mean of 226% of originally estimated costs (Babiarz, 1975, p. 2). And yet, another study by the Logistics Management Institute in 1971 on 139 programs found a mean of 150% of estimated costs, with a range from 79% to 284% of estimated costs (Fox, 1974, p. 364).

So historical information gathered in studies as shown above can be used to make statements concerning the probability distribution of the budgeted costs. From these probability statements, the variance of the distribution can then be determined. Using the data from the Logistics Management Institute study, a standard deviation can be determined. First, the probability statement is made that the probability that the cost is between 79% and 284% of the estimated cost is less than or equal to 98% ($P(.79 < \bar{x} < 2.84 : \mu = 1.5) \leq .98$). In other words:

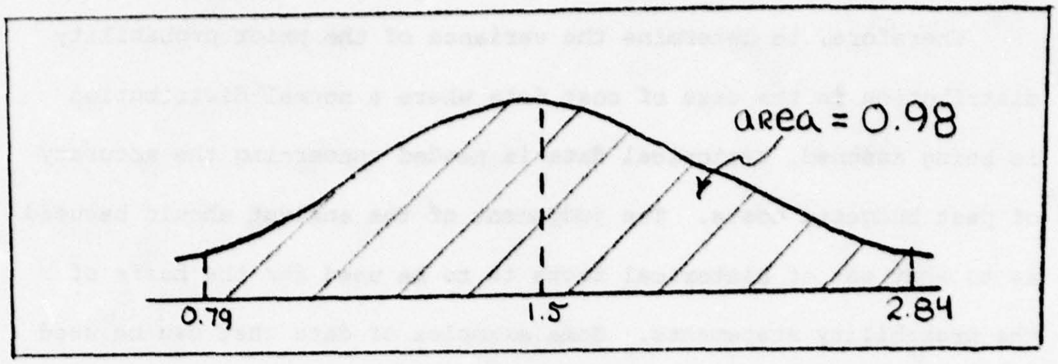


Figure 5. Variance Calculation

Then, using the formula $(\bar{x} - \mu) \div \sigma(\bar{x}) = Z$:

$$2.33\ \sigma_{\bar{x}} = 1.34$$

$$\sigma_{\bar{x}} = 0.575$$

<u>New Approach for the sample variance</u>: Developing the variance of the sample distribution in this specific case using cost data is more difficult than the variance of the prior probability distribution. The difficulty arises because in this sample of one the actual cost is the mean and there is no variance. A variance arises when the actual cost is extrapolated using the constant "c" to estimate the final cost.

Since the variance of the distribution is unknown, and there is no sample variance that can be substituted, the only method left is to make probability statements about the sample distribution. Thus, a variance is derived similar to the method used on the prior probability distribution variance. It is more difficult to determine the variance in this case because there is no explicit historical data available to make the probability statements as in the case of the prior probability distribution.

The probability statement that needs to be made concerns the uncertainty that $c\bar{x} = \mu_a$. The uncertainty is composed of two factors: 1) percentage of time to the end of the contract; and 2) accuracy of the budget to date. As the contract nears completion, the more accurate the mean that is derived from the sample distribution. In other words, it is intuitive that the closer the program gets to completion, the more weight should be placed on the sample distribution as opposed to the prior distribution (budgeted costs). This suggests that the sample distribution variance should get smaller as the program approaches completion.

The other piece of information that can be used is the accountant's variance reported in the cost performance report. Although it was shown previously that the variance on the report cannot be used directly, the accountant's variance can be used indirectly since it shows the amount that the program has deviated from the budgeted costs. It is then assumed that the larger the accountant's variance, the more weight that should be placed on the sample distribution. This suggests a smaller variance on the sample distribution as the accountant's variance gets larger.

How exactly to incorporate these two pieces of information in developing probability statements to derive a variance has to be left up to the individual program analyst. The method that was used in this research was to set the sample distribution variances for a given period of time, each new variance being smaller than the last one. The sample variance will be a percentage of the prior variance. As the program gets closer to its completion, the variances on the sample distribution get smaller, thus giving the sample distribution more weight than the prior distribution as the program progresses. In this research the variances were set so at the midpoint of the program the variance on the sample evidence was equal to the variance on the prior budgeted costs (Figure 4).

In this research, these percentages were modified by the program analyst if the accountant's variance from the cost performance report gets unusually large or small. This process is subjective and dependent on the views of the program analyst.

Example: 24 month program

Prior probability distribution facts:

Budgeted cost at completion (μ_o) = \$100,000
Standard deviation (σ_o) = 57.5% = 57,500

Sample distribution standard deviation:

First six months: $\sigma_a = 2.0\ \sigma_o$
Second six months: $\sigma_a = 1.5\ \sigma_o$
Third six months: $\sigma_a = 1.0\ \sigma_o$
Fourth six months: $\sigma_a = 0.5\ \sigma_o$

Figure 6. Sample Standard Deviation Calculations

Another subjective method that can be used to determine the sample variance is for the program analyst to update the subjective statements on the probability of the cost at completion made for the prior distribution. The analyst would update the probability statements used to determine the variances on the prior probability distribution to get a new probability statement for the sample distribution. The analyst could use the accountant's variance from the cost performance report and other pertinent information (i.e., changes in general and administration and labor rates) for inputs in updating the probability statements. From these updated statements a new variance can be calculated as before for use as the sample distribution variance.

The Use of Subjective Data

Some people feel uneasy in developing cost estimates using subjective cost information and statements. The probability statements used by the Bayesian method reflect the degree of rational belief held by the analyst in a given situation. The statements indicate the analyst's personal estimates of the likelihood of occurrence of the possible states. These probabilities usually reflect empirical and historical data to the extent incorporated in the beliefs of the analyst. Although two individuals might disagree on the probabilities assigned, if each has had similar experience, one would not expect substantial differences in the probabilities assigned.

It should also be remembered that this is an imperfect world, and that subjective probabilities are the only kind obtainable in many real world situations. In any actual problem objective, probabilities must be estimated on the information available, which is inevitably incomplete (Dyckman, 1969, p. 279-85).

Summary

This chapter dealt with Bayesian estimation and how the cost data output of a C/SCSC-validated accounting system can be used with this Bayesian estimation theory to develop continual estimates-at-completion for weapon system programs. The basis for Bayesian estimation is Bayes' theorem. Bayes' theorem describes conditional probabilities, and how the decision maker can update his probabilities by gathering additional information. Bayesian estimation goes one step further and provides a method of incorporating prior information that can be subjective in nature with sample evidence to produce a new estimate. When the prior and sample probability distributions are normally distributed, the resultant revised or posterior probability distribution will also be normally distributed.

El-Sabban applied this theory to the random variable cost. He described how the cost data from an accounting system that meets the Cost/Schedule Control Systems Criteria can be used to develop updated estimates-at-completion for a weapon system program. The author exposed several flaws in El-Sabban's application, and described a new approach in applying the Bayesian estimation technique that more accurately reflects the actual situation. Budgeted costs are used to form the prior probability distribution, and actual cost data from the program are used to revise or update this distribution, arriving at the posterior probability distribution. The mean of this new distribution is the new estimate-at-completion for the program. Chapter III ended with a discussion on how to determine the variances of both the prior and sample probability distributions, an integral part of the estimate-at-completion calculation. Chapter IV takes cost data from actual finished weapon system programs to determine if this Bayesian estimation technique is more accurate than two presently used estimation techniques.

IV. DATA ANALYSIS AND RESULTS

Chapter III answered the first research question: "Was the application of the Bayesian theory used by El-Sabban in his method correct?". It was shown that using cost as a random variable along with several reasonable assumptions allowed the use of Bayesian statistics to generate an estimate-at-completion. In Chapter IV, data from five actual programs are used to answer the second and third research questions: "Can the variance of the assumed distributions be calculated with enough accuracy that a useful estimate-at-completion can be generated?', and "Is the Bayesian method more accurate in generating estimates-at-completion than presently used methods?".

Data Collection Methodology

The data obtained from the five weapon system programs was taken directly from the monthly cost performance report that is submitted to the program office by the contractor. The cumulative actual costs to date, budgeted costs to date, and budget-at-completion were extracted from these reports. Only the top level (total program) data of the work breakdown structure was recorded. In actual operation, the Bayesian method, like all other methods, can be applied to data at all levels of the work breakdown structure (each part of the program).

The Bayesian method developed in Chapter III was applied to each program to generate continuous estimates-at-completion. The original budget was used for the first prior probability distribution, with the mean of the distribution being the budget-at-completion. Each succeeding month's data were used to update this prior distribution resulting in a new posterior distribution with the new (updated) mean being the new estimate-at-completion. Each month a new estimate-at-completion was generated, and this was plotted on a graph to show how the estimate-at-completion changed with the addition of each succeeding month's cost data.

The same data were used to develop estimates-at-completion using two other estimation techniques, the cumulative cost variance past performance factor and the non-linear regression method developed by Karsch (Karsch, 1974). As stated before, the cost variance past performance factor was selected because it is the most widely used method in developing estimates-at-completion for Department of Defense weapon system programs. The non-linear regression method was selected because it is the most sophisticated method developed to date. The results of each method on a program were plotted on the same graph, showing which method estimated closest to the actual final cost in any given month of the program.

For ease of calculation, a computer program using the Fortran language was developed by the author and is reprinted in Appendix B. This program was used only to speed up the data calculations because all the data (up to sixty months on one program) were processed at once. In an actual program office situation where the data comes in once a month, calculations could be done by hand.

Data

Data from five weapon system programs were used to test the Bayesian method and compare the results of the Bayesian method with the results from the other two methods. The first program cost performance data were taken from a report by A. Karsch (1974, p. 32). These data were taken from a completed research and development program at the Aeronautical Systems Division, Wright-Patterson AFB, Ohio. The data from Programs Two and Three were obtained through the Air Force Business Research Methods Office at Wright-Patterson AFB, and are also from actual research and development programs at the Aeronautical Systems Division. The data from Programs Four and Five were obtained from a report by Karsch, and represented data from two production programs (Karsch, 1976, p. 1). All five programs are typical acquisition programs that had additional work added during the course of the program. The first four programs were affected by varying degrees of cost escalation, and the fifth program experienced a slight underrun. The data for all five programs are presented in Appendix A.

Bayesian Method Modifications

Variance calculation: Several modifications of the Bayesian method were made during the course of the experiment. The first modification concerned the setting of variances for the sample distribution. The standard deviation of the prior distribution derived in Chapter III of 57.5% of the mean (budgeted cost-at-completion) was used. The mean of the sample distribution was $c \cdot \bar{x}_a$, or (BAC)·(ACWP)/BCWP. On the first attempt, the standard deviation of the sample distribution was set to decrease at set intervals in the program. The sample distribution standard deviation was calculated as a percentage of the original prior standard deviation, which in Program One was $41,055:

(Months 1-6)	$= 2.0\sigma =$	82,110
(Months 7-22)	$= 1.5\sigma =$	61,582
(Months 23-27)	$= 1.0\sigma =$	41,055
(Months 38-60	$= 0.5\sigma =$	20,058

Since the standard deviation determined the weight of the sample mean versus the prior mean, the objective here was to weight the sample mean equally with the prior mean after the 22nd month, and give the sample mean more weight after the 37th month.

This method of setting the sample standard deviation would have worked if the prior standard deviation was a constant $41,055 throughout the program. But the prior standard deviation itself is revised, and at each revision gets smaller. Therefore, this method of setting the sample standard deviation did not allow the sample mean to have enough weight in determining the revised mean (new estimate-at-completion). The result was that the Bayesian technique took a long time in recognizing the overrun condition in Program One, and was very late in predicting the actual final cost, especially in comparison with the other two techniques. Therefore, the method of setting the sample standard deviation was changed to the method presented in Chapter III, where the sample standard deviation is calculated as a percentage of the prior standard deviation for that month, not the original (first month) prior standard deviation.

<u>Increases in scope of work</u>: Another problem surfaced in the first run using the Bayesian technique. Many weapon system programs (including all five programs used in this experiment) have additions to their scope of work during the life of the program, which translates into an increase in the budget-at-completion. These changes or increases in the budget-at-completion are picked up one month late by the Bayesian method, biasing the revised mean (new estimate-at-completion). The Bayesian method is late in recognizing the increase because the prior probability distribution (last month's posterior probability distribution)

has no way of knowing about the increase in the budget-at-completion due to the additional work added to the program. So the prior probability mean had to be updated manually by adding the cost of the additional work directly to the mean of the prior probability distribution. This eliminated the biasing caused by the addition of new work. This can be seen in the computer program in Appendix B. Added to the prior probability mean is the expression "BAC - R". BAC is the new month's budget-at-completion and R is the budget-at-completion of the previous month. Of course, if there is no change, "BAC - R" will equal zero and the prior mean is unaffected.

Comparison of Methods

Program One: Program One is the program used by Karsch in the development of his non-linear regression model. Although the program was arbitrarily picked, it followed the prescribed curve, $b_1x^{b_2}$, very closely. Figure 7 shows the unconstrained non-linear regression method plotted with the cumulative variance past performance factor method and the Bayesian method. The non-linear regression method is by far the superior method with Program One, leading the past performance factor and Bayesian methods by as much as seven months in predicting the final actual cost. The Bayesian method lagged the past performance method slightly as expected until the 40th month (98% completion). The Bayesian method will always lag the past performance factor when there is an overrun (or underrun) that is increasing at an increasing rate, since the Bayesian method always gives at least some weight to the original budget (this phenomenon is further discussed later in this chapter).

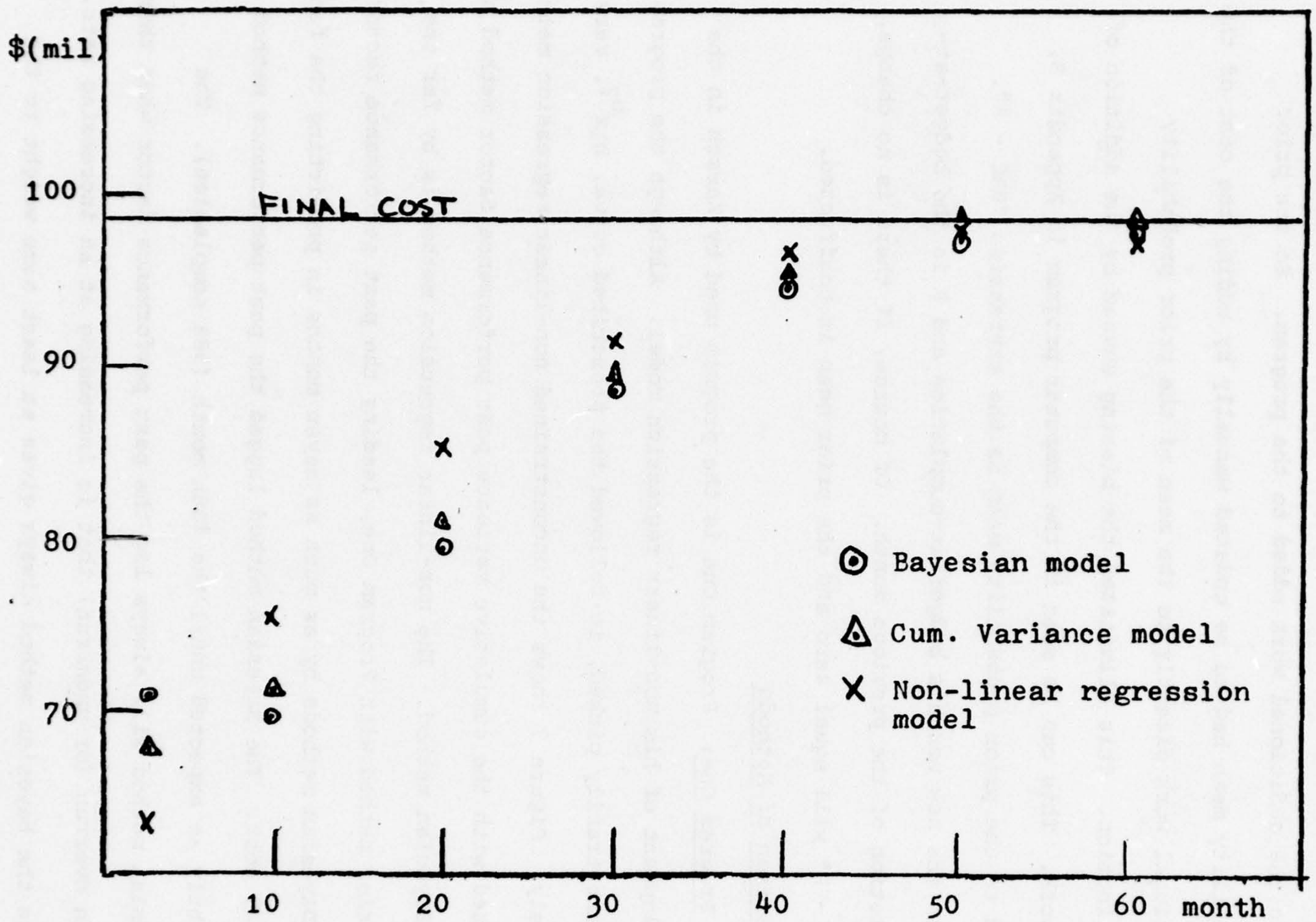


FIGURE 7. COMPARISON OF METHODS: PROGRAM ONE

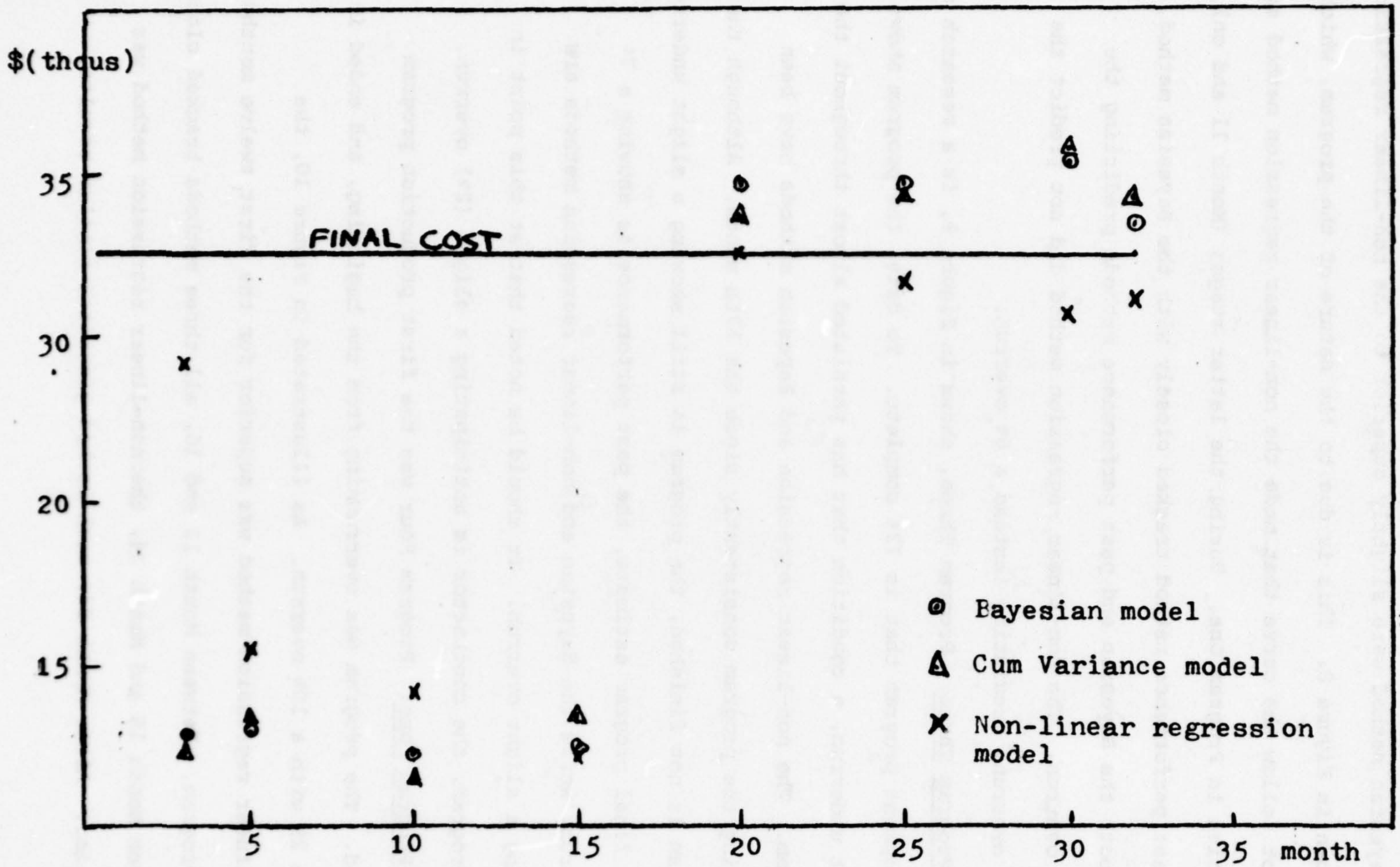


FIGURE 8. COMPARISON OF METHODS: PROGRAM TWO

Program Two: In Program Two, both the past performance factor and the Bayesian method were slightly superior to the non-linear regression as shown in Figure 8. This is due to the nature of the program, which did not follow the curve that made the non-linear regression method so effective in Program One. During the latter stages (Month 11 and on), the past performance method tracked closely with the Bayesian method, with both the Bayesian and past performance methods predicting the final overrun. The non-linear regression method did not predict the final overrun, predicting instead a 6% overrun.

Program Three: Program Three, shown in Figure 9, is a research and development program that is 73% complete. To date, the program shows a slight underrun, a condition that has persisted almost throughout the program. The non-linear regression and Bayesian methods have been tracking the program consistently since the 17th month. Although the program is not finished, the program is still showing a slight underrun. For a final program estimate, the past performance is showing a 7% underrun, while the Bayesian and non-linear regression methods are showing a slight overrun. It should be noted that at this point in the program, the contractor is anticipating a slight (2%) overrun.

Program Four: Program Four was the first production program tested. The program was overrunning from the beginning, and ended in Month 29 with a 10% overrun. As illustrated in Figure 10, the non-linear regression method was superior for the first twelve months of the program. Between Month 12 and 16, all three methods tracked closely. Between Month 16 and Month 24, the non-linear regression method was superior. After Month 24, each method generated similar results.

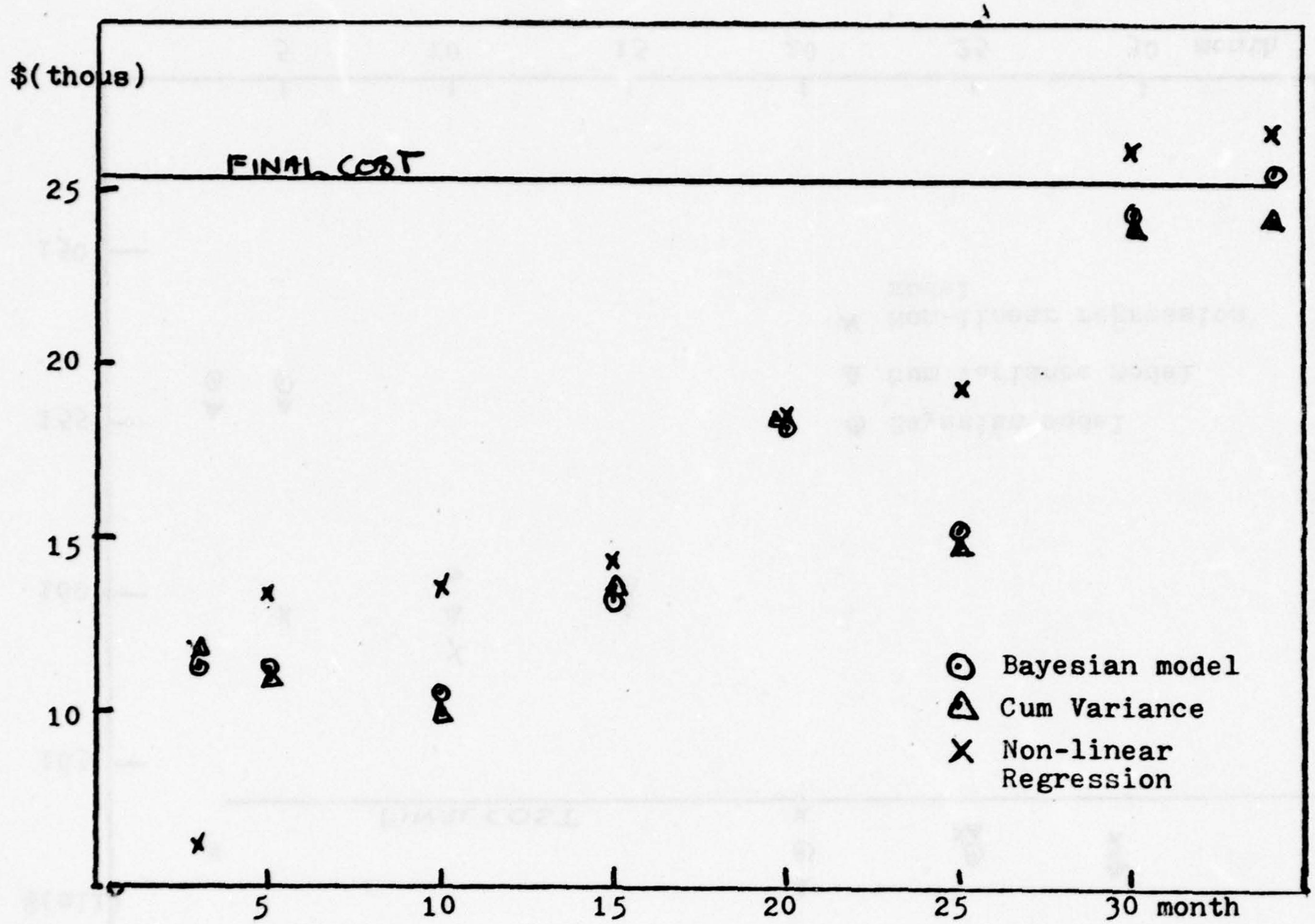


FIGURE 9. COMPARISON OF METHODS: PROGRAM THREE

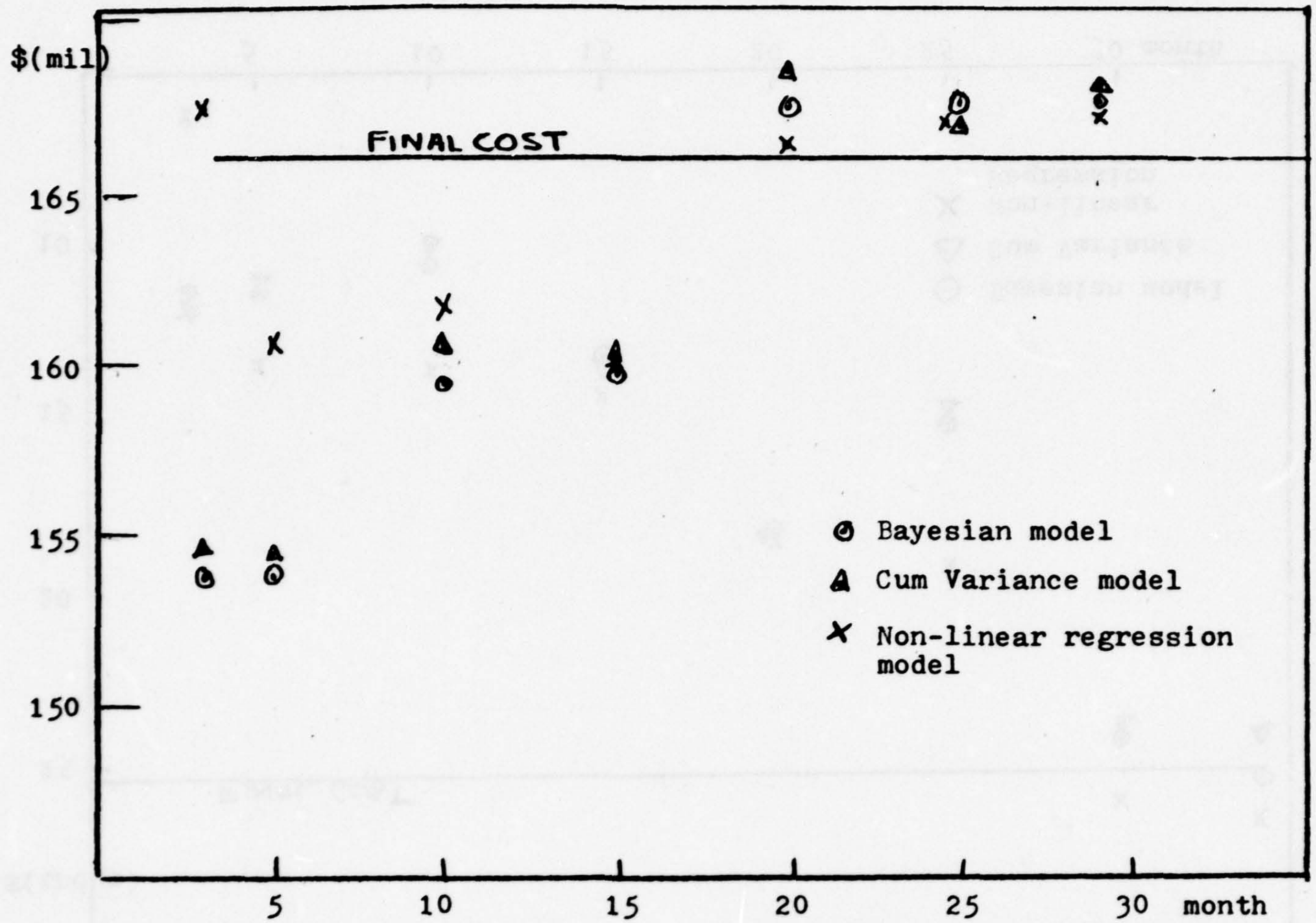


FIGURE 10. COMPARISON OF METHODS: PROGRAM FOUR

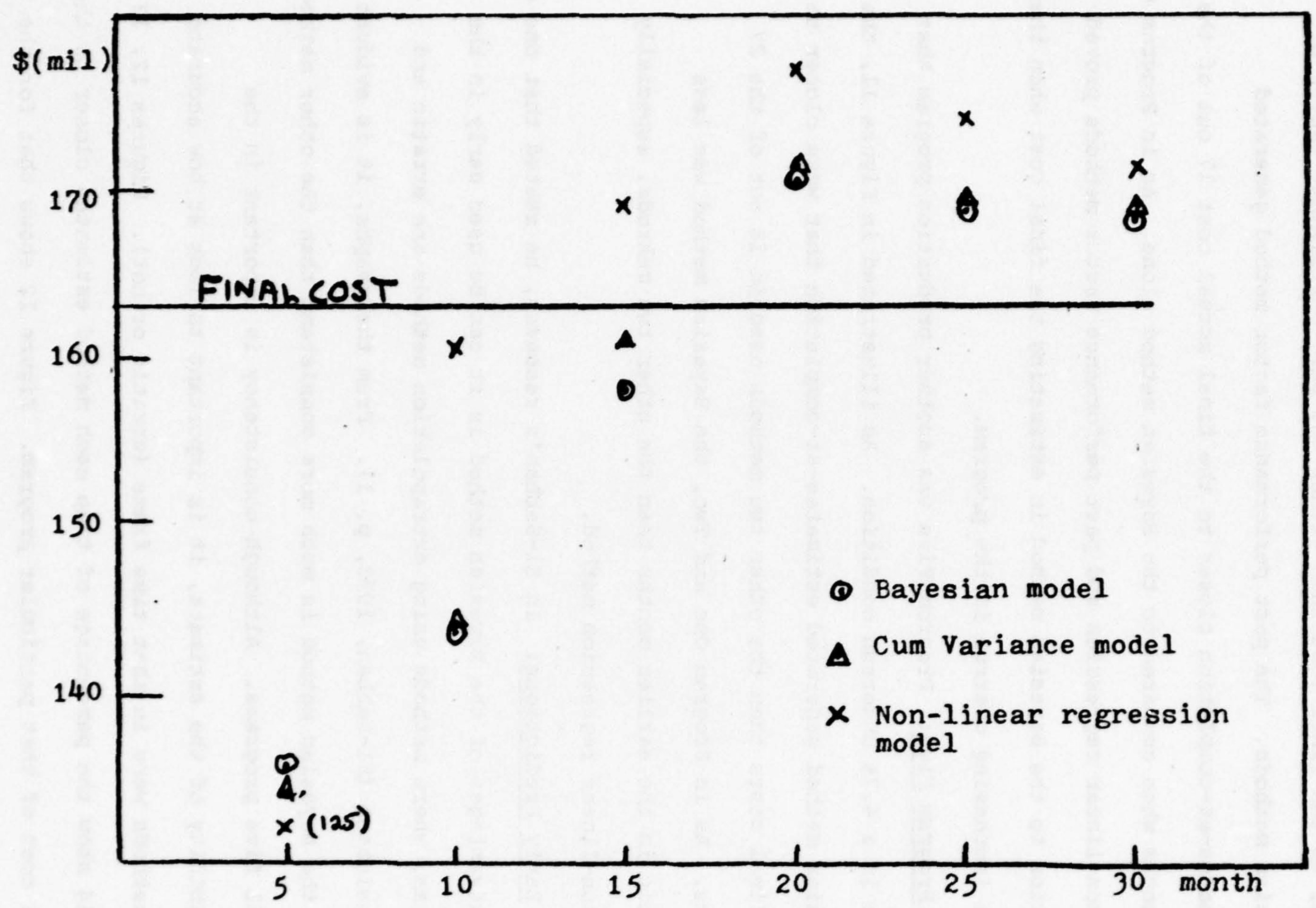


FIGURE 11. COMPARISON OF METHODS: PROGRAM FIVE

For 20 of the 29 months, the non-linear regression method estimated closer to the final cost than either the past performance factor or Bayesian methods. The past performance factor method generated estimates-at-completion closer to the final actual cost 17 out of the 29 months when compared to the Bayesian method alone. As in Program One, the non-linear regression and past performance factor methods proved superior to the Bayesian method in estimating the final cost when there is an increasing overrun in the program.

Program Five: Program Five was another production program that ended in a 4.7% underrun condition. As illustrated in Figure 11, the Bayesian method generated estimates-at-completion that were closer to the final costs than the other two methods used in 18 out of the 27 months. As in Program One and Two, the Bayesian method was less erratic in the earlier months than the other two mehtods, especially the non-linear regression method.

Early Predictions: In El-Sabban's research, he stated that one of the advantages of the Bayesian method is it can be used early in the program, where methods using extrapolation methods are erratic and inconsistent (El-Sabban, 1973, p. 1). From the graphs, it is evident that the Bayesian method is much more consistent than the other methods in all five programs. Although consistency is important in the credibility of the estimate, it is important to look at how accurate the methods were in that time frame (erratic or not). Figures 12, 13, and 14 show the percentage of time each method estimated closer to the final cost of that particular program. Figure 12 shows that for the total life of the program, the non-linear regression method was the superior estimating method in three out of the five programs, with the Bayesian

Method	Bayesian	Regression	Factor
Program			
One	7%	48%	45%
Two	35	31	34
Three	19	81	0
Four	15	70	15
Five	64	21	15

Percentage is the percent of time that the method estimated the final cost of the program closer to the actual final cost during the life of the program than the other two methods.

Figure 12. Summary of Data Analysis (entire program)

Method	Bayesian	Regression	Factor
Program			
One	42%	50%	8%
Two	50	50	0
Three	25	75	0
Four	0	100	0
Five	50	50	0

Percentage is the percent of time that the method estimated the final cost of the program closer to the actual final cost during the life of the program than the other two methods.

Figure 13. Summary of Data Analysis (first 12 months)

Method	Bayesian	Regression	Factor
Program			
One	100%	0%	0%
Two	83	0	17
Three	67	33	0
Four	0	100	0
Five	0	17	83

Percentage is the percent of time that the method estimated the final cost of the program closer to the actual final cost during the life of the program than the other two methods.

Figure 14. Summary of Data Analysis (first six months)

methods the superior method in the other two programs. Figure 13 shows that when only the first 12 months of the program are considered, the non-linear regression method is superior in three out of the five programs, with a tie between the non-linear regression and Bayesian method on the other two programs. When only the first six months of estimates are considered, the Bayesian method is superior on three out of the five programs, the non-linear regression method on one program, and the past performance factor on one program (Figure 14). This analysis tends to confirm El-Sabban's allegation that the Bayesian method is more consistent and accurate at the start of the program; but by the 12th month, the non-linear regression method overcomes the deficiency of small samples and is the more accurate method in three out of the five programs. The fluctuation is not as evident in the Bayesian method because the Bayesian method has the ability to use future budgeting information in the form of the prior probability distribution to smooth out these types of fluctuations.

Analysis of the Bayesian Method

Variance sensitivity: The major concern when using the Bayesian method is the subjectivity of the variance calculation, especially the variance of the sample mean. Theoretically, the most pleasing method of setting the variance would be to make probability statements each month concerning the accuracy of the sample distribution and its mean, and deriving the variance (and thus the standard deviation) from these probability statements. These statements could be made using historical data (data that indicated how accurate the sample mean is at that point in time in similar programs), or the statements could be made by the analyst using his personal knowledge of the program status. Although these procedures are theoretically pleasing, they are probably too complex for most program offices to use. Therefore, the author set the variances of the sample mean so that at the start of the program the sample mean has less weight than the prior mean, equal weight in the middle, and more weight than the prior mean near the end of the program.

Since this was rather arbitrary, a sensitivity analysis was done with data from Program One. The results were surprising. The Bayesian method is much less sensitive to the setting of the variance than the author had expected. When the variance of the sample distribution was set at half the variance originally set, the difference in estimates was 2% at most. In other words, the difference between the estimate with a "normal" sample variance and the estimate with the halved sample variance was 2%. When the variance was doubled the original sample variance (giving twice the weight to the prior mean as "normal"), the difference in estimates was less than 3%. Thus, the difference between estimates using the three sets of variances was less than 5%.

This analysis indicates that the setting of the variances is less critical than originally assumed by the author. This is a key point since the setting of the variances is the most subjective and weakest part of the Bayesian method. Although gross errors in setting the variances will distort the outcome, thus delaysing the point where the Bayesian method will predict an accurate estimate-at-completion, minor errors in setting the variance can be made with little harmful effect.

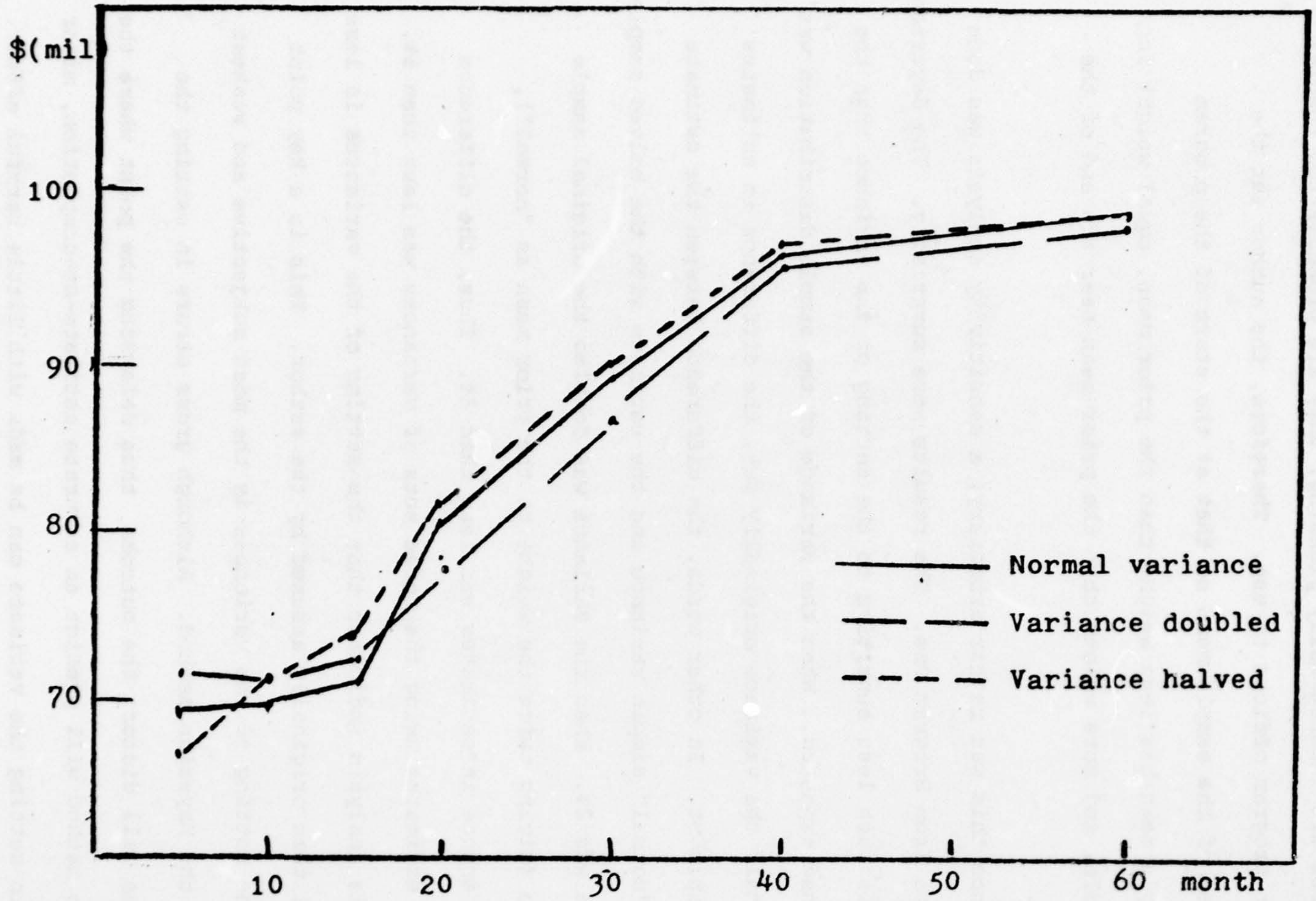


FIGURE 15. SAMPLE VARIANCE SENSITIVITY ANALYSIS

Prior Distribution Mean: When the mean of the original prior probability distribution was set at the outset of all five programs, the mean was set at 100% of the budgeted cost-at-completion. However, as indicated earlier in this paper, a study done by the Logistics Management Institute in 1971 discovered that the mean cost-at-completion was 150% of the original budget cost. Since the first program did experience the largest cost growth of the five programs, the Bayesian method was used on Program One with both the mean set at 100% and 150% of the budgeted costs for the prior probability distribution at the outset of the program. Figure 15 shows the result of changing from a 100% to a 150% mean.

Changing from the 100% to the 150% mean had little effect on Program One. Program One had a distinct underrun for the first ten months. This lengthy underrun nullified the effect the 150% mean had on biasing the final estimate-at-completion. Although Program One is just one program, an underrun at the beginning of the program is not unusual with weapon system programs. Karsch showed in his research that this behavior was predominant in his data samples, which encompassed aircraft, missile, and electronic system programs (Karsch, 1975, p. 12). The 150% mean was not attempted on the other programs, since they did not display large overruns at the end of the program.

Therefore, the use of the 150% mean is not recommended for two reasons: 1) typical early underruns that last as short as ten months negate the effect of the 150% mean; 2) better estimating and budgeting techniques, especially with the advent of C/SCSC, have negated the need to assume an inflated estimate-at-completion at the start of the program.

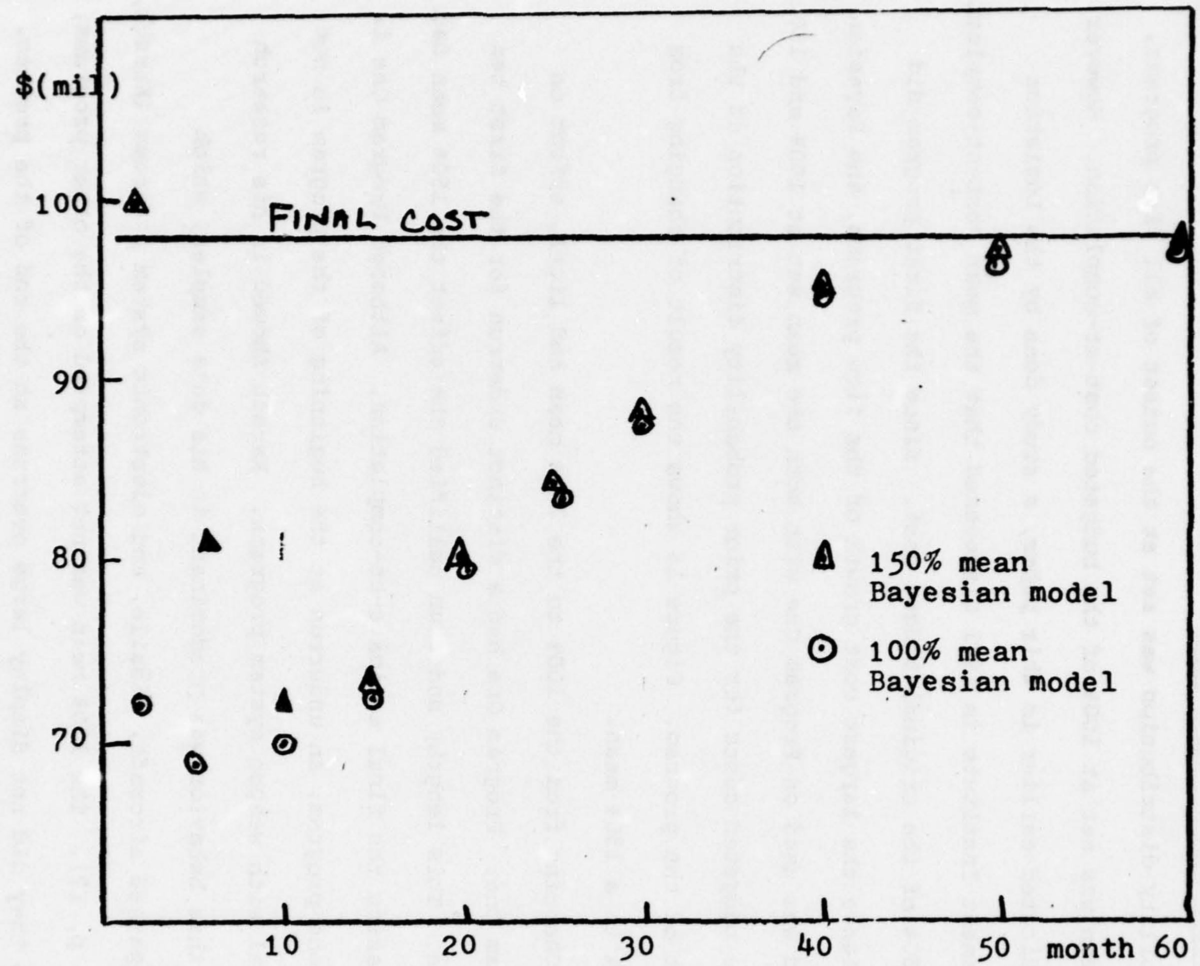


FIGURE 16. CHANGING PRIOR DISTRIBUTION MEAN

Modification of the Bayesian Model

In Karsch's research, he described a constrained approach to his non-linear regression model. The constrained approach fixes one of the parameters using an intelligent source, and the program is run letting only the second parameter vary (Karsch, 1974, p. 19). This method worked well when the actual outcome was used to fix the first parameter. As expected, this method worked better on Program One than any of the other methods used. Of course, this "constrained" method was essentially using perfect information. This is similar to holding the prior probability distribution mean constant at the actual outcome and letting only the sample mean vary in the Bayesian method.

A more logical method to constrain a parameter in regression analysis would be to use Bayesian analysis to update the constrained parameter. The parameter is constrained using an intelligent source, and then updated as new information is received using Bayesian updating. A method for applying this Bayesian technique to regression analysis is found in K. Sasaki's Statistics for Modern Business Decision Making (1968, p. 444).

A second method for combining non-linear regression analysis with Bayesian analysis is to use non-linear regression analysis to estimate the sample mean in the Bayesian model described in this paper. As stated in Chapter III, the Bayesian method uses the variable "c" to estimate the final cost. The variable "c" is the reciprocal of the percentage of work completed to date, or BAC/BCWP. This projection method is also the method used in the cumulative variance past performance factor method. In other words, the Bayesian technique in this paper takes a weighted average of the past performance factor method and the prior distribution mean to determine the new posterior distribution mean. This is why when there is an overrun that is increasing at an increasing rate, the Bayesian method will always lag the past performance factor method in estimating the actual final cost of the program.

If Karsch's findings are correct and the curve, $b_1 x^{b_2}$, is more appropriate to estimate final cost than a linear method, the sample mean could be estimated in the Bayesian model using non-linear regression analysis instead of the past performance factor method (BAC/BCWP):

Present Method: $$\mu = \frac{(c\bar{x})(h_a) + (\mu_o)(h_o)}{h_a + h_o}$$

Non-linear regression: $$\mu = \frac{(b_1 BAC^{b_2})(h_a) + (\mu_o)(h_o)}{h_a + h_o}$$

Figure 17 shows the result of using the non-linear regression method to determine the sample mean versus using the past performance factor method to estimate a sample mean. Program One data is used. Using the non-linear regression method, the final cost is estimated seven months sooner. Therefore, one modification to the Bayesian model presented in Chapter III is the inclusion of the non-linear regression method in determining the sample mean.

Summary

Chapter III described a Bayesian model that could be used to make estimates-at-completion on weapon system acquisition programs. The data used by this model is obtained from the monthly cost performance report. In Chapter IV, this model was tested using data from five weapon system programs at Wright-Patterson AFB, Ohio. The outcome of the Bayesian model was compared to two other estimation models presently used by the

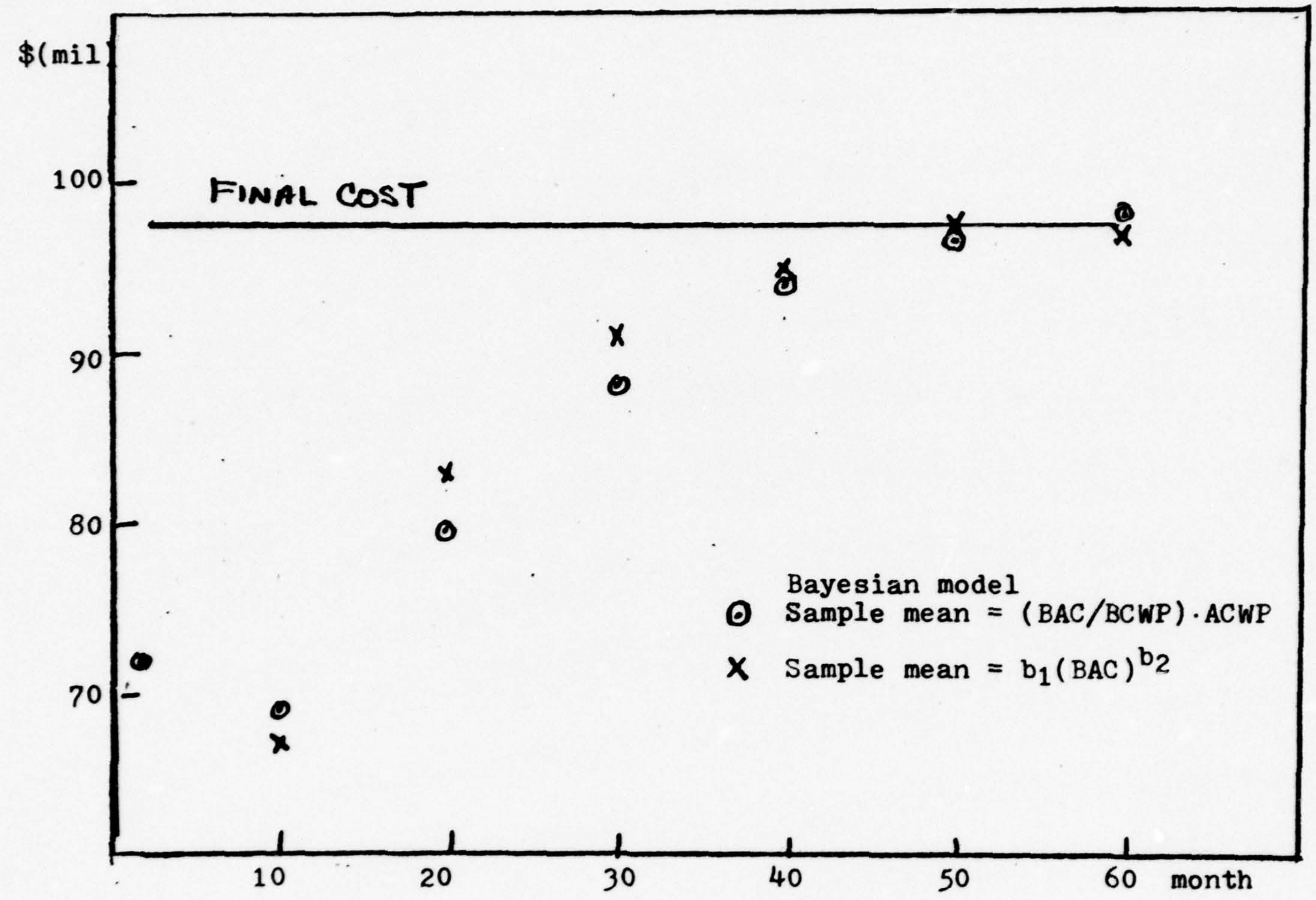


FIGURE 17. USING NON-LINEAR REGRESSION TO DETERMINE SAMPLE DISTRIBUTION MEAN

Department of Defense. The Bayesian model predicted the final cost the majority of the time on two of the five programs. The non-linear regression method predicted the final cost the majority of the time on the remaining three programs (the past performance factor method was never better than the other two methods in any of the five programs). Although in this comparison across five programs, the non-linear regression method was better than the Bayesian method, the Bayesian did show it is a viable estimation technique.

The Bayesian model displayed several advantages over the cumulative cost variance past performance factor method. First, the Bayesian model can be used at the start of the program where there is only a small amount of sample data available. The past performance factor method was very erratic and unreliable with small amounts of data. Second, the Bayesian method is not restricted to just a linear extrapolation as the past performance factor method is. The Bayesian model also can accept subjective judgments when developing an estimate-at-completion. The cumulative variance past performance factor method does have the advantage of predicting an overrun (or underrun) sooner if the overrun (underrun) condition continues to the end of the program at an increasing rate. This was the situation in Program One. As explained previously, this phenomenon happens because the past performance factor method does not give any weight to future budgeting as the Bayesian model does. Of course, most of this disadvantage can be eliminated in the Bayesian model if the analyst detects this increasing overrun (underrun) condition and gives a much smaller weight to future budgeting (i.e., the prior distribution mean).

The Bayesian model has similar advantages over the unconstrained non-linear regression method as it did the past performance factor method: inclusion of subjective information, use at the start of the program, and the inclusion of future budgeting information. The main advantage of the non-linear regression method is that it is not restricted to a linear extrapolation as is the past performance method. To eliminate most of the disadvantages of the non-linear regression method, Karsch introduced the constrained non-linear regression method. This allows the input of subjective information at the beginning of the program, and attempts to overcome the deficiency of small samples at the beginning of the program. However, as the author pointed out, the constrained parameter was selected using perfect information, and not allowed to vary. A method was proposed by the author to update this constrained parameter using Bayesian analysis.

Finally, several modifications were made to the Bayesian model described in Chapter III, including updating the prior probability distribution mean for increases in the budget, and the use of the non-linear regression method to estimate the sample mean. The latter proposal has the promise of combining the advantages of the Bayesian model with advantages of the non-linear regression method.

V. CONCLUSIONS AND RECOMMENDATIONS

The objective of this research was to extend the research begun by M. Zaki El-Sabban on a Bayesian technique to estimate costs-at-completion on weapon system programs. This objective was achieved by completing the model started by El-Sabban to a point where it can be used reliably in program offices that receive cost performance data on their programs. The model was evaluated using cost data from actual weapons system programs. The results of this evaluation indicate the model compares favorably with two models presently used by analysts in the Department of Defense to estimate costs-at-completion.

<u>Research Questions</u>

The first research question was:

"Was the application of the Bayesian theory used by El-Sabban in his method correct?".

This question was answered in Chapter III. One major fallacy and one major problem were uncovered with El-Sabban's method. The major fallacy concerned the construction of the prior and sample distributions. El-Sabban used actual cost to develop the prior distribution and budgeted costs to update the prior distribution. It was shown in Chapter III that the budgeted data should be used to construct the prior distribution, and updated to form the posterior distribution with actual costs received each month in the cost performance report. A major problem with El-Sabban's method was the lack of guidelines for determining the variances of the prior and sample distributions. This problem was overcome by using historical data to construct probability statements. The variance can then be derived from the probability statements.

The second research question concerned the use of those probability statements in the model:

"Can the variances of the assumed specific distributions be calculated with enough accuracy that a useful estimate-at-completion can be generated?".

This is a hard question to answer since the word "useful" is a very subjective term. In the five programs tested, the Bayesian model proved to be as reliable as the other two models tested. The variances on the sample distributions were set before the program was run using very simple rules. If probability statements (i.e., inputting intelligent subjective information) were used instead of these simple rules, the variances probably would have been even more representative of the situation. The sensitivity analysis done on Program One data indicated that large deviations, such as doubling the standard deviation, did not vary the result more than 3%. This indicates that variances can probably be estimated with enough accuracy to generate "useful" estimates-at-completion.

The final research question concerned the comparison of the Bayesian method with two presently used methods:

"Is the Bayesian method more accurate in generating estimates than presently used methods?".

Of the five tested, the non-linear regression method was the most accurate on three of the programs. The Bayesian method was the most accurate on the other two. Accuracy is defined as estimating the final cost closer than the other two methods the greatest percentage of the time during the life of the program. In this limited context, the answer to the research question is no. The Bayesian method was not more accurate than the non-linear regression method.

Conclusions

The author reached the conclusion that the Bayesian method is a viable technique in developing estimates-at-completion on weapon system programs. Of the five programs tested, the non-linear regression method was the most accurate, predicting the most accurate final cost the greater percentage of the time for three of the five programs examined. The Bayesian method predicted the most accurate final cost of the program the greater percentage of the time on the remaining two programs. The Bayesian approach did demonstrate the following advantages when compared with the non-linear regression and past performance factor methods:

1. It allows the use of subjective inputs from the analyst
2. It provides greater accuracy early in the program
3. It is not restricted to simple extrapolations of historical data
4. It can integrate prior knowledge information using formal statistics
5. Calculations are done without a computer
6. The Bayesian method does not discount the budgeting process, allowing the analyst to give some weight to the future budget.

This last point is very critical in a C/SCSC environment. A contractor that has a cost reporting system that follows the Cost/Schedule Control System Criteria puts a large emphasis on the budgeting process, including rebudgeting future work packages. When the contractor rebudgets, he will take into account that the program is underrunning or overrunning, and will adjust the budgeting process accordingly. Only the Bayesian method gives any weight to these future budgets.

It was also demonstrated that the Bayesian method can be integrated with other methods such as non-linear regression analysis to combine the advantages of both methods.

<u>Recommendations for Further Research</u>

This research, like most research, has generated as many questions as it has answered. The following are areas of research that could be explored to increase the accuracy of the Bayesian model:

1. Determine if the normal distribution is really applicable to cost, or if other distributions (e.g., the beta distribution) more accurately describe the distribution of the random variable cost.

2. Determine more precisely the variances of the budgeted cost data and the sample cost data throughout the program. This could be done by researching a large number of past programs to determine what the original weight should be on the budgeted data, and what weights should be placed on the actual cost data at different times during the program.

3. The third area of possible further research is the integration of the non-linear regression model with the Bayesian model. This paper described two methods in which non-linear regression analysis and Bayesian statistics might be integrated to combine the advantages of both. The first method was to use Bayesian analysis to continually update the constrained parameter in the constrained non-linear regression model developed by Arthur Karsch. The second method was to use non-linear regression analysis to determine the mean of the sample distribution in the Bayesian model developed in this research.

BIBLIOGRAPHY

A. References Cited

Barbriaz, A.S. and Peter Giedras. A Model to Predict Final Cost Growth in Weapon System Development Programs. Unpublished thesis, Wright-Patterson AFB: Air Force Institute of Tech, 1975.

Baumgartner, John S. "C/SCSC: Alive and Well", Defense Management Journal, 10, 33-35 (April 1974).

Busse, Daniel E. A Cost Performance Forecasting Model. Unpublished thesis. Maxwell Air Force Base, Ala: Air University, 1977.

Chase, Richard B. and Nicholas Aquilano. Production and Operations Management. Homewood, Ill: Richard D. Irwin, Inc., 1977.

Dyckman, T.R. Management Decision Making Under Uncertainty. London: MacMillan Co., 1969.

El-Sabban, M. Zaki. Forecast of Cost/Schedule Status Utilizing Cost Performance Reports of the Cost/Schedule Control Systems Criteria: A Bayesian Approach. Unpublished thesis. St. Louis, Mo.: U.S. Army Aviation Systems Command, January 1973. AD764576.

Fox, J. Ronald. Defense Management Journal, 10: (April 1974).

Fox, J. Ronald. "Funds Control Versus Costs Control", Army Logistician, 3: 4-7 (May-June 1971).

Freund, John E. Mathematical Statistics. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1971.

Holeman, J.B. "C/SCSC Analysis: The Time is Now", Defense Management Journal, 10: 39-41 (April 1974).

-------, A Product Improvement Method for Developing a Program Management Office Estimated Cost at Completion. Unpublished thesis. Fort Belvoir, Va: Defense Systems Management College, January 1975. AD A007125.

Karsch, O. Arthur. A Cost Performance Forecasting Concept and Model. Wright-Patterson AFB, Oh: Aeronautical Systems Division, November 1974. (Cost Research Report No. 117).

-------, Computer Program Input Instructions for Cost Performance Forecasting Model. Wright-Patterson AFB, Oh: Aeronautical Systems Division, February 1975. (Cost Research Report No. 117-A).

------. A Production Study Sequel to the Cost Performance Forecasting Concept and Model. Wright-Patterson AFB, Oh: Aeronautical Systems Division, August 1976. (Cost Research Report No. 132).

Kemps, Robert R. The Cost/Schedule Control Systems Criteria. Presentation. Office of the Assistant Secretary of Defense (Comptroller), December 1972.

------, "A Contractor Performance Measurement". Defense Industry Bulletin, 41-48 (Summer 1971).

Marella, Leonard S. The Effect of the Cost/Schedule Control Systems Criteria on Contractor Planning and Control. Unpublished dissertation. Fort Belvoir, Va: Defense Systems Management School, February 1973. AD774067.

Raiffa, H. and R. Schlaifer. Applied Statistics in Decision Theory. Cambridge, Mass: MIT press, 1968.

Sasaki, K. Statistics for Modern Decision Making. Belmont, Ca: Wadsworth Publishing Co., 1968.

Status, Trends and Projections. United States Army Management Engineering Training Agency, June 1974.

White, D. Decision Theory. Chicago: Aldine Publishing Co., 1971.

B. Official Documents

Contract Cost Performance, Funds Status, and Cost/Schedule Status Reports. Department of Defense Directive 7000.10. Washington, Department of Defense, 6 August 1974.

Cost/Schedule Control Systems Criteria Joint Implementation Guide. Air Force Systems Command/Air Force Logistics Command Pamphlet 173-5, 173-6, Army Material Command Pamphlet 37-5, Navy Material Command Pamphlet 5240. Washington: Department of Defense, 31 March 1972.

Performance Measurement for Selected Acquisitions. Department of Defense Directive 7000.2. Washington: Department of Defense, 25 April 1972.

C. Related Sources

Caldwell, Steven J. and Terry Earhart. Contractor Performance Measurement: Overrun Contracts with Selected Comments on Implementation of the C/SCSC. Unpublished thesis. Fort Belvoir, Va: Defense Systems Management School, April 1973.

Carlin, Gerald A. Cost Baselines in Excess of Contract Target Cost. Unpublished thesis. Maxwell AFB, Al: Air University, May 1975.

Cox, Wallace R. C/SCSC: Are We Getting Full Benefit by Max Use. Unpublished thesis. Fort Belvoir, Va: Defense Systems Management School, November 1974. ADA028401.

Durbron, Brian R. "C/SCSC Implementation Guide Reflects Evolution of the Program". Defense Management Journal, 10: April 1974.

Hofecker, W.J. Real Time Visibility of Value of Work Completed. Unpublished thesis. Fort Belvoir, Va: Defense Systems Management School, May 1975. LD34237A.

Lindsay, C.R. Decision Analysis for the Program Manager. Unpublished thesis. Fort Belvoir, Va: Defense Systems Management School, May 1974. LD32920A.

Ostdiek, Marion A. and Richard T. Estes. Cost/Schedule Control Systems Criteria: An Analysis of Managerial Utility. Unpublished thesis. Wright-Patterson AFB, Oh: Air Force Institute of Technology, 1975.

Slovic, Paul and Sarah Lichtenstein. "Comparison of Bayesian and Regression Approaches to the Study of Information Processing in Judgement". Organizational Behavior and Human Performance, 6: 649744 (1971).

Starch, S. The Use of Cost and Schedule Data By Program Managers. Unpublished thesis. Fort Belvcir, Va: Defense Systems Management School, May, 1975. LD34236A.

Zbylut, Robert S. A Case Study of the Usefulness of the Cost/Schedule Control System Criteria. Unpublished thesis. Wright-Patterson AFB, Oh: Air Force Institute of Technology, 1974.

APPENDIX A

Program One Data (Dollars in Thousands)

Month	ACWP	BCWP	BAC	Regression	Cum Variance	Bayesian
1	1263	1263	71901	---	---	71901
2	3213	3377	71901	61600	67600	71202
3	5163	5491	71901	63100	67200	70483
4	7113	7605	71901	63100	67200	69836
5	9063	9719	71901	64000	67000	69278
6	11013	11834	71901	64600	66900	68805
7	12963	13730	72055	67200	69600	68673
8	14913	15627	72208	69600	68900	68851
9	16893	17280	72208	73100	60600	69386
10	19205	19624	72239	74600	70700	69811
11	21616	21967	72147	75200	70900	70111
12	23837	24593	72799	74300	70600	70701
13	25716	26726	72813	72900	70000	70514
14	28738	29091	73439	74300	72500	71573
15	31108	30917	73538	75900	74000	72386
16	33310	32775	73538	77100	75000	73103
17	36132	35007	73602	78600	76000	74033
18	38958	37237	73659	80100	77100	75001
19	41836	40336	76950	84200	80000	78759
20	45355	43253	76950	84500	80700	79353
21	47965	45935	77050	84500	80500	79761
22	50893	48323	77050	84500	81200	80454
23	53904	50606	77598	85400	82700	81828
24	56293	52586	77801	85900	83300	82658
25	59001	54448	77802	86400	84300	83488
26	62021	57141	79054	88200	85900	85268
27	64208	59469	79353	88400	85700	85622
28	66501	61863	79823	88700	85800	85949
29	68682	64169	81507	90400	87200	87436
30	71573	66289	81789	90500	88300	88013
31	74508	68728	82241	91000	89100	88811
32	78408	71982	82618	91300	90000	89591
33	81105	73763	82329	91100	90500	89912
34	82978	75073	82081	90900	90700	90194
35	85641	76730	82270	91400	91900	91103
36	87636	77965	82295	91600	92500	91815
37	88919	79275	82399	91900	92400	92176
38	90428	80299	82491	92200	92900	92770
39	91134	81087	83496	93500	93800	93828
40	92096	83299	84275	94500	94200	94275
41	93247	82970	84538	94800	94100	94915
42	93877	83552	85972	95400	95500	95447
43	94719	83976	85089	95600	96000	95892

Month	ACWP	BCWP	BAC	Regression	Cum Variance	Bayesian
44	95228	84230	85126	95700	96200	96178
45	95555	84452	85058	95700	96300	96214
46	95895	84696	85126	95800	96400	96358
47	96218	84939	85295	96100	96600	96603
48	96453	84997	85295	96100	96800	96753
49	96486	85104	85461	96400	96900	96896
50	96565	85269	85539	96500	96800	96891
51	96787	85551	85686	96700	97000	96959
52	96896	85670	85686	96700	96900	96923
53	97219	85738	85777	96800	97300	97213
54	97513	85758	85778	96900	97600	97471
55	97581	85778	85778	96900	97600	97559
56	97611	85778	85778	96900	97600	97600
57	97864	85778	85778	97000	97900	97811
58	97882	85778	85778	97000	97900	97867
59	97885	85778	85778	97100	97900	97881
60	97887	85778	85778	97100	97900	97885

Program Two Data

Month	ACWP	BCWP	BAC	Regression	Cum Variance	Bayesian
1	228	179	17124	---	---	
2	513	483	17124	32444	18188	18061
3	915	922	17123	29145	16993	18086
4	1297	1250	17212	24557	17859	17866
5	2017	2578	17252	21531	18291	17954
6	2578	2535	17252	20010	17545	18021
7	3293	3155	17252	18803	18006	17966
8	3913	3912	17252	19051	17256	17978
9	4939	5166	17260	19212	16502	17756
10	5831	5956	17259	19121	16897	17383
11	6456	6595	17259	18994	16895	17228
12	7218	7202	17264	18804	17302	17126
13	8186	7932	17264	18310	17817	17185
14	9066	8620	17264	18049	18157	17379
15	9766	9101	17264	14483	18525	17618
16	10215	9709	17282	14593	18183	17897
17	11612	11178	32245	33243	33497	18058
18	12359	11834	32251	32953	33682	33259
19	13287	12647	32251	32861	33883	33677
20	13961	13306	32251	32677	33839	37888
21	15127	14285	32293	32234	34196	33977
22	15659	14876	32321	32323	34022	33999
23	16604	15616	32321	32083	34366	34183
24	17985	16817	32321	31876	34566	34438
25	19146	17877	32321	81628	34615	34556
26	21438	20077	32437	31783	34636	34609
27	22315	20832	32437	31489	34746	34700
28	23256	21656	32437	31484	34837	34789
29	24404	22442	32437	31221	35273	35112

Month	ACWP	BCWP	BAC	Regression	Cum Variance	Bayesian
30	25758	23300	32437	30920	35817	
31	26783	24028	32481	30968	36205	
32	30750	26290	32517		34068	

Program Three Data

Month	ACWP	BCWP	BAC	Regression	Cum Variance	Bayesian
1	161	128	12487	---	---	
2	229	212	12487	3595	11561	11185
3	303	276	12487	6048	11416	11320
4	535	402	12487	16604	9384	11228
5	688	616	12487	13633	11181	11346
6	851	736	12487	13723	10788	11313
7	1072	983	12487	12919	11440	11208
8	1424	1294	12462	12748	11316	11279
9	1735	1493	12462	13290	10726	11298
10	2071	1668	12462	13975	10036	11101
11	2563	2203	12462	14035	10716	10774
12	3100	2855	14058	14947	12951	10756
13	3471	3362	14058	14991	13616	12536
14	3773	3539	14058	14859	13186	12869
15	4574	4422	14058	14607	13591	12966
16	4890	4939	14058	14307	14202	13160
17	5421	5430	18775	18569	18806	13681
18	6079	5666	18775	18720	17497	18602
19	6479	6524	18646	18585	17998	18049
20	6853	6765	18646	18508	18412	14958
21	7490	7304	18646	18472	18185	18182
22	8341	7646	18646	18602	17092	18183
23	9149	7968	18646	18830	16239	14638
24	10007	8373	18675	19254	15624	16938
25	10726	10726	18675	19610	15259	15867
26	11530	11143	18675	19561	18050	15386
27	12278	12307	25431	26252	25490	18022
28	13171	12758	25431	26094	25635	18044
29	14033	13543	25431	26098	25546	25352
30	15002	14458	25431	26082	24509	24778
31	15983	15384	25431	26168	24496	24579
32	17224	16283	25525	26207	24129	24523
33	19665	18723	25525			24580

Program Four Data (Dollars in Thousands)

Month	ACWP	BCWP	BAC	Regression	Cum Variance	Bayesian
1	29529	28501	144007	---	---	145605
2	34694	33210	144007	---	---	148432
3	41547	39326	146371	167100	154600	153700
4	46663	44538	146296	158500	153500	153300
5	55015	51610	144666	160600	154200	153700
6	61368	57855	145337	159300	154200	154200
7	29810	65600	147577	159800	156000	155900
8	77369	73232	146178	157100	154400	154600
9	85209	80184	146490	156200	155700	155500
10	96090	89729	149890	171600	160500	159400
11	103604	97119	148460	160300	158800	158500
12	111725	104341	149161	160700	159700	159100
13	119309	112545	149293	159900	158300	158900
14	124272	118060	149551	159700	158700	158900
15	132190	123337	149575	160100	160300	159600
16	140879	130508	149576	160600	161500	160500
17	147617	134927	151851	164300	166900	164800
18	151692	138114	151900	165100	166800	165900
19	155551	140846	151933	165800	167800	166900
20	157389	142009	151772	166200	168200	167400
21	161053	144794	151704	166600	168700	168500
22	161630	145803	151833	167100	168300	168400
23	162102	146622	151666	167000	167700	167800
24	162644	146856	151389	166700	166500	166700
25	163280	148185	151414	166800	166800	166800
26	163373	148496	151415	166800	166600	166600
27	163957	148795	151581	167000	167000	166700
28	164315	149508	151569	166900	166600	166600
29	165471	149469	151540	167000	167800	167500

Program Five Data (Dollars in Thousands)

Month	ACWP	BCWP	BAC	Regression	Cum Variance	Bayesian
1	579	315	155606	---	---	195732
2	1056	1364	162421	---	---	164146
3	1722	2124	156266	25400	126700	132950
4	3741	4547	160075	114500	131700	132712
5	4872	5898	160080	125300	132200	132329
6	7286	9249	165667	113900	130500	131988
7	10911	13747	165791	121700	131600	131693
8	16210	18582	164741	161700	143700	141098
9	22168	25259	164819	166000	144600	143954
10	19412	33415	164526	160800	144800	144017
11	36921	42446	173747	160800	151100	152589
12	46660	53214	169956	154800	149000	148868
13	58105	63432	170088	162000	155800	151093
14	67501	72683	171214	166900	159000	155613
15	78506	83043	170402	169100	161100	157976
16	88671	92660	170780	169700	161700	160003
17	98907	103998	170743	169200	162400	161175
18	1095811	112909	171502	171300	166400	164190
19	119907	112535	173797	174600	170100	168277
20	127975	131719	176390	176700	171400	171121
21	135453	140560	175360	174100	169000	169189
22	142145	145772	175687	174200	171400	170955
23	145792	150203	175675	173700	170500	170601
24	142291	154001	175850	173400	170500	170532
25	153177	160380	177432	174000	160500	169984
26	156187	162313	177321	173400	170600	140479
27	157441	163965	177161	172800	170100	170153
28	158020	165422	177139	172300	160200	169396
29	161126	169223	176688	171300	168200	168376
30	163175	170803	176810	171100	168900	168830

Program One Data - Variance Sensitivity, Mean Sensitivity Analysis

Month	Variance Doubled	Variance Halved	150% Mean	Non-linear Reg. to Calculate Sample Mean
1	71901	71901	100661	---
2	71695	70155	94210	---
3	71455	68880	88889	69841
4	71207	68065	84561	68492
5	70962	67556	81058	67598
6	70724	67234	78229	66998
7	70609	67897	75245	67167
8	70592	68697	73448	68012
9	70592	69909	72569	69584
10	70633	70444	72024	71149
11	70577	70704	71615	72331
12	71227	71264	71942	73388
13	71125	70508	71377	73248
14	71893	72439	72363	73969
15	72202	73532	72963	74631
16	72449	74294	73502	75391
17	72871	75435	74330	76422
18	73346	76526	75224	77593
19	77283	81919	79926	81904
20	77624	81132	80161	82703
21	78007	80798	80351	83325
22	78635	81078	80749	83912
23	79987	82887	82250	84930
24	80850	83408	82971	85516
25	81542	84129	83640	85958
26	82454	84353	85622	87705
27	83789	85655	85949	88202
28	84562	86121	86101	88686
29	85016	88332	86636	90385
30	85956	88595	87754	90583
31	87048	89497	88909	91017
32	88014	90271	89827	91347
33	88227	90184	89886	91078
34	88478	90367	90057	90865
35	89336	91722	91129	91227
36	89995	92371	91841	91426
37	90586	92525	92241	91715
38	91833	92966	92857	92121
39	93842	94795	94649	93425
40	94796	95007	95063	94440

Month	Variance Doubled	Variance Halved	150% Mean	Non-linear Reg. to Calculate Sample Mean
41	95165	95272	95283	94780
42	95753	95894	95868	95366
43	95980	96086	96040	95576
44	96147	96268	96243	95682
45	96126	96174	96173	95682
46	96317	96431	96402	95789
47	96641	96782	96749	96072
48	96716	96790	96782	96094
49	96969	97050	97036	96372
50	96998	96959	96981	96490
51	97115	97087	97095	96687
52	97015	96924	96950	96697
53	97230	97334	97291	96797
54	97383	97524	97487	96879
55	97482	97577	97562	96895
56	97546	97609	97601	96899
57	97705	97849	97811	96979
58	97793	97880	97867	96995
59	97839	97884	97881	97079
60	97863	97886	97885	97095

APPENDIX B

<u>Fortran Bayesian Model Program</u>

```
      PROGRAM ESTIMATE (INPUT, OUTPUT, TAPE 5 = INPUT, TAPE 6 = OUTPUT)

      U = initial estimate for cost-at-completion

      W = initial variance

      R = initial BAC

 7    READ (5,1)M ACWP BCWP BAC

 1    FORMAT (I2, 3F5.0)

      IF (EOF(5))50, 10

10    C = BAC/BCWP

      IF (M.LT.1) GO TO 2

      IF (M.LT.7) GO TO 3

      IF (M.LT.22) GO TO 4

      IF (M.LT.38) GO TO 5

      T = 0.5*W

      GO TO 6

 3    T = 2*W

      GO TO 6

 4    T = 1.5*W

      GO TO 6

 5    T = W

 6    A = T*T*(U+(BAC-R))

      B = C*ACWP*W*W

      E = A+B

      D = W*W+T*T

      U1 = E/D

      W1 =SQRT(W12)
```

```
      W = W1
      U = U1
      PRINT 13, M, U, W
      R = BAC
      GO TO
 2    PRINT 40
40    FORMAT ("IMPROPER MONTH DATA")
13    FORMAT ("MONTH=", I2, 3X, "U=", F9.2,3X, "W=", F8.2)
50    CONTINUE
      END
```

VITA

Richard A. Hayes was born on March 13, 1951, in Rockville, Maryland. He received his Bachelor of Science in Physics from the United States Air Force Accademy in June 1973. His first assignment was in the Munitions System Project Office, Armament Development and Test Center, Eglin AFB, Florida. For two years, he served as a program manager on several programs. The third year he became a program analyst in the Program Control Division of the Munitions System Project Office. This job included analyzing cost performance reports on all programs with C/SCSC and CSSR. In June 1976, he was assigned to AFIT, School of Engineering. After graduating with a Master's degree in System Management in December 1977, he will be assigned to the Strategic Systems System Project Office at Wright-Patterson AFB as a program analyst.

Permanent Address: 3871 Roslin Avenue
Kettering, Ohio 45429

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE — READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFIT/GSM/77D-21
2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): AN EVALUATION OF A BAYESIAN APPROACH TO COMPUTE ESTIMATES-AT-COMPLETION FOR WEAPON SYSTEM PROGRAMS
5. TYPE OF REPORT & PERIOD COVERED: MS Thesis
6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Richard A. Hayes
8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Air Force Institute of Technology(AFIT-EN) Wright-Patterson AFB, Ohio 45433
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: December 1977
12. REPORT DATE: December 1977
13. NUMBER OF PAGES: 92
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)
15. SECURITY CLASS. (of this report): Unclassified
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report): Approved for public release; distribution unlimited.
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)
18. SUPPLEMENTARY NOTES: APPROVED FOR PUBLIC RELEASE AFR 190-17. JERRAL F. GUESS, CAPT, USAF Director of Information
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

C/SCSC	Bayesian statistics
Cost analysis	Cost performance report
Cost estimation	Estimation techniques

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The Bayesian model developed by the author to predict costs-at-completion on weapon system programs is an extension of research done by M. Zaki El-Sabban. The model assumes cost is a random variable and is normally distributed. Budgeted costs are used to develop the prior probability distribution. Actual cost information is used for the Bayesian updating of the probability distribution. The mean of the updated probability distribution

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

is the new estimated cost-at-completion for the program. The model was compared with a non-linear regression model and a linear extrapolation model on five weapon system programs. On three of the programs the non-linear regression model estimated the final cost the greater percentage of the time. On the remaining two programs the Bayesian model estimated the final cost the greater percentage of the time. The Bayesian model demonstrated several advantages over previous models: use at the beginning of the program, inclusion of subjective information, and giving weight to future program budgets.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)