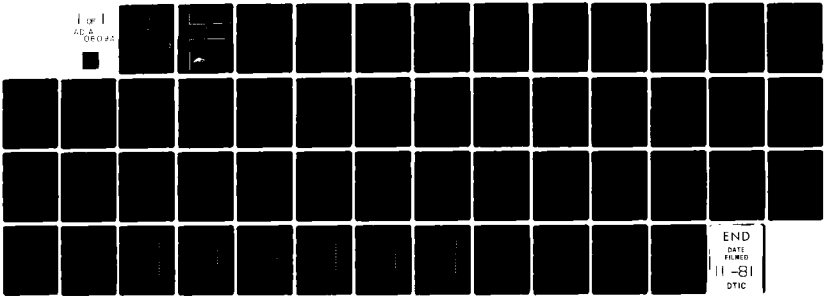
AD-A106 094 RAVEN SYSTEMS AND RESEARCH INC ATLANTA GA F/G 5/2
SOFTWARE MICRO RESOURCE ESTIMATION DATA COLLECTION STUDY.(U)
AUG 79 R C BARRIER, F W CHANG, W L HUGHEY DAAK70-78-D-0052
UNCLASSIFIED RAVEN-TP-2 NL



RESEARCH & DEVELOPMENT DIVISION
TECHNICAL PAPER RAVEN-TP-2
Published August 1979

LEVEL II

Software Micro Resource Estimation Data Collection Study.

by

Robert C. Barrier
Fiona W. Chang
W. Lee Hughey
Gaye B. Stewart

Raven Systems & Research, Inc.

Atlanta, Georgia

TECHNICAL PAPER
RAVEN-TP-2

August 1979

RAVEN
SYSTEMS &
RESEARCH, INC.

Software Micro Resource Estimation Data Collection Study

by

Robert C. Barrier
Fiona W. Chang
W. Lee Hughey
Gaye B. Stewart

Prepared for the U. S. Army
Institute for Research in
Management Information and
Computer Sciences
Contract No. DAAK70-78-D-0052

The views, opinions, and/or findings contained in this report are *those of the authors and should not be construed as an official* Department of the Army position, policy, or decision, unless so designated by other documentation.

ACKNOWLEDGMENTS

Raven Systems and Research, Inc., wishes to express appreciation to all the SGL personnel participating in the interviews. Special thanks are expressed to Colonel Black and Bob Smart for their support, to Gene Schwab and his group for their coordinating efforts and to Jim Beck for his contributions.

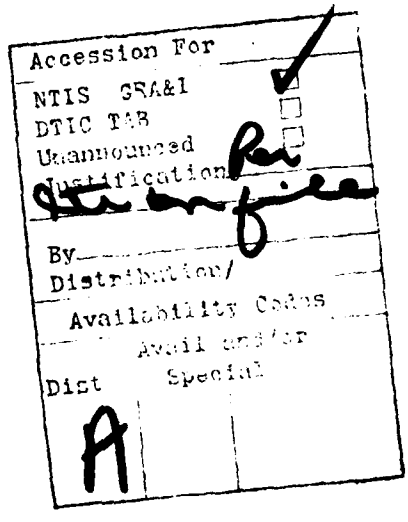


TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

LIST OF TABLES

LIST OF FIGURES

I. BACKGROUND

The U. S. Army Computer Systems Command (USACSC) designs, develops, tests and maintains Standard Army Multicommand Management Information Systems (STAMMIS) which are operative on over 200 computer installations throughout the world. With the standardization of these systems, the USACSC assumed responsibility to support all business applications at each of the Army installations and to be responsive to all user requirements.

Each of the U. S. Army Commands to which USACSC is responsible is known as a Proponent Agency. One such agency is the Army Logistics Command located at Ft. Lee, Virginia. Typical systems supported by CSC personnel at Support Group Lee (SGL) are SAILS (Standard Army Intermediate Level Supply), SAAS (Standard Army Ammunition System), SPS (Standard Port System), and IFS (Integrated Facilities System).

Committed to design and maintenance of each STAMMIS, CSC personnel at Ft. Lee continually modify the systems to meet current user needs. USACSC Regulation 18-21-1 governs the development of these modifications or System Change Packages (SCP's). From its origin with the user in the field to its implementation, the change request is managed very effectively through the use of these regulations. While USACSC-SGL has considerable experience in managing and implementing SCR's, their present technique for estimating the impact of these SCR's is under careful scrutiny.

An accurate estimate of the resources required to implement an SCR is prerequisite to formal modification of any STAMMIS. SGL uses microestimating (USACSC Form 50) for this purpose. While microestimating has worked well in the past, providing useful estimates for overall time,

cost, and resource requirements to implement an SCR, it appeared to need updating. As an outgrowth of this need, the U. S. Army Institute for Research in Management Information and Computer Science (AIRMICS) issued a task to Raven Systems and Research, Inc., to study and document SGL's current estimating procedure. This data collection study is the first phase in the development and automation of an updated resource estimation methodology.

Once Raven has completed the data collection study, personnel in the School of Industrial and Systems Engineering at Georgia Institute of Technology will recommend changes in the present estimating procedure.

II. PROCEDURES

The data collection task was divided into the seven subtasks listed below:

A. In order to develop a sampling design and data collection forms, it was necessary for Raven personnel to become familiar with SGL's operating environment. During this phase, Raven personnel:

   1. Visited SGL and made a preliminary analysis of the operating environment.
   2. Developed a preliminary data collection strategy based on the information obtained. This included development of a procedure for interviewing SGL personnel and the design of data collection forms for recording data.
   3. Coordinated the preliminary data collection strategy with SGL and obtained approval.

B. In order to refine the data collection strategy, Raven conducted a pilot test at Ft. Lee using the preliminary instruments. During this phase, Raven personnel:

1. Visited SGL and conducted interviews with a selected group of programmers, analysts, and managers.
2. Reviewed PMS and other management systems in order to plan for extracting relevant data during subtask 4.
3. Made a first revision of the data collection strategy based on the information gained in the pilot visit.

C. Upon completing all drafting processes, Raven made final editorial revision on the data collection forms and printed them prior to the on-site data collection subtask.

D. The fourth subtask consisted of on-site data collection at SGL. Raven personnel:

1. Conducted structured interviews of selected SGL personnel involved in any one of the following activities:
   (a) Microestimating (USACSC Form 50)
   (b) Developing SCP estimates and workplans
   (c) Using PMS as a management tool.
2. Conducted an in-process review upon returning to Atlanta from Ft. Lee. This IPR was for the purpose of soliciting suggestions for data analysis for subtask 6.

E. Raven then coded and entered the data on the AIRMICS PDP-11/70 using formats developed in cooperation with AIRMICS. This completed the data collection phase of the project.

F. At the end of the data collection phase, Raven analyzed the data collected for trends and other significant characteristics. In addition, a subjective study of the data was performed.

G. The final subtask of the Software Micro Resource Estimation Data Collection Study was the preparation of this final report.

III. DATA SOURCES

The fundamental management unit at SGL is a division, headed by a Division Chief. Each division has responsibility for one or more specific systems. For management purposes, divisions are formally divided into teams and are often informally further divided into groups. A division has 40 to 70 personnel in it, most of whom are programmers and analysts. Subdividing divisions into teams is based on functional areas. In some large systems, for example, teams are organized around cycles. In smaller systems, a team may be responsible for an entire system.

Three manhour accounting systems appeared to offer the most promising data for use in the microestimating study, PMS (Project Management System), REMARCS (Resource Management Accounting, Reporting, and Control System), and SAS (Status Accounting System).

A. PMS

PMS is a computer software management tool designed to aid the project manager in planning and scheduling, resource allocation, and project monitoring. This system, operating from a data base developed by the project manager, has the capability to schedule and allocate resources based or priority, availability, and/or network dependencies.

It has the capability to generate 30 basic reports, classified into 2 broad categories, planning and control.

The data contained in PMS suggests that most of the SGL divisions use it primarily as a manhour accounting system, with few attempts to take advantage of its other capabilities. Raven personnel prepared a report writer program to extract data items of interest and analyzed these data items as part of this study.

B. REMARCS

REMARCS provides the means for recording and reporting the expended manpower resources. The REMARCS system is mandated for use throughout the Computer Systems Command. Theoretically, personnel complete the forms on a daily basis and the forms are approved at the supervisory level. All CSC personnel record man-hours based on work measurement codes which indicate phases of particular jobs or tasks. Use of certain of these work measurement codes may also require entry of specific SCP numbers, EUCP numbers or SCR numbers or DPI code numbers of units. Thus the potential exists for tracking man-hours back to particular change requests. The guidelines for reporting REMARCS require that the hours reported by civilian personnel coincide with the hours reported for pay purposes. Military personnel's hours must be the number associated with the normal work week.

C. SAS

The Status Accounting System (SAS), identified as a possible data collection source of information about SCR's, is maintained by CSC and is a record of all SCR's received by CSC that have not been resolved.

Basic information on the SCR Status Report includes technical/functional designation, criticality category, origination data, problem description, status of the SCR, status date and estimated impact hours, if available.

D. Form 50's

USACSC regulation 18-21-1 requires that an impact estimate, or Form 50, (Figures 1A and 1B) be completed on every SCR not resolved through direct intervention. The proponent agency (PA) logs, reviews and forwards the SCR's to the Application System Developer (ASD) for impacting. The impact estimate is generally completed by the person responsible for the primary program or programs involved, with the summary information transferred to the original SCR.

Completion of the Form 50 results in two key items of information, net development time to complete the SCR, and a breakdown of time into tasks, using standard percentages appearing on the form. The net development time is computed according to a standard procedure, utilizing eight basic pieces of information about the SCR and the system it affects. Some factors included in the formula are function complexity of the affected program, level of effort required, available resources, other systems factors, and non-project factors. Review of the Form 50's provides data on the average size (in terms of effort) of SCR's at Fort Lee, and the accuracy of computations.

All divisions use the Form 50's, or a derivative, for impact estimating, with programmers and/or analysts actually completing the form. Divisions vary in the uses that they make of the Form 50's. Once the estimates are made, some groups basically never refer to them again; others use them for entry of information to PMS and for reviewing actual

IMPACT ESTIMATING

For use of this form, see USACSC-SGL Memo 18-1 (Chap 5); the proponent agency is USACSC Spt Gp, Ft Lee (QAO)

SYSTEM _____ ESTIMATOR'S NAME(S) _____

SCR/PROGRAM _____ DATE _____

This form is to be used for impacting man-days effort required for implementation of the above SCR/program. Standard factors are shown below. This form is to be attached to USACSC Form 6.

SECTION I

	Number	X Factor		
1. INPUT FILE FORMATS AFFECTED BY THIS SCR				
a. Number of card files	_____	X 1 =	_____	
b. Number of tape files	_____	X 1 =	_____	
c. Number of disk files	_____	X 1 =	_____	TOTAL _____
2. OUTPUT FILE FORMATS AFFECTED BY THIS SCR				
a. Number of card files	_____	X 1 =	_____	
b. Number of tape files	_____	X 1 =	_____	
c. Number of disk files	_____	X 1 =	_____	
d. Number of report formats	_____	X 1 =	_____	TOTAL

3. PROGRAM FUNCTIONS: NOTE: This table reflects the number of programs which include functions affected by the SCR, (e.g., An SCR may affect an edit-validation function in each of 3 programs. Two are simple, one is complex. Enter:
a. Edit-validation 4 X 2 = 8 8 X 1 = 8 12 X 0 = 0)

	SIMPLE Factor X Pgms		COMPLEX Factor X Pgms		VERY COMPLEX Factor X Pgms	
a. Edit-validation	4 X ____ =	____	8 X ____ =	____	12 X ____ =	____
b. Sort/merge process	2 X ____ =	____	3 X ____ =	____	4 X ____ =	____
c. Internal data manipulation	2 X ____ =	____	3 X ____ =	____	4 X ____ =	____
d. File search	2 X ____ =	____	3 X ____ =	____	4 X ____ =	____
e. Table look-up (internal or external)	3 X ____ =	____	5 X ____ =	____	7 X ____ =	____
f. Calculations	3 X ____ =	____	5 X ____ =	____	7 X ____ =	____
g. Utilities or subroutines	2 X ____ =	____	3 X ____ =	____	4 X ____ =	____
h. Job Control languages	1 X ____ =	____	2 X ____ =	____	3 X ____ =	____
Subtotals		____		____		____

Total of Program Functions _____

4. RESOURCES AVAILABLE FOR WORK ON THIS SCR

	Number	X Factor	
a. Lead Analyst (GS-13 Equivalent)	_____	X 0.75 =	_____
b. Senior Analyst (GS-12 Equivalent)	_____	X 1.25 =	_____
c. Journeyman Analyst (GS-11 Equivalent)	_____	X 1.75 =	_____
d. Analyst (GS-9 Equivalent)	_____	X 2.25 =	_____
e. Intern Analyst (GS-7 Equivalent)	_____	X 2.75 =	_____
f. Lead Programer (GS-13 Equivalent)	_____	X 0.75 =	_____
g. Senior Programer (GS-12 Equivalent)	_____	X 1.25 =	_____
h. Journeyman Programer (GS-11 Equivalent)	_____	X 1.75 =	_____
i. Programer (GS-9 Equivalent)	_____	X 2.25 =	_____
j. Intern Programer (GS-7 Equivalent)	_____	X 2.75 =	_____
No. people	_____	Sum	_____
Resource average = Sum ÷ Number people =			_____

5. JOB KNOWLEDGE REQUIRED FOR THIS SCR

	FACTOR
a. Limited	0.5
b. General	1.0
c. Detailed	1.5

_____

6. JOB KNOWLEDGE AVAILABLE FOR THIS SCR

	FACTOR
a. Limited	1.5
b. General	1.0
c. Detailed	0.5

_____

USACSC-FT LEE FORM 50
1 MAR 77

FIGURE 1B

7. PROGRAM TURN-AROUND TIME (Average)	FACTOR
a. Effective IAP Usage	0.6
b. More than once per day	0.8
c. Once per day	1.0
d. Less than once per day	1.2

_____

8. SYSTEM FACTOR	FACTOR
a. Developmental	2
b. Major change	3
c. Major modification	4
d. Minor modification	5
e. Maintenance	6
f. Minor technical change	7
g. JCL change only	8

_____

9. DOCUMENTATION CHANGES REQUIRED BY THIS SCR

Number of pages to be changed/added _____

10. COMPUTER TIME REQUIRED

Hours _____ *

SECTION II

NET DEVELOPMENT TIME

1. (1) Input Total _____ + (2) Output Total _____ + (3) Program Function Total _____ = (3a) Sub-Total _____

NOTE: If (3a) is zero, enter one.

2. Total from (3a) _____ X (4) Resources Average _____ X (5) Job Knowledge Required _____ X (6) Job Knowledge Available _____ X (7) Program Turn-around Factor _____ = (7a) Sub-total _____

3. Total from (7a) _____ X 2 ÷ (8) System Factor _____ = (9) Development Time _____ X (10) Other System Factor 1.8 X (11) Non-Project Factor (1.25 + 0.1) = 2.43 (12) Lost Time Factor _____ = (13) Net Development Time _____ man-days

Total of Column 13 is entered onto Line #1 of the SCR Estimate Summary and will be defined as Net Development Time on the SCR Estimate Summary.

SCR ESTIMATE SUMMARY

1. Net Development Time = _____ man days

a. Review and analysis	= NDT X 0.15 = _____	X 8 = _____ *		
b. Design	= NDT X 0.20 = _____	X 8 = _____ *		
c. Programing (including Level I testing)	= NDT X 0.35 = _____	X 8 = _____ *		
d. Testing (including Level II & III testing)	= NDT X 0.25 = _____	X 8 = _____ *		
e. Documentation (enter zero for none)	= NDT X 0.05 = _____	X 8 = _____ *		

2. Total project man-days (sum of 1a-e above) TPMD = _____ man-days X 8 = _____ man-hours

* Enter these figures in the appropriate blocks of the Impact Analysis section of USACSC Form 6, (System Change Request)

Time expended in preparing this estimate: _____ Man-hours

versus estimated resources. Recently, SGL has begun to require a post-SCP review in which each division will be required to compare estimates on Form 50's with actual manhours expended. This will probably extend the use of the form.

IV. PROCEDURES FOR EVALUATION OF DATA SOURCES

A. PMS

Raven evaluated the quality of the PMS data and the applicability of the data to the microestimating study through in-depth discussions with the PMS coordinator, interviews with PMS users, and a review of sample data. Also, Raven personnel conducted an extensive analysis of the PMS Users Guide and reviewed all PMS data forms, specifically the turnaround document.

B. REMARCS

Raven personnel examined the REMARCS data entry forms and the instructions manual, and conferred with personnel in charge of the system at Ft. Lee and Ft. Belvoir.

C. SAS

SAS was evaluated as a data source through reference to specific monthly status reports of System Change Requests and discussions with the SAS manager at Ft. Belvoir.

V. EVALUATION OF DATA SOURCES

A. PMS

The PMS data, while available on personnel from all six divisions, varies both in quality and detail across divisions. The level of usage

of the system for manhour reporting is high, while the sophistication of the users varies significantly and appears fairly low. A review of sample PMS data revealed that man-hour charges are not traceable to SCR's, as was originally suggested when PMS was identified as a data source. In general, time is usually reported by activity to a particular computer program. In theory, the PMS package appears to work very well as a management planning and scheduling tool. However, presently only one project manager is taking advantage of the system's capabilities. One factor contributing to PMS's low level of use may be the time and commitment required to learn to use the system effectively. One estimate was that it may require as much as six months to learn to use the system.

Further, PMS is a batch system written in COBOL that runs on an off-site computer. Usually two to four working days elapse between the time the PMS database is updated and the time the managers receive their reports. This time lapse discourages full use of the system as a planning, scheduling and monitoring tool. Also, since the system is non-interactive, it is not possible to query or update the database easily on an item by item basis. This also discourages aggressive usage of PMS.

B. REMARCS

A preliminary examination of the REMARCS data showed that it is used as a manhour accounting system, with all personnel basically reporting 40 hours per week. Also, the majority of personnel time and activities reported to the system are not keyed to SCR's. The data, as recorded, appeared to have no useful application to the micro-estimating study. In addition, the interviews with key SGL personnel

indicated no faith in the validity of REMARCS data. Finally, even if the data were applicable and valid, there are no effective report generation programs for the REMARCS system.

C. SAS

The information reported on SCR's in the Status Accounting System is not relevant to the present study. There is no history of SCR progress from one status to another. The current status and status date are the only items available which reflect actual SCR progress.

D. Form 50 Estimates

Each division maintains a file on the Form 50's. Raven personnel reviewed numerous forms and found that, in general, the data appeared very useful to the study. The blanks on the forms were all filled out, which suggested that the data was complete and it could be analyzed as originally projected.

VI. INTERVIEW PROCEDURES

The structured interview, chosen as one of the data gathering techniques, allowed for collection of information on the formal processes and interactions occurring among SGL personnel using the microestimating technique. After Raven personnel's initial visit to Ft. Lee, the preliminary interview forms were developed for the pilot test. The purpose of the pilot test was to collect information in order to refine the data collection strategy. The draft instrument contained 50 questions directed toward specific processes, resources and forms used in three general areas: (1) planning and scheduling, (2) resource estimation, and (3) manhour

accounting. The final data collection instruments contained fewer questions and were more general than the original pilot form, but they focused on the same areas.

Pilot interviews were conducted with five SGL personnel, representing two divisions and five different job classifications. The interview forms were then revised based on the results of the pilot interview.

In preparation for the final on-site interviews, Raven staff prepared a stratified random sample of potential interviewees using a random number table. The sample included programmers, analysts, team leaders and division chiefs from six divisions. Table 1 shows the breakdown by category.

TABLE 1

PERSONNEL INTERVIEWS

Classification/Grade*	No. of Persons Interviewed**	Total # of Persons by Classification* (6 divisions)
Supervisory Computer Specialist--GS14	6	6
Supervisory Computer Specialist--GS13	6	22
Computer Specialist--GS12	5	55
Computer Programmer--GS11	5	22
Computer Systems Analyst--GS11	4	16
Total	26	121

* Does not include vacant positions

** Military personnel job responsibilities are equated to civilian classifications

The final interviews, conducted over a one week period, included 26 persons. Raven personnel conducted the individual interviews, which lasted approximately 30 minutes to two hours, emphasizing the purpose of the study and anonymity of the responses. Insofar as possible, interviews in each division were blocked together in order to minimize interference with SGL operations.

Individuals indicated a willingness to share the processes they used to accomplish certain tasks. Basically, during the interview, sample printouts were examined and local forms were discussed to identify their purpose.

VII. TREATMENT OF DATA

A. Interview Data

The original purpose for using the structured interview format were: (1) to collect data on the operating environment in which the micro-estimating procedure is used, (2) to identify the inputs and procedures used to prepare line items on the Form 50's and (3) to document the context in which the manhour accounting system data is reported.

To convert the interview data to a usable form, the following steps were taken:

(1) The responses were converted to table format.

(2) The responses were summarized by job title and by division.

(3) This data was then summarized by job titles and across divisions.

(4) Finally, the data was summarized across job titles and across divisions.

B. Form 50 Data

The original purpose of collecting detailed Form 50 data was two-fold: (1) to determine what Form 50 factors had the highest correlations with actual resources (manhours) expended, and (2) to check the internal validity of the Form 50 itself. It was determined early in the overall data collection effort that it would not be possible to relate Form 50 data to actual resources expended. This is due to the fact that expended resource data for individual SCR's is unavailable once the SCR's are combined into an SCP. However, it was still possible to check the Form 50's for internal validity. In order to do this, the following procedure was implemented:

1. The investigators requested and received permission to examine the Form 50 files of the various SGL divisions. These files contain the original handwritten Form 50's prepared by programmers and/or analysts.
2. Seventy-four Form 50's, representing several different systems, were randomly selected for analysis. Each Form 50 factor was transcribed onto a worksheet, along with SCR identifying information. In the case of the "Program Complexity" and "Resource Available" factors, each subfactor was copied. The computed "Net Development Time" result was also transcribed from each Form 50.
3. Once the staff returned to Atlanta, "Program Complexity" and "Resources Available" factors were recomputed by two different persons who verified each other's work. Then the eight Form 50 factors and the computer results were loaded into the computer system for statistical analysis.

4. Statistical analysis consisted of computing standard descriptive data (mean, median, mode, etc.) on each field of the Form 50. In addition, the "Net Development Time" of each Form 50 was recomputed using the computer, and the recomputed time was compared with the net development time appearing on the Form 50 itself.
5. Since many of the Form 50's were found to contain computational errors, the investigators tested various procedures for simplifying the computation. These tests were made by (a) substituting a constant for one or more of the Form 50 variables, and (b) comparing the newly-computed "Net Development Time" with the correctly computed one. The goal of these experiments was to develop a simpler computational procedure which, if completed correctly, would more nearly match the corrected net development times than the current procedure (including the error rate) does.

C. PMS Data

The Project Management System (PMS) contains thousands of records on activities loaded into the system. Since these activities are identified by the individual project manager, their nomenclature may or may not conform to the standard terminology appearing on the Form 50. Consequently, no attempt was made to classify the activities. Instead, the investigators focused on activity length, resources estimated and expended, lead time, and meeting target dates. To do this, the following steps were carried out.

1. Each activity record on PMS contains 113 data fields. Any one of these may be displayed using the report writer facility of

PMS. However, the PMS report writer is limited to 132 characters per line of output. Consequently, it was necessary to extract the fields of interest (see Table 2) on two separate data files. One of these files contained "date" information, while the other contained "resource" information. These files were created on tape by SGL personnel.

TABLE 2

LIST OF FIELDS EXTRACTED FROM PMS

1. Project Identifier
2. Phase Identifier
3. Activity Identifier
4. Resource (Worker Identifier)
5. Activity Description
6. Setup Date
7. Original Resource Estimate
8. Original Finish Date Estimate
9. Previous Resource Estimate
10. Previous Finish Date Estimate
11. Revised (Last) Resource Estimate
12. Revised (Last) Finish Date Estimate
13. Actual Resources Used
14. Actual Finish Date

2. Although the two files were on tape, they were, in fact, line-by-line images of printouts. Consequently the first step of the PMS data processing was removing the headers, line feeds, etc., from the data. This work, and all subsequent data analysis, was computed on the AIRMICS PDP-11, using FORTRAN.

3. All dates on the PMS data files were then converted on Julian dates so that the number of days between two dates could be computed easily.
4. As a final preprocessing step, the two files were merged for ease of handling.
5. Actual analysis consisted of computation of the arithmetic means, maxima, and minima, of a number of combinations of data fields. Table 6 lists the results of those computations. Once maxima and minima were located, histogram intervals were established and data was prepared for developing the histograms shown as Figures 2 through 9 of this report.

Several comments should be made about data contained in the PMS files:

All data in PMS is self-reported. Consequently it is not subject to rigorous verification. For example, it was reported to the investigators that some workers keep activities "open" (rather than completing them) in order to avoid re-opening them if an unexpected "glitch" appears later.

It should be re-emphasized that PMS activities cannot be related directly to tasks outlined on an impact estimate. PMS activities are defined by programmers and managers to fit their own managerial needs. It is possible, for example, to miss any number of PMS activity targets and still broadcast an SCP on time. Conversely, meeting all PMS targets does not guarantee that broadcast dates will be met. The results given here should be viewed as suggestive rather than conclusive. However, the investigators feel that the results reported here, in general, reflect managerial practices throughout SGL.

VIII. RESULTS

A. Interview Data

An analysis of the interview data revealed several common items worth noting:

1. Form 50's are generally completed by programmers and analysts.
2. Schedules and plans are developed by Team Leaders and Division Chiefs, with very little input from personnel at the GS11 level or below. Tasks completed by lower-level persons are basically date-driven and completed in priority order.
3. The process for determining complexity of functions on the Form 50 was not consistent within or across divisions. However, some divisions do use locally developed Program Complexity Tables.
4. There is substantial variation across divisions in the processes used to complete various line items on the Form 50's.

Following is a summary by job classification of the major points made by the respondents.

Computer Programmers: GS11 (5 persons)

Form 50

-- One person indicated s/he did not complete the form frequently.

-- One person indicated that the time estimate was excessive for short duration tasks.

-- One person indicated that it was possible to work backwards on the Form 50, beginning with the estimate of the number of hours.

-- Three persons indicated that the Form 50 fit their time estimating needs very well, with one person noting a problem with the time-percentage distribution for documentation in the impact summary.

-- Two persons indicated that the Form 50 did not fit their time estimating needs very well. One noted that the time seemed to be off tremendously; while the other felt the form was not applicable to the types of changes made to their particular system.

-- Three of the five indicated that the Form 50 was used for SCRUB sessions; while one indicated that it was used to review resources needed. One person used it to review individual tasks for planning and felt that management used the summary to enter man-hours to PMS.

Planning and Scheduling

-- Basically, the responses from the programmers indicated that their input in setting target dates is through completion of the impact estimate form.

-- Four of the five people interviewed indicated that they did not estimate or plan their time in advance. Timelines were usually established at a higher level and they (the programmers) completed activities in priority order.

-- Four of the five interviewees indicated that dealing with EUCP's impacted time schedules. Other factors given by one or more persons as interfering with deadlines were: meetings, leave, machine downtime, travel to off-site computer, servicing other divisions' problems.

Manhour Reporting

-- Only three of the five interviewees indicated that they reported their time to PMS; while all said that they reported time to REMARCS.

-- One person indicated that they felt either REMARCS or PMS was needed, but not both.

Computer Systems Analyst: GS11 (4 persons)

Form 50

-- All of the Computer Systems Analysts indicated that they prepared Form 50's; with those from one division preparing only the analysis section since contractors serve as programmers.

-- Two of the four analysts indicated that the forms were used for SCP development (SCRUB sessions).

-- Two of the four analysts felt that the time estimates resulting from Form 50's were pretty good.

Planning and Scheduling

-- All of those interviewed indicated that the deadlines were set by someone else. Their input to setting target dates was through completion of the impact estimate.

-- Tasks given by one or more personnel as interfering with time schedules were

  -- Addition of SCR's to an SCP after broadcast date
  -- EUCP's
  -- Customer assistance
  -- Functional guidance unclear
  -- Contractor turnaround time

Manhour Reporting

-- All but one indicated that they reported their time to PMS

-- All stated that they reported their time to REMARCS

Team Leaders: GS13 (6 persons)

Team Leaders

-- All of the team leaders indicated that they did not usually prepare Form 50's; however most indicated that they reviewed the estimates made by others.

-- All of the team leaders indicated that they used Form 50's in negotiations at SCRUB sessions.

-- All of the team leaders indicated a degree of satisfaction with the Form 50 for use at SCRUB sessions. However, two persons indicated that the microestimating procedure is not good for estimating calendar time for tasks that are very short.

Planning and Scheduling

-- All of the team leaders except one indicated that they used a planning and scheduling system, PMS, PERT Networks, etc., for setting deadlines and monitoring progress.

-- Indicated by one or more persons as impacting or distorting the plan were:

  -- EUCP's

  -- Addition of SCR's to SCP after broadcast date

  -- Unclear functional specifications

  -- Field validation test taking longer than originally projected

-- Two of the team leaders indicated that they made daily written reports to their division chief; while one attended daily meetings conducted by the division chief.

Manhour Reporting

-- All team leaders responded that they reported their time to both PMS and REMARCS.

Division Chiefs: GS14 (6 persons)

Form 50's

-- One division chief indicated that SCR's were not impacted until after a pre-SCRUB priority setting session with the Proponent Agency.
-- All division chiefs used the Form 50's at SCRUB sessions.
-- The time lapse between completion of the impact estimate and actual work beginning date was reported as a concern by two division chiefs.
-- All the division chiefs indicated that they reviewed impact estimates.
-- Flaws in the Form 50 identified by one of the division chiefs were:
    -- The form does not have adequate explanation for the data used in the computation.
    -- The form does not have a place to write preliminary directions on how to solve the problem identified by the change.

Planning and Scheduling

-- Some things identified by one or more of the division chiefs as affecting schedules were:

-- Staff turnover

-- Inadequate or misleading functional guidance

-- EUCP's

-- Inadequate or improper testing/testbed

-- Computer availability

Needs

Some needs identified by members of the group were:

-- An estimating technique for developmental projects.

-- Training in the use of microestimation.

-- Profiles for function complexity for each program.

-- A clear process for making composite estimates.

B. Results--Form 50's

Table 3 presents the results of the analysis of the data fields on the 74 selected Form 50's. Several items are notable.

1. Median program turnaround was reported as "less than one per day." This figure reflects very slow turnaround, yet it appears to be in almost universal use at SGL. (Fifty-six percent of the Form 50's examined used it.)
2. The mean Net Development Time (correctly computed) was 19.57 man-days (157 man-hours). This is consistent with a modal system factor of 5 (minor modification). It would appear that SGL's systems are extremely mature, and that significant

system changes are relatively rare. Interviews with SGL personnel tend to support this conclusion. In fact, the largest system change impact estimate in the group of Form 50's selected was less than 750 manhours.

TABLE 3

RESULTS OF ANALYSIS OF DATA FIELDS
APPEARING ON 74 RANDOMLY SELECTED FORM 50's

	Mean	Median	Standard Deviation	Maximum	Minimum
Input File Formats Affected	0.48	0	1.408	8	0
Output File Formats Affected	0.6	0	1.115	5	0
Program Functions (computed)	9.53	4	20.800	150	1
Resources Available	1.91	2	0.309	2.5	1.25
Job Knowledge Required	1.03	1	0.317	1.5	0.5
Job Knowledge Available	1.06	1	0.246	1.5	0.5
Program Turnaround Factor	1.16	1.2	0.282	2.4	0
System Factor	5.4	5	1.325	8	1
Number of Pages to be Changed/Added	8.89	0	30.129	200	0
Computer Time Required	8.35	5	17.036	120	0
Net Development Time (appearing on Form 50)	20.30	15	18.692	79	0.5
Net Development Time (recomputed from Form 50 Data)	19.57	13.6	19.085	98.415	0

3. One of the surprising findings of the study of the Form 50 data was that almost half of the Form 50 estimates contained arithmetic errors resulting in misestimates of eight manhours or more. These errors appeared to derive primarily from two sources: (a) misunderstanding of the "other system," "non-project" and "lost time" factors; and (b) incorrect computation of the "resources available."

After observing the large number of arithmetic errors in the Form 50's, several experiments were performed to see if it would be possible to reduce the computational complexity of the form without materially affecting its results. It was found that both "resources available" and "job knowledge available" could be replaced by constants. This procedure is consistent with interview data. Interviewees stated that estimates were frequently done long before the actual job was accomplished. Therefore, the actual resources that would be available for the job were unknown. As a consequence, many of the interviewees used arbitrarily chosen numbers for the "resources available" and "job knowledge available" fields. Table 4 represents the original estimate, the correctly computed estimate, and the estimate derived by eliminating the "resources available" and "job knowledge available" fields. If the correctly computed estimate is considered to be the "correct" one, the average error of the original estimates is 5.24 days, while the average error of the newly derived estimating procedure is 4.25 days.

C. Results--PMS

For the purpose of analysis, all PMS data was divided into two files. (1) SAILS AB, and (2) all other projects. This was done because the SAILS AB

TABLE 4

FORM 50 RESULTS

Correctly Computed Net Development Time	Net Development Time Computed by SGL Personnel	Error	Net Development Time Using Simplified Algorithm	Error
5.06	5.00	-.06	3.33	-1.73
2.25	3.00	.75	4.45	2.20
14.58	15.00	.42	10.00	-4.58
8.16	9.30	1.14	9.41	1.25
9.19	9.20	.01	21.17	11.98
26.24	47.54	21.30	17.64	-8.60
42.52	42.50	-.02	42.87	.35
29.52	29.50	-.02	17.64	-11.88
25.19	50.00	24.81	25.40	.21
6.30	6.30	.00	6.35	.05
19.68	29.60	9.92	10.58	-9.10
10.21	10.20	-.01	11.76	1.55
19.68	35.43	15.75	12.00	-7.68
9.11	16.40	7.29	7.06	-2.05
15.55	27.99	12.44	15.99	.44
13.61	24.49	10.88	13.99	.38
48.60	17.49	-31.11	24.50	-24.10
5.10	10.56	5.46	6.00	.90
13.61	24.49	10.88	13.99	.38
9.72	17.49	7.77	10.00	.28
8.75	7.00	-1.75	14.11	5.36
3.75	3.50	-.25	3.29	-.46
13.12	13.10	-.02	11.76	-1.36
14.58	26.24	11.66	12.00	-2.58
16.33	16.30	-.03	18.82	2.49
3.40	3.40	.00	3.92	.52
1.25	1.24	-.01	1.65	.40
4.08	4.03	.00	4.70	.62
4.37	7.80	3.43	4.00	-.37
4.37	2.43	-1.94	3.92	-.45
55.40	61.24	5.84	55.86	.46
2.92	2.90	-.02	4.70	1.78
33.85	22.90	-10.95	26.46	-7.39
18.50	19.00	.50	26.46	7.96
49.21	79.00	29.79	44.98	-4.23
69.98	70.00	.02	70.91	.93
9.33	9.30	-.03	9.41	.08
7.00	7.00	.00	7.06	.06
26.24	2.91	-23.33	23.52	-2.72
98.42	75.80	-22.62	88.20	-10.22
7.78	10.70	2.92	10.00	2.22
35.69	27.58	-8.11	39.98	4.29
.61	.50	-.11	1.37	.76
18.37	18.10	-.27	42.34	23.97
38.88	4.40	-34.48	3.92	-34.96
25.52	23.60	-1.92	28.22	2.70
40.40	37.40	-3.00	44.69	4.29
63.79	59.10	-4.69	70.56	6.77

TABLE 4

FORM 50 RESULTS - Continued

Correctly Computed Net Development Time	Net Development Time Computed by SGL Personnel	Error	Net Development Time Using Simplified Algorithm	Error
38.27	38.27	.00	34.30	-3.97
16.45	16.56	.11	23.52	7.07
18.50	18.60	.10	26.46	7.96
6.85	7.30	.45	9.41	2.56
1.28	1.80	.52	1.76	.48
9.37	9.37	.00	8.23	-1.14
10.50	10.50	.00	9.41	-1.09
18.37	18.37	.00	21.17	2.80
19.68	37.18	17.50	12.00	-7.68
2.19	3.93	1.74	2.00	-.19
3.40	2.43	-.97	3.92	.52
.85	.85	.00	2.00	1.15
52.49	52.49	.00	52.92	.43
22.96	23.30	.34	13.72	-9.24
13.12	13.10	-.02	21.17	8.05
52.49	52.30	-.19	56.45	3.96
12.25	12.25	.00	14.11	1.86
31.89	32.00	.11	24.50	-7.39
52.49	32.80	-19.69	47.04	-5.45
28.07	28.07	.00	32.34	4.27
3.94	15.80	11.86	14.11	10.17
5.60	6.00	.40	6.00	.40
14.00	12.00	-2.00	9.80	-4.20
10.21	10.20	-.01	11.76	1.55
10.40	10.90	.50	10.98	.58
2.46	6.20	3.74	2.67	.21
		5.24		4.25

group was identified as being the most effective user of PMS. (Both interview data and PMS data support this conclusion.) The two files were of approximately equal size.

Prior to presenting the results, some explanation of PMS terminology is in order. The investigators examined two types of data fields: (1) dates and (2) resources. As has been discussed previously, dates were converted to Julian style. Resources are measured in man-days. Each resource and date may go through many revisions. PMS retains several of those revised resource estimates and dates in its files. They will be called:

1. Earliest--The very first estimate made of the target date or resource for a particular activity.
2. Previous--The second-most-recent estimate of the date or resource.
3. Last--The most recent estimate of the date or resource.
4. Actual--The actual date on which the activity was completed, or the actual man-days used on the activity.

On examination of the data, it was found that only a relatively small number of activity estimates progressed through as many as two revisions. As a result, only 27 percent of the records had a previous finish date or previous resource estimate. Table 5 shows this graphically.

TABLE 5

SUMMARY OF PMS RECORDS WITH REVISED ESTIMATES

	Number of Records	Number of Records with Previous Finish Date/ Resource Estimate
Total File	1879	514
SAILS AB	821	498
All Others	1058	16

It will be noted that of the 514 records with more than one revision (thus having a previous finish date/resource estimate) 498 (96.8%) were in the SAILS AB file. This suggests that the SAILS AB team is using the PMS system more aggressively as a management tool than other units.

Since "previous" estimates were confined almost entirely to SAILS AB, they were eliminated from further consideration. This left three fields: (1) the earliest estimate, (2) the late estimate, and (3) the actual figure. From this data, the investigators attempted to answer two questions:

A. Do managers consistently either overestimate or underestimate target dates or resource requirements?

B. Are the late estimates (closer to the actual dates) any better than the earlier ones?

Figures 2 through 9 attempt to illustrate the answers to these questions. Figures 2 and 3 show early and last completion date estimates for SAILS AB activities compared with the actual completion date. It will be noted that in the early estimate the modal error is from 1 to 20 days late. (The mean error is 14.52 days late.) Revision, however (Figure 3), leads to a significant increase in the accuracy of target date estimates. At this point over half of the SAILS AB estimates are exactly on target (zero days error). The mean error of the last estimate is only 3.92 days late, a substantial improvement over the earliest estimate. It should be noted, however, that both estimates are somewhat low.

Figures 4 and 5 make the same comparison for all other units. Figure 4 shows that there is not a well-defined curve of estimate errors, although the modal error in estimate between the early completion date estimate and

the actual completion date is 1-20 days late. The mean completion date error is 43.85 days late.

Like SAILS AB, the last estimate of the completion date for all other units shows a marked improvement (Figure 5). However, the mean completion date estimate is still 17.77 days late.

In summary, both SAILS AB and all other units improve their target date estimates over time. However, SAILS AB reduces its error by a factor of 3.7, while the other units reduce their error only by a factor of 2.4. In addition, SAILS AB's first estimate shows less error than all other units' last estimates. Both SAILS AB and other units tend to be optimistic about target dates.

Figures 6 and 7 illustrate the early and late resource estimate errors of SAILS AB. The modal early estimate error (Figure 6) is to underestimate resources required by 1 to 19 days. The mean early estimate is 3.02 mandays less than the actual resources expended.

Figure 7 shows the increase in the accuracy of the resource estimate for SAILS AB at last estimate time. The modal estimate error is now zero, and the mean estimate error is 3.83 mandays less than the actual resources expended.

Although the mean resource estimate error increases from the early estimate to the last estimate in SAILS AB, the dispersion of the error is reduced substantially. SAILS AB tends to consistently underestimate the resources required to perform a given task. Mean actual resources needed to complete a task for SAILS AB are 22.43 man-days. The mean early underestimate of resources required is 13.5 percent. The mean late underestimate of resources required is 17.1 percent.

Figures 8 and 9 illustrate early and late resource estimate errors for all other units. While the modal error of the early estimate (Figure 8) is zero days, the mean error is an overestimate of 4.52 days. This pattern is essentially unchanged in late estimate, Figure 9. The mean error of the early estimate has changed to 5.06 days overestimate.

In summary, both SAILS AB and other units appear to do a reasonably accurate job of forecasting resources needed to complete a specific activity. All errors were on the order of $\pm$ 15 percent. Neither group showed substantial improvement in resource estimates over time.

IX. CONCLUSIONS AND RECOMMENDATIONS

~~Although~~ Raven's task was primarily one of data collection, several observations regarding the data are in order.

1. Virtually all lower-level people felt that they had little or no input into the timeline setting process. This may place them in a date-driven mode where meeting target dates may take precedence over quality control.
2. Of all the SGL Form 50's examined, none had estimates in excess of 1000 manhours. In talking with another agency (ALMSA) Raven discovered that they did not use microestimating on jobs of less than 1000 manhours. If 1000 manhours were used as a cutoff point, only a very small percentage of SGL's system change requests would ever be impacted using microestimating.
3. The computational error rate on the Form 50's is remarkable. It is recommended that two steps be taken to improve this error rate.

a. Give higher management visibility to Form 50's. While some of the errors were subtle, many were obvious on even cursory examination. For example, many of the errors consisted of using 4.38 for a system factor instead of 2.43. This nearly doubles the estimated manhours.

b. Redesign the Form 50 to make the computations easier. Raven has developed a redesigned Form which appears as Appendix A of this paper. Hopefully, the new Form will reduce computational errors.

4. A frequent complaint heard at SGL was that the person doing the work did not have access (due to transfers, resignations, etc.) to the person completing the Form 50. Consequently, the basis for the original estimate was unclear. It is recommended that the newly-designed Form 50 have a section for a narrative of the change in order to provide historical data on what was intended.

5. One of the surprises of this study was that, although PMS and REMARCS regularly collect manpower data, SGL currently has no program for collecting manpower data which can be traced back to SCR's. Therefore, there is no method of actually checking to see if Form 50 estimates are accurate. It is recommended that SGL, in cooperation with AIRMICS, develop and implement a data collection strategy aimed at verifying resource estimation procedures.

6. While PMS is an extremely powerful management tool, its slow turnaround time severely limits its usefulness. It is recommended

that the Decision Support System, currently under development by AIRMICS, incorporate a scheduling/resource allocation program (such as PMS) which will allow managers to schedule and plan in an interactive environment.

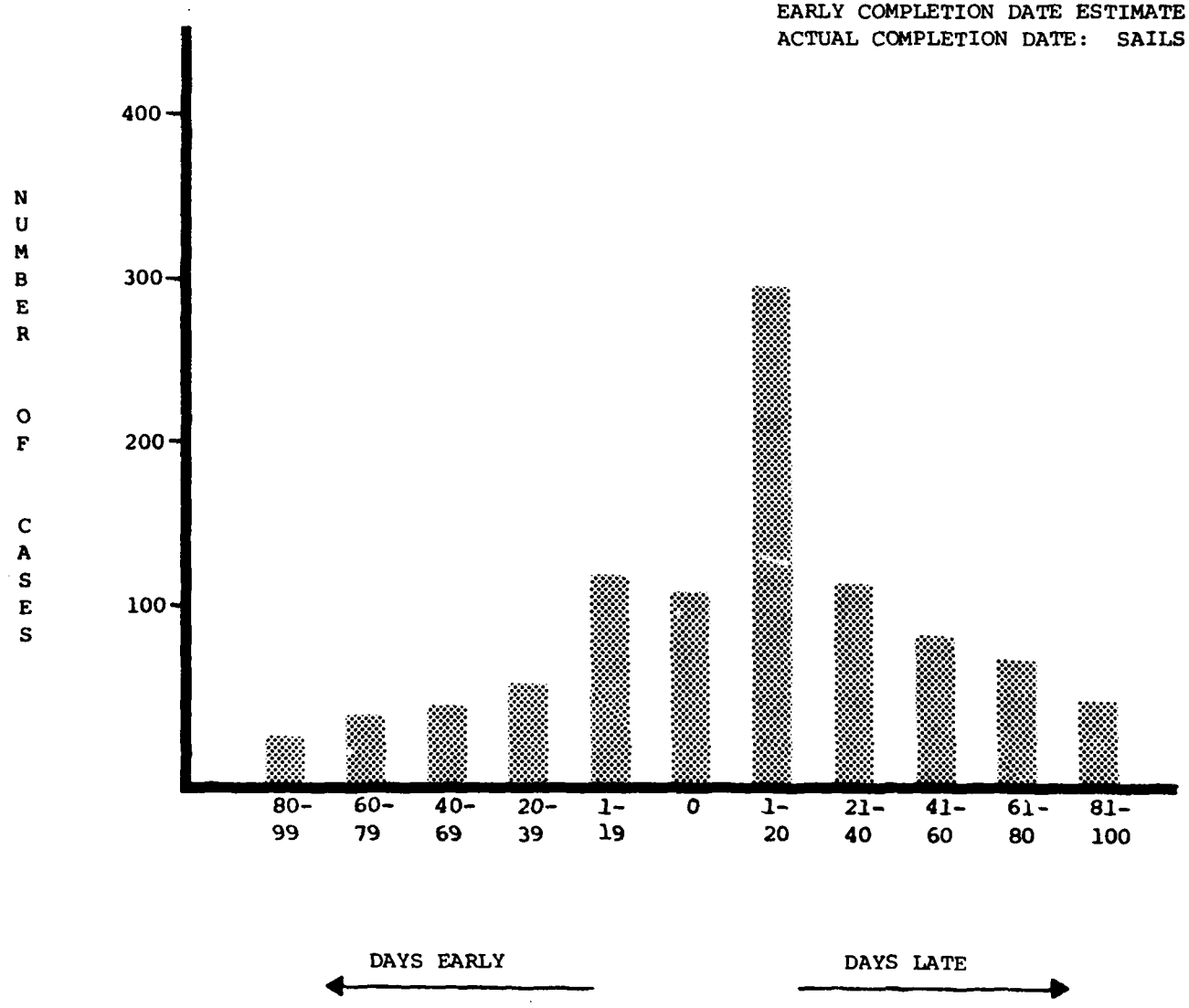

EARLY COMPLETION DATE ESTIMATE VS.
ACTUAL COMPLETION DATE: SAILS AB
NUMBER OF CASES
400
300
200
100
80-99
60-79
40-69
20-39
1-19
0
1-20
21-40
41-60
61-80
81-100
DAYS EARLY
DAYS LATE

FIGURE 2

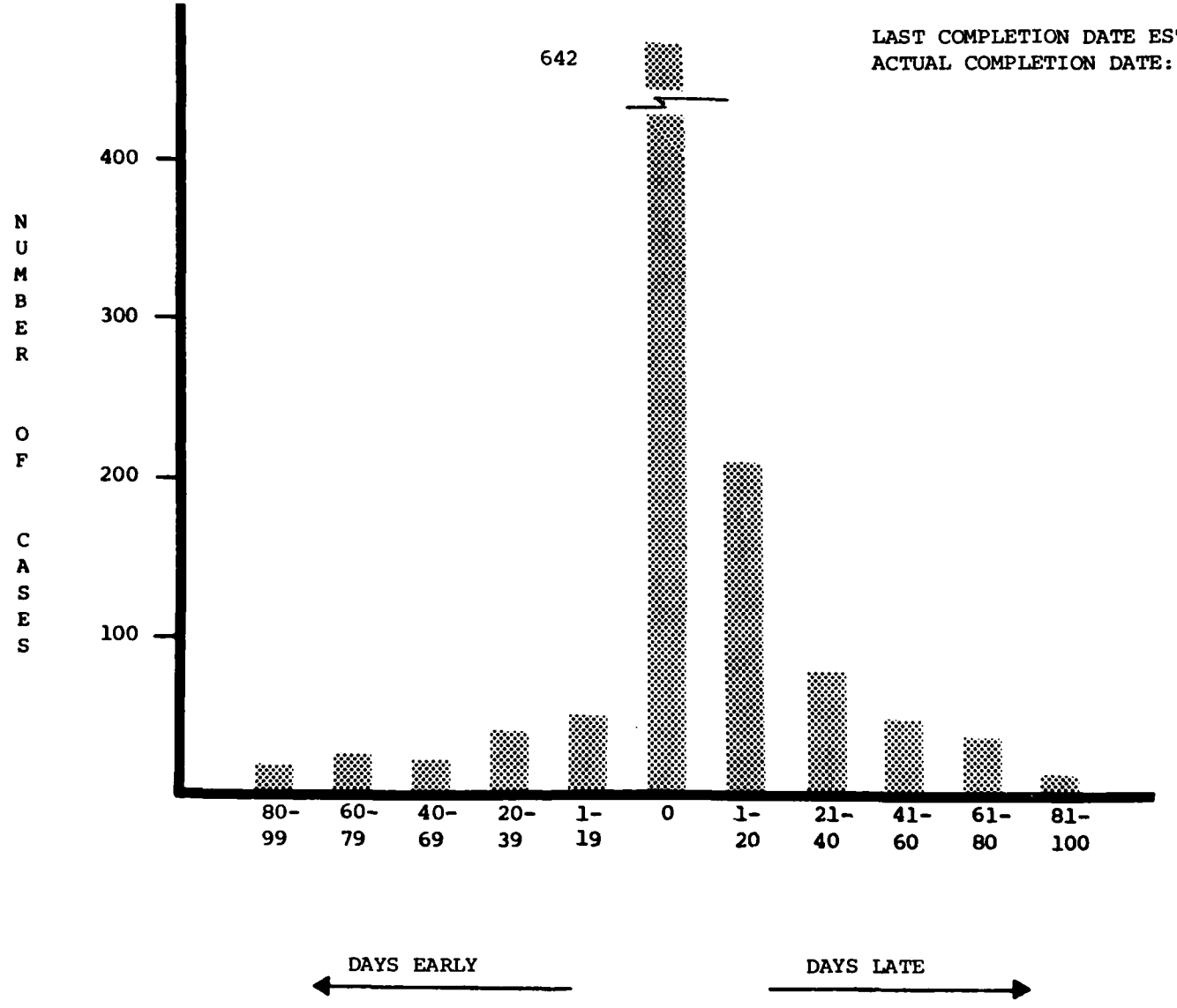

LAST COMPLETION DATE ESTIMATE VS.
ACTUAL COMPLETION DATE: SAILS AB
642
400
300
200
100
NUMBER OF CASES
80-99
60-79
40-69
20-39
1-19
0
1-20
21-40
41-60
61-80
81-100
DAYS EARLY
DAYS LATE

FIGURE 3

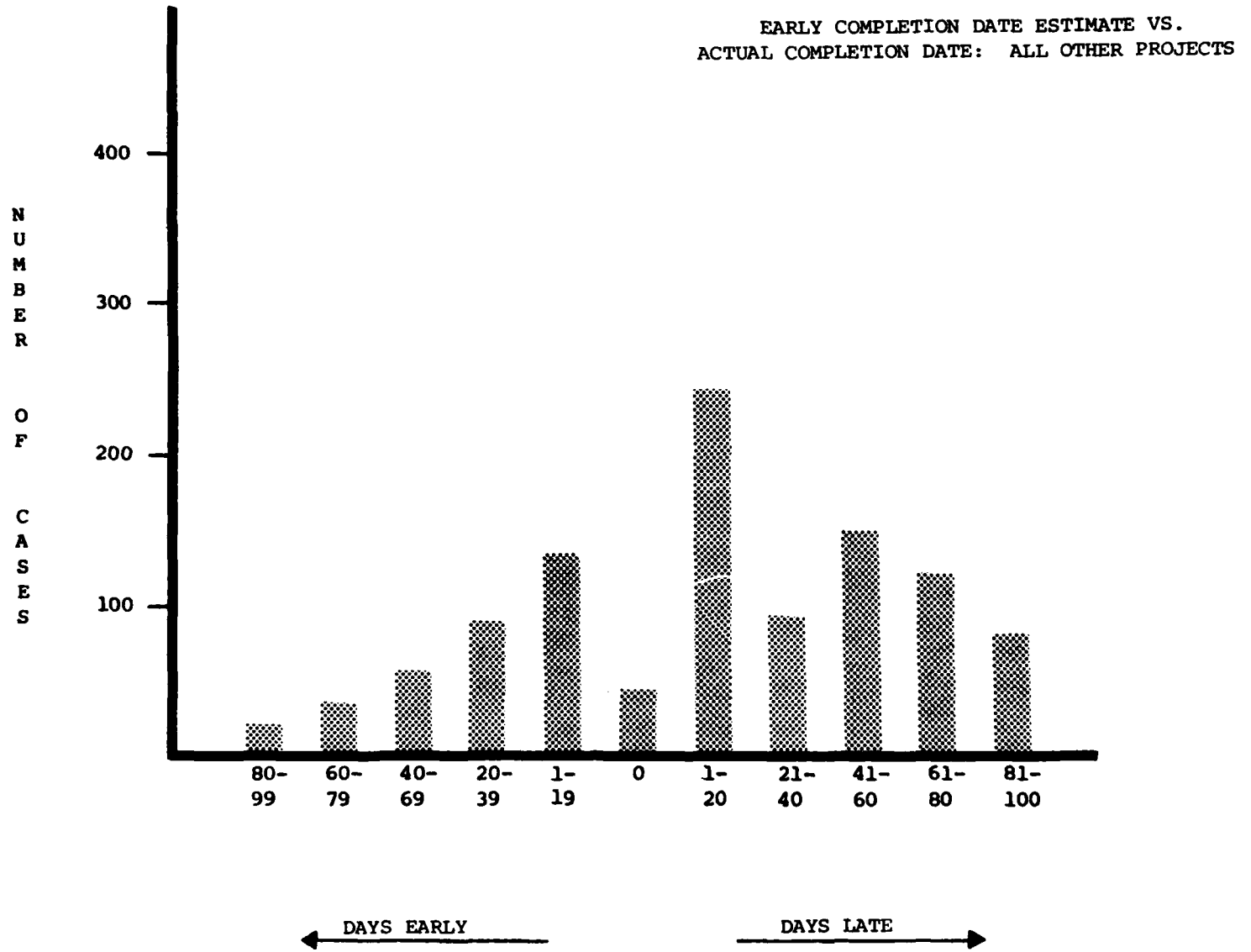

EARLY COMPLETION DATE ESTIMATE VS.
ACTUAL COMPLETION DATE: ALL OTHER PROJECTS
NUMBER OF CASES
400
300
200
100
80-99
60-79
40-69
20-39
1-19
0
1-20
21-40
41-60
61-80
81-100
DAYS EARLY
DAYS LATE

FIGURE 4

LAST COMPLETION DATE ESTIMATE VS.
ACTUAL COMPLETION DATE: ALL OTHER PROJECTS

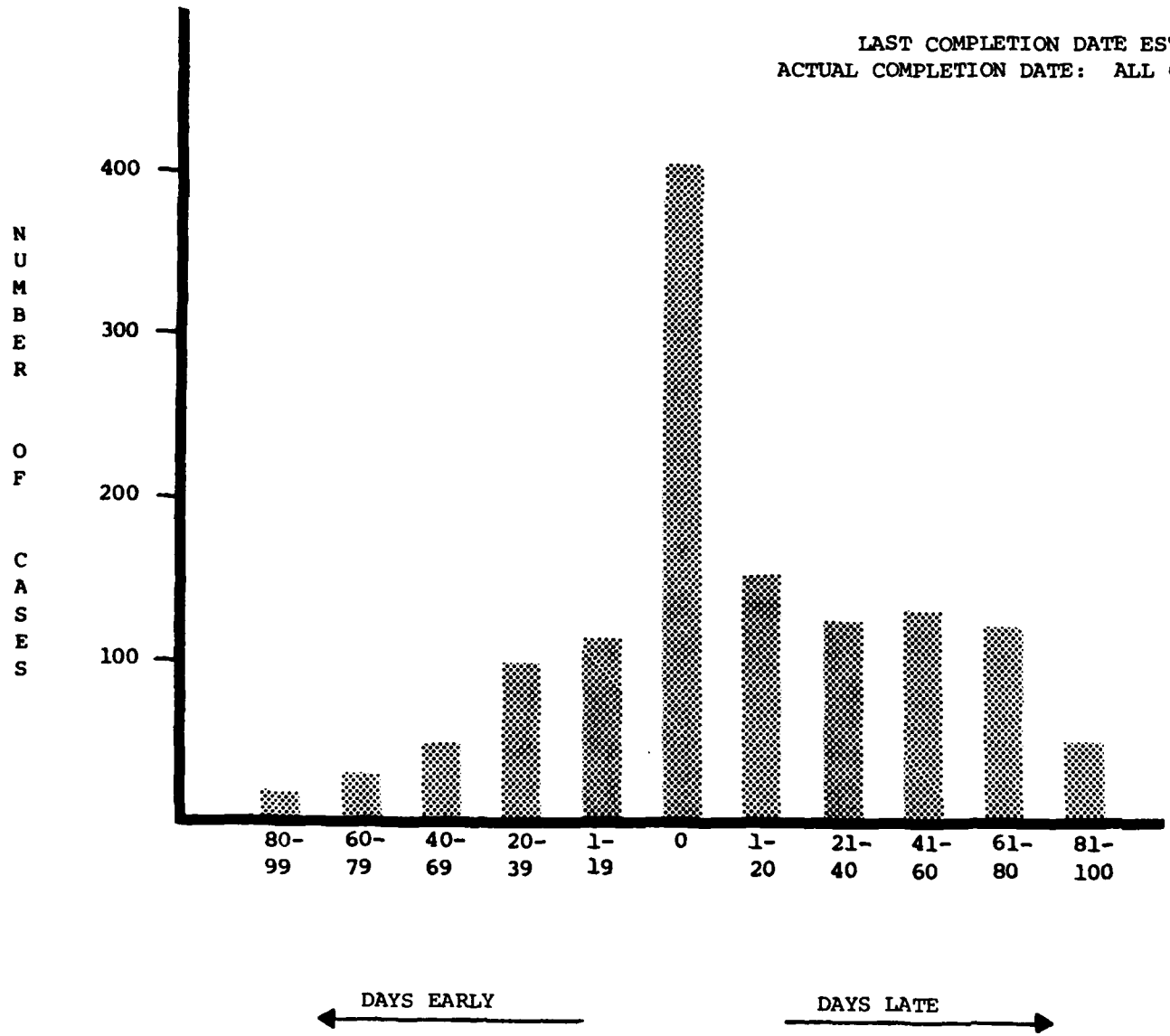


FIGURE 5

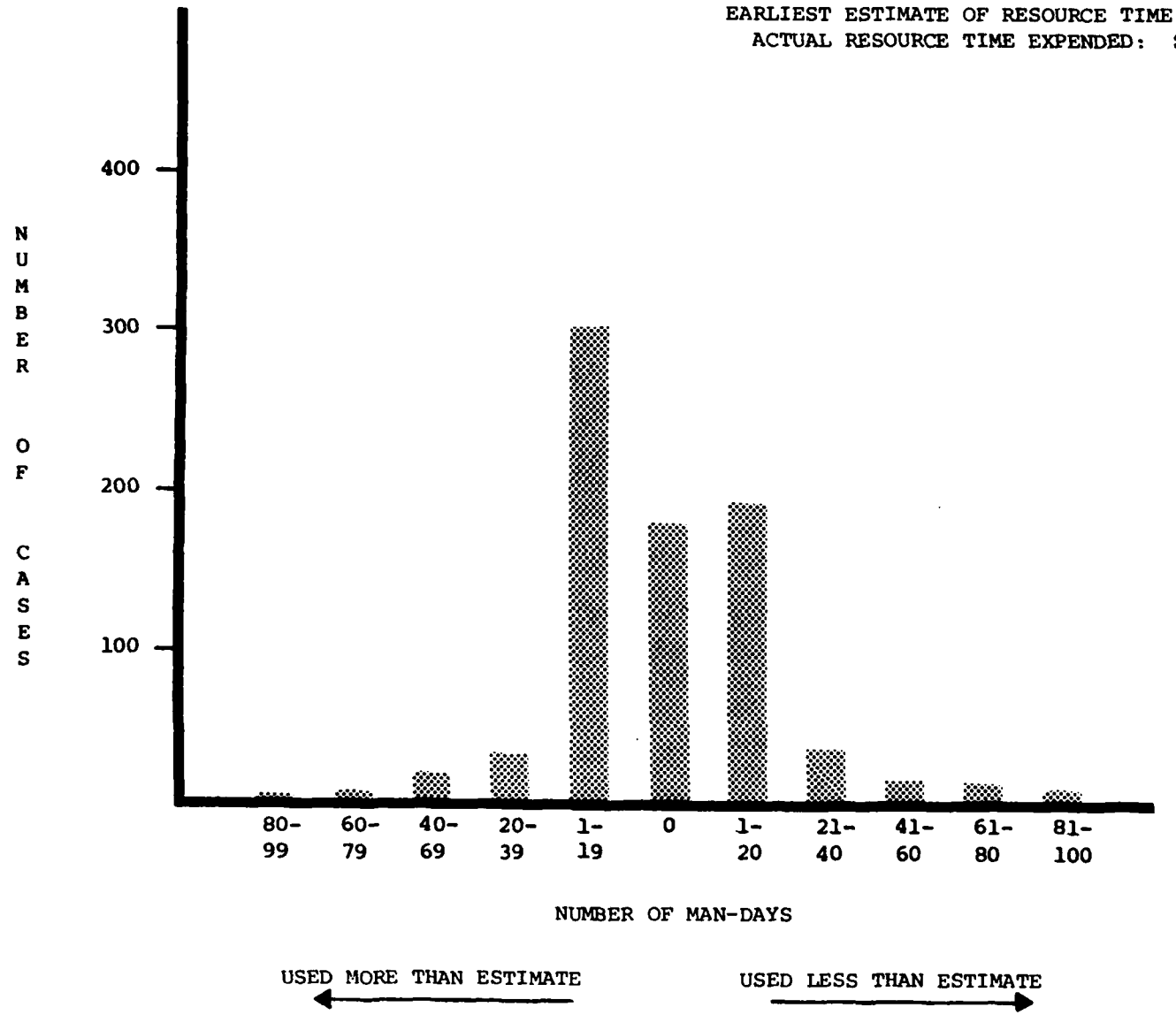

EARLIEST ESTIMATE OF RESOURCE TIME REQUIRED VS.
ACTUAL RESOURCE TIME EXPENDED: SAILS AB
NUMBER OF CASES
400
300
200
100
80-99
60-79
40-69
20-39
1-19
0
1-20
21-40
41-60
61-80
81-100
NUMBER OF MAN-DAYS
USED MORE THAN ESTIMATE
USED LESS THAN ESTIMATE

FIGURE 6

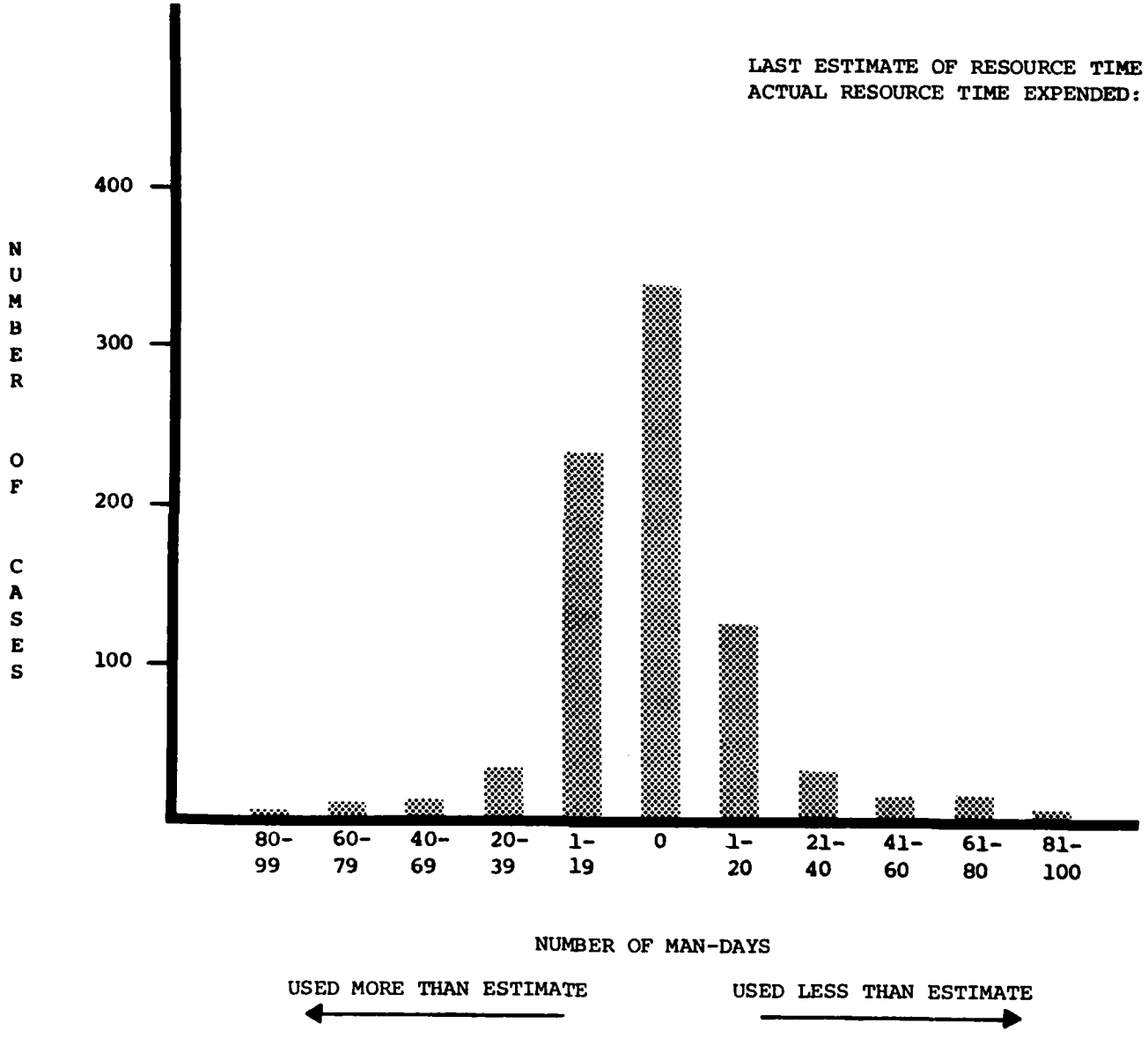

LAST ESTIMATE OF RESOURCE TIME REQUIRED VS.
ACTUAL RESOURCE TIME EXPENDED: SAILS AB
NUMBER OF CASES
400
300
200
100
80-99
60-79
40-69
20-39
1-19
0
1-20
21-40
41-60
61-80
81-100
NUMBER OF MAN-DAYS
USED MORE THAN ESTIMATE
USED LESS THAN ESTIMATE

FIGURE 7

EARLIEST ESTIMATE OF RESOURCE TIME REQUIRED VS.
ACTUAL RESOURCE TIME EXPENDED: ALL OTHER PROJECTS

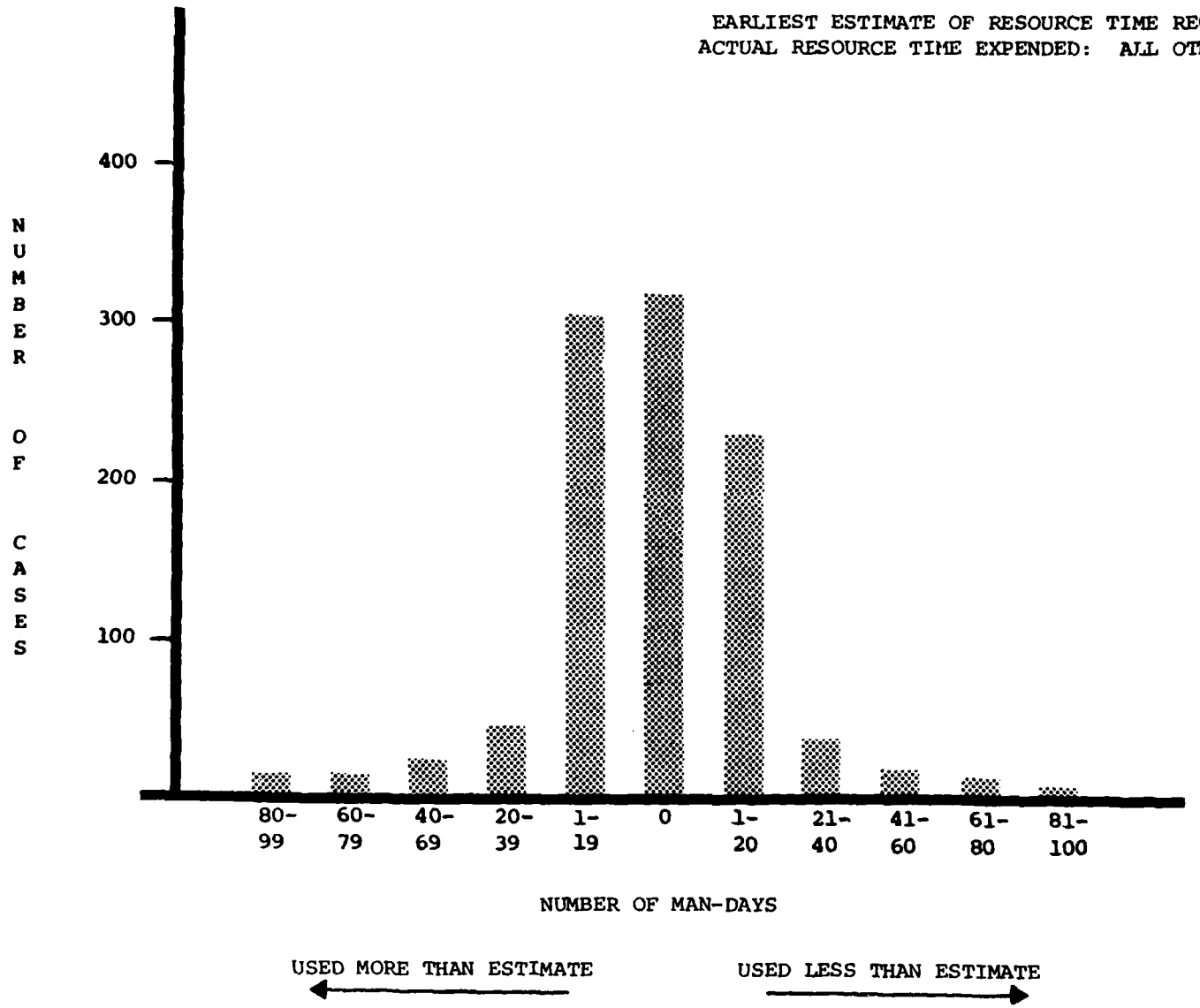


FIGURE 8

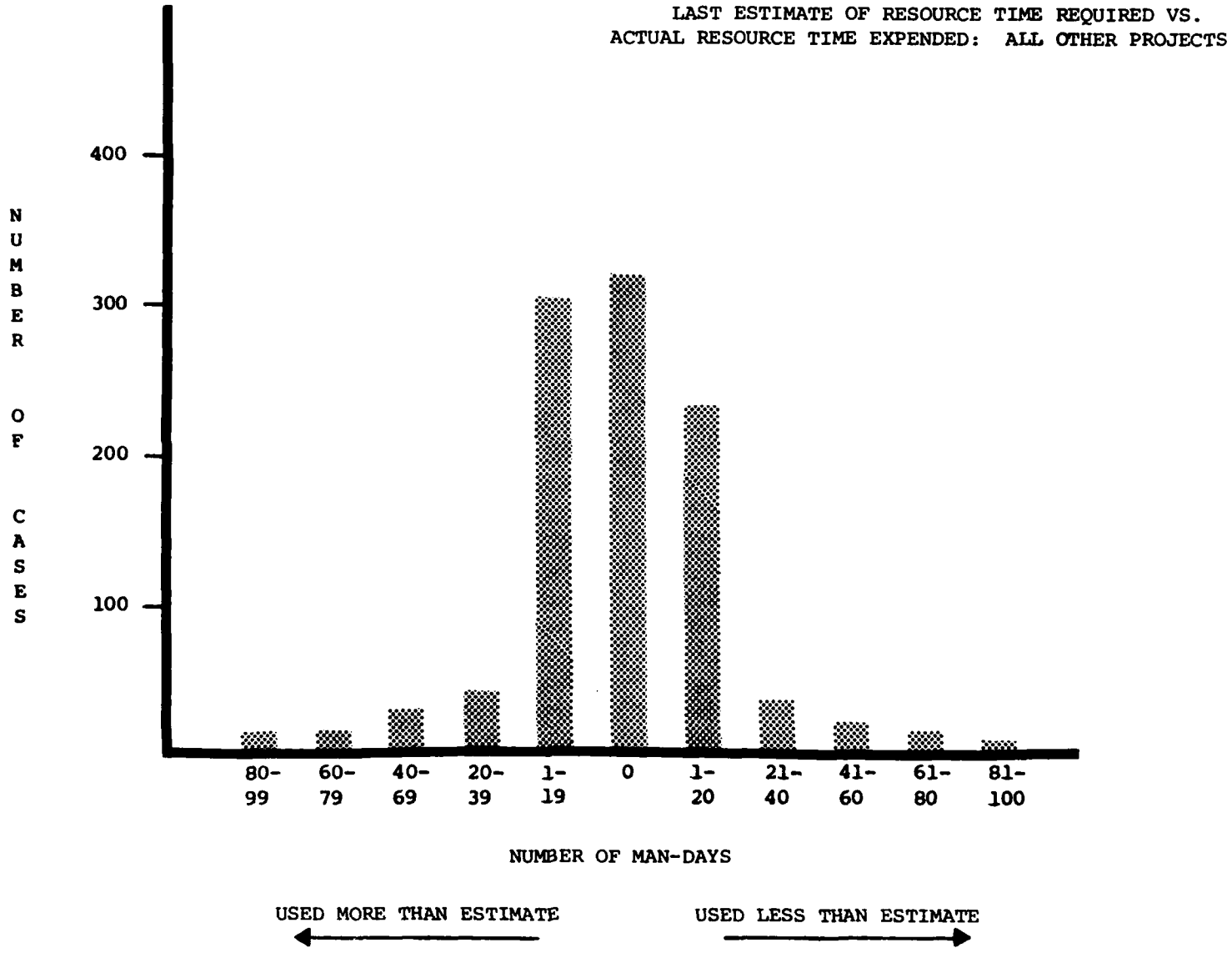

LAST ESTIMATE OF RESOURCE TIME REQUIRED VS.
ACTUAL RESOURCE TIME EXPENDED: ALL OTHER PROJECTS
NUMBER OF CASES
400
300
200
100
80-99
60-79
40-69
20-39
1-19
0
1-20
21-40
41-60
61-80
81-100
NUMBER OF MAN-DAYS
USED MORE THAN ESTIMATE
USED LESS THAN ESTIMATE

FIGURE 9

TABLE 6

DESCRIPTIVE STATISTICS OF DATA ITEMS

	SAILS AB				NON-SAILS AB			
	N	X	Maximum	Minimum	N	X	Maximum	Minimum
Original Resource Estimate	1086	18.41	490.80	0.0	1390	24.54	500.0	0.0
Previous Resource Estimate	674	16.85	998.10	0.0	18	61.94	472.0	0.0
Last Resource Estimate	1019	17.77	413.00	0.0	1390	25.04	500.0	0.0
Actual Resources Expended	821	22.43	471.00	0.0	1058	20.92	525.0	0.0
Estimated Lead Time (Original Finish Date Estimate--Setup Date)	1286	53.79	0.0	304.0	1384	65.90	1.0	860.0
Actual Lead Time (Actual Finish Date--Setup Date)	1286	68.30	0.0	308.0	1384	110.02	259.0	434.0
Error in Original Time Estimate (Actual Finish Date--Original Estimate)	1286	14.52	289.0	-248.0	1384	43.85	411.0	-699.0
Error In Latest Time Estimate (Actual Finish Date--Last Estimate)	1286	3.92	283.0	-165.0	1384	17.77	291.0	-699.0
Error In Original Resource Estimate (Actual Resources--Original Estimate)	821	3.02	414.0	-286.40	1058	-4.52	445.0	-466.0
Error in Last Resource Estimate (Actual Resources--Last Resource Estimate)	821	3.83	321.0	-192.0	1058	-5.06	445.0	-466.0

APPENDIX A

IMPACT ESTIMATING

System _____ Estimator's Name _____

SCR Number _____ Program _____ Date _____

1. Total number of file formats affected by this SCR. (1) _____

2. Program Functions. NOTE: This table reflects the number of functions in the program affected by the SCR. (e.g., There are 3 edits, two simple, one complex; and one very complex sort/merge. Enter edit/validation 4 x 2 = 8, 8 x 1 = 8, and sort/merge 4 x 1 = 4 Total = 20.)

	Simple	Complex	Very Complex
Edit/Validation	4 x ___	8 x ___	12 x ___
Sort/Merge Process	2 x ___	3 x ___	4 x ___
Internal Data Manipulation	2 x ___	3 x ___	4 x ___
File Search	2 x ___	3 x ___	4 x ___
Table Lookup	3 x ___	5 x ___	7 x ___
Calculations	3 x ___	5 x ___	7 x ___
Utilities/Subroutines	2 x ___	3 x ___	4 x ___
Job Control Language	1 x ___	2 x ___	3 x ___
Totals	___ +	___ +	___ =

(2) _____

3. Job knowledge required to program this SCR. (Limited = 1; General = 2; Detailed = 3) (3) _____

4. Program turnaround factor (Effective IAP usage = 0.6; more than once per day = 0.8; once per day = 1.0; less than once per day = 1.2) (4) _____

5. Extent of Change (5) _____

Developmental	= 1.0	Minor Modification	= .40
Major Change	= .67	Maintenance	= .34
Major Modification	= .50	Minor Technical Change	= .28
		JCL Change Only	= .24

DEVELOPMENT TIME

(1) Number of File Formats Affected		(2) Program Function Total		(2A) Subtotal
_____	+	_____	=	_____

(2A) Subtotal (From 2A)		(3) Job Knowledge Required		(4) Program Turnaround		(5) Extent of Change	=	DIRECT Development Time		Non-Project Factor	=	GROSS Development Time
_____	x	_____	x	_____	x	_____	=	_____	x	2.45	=	_____

NARRATIVE

1. Files Affected

_____ _____ _____ _____

_____ _____ _____ _____

2. Description of program changes on which this estimate is based:

SCR ESTIMATE SUMMARY

1. Gross Development Time = _____ man-days (from front)

a. Review and Analysis	GDT x 0.15	_____ x 8 =	_____*
b. Design	GDT x 0.20	_____ x 8 =	_____*
c. Programming (including Level I testing)	GDT x 0.35	_____ x 8 =	_____*
d. Testing (including Level II and III testing	GDT x 0.25	_____ x 8 =	_____*
e. Documentation (enter zero for none)	GDT x 0.05	_____ x 8 =	_____*

2. Total Project Mandays (sum of a-e above)
TPMD = _____ mandays x 8 = _____manhours

3. Estimated Machine Hours _____*

* Enter these figures in the appropriate blocks of the Input Analysis Section of USACSC form 6, System Change Request)