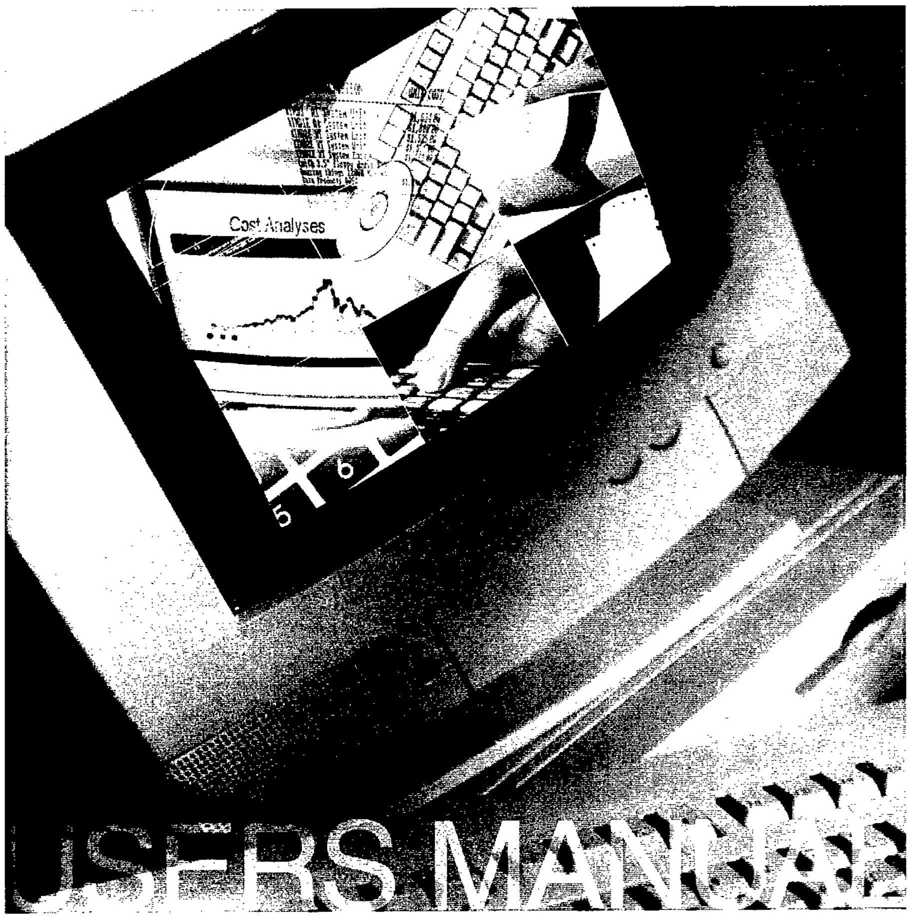
Logistics Management Institute

The Rayleigh Analyzer®

Volume II—Users Manual

AT902C2



John Dukovich
Scott Houser
David A. Lee

LMI®

Logistics Management Institute

The Rayleigh Analyzer®

Volume II—Users Manual

AT902C2

October 1999

John Dukovich
Scott Houser
David A. Lee

Prepared pursuant to Department of Defense Contract DASW01-99-F-1035. The views expressed here are those of the Logistics Management Institute at the time of issue but not necessarily those of the Department of Defense. Permission to quote or reproduce any part except for government purposes must be obtained from the Logistics Management Institute.

LOGISTICS MANAGEMENT INSTITUTE
2000 CORPORATE RIDGE
MCLEAN, VIRGINIA 22102-7805

Contents

FIGURES

TABLES

Chapter 1
Overview

This volume is the second in a series of two. The first volume described the theory and application of the Rayleigh statistical model. This volume is the users manual for operating the *Rayleigh Analyzer*® model.[1]

In this chapter we discuss the background of the *Rayleigh Analyzer* model, briefly describe the previous version of the model, and introduce the latest model version.

BACKGROUND

For several years, the Logistics Management Institute (LMI) has been providing a Microsoft Excel-based tool, the *Rayleigh Analyzer*, for analyzing and estimating development program costs and schedules. The tool is based on the Norden-Rayleigh statistical model. Volume I of this series, *Theory and Applications*, discusses the theoretical details of the model.[2] In this volume, we discuss implementing the theory in the latest version of a computer-based tool that is available to cost analysts and other personnel interested in evaluating the costs and schedules of development programs.

The *Rayleigh Analyzer* generates statistics of completion time and total cost for development programs, conditioned on actual cost of work performed (ACWP) data and on the Norden-Rayleigh statistical model. The *Rayleigh Analyzer* uses multiple model adaptive estimation (MMAE) a widely-used system identification method that employs Kalman filters, to identify the parameters of the Rayleigh curve that best models the data. Gallagher and Lee describe the application of MMAE fully.[3]

The purpose of the *Rayleigh Analyzer* is to obtain better estimates of the total cost and completion time on development contracts. By applying it to a current contract, one can estimate its total cost and completion time.

Our research was sponsored by the Office of the Secretary of Defense, Acquisition Program Integration (API). API staff routinely use the model to analyze current U.S. Department of Defense (DoD) research and development (R&D) programs.

[1] *Rayleigh Analyzer* is a registered trademark of the Logistics Management Institute.

[2] Dukovich, J., Houser, S., and Lee, D., *The Rayleigh Analyzer—Volume I: Theory and Applications*, LMI report AT902C1, October 1999.

[3] M. A. Gallagher and D. A. Lee, "Final-Cost Estimates for Research & Development Programs Conditioned on Realized Costs," *Military Operations Research*, Volume 2, Number 2, 1996, pp. 51-65.

Previous Model Release

The previous version of the model—released in March 1998—allowed for two types of analyses.[4] In the first type of analysis, the analyst would choose a contract of interest, run the model, then view estimates of completion time and total cost. In the second type of analysis, the analyst would fix the completion time to estimate the total cost on the basis of that constraint.

Various resulting graphs were available to users, including best-fit curves; marginal cumulative distribution functions (CDFs) for total cost and completion time, both "raw" and smoothed; and confidence limits.

Adding another development program to the *Rayleigh Analyzer* was difficult, as was updating program cost histories using a database. Neither process was automated.

Current Model Release

Our client in the API office requested that LMI investigate improving the *Rayleigh Analyzer*. First, API requested that we automate the process for updating the model with new data. API also wanted to investigate the effect of acquisition reform on the Rayleigh statistical model. In addition, LMI investigated to discover the programmatic effects that deviated from the Rayleigh distribution, but that could be modeled in the *Rayleigh Analyzer*. The effects include expanding the requirements and restructuring contracts.

LMI interviewed a cross section of representatives of programs that were influenced by acquisition reform. Researchers analyzed data from the programs and asked program representatives to comment on the applicability of the *Rayleigh Analyzer* to their programs. Each representative also had the opportunity to comment on the model's accuracy, and to name programmatic effects that may have influenced spending profiles. Volume I discusses the results of our interviews.

This updated version of the model has several improvements and new features, some resulting from issues raised during the interviews. One automation improvement includes the data update feature, which is now a one-button-click process. In addition, one has the option of defining unique programs or analyzing temporary data that is independent of the main database. Also, we added several new output features, including both then-year and constant-dollar charts.

Our investigation also found that the Rayleigh distribution still applies to most post-acquisition-reform programs. Our researchers modified the *Rayleigh Analyzer* to account for events, identified during program office interviews, that could affect the fundamental shape of the Rayleigh model. A change in scope midway

[4] Lee, D. A. and Dukovich, J. A., *Using the Rayleigh Analyzer®—Beta Test Version*, LMI report AT701C1, March 1998.

through the execution of a contract is an example of such an event. We now have a tool that combines several Rayleigh curves and effectively models such events.

Chapter 2

Model Installation

WEBPAGE DOWNLOAD

The World Wide Web is the primary means of disseminating this version of the *Rayleigh Analyzer*. It can be downloaded from

`http://alpha.lmi.org/rayleigh/`

The web page contains a short description of the model, installation instructions, and information on how to contact the developers for questions or comments.

The file is a self-extracting executable program. Once downloaded to your hard drive and run, the program extracts several necessary files. Please follow the installation instructions in the file named "install.txt." Because the *Rayleigh Analyzer* uses Excel's Add-in structure and the Solver tool, **it is very important to follow the installation instructions exactly**. The next subsection contains the installation instructions. They are identical to those included with the software.

If you are unable to access the World Wide Web, you can request a diskette containing the model. Send written requests, including your name, title, organization, mailing address, and telephone number to:

Resource Analysis Group
Logistics Management Institute
2000 Corporate Ridge
McLean, VA 22102-7805
USA

We are interested in your comments and feedback, which you can provide either via the World Wide Web or by using the above mailing address.

If you are not using a 32-bit operating system (such as Windows 95/98), the *Rayleigh Analyzer* will not function. The version of the *Rayleigh Analyzer* described in this manual was designed for Excel 97 operating on Windows 95, Windows 98, or Windows NT 3.51.

INSTALLING THE MODEL

The *Rayleigh Analyzer* uses Excel's Add-in structure and the Solver tool. Therefore, it is **very important to follow these installation instructions exactly**. The following instructions are identical to those in the file *install.txt,* which is included in the download package located at http://alpha.lmi.org/rayleigh/.

1. After downloading, locate the self-extracting file on your hard drive (it will be named "Ray99.exe").
2. Double-click on the file. (This will launch the WinZip® self-extractor. No additional software is required for this operation.)
3. Make sure the text box labeled "Unzip to Folder" contains the folder "c:\api"—if not, type "c:\api" into the text box. *The Rayleigh Analyzer can only operate from this working directory.*
4. Click the "Unzip" button.
5. Click the "Close" button.
6. Launch Excel.
7. Go to the "Tools" menu and select "Add-ins…"
8. Click the "Browse" button, select folder c:\api and choose the file "Rayleigh_Analyzer.xla."
9. Click "OK" on the browse box.
10. While still in the Add-in dialog box, also check the box next to the file "Solver Add-In." If that file is not listed in the Add-Ins dialog box, click "Browse" and locate the drive, folder, and file name for the "Solver.xla" add-in—usually located in the Office\Library\Solver folder. (You may need to use the Windows Explorer to find the file. If that does not work, you may need to re-run the MS Office Setup program to install the file.)
11. Click "OK" on the Add-ins box.
12. Go to the "Tools" menu and select "Macro."
13. Select "Visual Basic Editor" from the menu.
14. In the Visual Basic Editor, go to "Tools" and select "References…"
15. Check the box next to "Solver.xls." (Some systems may show "Solver.xla" — select whichever one is displayed.)
16. Click "OK" on the references dialog box.
17. Close the Visual Basic Editor.

18. Now you can launch Rayleigh_Analyzer.xls as you would normally open an Excel file (e.g., two options are to use the File/Open menu in Excel, or to double click on the file in Windows Explorer).

REQUIRED FILES

The self-extracting file uncompresses a series of files and installs them on your hard drive. This subsection lists an inventory of the files.

The files and the required directories are:

`c:\api\Rayleigh_Analyzer.xls`	The *Rayleigh Analyzer*®
`c:\api\Rayleigh_Analyzer.xla`	The Excel Add-in
`c:\api\mmaedls.exe`	MMAE executable routine
`c:\api\data.dqy`	MS Query file

The following are the raw data files that are used during the automated data update:

```
c:\api\contract.dbf
c:\api\dcontrct.dbf
c:\api\dcontrct.dbt
c:\api\hist_pno.dbf
```

DoD cost analysts have access to the actual DBF files. However, because of the proprietary nature of the cost information in the files, a generic set of DBF files is given to non-DoD users as part of the self-extracting suite. DoD users with access to the actual DBF files should delete the generic files.

Chapter 3

Basic Model Operation

INTRODUCTION

To open the *Rayleigh Analyzer*, launch the Excel workbook named "Rayleigh_Analyzer.xls."

Select individual development programs from a list of those in the database. You also may choose to add new data or modify existing data. After running the basic model, several graphical output formats are available, including worksheets and line graphs. Experienced users may choose to use the advanced analyses. The following subsections describe the model's operation in detail.

THE START WORKSHEET

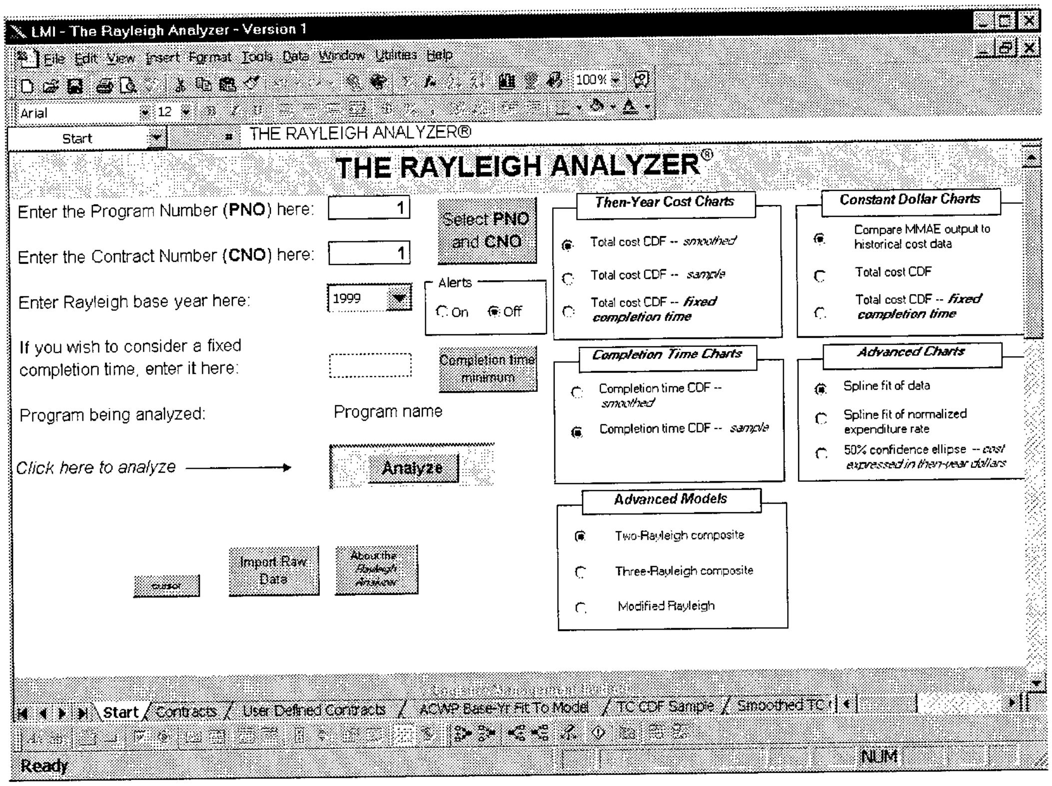
The model is operated from the Start worksheet. The Start worksheet has "The Rayleigh Analyzer®" banner across the top of the page, shows the worksheet tab labeled Start, and has several data input boxes and macro buttons. If you find that you are not on the Start worksheet, methods of navigating to it are discussed below. This worksheet appears as shown in Figure 3-1.

The left-hand side of the worksheet is the area for user input. After selecting the program of interest, you click the "Analyze" button to launch the *Rayleigh Analyzer*. After the model completes its computations, then you can select the results to review, navigating by a series of buttons on the right-hand side of the worksheet. The results are organized into types of results (then-year, constant dollar, completion time, and advanced).

To run the advanced analyses, you should select the button for one of the three advanced analyses (two-Raleigh composite, three-Rayleigh composite, or modified Rayleigh). By selecting an advanced analysis, you will be presented with a worksheet appropriate to the analysis. We advise that only those users who are familiar with programmatic details of specific contracts use these analyses, because the results are easily misinterpreted and may not be valid in some situations.

Most other worksheets are equipped with a "Return to Start" button for returning to the Start worksheet. As an alternative, you can use the scroll arrows on the lower left-hand corner of the Excel window or use the worksheet navigation drop-down box, discussed later.

Figure 3-1. The Start Worksheet



Details of operating the *Rayleigh Analyzer* are discussed in the following subsections.

Entering Contract Data

Contract data can be entered into the *Rayleigh Analyzer* in two ways. One method is for manually entering data. Another method is automated and requires that you to have access to several database files. Each method is discussed in the following subsections.

Entering Data Manually

A simple and straightforward way for entering new contract data is to go to the "User Defined Contracts" worksheet. Simply enter the data for each of the nine fields (columns) for as many rows as required. Each row represents a monthly or quarterly cost report. The fields are defined as follows:

- *PNO*—the program number. This is an integer of any value, though normally limited to three digits. If you enter a PNO/CNO combination that is a duplication of a combination already in the Contract Analysis System (CAS) database, an error will occur if you try to analyze the combination.

For those using the generic CAS data files, this won't be much of an issue, as there are only three PNO/CNO example combinations given (these can be seen on the "Contracts" worksheet).

- *PNASHORT*—an abbreviated version of the program name in text format. For example, if the true program name is F-22 Raptor, the PNASHORT might be simply F-22.
- *CNO*—the contract number. This is an integer, usually sequentially numbered, beginning with 1. A PNO may have several CNOs associated with it. For example, the F-22 aircraft PNO may have separate CNOs for the airframe, engines, and avionics.
- *CDES*—the contract description. This text field is the full name of the contract. (The field PNASHORT is often an abbreviated version of this field.)
- *RPD*—the report date. This is a date field that represents the date as of when the reported funds were expended.
- *WSDATE*—the work start date. This is another date field that represents when funds are first expended on the contract. For a given PNO/CNO combination, the WSDATE will be identical for each row.
- *ACWP*—the actual cost of work performed. This is a real number that represents the funds expended on the program, as of the RPD. The value is expressed in terms of millions of then-year U.S. dollars.
- *PPHASE*—the program phase. Since the *Rayleigh Analyzer* was designed to model development programs, this field should be completed with the term "development." Although there are many other equivalent terms, using the consistent term "development" will alleviate confusion.
- *Service*—the branch of military service responsible for the program. The model uses the following codes: A = Army; F = Air Force; N = Navy. The purpose of this field is to ensure that the appropriate inflation indices are used. If you analyze a non-DoD program, one of the three codes must be used, although no substantial difference will be noticed among the three codes selected.

Once you have included the new information to be analyzed, you must click on the "Update PNOs/CNOs" button to add the PNO/CNO combinations to the PNO/CNO dialog box on the Start worksheet.

Importing Data Automatically

In the Model Installation section, three database file names were listed. Those files (contract.dbf, dcontrct.dbf, and hist_pno.dbf) are generated by the CAS database program. CAS contains monthly cost performance report (CPR) data from all unclassified DoD programs. Many DoD analysts have access to CAS and the three files that it generates.

Because CAS data are updated monthly, and the *Rayleigh Analyzer* uses a subset of the CAS data for its analyses, we developed an automated update feature that takes the relevant data fields from CAS and imports them into the *Rayleigh Analyzer*.

For users who do not have access to CAS, we provided three generic files—with the same filenames—that can be used instead. Users who have access to CAS should delete the generic files immediately to prevent them from overwriting the CAS files.

First, you must ensure that the three database files are loaded into the c:\api subdirectory on your computer. The next step is simply to click on the "Import Raw Data" button on the Start worksheet. The update will take several minutes on current-technology computers.

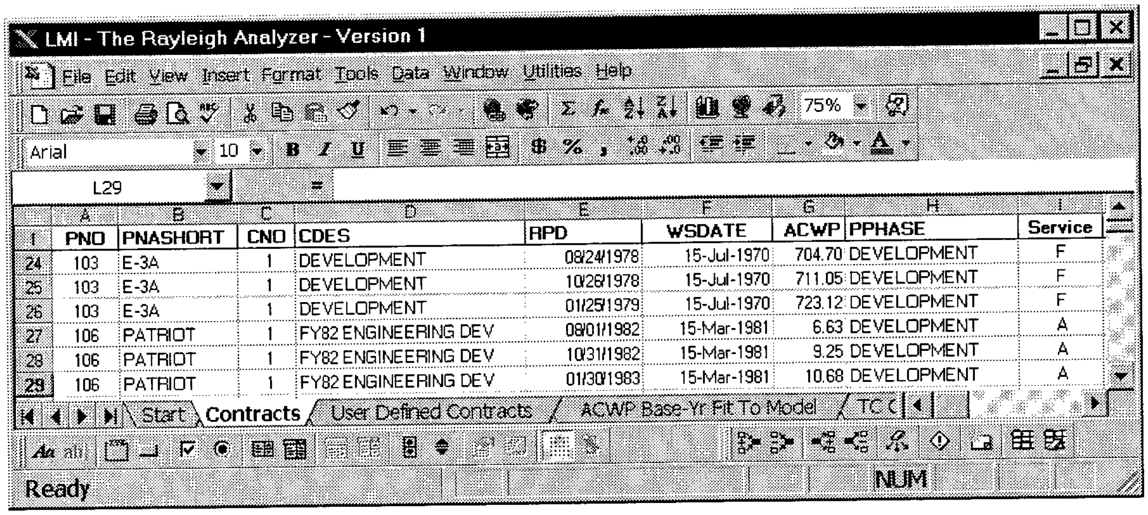
Warning: the automated update feature will overwrite data loaded onto the Contracts spreadsheet.

During the update, the model will ask whether you want to preprocess the data. Preprocessing takes additional time, but eliminates several types of database errors that could occur when operating the *Rayleigh Analyzer*. We recommend you select the preprocessing option. After preprocessing the data, the model will ask for a minimum number of cost reports, and provide a default of eight. This means that the *Rayleigh Analyzer* will not load any contract with less than eight completed cost reports into the Contracts worksheet. The MMAE algorithm requires a minimum of three cost reports simply to run; depending on the overall size of the program, several additional points may be needed to provide reasonably accurate output. Culling out contracts that have less than a minimum number of cost reports eliminates the possibility of applying the analysis prematurely.

After the update is completed, a dialog box will ask if you want to regenerate the PNO/CNO drop-down box lists. If you are sure that no new PNO and CNO identifiers have been added since the last update, clicking "No" is fine. If you know that there were additions to the PNO and CNO identifiers, or are not sure of additions, we recommend clicking on "Yes" to ensure that all of the identifiers are available for use. Updating the PNO and CNO identifiers will take several more minutes.

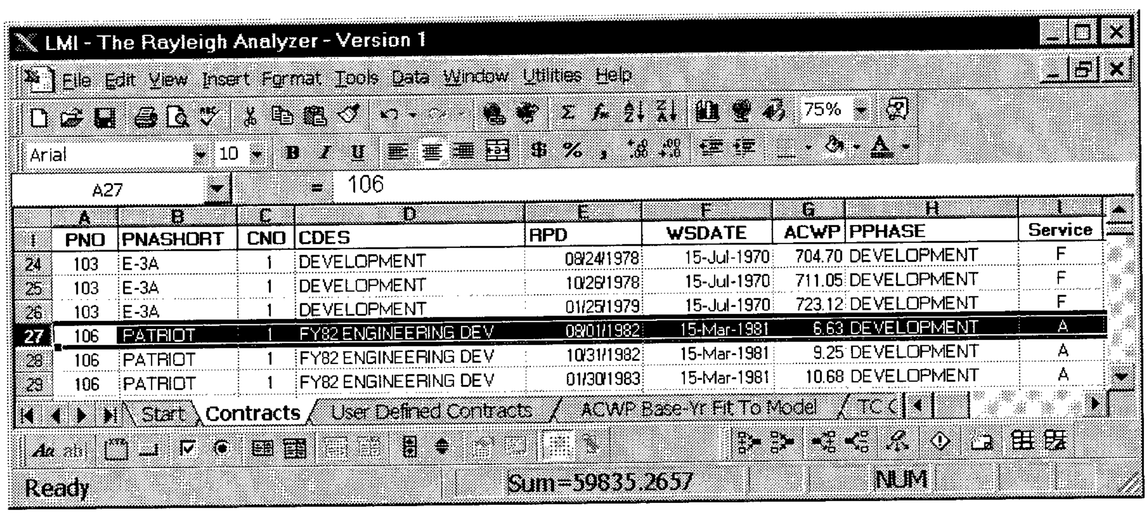
You may choose to include additional records to existing PNO/CNO combinations. Records can be added by scrolling to the appropriate area on the Contracts worksheet or the User Defined Contracts worksheet. The easiest way to add a new row of data is to highlight the entire row below the last row in the series. For example, we'll briefly explain the process for adding a data point for the E-3A program just before the Patriot data, as illustrated in Figure 3-2 (ACWP figures have been altered for sensitivity reasons).

Figure 3-2. Example Contracts Worksheet



First, select the entire row 27. This is done by clicking on the "27" row label on the left-hand side of the worksheet. The result is that the entire row is highlighted, as shown in Figure 3-3.

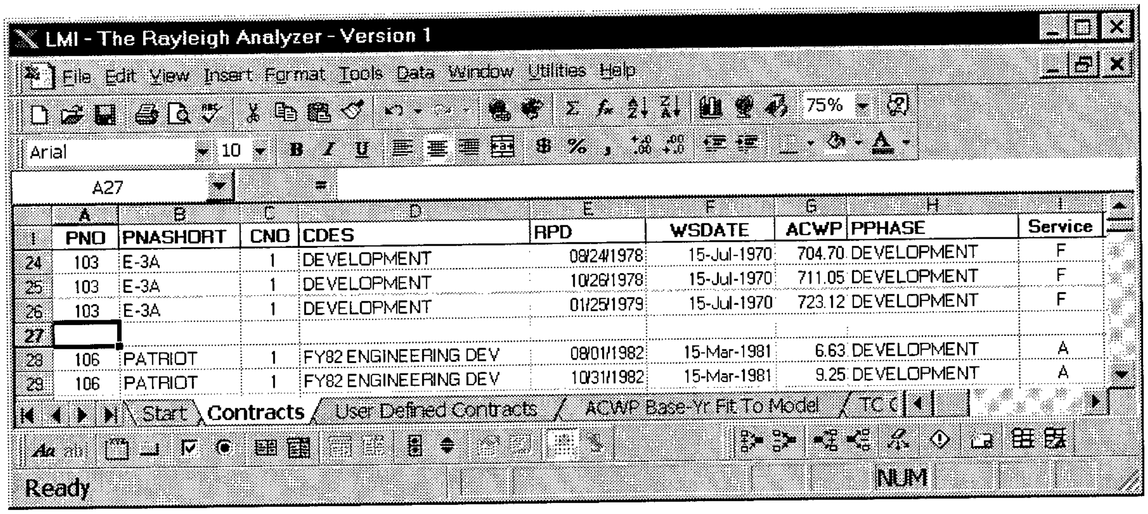
Figure 3-3. Selecting an Entire Row



Next, insert an entire row. You can do this in two ways. Either use the menu bar and choose Insert\Row, or "right-click" (use the right mouse button) on the row and choose "Insert" off of the pop-up menu.

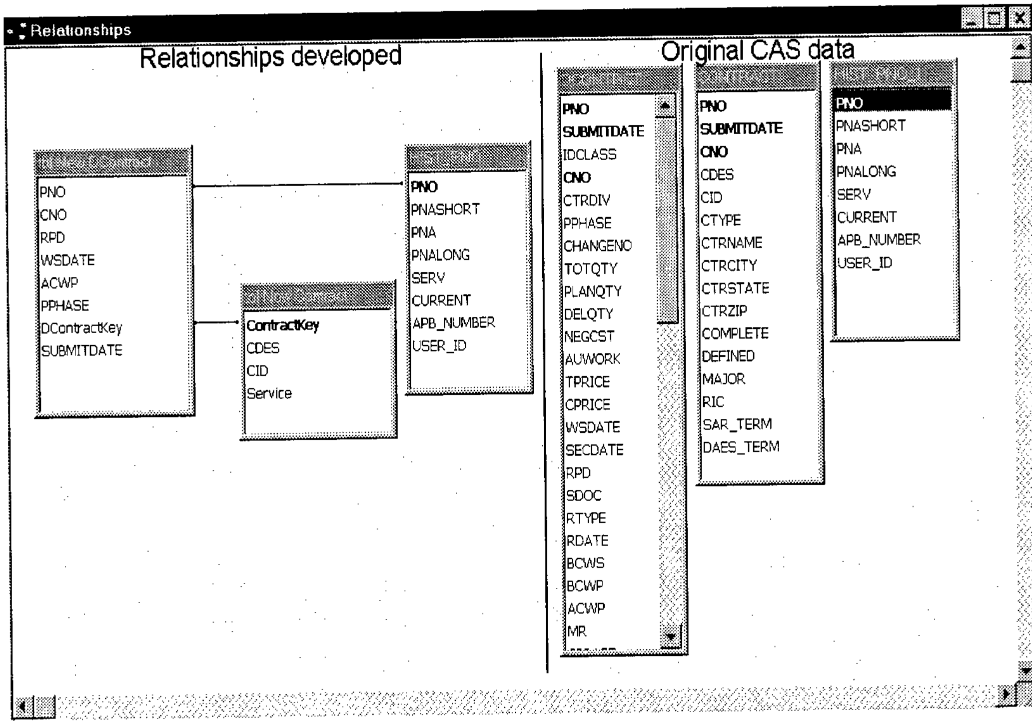
The result of the procedure is illustrated in Figure 3-4. Notice the data that were in row 27 and below have shifted down one row, leaving row 27 blank. Then simply type in the additional data. The only fields that should change from the preceding row are the RPD and the ACWP. (All of the fields are defined in the preceding subsection, "Entering Data Manually.")

Figure 3-4. A Blank Row Has Been Inserted



If you want to add a new program or one not contained in the CAS database, there are two options. The first is to modify each of the three database files (using Excel, Access, or another database application) and enter the appropriate data, ensuring that all three files are updated appropriately. Then the automated update process can be run as described previously. However, you will need a more detailed understanding of the relationships between the three database files and the fields that must be filled than we can provide in this manual. Figure 3-5 illustrates the complexity of the data. The right-hand side of the figure gives an idea of the CAS datatables and fields. Those tables need to be reduced and linked similar to the three tables on the left-hand side of the figure.

Figure 3-5. CAS Data and Relationships



Because of this complexity, we recommend entering the data manually to include additional programs in the *Rayleigh Analyzer*, as described in the previous subsection.

Basic Model Operation

The following subsections describe how to set up an analysis, run the model, and review the results.

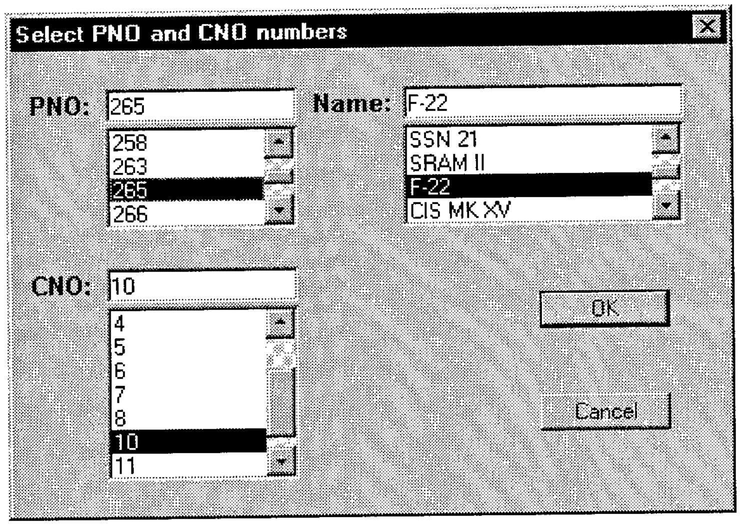
Setting Up an Analysis

Setting up an analysis requires several steps.

Choosing a Program to Analyze

If you know the PNO and CNO for the development program that you wish to analyze, type those numbers in the indicated cells of the Start worksheet (as shown in Figure 3-1). In addition, a dialog box with all applicable PNOs and CNOs is available to assist your data entry. Clicking on the "Select PNO and CNO" button displays the dialog box shown in Figure 3-6.

Figure 3-6. PNO and CNO Dialog Box

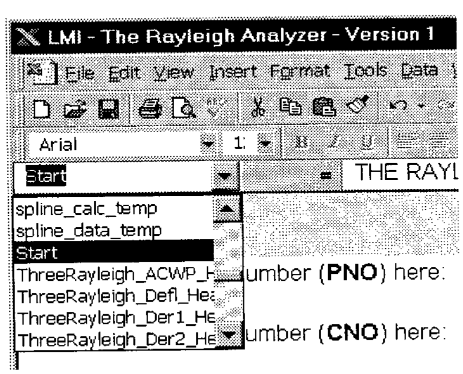


You may select the PNO either by clicking on the numbers listed in the PNO drop-down box or by choosing the program name from the "Name" list (as shown on the upper right of Figure 3-6). The two lists ("PNO" and "Name") are synchronized so that choosing a PNO number references the associated program name, and vice versa. After choosing a PNO or program name, the CNO list automatically updates to reflect the available CNOs for that PNO.

Each PNO may have one or more CNOs. Usually, the CNOs are sequentially numbered contracts under each program. You must be familiar with a specific program and contract to know which CNO should be selected. Browsing the Contracts worksheet may indicate of the differences between CNOs, if the associated "PNASHORT" or "CDES" descriptions are meaningful.

Once you have selected a PNO and CNO, click on the "OK" button to return to the Start worksheet and proceed with the directions in the subsequent paragraphs of this section. You can easily navigate to the Start and Contracts worksheets from anywhere in the workbook by selecting them from the drop-down box in the upper left corner of the workbook menu bar (Figure 3-7).

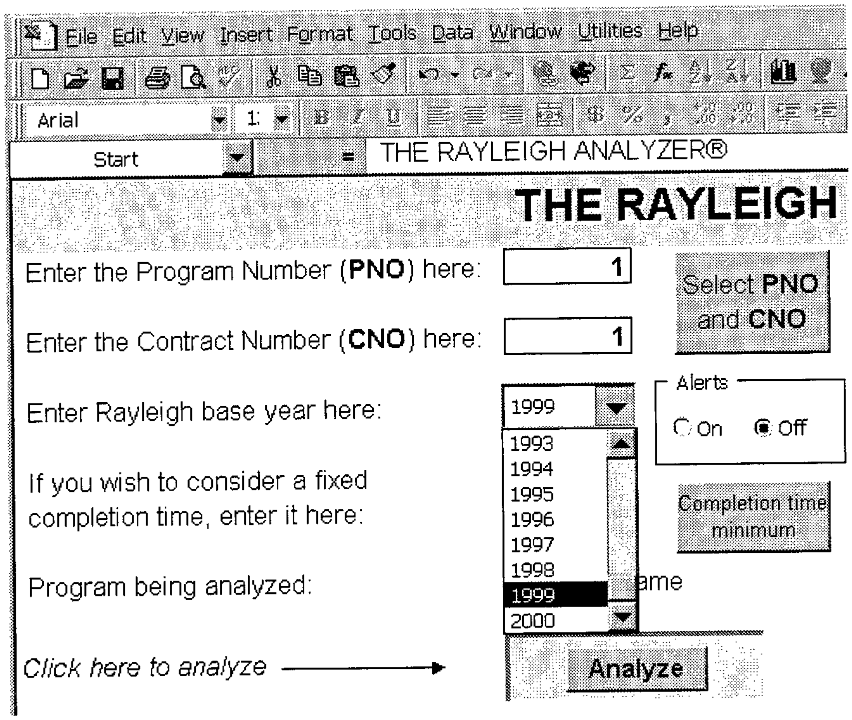
Figure 3-7. Worksheet Navigation Drop-Down Box



CHOOSING A BASE YEAR

The Rayleigh model works with constant or then-year dollars. Computations take place in constant dollars, but are displayed in both constant and then-year terms. Specify your choice for the constant dollar base year by making a selection from the drop-down box on the Start worksheet. The base year need not be the same as the contract's official base year, and has no effect on then-year results.

Figure 3-8. Base Year Drop-Down Box



FIXING THE COMPLETION TIME (OPTIONAL)

If you wish to fix the completion time for a contract being analyzed, type that time (in years—decimals may be used) in the indicated cell (noted by a dotted outline in Figure 3-1, since it is an optional field).

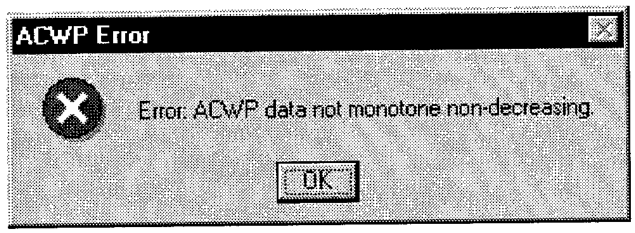
If you enter a number of years that is too small, you will get an error message. Clicking on the button labeled "Completion time minimum" will produce a message box displaying the minimum value of years that can be entered. *That button only provides valid results if it is clicked after entering a fixed completion time and running the model.*

Entering a fixed completion time causes the *Rayleigh Analyzer* to compute total cost statistics with the completion time held constant. Several charts provide fixed-time results, as discussed in a later section.

DISPLAYING ALERTS

There is an "on/off" toggle button you can select to determine if you want to receive "alerts." A series of alerts can notify users of potential data discrepancies and other ambiguities. The alerts are notifications, but do not stop model operation. An example of such an alert is shown in Figure 3-9.

Figure 3-9. Example Alert



We recommend that new users turn on the alerts for an initial period of time to become familiar with the types of alerts their data are causing. However, each alert requires that you manually close a dialog box, so continually displaying them can become tedious. After the initial period of use, you may want to toggle the button off to minimize user interaction when the model is operating.

DATA MANIPULATION GUIDELINES

The previous subsections described the limit of input data and selections for you to enter (except for the Advanced Models, discussed later). There is a line of text beneath the optional fixed completion time cell that may look like it requires input, but it does not. That line reads, "Program being analyzed:"—it is a field that is automatically determined by the system. This will become apparent in a subsequent section of this manual, titled "Example."

Some users may want to view the contracts data that underlie the analysis. The raw data used for generating the lists of PNOs and CNOs can be browsed by activating the Contracts worksheet. If you have defined additional contract data, those data will be visible on the User Defined Contracts worksheet. There is also a 2×2 matrix of valid PNO/CNO combinations on the "Matrix" worksheet. A "1" in a matrix cell indicates a valid PNO/CNO combination, corresponding to the associated row and column headers, respectively. You can quickly navigate to the Start, Contracts, User Defined Contracts, and Matrix worksheets by using the navigation tool shown in Figure 3-7.

On the Contracts, User Defined Contracts, and Matrix worksheets, you may alter the data. That may be useful for users who need to make deliberate changes or modifications to the data. However, all data changes will become permanent if they are saved. If an inadvertent change is made, you can revert to the original files on the Website for reference.

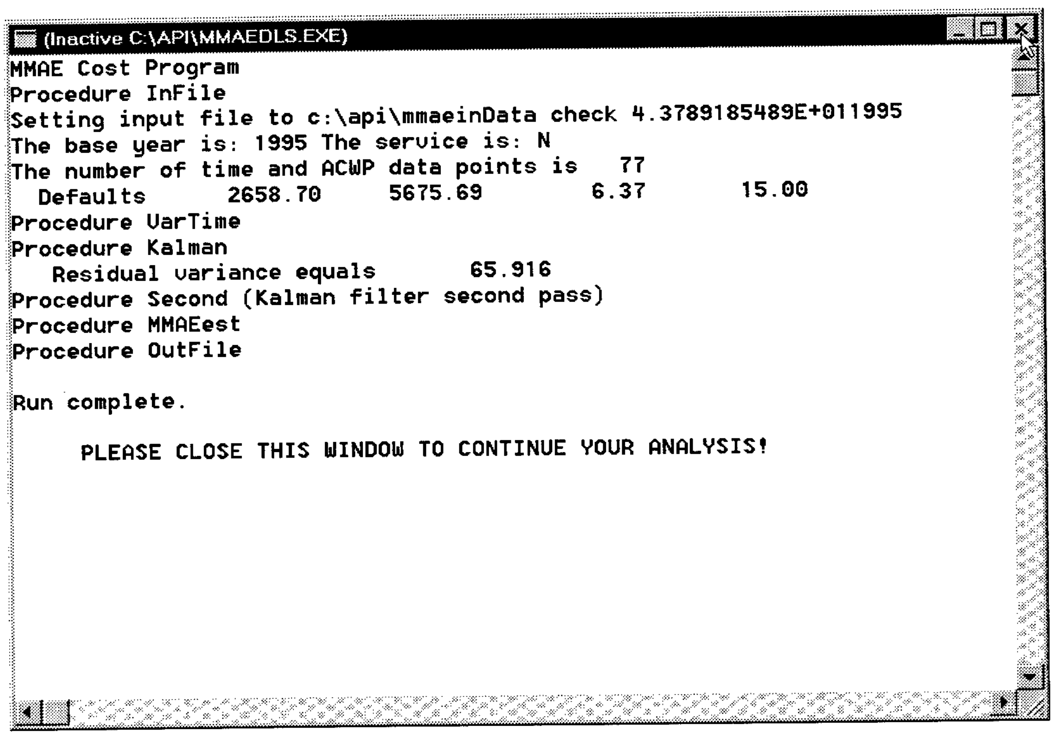
We recommend users only add and alter data on the User Defined Contracts worksheet, unless you are sure you want the changes on other worksheets to become permanent.

Initiating the Analysis

Once your desired parameters are entered on the Start worksheet, begin the analysis by clicking on the macro button Analyze on the Start worksheet. The system begins the analysis.

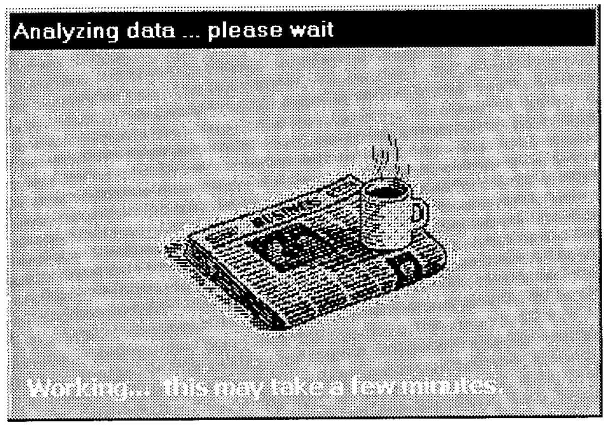
As prompted by a message box, close the "mmaedls.exe" window when that screen displays the "Run complete" message followed by "Please close this window to continue your analysis!" To close the window, click on the close box (×) in the upper right-hand corner, as shown in Figure 3-10. *Do not close the window prematurely.*

Figure 3-10. MMAE Window Needs to be Manually Closed



Next, a dialog box, as shown in Figure 3-11, will appear, asking you to wait while the calculations take place.

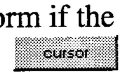
Figure 3-11. Dialog Box During Analysis



At the end of the analysis, the dialog box will disappear and the system will return you to the Start worksheet. *Please do not attempt to close the dialog box until the analysis has completed.*

Model Operation Hints

Several general procedures must be followed when operating the *Rayleigh Analyzer*. The tips below will provide you with some basic steps for ensuring that the model operates smoothly.

- Run the analysis first without choosing a fixed completion time.
- If new PNO/CNO combinations have been added, be sure to click the Update PNO/CNO list button on the User Defined Contracts worksheet.
- If an error occurs while the program is executing and the cursor changes to an hourglass, the cursor should revert back to its default form if the model is run again. As an alternative, you can click on the button 

on the lower left-hand side of the Start worksheet.
- New users may want to run the first few analyses with the "Display Alerts" option toggled "on" to get the feel for the types of errors that may be present in the data. However, the analysis will run more quickly—and with less user interaction—if the alerts are turned off.

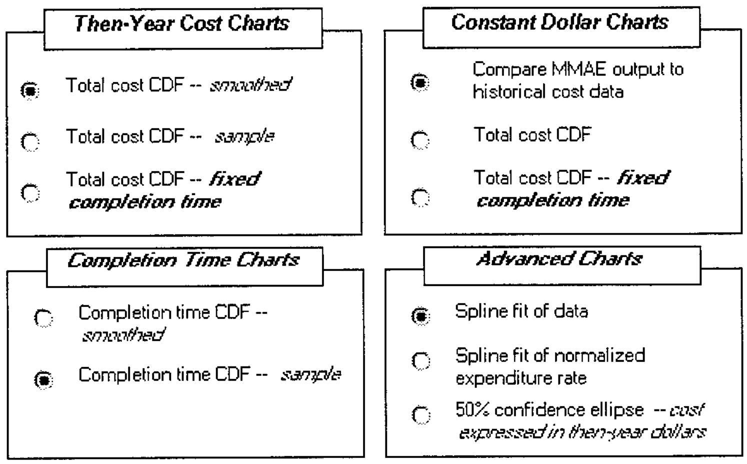
Viewing the Results

In this subsection, we discuss the organization of model results and describe each chart in detail. Advanced charts are discussed in later subsections.

ORGANIZATION OF RESULTS

When the analysis is complete, various graphical results are available. Use the buttons on the right-hand side of the Start worksheet to select individual charts to view, as shown in Figure 3-12.

Figure 3-12. Chart Selection Buttons



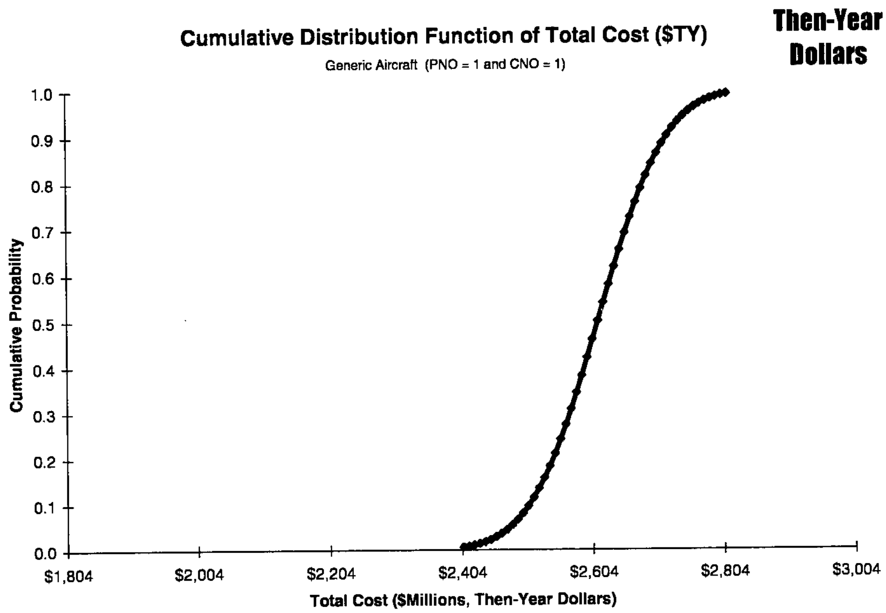
Excel's worksheet tabs also can be used to navigate to each chart, although worksheet names may have slightly different descriptions than those shown in Figure 3-12.

In the next subsections, we describe each chart in detail. The charts are organized by the categories shown in Figure 3-12.

THEN-YEAR COST CHARTS

Clicking on the button labeled "Total cost CDF—smoothed" displays the chart similar to the one shown in Figure 3-13. The chart is of a smoothed, then-year dollar, CDF for the total cost expenditure profile. The smoothed version is constructed on the basis of the assumption that the marginal distribution functions of total cost and completion time are normal. This graph is suitable for a high-level presentation. The smoothing may, however, introduce distortions, and should not be used to obtain actual estimates. The smoothed curve may assign small positive probabilities to values of total cost close to, or even less than, the values of the last data set.

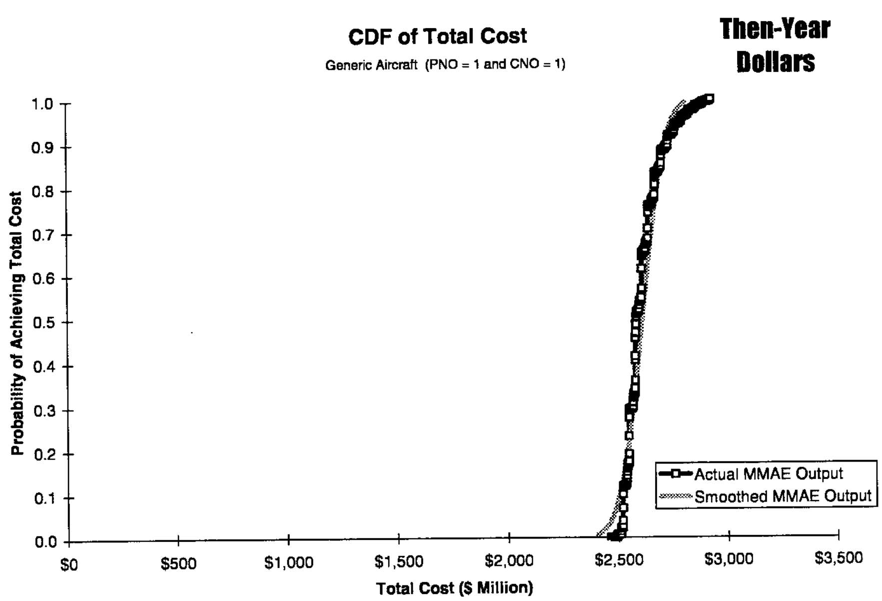
Figure 3-13. CDF of Total Cost, Then-Year, Smoothed



The smoothed graph may be more suitable for high-level presentations. The "raw" version of this graph is discussed next.

Clicking on the button labeled "Total cost CDF—sample" displays a sample marginal CDF of total cost chart, similar to the one shown in Figure 3-14.

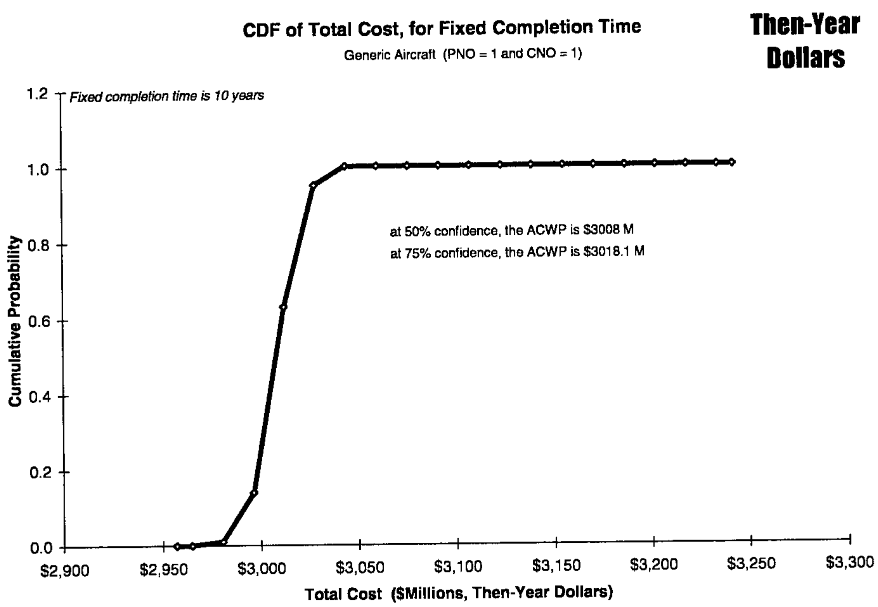
Figure 3-14. CDF of Total Cost, Then-Year, Sample



Because the set of Kalman filters is finite, the sample CDF may have a noticeably "grainy" structure. The smoothed version (shown in Figure 3-13) may be more suitable for presentations than the sample CDF (shown in Figure 3-14), since "jumps" in the sample curve may be distracting and may be difficult to explain in high-level terms. Analysts should use the sample graph in Figure 3-14 for obtaining estimates.

The next button, labeled "Total cost CDF—fixed completion time," takes us to a chart that is blank if a fixed completion time was not entered on the Start worksheet. When you enter a fixed completion time, a chart similar to the one shown in Figure 3-15 is displayed.

Figure 3-15. CDF of Total Cost, Fixed Completion Time, Then-Year



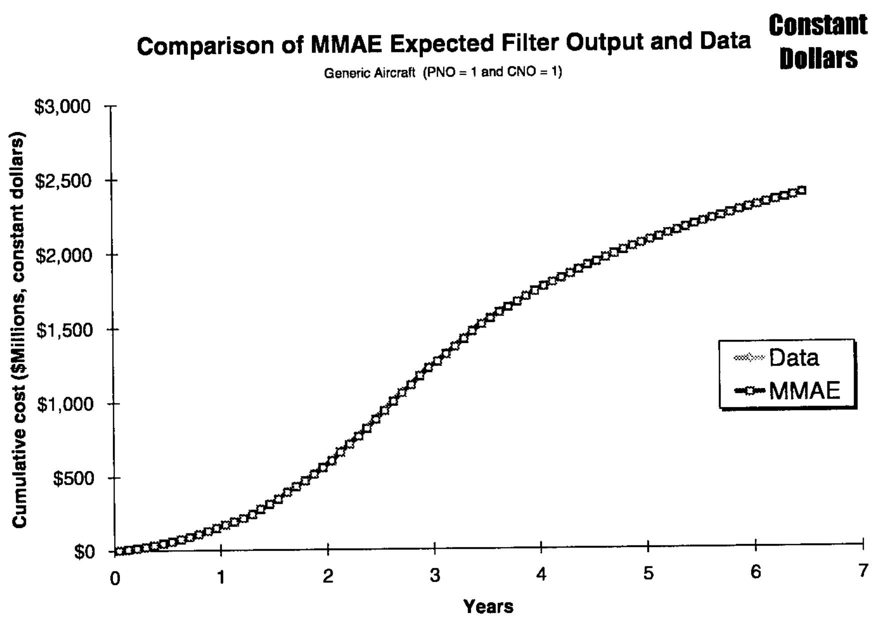
For the example program shown in Figure 3-15 a completion time of ten years was chosen. (Ten years was chosen since later in this subsection we will find the *Rayleigh Analyzer* predicts that the program will be complete in a little more than ten years.)

Constant-Dollar Charts

There are three charts in the constant-dollar group. In this subsection, we briefly discuss each chart.

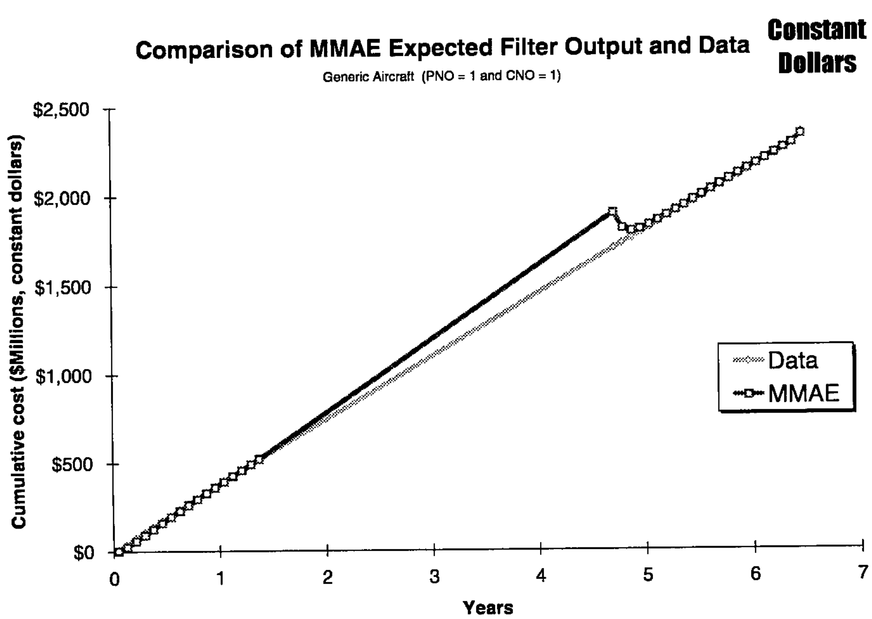
The button labeled "Compare MMAE output to historical cost data" displays a chart showing the ACWP historical data and the MMAE output for the Rayleigh curve (expected values of total cost and completion time). This type of chart, an example of which is shown in Figure 3-16, represents the output of an optimal Kalman filter based upon the Rayleigh model. In the example, two lines are difficult to see, meaning the fit is very good and the development program being modeled is very Rayleigh-like.

Figure 3-16. Comparison of MMAE Fit to Data, Constant Dollars



If a chart shows a separation of the two lines, the data do not exhibit strong Rayleigh characteristics. Figure 3-17 shows an exaggerated diversion between MMAE results and historical data.

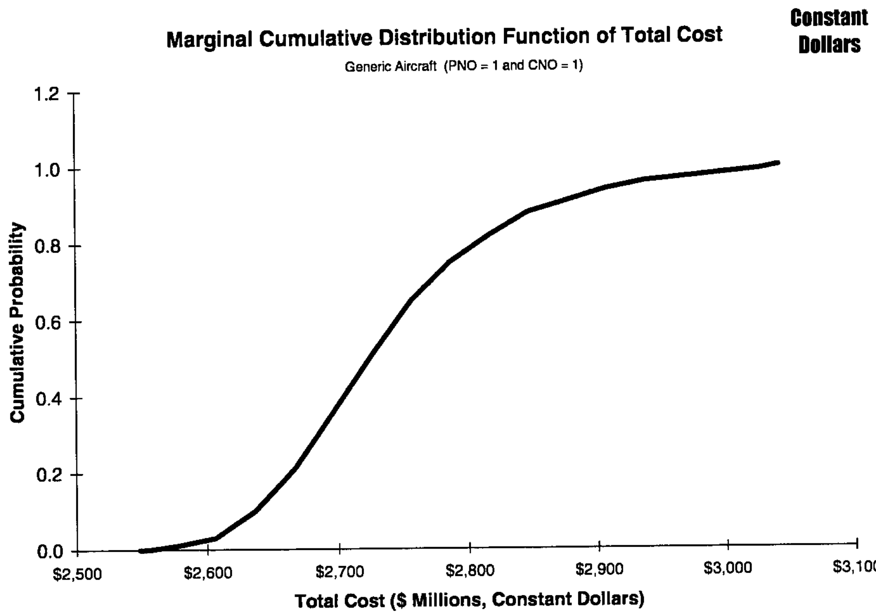
Figure 3-17. Example of Poor MMAE Fit



In such a situation, the analyst might want to investigate deeper into the program to understand what could be causing the behavior. Possibly, an advanced analysis (discussed in the next chapter) may provide insight.

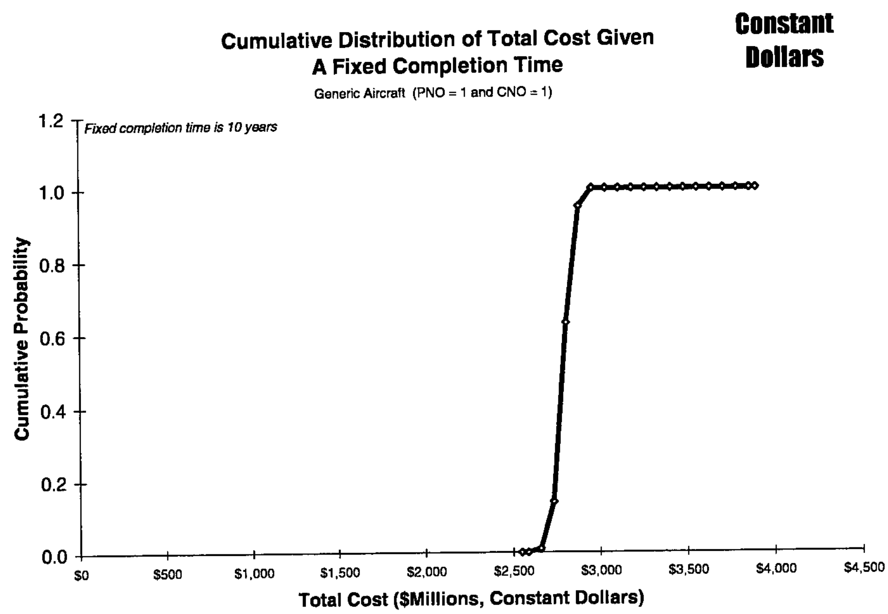
The "Total cost CDF" button, in the constant-dollar chart group, displays the following chart.

Figure 3-18. Marginal CDF of Total Cost, Constant Dollars



Using ten years as the fixed completion time (as we did in the then-year example) produces the third chart of the constant-dollar group, shown in Figure 3-19. The figure shows the CDF of constant-dollar total cost, conditioned on the ACWP data and the specified completion time. The graph would be blank if no fixed completion time was entered on the Start worksheet.

Figure 3-19. CDF of Total Cost, Fixed Completion Time, Constant Dollars

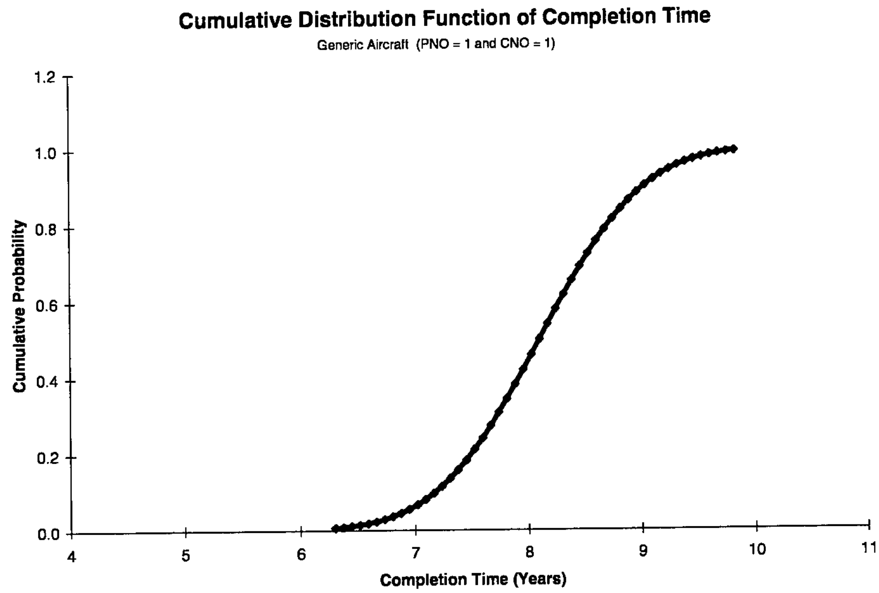


The above constant-dollar results of approximately $2.7 million can be compared to the then-year results of $3.0 million, given in Figure 3-15. This makes sense, since inflation effects are removed from the constant-dollar analysis.

Completion Time Charts

Two completion time charts are available. The first is a smoothed version of the second. The smoothed version is made using the assumption that the marginal distribution functions of total cost and completion time are normal. The smoothed version may be more suitable for presentations than the sample CDF because questions may arise about reasons for particular "jumps" in the sample curve. The smoothing may, however, introduce distortions. The smoothed curve may assign small positive probabilities to values of completion time close to, or even less than, the values of the last data set. Figure 3-20 is an example of a smoothed CDF.

Figure 3-20. CDF of Completion Time, Smoothed



The contract depicted in Figure 3-20 has accumulated more than six years of data, so the minimum completion time is given as approximately 6.25 years. However, the *Rayleigh Analyzer* is estimating the completion time to occur somewhere near the ten-year point (this is the reason we used ten years as the fixed completion time in a previous example).

Figure 3-21 shows the chart, "CDF of Completion Time Sample."

Figure 3-21. CDF of Completion Time, Sample

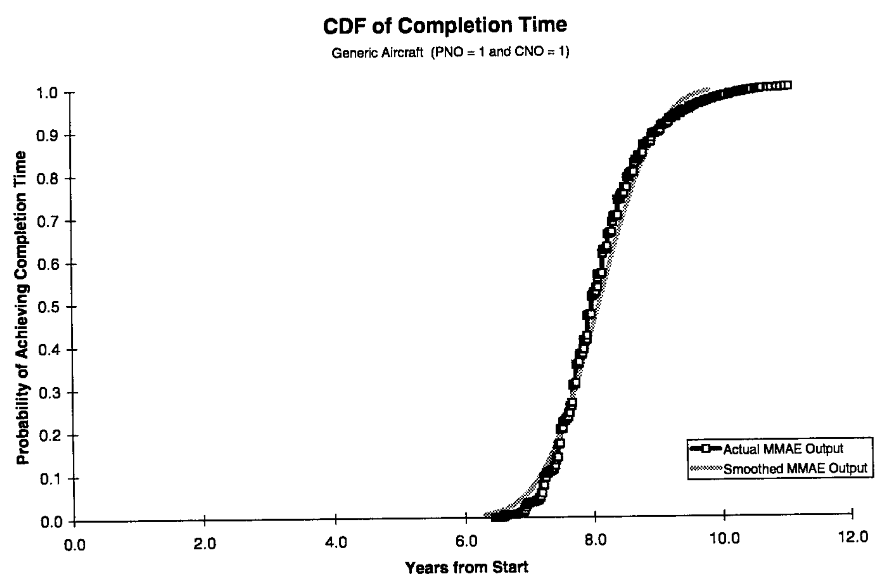


Figure 3-21 is the sample marginal CDFs of completion time. Because the set of Kalman filters is finite, the sample CDF may have a noticeably “grainy” structure. That is the reason it may sometimes be more palatable to show the smoothed output during high-level presentations. However, notice that the smoothed output in Figure 3-20 ends at about the ten-year point, but the sample output in Figure 3-21 tails off somewhere beyond eleven years. The *Rayleigh Analyzer* estimates that there is some small probability that the contract could be completed at that point, but the analyst has a high degree of confidence that ten years would be sufficient.

ADVANCED CHARTS

Three advanced charts are available. Also, three advanced analyses are used to create three additional advanced charts. We recommend experienced users investigate the analyses that are used to develop the advanced charts. Those analyses and charts are described in the next chapter, Advanced Model Operation.

Chapter 4

Advanced Model Operation

In the previous chapter, we discussed the basic operation of the *Rayleigh Analyzer* model. However, sometimes analysts will encounter contract data that do not always lend themselves to straightforward Rayleigh analysis. In this chapter we discuss the advanced features, which are all new to this version of the *Rayleigh Analyzer.*

POLYNOMIAL SPLINES

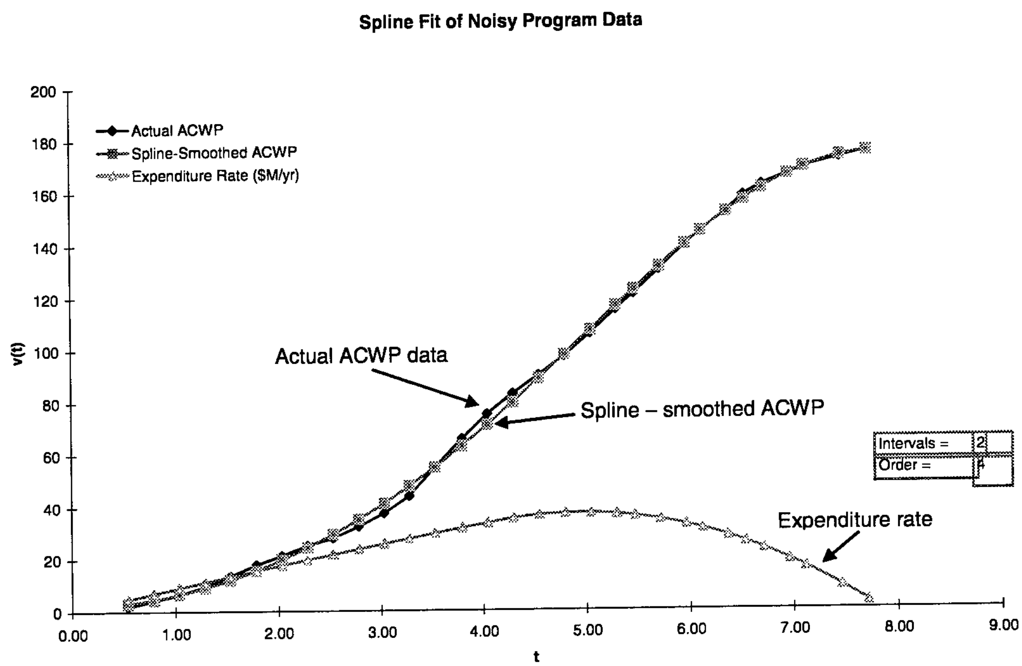
During an analysis of contract data, an analyst may encounter a few cases in which contracts are structured in ways that do not meet with the assumptions for applying the Rayleigh model. Determining if the contract meets the assumptions is difficult solely on the basis of a set of periodic cost reports. It is helpful to fit the data with a *non-parametric* curve that can be used for comparing with a Rayleigh curve and its derivative. A non-parametric curve must be flexible enough to capture the features of the data faithfully, yet stable enough to avoid inserting artifacts such as "phantom" peaks and troughs into the data. In addition, because the cost data may be noisy and contain local oddities, such as accounting adjustments and one-time purchases, curves that can provide "local control"—that is, isolate the effect of any single data point to a small region of the curve—would be beneficial. *Polynomial splines* are curves that have these features.

Polynomial splines are piecewise polynomial functions. A polynomial is an algebraic function that maps the set of real numbers onto itself. The theory behind polynomial splines is discussed in more detail in Volume I, Chapter 4. In this user guide, we will discuss the operation of the polynomial spline feature of the *Rayleigh Analyzer*.

Example of Spline Functionality

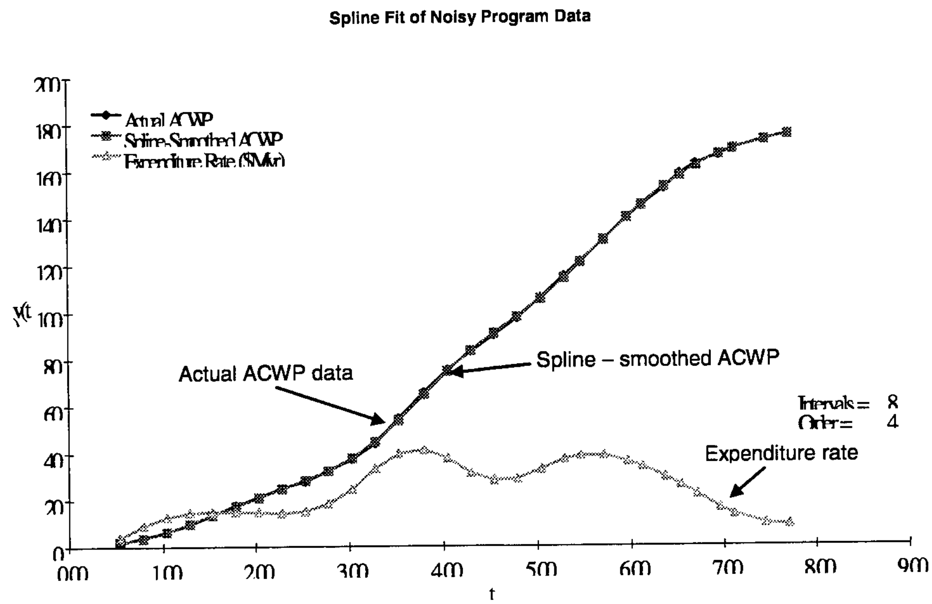
First, we use real program data to demonstrate how the number of intervals influences the quality of the spline fit and the accuracy of its derivative. After going through a brief example of how splines can be used, we discuss how to operate the *Rayleigh Analyzer* to use splines to enhance insight of the program. Figure 4-1 represents a missile program with noisy data and a two-interval spline used to model the data.

Figure 4-1. Spline Fit of Noisy Program Data—Two Intervals



In Figure 4-1, the spline fit uses too few intervals to fit the noisy data. The spline fit does not capture a slight dip in the actual data near the t = 3.00 range, and does not capture a subsequent peak near the t = 4.00 range. These discrepancies do not seem to be too large when looking at the data. However, the derivative of the spline fit (the expenditure rate curve) shows a very gradual, smooth curve with a single peak. This shape differs greatly from the expected shape; the ripples in the actual data imply that even larger ripples should be present in the derivative. Hence, the derivative should display pronounced perturbations at the t = 3.00 and t = 4.00 positions. Two intervals are, therefore, insufficient for capturing the actual features of a data set of this size. In Figure 4-2, we see the effect of increasing the number of intervals to eight.

Figure 4-2. Spline Fit of Noisy Program Data—Eight Intervals



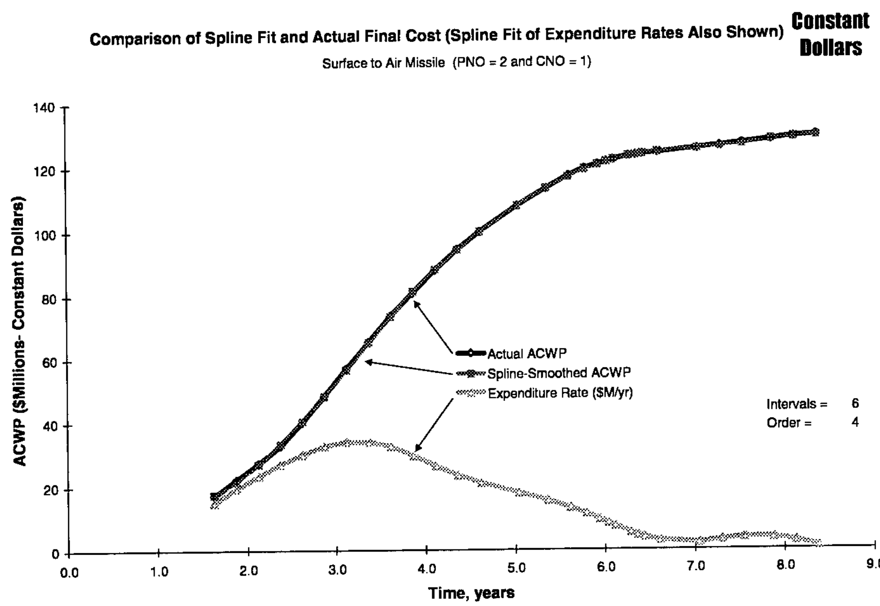
With eight intervals, the spline has sufficient flexibility to closely match the data's local peaks and troughs. As a result, three distinct peaks are evident. (According to the program office for this contract, these peaks match actual events in the development project that caused temporarily accelerated spending.)

Because increasing the intervals from two to eight improves the spline fit and the derivative curve, increasing the number of intervals more is tempting. However, using too many intervals makes the derivative curve much more sensitive to noise. In fact, distinct peaks evident in lower-interval fits can become obscured by noise in high-interval fits. An example of this phenomenon is presented in Volume I, Chapter 4; a mathematically rigorous explanation of this phenomenon is in Appendix C of Volume I.

Using the Spline Tool

The spline analysis is conducted in conjunction with each basic Rayleigh analysis (when the "Analyze" button is clicked). However, most users will not recognize that this is occurring in the background. Also, the number of spline intervals and the order of the polynomial used will be simply those that were used in the previous analysis (the default values are 8 intervals and a fourth order polynomial). In conjunction with any basic analysis, simply click on the "Spline Fit of Data" button, in the advanced charts category, to view the spline smoothed fit of the current data. The result will be a chart similar to the one displayed in Figure 4-3.

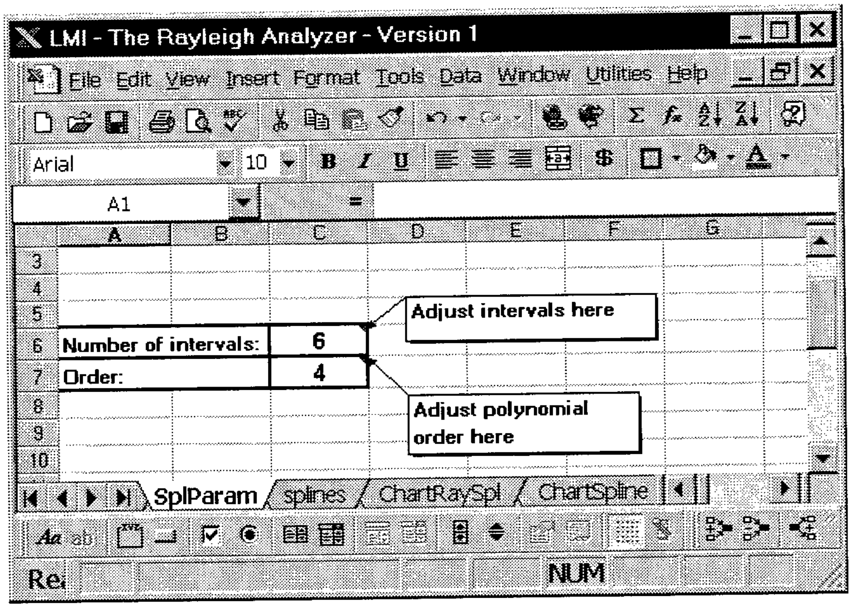
Figure 4-3. Comparison of Spline Fit and Total Cost



The above graph, located on the chart sheet labeled "ChartRaySpl," is the same format as shown in Figure 4-1 and Figure 4-2. In Figure 4-3, six intervals and a fourth-order polynomial were used. The data and calculations used to build this graph are on the worksheet labeled "splines," preceding this chart.

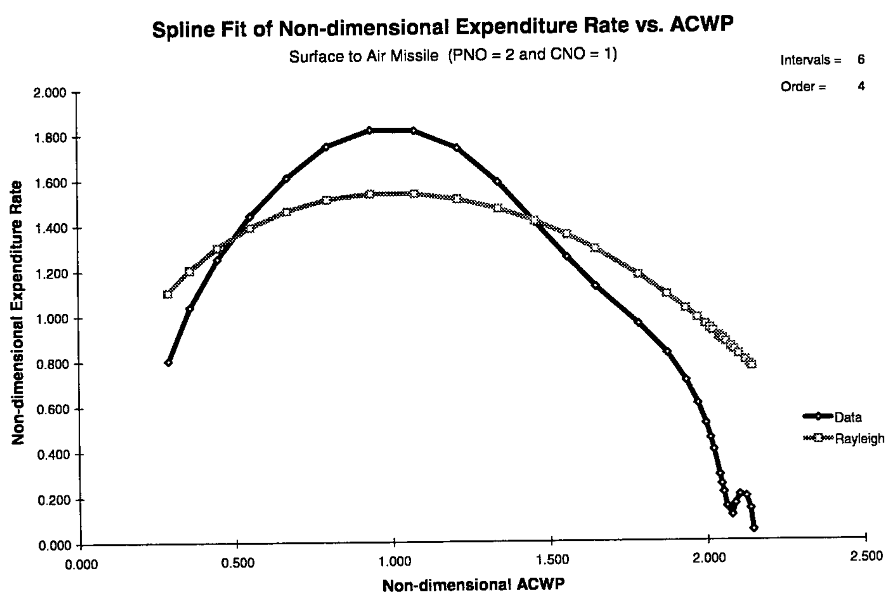
If you are not happy with the display, you can adjust the parameters by navigating to the worksheet labeled "SplParam," preceding the Splines worksheet. On that worksheet, illustrated in Figure 4-4, one cell is for adjusting the number of intervals and one is for adjusting the order of the polynomial. If you decide to adjust the spline parameters, re-running the entire *Rayleigh Analyzer* program is not necessary. If you are simply interested in exploring different Spline parameters, you can go to the Splines worksheet and click on the button labeled "Update Splines" to run the spline-only analysis. (Figure 4-4 shows the four spline worksheets and their relative positions.)

Figure 4-4. Spline Parameter Adjustment



Another spline curve compares the expenditure rates of the spline and the Rayleigh curves. The rates are the derivatives of the respective curves and are computed on a non-dimensionalized (normalized) basis, so that they can be directly compared. Figure 4-5 shows this comparison.

Figure 4-5. Comparison of Non-Dimensional Spline and Rayleigh Derivatives



The chart shown in Figure 4-5 is on the chart sheet labeled "ChartSpline." That chart also can be accessed from the Start worksheet by selecting the advanced chart button labeled "Spline Fit of Normalized Expenditure Rate."

Please keep in mind that this is a very demanding analysis—there are no adjustable parameters, so do not always expect to see close agreement between the spline and Rayleigh derivatives. When considering the goodness of fit using this graph, you should consider only the gross topological features.

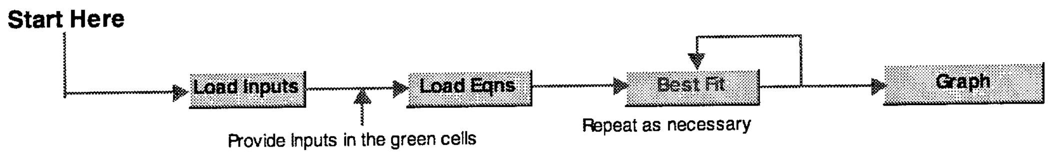
MULTIPLE RAYLEIGH CURVE OPTIMIZERS

This latest version of the *Rayleigh Analyzer* contains three new models that address specific weaknesses of the Rayleigh model by using two or more Rayleigh curves to approximate the data. Each model handles a unique set of circumstances of a development project; you should use only one of the models when the project that you are analyzing fits its specific circumstances. Even when the models are used properly, a single Rayleigh curve may provide a better fit than any of these models for a given development project.

There are three advanced models: a two-Rayleigh composite curve, a three-Rayleigh composite curve, and a modified Rayleigh composite curve. The two-Rayleigh curve handles development projects that have two distinct spending peaks. The two peaks are due to some types of contract expansions, contract restructuring, or other major disturbances in the development project. A three-Rayleigh curve addresses a project that has three distinct activities, all of which can be modeled by a distinct Rayleigh curve: for example, a project with conceptual design, prototyping, and testing phases. Most development projects have at least three such phases; however, sometimes the phases are more pronounced in some projects than in others. The modified Rayleigh curve addresses an expansion in project scope that depends heavily on the work already done on the project. In other words, the expansion is a "hop" from an original Rayleigh curve to a new Rayleigh curve with a larger peak expenditure rate and peak accumulated cost. The circumstances in which this can happen are fairly rare, so extra care must be taken in using this model.

The multiple Rayleigh models work by using Microsoft Excel's internal optimizer to fit the composite curve to ACWP data from a specific project. You can choose the project by entering in the PNO and CNO as you normally would in the Start worksheet. The advanced models do not have a fixed completion time capability, so including a completion time is not necessary. Once the desired program has been entered at the Start worksheet, you then select the desired model from the Advanced Model radio buttons on the Start worksheet. Each model requires that you follow a simple sequence to build results. The button layout of each model's spreadsheet shows this process graphically (Figure 4-6).

Figure 4-6. Multiple Rayleigh Worksheet Button Layout



The first step in running an advanced model is to press the "Load Inputs" button. This button loads the constant-dollar ACWP and time data for the desired project into the advanced model spreadsheet. After this step, you must enter a set of parameters into the green cells at the top-center of the model spreadsheet. These cells may include the expected time of peak spending, total project cost in constant-year dollars, and total time in years. The optimizer accepts these values as initial guesses when fitting the model curve to the data. After you input this information, you must then press the "Load Eqns" button to load the model equations into the spreadsheet. Then you must click the "Best Fit" button to start an optimization. Once the optimization is complete, pressing the "Graph" button will generate a graph of the data, the model curve, and the expenditure rate of the model in then-year dollars. The then-year results are not generated until you press the "Graph" button to reduce computational overhead during optimization.

You should evaluate the results by looking at the least mean squares (LMS) value located near the top of the spreadsheet; the smaller the value, in general, the better the fit. As with any optimization, however, the result could converge to an answer that doesn't make sense. You should, therefore, use care and common sense when interpreting the results. For example, if a two-Rayleigh curve shows start times and spending peaks that do not correspond to those of a major contract event, the optimizer may have converged to the wrong solution. In this event, you should adjust the parameters of the model and try again.

Two-Rayleigh Optimized Curve Fitting

Two-Rayleigh composite curves are used where a development program undergoes some change of contract structure that affects both the spending patterns of the project and the direction of the development. This structure change may be an expansion of the contract that adds a new set of requirements, or a major modification to the existing requirements that forces some change in the project's direction. In response to this change, the spending patterns will exhibit two distinct phases: an initial phase, which follows a Rayleigh curve from the beginning of the contract, and a restructured phase in which the development team reacts to the change by starting a new development effort on top of the initial effort. The two-Rayleigh composite curve models the restructured phase as its own Rayleigh curve on the assumption that the new effort prompted by the restructuring is relatively independent of the initial effort.

This model should be used when you know that a major restructuring has occurred, resulting in new development work that is essentially independent of the work done before the restructuring. If you believe that the new work builds seamlessly on the work completed before the restructuring, then you should use the modified Rayleigh model instead of the two-Rayleigh model. In general, we have found that major restructuring involves a significant amount of work that cannot build on previous work, so the two-Rayleigh model will apply more often than the modified Rayleigh model. To use this model effectively, you should have an approximate idea of when the contract was restructured, and approximately when the peak expenditure rate for that restructuring has or will occur. You do not need to be certain of the times, but knowing this will enable you to intelligently interpret the results of the model optimization.

To use the two-Rayleigh model, you should select a PNO, CNO, and base year on the Start worksheet as if running a normal MMAE model. Rather than pressing the "Analyze" button, you should select the two-Rayleigh worksheet from the Start worksheet radio button menu, and press the "Load Input" button on the two-Raleigh worksheet (Figure 4-6). When this is done the spreadsheet will be ready to begin an analysis of the desired development project.

Once the desired project has been loaded, you must type appropriate values for the project in the green cells at the top of the worksheet. Table 4-1 below shows these cells.

Table 4-1. Adjustable Parameters for the Two-Rayleigh Optimized Model Sheet

	Curve 1	Curve 2
Total Cost (d):	**38.242**	**27.976**
Offset Time:	**0.000**	**1.137**
Peak Time:	**1.322**	**2.568**
α:	0.286	0.244
Rough Est.		
Comp. Time:	**5.000**	
LMS:	0.401	

These cells have the following meanings:

- *Total cost (d)*: the total cumulative cost associated with each curve. The two values will add up to the total cost of the development project in constant-year dollars. This parameter also corresponds to the ***d*** parameter in the Norden-Rayleigh equation

$$R(t) = d(1 - e^{-\alpha t^2})$$

- *Offset time*: the time at which each Rayleigh curve will start in relation to the start of work for the project being analyzed. The first curve will always start at zero; the second curve will start at some time after zero.

- *Peak time*: the time at which each curve reaches its peak expenditure rate in relation to the start of work.

- α: the time constant parameter of the Norden-Rayleigh equation. These values are calculated on the basis of the peak time of the curves, and are provided for your convenience.

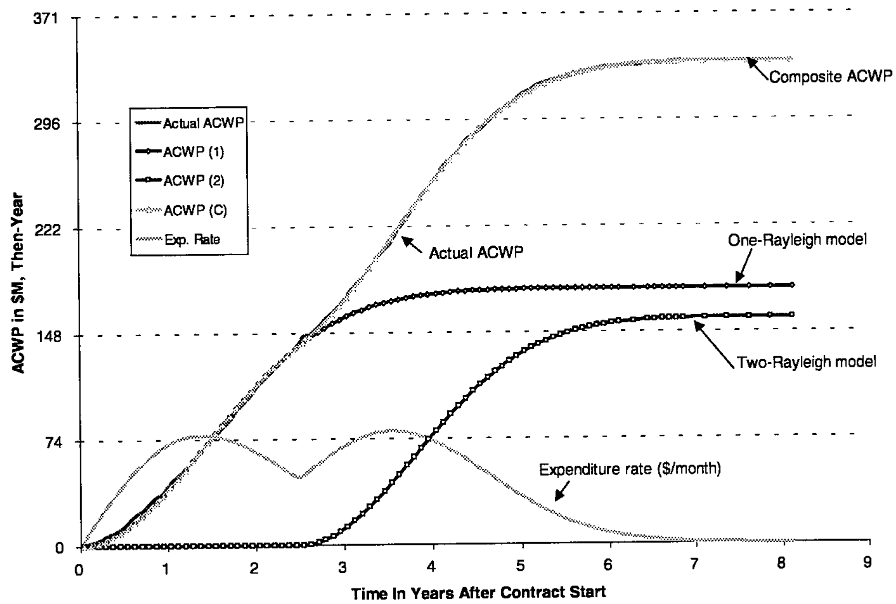
- *Rough estimated completion time*: an estimated completion time for the project expressed in years from the start of work. The spreadsheet uses this value to extend the data points provided by the project's cost reports. The spreadsheet will create a new data point for each quarter year beyond the input completion time to 1.5 times the estimated completion time. For example, if a project has four years of input data and you estimate completion at six years, the spreadsheet will create data points for each quarter between 4.0 and 9.0 years.

- *LMS—least mean squares*: The value represents the average of the least mean square of residuals between the actual cost data from the project and the fit cost data from the composite curve. The optimizer minimizes this value to fit the model to the data. If this value is high relative to the project's cost, you should adjust the parameters listed above and try again.

Once you have adjusted the parameters listed above, the next step is to press the "Best Fit" button to fit the model to the data. The spreadsheet will show the final parameters for each of the two-Rayleigh curves in the cells described above. If the results are valid, then the total cost of the two curves should add to the approximate cost of the project, and the individual costs should make sense considering the financial impact of the project modification. In addition, the second curve's offset (start) time should correspond roughly to the time at which the modification occurred, and the peak times should correspond to the times at which the development efforts reached a maximum. If the data do not reflect this, the optimizer could have found an invalid solution to the problem posed. You should readjust the input parameters and rerun the curve fitting routine as many times as necessary to find a valid solution.

Once you have satisfactory results, you can visualize the results by pressing the "Graph" button. The graphing routine generates the expenditure rates for the optimized curves, converts the curves and the expenditure rates to then-year dollars, and graphs them (Figure 4-7). The graph shows the actual data and the optimized two-Rayleigh composite curve fit of the data. It also shows each of the two curves that sum to the composite curve, and the expenditure rates as calculated using the composite curve. All of these values are in then-year dollars.

Figure 4-7. Results of a Two-Rayleigh Model Analysis



Three-Rayleigh Optimized Curve Fitting

The three-Rayleigh model addresses the case in which a project has three distinct phases—for example, conceptual design, prototyping, and test—that each follow a Rayleigh model. The spending history of the project as a whole, therefore, resembles a composite of the Rayleigh curve used to model each phase rather than a single Rayleigh curve.

You should be aware that the project to be analyzed has three distinct phases explicitly built into the project's schedule. You should also have an idea of the approximate size of each phase in terms of total dollars and the times at which each phase starts and reaches its peak expenditure rate.

To use the three-Rayleigh model, you should select a PNO, CNO, and base year on the Start worksheet as if running a normal MMAE model. Rather than pressing the "Analyze" button, you should select the three-Rayleigh worksheet from the Start worksheet radio button menu, and press the "Load Input" button on the three-Raleigh spreadsheet (Figure 4-6). When this is done, the spreadsheet will be ready to begin an analysis of the desired development project.

Once the desired project has been loaded, you must type appropriate values for the project in the green cells at the top of the worksheet. Table 4-2 shows these cells.

Table 4-2. Adjustable Parameters for the Three-Rayleigh Optimized Model Sheet

	Curve 1	Curve 2	Curve 3
Total Cost (d):	**45.592**	**27.232**	**15.334**
Offset Time:	**0.000**	**1.018**	**1.018**
Peak Time:	**0.828**	**1.536**	**1.544**
α:	0.729	0.212	0.210
Rough Est. Comp. Time:	**3.000**		
LMS:	1.916		

These cells have the following meanings:

- *Total cost (d)*: the total cumulative cost associated with each curve. The three values add up to the total cost of the development project in constant year dollars. This parameter also corresponds to the ***d*** parameter in the Norden-Rayleigh equation

$$R(t) = d(1 - e^{-\alpha t^2})$$

- *Offset time*: the time at which each Rayleigh curve will start in relation to the start of work for the project being analyzed. The first curve will always start at zero; the second curve will start at some time after zero, and the third curve will start at some time after the second curve.

- *Peak time*: the time at which each curve reaches its peak expenditure rate in relation to the start of work.

- α: the time constant parameter of the Norden-Rayleigh equation. These values are calculated on the basis of the peak time of the curves, and are provided for your convenience.

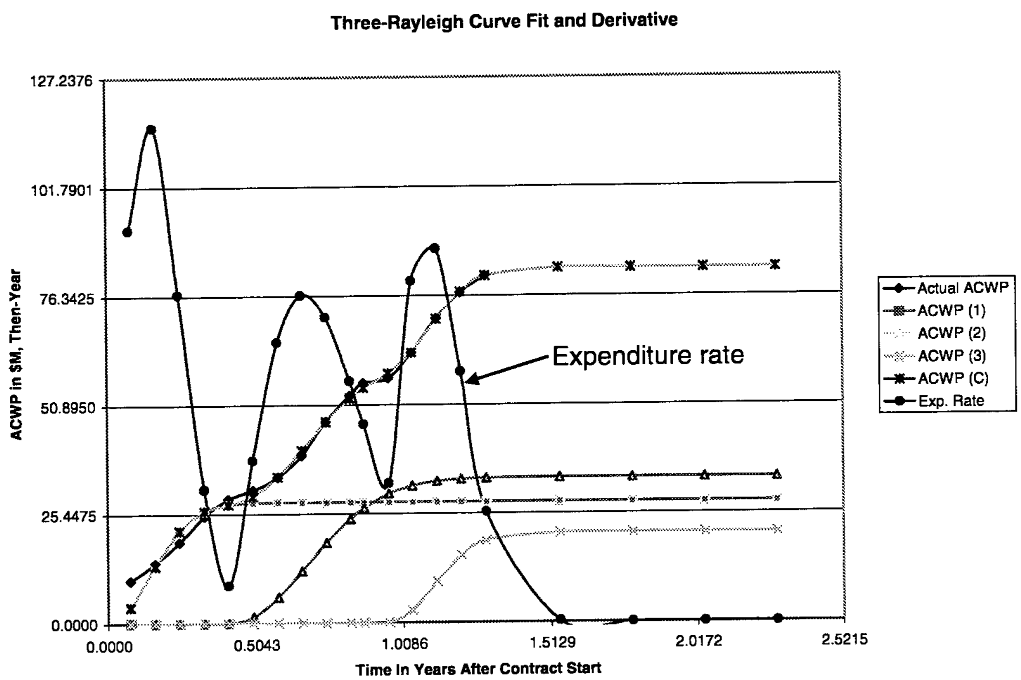
- *Rough estimated completion time*: an estimated completion time for the project expressed in years from the start of work. The spreadsheet uses this value to extend the data points provided by the project's cost reports. The spreadsheet will create a new data point for each quarter year beyond the input completion time to 1.5 times the estimated completion time. For example, if a project has four years of input data and you estimate completion at six years, the spreadsheet will create data points for each quarter between 4.0 and 9.0 years.

- *LMS—least mean squares*: The value represents the average of the least mean square of residuals between the actual cost data from the project and the fit cost data from the composite curve. The optimizer minimizes this value to fit the model to the data. If this value is high relative to the project's cost, you should adjust the parameters listed above and try again.

Once you have adjusted the parameters listed above, the next step is to press the "Best Fit" button to fit the model to the data. The spreadsheet will show the final parameters for each of the three-Rayleigh curves in the cells described above. If the results are valid, then the total cost of the curves should add to the approximate cost of the project, and the individual cost of each curve should correspond with the expected cost and spending history of each phase of the contract. If the data do not reflect this, the optimizer could have found an invalid solution to the problem posed. You should readjust the input parameters and rerun the curve fitting routine as many times as necessary to find a valid solution.

Once you have satisfactory results, you can visualize the results by pressing the "Graph" button. The graphing routine generates the expenditure rates for the optimized curves, converts the curves and the expenditure rates to then-year dollars, and graphs them (Figure 4-8). The graph shows the actual data and the optimized three-Rayleigh composite curve fit of the data. It also shows each of the three curves that sum to the composite curve, and the expenditure rates as calculated using the composite curve. All of these values are in then-year dollars.

Figure 4-8. Results of a Three-Rayleigh Model Analysis



The results shown in the graph above come from a contrived example that makes the three phases of the project more distinct than they might otherwise be. In this case, the expenditure rate curve (the wildly oscillating curve) shows three sharp peaks that correspond to the times at which each of the three phases reach their maximum spending rate. The three phases in both the raw data curve and the composite Rayleigh curve used to fit the data can be seen. Most projects do not show three pronounced phases like this example does, so a graph that does not show these features may still be valid. The key question to ask yourself is: "Do each of the three curves that make up the composite curve start about when they should, and reach their peak expenditure rate about when they should, according to what I know about the project?" If the answer to this question is "yes," then the three-Rayleigh composite model has, in all probability, provided an accurate estimate.

Modified Rayleigh Optimized Curve Fitting

The modified Rayleigh curve addresses projects that have a restructuring event—typically an expansion in requirements—that causes the development team to accelerate their efforts to some degree. If a project follows the modified Rayleigh model, then the spending patterns of that contract follow a Rayleigh curve with a given ending time and cost at first, then transition to a new Rayleigh curve with the same ending time, but with a significantly higher cost. This can happen only when the development project is expanded significantly and the existing development team can immediately begin work at an accelerated rate without needing new ramp-up activities.

This particular model corresponds to circumstances that are more rare than a single Rayleigh curve or a two-Rayleigh curve model. The modified Rayleigh curve best fits projects in which the existing development team is able to rapidly adapt to a significantly accelerated schedule with existing personnel and existing resources. Without significant acceleration in schedule, the program will follow a single Rayleigh curve quite well. With a significant change in personnel or resources, the additional ramp-up effort required to use the assets effectively will give the project a two-Rayleigh characteristic. Because the modified Rayleigh model tackles the none-too-expansive gray area between the one-Rayleigh or two-Rayleigh models, you also should evaluate the project with both of these models when evaluating it with the modified Rayleigh model.

You should be aware of a major restructuring of the contract for the project before choosing the modified Rayleigh model. You also should believe that the additional work tacked onto the contract builds on existing work without requiring significant integration.

To use the modified Rayleigh model, you should select a PNO, CNO, and base year on the Start worksheet as if running a normal MMAE model. Rather than pressing the "Analyze" button, you should select the modified Rayleigh worksheet from the Start worksheet radio button menu, and press the "Load Input" button on the modified Raleigh worksheet (Figure 4-6). When this is done, the spreadsheet will be ready to begin an analysis of the desired development project.

Once the desired project has been loaded, you must type appropriate values for the project in the green cells at the top of the worksheet. Table 4-3 shows these cells.

Table 4-3. Adjustable Parameters for the Modified Rayleigh Optimized Model Sheet

	Starting Curve	Finishing Curve
Total Cost (d):	**184.0084**	**234.3819**
Offset Times:	**0.0000**	1.8000
Peak Time:	**1.4488**	**2.5691**
α:	0.2382	0.8453
Rough Est. Comp. Time:	**6.0000**	
LMS:	12.641	

These cells have the following meanings:

- *Total cost (d)*: the total cumulative cost associated with each curve. This parameter also corresponds to the ***d*** parameter in the Norden-Rayleigh equation

$$R(t) = d(1 - e^{-\alpha t^2})$$

  The total cost for Curve 1 should be the estimated cost of the project before the restructuring occurred. The total cost of the second curve should be the estimated cost of the project after the contract was restructured. For example, if emergent requirements add 50 percent in cost to the original contract, then the total cost of Curve 2 should be 1.5 times that of the original.

- *Offset time*: the time at which each Rayleigh curve will start in relation to the start of work for the project being analyzed. The first curve will always start at zero; the second curve will start at some time after zero, depending on when the contract is restructured. This is *not* the point at which the contract is restructured itself; rather, it is a point that is calculated internally by the spreadsheet, and is provided for your convenience. The transition time itself is not explicitly calculated, but should be apparent from the graph of the model.

- *Peak time*: the time at which each curve reaches its peak expenditure rate in relation to the start of work. We hope that these times will be evident because of your knowledge of the contract or after inspection of the cost data; however, this may not always be true. You can reasonably estimate the first curve's peak expenditure rate time by dividing the expected completion time by three. The second curve's peak time should always be larger than the first curve's peak time.

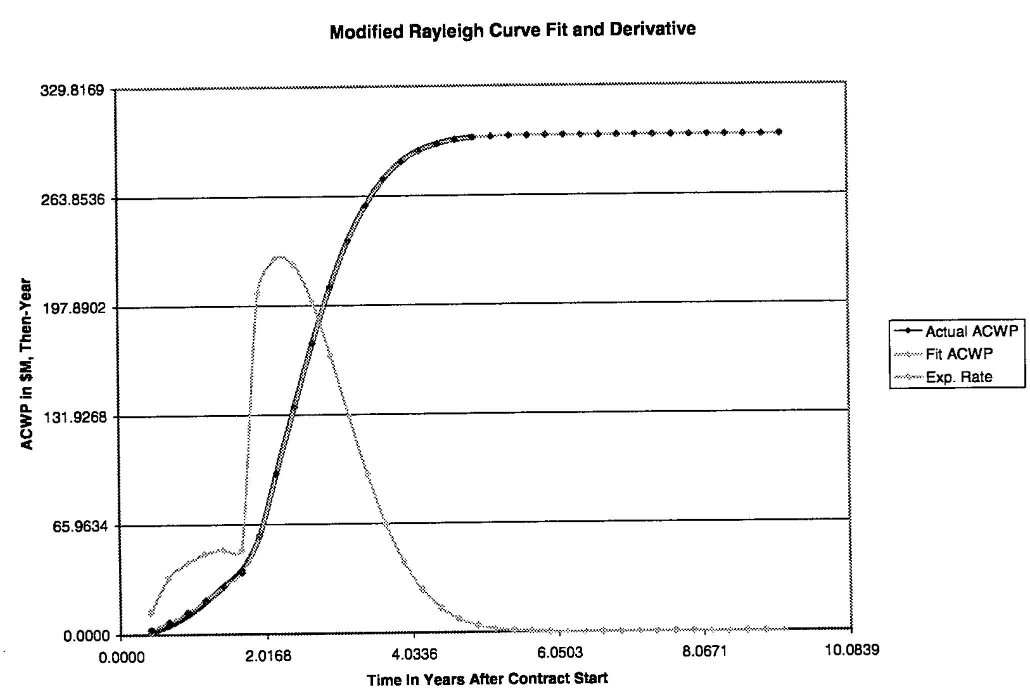
- α: the time constant parameter of the Norden-Rayleigh equation. These values are calculated on the basis of the peak time of the curves, and are provided for your convenience.

- *Rough estimated completion time*: an estimated completion time for the project expressed in years from the start of work. The spreadsheet uses this value to extend the data points provided by the project's cost reports. The spreadsheet will create a new data point for each quarter year beyond the input completion time to 1.5 times the estimated completion time. For example, if a project has four years of input data and you estimate completion at six years, the spreadsheet will create data points for each quarter between 4.0 and 9.0 years.

- *LMS—least mean squares*: The value represents the average of the least mean square of residuals between the actual cost data from the project and the fit cost data from the composite curve. The optimizer minimizes this value to fit the model to the data. If this value is high relative to the project's cost, you should adjust the parameters listed above and try again.

Once you have adjusted the parameters listed above, the next step is to press the "Best Fit" button to fit the model to the data. The spreadsheet will show the final parameters for each modified curve in the cells shown above. If the results are valid, you should be able to see a sharp bend in the graph where the composite curve shifts abruptly from one Rayleigh curve to another (Figure 4-9). This shift should occur early in the data set; changes that occur later in the project will be more likely to require additional development efforts to react to the changes and integrate them into the existing project efforts, and will, therefore, be more likely to follow a two-Rayleigh model.

Figure 4-9. Results of a Modified Rayleigh Model Optimization

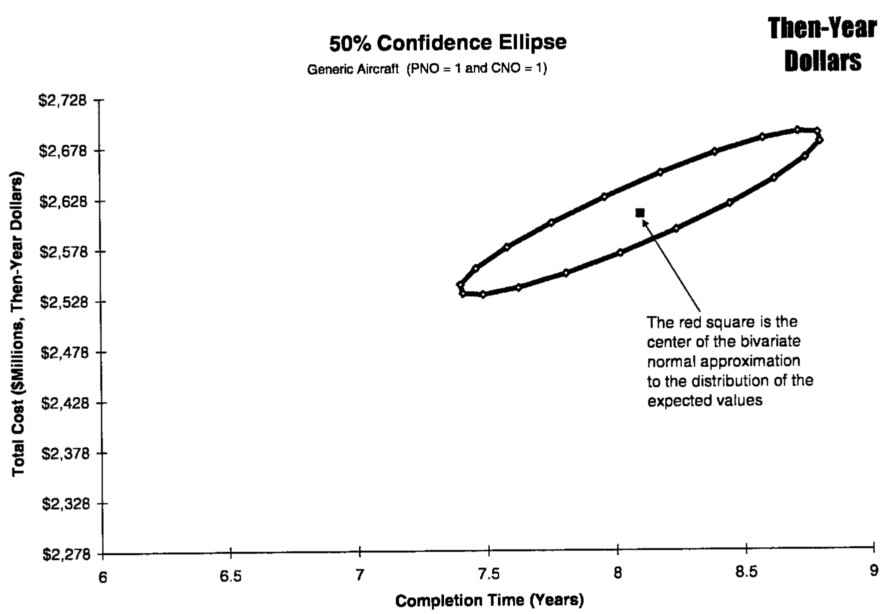


The graph shows the actual data and the optimized modified Rayleigh composite curve fit of the data and the expenditure rates of the project as calculated using the composite-curve. All of these values are in then-year dollars. Notice that the expenditure rate graph shows a sharp bend followed by a radical increase in value at about 1.75 years beyond the start of the contract. This is where the project transitions from the initial set of requirements to an expanded set of requirements. The authors generated this data set as an example of a modified Rayleigh model; the set was contrived so the new curve would show a total cost of three times the first. This example demonstrates two important points: 1) the transition point may not always be clear from the actual data or the model itself, so the expenditure rate curve should be used to determine where it is; and 2) even for this relatively extreme example, the actual data can be fit fairly well by a single Rayleigh curve.

Confidence Limit

Total cost and completion time are correlated random variables. Generally, understanding their relationship is helpful. One way to do this is to view their joint probability distribution. The chart labeled "50% Confidence Ellipse" gives such a view. Figure 4-10 shows an example.

Figure 4-10. Joint Probability Confidence Limit



The ellipse is a constant-probability contour of the bivariate normal approximation to the joint distribution of *completion time* and *total cost.* There is a 50 percent probability that the pair (completion time, total cost) lies inside this ellipse. The expected value of the pair is shown for reference.

The chart shown in Figure 4-10 is accessed by clicking on the advanced chart button labeled "50% confidence ellipse."

UPDATING INFLATION INDICES

The conversion between constant dollars and then-year dollars is accomplished by the conventional method of applying inflation indices. Each of the three military services publishes their own indices. This version of the *Rayleigh Analyzer* uses the indices published in FY1999. For the model to reflect the inflation rates published by each service, the indices in the model need to be updated annually. There are three hidden worksheets containing the inflation indices, one for each service. The World Wide Web locations where the indices were obtained are listed on each worksheet. We do not recommend that users attempt to update the indices, because a linear regression model is applied to them so they can be projected far enough into the future to cover time periods needed for Rayleigh analyses. Our intention is to update the *Rayleigh Analyzer* with the proper inflation indices annually.

Appendix A
Abbreviations

ACWP	Actual cost of work performed
API	Acquisition Program Integration
CAS	Contract Analysis System
CDES	Contract description
CDF	Cumulative distribution function
CNO	Contract number
CPR	Cost performance report
LMS	Least mean squares
MMAE	Multiple Model Adaptive Estimation
PNO	Program number
PPHASE	Program phase
RPD	Report date
WSDATE	Work start date

REPORT DOCUMENTATION PAGE

Form Approved
OPM No.0704-0188

1. AGENCY USE ONLY ***(Leave Blank)***

2. REPORT DATE
Oct 99

3. REPORT TYPE AND DATES COVERED
Final

4. TITLE AND SUBTITLE
The Rayleigh Analyzer®: Volume II—Users Manual

5. FUNDING NUMBERS
C DASW01-99-F-1035
PE 0902198D

6. AUTHOR(S)
John Dukovich
Scott Houser
David A. Lee, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Logistics Management Institute
2000 Corporate Ridge
McLean, VA 22102-7805

8. PERFORMING ORGANIZATION REPORT NUMBER
LMI- AT902C2

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of the Under Secretary of Defense (Acquisition and Technology)
Acquisition Program Integration Directorate
Pentagon
Washington, DC 20301-1800

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
A: Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT ***(Maximum 200 words)***

Volume I of the two-volume series presented theory and applications of the Rayleigh statistical model. Volume II is a users manual to assist in the operation of the Rayleigh Analyzer® model. The Rayleigh Analyzer is meant to be used by cost analysts and project managers in program offices, as well as analysts in central monitoring offices. The Rayleigh Analyzer generates statistics of completion time and final cost for development programs, conditioned on actual cost of work performed (ACWP) data and on the Rayleigh model for the time variation of ACWP in development programs.

The Rayleigh Analyzer is a spreadsheet-based model. It has been available for several years to analysts throughout the U.S. government. The version presented in this manual was improved to incorporate many automation features. More importantly, the model now incorporates several advanced analytical techniques of combining Rayleigh curves to model programs that may not follow the traditional Rayleigh profile.

14. SUBJECT TERMS
Cost analysis, schedule analysis, cost estimating, statistical analysis, research and development, project management

15. NUMBER OF PAGES
50

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL