Calhoun: The NPS Institutional Archive

Theses and Dissertations | Thesis Collection

1982

An application of Rayleigh curve theory to contract cost estimation and control.

Watkins, Harry.

Monterey, California. Naval Postgraduate School

http://hdl.handle.net/10945/20073

Calhoun is a project of the Dudley Knox Library at NPS, furthering the precepts and goals of open government and government transparency. All information contained herein has been approved for release by the NPS Public Affairs Officer.

Dudley Knox Library / Naval Postgraduate School
411 Dyer Road / 1 University Circle
Monterey, California USA 93943

http://www.nps.edu/library

SCHOOL

NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

AN APPLICATION OF RAYLEIGH CURVE THEORY
TO CONTRACT COST ESTIMATION AND CONTROL

by

Harry Watkins III

March, 1982

Thesis Advisor: Dan C. Boger

Approved for public release; distribution unlimited

T204929

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)

An Application of Rayleigh Curve Theory To Contract Cost Estimation and Control

5. TYPE OF REPORT & PERIOD COVERED

Master's Thesis ; March, 1982

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)

Harry Watkins III

8. CONTRACT OR GRANT NUMBER(s)

9. PERFORMING ORGANIZATION NAME AND ADDRESS

Naval Postgraduate School
Monterey, California 93940

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS

Naval Postgraduate School
Monterey, California 93940

12. REPORT DATE

March, 1982

13. NUMBER OF PAGES

83

14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)

15. SECURITY CLASS. (of this report)

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

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Contract cost estimation
Minitab

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

Cost growth is a major problem in defense systems acquisition. Since 1969 the DOD has underestimated the ultimate costs of major systems by more than 50 percent. Consequently, the importance of contract costs has risen greatly in recent years to the point that costs are now officially equated to technical performance in importance. A body of knowledge of the structure and models of the behavior of contract costs and contract

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

performance within DOD is desired. This paper develops a simplified methodology for the systematic analysis and prediction of cost and schedule variables from an existing data base. The methodology is applied to actual DOD contract data using the interactive computing system MINITAB.

Approved for public release; distribution unlimited.

An Application of Rayleigh Curve Theory To Contract Cost Estimation and Control

by

Harry Watkins III
Lieutenant, United States Navy
B.S., University of Louisville, 1976

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
March 1982

c.1

DUDLEY KNOX LIBRARY
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CALIF. 93940

ABSTRACT

Cost growth is a major problem in defense systems acquisition. Since 1969 the DOD has underestimated the ultimate costs of major systems by more than 50 percent. Consequently, the importance of contract costs has risen greatly in recent years to the point that costs are now officially equated to technical performance in importance. A body of knowledge of the structure and models of the behavior of contract costs and contract performance within DOD is desired. This paper develops a simplified methodology for the systematic analysis and prediction of cost and schedule variables from an existing data base. The methodology is applied to actual DOD contract data using the interactive computing system MINITAB.

TABLE OF CONTENTS

I. INTRODUCTION - 10

A. GENERAL OVERVIEW - - - - - - - - - - - - - - - - - - 10

B. PROGRAM MANAGEMENT - - - - - - - - - - - - - - - - - 11

1. Historical Perspective - - - - - - - - - - - - 12

2. Risks - 13

3. DSARC - 14

4. Cost/Schedule Control System Criteria - - - - 15

C. COST GROWTH IN DEFENSE CONTRACTS - - - - - - - - - 17

D. NATURE OF THE PROBLEM - - - - - - - - - - - - - - - 18

II. METHODOLOGY - 20

A. GENERAL - 20

B. RAYLEIGH-NORDEN CURVE THEORY - - - - - - - - - - - 21

C. SCATTER PLOT - - - - - - - - - - - - - - - - - - - 24

D. REGRESSION - 26

E. SUMMARY - 27

III. STRUCTURE OF DATA BASE - - - - - - - - - - - - - - - 28

IV. A FORMAT FOR MINITAB APPLICATION - - - - - - - - - - 35

A. MINITAB - 35

B. ANALYTICAL COMMANDS - - - - - - - - - - - - - - - - 35

1. Regress - 36

2. Plot - 37

C. SUPPORTING COMMANDS - - - - - - - - - - - - - - - - - 41

1. Join - 41

2. Pick - 43

D. SUMMARY - 44

V. DATA ANALYSIS - 46

A. THE ADAPTED MODEL - - - - - - - - - - - - - - - - - - 47

B. TRANSFORMATION OF THE MODEL - - - - - - - - - - - - - 48

C. CONTRACT CONTROL AND ESTIMATION - - - - - - - - - - 50

D. ANALYSIS OF RESULTS - - - - - - - - - - - - - - - - - 58

E. INFLATION - 64

F. SUMMARY RESULTS - - - - - - - - - - - - - - - - - - - 69

1. Helicopter Engine - - - - - - - - - - - - - - - - - 69

2. Cruise Missile - - - - - - - - - - - - - - - - - - 74

VI. CONCLUSION - 79

A. GENERAL - 79

B. SUMMARY - 80

LIST OF REFERENCES - 82

DISTRIBUTION LIST - 83

LIST OF TABLES

1. Minitab Column Format - - - - - - - - - - - - - - - - 31

2. IMPLICIT PRICE DEFLATORS (CY 1972 = 100) - - - - - - 32

3. MINITAB DATA COLUMNS - - - - - - - - - - - - - - - - - 33

4. Definitions of Terms in 'REGRESS' Results - - - - - - 39

5. TRANSFORMED DATA - - - - - - - - - - - - - - - - - - - 50

6. INPUT DATA FOR INITIAL ESTIMATE - - - - - - - - - - - 54

7. INPUT DATA AFTER 5 QUARTERS INTO CONTRACT - - - - - - 56

8. OUTPUT DATA - 60

9. DEFLATED INPUT VALUES - - - - - - - - - - - - - - - - 66

10. OUTPUT DATA (DEFLATED) - - - - - - - - - - - - - - - 67

11. Input Data - 71

12. Output Data (Helicopter Engine) - - - - - - - - - - - 71

13. INPUT DATA (Cruise Missile) - - - - - - - - - - - - - 75

14. OUTPUT DATA (Cruise Missile) - - - - - - - - - - - - 76

LIST OF FIGURES

1. Current manpower utilization curve - - - - - - - - - 23

2. Life-cycle method of project control (Early warning technique) - 25

3. Sample Results of 'REGRESS' Command - - - - - - - - - 38

4. Sample Results of 'PLOT' Command - - - - - - - - - - 42

5. Sample Results of 'MPLOT' Command - - - - - - - - - - 43

6. RESULTS OF 'JOIN' COMMAND - - - - - - - - - - - - - - 45

7. INACWP VERSUS PERIOD - - - - - - - - - - - - - - - - - 47

8. LINEAR FORM OF RAYLEIGH-NORDEN CURVE (Ln(Y/T) versus TIME Squared) - - - - - - - - - - - - - - - - - 49

9. REGRESSION STATISTICAL RESULTS (ONE YEAR INTO CONTRACT) - 55

10. GRAPHICAL RESULTS OF REGRESSION AFTER 4 QUARTERS - 56

11. GRAPHICAL RESULTS OF REGRESSION AFTER 5 QUARTERS - 57

12. REGRESSION RESULTS AFTER 5 QUARTERS - - - - - - - - - 58

13. REGRESSION RESULTS AFTER 14 QUARTERS - - - - - - - - 61

14. REGRESSION RESULTS AFTER 14 QUARTERS (Graphical Presentation) - 62

15. REGRESSION RESULTS (DEFLATED) - - - - - - - - - - - - 68

16. REGRESSION RESULTS (DEFLATED) (GRAPHICAL PRESENTATION) - 69

17. INACWP versus PERIOD (Helicopter Engine) - - - - - 72

18. Regression Results (Helicopter Engine) - - - - - - - 73

19. INACWP versus PERIOD (Cruise Missile) - - - - - - - - 74

20. REGRESSION RESULTS (Cruise Missile) - - - - - - - - - 77

I. INTRODUCTION

A. GENERAL OVERVIEW

Cost growth is a major problem in defense systems acquisition. In recent congressional hearings, Jack Brooks, Chairman of the House Government Operations Committee, stated: "Since 1969 the DOD has underestimated costs of all major systems by more than 50 percent"[Ref. 1]. The annual cost growth as agreed to by the Navy for authorized shipbuilding in years 1971-1975, (excluding possible future contractor recovery on outstanding claims) rose to almost $2 billion - a cost growth, for 1975 approved programs, of around 50 per cent [Ref. 2]. What the Brooks' committee was investigating was growth within a program -- from the time of its initial planning estimate to actual delivery of production systems. There is another aspect of cost growth, namely the trend of increasing unit costs from one generation of systems to the next. As an example, the Secretary of the Air Force compared the respective costs of a World War II fighter, the P-38 at $ 200,000, and the new F-15, at $ 20 million. No single factor can be identified as the cause for the increased cost of military systems.

Investigations of cost growth have identified inflation, technical changes, quantity decreases, overoptimism and "buy-ins," and reduced DOD budgets as major causes of cost growth and resultant increased unit cost of systems. Production management has been criticized for its lack of adequate control of contractors. Yet, even in aggregate, these well-known causes do not provide a complete picture.

B. PROGRAM MANAGEMENT

There are three basic goals in the management of every defense project. The foremost is technical performance. The second reason for project management is the value of time. This relates notably to whether or not a system is available for operational readiness. It relates also to costs. Inflation obviously increases costs according to project duration. Less obvious is the fact that stretching out the development and production of a system usually reduces its potential operational life. The third basic objective for project management is lower system acquisition costs.

A fair measure of effectiveness of the management of defense programs, however, is a comparison of each program's status versus the plan for that particular project. Past and present efforts to improve project management show great

improvement in meeting cost goals, particularly, "target cost," rather than something between target and ceiling. "Target cost" is the goal sought on each major program. In a little over two decades, many project management techniques have been tried - some have been rejected (like PERT-cost), and others have been refined. Project management is now a rapidly maturing field of management.

1. Historical Perspective

The history of cost growth for individual defense (weapons) programs is abundant. Dramatic overruns on the F-111, the C-5A, and the SRAM missile made headlines in the late 1960's just at the time when public concern over the political, social and macroeconomic impacts of the military-industrial complex had begun to surface [Ref. 3].

Such results in defense systems acquisitions have brought much criticism to the Department of Defense for the way in which DOD has, in the past, managed the acquisition of major defense systems. Costs initially were not a major reason for adopting project management, as evidenced by early widespread use of cost-type contracts in which all or most of the cost risk was borne by the government. But the importance of costs has risen greatly in recent years to the point that costs are now officially equated to technical

performance in importance. This importance has led to many design-to-cost-type contracts. It is important to recognize the nature of the management problems involved in bringing into our operational inventory a major weapons system for our national defense at a target cost.

2. Risks

The first of several categories of risk facing program managers concerns acquisition time for a weapon system, the length of which can span five to seven years, or sometimes longer. Reaching as we are so far into the future, DOD managers must effectively deal with the risks of making cost projections over this time span, and ensure accountability for such projections.

A second category of risk is the fact that a major weapon system involves nearly every field of technology. DOD managers must resolve the nature and amount of forecasted growth in the technologies that are included in a major system that will be operational for far into the future.

The ever-changing levels of capability of our adversaries present another form of risk that must be effectively dealt with. The weapon system must also be designed to meet a forecasted threat derived from such

variables as support cf our increasing international commitments, obsolescence of our current weapons systems, changing enemy objectives, greater weapons systems effectiveness, or a combination of any or all of these. Our management process must be responsive to the need of effectively resolving the risks presented by this facet of this dynamic environment.

Another category of risk that management must face is that, in developing and producing a weapon system that has not existed before, provision must be made for the proper identification and timely resolution of the many uncertainties that experience indicates in such an effort. Our management process must be able to prevent or minimize degradation in cost, schedule, or performance of the weapon system as these uncertainties are resolved.

3. DSARC

As a means of providing management overview for timely decision making, the Secretary of Defense has established the Defense Systems Acquisition Review Council (DSARC). The mission of the council is to review major weapon system acquistion programs at appropriate milestone points in their life cycle. These reviews are to permit coordinated evaluation and deliberation among senior

managers, to assure that the advice given the Secretary of Defense is as complete and objective as possible prior to a decision to proceed to the next step of the system's life cycle. The three basic milestone points are:

1. When initiation of contract definition is proposed,
2. When transition from contract definition to full-scale development is proposed,
3. When transition from the development phase into production for service deployment is proposed.

Thus, it can be seen that before a major system can progress through its life cycle, senior DOD managers must determine that satisfactory progress has been made and is expected to continue, in accordance with the original and updated plans for accomplishing the acquisition of the system.

4. Cost/Schedule Control System Criteria

The Cost/Schedule Control System Criteria (C/SCSC) are a set of criteria or standards that a contractor's management system, whatever it may be, must meet in undertaking development of a major defense program. Essentially what C/SCSC does is ensure that data provided by a contractor, such as his monthly cost performance report, is accurate and timely. Cost and schedule deviations can

then be traced to their sources and action taken. There are some thirty-five of these criteria in the C/SCSC Joint Implementation Guide [Ref. 4]. They define the standards a contractor's management system must meet regarding organization, planning and budgeting, accounting, analysis, revision and access to data.

C/SCSC introduced the concept of earned value which enables program managers in industry and the government to determine, with considerable accuracy, the cost status of their programs and supplements. They do this by comparing budgeted cost of work performed (BCWP) and actual cost of work performed (ACWP). By comparing BCWP with the budgeted cost of work scheduled (BCWS), they can measure actual and planned progress in terms of the cost required to come from a behind-schedule position to on schedule. The Acquisition Management Information Division of the Office of the Assistant Secretary of Defense (Comptroller) analyzes BCWP, ACWP, and BCWP data to monitor the performance and progress of major defense acquisition programs [Ref. 5].

Several benefits result from C/SCSC. First, it assists the program manager in keeping within his target costs. Other benefits are identification of problems not previously recognized, the ability to trace problems to

their source, and objective rather than subjective assessment of program status.

C. COST GROWTH IN DEFENSE CONTRACTS

Control of costs and schedules is a dominant concern of defense systems acquisition management in both government and industry. The Comptroller General's Reports to Congress since 1975 show that each year major defense systems cost more than their base-line estimates.

While the costs increase, the schedules slip, and vice versa. The Defense Science Board reports that the time from Milestone 0 to Milestone 1 has lengthened significantly from less than 2 years pricr to 1950 to nearly 5 years in 1974 [Ref. 6].

These sizable cost increases and schedule slippages cannot be attributed to any scarcity of data. Results of an extensive survey conducted by the National Security Industrial Association (NSIA) to develop a Cost/Schedule Systems Compendium indicate that about 1,056,000 pages of cost account documentation are created each month by DOD contractors in order to satisfy government requirements, averaging 2,672 pages for each C/SCSC application [Ref. 7].

Despite the size of such reports, Defense Secretary Caspar Weinberger and Budget Director David Stockman have

found it necessary to assure Congress that rigorous program management will be pursued. In a joint letter to Senate Budget Committee Chairman Pete Domenici, the two cabinet officers wrote: "Furthermore, we want to assure you that as part of our overall effort to constrain federal spending growth and eliminate the deficit by FY 1984, we plan to impose rigorous program management responsibilities on all agencies, including the Department of Defense, to assure that outlays do not exceed estimates." [Ref. 7].

D. NATURE OF THE PROBLEM

It is apparent that several successful mechanisms are currently utilized to objectively monitor contractor performance on major defense systems at all levels in DOD. Yet, control of program cost growth continues to elude the majority of those tasked with that control, the program managers. There exists the possibility that, while C/SCSC has standardized the contractor submission of and the government collection of pertinent contract cost and schedule data, the lack of standardized utilization of this data by program cost managers may be a major reason that costs are difficult to control. Improved standardized cost monitoring methods could be derived for implementation such as done in this research.

Although C/SCSC and Contractor Performance Reports (CPR) help us in management of our major contracts, the potential for overruns on the non-major contracts is always present, and such overruns in the aggregate can easily equate to the growth of a major contract or two. Proper utilization of the tool for non-major contracts, the Cost/Schedule Status Report (C/SSR) , a tailored-down CPR, is necessary for successful control of these contracts.

Although this paper specifically addresses the utilization of C/SSR data in contract cost estimation, the use of other contractor-generated cost and schedule data also needs to be studied in greater detail.

II. METHODOLOGY

A. GENERAL

Past and present program management efforts to control defense systems contract costs have certainly not been totally successful. There are instances of actual costs of several times the initially budgeted cost, and a time to initial operational capability sometimes several years longer than planned. A better method of forecasting costs is needed so that those costs can be better controlled. An accurate forecast of tomorrow's actuals (such as ACWP) compared with corresponding plans (such as BCWP) can provide the expected cost excess (BCWP-ACWP). Current values of ACWP and BCWP are standard data elements in the contractor-supplied cost performance reports.

To begin to gain a clear understanding of contract cost forecasting, the researcher must be able to understand trends and relationships as they exist in the data. Second, he must have some means of testing hypotheses to predict future outcomes. The visualization can be accomplished through the use of graphs, equations, histograms, etc., or, as developed in this study, through the plotting of contract

cost data. Of the several methods available for quantitatively determining relationships among data, a basic tool for studying relations among variables is regression. Regression may provide the necessary coefficients to predict future cost and schedule outcomes.

The methodology proposed here provides a basic framework within which contract costs and schedule estimates can be systematically reviewed, studied, updated and forecasted. The methodology is as straightforward as possible in order to insure the widest possible application. Behavior of weapon system acquisition contracts can easily be assimilated within the framework provided. The methodology proposed by this researcher utilizes the application of the Rayleigh-Norden curve to contract cost data. Other methodological support techniques used by this researcher are scatter plots and linear regressions.

B. RAYLEIGH-NORDEN CURVE THEORY

Studies have indicated that there are regular patterns of manpower buildup and phase-out in complex projects [Ref. 8]. This life-cycle pattern happens to follow the distribution formulated by Lord Rayleigh to describe other phenomena [Ref. 9]. Norden used the model to describe the quantitative behavior of the various cycles of research and

development projects each of which had a homogeneous character. Accordingly, it is appropriate to call the model the Rayleigh-Norden Model.

These patterns have been described by a number of mathematical functions, generally in the family of exponential, gamma, beta, or logistic curves, by several researchers. In the life-cycle model, curves are fitted to a small number of successive "cycles" of work which occur during the life of a project. The cycles do not depend on the nature of nor the content of the project, but appear to be a function of the way groups of engineers and scientists tackle complex technological development problems. Each cycle (Fig. 1) can be described by a comparatively simple first-order differential equation, (as symbolized by the • above the y):

$$\dot{y} = 2Kate^{-at^2} ,$$

where

1. $\dot{y}$ = manpower utilized during each time period,
2. K = total cumulative manpower utilized by the end of the project,
3. a = shape parameter (governing time to peak manpower),
4. t = elapsed time from start of cycle.

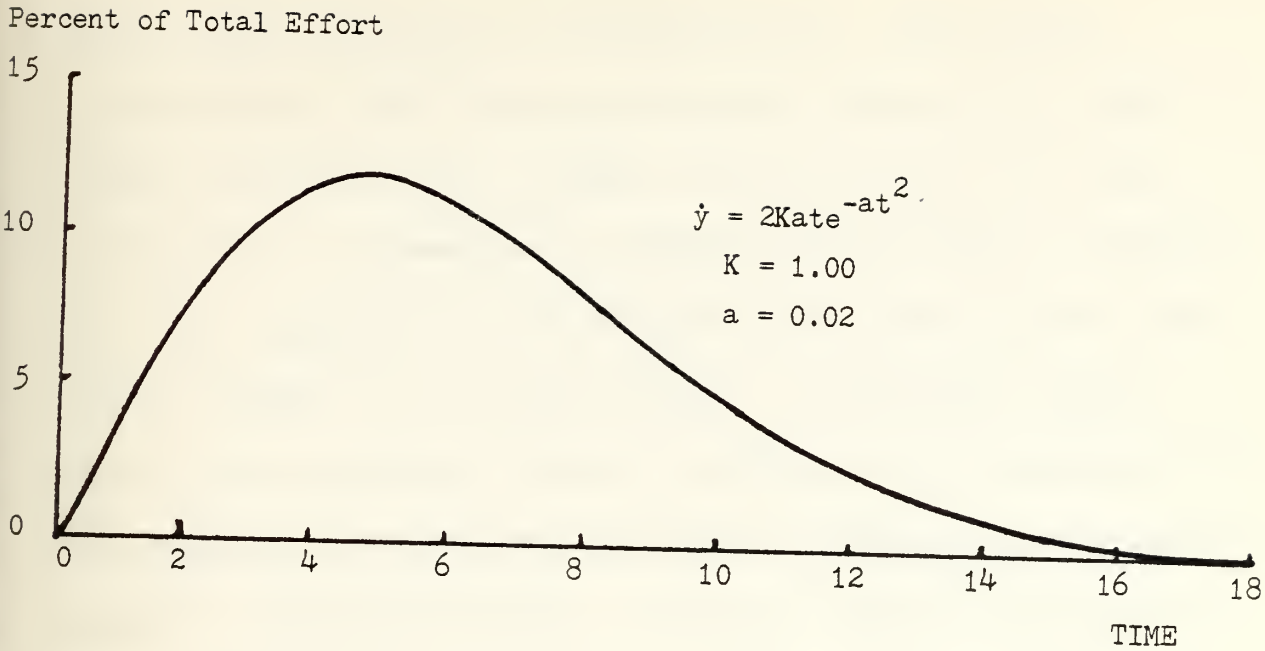

Figure 1: Current manpower utilization curve

Thus the cycles can be represented by a series of curves from the same family relating manpower used each month to elapsed calendar time, and differing only in relative size and proportions. The single parameter governing the shape of the curves can be thought of as a measure of the importance of the project (Fig. 2). Sharply peaked manpower buildups correspond to rush projects, while shallower curves are associated with stretched-out projects.

The mathematical model of project manpower consists of the equation for each cycle plus a linking function which specifies the relative sizes and durations of the cycles and

their lags or spacing in calendar time. The linking relationships have been encouragingly stable over a wide range of projects and a number of years [Ref. 8]. These results make it possible to develop projections of manpower and time requirements for comparable projects, given a few actual points on the early cycles. In addition, early warning of significant departures from current schedules can be obtained by monitoring the actual progress of a project against a prior projection. Since an overrun in an early cycle cascades through all the subsequent cycles, this 'early warning' potential of life-cycle analysis is applicable, particularly when making future projections. Further details of this model and its adaptation and application to defense contract cost estimation is presented in Chapter V.

C. SCATTER PLOT

The scatter diagram provides a quick and flexible means for displaying variables. Because the researcher in cost estimation is often interested in what happens to selected variables over time, the time series display of a scatter plot is invaluable. The researcher can study a scatter diagram unhampered by the clutter which accompanies most plots. Because trends are normally the point of interest,

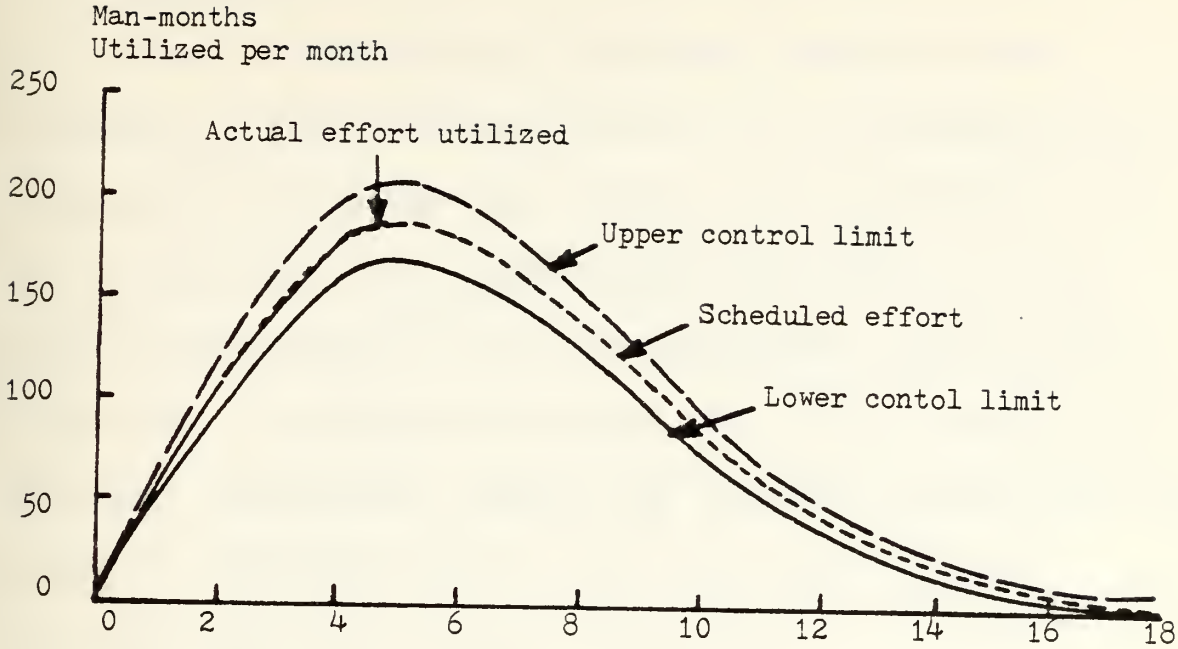

Figure 2: Life-cycle method of project control
(Early warning technique)

the simplified picture provides what the researcher needs to observe trends. Additionally, various relationships among data such as ratios, percentages, and transformations must be computed. In a computationally interactive mode, when a researcher happens upon a significant relationship, he must not be hampered by excessive formatting or computational requirements in the attempt to establish that relationship. The basic simplicity of scatter plotting and the lack of an individual format are proposed in this thesis as the appropriate choice for interactive display of data.

D. REGRESSION

The use of regression in studying relations among variables is widely accepted [Ref. 10]. The application of regression is proposed here because the observations, i.e., the contract cost data (ACWP,target cost,BCWP) and time, occur simultaneously in the nonexperimental, uncontrolled situations in contractor's plants. In order to explore the empirical relationship between the Rayleigh-Norden model and selected historical contract cost data, the specific technique required is linear regression.

Let Y be a given dependent variable and X a vector of K independent variables which combine linearly to yield Y. For n sample observations the model is:

$$Y = B_0 + B_1 X_1 + B_2 X_2 + \ldots + B_k X_k + u$$

The error term, u, is due to two factors. First, the selected model is a simplification and many variables expected to have minor impact may be omitted. Second, even if a selected theoretical relationship is exact, errors of measurement in real observations will produce errors in the observed relationship. The supposition of linearity of contract cost and the Rayleigh-Norden model is based on a transformation of both the data and the model and is

presented in Chapter V. The basic assumptions of the model are:

1. $E(u) = 0$,
2. $E(u^2) = \sigma^2$,
3. The X's are either nonrandom or, if random, are independent of u.

The above assumptions hold for $i = 1,2,..,n$, where n is the number of observations.

E. SUMMARY

An introductory explanation of Rayleigh curve theory has been presented in this chapter. The methodology proposed by this researcher utilizes the Rayleigh-Norden model and selected historical cost data. Further discussion describes the importance of scatter plots and linear regression techniques in determining significant relationships between the data and the model.

A description of the data base used is presented in Chapter III. Application of the Rayleigh-Norden model to the data base using an interactive computing system is presented in Chapter V.

III. STRUCTURE OF DATA BASE

Schedule and cost data for thirty DOD contracts in the form of a summary of quarterly C/SSR's were obtained from a DOD agency. The cost items were expressed in current dollars as of the report dates. Contracts with start dates prior to January 1977, the implementation date of C/SSR, could not be used since the initial report period duration exceeded three months and thus, created a major anomaly. A few contracts researched were missing quarterly reports and linear interpolation was used by this researcher to provide the missing data points.

An individual data file or worksheet, per contract, containing the cost and schedule items of interest, was constructed in the Minitab environment. Table 1 demonstrates the column format of Minitab for one contract. The length of the columns of data was determined by the number of reports for each contract. These files served as the basis for all subsequent operations on and manipulations of the contract data.

A general knowledge of the statistical computing system Minitab is assumed in the following comments. Readers are referred to Ryan, Joiner and Ryan [Ref. 11] for specific

format and worksheet management. Chapter IV contains a brief introduction to the format of, and details for, the specific utilization of Minitab in this analysis.

Any selected data file could be directly accessed by entering the command RETRIEVE 'DAT"I"', where I represented the contract number as coded by the researcher. The possible values of I were the integers 1 through 30.

Due to the comparative nature of this study it was necessary to investigate the impact of inflation in a robust manner. A column named 'DEFLAT' was constructed which had as each of its components an index applicable for the corresponding quarter. The selected indices shown in Table 2 were obtained from the Bureau of Economic Analysis of the United States Department of Commerce. They represent seasonally adjusted implicit price deflators (base year 1972 = 100) for DOD purchases of goods and services by category of contract, i.e., aircraft, missiles,and ship construction. See Chapter I for a discussion of the role inflation plays in contract cost escalation.

Table 3 provides a list of the columns which were constructed in the worksheet. Unless otherwise indicated, the number of rows in each column is determined by the number of quarterly reports submitted from the start date of

the contract through and including cycle year 1980. For example, row 1 pertains to cost and schedule data for the first quarter after the commencement of work authorized by the contract.

The data base, as constructed by this researcher, requires a one-time unformatted entry of contract cost and schedule data into computer files. Chapter IV describes the interactive capability of Minitab for analysis of these files.

Table 1: Minitab Column Format

```
minitab

MINITAB RELEASE 81.1 *** COPYRIGHT - PENN STATE UNIV. 1981
FEB. 14, 1982 *** NAVAL POSTGRADUATE SCHOOL (31 AUG 81) VER. 3
STORAGE AVAILABLE  40000
--
retrieve 'dat8'
--
print c1-c12
```

COLUMN ROW	PERIOD	RBCWS	RBCWP	ACWP	MGTRES	TARGET
1	1.	0.0	2.000	2.000	6.70000	117.200
2	2.	12.700	9.700	8.800	0.0	107.200
3	3.	19.100	15.600	17.000	4.70000	110.300
4	4.	33.600	30.700	30.900	3.90000	110.900
5	5.	54.100	44.400	45.500	3.80000	114.600
6	6.	63.600	60.000	61.600	3.60000	114.900
7	7.	80.400	73.700	74.800	3.20000	117.200
8	8.	101.700	95.500	94.200	2.70000	120.900
9	9.	110.000	102.600	99.600	0.90000	121.600
10	10.	113.700	109.300	107.000	0.30000	128.500
11	11.	117.000	115.500	114.200	0.30000	129.300
12	12.	120.100	119.800	119.100	0.10000	132.800
13	13.	123.800	123.700	123.200	0.30000	132.800

COLUMN ROW	ACW/BCWP	A	INACWP	IMPDEF	DEFLAT	IDACWP
1	1.00000	117.200	2.0000	130.600	0.765697	1.5314
2	0.90722	97.254	6.8000	131.100	0.762777	5.1869
3	1.08974	120.198	8.2000	132.500	0.754717	6.1887
4	1.00651	111.622	13.9000	135.900	0.735835	10.2281
5	1.02477	117.439	14.6000	138.200	0.723589	10.5644
6	1.02667	117.964	16.1000	142.000	0.704225	11.3380
7	1.01493	118.949	13.2000	142.700	0.700771	9.2502
8	0.98639	119.254	19.4000	147.400	0.678426	13.1615
9	0.97076	118.044	5.4000	150.200	0.665779	3.5952
10	0.97896	125.796	7.4000	151.700	0.659196	4.8780
11	0.98874	127.845	7.2000	152.200	0.657030	4.7306
12	0.99416	132.024	4.9000	160.400	0.623441	3.0549
13	0.99596	132.263	4.1000	162.300	0.616143	2.5262

Table 2: IMPLICIT PRICE DEFLATORS (CY 1972 = 100)

YEAR	QUARTER	AIRCRAFT	MISSILES	SHIPS
1976	I	116.6	112.5	146.0
	II	119.5	113.9	147.1
	III	116.8	115.0	149.2
	IV	117.6	112.6	155.0
1977	I	126.3	115.6	160.2
	II	131.6	122.3	165.0
	III	135.3	119.1	165.9
	IV	140.0	124.6	170.4
1978	I	144.0	136.6	176.9
	II	144.8	133.2	178.9
	III	148.2	135.1	185.3
	IV	150.4	140.8	185.2
1979	I	158.0	133.5	185.1
	II	160.3	141.6	187.3
	III	169.6	147.4	198.3
	IV	166.2	149.8	200.4
1980	I	174.6	156.1	200.6
	II	177.4	153.9	209.4
	III	186.4	156.7	216.7
	IV	184.1	162.3	222.6

Table 3: MINITAB DATA COLUMNS

Column Name	Contents
A	Product of ratio of ACWP to BCWP and target cost ((ACWP/BCWP) X TARGET)
ACWP	Actual cost of work performed (cumulative)
ACW/BCWP	Ratio of ACWP to BCWP
DEFLAT	Deflation index, deflates reported contract cost data to 1972 constant dollars. DEFLAT = 100 / IMPDEF
DTARGET	Deflated target cost of contract
IDACWP	Incremental (quarterly) deflated actual cost of work performed
IMPDEF	Implicit price deflator (base year 1972)
K	Target cost of contract, predicted by the model
LN(Y/T)	Natural logarithm of (Y/T), the predictor variable in regression
INACWP	Incremental (quarterly) actual cost of work performed
MGTRES	Management reserve
PERIOD	Report period after commencement of contract authorized work
RBCWP	Reported budgeted cost of work performed (cumulative)
RBCWS	Reported budgeted cost of work scheduled

	(cumulative)
RSQUARE	Measure of the fit of the model to the data, coefficient of determination
TARGET	Target cost of contract
Td	Quarter in which maximum INACWP occurs, predicted by the model
TSQUARE	Period squared, the explanatory variable in the regression
Y/T	Ratio of INACWP to Period
YHAT	Incremental (quarterly) actual cost of work performed, predicted by the model

IV. A FORMAT FOR MINITAB APPLICATION

A. MINITAB

The purpose of this thesis is to demonstrate the application of a software cost estimation methodology to DOD weapon systems contract cost data; however, the selection of a computer based statistical analysis system for use by this researcher deserves some explanation.

Minitab is a general purpose statistical computing system, designed particularly for students and researchers who have no previous experience with computers. While it is fairly powerful, very flexible and easy to use, it is especially useful for exploring data in the early phases of analysis, for plotting and for regression analysis. The ability to manipulate columns in a single program step vice the looping and multiple steps required in a language such as FORTRAN provides the programmer/user tremendous power.

B. ANALYTICAL COMMANDS

A relatively small number of Minitab commands is sufficient to analyze DOD contract costs in the context of the methodology presented in this paper. Two principal commands, REGRESS and PLOT (MPLOT), in conjunction with two

supporting commands JOIN and PICK, provide sufficient computational capability while also providing a means to graphically present data. What follows is a user-oriented synopsis of each of these commands. The analytical commands are the principal tools used in the study and graphical presentation of contract cost and schedule data.

1. Regress

Syntax: REGRESS y in C using K predictors in C,...,C

Parameters:

1. K - The number of predictor variables
2. C - the column numbers of the y variables and the K predictor variables

Description : In its basic capability to do linear regression analysis, REGRESS relates the dependent variable, Y, to the independent variables, X's. Basic output consists of the regression equation, standard deviation of the regression coefficients, r-square, r-square adjusted, ANOVA table and Durbin-Watson statistic. Additionally, if specified in the REGRESS command, residuals, predictors (fitted Y values), and the regression coefficients may be stored in user defined columns for use in analysis. This optional storage allows for simplified plotting of the usual

plots of residuals versus the predicted values, residuals versus each X variable and residuals versus time order. An example of the results of REGRESS is presented in Figure 3, while the definitions of the REGRESS output statistics appear in Table 4.

2. Plot

Syntax : PLOT y in C vs x in C

Parameters :

1. WIDTH - Controls the horizontal size of the display (30-100 spaces). (Default = 50 spaces)

2. HEIGHT - Controls the vertical size of the display (15-400 lines), (Default = approximately one-half the width)

3. WIDTH and HEIGHT - Controls both the horizontal and vertical size of the display with single command. (Default = 50 spaces wide, 25 lines high)

Description : PLOT produces a two dimensional scatter plot of the data contained in the specified columns. The pairs (x,y) are plotted with the symbol '*' unless two or more points fall on the same spot. In this case a count of points falling on the same spot is given. A '+' is given if the count is over nine (9). Figure 4 displays a sample output of PLOT. Divisions on the axes of the plot are explicitly printed.

```
THE REGRESSION EQUATION IS
Y =    2.32 -0.0503 X1

                                        ST. DEV.     T-RATIO =
      COLUMN      COEFFICIENT           OF COEF.     COEF/S.D.
      --              2.3162             0.2428          9.54
X1    TSQUARE       -0.05035            0.01735        -2.90

THE ST. DEV. OF Y ABOUT REGRESSION LINE IS
S = 0.3355
WITH (   5- 2) =   3 DEGREES OF FREEDOM

R-SQUARED = 73.7 PERCENT
R-SQUARED = 65.0 PERCENT, ADJUSTED FOR D.F.

ANALYSIS OF VARIANCE

 DUE TO      DF            SS        MS=SS/DF
REGRESSION    1         0.9480         0.9480
RESIDUAL      3         0.3377         0.1126
TOTAL         4         1.2857


DURBIN-WATSON STATISTIC = 1.96
```

Figure 3: Sample Results of 'REGRESS' Command

Table 4: Definitions of Terms in 'REGRESS' Results

TERM	DEFINITION
Analysis of variance	The analysis of variance table gives the following sums of squares (SS). Let n = number of observations. The TOTAL SS = SUM((Y-Ybar)-squared), and has n-1 degrees of freedom. The SS DUE TO REGRESSION is SUM((Yhat-Ybar)-squared) where Yhat is the predicted value of Y. The SS RESIDUAL is SUM((Y-Yhat)-squared). The last column gives the mean squares which are useful for various F-tests.
Degrees of freedom	The number of degrees of freedom for the sum of squared residuals is the number of observations minus the number of coefficients in the regression equation. This degrees of freedom is used in t-tests and confidence intervals and in F-tests. Each regression coefficient has 1 degree of freedom associated with it.
Durbin-Watson statistic	Used to test for autocorrelation in the data.
F-tests	F-tests for the significance of regression, etc, are easily carried out using the MS values in the analysis of variance tables.
Mean square (MS)	Sum of squares divided by its degrees of freedom.
Predicted y- value	The value produced by substituting

	the x-value(s) into the fitted regression equation.
Regression equation	The equation found by the REGRESS command which fits the data best, according to the least squares criterion.
Residual	The difference between the observed and predicted y value, i.e., y - (b0 + b1 X1 + ...). This difference is also called the error or deviation.
R-squared	A measure of how well the regression equation fits the data, with 100% indicating a perfect fit. Defined by 100(SS due to regression)/(SS total).
Standard deviation of coefficient	Since the coefficients are determined by the data, they are random variables. The estimate of their standard deviation is printed in the table of coefficients.
Standard deviation of predicted y values	Since the fitted regression equation is determined by the data, the predicted values are random variables. The standard deviation of these is printed out as an indication of uncertainty and for use in forming prediction and confidence intervals.
Standard deviation of y about regression line	This is an estimate of sigma. It is defined by s = sqrt(MS(error)). The degrees of freedom associated with s is the same as for SS(error), n-p, where p is the number of coefficients

	in the equation.
Tests	Tests of significance of the coefficients are based on the t-ratio in the table of coefficients. F-tests for the significance of regression, etc., can be done easily by using MS values in the analysis of variance tables.
T-ratio	The ratio coefficient/(est. st. dev. of coef.) is used as a test statistic for testing the hypothesis that the true (population) coefficient is 0.

MPLOT produces multiple plots on the same axes. The first pair of columns are plotted with the symbol 'A' , the second pair with 'B' and so on. If several points fall on the same spot, a count is given as in PLOT. An example of the results of MPLOT is displayed in Figure 5.

C. SUPPORTING COMMANDS

The supporting commands provide greater flexibility for data manipulation and format for interactive output.

1. Join

Syntax: JOIN E to the bottom of E (to the bottom of E,...,to E) put into C. (E denotes either a stored constant, a number, or a stored column)

Description : The command JOIN is used to create a new data

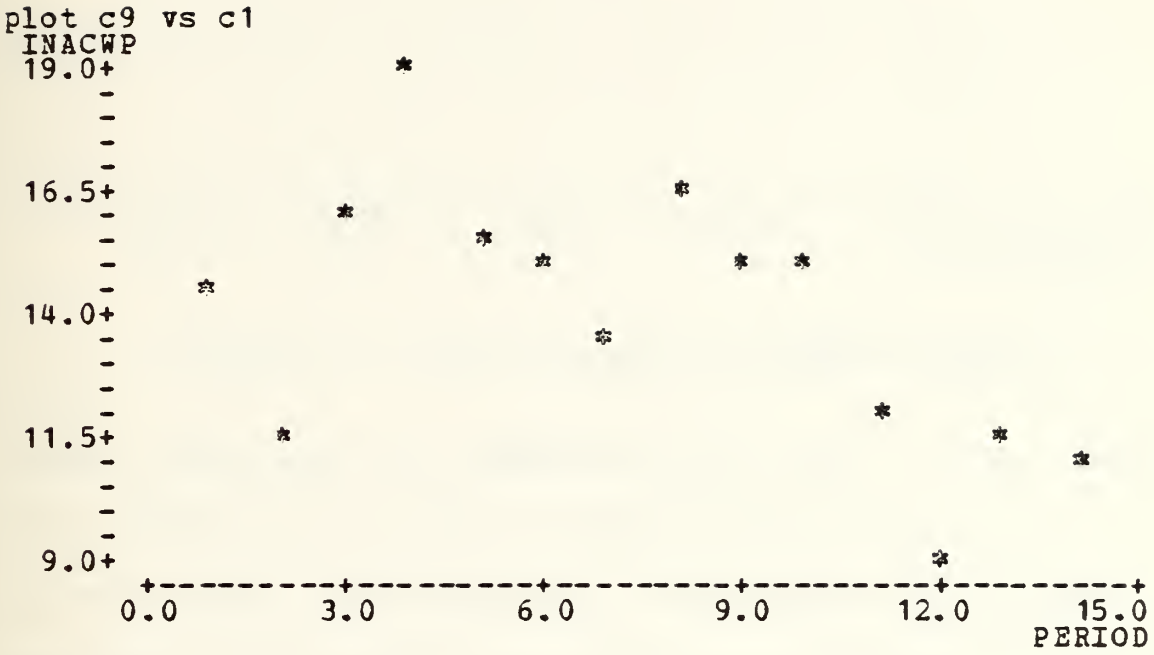

Figure 4: Sample Results of 'PLOT' Command

```
mplot c17 vs c16 and c19 vs c16
 Ln(y/t)
 2.70+     A
     -
     -                          A = ACTUAL
     -                          B = PREDICTED
     -
 2.40+
     -     B
     -
     -
     -           B
 2.10+
     -
     -
     -
 1.80+                     B
     -           A
     -                     A
     -
     -                                         A
 1.50+
     -                                         B
     -
     -
     -
 1.20+
       +---------+---------+---------+---------+---------+
       0.0       4.0       8.0       12.0      16.0      20.0
                                                   TSQUARE
```

Figure 5: Sample Results of 'MPLOT' Command

column consisting of E augmented by E1, E2, .. , EN. Since E is defined to be either a number(scalar) or a column(vector), conformability in the usual sense is not necessary. The command JOIN provides an ability to readily and handily increase the number of data points in a column (Fig. 6) to be regressed by REGRESS and/or plotted by PLOT or MPLOT.

2. <u>Pick</u>

Syntax : PICK rows K thru K of C, put into C

Description : The command PICK is used to create a new data

column, of equal or shorter length, from an existing data column. It provides flexibility to the interactive user in regressing and/or plotting partial columns of data without the hardship of re-entering the interesting data by conventional methods.

D. SUMMARY

This chapter presented detailed descriptions and results of Minitab commands extensively used in this research. REGRESS describes the relationship between modelled independent and dependent variable data columns. PLOT and MPLOT produce desired scatter plots of data columns. PICK and JOIN are useful for the necessary manipulation of the data.

```
print c23 c24
COLUMN YINTCEPT        SLOPE
COUNT        10           10
ROW
 1      1.00000      9.00000
 2      2.24182     -0.04105
 3      2.16964     -0.03435
 4      2.05769     -0.02630
 5      1.97806     -0.02174
 6      1.90445     -0.01830
 7      1.85665     -0.01644
 8      1.83264     -0.01566
 9      1.76162     -0.01366
10      1.69895     -0.01214
--
join c23 2.3162,c23
--
join c24 -0.05035,c24
--
print c23 c24
COLUMN YINTCEPT        SLOPE
COUNT        11           11
ROW
 1      2.31620     -0.05035
 2      1.00000      9.00000
 3      2.24182     -0.04105
 4      2.16964     -0.03435
 5      2.05769     -0.02630
 6      1.97806     -0.02174
 7      1.90445     -0.01830
 8      1.85665     -0.01644
 9      1.83264     -0.01566
10      1.76162     -0.01366
11      1.69895     -0.01214
```

Figure 6: RESULTS OF 'JOIN' COMMAND

V. DATA ANALYSIS

Illustrated in this chapter is the adaptation and application of the proposed methodology described in Chapter II to actual cost data of various defense procurement contracts.

In the initial stages of investigation, the researcher extensively applied the Minitab command PLOT to the raw contract data, specifically INACWP versus PERIOD, in an effort to ascertain whether or not these variables of interest visually displayed characteristics of the Rayleigh-Norden curve. Visual inspection of such plots (sample in Figure 7) and comparison of these with Norden's curve as displayed in Chapter II, Figure 1, revealed graphic similarity, raising the possibility that the Rayleigh-Norden model could be applicable to this problem area.

Only three of the thirty contracts comprising the data base were suitable for investigation by the researcher for application to the proposed methodology due to some apparent irregularities in cost reporting, as cited in Chapter III. However, this sample should provide an insight into the ease with which relationships can be ascertained and examined using a simple interactive methodology.

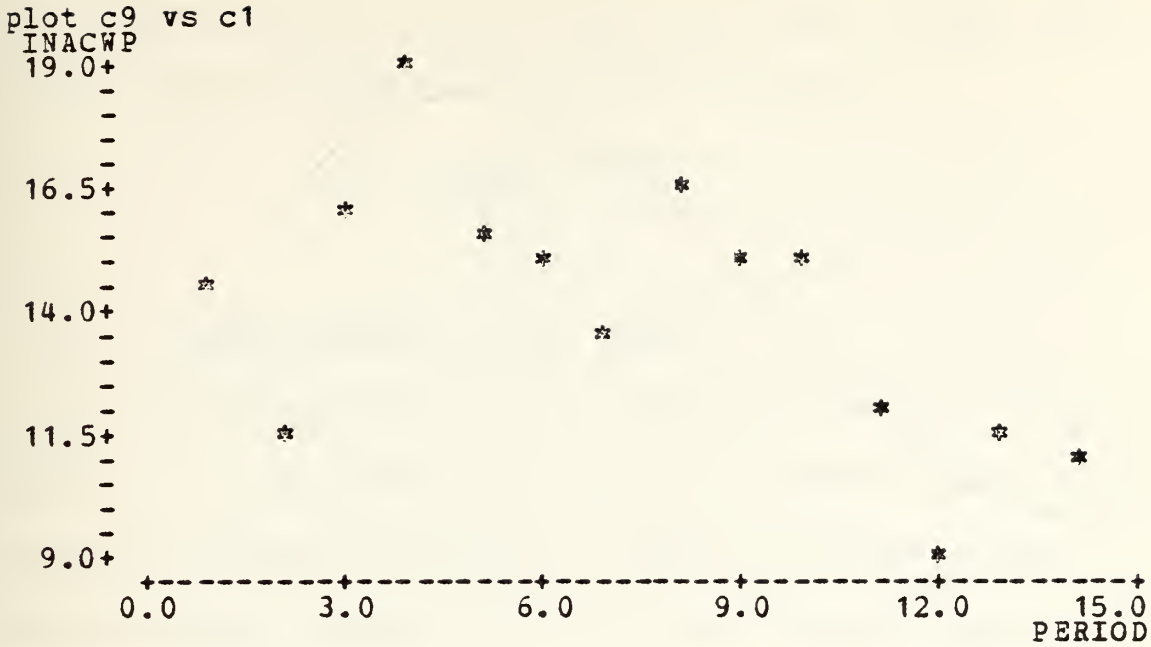

Figure 7: INACWP VERSUS PERIOD

A. THE ADAPTED MODEL

The Rayleigh-Norden equation is adapted for use by substituting contract cost data and terminology for the variables Norden utilized in his software development research efforts. The resultant equation is:

$$\dot{y} = 2Kate^{-at^2} \text{, where}$$

1. $\dot{y}$ = ACWP during each report period (INACWP),
2. K = total cumulative contract cost (target-cost) by the end of the project,
3. a = shape parameter (governing time to peak ACWP), and
4. t = elapsed time from start of contract.

By substituting $t_d^2 = 1/2a$, where t_d is time of peak ACWP, yields the following form of above equation:

$$\dot{y} = K/t_d^2 t e^{-t^2/2t_d^2}.$$

B. TRANSFORMATION OF THE MODEL

The simplest way to follow and track the time varying behavior of a contract is to plot the INACWP at any instant in time, as shown in Figure 7. Using this data stream transforms the problem into on of time series analysis. The time series problem is most easily solved by turning the characteristic Rayleigh behavior into a straight line. One can fit the actual contract cost data to get a revised estimate of future resource consumption: target cost, contract duration, future ACWP, and their tolerances. The results of transforming the Rayleigh-Norden curve into a linear form is illustrated in Figure 8. The equations underlying this natural logarithmic transformation are summarized as follows:

$$\dot{y} = (K/t_d^2)\, t e^{-t/2t_d^2},$$

Dividing by t yields

$$\dot{y}/t = (K/t_d^2)\, e^{-t^2/2t_d^2}$$

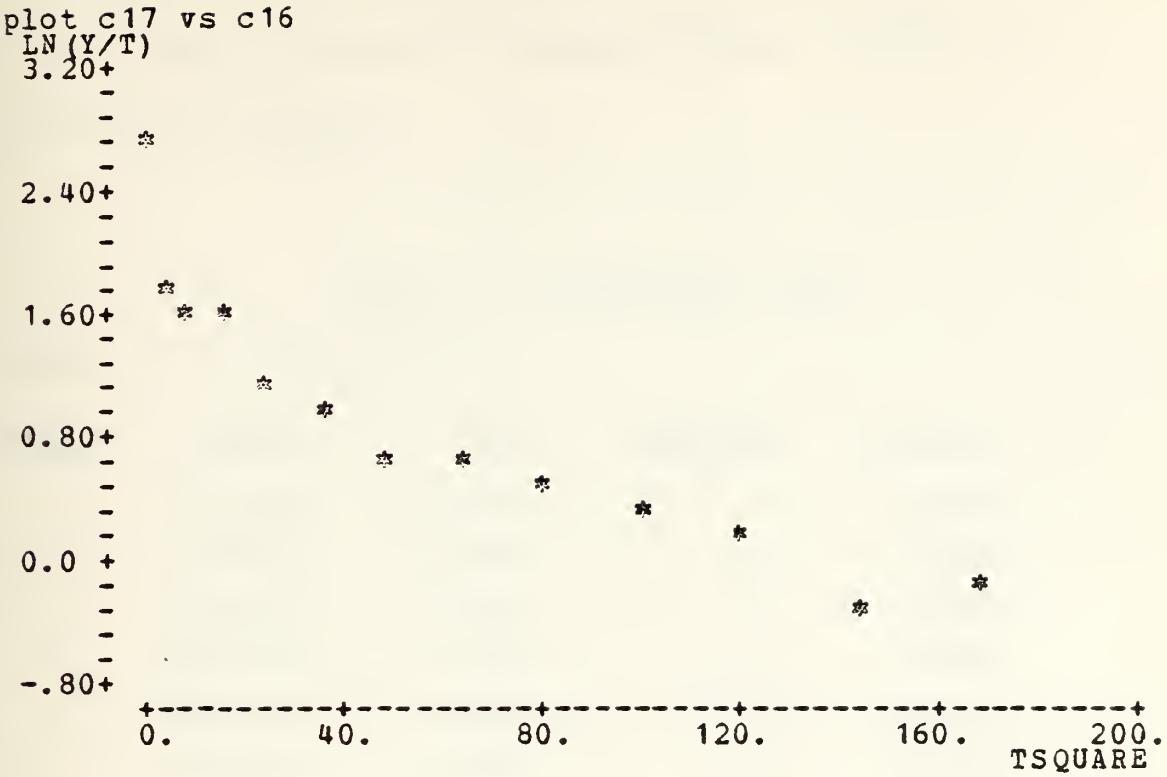

Figure 8: LINEAR FORM OF RAYLEIGH-NORDEN CURVE
(Ln(Y/T) versus TIME Squared)

and taking natural logarithims yields

$$Ln(\dot{y}/t) = Ln(K/t_d^2) + (-1/2t_d^2)t^2 ,$$

which is the familiar linear form

$$Y = A + BX,$$

but expressed as a function of t^2.

The ease with which column arithmetic can be performed in Minitab was exploited in transforming the contract data for later use in regression. The resultant transformed data

representing a helicopter weapons system contract is displayed in Table 5.

Table 5: TRANSFORMED DATA

print c1 c9 c14 c16 c17

PERIOD	INACWP	Y/T	TSQUARE	LN(Y/T)
1	14.6000	14.6000	1.	2.68102
2	11.7000	5.8500	4.	1.76644
3	16.0000	5.3333	9.	1.67398
4	18.9000	4.7250	16.	1.55287
5	15.6000	3.1200	25.	1.13783
6	14.8000	2.4667	36.	0.90287
7	13.6000	1.9429	49.	0.66416
8	16.4000	2.0500	64.	0.71784
9	15.0000	1.6667	81.	0.51083
10	14.8000	1.4800	100.	0.39204
11	12.2000	1.1091	121.	0.10354
12	9.0000	0.7500	144.	-0.28768
13	11.7000	0.9000	169.	-0.10536
14	10.9000	0.7786	196	-0.25029

C. CONTRACT CONTROL AND ESTIMATION

One can learn the characteristics of the adapted Rayleigh-Norden curve for a particular contract by analyzing early INACWP cost data. Then subsequent data can be used to predict a new curve, which may differ from the one originally projected. This gives a forecast of final cost

which can warn of cost growth and change in other contract management parameters.

As work proceeds during the contract, uncertainty about the management parameters decreases. At the end of the project, all the parameters are precisely known, but this is, of course, too late to be useful for estimating or control purposes. However, use of the early and subsequent forecasts can approximate the final actual target cost figure long before the end of the contract.

To determine the early characteristics of the contract, the first four reported INACWP values in transformed format were linearly analyzed. This selection of data (Table 6) is facilitated by using the PICK command described in Chapter IV, and represents one year's progression of the contract. Regressing $Ln(y/t)$ on t^2 using the command REGRESS provides information which can be used for contract control and estimation. From the regression results in Figure 9, the slope (X-coefficient) can be used to compute the period in which maximum INACWP occurs, t_d. The intercept, $Ln(K/t^2)$, given the value of t just obtained, can be used to determine the estimated value of total cumulative contract cost, K, and both the slope and intercept can be used to project next quarter's INACWP, now defined as Yhat. A

graphic representation of the regression results is displayed in Figure 10.

The algebra supporting the above discussion is as follows:

To solve for t_d begin with

$$\text{slope} = -1/2t_d^2 = B,$$

taking reciprocals yields

$$-2t_d^2 = 1/B,$$

dividing by -2 yields

$$t_d^2 = -1/2B,$$

and taking the square root gives

$$t_d = \text{SQRT}\ (-1/2B).$$

To solve for K, begin with

$$\text{Intercept} = \text{Ln}(K/t_d^2) = A,$$

taking the exponential yields

$$\exp\ (\text{Ln}(K/t_d^2) = \exp\ (A),$$

solving yields

$$K/t_d^2 = e^A,$$

and multiplying by t_d^2 gives

$$K = e^{A} t_d^2 .$$

To solve for Yhat, begin with

$$Ln(Yhat/t) = A + Bt^2,$$

taking the exponential yields

$$Yhat/t = e^{(A + Bt^2)},$$

and multiplying by t gives

$$Yhat = te^{(A + Bt^2)} .$$

The resultant important management parameter values for this data are:

1. t_d = 2.85761 quarters
2. K = $ 88.038 million
3. Yhat = $ 11.6638 million (next quarter's INAWCP)

Recall that these parameter values are based on data for the first year of the contract. A quarter later, upon receipt of the next contractor-supplied C/SSR report, the actual data points are available and are added to the data from Table 6 to yield Table 7. An additional dot is added to the graph in Figure 10 (Fig. 11) and the best straight line is again fitted. Linear regression statistics (Fig. 12) reveal that both the intercept and the slope have

Table 6: INPUT DATA FOR INITIAL ESTIMATE

PERIOD	INACWP	Y/T	TSQUARE	LN(Y/T)
1	14.6000	14.6000	1.	2.68102
2	11.7000	5.8500	4.	1.76644
3	16.0000	5.3333	9.	1.67398
4	18.9000	4.7250	16.	1.55287

changed, thus resulting in new values of the contract management parameters:

1. t_d = 3.15127 quarters
2. K = $ 100.666 million
3. Yhat = $ 9.9280 million.

```
regress y in c17 using 1 predictor in c16,(store st.
residuals in c18 (pred y in c19(coef in c20)))

THE REGRESSION EQUATION IS
Y =    2.38 -0.0612 X1

                                    ST. DEV.     T-RATIO =
      COLUMN      COEFFICIENT       OF COEF.     COEF/S.D.
      --              2.3778         0.3284          7.24
X1    C16           -0.06123        0.03491         -1.75

THE ST. DEV. OF Y ABOUT REGRESSION LINE IS
S = 0.3965
WITH (   4- 2) =   2 DEGREES OF FREEDOM

R-SQUARED = 60.6 PERCENT
R-SQUARED = 40.9 PERCENT, ADJUSTED FOR D.F.

ANALYSIS OF VARIANCE

 DUE TO      DF            SS       MS=SS/DF
REGRESSION    1        0.4837         0.4837
RESIDUAL      2        0.3144         0.1572
TOTAL         3        0.7980
```

Figure 9: REGRESSION STATISTICAL RESULTS
(ONE YEAR INTO CONTRACT)

```
mplot c17 vs c16 and c19 vs c16
 Ln(y/t)
 2.70+     A
     -                                A = ACTUAL
     -
     -                                B = PREDICTED
     -
 2.40+
     -     B
     -
     -
     -            B
 2.10+
     -
     -
     -
     -
 1.80+                        B
     -            A
     -                        A
     -
     -                                                A
 1.50+
     -                                                B
     -
     -
     -
 1.20+
       +---------+---------+---------+---------+---------+---------+
        0.0       4.0       8.0      12.0      16.0      20.0
                                                      TSQUARE
```

Figure 10: GRAPHICAL RESULTS OF REGRESSION AFTER 4 QUARTERS

Table 7: INPUT DATA AFTER 5 QUARTERS INTO CONTRACT

PERIOD	INACWP	Y/T	TSQUARE	LN(Y/T)
1	14.6000	14.6000	1.	2.68102
2	11.7000	5.8500	4.	1.76644
3	16.0000	5.3333	9.	1.67398
4	18.9000	4.7250	16.	1.55287
5	15.6000	3.1200	25.	1.13783

```
MPLOT C17 VS C16 AND C19 VS C16
 LN(Y/T)
 2.80+
     -    A
     -
     -
     -                                A = ACTUAL
 2.40+                                B = PREDICTED
     -
     -    B
     -
     -         B
 2.00+
     -
     -                   B
     -         A
     -                   A
 1.60+
     -                              2
     -
     -
     -
 1.20+
     -                                                  A
     -                                                  B
     -
     -
 0.80+
       +---------+---------+---------+---------+---------+
      0.0       6.0       12.0      18.0      24.0      30.0
                                                     TSQUARE
```

Figure 11: GRAPHICAL RESULTS OF REGRESSION AFTER 5 QUARTERS

THE REGRESSION EQUATION IS

Y = 2.32 -0.0503 X1

	COLUMN	COEFFICIENT	ST. DEV. OF COEF.	T-RATIO = COEF/S.D.
	--	2.3162	0.2428	9.54
X1	TSQUARE	-0.05035	0.01735	-2.90

THE ST. DEV. OF Y ABOUT REGRESSION LINE IS

S = 0.3355

WITH (5- 2) = 3 DEGREES OF FREEDOM

R-SQUARED = 73.7 PERCENT

R-SQUARED = 65.0 PERCENT, ADJUSTED FOR D.F.

ANALYSIS OF VARIANCE

DUE TO	DF	SS	MS=SS/DF
REGRESSION	1	0.9480	0.9480
RESIDUAL	3	0.3377	0.1126
TOTAL	4	1.2857	

DURBIN-WATSON STATISTIC = 1.96

Figure 12: REGRESSION RESULTS AFTER 5 QUARTERS

D. ANALYSIS OF RESULTS

This technique, continued iteratively for fourteen quarters reveals rather startling results. Using the output DATA IN TABLE 8 (RSQUARE, PREDICTED AND ACTUAL COST parameters), the regression results on this data (Fig. 13),

and the plot in Figure 14, this researcher asserts the following:

1. As expected, the initially projected contract parameters will differ from the actual values; however, the differences are not usually significant,

2. Control usually gets better as each new data point is added, because knowledge about system behavior becomes more precise as the project proceeds,

3. The technique is adaptive - it indicates present status of the contract,

4. The technique is predictive - it tells where the contract is going.

These assertions by this researcher are supported primarily by the less than one percent error between the projected target value, (K), and the actual TARGET after thirteen quarters of application of this methodology (Table 7). At this point in time of the contract, eighty-six percent of the targeted cost ($ 213.3 million) , had been expended, (represented by ACWP, $ 184.3 million). Eight quarters remained until the estimated contract completion date. While a universally acceptable definition of the term 'early', as it relates to contract control, could be subject to a great deal of debate by many,

this researcher claims that reasonably accurate knowledge about contract parameters when less than two-thirds through the length of the contract can be classified as 'early'. It is acknowledged that it may then be hard to change the final results by more than a small amount, however.

Table 8: OUTPUT DATA

print c1 c6 c33 c9 c29 c31

PERIOD	R-SQUARE	TARGET	K	INACWP	YHAT	TD
1		0.0		14.6000		
2		208.000		11.7000		
3		214.100		16.0000		
4	60.6	210.300	88.038	18.9000		2.85761
5	73.7	208.900	100.666	15.6000	11.6638	3.15127
6	79.3	205.700	114.610	14.8000	9.9280	3.48985
7	82.5	206.000	127.457	13.6000	8.8117	3.81550
8	77.3	212.500	148.836	16.4000	7.7755	4.36046
9	76.9	212.300	166.284	15.0000	8.3717	4.79618
10	76.5	212.000	183.510	14.8000	8.2239	5.22739
11	79.3	214.500	194.688	12.2000	8.0712	5.51448
12	83.8	211.500	199.622	9.0000	7.1985	5.65135
13	82.2	213.300	213.036	11.7000	5.7651	6.04916
14	81.5	225.400	225.155	10.9000	5.5987	6.41681
15					5.3377	

```
regress c15 1 c13,(c18(c19(c20)))

 THE REGRESSION EQUATION IS
 Y =     1.70 -0.0121 X1

                                          ST. DEV.     T-RATIO =
       COLUMN      COEFFICIENT            OF COEF.     COEF/S.D.
       --              1.6989              0.1595         10.65
 X1    C13          -0.012143            0.001670         -7.27

 THE ST. DEV. OF Y ABOUT REGRESSION LINE IS
 S = 0.3885
 WITH (  14- 2) =  12 DEGREES OF FREEDOM

 R-SQUARED = 81.5 PERCENT
 R-SQUARED = 80.0 PERCENT, ADJUSTED FOR D.F.

 ANALYSIS OF VARIANCE

  DUE TO      DF            SS       MS=SS/DF
 REGRESSION    1        7.9773         7.9773
 RESIDUAL     12        1.8115         0.1510
 TOTAL        13        9.7888

           X1            Y     PRED. Y   ST.DEV.
 ROW      C13          C15       VALUE   PRED. Y   RESIDUAL
  1         1        2.681       1.687     0.158      0.994
 14       196       -0.250      -0.681     0.231      0.431


 DURBIN-WATSON STATISTIC = 0.68
```

Figure 13: REGRESSION RESULTS AFTER 14 QUARTERS

```
mplot c15 vs c13 and c19 vs c13
 Ln(y/t)
 2.70+ A
     -
     -
     -
     -
 1.80+  A
     -  BB2A                          A = ACTUAL
     -     B B                        B = PREDICTED
     -        B
     -       A    B
 0.90+          A     B
     -             A  A   B
     -                    A    B
     -                         A
     -                              2
 0.0 +                                  B
     -                                       A     A
     -                                  A    B
     -
     -                                            B
 -.90+
       +---------+---------+---------+---------+---------+
       0.        50.       100.      150.      200.      250.
                                                   TSQUARE
```

Figure 14: REGRESSION RESULTS AFTER 14 QUARTERS
(Graphical Presentation)

The use of r-square as a measure of the explanatory power of the regression, in particular, as a measure of how well the estimated model fits the available data, is an acceptable standard in statistics [Ref. 12]. R-square values tend to be high when using large sample sizes of time-series data, and a value of ninety percent or higher is usually expected [Ref. 10]. While the resultant R-SQUARE value after thirteen quarters (a relatively small sample size), is 82.2 percent (Table 8), short of the expected value, the difference is not great enough to discard this model from consideration for applicability.

The proposed model falls short in its forecast of Yhat, the next quarter's INACWP. A comparison of the two columns in Table 8 reveals the differences.However, the predicted values are within the 95 percent confidence interval for actual INACWP values.

The parameter t_d projects the period in which the INACWP is maximum. After fourteen quarters, the value of t_d is 6.41681 quarters. However, the actual maximum INACWP occurred in quarter four.

The Durbin-Watson statistic tests for the presence of autocorrelation, in which the stochastic disturbance terms (u) of the regression are not independent of one another. The problem of autocorrelation is a frequent, if not typical, one when using time series data. The stochastic disturbance term at one observation will be related to the disturbance term of nearby observations. The value of the Durbin-Watson statistic (Fig. 13) is 0.68 and indicates the presence of first-order autocorrelation in this model [Ref. 10]. Therefore, the least-square estimators, the regression coefficients, are linear, unbiased, and consistent. However, the usual t-test of significance of coefficients and the F-tests of the significance of the entire regression will, in general, be biased. While there are possible

methods to treat first-order autocorrelation, such treatment methods were not pursued in this paper.Methods are available in [Ref. 10].

E. INFLATION

Because its exact values are unknown, inflation degrades the efforts of contract management to control project outcomes. An attempt was made to investigate whether or not the proposed methodology could infer a degree of measurement of the impact of inflationary uncertainty.

To measure the impact of inflation, the proposed methodology was applied to the data used in the previous section after deflation to constant 1972 dollars (Table 9). The hypothesized standard of measure of inflation impact is that an improved model fit (increase in r-square) and improved contract management parameter values (when compared to deflated actual values),would imply a negative impact of inflation on control and estimation.

The results of the application of the proposed methodology are displayed in Table 10, Figure 15, and Figure 16. While r-square slightly increased in value from 81.5 percent (Table 8) to 82.7 percent(Table 10) after fourteen quarters, the average differences in IDACWP and DYHAT are less than the average differences in INACWP and YHAT. DK,

the estimated parameter of primary concern, reaches the actual TARGET value after twelve quarters but thereafter, continues to increase through the fourteenth quarter. Dt_d is 6.10847 quarters compared to the actual maximum IDACWP occurring in the fourth quarter. The Durbin-Watson statistic value of 0.66 (Fig. 16) indicates the presence of first-order autocorrelation.

The improved fit of the model to the data, and the improved estimates of the parameter values(excepting target cost), demonstrate a potential capability of this methodology to effectively measure the impact of inflation.

Table 9: DEFLATED INPUT VALUES

print c10 c11 c12 c52 c51

PERIOD	IMPDEF	DEFLAT	IDACWP	DACWP	DTARGET
1	131.600	0.759878	11.0942	11.094	0.0
2	135.300	0.739098	8.6474	19.438	153.732
3	140.000	0.714286	11.4286	30.214	152.929
4	144.000	0.694444	13.1250	42.500	146.042
5	144.800	0.690608	10.7735	53.039	144.268
6	148.200	0.674764	9.9865	61.808	138.799
7	150.400	0.664894	9.0426	69.947	136.968
8	147.000	0.680272	11.1565	82.721	144.558
9	158.000	0.632911	9.4937	86.456	134.367
10	160.300	0.623830	9.2327	94.448	132.252
11	169.600	0.589623	7.1934	96.462	126.474
12	166.200	0.601685	5.4152	103.851	127.256
13	163.700	0.610874	7.1472	112.584	130.299
14	174.700	0.572410	6.2393	111.734	129.021

Table 10: OUTPUT DATA (DEFLATED)

print c34 c51 c44 c12 c50 c38

PERIOD	RSQUARE	DTARGET	DK	IDACWP	DYHAT	DTD8
1		0.0		11.0942		
2		153.732		8.6474		
3		152.929		11.4286		
4	64.1000	146.042	60.851	13.1250		2.72976
5	75.5000	144.268	70.111	10.7735	7.62886	3.03449
6	80.5000	138.799	79.510	9.9865	6.46860	3.36337
7	83.2000	136.968	87.917	9.0426	5.64129	3.68105
8	77.2000	144.558	102.761	11.1565	4.89319	4.22577
9	77.8000	134.367	113.025	9.4937	5.36139	4.61266
10	77.9000	132.252	122.594	9.2327	5.06618	5.00000
11	81.2000	126.474	129.436	7.1934	4.79655	5.24142
12	85.1000	127.256	132.913	5.4152	4.11282	5.39164
13	83.1000	130.299	140.690	7.1472	3.24836	5.77350
14	82.7000	129.021	146.841	6.2393	3.12383	6.10847
15					2.89520	

```
regress y c43 1 c42,(c44(c45(c46)))

 THE REGRESSION EQUATION IS
 Y =     1.37 -0.0134 X1

                                          ST. DEV.  T-RATIO =
       COLUMN      COEFFICIENT            OF COEF.  COEF/S.D.
       --              1.3723              0.1688       8.13
 X1    T²          -0.013403           0.001768      -7.58

 THE ST. DEV. OF Y ABOUT REGRESSION LINE IS
 S = 0.4112
 WITH (  14- 2) =  12 DEGREES OF FREEDOM

 R-SQUARED = 82.7 PERCENT
 R-SQUARED = 81.3 PERCENT, ADJUSTED FOR D.F.

 ANALYSIS OF VARIANCE

  DUE TO      DF            SS       MS=SS/DF
 REGRESSION    1        9.7183         9.7183
 RESIDUAL     12        2.0287         0.1691
 TOTAL        13       11.7471

           X1            Y     PRED. Y    ST.DEV.
 ROW       T²    LN(DY/T)       VALUE    PRED. Y   RESIDUAL
  1         1       2.406       1.359      0.167      1.047
 14       196      -0.808      -1.255      0.244      0.446


 DURBIN-WATSON STATISTIC = 0.66
```

Figure 15: REGRESSION RESULTS (DEFLATED)

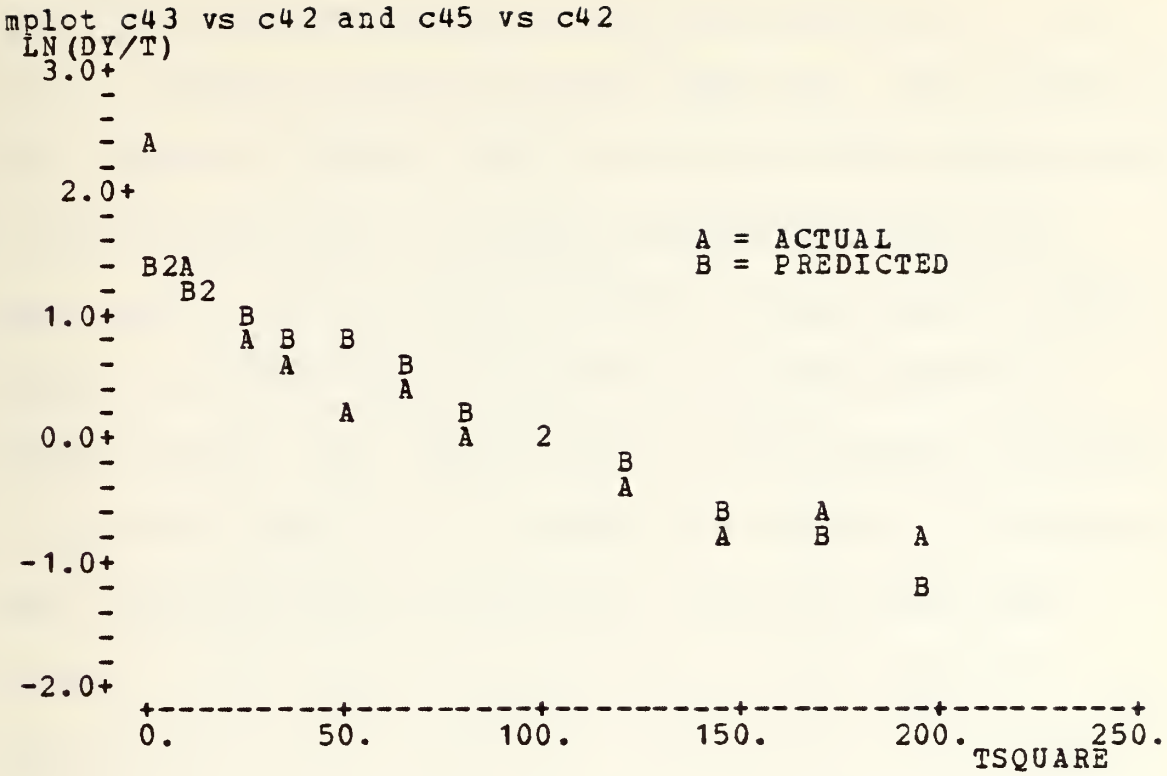

Figure 16: REGRESSION RESULTS (DEFLATED)
(GRAPHICAL PRESENTATION)

F. SUMMARY RESULTS

Presented in this section are summary results of the application of the proposed methodology to a helicopter engine contract and to a cruise missile contract, both of which were 92 percent complete.

1. Helicopter Engine

The significant input and output data, and regression results are displayed in Table 11, Table 12, Figure 17, and Figure 18. The projected target cost, K, was

less than 9 percent in error after nine quarters and less than 3 percent in error after ten quarters. The estimate of next quarter's INACWP, YHAT, is closest to the actual INACWP in quarter ten, at a 12 percent difference. After ten quarters, t_d, the projected period in which maximum INACWP occurs is 4.6225, while the actual maximum occurred in quarter eight. An r-square value of 86.8 percent indicated a relatively good fit of the model to the data. As expected when regressing time series data, autocorrelation was present, thus, invalidating the use of tests for the significance of the regression and regression coefficients.

Table 11: Input Data

print c1 c9 c14 c13 c15

PERIOD	INACWP	Y/T	TSQUARE	LN(Y/T)
1	2.60000	2.60000	1.	0.95551
2	2.70000	1.35000	4.	0.30010
3	2.50000	0.83333	9.	-0.18232
4	3.10000	0.77500	16.	-0.25489
5	3.10000	0.62000	25.	-0.47804
6	3.40000	0.56667	36.	-0.56798
7	3.10000	0.44286	49.	-0.81451
8	4.20000	0.52500	64.	-0.64436
9	1.80000	0.20000	81.	-1.60944
10	1.30000	0.13000	100.	-2.04022

Table 12: Output Data (Helicopter Engine)

print c35 c6 c33 c9 c29 c31

PERIOD	RSQUARE	TARGET	K	INACWP	YHAT	TD
1		28.7000		2.60000		
2		28.7000		2.70000		
3		28.7000		2.50000		
4	78.1000	28.7000	14.3456	3.10000		2.57855
5	75.9000	28.7000	18.3023	3.10000	1.64615	3.11588
6	71.9000	28.7000	23.0065	3.40000	1.77138	3.71135
7	74.7000	28.7000	26.4437	3.10000	1.97430	4.13096
8	64.8000	28.8000	33.0709	4.20000	1.90074	4.90290
9	78.9000	28.8000	31.3195	1.80000	2.29659	4.69323
10	86.8000	30.1000	30.8727	1.30000	1.46901	4.62250
11					0.93660	

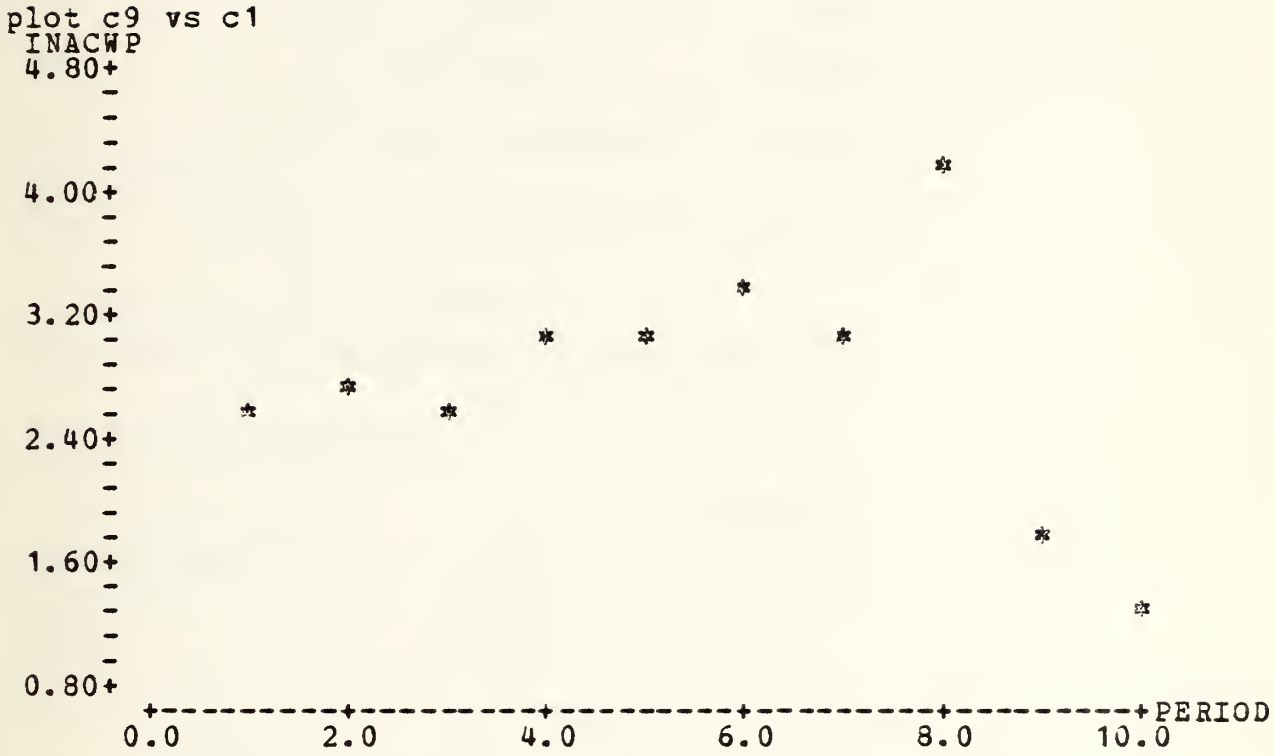

Figure 17: INACWP versus PERIOD (Helicopter Engine)

```
REGRESS Y C15 1 C13,(C18(C19(C20)))

 THE REGRESSION EQUATION IS
 Y =   0.368 -0.0234 X1

                                          ST. DEV.    T-RATIO =
       COLUMN      COEFFICIENT            OF COEF.   COEF/S.D.
       --              0.3677              0.1625        2.26
 X1    C13          -0.023411            0.003228       -7.25

 THE ST. DEV. OF Y ABOUT REGRESSION LINE IS
 S = 0.3309
 WITH (  10- 2) =   8 DEGREES OF FREEDOM

 R-SQUARED = 86.8 PERCENT
 R-SQUARED = 85.1 PERCENT, ADJUSTED FOR D.F.

 ANALYSIS OF VARIANCE

  DUE TO      DF            SS       MS=SS/DF
 REGRESSION    1         5.7603         5.7603
 RESIDUAL      8         0.8761         0.1095
 TOTAL         9         6.6363

           X1            Y     PRED. Y    ST.DEV.
 ROW      C13          C15       VALUE    PRED. Y    RESIDUAL      ST.RES.
  1         1        0.956       0.344      0.160       0.611        2.11R
 10       100       -2.040      -1.973      0.224      -0.067       -0.27 X
```

Figure 18: Regression Results (Helicopter Engine)

2. Cruise Missile

A plot of INACWP versus PERIOD (Fig. 19), the input data (Table 13), the output data (Table 14), and the regression results (Fig. 20) were reviewed to summarize the results of applying the proposed methodology to a cruise missile contract.

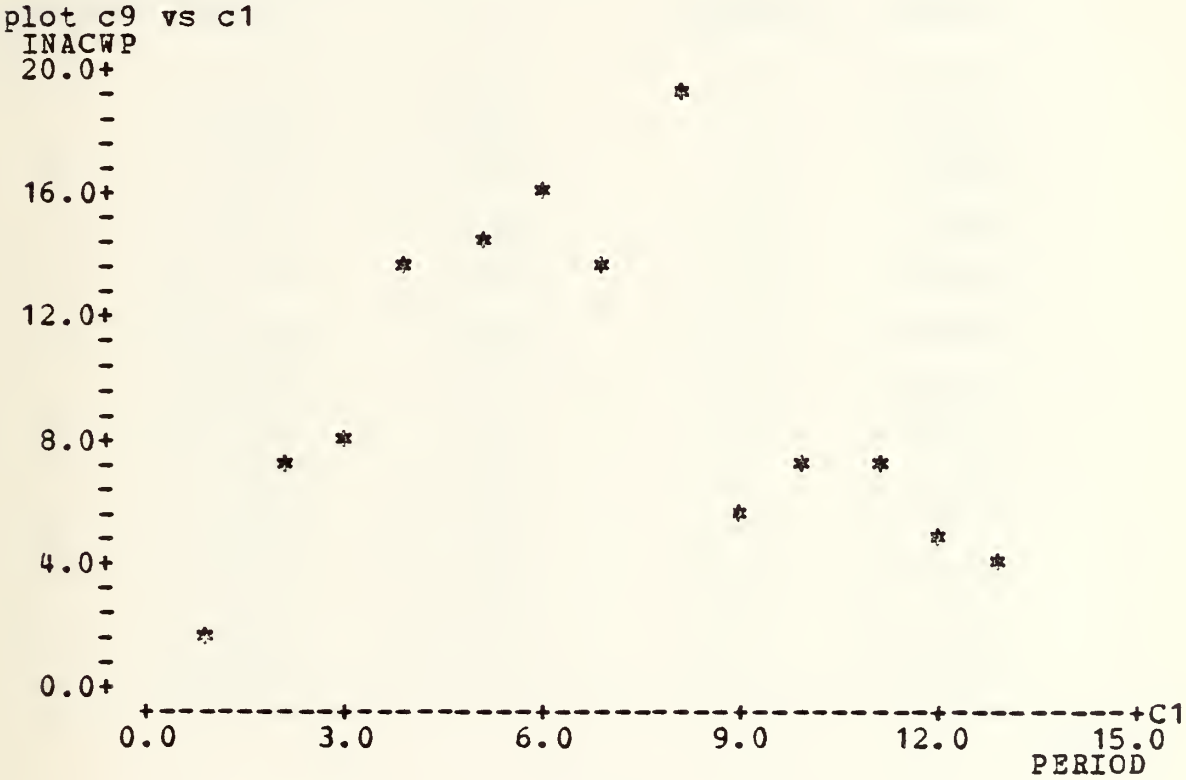

Figure 19: INACWP versus PERIOD (Cruise Missile)

Table 13: INPUT DATA (Cruise Missile)

print c1 c9 c14 c16 c17

PERIOD	INACWP	Y/T	TSQUARE	LN(Y/T)
1	2.0000	2.00000	1.	0.69315
2	6.8000	3.40000	4.	1.22377
3	8.2000	2.73333	9.	1.00552
4	13.9000	3.47500	16.	1.24559
5	14.6000	2.92000	25.	1.07158
6	16.1000	2.68333	36.	0.98706
7	13.2000	1.88571	49.	0.63431
8	19.4000	2.42500	64.	0.88583
9	5.4000	0.60000	81.	-0.51083
10	7.4000	0.74000	100.	-0.30111
11	7.2000	0.65455	121.	-0.42381
12	4.9000	0.40833	144.	-0.89567
13	4.1000	0.31538	169.	-1.15396

Table 14: OUTPUT DATA (Cruise Missile)

print c35 c6 c33 c9 c29 c31

PERIOD	RSQUARE	TARGET	K	INACWP	YHAT	TD
1		117.200		2.0000		
2		107.200		6.8000		
3		110.300		8.2000		
4	42.9000	110.900	-45.666	13.9000		
5	18.7000	114.600	-129.034	14.6000	22.1855	
6	3.3000	114.900	-501.376	16.1000	21.8928	
7	13.4000	117.200	300.478	13.2000	21.6328	10.1015
8	13.3000	120.900	400.885	19.4000	17.2160	11.7851
9	50.9000	121.600	125.780	5.4000	19.4068	6.0634
10	66.3000	128.500	120.733	7.4000	8.7810	5.8521
11	74.7000	129.300	122.899	7.2000	6.6280	5.9339
12	82.4000	132.800	120.733	4.9000	5.4201	5.8521
13	87.2000	132.800	121.192	4.1000	3.8867	5.8926
14					2.9057	

```
regress y c17 1 c16,(c18(c19(c20)))

THE REGRESSION EQUATION IS
 Y =     1.25 -0.0144 X1

                                        ST. DEV.     T-RATIO =
       COLUMN      COEFFICIENT          OF COEF.     COEF/S.D.
       --              1.2515            0.1382          9.05
 X1    C16          -0.014418          0.001668        -8.64

 THE ST. DEV. OF Y ABOUT REGRESSION LINE IS
 S = 0.3238
 WITH (  13- 2) =  11 DEGREES OF FREEDOM

 R-SQUARED = 87.2 PERCENT
 R-SQUARED = 86.0 PERCENT, ADJUSTED FOR D.F.

 ANALYSIS OF VARIANCE

  DUE TO      DF            SS       MS=SS/DF
 REGRESSION    1         7.8312        7.8312
 RESIDUAL     11         1.1530        0.1048
 TOTAL        12         8.9842

           X1            Y     PRED. Y    ST.DEV.
 ROW      C16          C17       VALUE    PRED. Y   RESIDUAL
 13       169      -1.1540     -1.1851     0.1983     0.0311


 DURBIN-WATSON STATISTIC = 2.03
```

Figure 20: REGRESSION RESULTS (Cruise Missile)

A close look at the RSQUARE and K columns of Table 14 indicated that the proposed model experienced a very poor fit to the data as well as a poor predictive ability in

quarters four through six. The negative K values, the very low RSQUARE values, and the absence of t_d values resulted from the large increases in INACWP during the first six quarters. Regressing Ln(y/t) on t^2 in quarters four through six yielded positive slope coefficients which in turn resulted in negative values of t_d^2. Since K is the product of t_d^2 and the exponential of A (the intercept), and t_d is the square root of t_d^2, t_d^2 being negative indeed causes these abnormal projection results (see the derivatives in Section C). After thirteen quarters, K differed from the actual TARGET value by 9 percent, YHAT differed from INACWP by 6 percent, and t projected the maximum INACWP to occur in quarter 5.89, while the actual maximum occured in quarter eight. The RSQUARE value after thirteen quarters was 87.2 percent.

VI. CONCLUSION

A. GENERAL

A straightforward, simplified methodology was presented for the study of the behavior of contract costs. The interactive capability of MINITAB was exploited to provide a means of rapidly manipulating selected data (in its reported format, i.e.,columns) for analysis and quickly testing emergent hypotheses. The tools proposed, scatter plotting and regression, as well as the model proposed, the Rayleigh-Norden model, were demonstrated using actual data.

Several relationships were hypothesized and although only a few contracts were suitable for application of the proposed methodology, the relationships were statistically significant. The use of current quarterly ACWP(INACWP) and elapsed time (TSQUARE) were found to be good input parameters to the adapted Rayleigh-Norden model producing good predictors for total contract costs(K), and for next quarter's INACWP(Yhat). Resulting predictors for t_d proved to be surprisingly poor. R-square values in all contracts demonstrating Rayleigh-Norden shape characteristics indicated close fitting of the data to the proposed model.

B. SUMMARY

The impact of uncontrolled cost growth on the acquisition of defense weapons systems has directed research effort toward gaining a clear understanding of the control and estimation of that cost growth. The basis for such understanding lies in analysis of the behavior of acquisition contracts, especially in aggregate at the systems level.

This thesis has presented a broadly applicable and simple methodology, using the interactive capability of the MINITAB computing system, for conducting an analysis of both current and deflated quarterly contract ACWP variables. Predictability attained through model fitting, scatter diagramming, and regression analysis will however, depend upon the availability of relevant, precisely-defined data accruing from standard reporting practices.

It is the author's opinion that much of the behavior of contract performance and costs can be learned from existing data. The further application of this proposed methodology to a greater number and a wider variety of DOD contracts should increase the practitioner's knowledge of the behavior of contract performance and costs. A speculative assessment is that the behavior of performamnce and costs will be, to a

great extent, contractor dependent. The possible growth of knowledge in conjunction with a possible refinement of this methodology may provide cost estimators, program managers, and DOD decision makers a tool with which to better resolve their ultimate management problem, i.e., the estimation and control of major programs. Major programs are composed of a multitude of contracts of the type addressed in this thesis, as well as many more sub-contracts, all of which require close monitoring if costs and schedules are to be controlled. The estimation and control of aggregate contracts are issues worthy of expanded research.

LIST OF REFERENCES

1. Bryan, Noreen S. and Clark,Rolf, "Is Cost Growth Being Reinforced?" Concepts,v.4, no.2, pp. 105, 1981.

2. Gansler, Jaques S., The Defense Industry, The Massachusetts Institute of Technology, 1981.

3. Baumgautner, J. Stanley, Systems Management, The Bureau of National Affairs Inc., 1979.

4. "Cost/Schedule Control Systems Criteria Joint Implementation Guide," AMCP 37-5, NAVMAT P-5240, AFSCP/AFLCP 173-5, March 31, 1972.

5. Christle, Gary E.,Automation of Program/Project Cost Reports Within DOD, paper presented at DOD/FAI Acquisition Research Symposium, 9th, Annapolis, Maryland, 9-11 June 1980.

6. Defense Science Board, "Summer Study of 1977", Department of Defense, Washington, D.C., March 1978.

7. Chacko, Gearge K., "Improving Cost and Schedule Controls Through Adaptive Forecasting", Concepts, Summer 1981 v.4, no. 3, pp. 74-75.

8. Norden, Peter V., "Useful Tools for Project Management", pp. 216- 225, in Quantitative Management: Software Cost Estimating, Putnam, Lawrence H. and Wolvertonm Ray W., IEEE, New York, N.Y., 1977.

9. Putnam, Lawrence H., "A General Empirical Solution to the Macro Software Sizing and Estimating Problem",pp. 138 -154, in Software Cost Estimating and Life-Cycle Control: Getting the Software Numbers, by Putnam, Lawrence H., Computer Society Press, 1980.

10. Intriligator, Michael D., Econometric Models, Techniques, and Applications, Prentice-Hall, 1978.

11. Ryan, Thomas A., Jr., Joiner, Bryan L., and Ryan, Barbara F., Minitab Reference Manual, Duxbury, 1981

12. Draper, N.R. and Smith, H. Applied Regression Anaylsis,Wiley,1966.

13. DeGroot, Morris H., Probability and Statistics,Addison-Wesley,1975.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Technical Information Center Cameron Station Alexandria, Virginia 22314	2
2.	Library, Code 0142 Naval Postgraduate School Monterey, California 93940	2
3.	Asst Professor D.C. Boger, Code 54 Bk Department of Administrative Sciences Naval Postgraduate School Monterey, California 93940	2
4.	Professor M.G. Sovereign, Code 55 Z0 Department of Operations Research Naval Postgraduate School Monterey, California 93940	4
5.	Lieutenant Harry Watkins III, USN 3217 Kalahari N.E. Canton, Ohio 44705	2
6.	Dr. T.C. Varley, Code 411OR Office of Naval Research 800 N. Quincy St. Arlington, Virginia 22217	1
7.	Dr. J.Augusta, Director Navy Center for Acquisition Research Code 75 Naval Postgraduate School Monterey, California 93940	1
8.	Mr. J. Kammerer DASD (Cost and Audit) Rm. 3A336 Pentagon Washington, D.C. 20301	1
9.	Professor Kneale T. Marshall Department Chairman, Code 55 Department of Operations Research Naval Postgraduate School Monterey, California 93940	1

Thesis
W233
c.1

Watkins

An application of Rayleigh curve theory to contract cost estimation and control.

thesW233
An application of Rayleigh curve theory
3 2768 000 99481 8
DUDLEY KNOX LIBRARY