

1993 RESEARCH REPORTS

NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

JOHN F. KENNEDY SPACE CENTER

UNIVERSITY OF CENTRAL FLORIDA

EDITORS:

Dr. E. Ramon Hosler
Professor of Engineering
University of Central Florida

Ms. Carol Valdes/Mr. Tom Brown
Systems Training and Employee Development Branch
John F. Kennedy Space Center

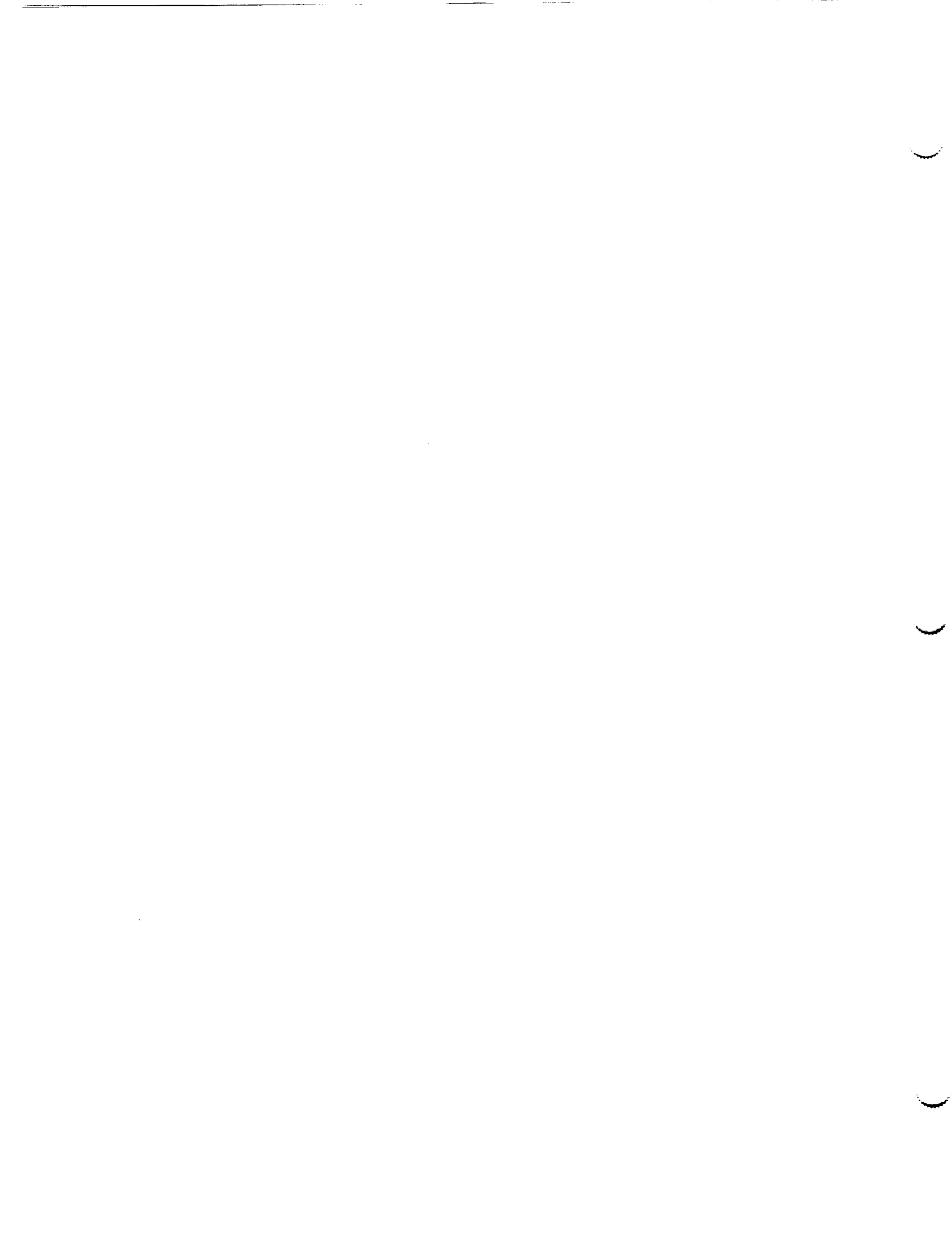
PREPARED FOR:

John F. Kennedy Space Center
Merritt Island, Florida

NASA Grant NGT-60002 Supplement: 11

Contractor Report No. CR-194678

October 1993



PREFACE

This document is a collection of technical reports on research conducted by the participants in the 1993 NASA/ASEE Summer Faculty Fellowship Program at Kennedy Space Center (KSC). This was the ninth year that a NASA/ASEE program has been conducted at KSC. The 1993 program was administered by the University of Central Florida in cooperation with KSC. The program was operated under the auspices of the American Society for Engineering Education (ASEE) with sponsorship and funding from the Office of Educational Affairs, NASA Headquarters, Washington, D.C. The KSC program was one of nine such Aeronautics and Space Research Programs funded by NASA Headquarters in 1993. The basic common objectives of the NASA/ASEE Summer Faculty Fellowship Program are:

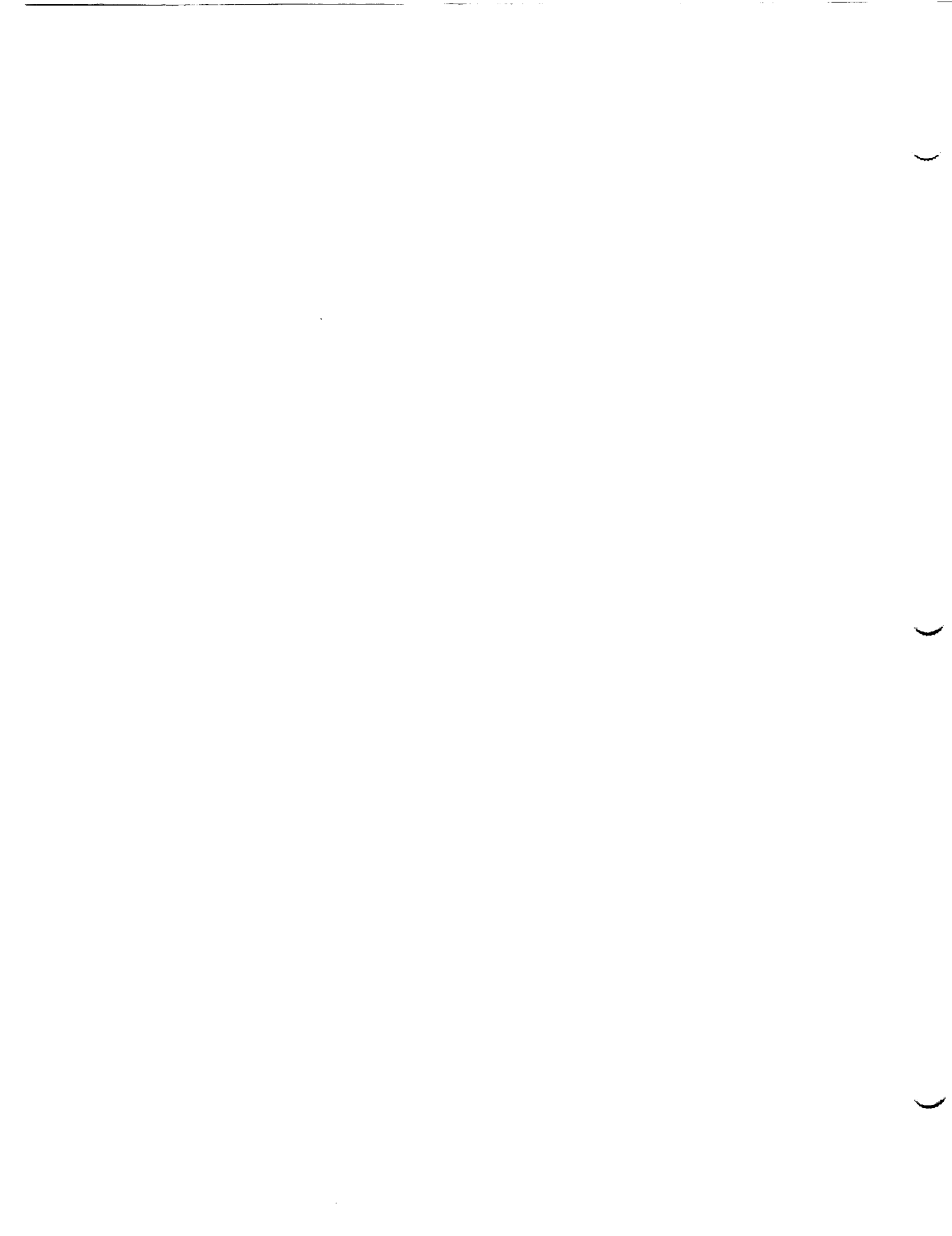
- a. To further the professional knowledge of qualified engineering and science faculty members;
- b. To stimulate an exchange of ideas between participants and NASA;
- c. To enrich and refresh the research and teaching activities of participants' institutions; and,
- d. To contribute to the research objectives of the NASA centers.

The KSC Faculty Fellows spent ten weeks (June 1 through August 6, 1993) working with NASA scientists and engineers on research of mutual interest to the University faculty member and the NASA colleague. The editors of this document were responsible for selecting appropriately qualified faculty to address some of the many problems of current interest to NASA/KSC. A separate document reports on the administration aspects of the 1993 program. The NASA/ASEE program is intended to be a two-year program to allow in-depth research by the University faculty member. In most cases a faculty member has developed a close working relationship with a particular NASA group that has provided funding beyond the two-year limit.

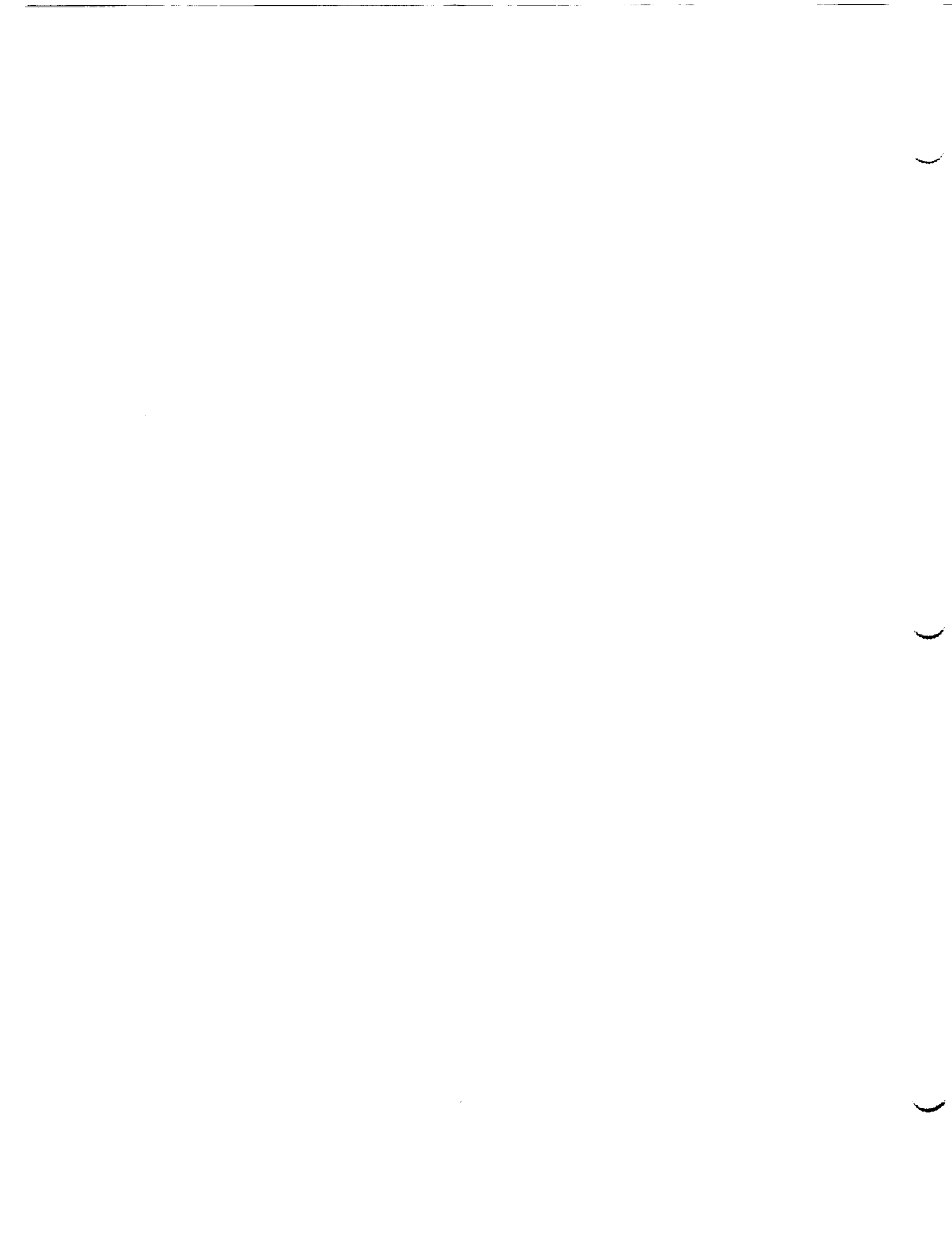


TABLE OF CONTENTS

		<u>PAGE</u>
1.	BATES, Harry E. "Wide Band Fiber-Optic Communications"	1 - 1
2.	BYERS, Robert M. "Control of a Serpentine Manipulator with Collision Avoidance"	27 - 2
3.	CALLE, Luz M. "Evaluation of Inorganic Zinc-Rich Primers Using Electrochemical Impedance Spectroscopy (EIS)"	47 - 3
4.	CENTENO, Martha A. "A Framework for SFC/DC System Data Analysis"	77 - 4
5.	DOTY, Keith L. "Mass Properties Measurement System Dynamics"	91 - 5
	"Mass Properties Measurement System: Dynamics and Statics Measurements"	111 - 6
6.	ENGEL, Bernard A. "A Prototype Expert/Information System for Examining Environmental Risks of KSC Activities"	135 - 7
7.	FORBES, Gregory S. "Lightning Studies Using LDAR and LLP Data"	165 - 8
8.	GEJJI, Raghvendra R. "Reduced Cost Alternatives to Premise Wiring Using ATM & Microcellular Technologies"	195 - 9
9.	JOGLEKAR, Prafulla N. "Cost Benefits of Advanced Software: A Review of Methodology Used at Kennedy Space Center"	229 - 10
10.	MADSEN, Brooks C. "Chemical Characterization of Some Aerobic Liquids in 'CELSS'"	269 - 11



11.	MALLADI, Narasimha S. "Rack Insertion End Effector (RIEE) Automation"	301 - 13
12.	MEHTA, Narinder K. "FTIR Quantification of Industrial Hydraulic Fluids in Perchloroethylene"	327 - 13
13.	ROSA, Ferdinand "Stress Corrosion Cracking Properties of 15-5PH Steel"	375 - 14
14.	RUSSELL, John M. "Interpretation of Discrepancies in Mass Spectroscopy Data Obtained from Different Experimental Configurations"	403 - 15
15.	SEPULVEDA, Jose A. "KSC Management Training System Project"	433 - 16
16.	SHERIF, Sherif A. *** Report Not Available at Publication Time ***	465/466 - 17
17.	THOMAS, Stan J. "Expanding the KATE Toolbox"	467 - 17
18.	WOHLMAN, Richard A. "Hurricane Risk Assessment: Rollback or Ride Out"	487 - 18
19.	ZEYTINOGLU, Nuri "Flame Trench Analysis of NLS Vehicles"	521 - 19
20.	ZOBRIST, George W. "Switch Configuration for Migration to Optical Fiber Network"	551 - 20



1994018758

N94-28231

1993 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA

51-32
197187
p. 25

WIDE BAND FIBER-OPTIC COMMUNICATIONS

PREPARED BY:	Dr. Harry E. Bates
ACADEMIC RANK:	Professor
UNIVERSITY AND DEPARTMENT:	Towson State University Department of Physics
NASA/KSC	
DIVISION:	Electronic Systems
BRANCH:	Communications
NASA COLLEAGUE:	Po T. Huang
DATE:	August 20, 1993
CONTRACT NUMBER:	University of Central Florida NASA-NGT-60002 Supplement: 11

ACKNOWLEDGEMENTS

This research effort would not have been possible without the assistance, cooperation and careful planning of a number of people. Larry Hand and Po T. Huang suggested the problem and provided valuable support along the way. Po T. Huang was always available to answer questions and provide literature references. He also gave guidance and support in more ways than can be described here. All levels of Engineering management above Po and Larry beginning with Perry Rodgers provided valuable assistance and showed genuine interest in the project. Everyone was willing to listen and suggest solutions to problems. Special thanks to Fred McKenzie for helping to provide computer support and introductions to people who could help with other problems. INET personnel were most helpful in the laboratory. Special thanks to F. Houston Galloway and Bob Swindle who worked closely in all laboratory phases of this project. Thanks to Tim Shepard, co-op student, who participated in a number of experiments and was always willing to lend a hand in the lab. Also, thanks to David Wedekine who helped with software problems on the Macintosh IIx. Last but not least I would like to thank Dr E. R. Hosler of UCF who managed the Summer Faculty Program and provided a great deal of support and assistance.

ABSTRACT

A number of optical communication lines are now in use at the Kennedy Space Center (KSC) for the transmission of voice, computer data and video signals. At the present time most of these channels utilize a single carrier wavelength centered near 1300 nm. As a result of previous work the bandwidth capacity of a number of these channels is being increased by transmitting another signal in the 1550 nm region on the same fiber. This is accomplished by means of wavelength division multiplexing (WDM). It is therefore important to understand the bandwidth properties of the installed fiber plant. This work developed new procedures for measuring the bandwidth of fibers in both the 1300nm and 1550nm region. In addition, a preliminary study of fiber links terminating in the Engineering Development Laboratory was completed.

SUMMARY

A number of multimode optical communication links are now in use at the Kennedy Space Center (KSC) for the transmission of voice, computer data and video signals. At the present time most of these channels utilize a single carrier wavelength centered near 1300 nm. As a result of previous work the bandwidth capacity of a number of these channels is being increased by transmitting another signal in the 1550 nm region on the same fiber. This is accomplished by means of wavelength division multiplexing (WDM).

The main goal of this experimental program was to characterize the bandwidth of the fibers used in these systems in the 1550 nm window.

During this project, new test procedure was developed to measure bandwidth using the HP 8702 Lightwave Component Analyzer System and associated components. The system was used to determine the bandwidth of a small number of fiber optics communications links terminating in the Engineering Development Laboratory at KSC.

Using the new test procedure, bandwidth measurements were accomplished on all multimode fiber test links terminating in the engineering development laboratory. These measurements showed that the 1550 LED bandwidth distance product of $.42 \pm .01$ GHz*km.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS	6
LIST OF TABLES	7
ABBREVIATIONS AND ACRONYMS LIST	8
I. INTRODUCTION	9
II. ANALYSIS OF EXISTING DATA	10
III. LIGHTWAVE ANALYZER SYSTEM AND PROCEDURES ANALYSIS...	11
IV. BANDWIDTH OF KSC MULTIMODE FIBER	21
V. CONCLUSIONS	23
REFERENCES	24

LIST OF ILLUSTRATIONS

- Figure 2-1. Lightwave Analyzer Setup Used to Collect Data
- Figure 2-2. Spectrum of Link from CDSC to HQ using 1530 nm LED
- Figure 3-1. Optical spectrum of 1.3 micrometer LED E/O converter
- Figure 3-2. Optical spectrum of 1.5 micrometer LED E/O converter
- Figure 3-3. Optical spectrum of 1.3 micrometer Laser E/O converter
- Figure 3-4. Optical spectrum of 1.5 micrometer Laser E/O converter
- Figure 3-5. Optical calibration spectra
- Figure 3-6. 1530 nm Laser E/O-Short Link-O/E Response Spectrum
- Figure 3-7. 1530 nm LED E/O-Short Link-O/E Response Spectrum
- Figure 3-8. Improved equipment setup for bandwidth tests
- Figure 4-1 Typical Link from Fiber Optic Lab to CDIC and Back
- Figure 4-2. Sample Spectrum EDL Loop 7-8 1550nm LED E/O source

LIST OF TABLES

- Table 3-1. Cutoff signal levels below which no measurement should be attempted without changing the procedure.
- Table 4-1. Bandwidth measurements EDL links using 1300nm laser
- Table 4-2. Bandwidth measurements EDL links using 1550nm laser
- Table 4-3. Bandwidth measurements EDL links using 1300nm LED
- Table 4-4. Bandwidth measurements EDL links using 1550nm LED

ABBREVIATIONS AND ACRONYMS LIST

CDSC	Communications Distribution and Switching Center
OSA	Optical Spectrum Analyzer
LED	Light Emitting Diode
KSC	Kennedy Space Center
WDM	Wavelength Division Multiplexer or Demultiplexer

I

INTRODUCTION

A number of multimode optical communication links are now in use at the Kennedy Space Center (KSC) for the transmission of voice, computer data and video signals. At the present time most of these channels utilize a single carrier wavelength centered near 1300 nm. As a result of previous work the bandwidth capacity of a number of these channels is being increased by transmitting another signal in the 1550 nm region on the same fiber. This is accomplished by means of wavelength division multiplexing (WDM).

The main goal of this experimental program was to characterize the bandwidth of the fibers used in these systems in the 1550 nm window.

Practical considerations such as modal dispersion, material (chromatic) dispersion, detector rise time, modulation limits of the receiver...etc. limit the usable bandwidth.¹ The fibers have the capability of transmitting a multiplicity of signals simultaneously in each of two separate bands (1300nm and 1550 nm)^{2,3}.

It is important to mention that this study has not determined the properties of the overall fiber optic plant at KSC with any great statistical certainty.

II

ANALYSIS OF EXISTING DATA

2.1 SAMPLE OF DATA

A number of RF spectra were collected using the HP8702 Lightwave Component Analyzer. The experimental arrangement used is shown in Figure 2-1 and a sample of the spectra collected using a 1530nm LED as the E/O converter is shown in Figure 2-2. Four different E/O sources can be employed LEDs and lasers with output wavelengths in the 1.5 micrometer or 1.3 micrometer windows.

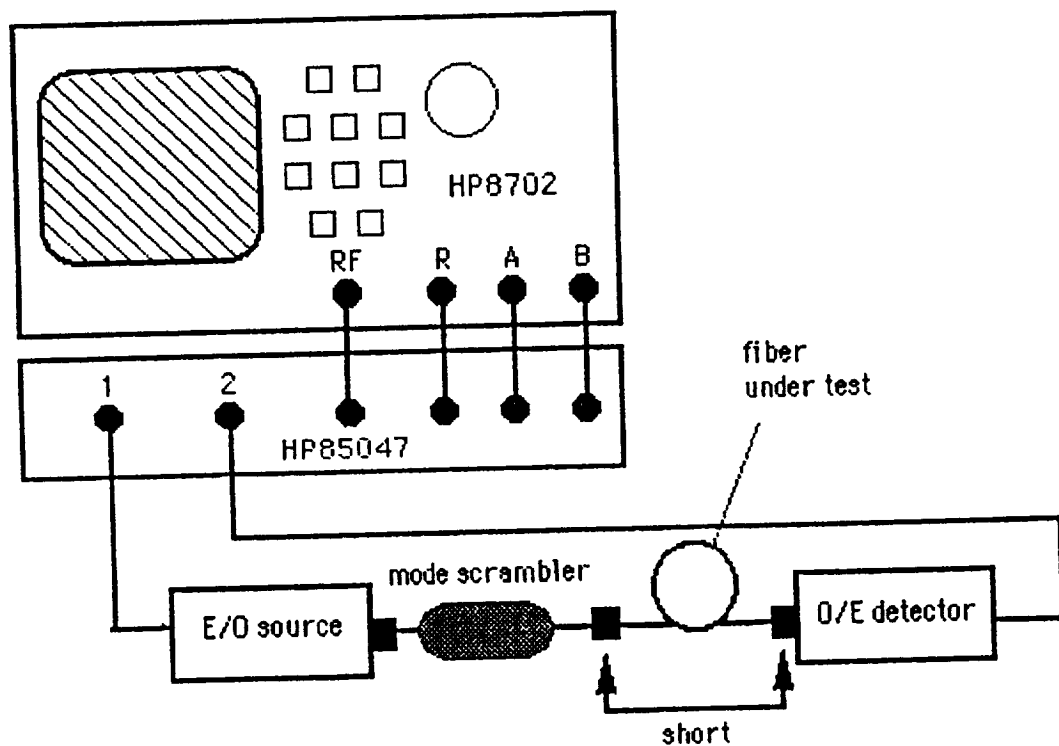


Figure 2-1. Lightwave Analyzer Setup Used to Collect Data

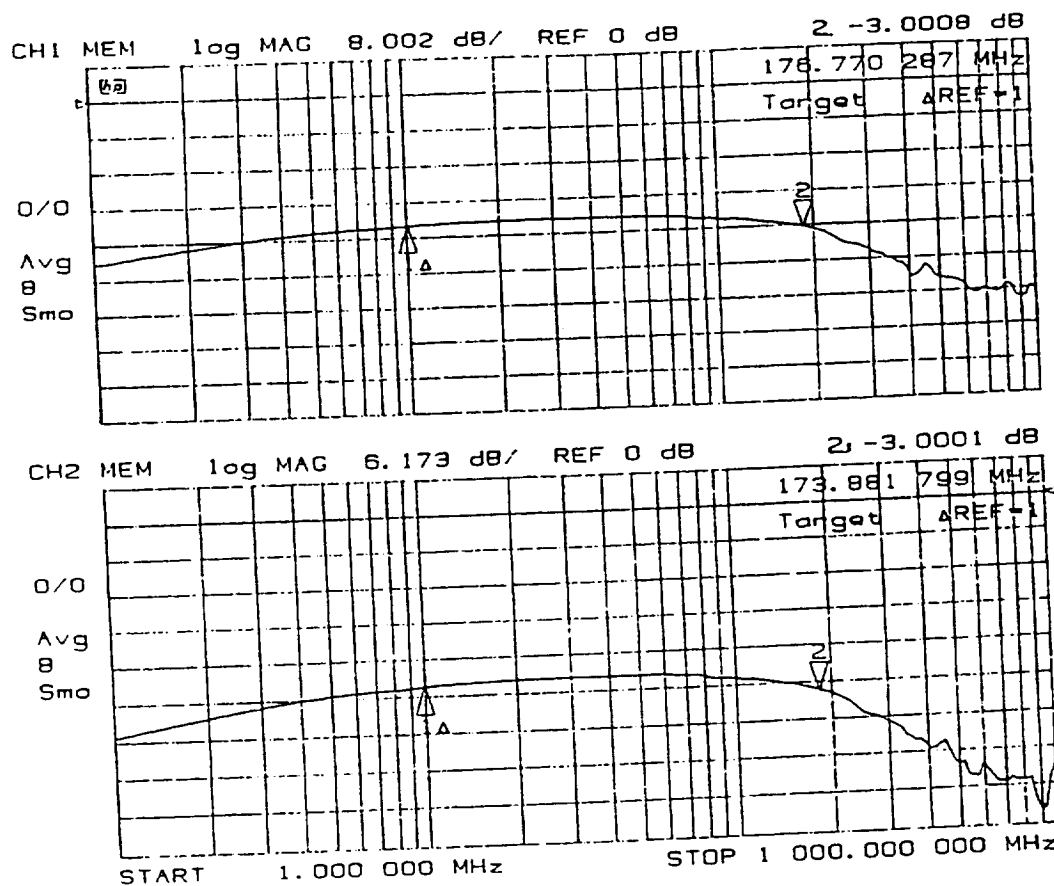


Figure 2-2. Spectrum of Link from CDSC to HQ using 1530 nm LED

A preliminary analysis of these results revealed fundamental problems with the data. Most spectra showed a 6 to 10 dB roll on in the low frequency region, a relatively flat area in the mid frequency range and a roll off at the high frequency end. There is no physical reason for the spectrum of an optical fiber link to roll on. A fiber acts like a low pass filter and should exhibit the best transmission at the lowest frequencies. Therefore an analysis of each element of the lightwave analyzer as well as test procedures was initiated.

III

LIGHTWAVE ANALYZER SYSTEM AND PROCEDURES ANALYSIS

3.1 Optical Spectra of E/O Converters

An Anaritsu optical spectrum analyzer was available in the laboratory to be used to perform spectral analysis of coherent and incoherent sources. This piece of test equipment was equipped with an IEEE 488 computer interface which provides for bidirectional computer communication. This equipment was interfaced with a Macintosh IIx computer to enable the efficient collection of spectral data in machine readable form during a project in the Summer of 1989. The analyzer was used to collect spectra of the E/O converters.

3.1.1 LED SPECTRA. The optical spectra of the LED-type E/O converters were collected. They are shown in Figures 3-1 and 3-2 below:

Spectrum of 1.3 micrometer LED

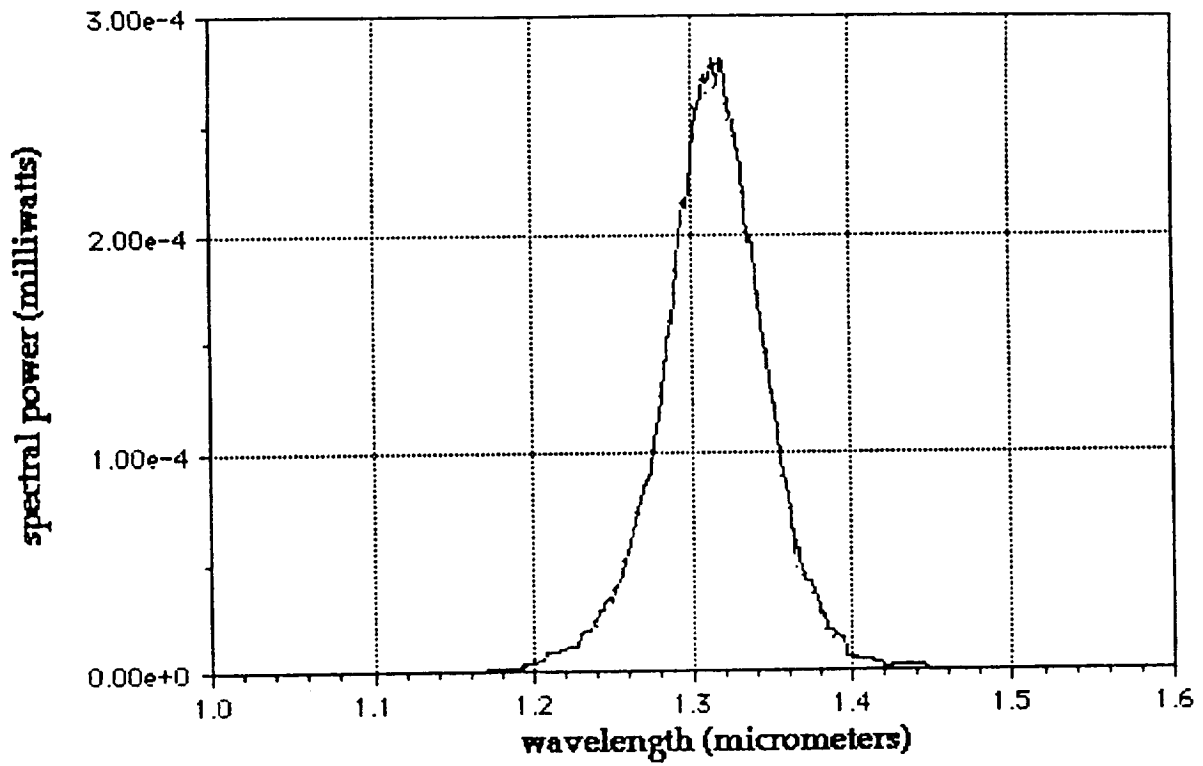


Figure 3-1. Optical spectrum of 1.3 micrometer LED E/O converter

Spectrum of 1.53 micrometer LED

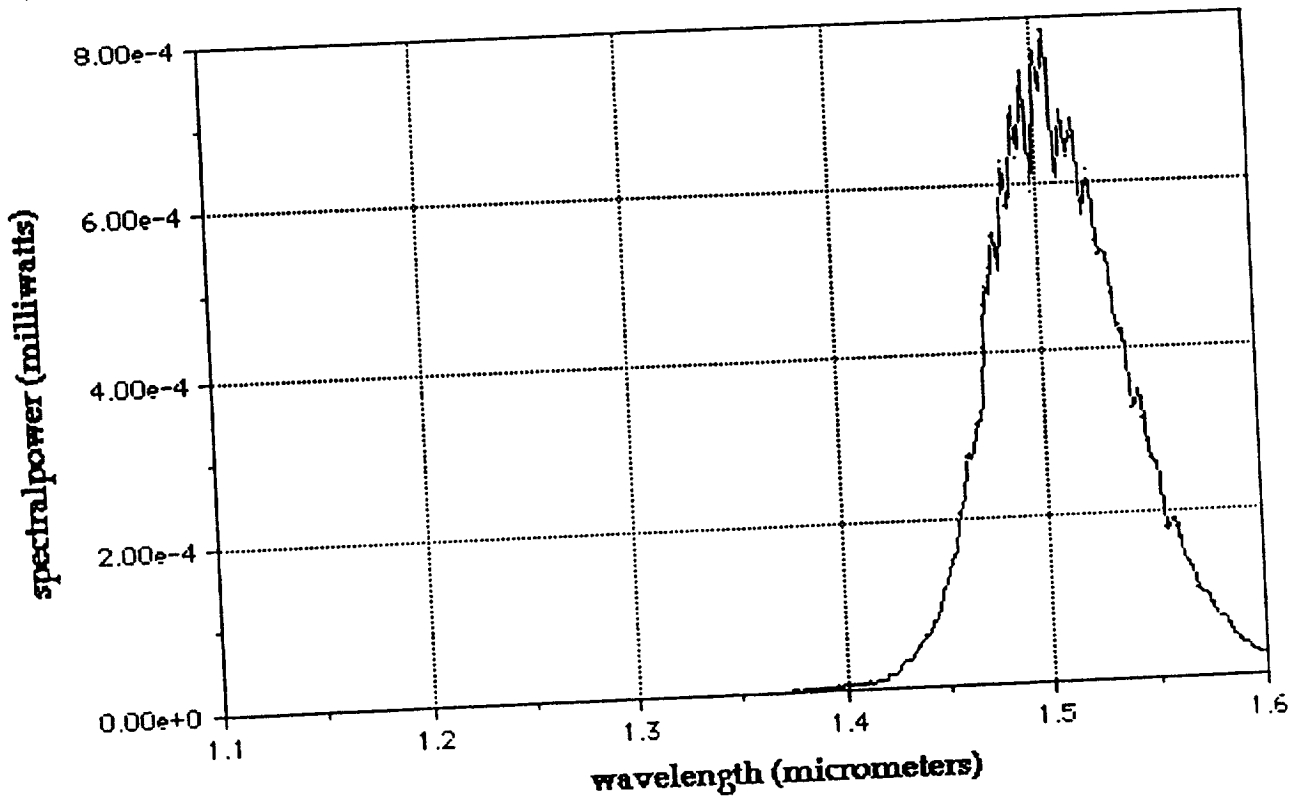


Figure 3-2. Optical spectrum of 1.5 micrometer LED E/O converter

3.1.2 LASER SPECTRA. The optical spectra of the LASER-type E/O converters were collected. They are shown in Figures 3-3 and 3-4 below:

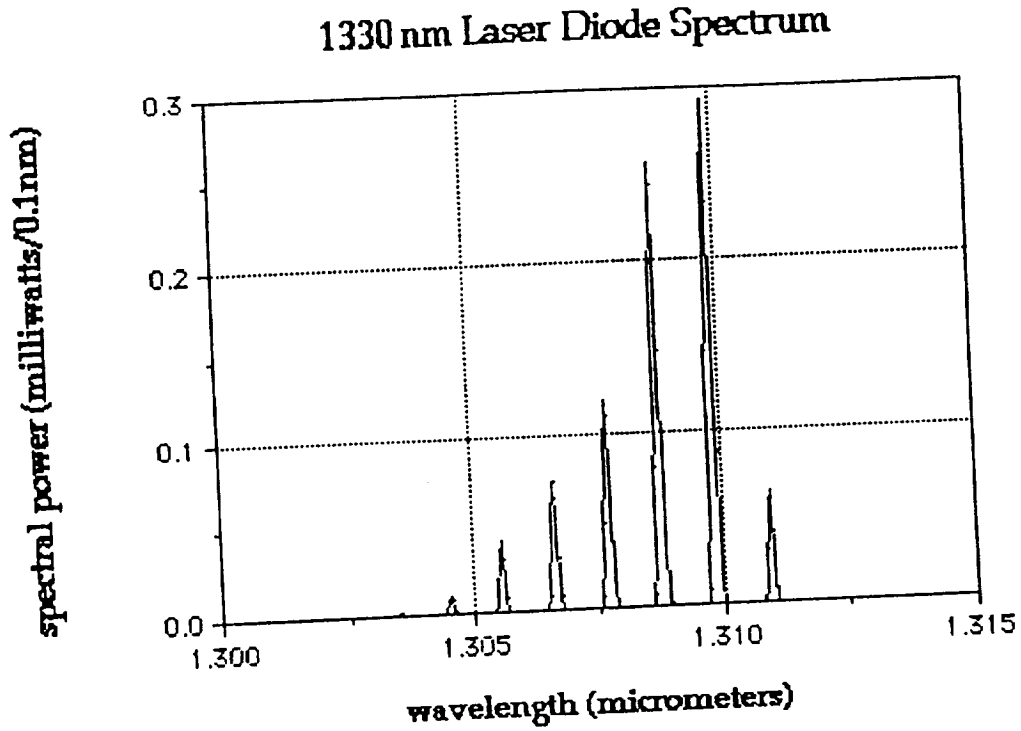


Figure 3-3. Optical spectrum of 1.3 micrometer Laser E/O converter

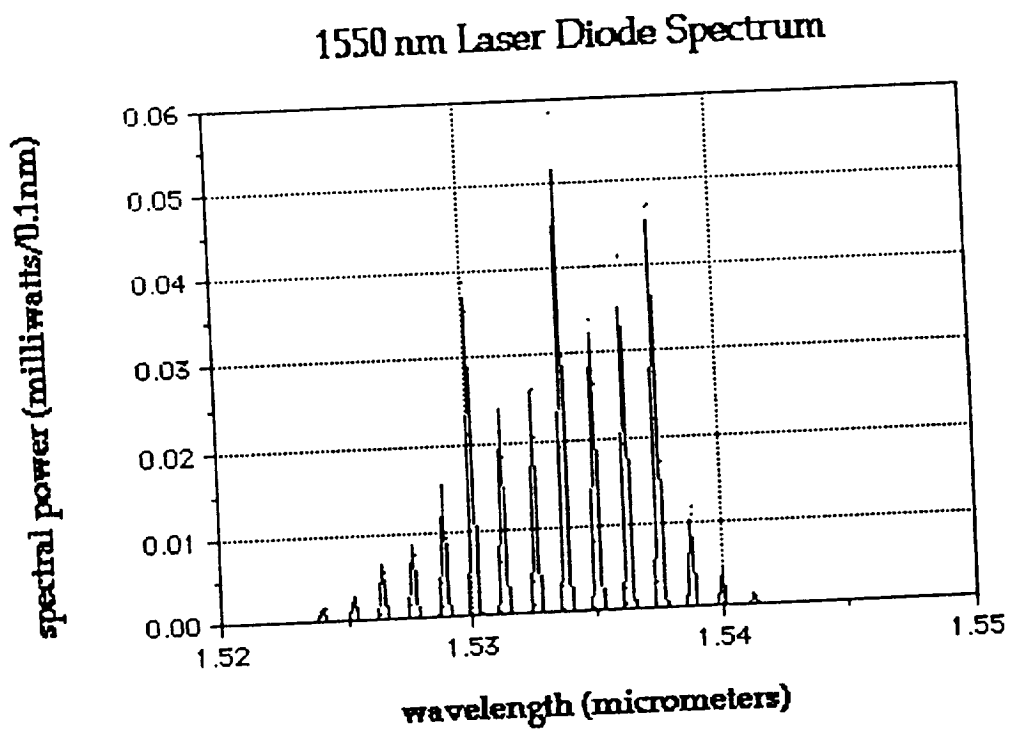


Figure 3-4. Optical spectrum of 1.5 micrometer Laser E/O converter

3.1.3 OPTICAL CALIBRATION CHECK OF SPECTRUM ANALYZER. The OSA was checked by using a HeNe laser source. A graph of the spectra of two LED's and the sharp peak of the 0.6328 micrometer laser is shown below. All spectra collected were normal.

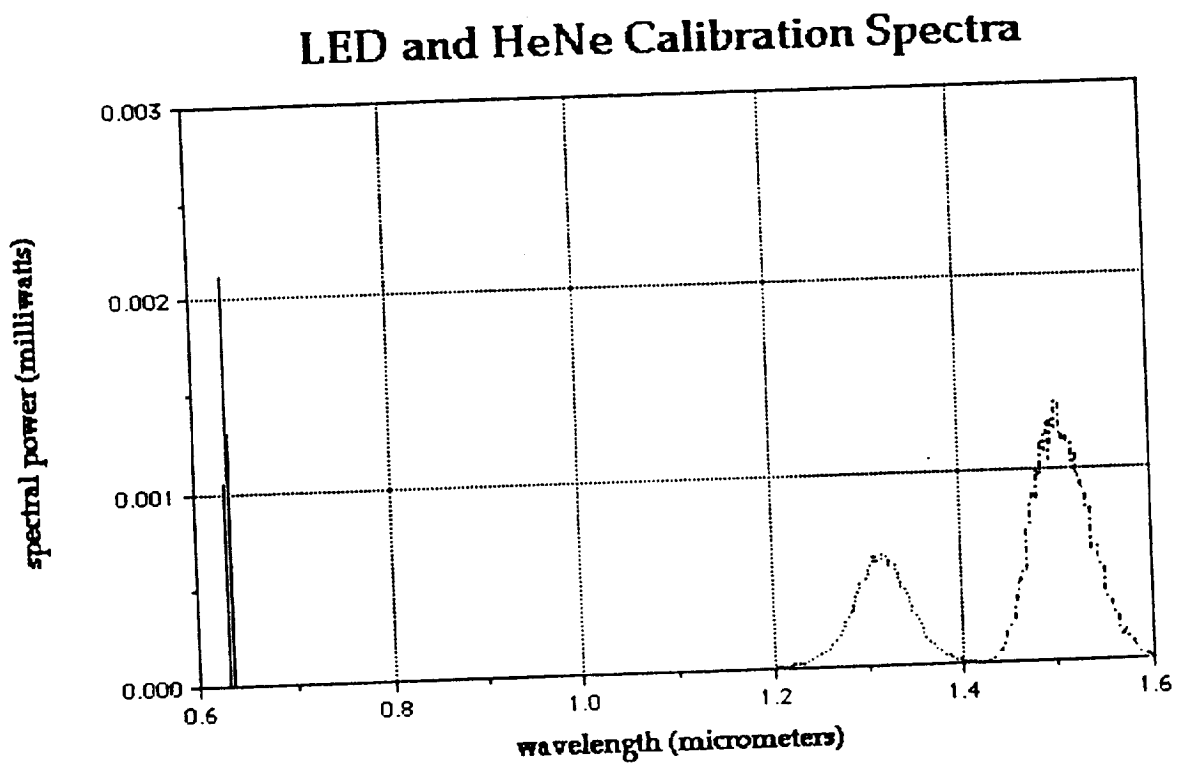


Figure 3-5. Optical calibration spectra

3.2 BANDWIDTH OF E/O O/E CONVERTERS

The bandwidth of the E/O - short fiber- O/E link was checked referenced to a short RF cable. Example RF spectra are shown for the 1.5 micrometer laser source and the 1.5 micrometer LED source in Figures 3-6 and 3-7 below. The lasers had the same general characteristics and the LEDs had the same general characteristics. The lasers generally had a low-pass filter type response and the LEDs showed a marked roll on at low frequencies and then a roll off at high frequencies.

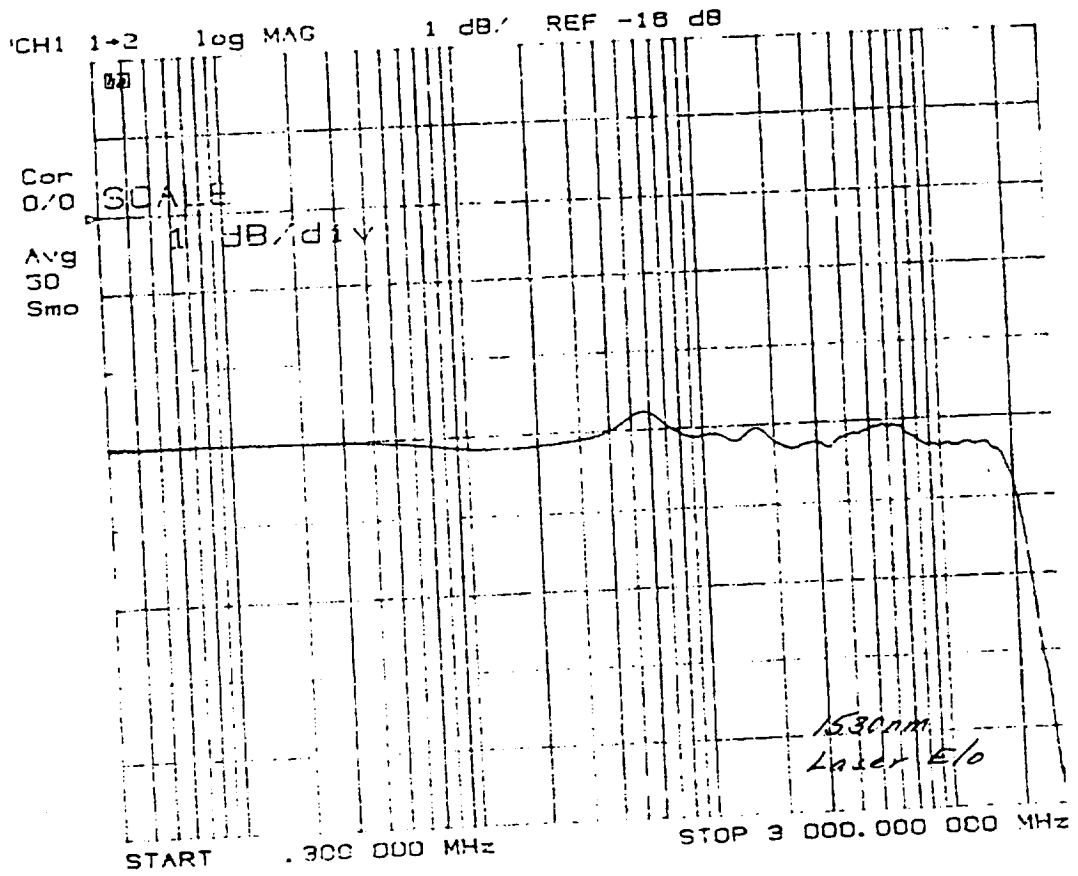


Figure 3-6. 1530 nm Laser E/O-Short Link-O/E Response Spectrum

ORIGINAL PAGE IS OF POOR QUALITY

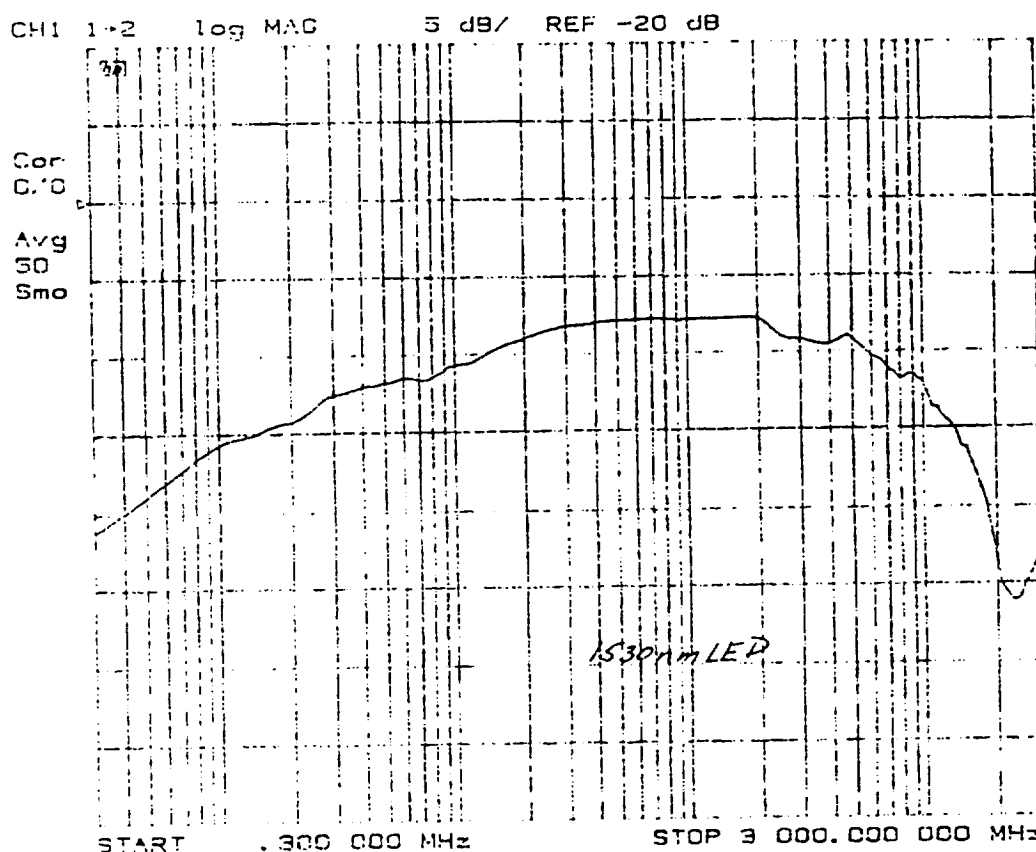


Figure 3-7. 1530 nm LED E/O-Short Link-O/E Response Spectrum

3.3 PROCEDURES ANALYSIS

The operating procedure and equipment setup used for collecting data was extensively analyzed. A number of problems were identified and corrected. Most of the problems could be traced to misleading equipment indications and poor human-machine interface software. Each item discovered together with the fix for the item is discussed in the paragraphs below:

3.3.1 DONE COMMAND USE AFTER CALIBRATION. The test procedure includes a step where a short fiber link is attached between the E/O - O/E system and a sequence of soft keys are depressed to memorize the thru response of the basic system elements. In the original procedure, the sequence of key strokes was CAL,CAL. MENU, RESPONSE, THRU. Then another key on the front panel was pushed. (Please note

that labels in bold indicate front panel keys and small caps indicate soft key inputs). A sequence of soft keys appear on the right side of the CRT and get various screen labels depending on the mode of the analyzer. The DONE soft key at the bottom of the screen must be pressed in order for the instrument to accept and memorize the reference loop. Since this was not done in the collection of data, the spectra did not reflect just the response of the fiber, but also the response of other system elements that should have been compensated for. Therefore the original data is useless.

3.3.2 AVERAGING DURING DATA COLLECTION AND CALIBRATION. If data is to be collected using averaging, then the reference short loop response calibration should be done with the same level of averaging. If averaging is used, the DONE softkey must be pressed within one second of completing averaging or the instrument incorrectly restarts the calibration averaging procedure. This is a software flaw in the instrument.

3.3.3 USE OF PORT B FOR INPUT PORT. The diagram displayed by the instrument during the guided setup phase of operation is incorrect. The best dynamic range is obtained by utilizing port B as the input port from the network instead of port 2 as shown in the guided setup. This gives an improvement in dynamic range of about 10 dB. An improved equipment setup diagram is shown below in Figure 3-8:

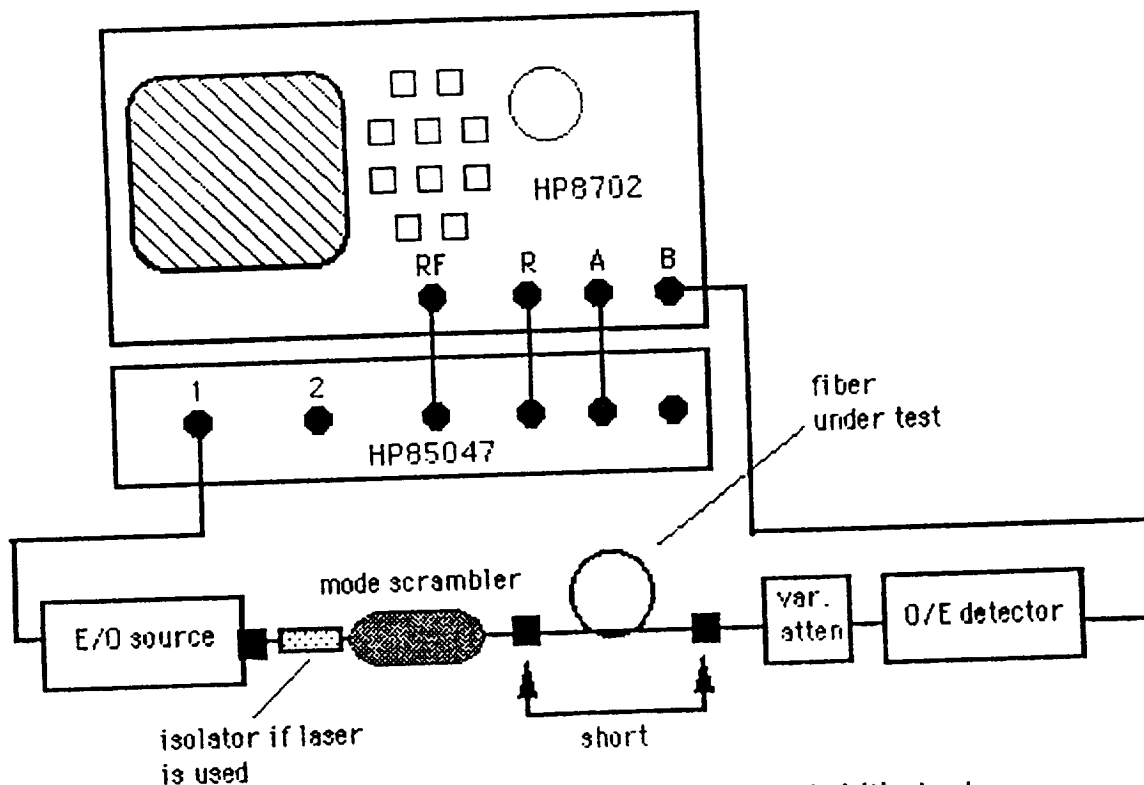


Figure 3-8. Improved equipment setup for bandwidth tests

3.3.4 SINGLE CHANNEL VS DUAL CHANNEL OPERATION. Improved graphs and more reliable operation can be obtained if single channel operation is used.

3.3.5 SIGNAL LEVEL TEST. With an averaging factor of 16, an acceptable signal level from the E/O converter thru the test fiber to the receiver is different for each different E/O converter. It is recommended that acceptable signal levels be established for each converter and wavelength region and that a test be made of the signal level thru the loop prior to making bandwidth measurements. If the signal level is not sufficient, bandwidth measurements should not be attempted. A conservative set of cutoff points is given in Table 3-1 below.

Table 3-1. Cutoff signal levels below which no measurement should be attempted without changing the procedure.

E/O Converter	MINIMUM SIGNAL LEVEL
1550 LED	-34 dBm
1550 LASER	-32 dBm
1300 LED	-26 dBm
1300 LASER	-33 dBm

3.3.6 INTEGER REFERENCE LEVEL CONTROL. Vertical reference level control should be set to integer values in order to make it easier to read the output graphs.

3.3.7 DB PER DIVISION FIXED VERTICAL SCALE. Vertical scale/div setting should be set to integer values in order to make it easier to read the output graphs. I recommend 1 dB/division.

IV

BANDWIDTH OF KSC MULTIMODE FIBER

4.1 TYPICAL FIBER TEST LINK

The Fiber Optics Laboratory in the EDL building has a series of multimode fiber optic links to the CDSC building. See Figure 4 -1 below for a diagram typical of any of the 2.4 km links marked (5-12). The bandwidth of pairs of these links was measured using the improved procedure discussed in the previous section. The bandwidth of the 6.3 km links (1-4) were also measured. They are routed from the EDL to the Banana River Repeater Station.

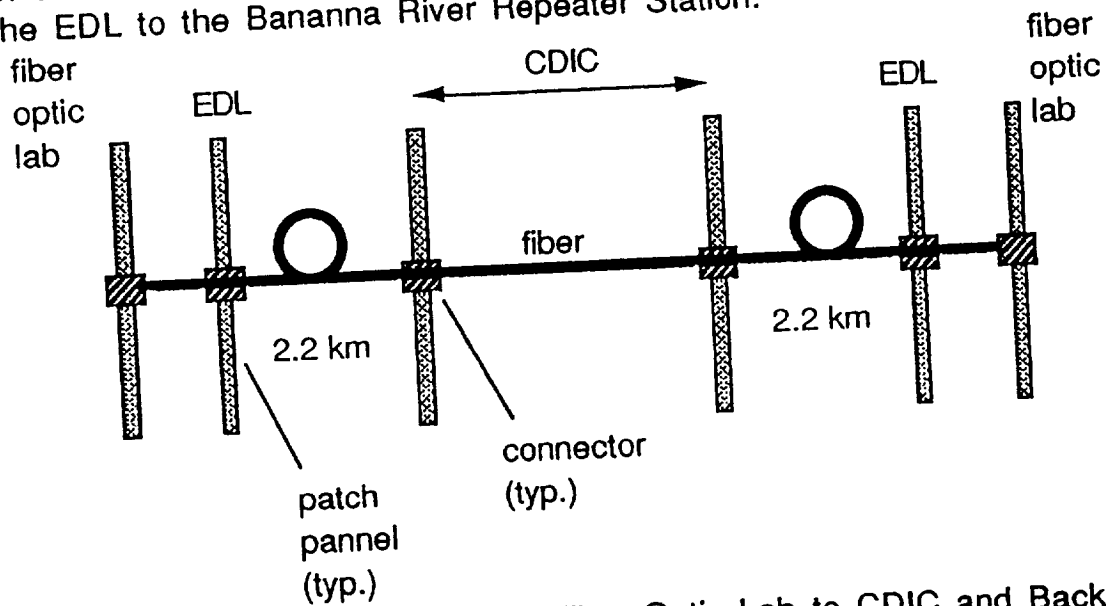


Figure 4-1 Typical Link from Fiber Optic Lab to CDIC and Back

Tables 4-1 thru 4-4 detail the results of the bandwidth measurements.

Table 4-1. Bandwidth measurements EDL links using 1300nm laser

	Length	Forward	Reverse	Avg.	BW*Dist.
	(km)	(MHz)	(MHz)	(MHz)	(GHz*km)
loop 1-2 to BRRS	12.6	369	414	391.5	4.93
loop 3-4 to BRRS	12.6	230	241	235.5	2.97
loop 5-6 to CDSC	4.8	640	653	646.5	3.10
loop 7-8 to CDSC	4.8	399	459	429	2.06
loop 9-10 to CDSC	4.8	494	509	501.5	2.41
loop 11-12 to CDSC	4.8	344	336	340	1.63

Table 4-2. Bandwidth measurements EDL links using 1550nm laser

	Length	Forward	Reverse	Avg.	BW*Dist.
	(km)	(MHz)	(MHz)	(MHz)	(GHz*km)
loop 1-2 to BRRS	12.6	could not measure-----			
loop 3-4 to BRRS	12.6	97	102	99.5	1.25
loop 5-6 to CDSC	4.8	275	260	267.5	1.28
loop 7-8 to CDSC	4.8	181	261	221	1.06
loop 9-10 to CDSC	4.8	262	198	230	1.10
loop 11-12 to CDSC	4.8	179	169	174	0.84

Table 4-3. Bandwidth measurements EDL links using 1300nm LED

	Length	Forward	Reverse	Avg.	BW*Dist.
	(km)	(MHz)	(MHz)	(MHz)	(GHz*km)
loop 1-2 to BRRS	12.6	252	242	247	3.11
loop 3-4 to BRRS	12.6	198	196	197	2.48
loop 5-6 to CDSC	4.8	470	479	474.5	2.28
loop 7-8 to CDSC	4.8	345	330	337.5	1.62
loop 9-10 to CDSC	4.8	384	371	377.5	1.81
loop 11-12 to CDSC	4.8	281	273	277	1.33

Table 4-4. Bandwidth measurements EDL links using 1550nm LED

	Length (km)	Forward (MHz)	Reverse (MHz)	Avg. (MHz)	BW*Dist. (GHz*km)
loop 1-2 to BRRS	12.6	34.1	33.6	33.85	0.43
loop 3-4 to BRRS	12.6	34.2	34.1	34.15	0.43
loop 5-6 to CDSC	4.8	89	89.7	89.35	0.43
loop 7-8 to CDSC	4.8	86.6	89.2	87.9	0.42
loop 9-10 to CDSC	4.8	91.3	91.2	91.25	0.44
loop 11-12 to CDSC	4.8	82.9	82.7	82.8	0.40

An example spectrum collected using the 1550nm LED E/O source for loop 7-8 is shown in Figure 4-2 below.

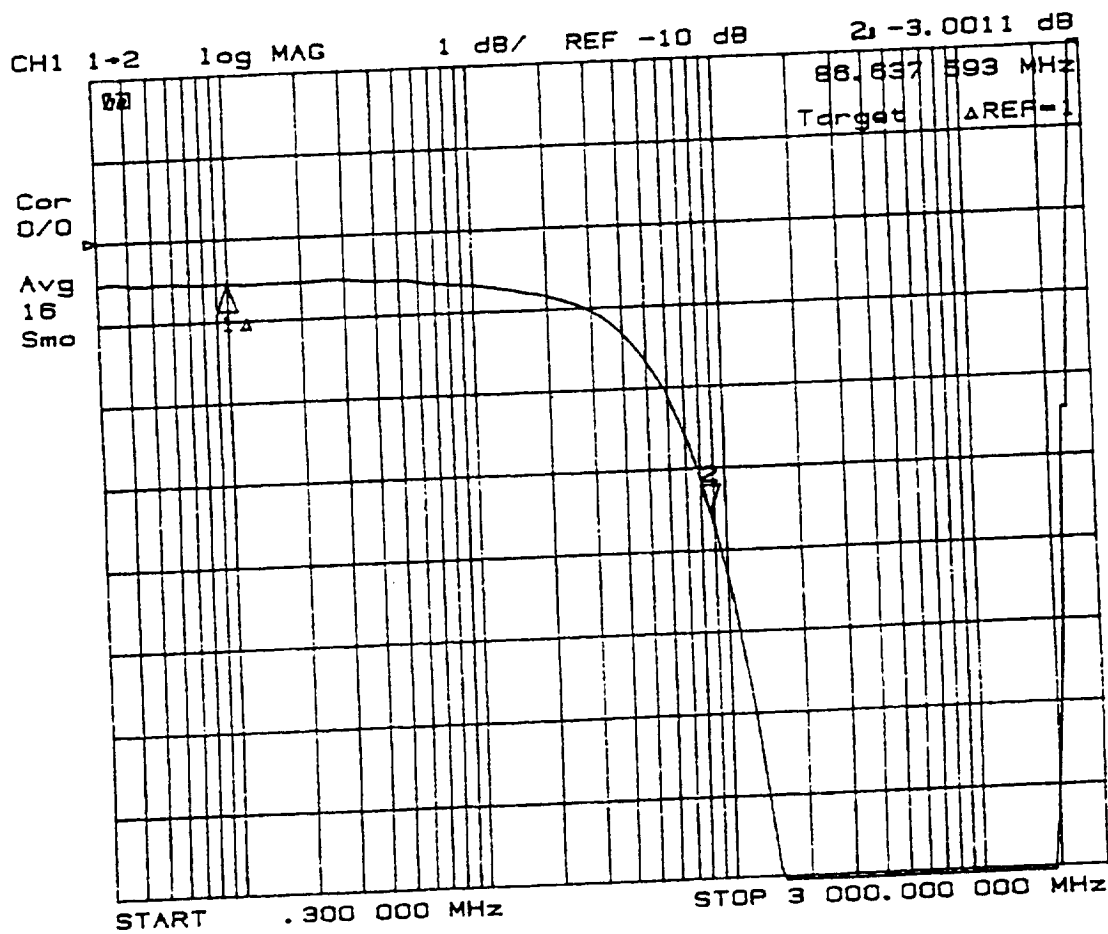


Figure 4-2. Sample Spectrum EDL Loop 7-8 1550nm LED E/O source

V

CONCLUSIONS

Significant results were achieved. The original data to be analyzed was found to be unusable. The test procedure was updated and refined. Using this new procedure, bandwidth measurements were accomplished on all multimode fiber test links terminating in the engineering development laboratory. These measurements show that the 1550 LED bandwidth distance product is $.42 \pm .01$ Ghz*km.

Further work should be done to sample bandwidth of the multimode fiber links in the 1550 nm window of the rest of the fiber plant at KSC using the new test procedure.

REFERENCES

¹Winzer and Reichelt. "Wavelength Division Multiplex Transmission over Multimode Optical Fibers: Comparison of Multiplexing Principles" Siemens Forsch.vol 9, number 4 (1980)

²N. A. Olsen et al., 68.3 km Transmission With 1.37 Tbit Capacity using Wavelength Division Multiplexing of Ten Single Frequency Lasers at 1.5 micrometers", Electronics Letters Vol. 21, Number 3 (January 1985)

³J. M. Senior et al. "Devices for wavelength multiplexing and demultiplexing", IEE Proceedings, Vol. 136, No. 3 (June 1989)



1993 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

N 9 4 - 2 8 2 3 2

1994018759

52-37

197188

P. 20

JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA

CONTROL OF A SERPENTINE MANIPULATOR WITH COLLISION AVOIDANCE

PREPARED BY:

Dr. Robert M. Byers

ACADEMIC RANK:

Assistant Professor

UNIVERSITY AND DEPARTMENT:

University of Central Florida
Department of Mechanical and
Aerospace Engineering

NASA/KSC

DIVISION:

Mechanical Engineering

BRANCH:

Special Projects

NASA COLLEAGUE:

Bill Jones
Gabor Tamasi
Todd Graham

DATE:

August 10, 1993

CONTRACT NUMBER:

University of Central Florida
NASA-NGT-60002 Supplement: 11

ACKNOWLEDGMENTS

The author gratefully acknowledges the support of the University of Central Florida and the NASA personnel who made the Summer Faculty Fellowship Program a productive and enjoyable experience. Especially noteworthy are the efforts of Dr. Ray Hosler and Kari Stiles, who rode herd on the Faculty Fellows and kept everything running smoothly. In addition, I wish to thank Gabor Tamasi, Bill Jones, and Todd Graham for their encouragement and patience. I look forward to working with them in the future.

ABSTRACT

The robotics lab at the Kennedy Space Center is investigating the possibility of using a "serpentine" manipulator for Shuttle inspection and payload processing. Serpentine manipulators are characterized by a large number of degrees of freedom giving them a high degree of redundancy. This redundancy allows them to be used to reach confined areas while avoiding collisions with their environment. In this paper, the author describes a new approach to controlling the joint rates for an n degree of freedom robot such that it moves its end effector to a desired position while simultaneously avoiding collision of any part of the robot arm with obstacles. Joint rates which move the end effector toward the target are found via a Lyapunov stability function. The gradient of an obstacle cost function indicates the direction toward obstacle collision in the joint space. The component of the end effector joint rates orthogonal to the obstacle gradient becomes the commanded joint rates. A notional eleven DOF model is used to numerically demonstrate the efficacy of the control law.

TABLE OF CONTENTS

- I. INTRODUCTION**
- II. MANIPULATOR KINEMATICS**
- III. INVERSE KINEMATICS**
- IV. LYAPUNOV STABILITY FOR END EFFECTOR TRAJECTORY**
- V. OBSTACLE AVOIDANCE**
- VI. LIMITATIONS ON JOINT RATES AND DEFLECTIONS**
- VII. CONCLUSIONS AND RECOMMENDATIONS**
- VIII. APPENDIX: *MATHEMATICA* PROGRAM FOR ROBOT SIMULATION**
- XI. REFERENCES**

I. INTRODUCTION

The range of motion achievable by a robot manipulator's end effector is a function of the number and type of joints or degrees of freedom it possesses. Any degrees of freedom in excess of the minimum number required to reach an arbitrary end effector position and orientation within the workspace are considered "redundant". Commercial manipulators typically possess six or fewer DOF for primarily "anthropomorphic" tasks such as industrial assembly and are therefore not redundant.

There are some tasks for which such standard manipulators are not well suited, such as those requiring an extended reach in a confined workspace. For that reason, so-called "serpentine" manipulators have attracted interest. Their designation and appearance (Fig. 1) suggest the long reach and dexterity associated with snakes or tentacles. They achieve this snake-like ability by possessing a high degree of redundancy. This redundancy allows them, theoretically, to "wriggle" an end effector into a confined or difficult to reach point while allowing the robot arm to be configured in such a way as to not contact the surrounding environment.

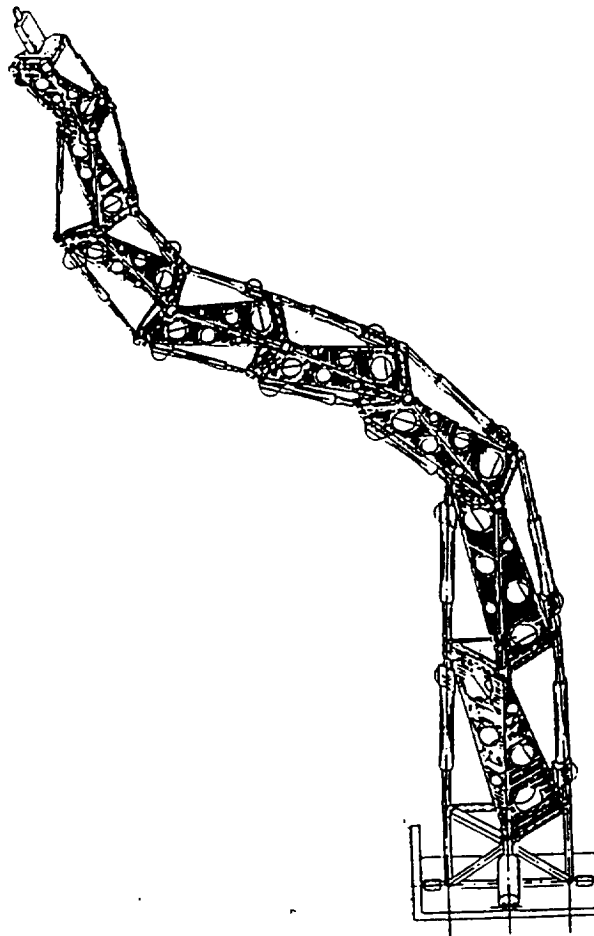


Figure 1
Serpentine Manipulator

One possible application is the inspection and processing of shuttle orbiter payloads in the Payload Changeout Room. During final launch preparations, tasks such as connecting/disconnecting umbilicals, removal of lens covers, or visual inspections must be carried out on Shuttle payloads. It is difficult and sometimes treacherous for technicians to see or reach many of the points at which these processes must take place and the payload itself may be put at risk. A serpentine robot is currently under development at KSC to study the feasibility for its use for such tasks [1].

There are two traditional approaches to controlling robot motion: to determine the dynamical equations of motion for each of the joints and generating the required torque for desired end-effector motion, or to control the joint velocities in response to the robot kinematics.

The complexity of serpentine motion, coupled with collision avoidance requirements, typically dictate relatively slow motion. This usually renders the dynamics of the robot arm negligible. Therefore, only the kinematics of the serpentine motion need be addressed.

Several approaches to controlling redundant manipulators for collision avoidance have been suggested. Maciejewski and Klein [2], Nakamura [3] and Wegerif, et al [4] make use of the pseudo-inverse and some variations of null-motion. Sciavicco, and Siciliano [5] make use of Lyapunov stability and an augmented configuration space to track a prescribed trajectory and incorporate obstacle avoidance. Alternatively, Pasch [1], and Asano [6] prescribe an end effector path and cause each joint to follow it in a "follow-the-leader" mode. All of these methods require that at least the end effector's trajectory and velocity be prescribed. This presumes that a clear path for the end effector is easily determined. Only [4] allows for the end effector to deviate from the prescribed path but only as an emergency measure.

In this paper, the author presents an alternative method for determining an acceptable robot trajectory which allows the end effector's path, as well as the entrained link's to be free to move around obstacles.

II. MANIPULATOR KINEMATICS

The position and orientation of the end effector $\underline{r} \in R^m$ is described, using the standard Denavit-Hartenberg convention as a function of the generalized vector of joint displacements $\underline{q} \in R^n$ for a manipulator with n joints.

$$\underline{r} = \underline{r}(\underline{q}(t), t) \quad (1)$$

At present, only the end effector's position is of interest, giving $m = 3$. The Denavit-Hartenberg convention, (described by Craig [7]), permits a single degree of freedom for each joint. The rotation matrix which relates the i th link frame to the $(i-1)$ th frame is given by:

$${}^{i-1}R = \begin{bmatrix} \cos \theta_i & -\sin \theta_i \cos \alpha_i & \sin \theta_i \sin \alpha_i \\ \sin \theta_i & \cos \theta_i \cos \alpha_i & -\cos \theta_i \sin \alpha_i \\ 0 & \sin \alpha_i & \cos \alpha_i \end{bmatrix} \quad (2)$$

The angle α_i reflects the rotation of the i th joint frame about the local x axis with respect to the $(i-1)$ th frame. The angle θ_i is the rotation angle of the i th joint and corresponds to a component of the vector \underline{q} . The end of the i th link is located by the vector

$$\underline{r}_i = \underline{r}_{i-1} + {}^{i-1}R \begin{bmatrix} l_i & 0 & 0 \end{bmatrix}^T \quad (3)$$

so that the end effector is located by:

$$\underline{r} = \sum_{i=0}^n \underline{r}_i \quad (4)$$

Serpentine manipulators are characterized by a high degree of redundancy and by a distinctive structure. Revolute joints are essentially universal joints possessing two degrees of freedom. This permits any link to have an arbitrary direction in space. Two degrees of freedom are modeled in the D-H convention by incorporating links of zero length. Following the D-H convention, the link frames are alternately rotated about the link $\pm x$ axes with every other link having zero length. Equation (2) becomes:

$$\alpha_i = \begin{cases} -\pi/2, & i \text{ even} \\ \pi/2, & i \text{ odd} \end{cases} \Rightarrow \begin{matrix} \text{odd} \\ \text{even} \end{matrix} R = \begin{bmatrix} \cos \theta_i & 0 & -\sin \theta_i \\ \sin \theta_i & 0 & \cos \theta_i \\ 0 & -1 & 0 \end{bmatrix}, \begin{matrix} \text{even} \\ \text{odd} \end{matrix} R = \begin{bmatrix} \cos \theta_i & 0 & \sin \theta_i \\ \sin \theta_i & 0 & -\cos \theta_i \\ 0 & 1 & 0 \end{bmatrix} \quad (5)$$

The location of the link ends and the end effector is then found by Eq. (3). In this paper, for simulation purposes, a notional 11 DOF manipulator is used. It consists of five links each of unit length, each with 2 DOF and a base that can translate in the base frame x direction. Numerical simulations and their graphical presentation were facilitated by *Mathematica* [8], a symbolic mathematics software package.

III. INVERSE KINEMATICS

The velocity of the end effector is calculated by :

$$\frac{dr}{dt} = \left(\frac{\partial r}{\partial \underline{q}} \right) \frac{dq}{dt} \quad (6)$$

where the coefficient matrix, $(\partial r / \partial \underline{q}) = J$, is the *Jacobian* matrix.

To achieve a desired trajectory for the end effector, when $n > 3$, the differential equation for the generalized joint displacements may be solved:

$$\dot{\underline{q}} = J^* \dot{\underline{r}} \quad (7)$$

where J^* is the Moore-Penrose pseudo inverse given by:

$$J^* = J^T (JJ^T)^{-1} \quad (8)$$

Eq. (7) yields the minimum norm solution for $\dot{\underline{q}}$. This solution assumes an appropriate scaling metric as discussed by Doty, et al [9]. In fact, because in a redundant manipulator, $n > m$, there are infinitely many solutions to Eq. (6). The joint rates are a function of those rates which cause the end effector to move, $\dot{\underline{q}}_R$ and "null rates", $\dot{\underline{q}}_N$ which do not. That is, $\dot{\underline{q}} = \dot{\underline{q}}_R + \dot{\underline{q}}_N$ where $\dot{\underline{q}}_R$ is the minimum norm motion given by Eq. (7) and $\dot{\underline{q}}_N$ is given by:

$$\dot{\underline{q}}_N = (E_n - J^* J) \underline{\mu} \quad (9)$$

where E_n is the $n \times n$ identity matrix and $\underline{\mu} \in R^n$ is an arbitrary vector. It is a simple matter to confirm that $\dot{\underline{q}}_R$ and $\dot{\underline{q}}_N$ are orthogonal vectors by taking their inner product. Alternatively, multiplying Eq. (8) by J yields the null vector.

The selection of $\underline{\mu}$ generates one of an infinite number of joint rate combinations which move all of the links but do not cause motion in the end effector. There have been several control laws suggested which make use of the null motion to avoid obstacles. Several of these are recapped by Nakamura [3] in some detail. A major shortcoming of these methods is the requirement to prescribe the end effector path and velocity. Not only can this be a difficult task in itself for a complex workspace, but, in some situations, it proscribes joint motion which could avoid collisions.

IV. LYAPUNOV STABILITY APPROACH TO END EFFECTOR TRAJECTORY

In contrast to the pseudo-inverse approach, the author has adopted a Lyapunov stability measure similar to that in [5]. However, whereas the end effector in [5] tracks a prescribed end effector path, here only the final end effector position is required. The error vector \underline{e} is defined:

$$\underline{e} = \underline{r}_T - \underline{r} \quad (10)$$

where $\underline{r}_T \in R^3$ is the vector locating the target with respect to the robot base frame. For a fixed target

$$\dot{\underline{e}} = -\dot{\underline{r}} \quad (11)$$

The Lyapunov scalar function v is defined

$$v = \frac{1}{2} \underline{e}^T \underline{e} \quad (12)$$

Because v is a positive scalar (related to the error's magnitude squared), then, if the time derivative of v is negative, \underline{e} will go to zero as time approaches infinity. Taking the time derivative gives:

$$\begin{aligned} \dot{v} &= \underline{e}^T \dot{\underline{e}} \\ &= -\underline{e}^T (J \dot{\underline{q}}) \end{aligned} \quad (13)$$

An obvious selection for $\dot{\underline{q}}_R$ is to make its elements proportional to the elements of $J^T \underline{e}$. One such solution is given by

$$\dot{\underline{q}}_R = M \left(\frac{J^T \underline{e}}{\|J^T \underline{e}\|} \right) \quad (14)$$

where M is a positive definite matrix of dimension $n \times n$. The computational simplicity of Eq. (14) contrasts starkly with the complexity of computing the motion using the pseudo-inverse approach in Eq. (7). In addition, since $\dot{\underline{q}}_R$ is a unit vector scaled by M and no matrices must be inverted, the control law works well even in the vicinity of joint singularities.

The importance of an appropriate metric must be emphasized. Doty, et al [9] show that results may be obtained which are non-invariant with respect to choice of

reference frame or dimensional units used to describe the problem. In Eq. (14) $\dot{\underline{q}}_R$ has the dimensions of *radians/time* for revolute joints and *length/time* for prismatic joints. On the right hand side of Eq. (14), J has dimensions of *length/radians* and *length/length* for revolute and prismatic joints, respectively, while the error vector has the units of *length*.

Clearly the units of $J^T \underline{e}$ are incompatible with $\dot{\underline{q}}_R$. Normalizing $J^T \underline{e}$ as a unit vector renders it dimensionless and the matrix M serves to give the appropriate unit metric. In addition, the selection of the magnitude of the elements of M may be used to emphasize the motion of some joints over that of others. This aspect will be discussed later as a means of avoiding excessive joint rates or deflections.

Figure 2 shows three simulations of robot motion for different values of a diagonal matrix M . For all the maneuvers, the manipulator has an initial end effector position $\underline{r}(0)=[3, 2, 0]$, shown in Fig. 2(a), and moves to a final end effector position of $\underline{r}(t_f)=[-3, 2, 0]$, indicated by the dot in the upper left portion of the workspace. In Fig. 2(b) the final configuration and the end effector path are shown for M equal to the identity matrix. That is, all of the joint rates are equally weighted. The end effector trajectory resembles a damped sinusoid.

In Fig. 2(c) the motion of the last two revolute joints is given a weight of ten times greater than the other nine joints. This causes the manipulator to attempt to reach the target primarily by moving these two joints, which at one point causes a near singularity. This is evidenced by the abrupt direction change of the end effector. Because of a fairly large step size in the *Mathematica* program, the final conditions are not satisfied exactly.

In Fig. 2(d) the motion of the translating base and the first revolute joint are emphasized by a factor of ten. The motion to the target is accomplished almost exclusively by the motion of these two joints.

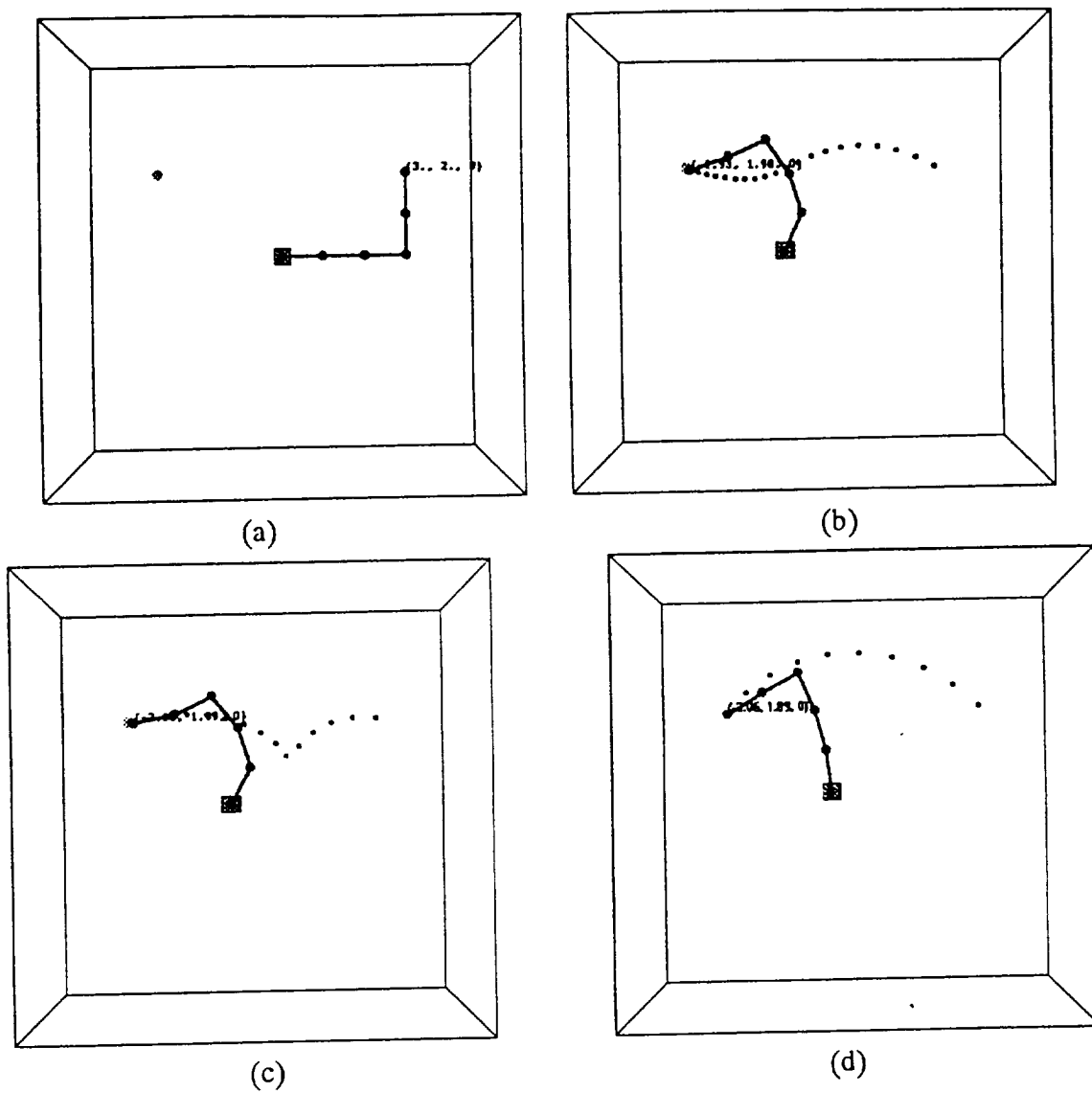


Figure 2.
End Effector Trajectory Determined By Lyapunov Function
(a) Initial Configuration
(b) Final Configuration with Equal Weights on Joint Rates
(c) Final Configuration with Last Two Joint Rates Emphasized
(d) Final Configuration with First Two Joint Rates Emphasized

V. OBSTACLE AVOIDANCE

As noted before, most other proposed methods of obstacle avoidance presume a prescribed end effector path that is obstacle free. This can be an important constraint because it may require detailed knowledge of the work space or excessively complicated path planing. By allowing the effector to seek its own path, the overall manipulator configuration becomes much more robust in its ability to avoid obstacles.

As suggested by Khatib[10], each obstacle is assigned a cost function. Figure 3 shows a representative manipulator arm with obstacle avoidance points p_i , $i = 1, \dots, n_p$, where $p_i = [x_i \ y_i \ z_i]$, identified along it. A likely location for such points would be the manipulator joints and the link mid points but they may be dictated by sensor location or other criteria. In this paper, obstacles are assumed to be rectangular parallelepipeds with their center points o_j , and with dimensions $2a_j$, $2b_j$, $2c_j$, $j=1, \dots, n_o$.

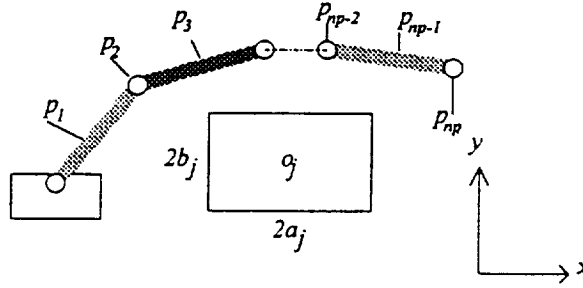


Figure 3.
Typical Designated Obstacle Avoidance Points

The cost function for the j_{th} obstacle with respect to the i_{th} obstacle avoidance point is given by the super-ellipsoid

$$C_j(p_i) = \left(\frac{x_i - x_j}{a_j} \right)^8 + \left(\frac{y_i - y_j}{b_j} \right)^8 + \left(\frac{z_i - z_j}{c_j} \right)^8 \quad (15)$$

Contact with the surface of the obstacle by the i_{th} point is approximated by $C_j(p_i) = 1$. A potential function is defined by:

$$P = \sum_j^{n_o} \sum_i^{n_p} [C_j(p_i) - 1]^{-1} \quad (16)$$

which guarantees that the cost becomes infinite before actual contact is made with the obstacle. The gradient of the potential function is

$$\underline{\mu} = \frac{\partial P}{\partial \underline{q}} \quad (17)$$

The time rate of change of P can thus be expressed

$$\frac{dP}{dt} = \underline{\mu}^T \dot{\underline{q}} \quad (18)$$

As one might expect, joint rates generated by Eq. (14) may also adversely change the proximity to obstacles. However, if the only requirement on the joint rates is that the error vector be decreased over every sub-interval, then an infinite number of trajectories may be found which accomplish this. Assuming that at least one unobstructed trajectory exists, one possible solution is to find the component of $\dot{\underline{q}}_R$ from Eq. (14) which is orthogonal to $\underline{\mu}$. This results in $\dot{P} = 0$. The Gram-Schmidt procedure described by Luenberger [11], subtracts from $\dot{\underline{q}}_R$ its projection in the $\underline{\mu}$ direction. This results in the commanded joint rates:

$$\dot{\underline{q}} = \dot{\underline{q}}_R - \left[\left(\dot{\underline{q}}_R \right)^T \hat{\underline{\mu}} \right] \hat{\underline{\mu}} \quad (19)$$

where $\hat{\underline{\mu}}$ is a unit vector parallel to $\underline{\mu}$. A three dimensional analogy is shown in Figure 4. Moving toward a destination, a traveler's path intersects a portion of a hill. The traveler's location is analogous to the robot's current joint configuration; the hill is an obstacle. Going uphill in the direction of the gradient, increases the cost. By moving along a contour line of constant cost, orthogonal to the slope, the traveler may simultaneously move closer to his destination without going uphill. Eventually a point is reached where the target is downhill.

In the three dimensional analogous state, it is easy to see that if the destination lies exactly opposite of the summit from the present position, forward motion eventually becomes impossible (the cost becomes infinite). Motion stops for the robot in the event that no motion whatever will move the end effector closer to the target without colliding with an obstacle. In the n dimensional joint space however, this possibility recedes as n becomes large. That is, although this algorithm does not explicitly require redundancy, redundancy increases its robustness.

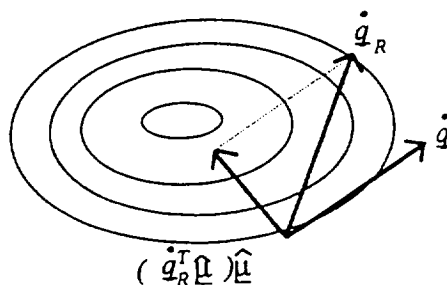


Figure 4.
Three Dimensional Analogy of Obstacle Avoidance

The control embodied in Eqs. (14) and (19) is demonstrated in the simulation shown in Figure 5. Because the interest in serpentine manipulators is largely due to the potential for reaching targets in constricted areas, in this example, the target is located "down the hall and around the corner" with the walls modeled by three obstacles. A collision free trajectory is generated by the method described above. Once again the end effector trajectory is non-intuitive.

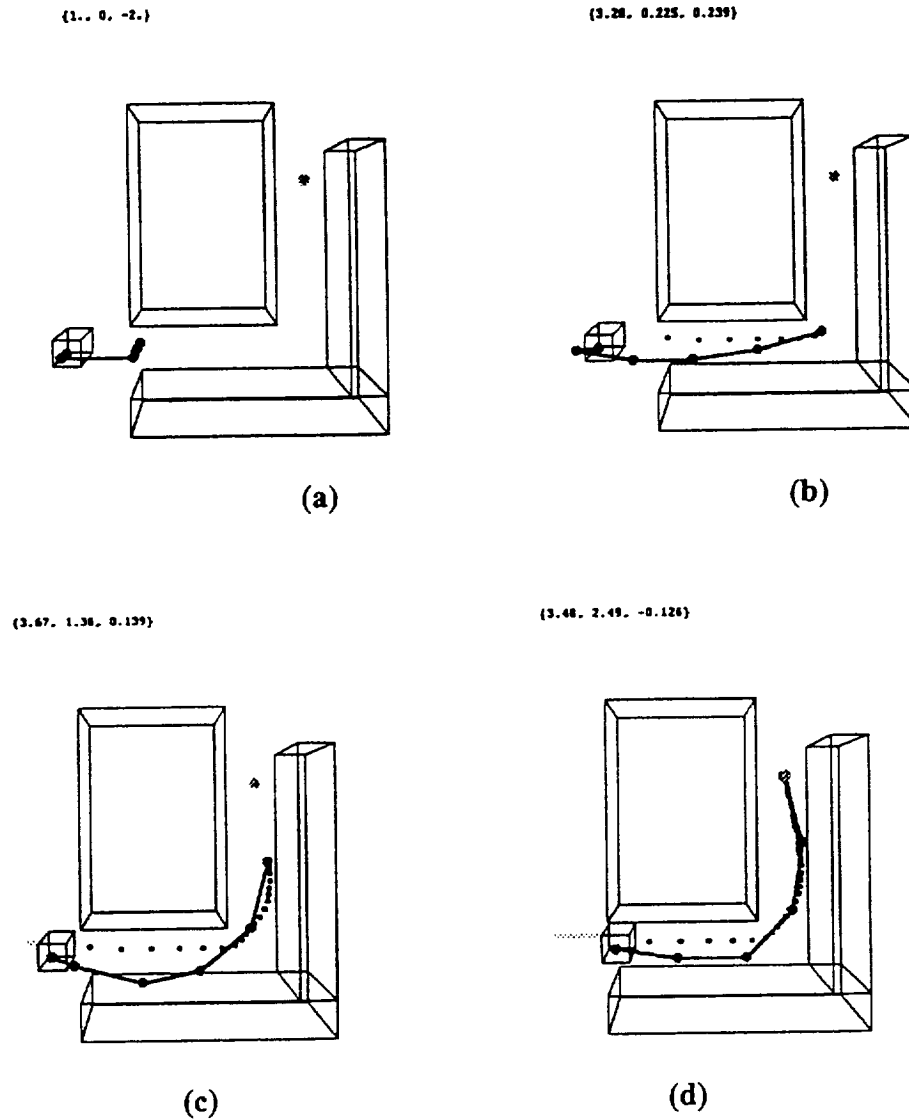


Figure 5
Obstacle Avoidance Trajectory
 (a) Initial Configuration
 (b),(c) Intermediate Configurations
 (d) Final Configuration

VI. LIMITATIONS ON JOINT RATES AND DEFLECTIONS

In addition to avoiding obstacles, manipulator arms are frequently limited in the magnitude of the joint deflections which can be achieved. In addition, the joint rates are usually limited by the manipulator architecture. While not explicitly examined in the author's research, some possible solutions are suggested.

It has been demonstrated that the joint rates can be influenced by the weighting of the elements of the matrix M . The most straightforward approach is to weight each joint rate independently by making M diagonal. The weight on each joint rate may be made a function of its current deflection and commanded joint rate.

It is useful to think of the diagonal elements of M as the stiffness coefficients of n non-linear spring. The deflection of the i th joint is bounded by $q_{i\min} \leq q_i \leq q_{i\max}$. Defining:

$$\begin{aligned}\Delta_i &= q_{i\max} - q_{i\min} \\ \Gamma_i &= q_{i\max} + q_{i\min} \\ f_i &= \frac{2q_i - \Gamma_i}{\Delta_i} \\ \eta_i &= \text{sign}[(J^T \underline{e})_i]\end{aligned}\tag{20}$$

The diagonal elements of M are defined:

$$m_{ii} = k_i(1 - \eta_i f_i), \quad i = 1, \dots, n\tag{21}$$

where $k_i \leq \dot{q}_{i\max} / 2$.

This function guarantees that the maximum allowable joint rate is never exceeded and motion away from the maximum deflections is encouraged while motion toward maximum deflection is discourage. This approach has not been implemented in any of the examples in this paper and requires further verification.

VII. CONCLUSIONS AND RECOMMENDATIONS

A method for moving a serpentine manipulator's end effector to a target in a constricted area while avoiding collision's of the manipulator's arm with the surrounding workspace has been demonstrated. This method has the virtues of being computationally straightforward. It is robust in the vicinity of singularities and multiple obstacles. While it must be emphasized that this paper discusses only very preliminary results, the algorithm appears to have great potential for successful implementation for achieving numerous robot tasks.

Although the algorithm appears to be fairly versatile, the ability of the manipulator to reach a target can be sensitive to its initial configuration relative to the target. While path planning is not explicitly required, it is necessary to orient the robot with respect to the target so that a likely path is unambiguous. In addition, it has been observed that there are a number of cases in which the robot will not be able to reach the target. It is possible for the end effector to arrive a point where no further forward motion is possible. This is the case where a wide, flat obstacle is approached and only moving away from the target will eventually result in a configuration in which forward motion. Additional heuristics need to be developed to address this possibility.

The scaling matrix requires further research. While the suggested configuration works adequately, there is room for improvement. Further research is especially necessary in properly scaling the control vector in order to avoid joint rate and deflection limitations. These limitations, although addressed in this paper, should be further investigated in the context of a realistic robot architecture.

Finally, although *Mathematica* is a versatile programming tool, it is far too slow for practical numerical integration. For implementation on an actual robot, programming in *C* language is recommended. *Mathematica* may be linked to a *C* code to evaluate symbolically complicated expressions such as the Jacobian matrix or the obstacle gradient vector.

VIII. APPENDIX: MATHEMATICA PROGRAM FOR ROBOT SIMULATION

Mathematica was used to produce the simulations in this paper. The parameters in the following program are those for the simulation depicted in Fig. 5. *Mathematica* commands are shown in Courier font. Explanatory comments have been added in Times Roman font.

```

ClearAll[x, theta, al, p, r0, J, T0];
Clear[n, d, obsnum];
JacobianMatrix[funcs_List, vars_List]:=Outer[D, funcs, vars];
Norm[vars_List]:=Sqrt[vars.vars];
UnitVector[vars_List]:=vars/Sqrt[vars.vars];
al:=Array[alpha, n]; (*Vector of frame rotation angles*)
x:=Append[Array[theta, n], d]; (*Generalized Vector of joint displacements.
The translational displacement of the base is given by d.*)
p:=Array[l, n]; (*Vector of link lengths*)

(*R[n] defines the rotation matrix relating the nth link frame to the (n-1)th frame using
standard Denavitt-Hartenberg convention *)

R[n_]:=({Cos[x[[n]]], -Cos[al[[n]] Sin[x[[n]]], Sin[al[[n]]]
Sin[x[[n]]},
      {Sin[x[[n]]], Cos[al[[n]] Cos[x[[n]]], -Sin[al[[n]]]
Cos[x[[n]]},
      {0, Sin[al[[n]]], Cos[al[[n]]}});

(*T0[n] gives the orientation of the nth link with respect to the base frame*)
T0[n_]:=T0[n]=T0[n-1].R[n];
T0[0]=IdentityMatrix[3];
R[0]=IdentityMatrix[3];
r0[n_]:=r0[n]=r0[n-1]+T0[n].{p[[n]], 0, 0}; (*Endpoint of nth link*)

mp[n_]:=mp[n]=r0[n-1]+T0[n].{p[[n]]/2, 0, 0}; (*Midpoint of nth link*)
r0[0]={d, 0, 0};
mp[0]={d, 0, 0};
(*Manipulator architecture is defined by n, al, and p *)
n=10;
al={Pi/2, -Pi/2, Pi/2, -Pi/2, Pi/2, -Pi/2, Pi/2, -Pi/2, Pi/2, -Pi/2};
p={0, 1, 0, 1, 0, 1, 0, 1, 0, 1};
joints=Table[Point[r0[i]], {i, 0, n}]; (*Table of joint coordinates; for
plotting purposes*)
arm=Line[Table[r0[i], {i, 0, n}]]; (Line from joint to joint; for plotting
purposes*)
J=JacobianMatrix[r0[n], x];

```

(*Obstacles are depicted as rectangular solids with six coordinates: first three values are coordinates of mass center. Second three values are x,y,z dimensions*)

```
obsnum=3;
```

```
ob[1]={2,2,-.5,2,3,5};
```

```
ob[2]={2.75,-.75,-.5,3.5,.5,5};
```

```
ob[3]={4.25,1.25,-.5,.5,3.5,5};
```

(*obstacleshape is a function which draws a rectangular solid in the plot to represent each obstacle*)

```
obstacleshape[k_]:=
```

```
Cuboid[{(ob[k][[1]]-ob[k][[4]]/2),
```

```
(ob[k][[2]]-ob[k][[5]]/2), (ob[k][[3]]-ob[k][[6]]/2)},
```

```
{(ob[k][[1]]+ob[k][[4]]/2),
```

```
(ob[k][[2]]+ob[k][[5]]/2), (ob[k][[3]]+ob[k][[6]]/2)}];
```

(*cost is the potential function*)

```
cost= (Sum[Sum[
```

```
1/(Sum[{(mp[i][[j]]-ob[k][[j]])/(ob[k][[j+3]]/2)}^8,
```

```
{j,1,3}]-1.1), {i,0,n}]+
```

```
1/(Sum[{(r0[n][[j]]-
```

```
ob[k][[j]])/(ob[k][[j+3]]/2)}^8, {j,1,3}]-1.1),
```

```
{k,1,obsnum}]);
```

(* mu is the obstacle gradient vector*)

```
mu=Table[D[cost,x[[i]]], {i,1,n+1}];
```

```
step=.05;(*step size*)
```

```
imax=100;(*maximum number of steps*)
```

```
target={3.5,2.5,0};
```

```
d=0;
```

```
theta[1]=0;
```

```
theta[2]=N[Pi/2];
```

```
theta[3]=0;
```

```
theta[4]=-N[Pi/2];
```

```
theta[5]=0;
```

```
theta[6]=-N[Pi/2];
```

```
theta[7]=0;
```

```
theta[8]=0;
```

```
theta[9]=0;
```

```
theta[10]=0;
```

```
pts={Point[r0[n]]};
```

```
base={Point[{d,0,-3]}];
```

```
Metric=DiagonalMatrix[1,1,1,1,1,1,1,1,1,1];
```

(* Robotsim is a function which performs the actual simulation. It is largely devoted to drawing the graphic images of the robot motion. It is started by compiling the program and then typing "Robotsim"*)

```
Robotsim:=
For[i=0,i<=imax,i++,
If[EvenQ[i], path=Append[pts,Point[r0[n]]];
track=Append[base,Point[{d,0,-3}]];
Show[
Graphics3D[({AbsoluteThickness[2],arm,
Cuboid[{d-.2,-.2,-3},{d+.2,.2,0]}},
{RGBColor[1,0,0],PointSize[.02],joints},
{RGBColor[0,0,1],AbsoluteThickness[1],Table[path]},
{RGBColor[1,1,0],AbsoluteThickness[1],Table[track]},
{RGBColor[0,1,0],PointSize[.02],Point[target]},
{Table[obstacleshape[i],{i,1,obsnum}]},
{Text[NumberForm[r0[n],3],{0,5,0},{-1,1}]}},
Boxed->False, ViewPoint->{0.000,0.000,3.384},PlotRange->{{-
1,5},{-1,5},{-3,5}}];
pts=path;
base=track;
Print[i];
If[Sqrt[(target-r0[n]).(target-r0[n])]<.001,i=imax];
(* The last six lines of code perform a fairly crude numerical integration with a first order
Euler's method*)
vr=UnitVector[(target-r0[n]).J.Metric];
nu=UnitVector[mu];
v=UnitVector[vr- vr.nu nu];
For[j=1,j<=n,j++,
new[j]=theta[j]+step v[[j]];theta[j]=new[j]];
new[n+1]=d+step v[[n+1]];d=new[n+1]]];
```

IX. REFERENCES

- [1] Pasch, K., "Self-Contained Deployable Serpentine Truss for Pre launch Access of Space Shuttle Orbiter Payloads", NAS -2659-FM-9106-387, Final Report, Contract No. NAS10-11659. NASA, Kennedy Space Center, FL.
- [2] Maciejewski, A., and Klein, C., "Obstacle Avoidance for Kinematically Redundant Manipulators in Dynamically Varying Environments", *The International Journal for Robotics Research*, Vol. 4, No. 3, Fall 1985, pp. 109-117.
- [3] Nakamura, Y., *Advanced Robotics, Redundancy and Optimization*, Addison-Wesley, Publishing Co., Inc., Redwood City, CA, 1991.
- [4] Wegerif, D., Rosinski, D., and Parton, W., "Results of Proximity Sensing Research for Real-Time Collision Avoidance of Articulated Robots Working Near the Space Shuttle", *Proceedings of the 6th Annual Conference on Recent Advances in Robotics*, University of Florida, Gainesville, FL, 19-20 April 1993.
- [5] Sciavicco, L., and Siciliano, B., "A Solution Algorithm to the Inverse Kinematic Problem for Redundant Manipulator", *IEEE Journal of Robotics and Automation*, Vol. 4, No. 4, Aug. 1988, pp. 403-410.
- [6] Asano, K., et al, "Multijoint Inspection Robot", *IEEE Transactions on Industrial Electronics*, Vol. IE-30, No. 3, August 1983, pp. 277-281.
- [7] Craig, J., *Introduction to Robotics*, 2nd Ed., Addison-Wesley, Publishing Co., Inc., Redwood City, CA, 1989.
- [8] Wolfram, S., *Mathematica*, 2nd Ed., Addison-Wesley Publishing Co., Inc., Redwood City, CA, 1991.
- [9] Doty, K., Melchiorri, C., & Bonivento, C., "A Theory of Generalized Inverses Applied to Robotics", *The International Journal of Robotics Research*, Vol. 12., No. 1, Feb. 1993, pp. 1-19.
- [10] Khatib, O. and Le Maitre, J.-F., "Dynamic Control of Manipulators Operating in a Complex Environment", *Proc. 3rd Int. CISM-IFTOMM Symp.*, pp. 267-282.
- [11] Luenberger, D., *Optimization by Vector Space Methods*, John Wiley & Sons, Inc., New York, 1969.
- [12] Nakamura, Y., and Hanafusa, H., "Optimal Redundancy Control of Robot Manipulators", *International Journal of Robotics Research*, Vol. 6, No. 1., Spring 1987, pp. 32-42.

1994018760

N94-28233

1993 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA

53-25

197189

P-30

EVALUATION OF INORGANIC ZINC-RICH PRIMERS USING
ELECTROCHEMICAL IMPEDANCE SPECTROSCOPY (EIS)

PREPARED BY:	Dr. Luz M. Calle
ACADEMIC RANK:	Associate Professor
UNIVERSITY AND DEPARTMENT:	Randolph-Macon Woman's College Department of Chemistry
NASA/KSC	
DIVISION:	Material Science
BRANCH:	Failure Analysis and Materials Evaluation
NASA COLLEAGUE:	Louis G. MacDowell III
DATE:	August 13, 1993
CONTRACT NUMBER:	University of Central Florida NASA-NGT-60002 Supplement: 11

ACKNOWLEDGMENT

I would like to express my appreciation to NASA/ASEE for providing me with the wonderful opportunity to spend this summer engaged in what I consider a fascinating area of research. The enthusiasm and professionalism of E. R. Hosler, the program director, and of Kari Styles, the administrative assistant, made participating in the program a very pleasurable and rewarding experience. I would also like to thank my NASA colleague, Louis G. MacDowell, III, for his expert guidance and support throughout the duration of the project. I am also indebted to Irby Moore, Cole Bryan, and specially , Scott Murray and the rest of the materials section for making me feel welcome and for always being there when I needed help. The experimental part of this project could not have been accomplished without the cooperation of Hal Novak from USBI who let me borrow his lock-in amplifier when the one I was using had to be sent away for repair. I am also grateful to Brian Fordham, a Sharp student at KSC this summer, for his technical assistance. Last, though by no means least, I wish to thank my husband Carlos for patiently coping with the inconveniences of moving to Florida with me during this summer.

ABSTRACT

This investigation explores the use of Electrochemical Impedance spectroscopy (EIS) in combination with beach exposure as a short term method for analyzing the performance of twenty-one zinc-rich primers. The twenty-one zinc-rich primers were: Carboline CZ-11, Ameron Devoe-Marine Catha-Coat 304, Briner V-65, Ameron D-21-9, Sherwin Williams Zinc Clad II, Carboline CZ-D7, Ameron D-4, Dupont Ganicin 347WB, Porter TQ-4374H, Inorganic Coatings IC-531, Subox Galvanox IV, Southern Coatings Chemtec 600, Glidden Glidzinc 5530, Byco SP-101, Tnemec 90E-75, Devoe Catha-Coat 302H, Glidden Glidzinc 5536, Koppers 701, Ameron D-21-5, Coronado 935-152, and Subox Galvanox V. Data were also collected on galvanized steel for comparison purposes. A library of Bode magnitude plots was generated for each coating including curves for the initial time and after each week of atmospheric exposure at the Beach Corrosion Test Site near the Space Shuttle launch pad at the Kennedy Space Center for up to three weeks. An examination of the variation of the Bode magnitude plots with atmospheric exposure revealed no clearly identifiable trend at this point that could distinguish between the good and the poor coatings. The test will be continued by including EIS measurements after six months and one year of atmospheric exposure.

SUMMARY

This investigation explores the use of Electrochemical Impedance spectroscopy (EIS) in combination with beach exposure as a short term method for analyzing the performance of twenty-one zinc-rich primers. The twenty-one zinc-rich primers were: Carboline CZ-11, Ameron Devoe-Marine Catha-Coat 304, Briner V-65, Ameron D-21-9, Sherwin Williams Zinc Clad II, Carboline CZ-D7, Ameron D-4, Dupont Ganicin 347WB, Porter TQ-4374H, Inorganic Coatings IC-531, Subox Galvanox IV, Southern Coatings Chemtec 600, Glidden Glidzinc 5530, Byco SP-101, Tnemec 90E-75, Devoe Catha-Coat 302H, Glidden Glidzinc 5536, Koppers 701, Ameron D-21-5, Coronado 935-152, and Subox Galvanox V. Data were also collected on galvanized steel for comparison purposes.

Electrochemical Measurements were performed on 4-inch x 6-inch x 1/8 inch test panels. The panel included in this investigation had been rated previously to determine their degree of corrosion performance on a scale of 1 to 10, with 10 being the highest rating. The twenty-one primers chosen represent a wide range of performance. AC impedance data were collected after one hour immersion time in 3.55% NaCl. Data were gathered in the frequency range from 100 kHz to 0.01 Hz.

A library of Bode magnitude plots was generated for each coating including curves for the initial time and after each week of atmospheric exposure for up to three weeks. An examination of the variation of the Bode magnitude plots with atmospheric exposure revealed no clearly identifiable trend at this point that could distinguish between the good and the poor coatings. The test will be continued by including EIS measurements after six months and after one year of atmospheric exposure.

TABLE OF CONTENTS

<u>Section</u>	<u>Title</u>	<u>Page</u>
I.	INTRODUCTION	7
II.	MATERIALS AND METHODS	8
III.	RESULTS AND DISCUSSION	12
IV.	REFERENCES	29

LIST OF ILLUSTRATIONS

<u>Figure</u>	<u>Title</u>	<u>Page</u>
1	Exploded View of Model K0235 Flat Cell	14
2	Corrosion Potential in Volts for Coatings A-H	15
3	Corrosion Potential in Volts for Coatings I-O	15
4	Corrosion Potential in Volts for Coatings P-U and Galvanized Steel	17
5	Bode Magnitude Plot for Carboline CZ-11	18
6	Bode Magnitude Plot for Ameron Devoe-Marine Catha-Coat 304	18
7	Bode Magnitude Plot for Briner V-65	19
8	Bode Magnitude Plot for Ameron D-21-9	19
9	Bode Magnitude Plot for Sherwin Williams Zinc Clad II	20
10	Bode Magnitude Plot for Carboline CZ-D7	20
11	Bode Magnitude Plot for Ameron D-4	21
12	Bode Magnitude Plot for Dupont Ganicin	21
13	Bode Magnitude Plot for Porter TQ-4374H	22
14	Bode Magnitude Plot for Inorganic Coatings IC-531	22
15	Bode Magnitude Plot for Subox Galvanox IV	23
16	Bode Magnitude Plot for Southern Coatings Chemtec 600	23
17	Bode Magnitude Plot for Glidden Glidzinc 5530	24
18	Bode Magnitude Plot for Byco SP-101	24
19	Bode Magnitude Plot for Tnemec 90E-75	25
20	Bode Magnitude Plot for Devoe Catha-Coat 302H	25
21	Bode Magnitude Plot for Glidden Glidzinc 5536	26
22	Bode Magnitude Plot for Koppers 701	26
23	Bode Magnitude Plot for Ameron D-21-5	27
24	Bode Magnitude Plot for Coronado 935-152	27
25	Bode Magnitude Plot for Subox Galvanox V	28
26	Bode Magnitude Plot for Galvanized steel	28

LIST OF TABLES

<u>Table</u>	<u>Title</u>	<u>Page</u>
1	Materials Tested, Code, Type, Rust Grade Evaluation, and Zinc Content, and Thickness	9
2	Weather Conditions During the First Three Weeks of Exposure	10
3	Corrosion Potential in Volts	13

I. INTRODUCTION

For over 25 years, the process of coating testing at the Kennedy Space Center (KSC) to identify materials for the long term protection of carbon steel structures has taken place at the Beach Corrosion Test Site near the launch pad. The atmosphere at the launch pad is highly corrosive due to the proximity of the ocean, high heat from rocket exhaust, and now with the Space Shuttle, the acidic combustion products of the Solid Rocket Boosters (SRB's). As a result of previous tests, zinc-rich primers were identified as the best choice to provide long term protection of launch structures and ground support equipment at KSC^{1,2,3}.

Exposure at the beach site provides very valuable data but it takes a long time. Testing requires 18 months for preliminary approval and continued good performance for 5 years for final approval. This investigation explores the use of Electrochemical Impedance spectroscopy (EIS) in combination with beach exposure as a short term methods for analyzing the performance of zinc-rich primers. EIS has been used before on bare metals⁴.and on zinc-rich primers⁵. The samples tested were flat, 0.16 cm (1/16 inch) thick coupons, 1.59 cm (5/8 inch) in diameter and did not include atmospheric exposure.

II. MATERIALS AND METHODS

A model 378 Electrochemical Impedance system manufactured by EG&G Princeton Applied Research Corporation was used for all electrochemical measurements. Electrochemical Measurements were performed on 4-inch x 6-inch x 1/8 inch test panels. Twenty-one panels were coated with the inorganic zinc-rich primers listed in Table 1. The panel rating corresponds to the degree of corrosion judged on a scale of 1 to 10, with 10 being the highest rating⁶. The 21 primers chosen represent a wide range of performance. Also listed on the table is the zinc content, and the thickness when available.

The panels were mounted on an EG&G Princeton Applied Research Model KO235 Flat Cell with a working electrode area of 1 cm². The flat cell (see Figure 1) consists of a flat cylinder clamped horizontally between two end plates. One end plate houses the working electrode and the other houses the platinum counter electrode. The reference silver/silver chloride electrode is housed in a Luggin well with a Luggin capillary protruding from the bottom of the well.

AC impedance data were collected after one hour immersion time in aerated 3.55% NaCl. Data were gathered in the frequency range from 100 kHz to 0.01 Hz using the Auto Execute option of the M388 electrochemical analysis software. Three experiments were performed in a sequence covering the specified frequency range and the data were automatically merged and saved. The frequency ranges for the three experiments were 100 kHz to 5 Hz, 10 Hz to 0.1 Hz, and 0.1 Hz to 0.01 Hz. The ac amplitude was 10 mV. After each set of measurements, the panels were mounted on a rack and taken to the beach corrosion test site and left there for a week. At the end of the week, the panels were brought back to the laboratory for a new set of measurements and the cycle was repeated. Although the test will continue for as long as it is necessary to identify a trend, this report only includes data up to the third week of exposure.

TABLE 1. MATERIALS TESTED, TYPE, RUST GRADE EVALUATION, ZINC CONTENT, AND THICKNESS

COATING	CODE	TYPE	RATING	% ZINC	COATING THICKNESS (MILS)
AMERON D-21-9	A	SB	10.00	N/A	6.0
DEVOE CATHACOAT 302H	B	SB	6.32	N/A	4.0
SHERWIN WILLIAMS ZINC CLAD II	C	SB	10.00	N/A	5.0
AMERON D-4	D	WB	9.41	N/A	4.5
AMERON D-21-5	E	WB	6.38	N/A	4.5
BRINER V-65	F	WB	10.00	N/A	3.0
CARBOLINE CZ-D7	G	WB	9.94	N/A	3.5
DuPONT GANICIN 347 WB	H	WB	10.00	86	3.5
INORGANIC COATINGS IC-531	I	WB	9.79	90	3.5
PORTER TQ-4374H	J	WB	9.10	N/A	3.0
SOUTHERN COATINGS CHEMTEC 600	K	WB	9.10	90	3.0
SUBOX GALVANOX IV	L	SB	8.75	N/A	3.5
BYCO SP-101	M	SB	8.13	N/A	3.5
CARBOLINE CZ-11	N	SB	10.00	N/A	5.0
CORONADO 935-152	O	SB	5.25	84	4.5
DEVOE-MARINE CATHA-COAT 304	P	SB	10.00	N/A	5.0
GLIDDEN GLID-ZINC 5530	Q	SB	8.50	86	5.0
GLIDDEN GLID-ZINC 5536	R	SB	2.00	76	4.0
KOPPER 701	S	SB	7.00	N/A	4.0
SUBOX GALVANOX V	T	SB	2.75	N/A	4.0
TNEMEC 90E-75	U	SB	7.00	76	3.5
GALVANIZED STEEL	Z				
WB = Water-based (inorganic silicate)					
SB = Solvent-based (ethyl silicate)					

A summary of the weather conditions during the first three weeks of exposure appears in Table 2.

TABLE 2. WEATHER CONDITIONS DURING THE FIRST THREE WEEKS OF EXPOSURE

WEEK 1	TEMPERATURE	BAROMETER (IN)	% HUMIDITY	RAINFALL (IN)	WINDS (MPH)
Day 1	H 90 (32C), L73 (23C)	30.08	79	0.00	E 11
Day 2	H 88 (31C), L 72 (22C)	30.08	72	0.00	S E9
Day 3	H 87 (31C), L 72 (22C)	30.08	63	0.00	E 9
Day 4	H 88 (31C), L 76 (25C)	30.10	65	0.00	E 11
Day 5	H 87 (31C), L 70 (21C)	30.07	72	0.00	SE 11
Day 6	H 87 (31C), L 72 (22C)	30.05	70	0.00	E 11
Day 7	H 86 (30C), L74 (24C)	30.01	73	0.00	SE 11
WEEK 2					
Day 1	H 82 (28C), L72 (22C)	30.03	67	0.00	N 7
Day 2	H 87 (31C), L 73 (23C)	30.15	72	0.20	SE 9
Day 3	H 84 (29C), L 68 (20C)	30.13	76	Trace	SE 8
Day 4	H 87 (31C), L 78 (26C)	30.08	78	0.30	NW 6
Day 5	H 86 (30C), L 71 (22C)	30.06	83	0.10	E 7
Day 6	H 83 (29C), L 70 (21C)	30.08	79	1.00	Calm
Day 7	H 86 (30C), L 70 (21C)	30.07	85	1.10	S 9

TABLE 2. WEATHER CONDITIONS DURING THE FIRST THREE WEEKS OF EXPOSURE (CONTINUED)

WEEK 3	TEMPERATURE	BAROMETER (IN)	% HUMIDITY	RAINFALL (IN)	WINDS (MPH)
Day 1	H 88 (31C), L 71 (23C)	30.12	78	Trace	SE 9
Day 2	H 86 (30C), L 73 (20C)	30.18	89	Trace	SE 7
Day 3	H 86 (30C), L 71 (20C)	30.13	76	Trace	SE 7
Day 4	H 92 (34C), L 70 (21C)	30.04	81	0.00	SE 9
Day 5	H 92 (34C), L 77 (25C)	30.02	71	0.00	SW 9
Day 6	H 92 (34C), L 78 (26C)	30.06	73	0.00	SW 5
Day 7	H 92 (34C), L 74 (24C)	30.05	65	0.00	NW 7

III. RESULTS AND DISCUSSION

Corrosion potential values in volts are shown on Table 3. Graphical presentation of the variation of the corrosion potential over the first three weeks of exposure is presented in Figures 2-4. No identifiable trend was observed that would differentiate at this point between the good coatings (those having a corrosion rating near 10) and the poor ones (those having a rating closer to 1).

Data generated from impedance measurements are summarized in Figures 5-25 for each of the coatings. The sets of data for each coating are given as Bode magnitude plots, where the logarithm of the absolute value of the impedance ($\log |Z|$) was plotted versus the logarithm of the frequency. Figure 26 is a Bode magnitude plot of galvanized steel which was included in the study for comparison purposes. Each figure includes the initial data and the Bode magnitude plots obtained after each week of atmospheric exposure. The plots indicate that changes are taking place in the coatings. However, no clear trend is evident at this point that will allow to distinguish between a good coating and a poor one. It is possible that a trend will develop with time. The measurements will be performed again after six months of exposure and after one year of exposure.

Time limitations did not permit to analyze the data in the Nyquist format. Future work will focus on exploring the best parameters to identify a trend.

TABLE 3. CORROSION POTENTIAL

COATING	CORROSION POTENTIAL IN VOLTS				
	CODE	INITIAL	ONE-WEEK	TWO-WEEK	THREE-WEEK
AMERON D-21-9	A	-1.029	-1.025	-0.980	-0.947
DEVOE CATHACOAT 302H	B	-0.004	-0.894	-0.143	-0.328
SHERWIN WILLIAMS ZINC CLAD II	C	-1.015	-1.024	-0.461	-0.923
AMERON D-4	D	-0.990	-0.899	-0.767	-0.816
AMERON D-21-5	E	-0.249	-0.552	-0.575	-0.155
BRINER V-65	F	-1.009	-1.030	-0.995	-0.976
CARBOLINE CZ-D7	G	-0.777	-0.117	-0.829	-0.869
DuPONT GANICIN 347 WB	H	-0.895	-0.758	-0.968	-0.903
INORGANIC COATINGS IC-531	I	-0.846	-0.725	-0.881	-0.867
PORTER TQ-4374H	J	-0.992	-0.982	-0.944	-0.935
SOUTHERN COATINGS CHEMTEC 600	K	-0.524	-0.571	-0.560	-0.616
SUBOX GALVANOX IV	L	-0.988	-0.999	-0.979	-0.935
BYCO SP-101	M	-1.013	-1.025	-0.964	-0.965
CARBOLINE CZ-11	N	-1.051	-1.021	-1.008	-0.949
CORONADO 935-152	O	-1.035	-1.022	-0.956	-0.926
DEVOE-MARINE CATHA-COAT 304	P	-0.987	-0.995	-0.939	-0.930
GLIDDEN GLID-ZINC 5530	Q	-1.016	-1.011	-0.957	-0.929
GLIDDEN GLID-ZINC 5536	R	-0.768	-1.015	-0.490	-0.771
KOPPER 701	S	-0.975	-0.999	-0.937	-0.950
SUBOX GALVANOX V	T	-1.002	-1.020	-0.244	-0.584
TNEMEC 90E-75	U	-0.983	-0.686	-0.861	-0.773
GALVANIZED STEEL	Z	-1.024	-0.966	-0.959	

Figure 1. Exploded View of Model K0235 Flat Cell

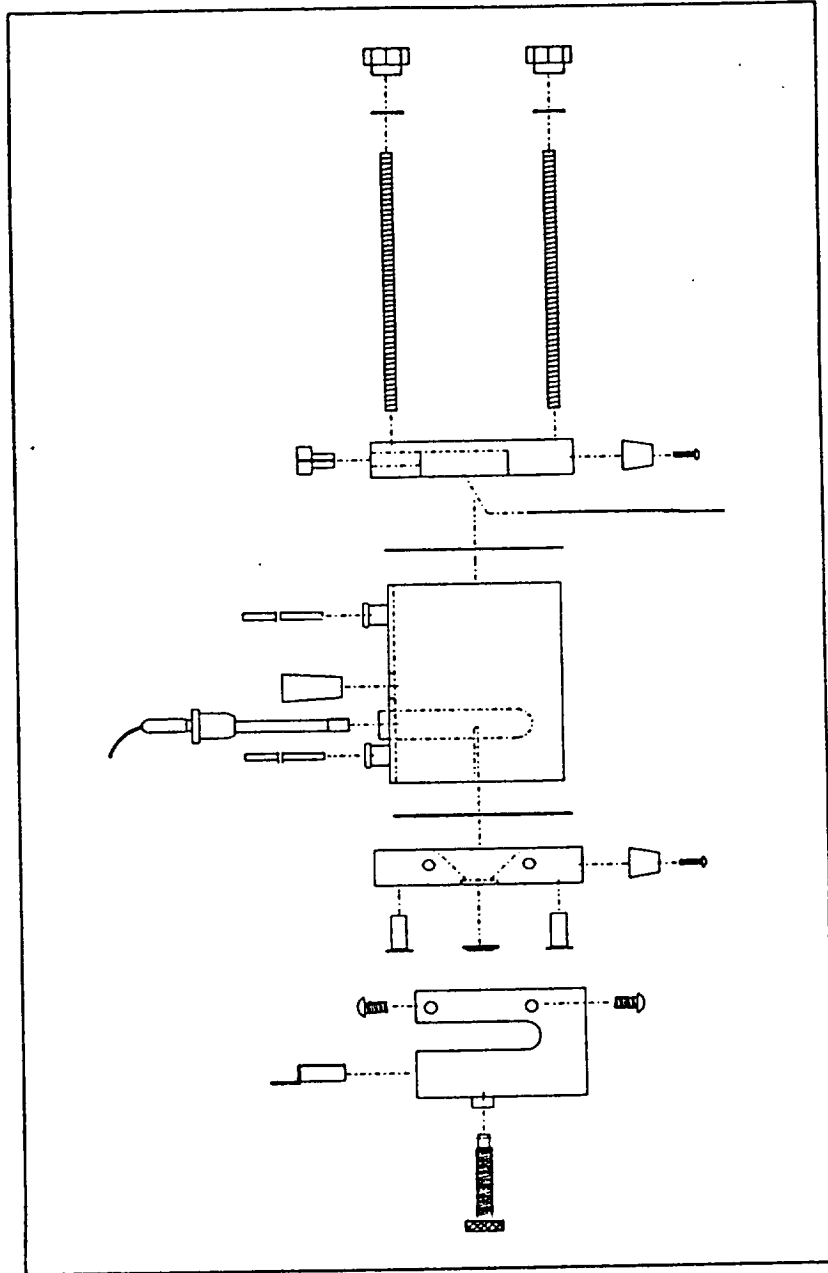


Figure 2. Corrosion Potential in Volts for Coatings A-H

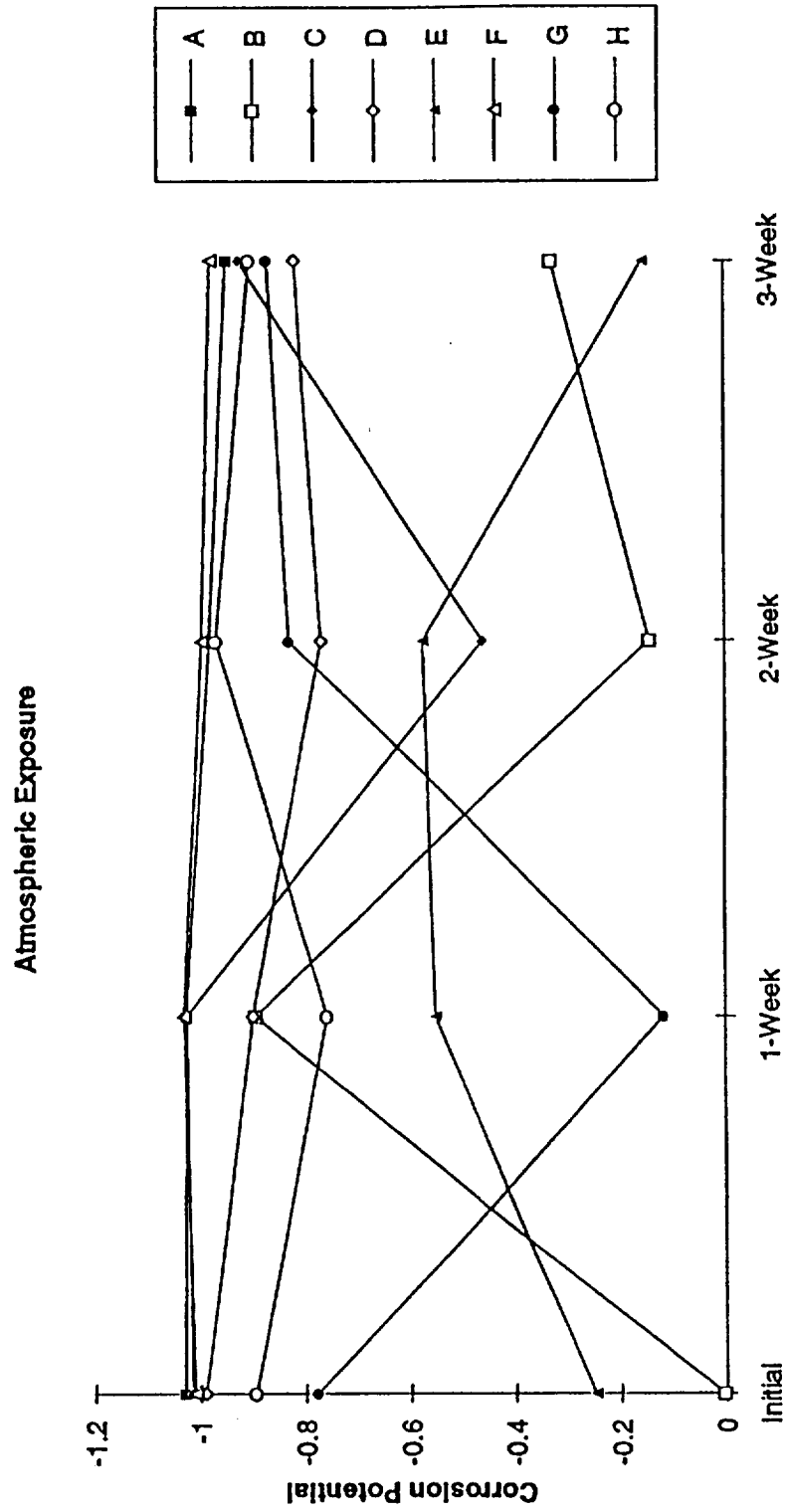


Figure 3. Corrosion Potential in Volts for Coatings I-O

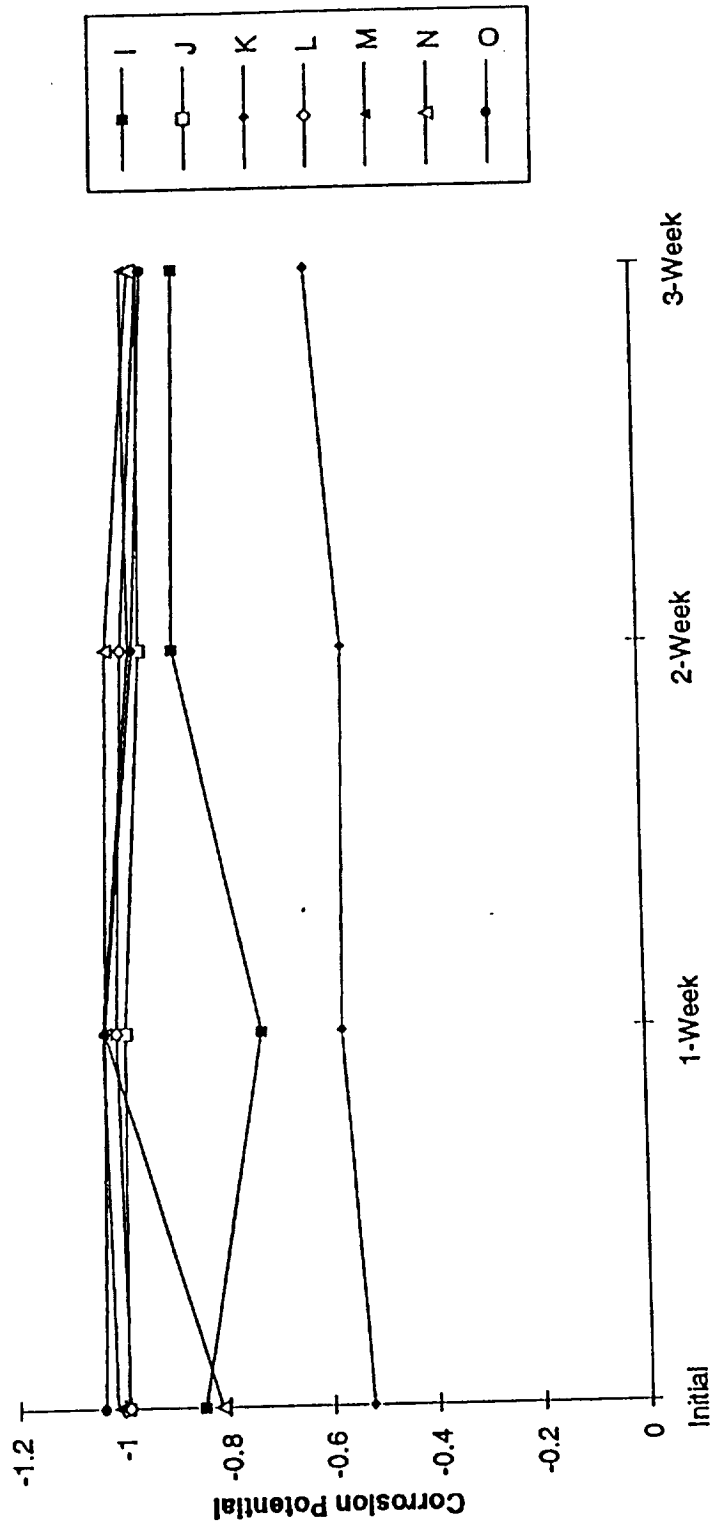
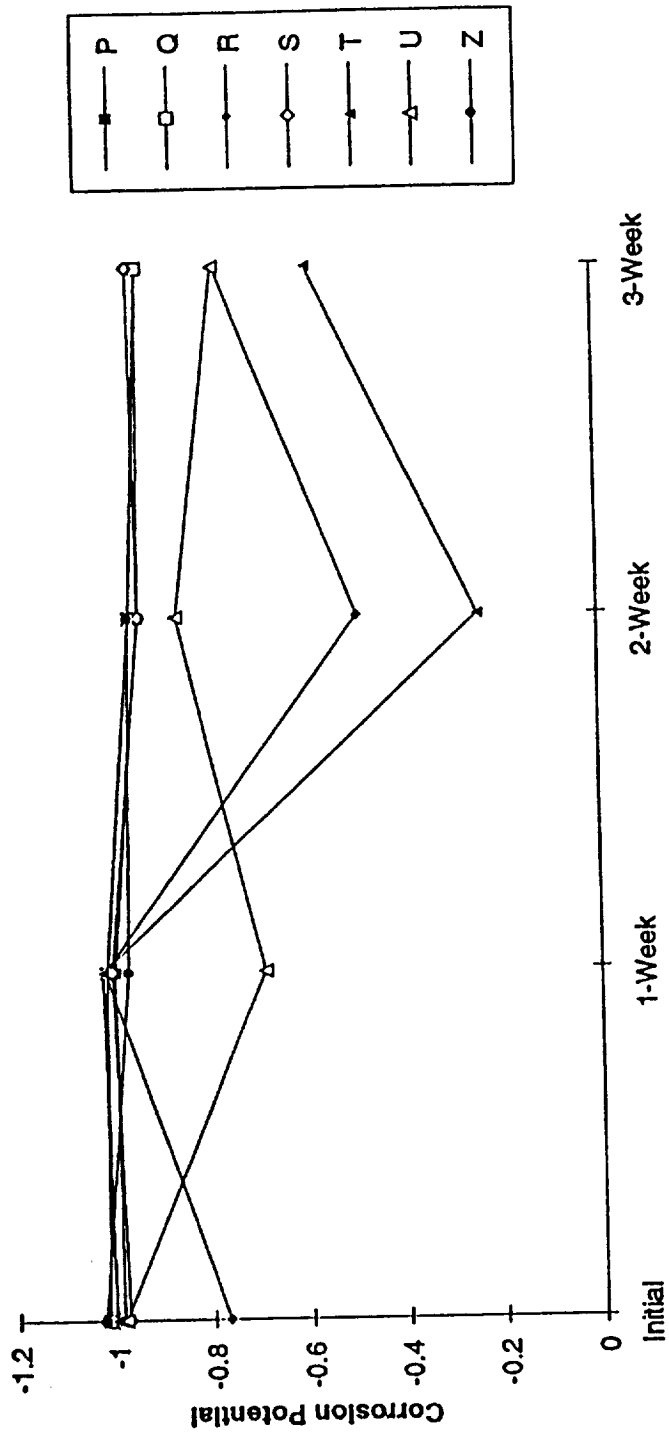


Figure 4. Corrosion Potential in Volts for Coatings P-U and Galvanized Steel



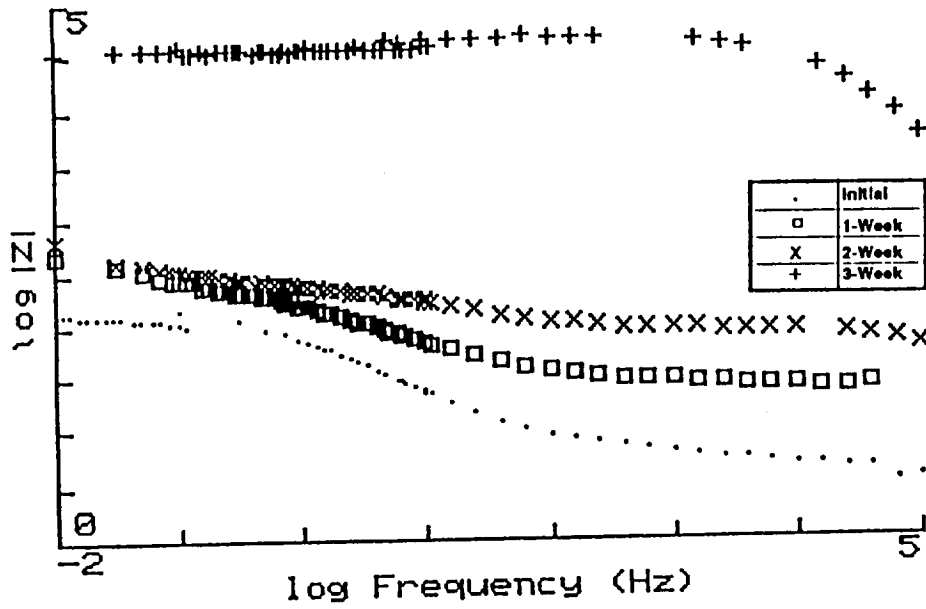


Figure 5. Bode Magnitude Plot for Carboline CZ-11

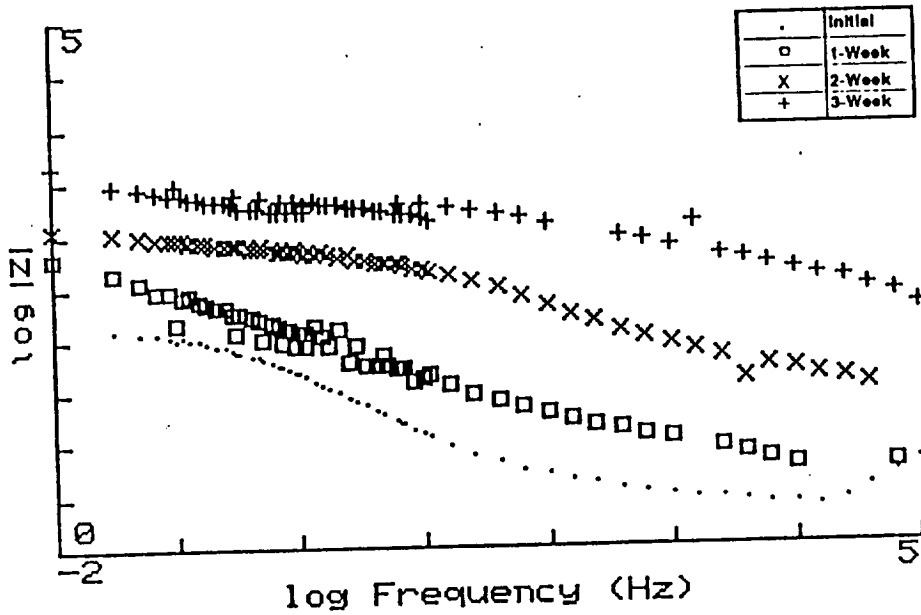


Figure 6. Bode Magnitude Plot for Ameron Devoe-Marine Catha-Coat 304

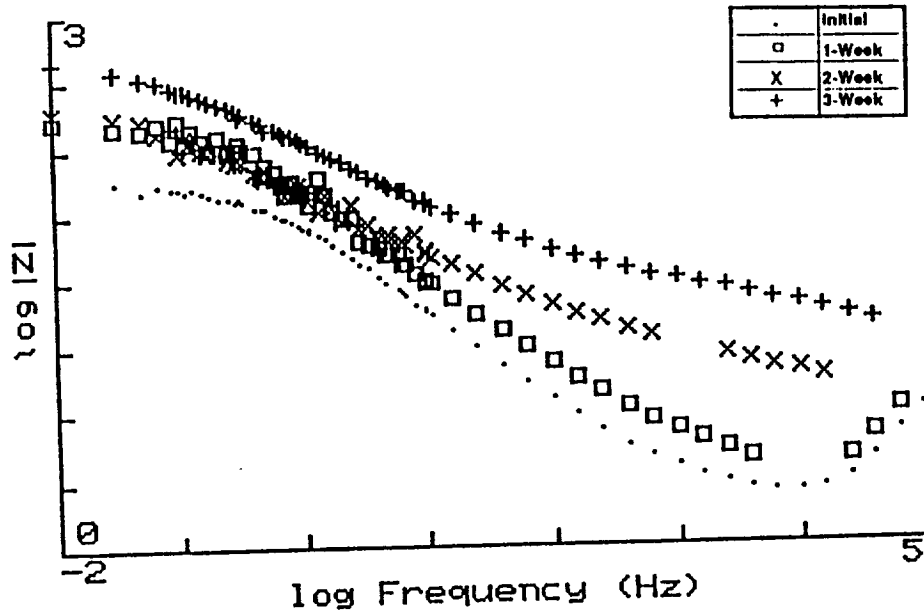


Figure 7. Bode Magnitude Plot for Briner V-65

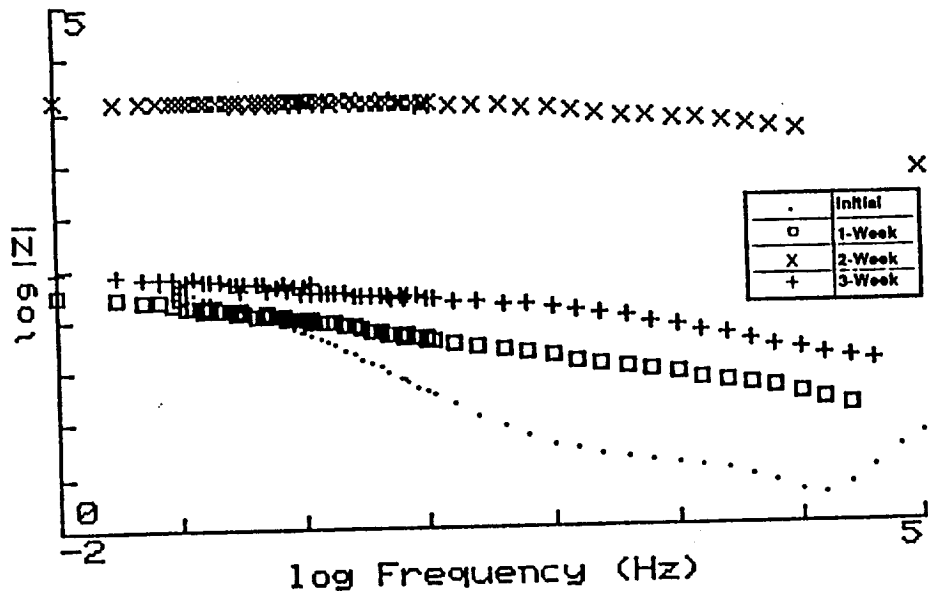


Figure 8. Bode Magnitude Plot for Ameron D-21-9

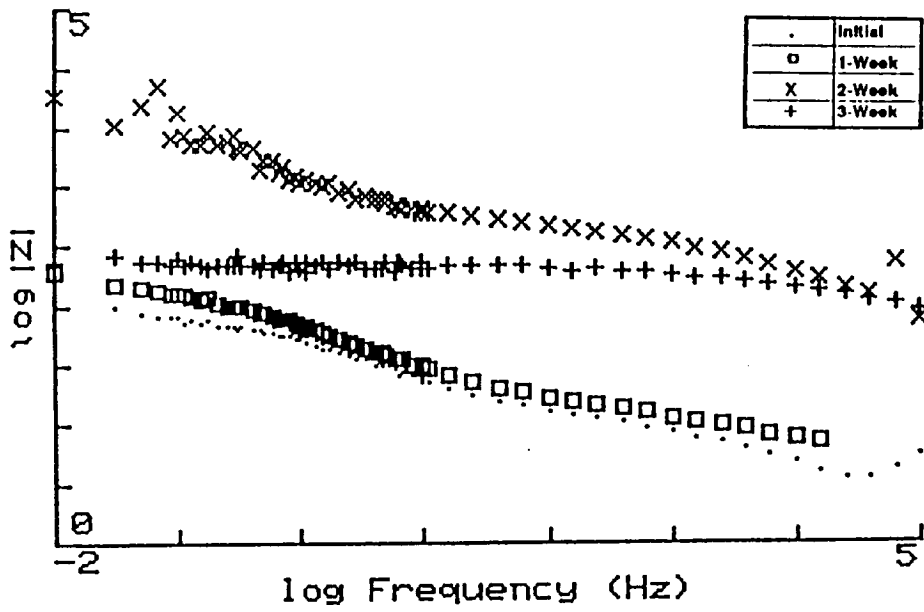


Figure 9. Bode Magnitude Plot for Sherwin Williams Zinc Clad II

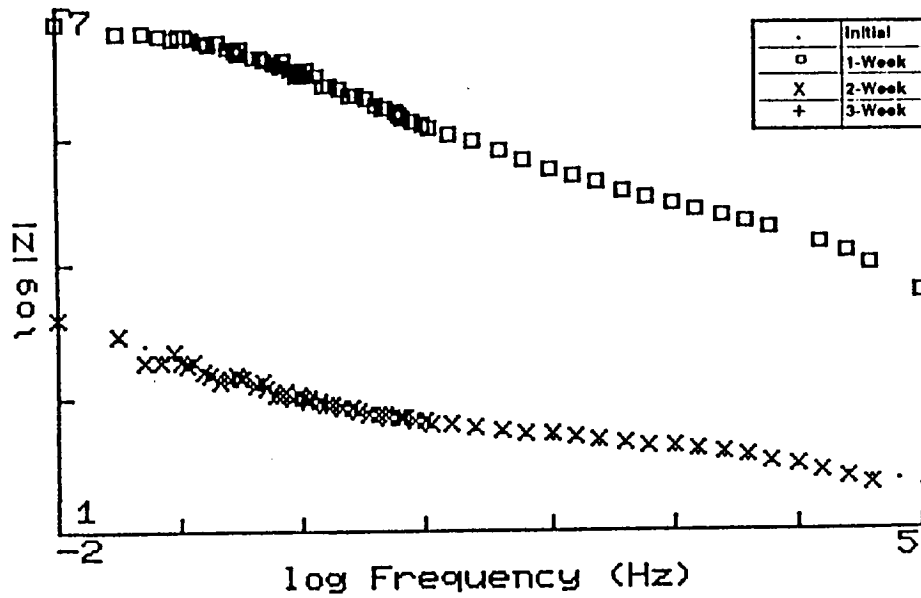


Figure 10. Bode Magnitude Plot for Carboline CZ-D7

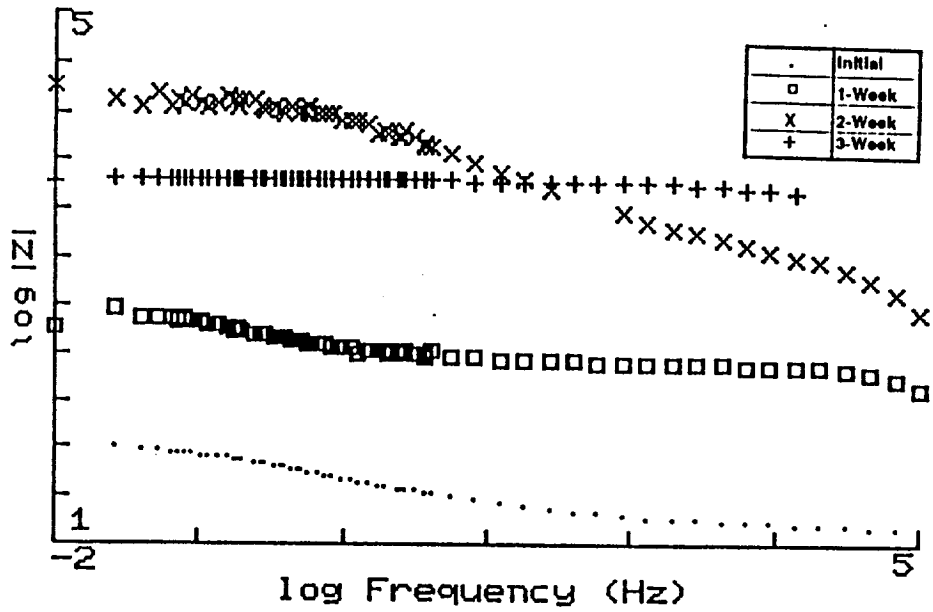


Figure 11. Bode Magnitude Plot for Ameron D-4

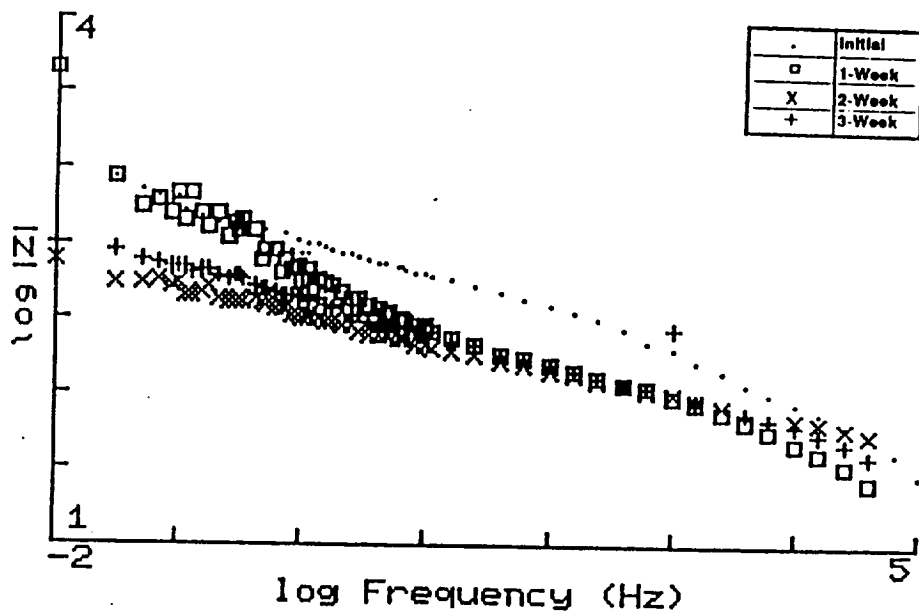


Figure 12. Bode Magnitude Plot for Dupont Ganicin

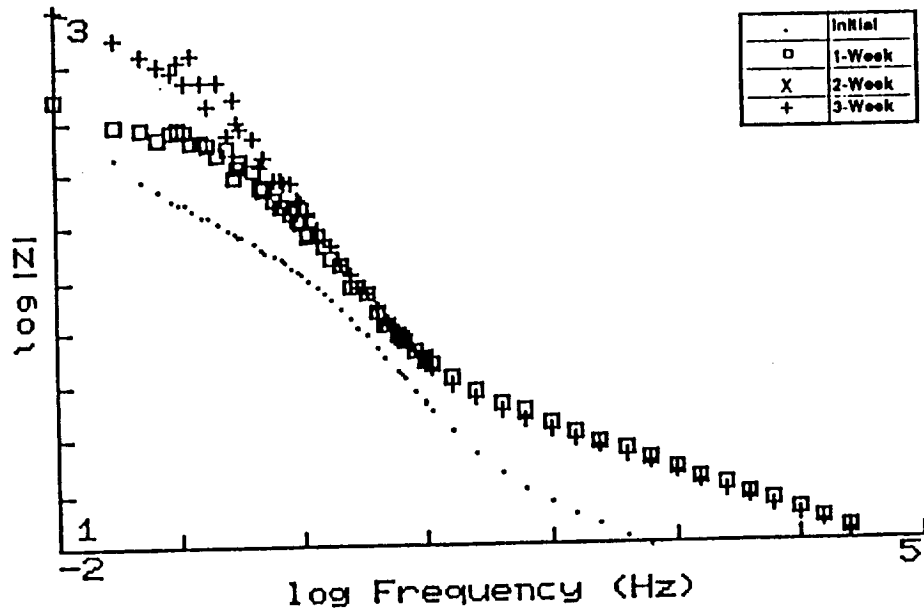


Figure 13. Bode Magnitude Plot for Porter TQ-4374H

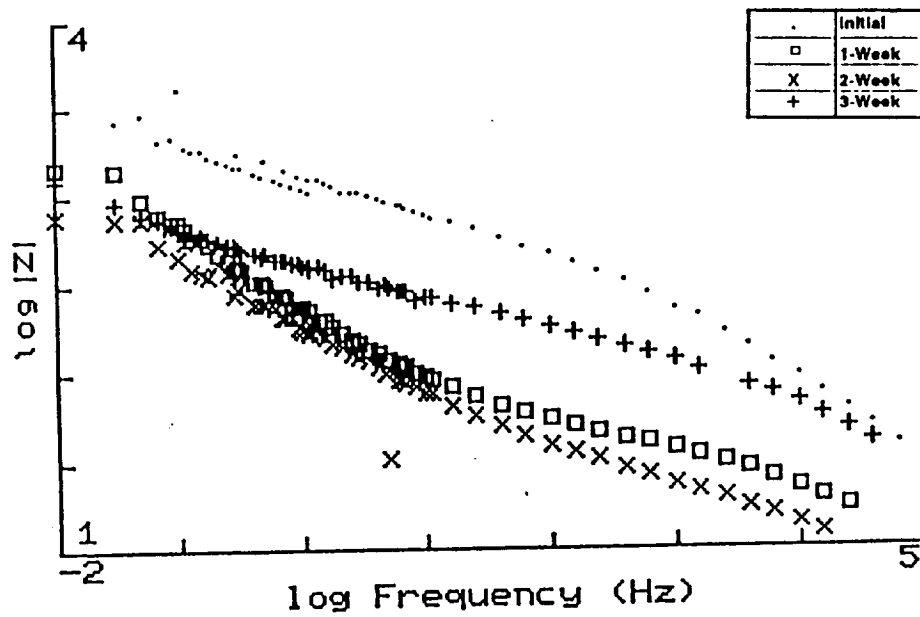


Figure 14. Bode Magnitude Plot for Inorganic Coatings IC-531

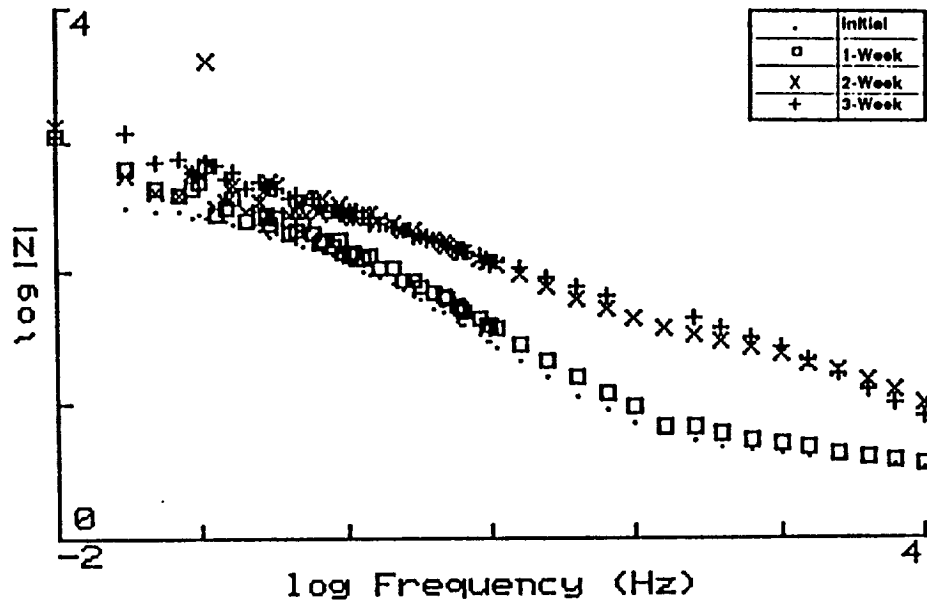


Figure 15. Bode Magnitude Plot for Subox Galvanox IV

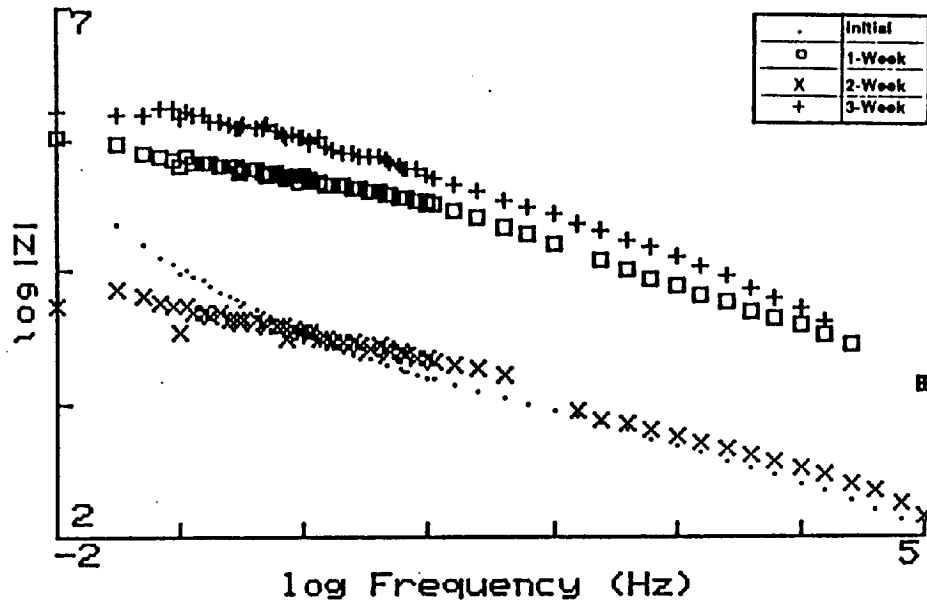


Figure 16. Bode Magnitude Plot for Southern Coatings Chemtec 600

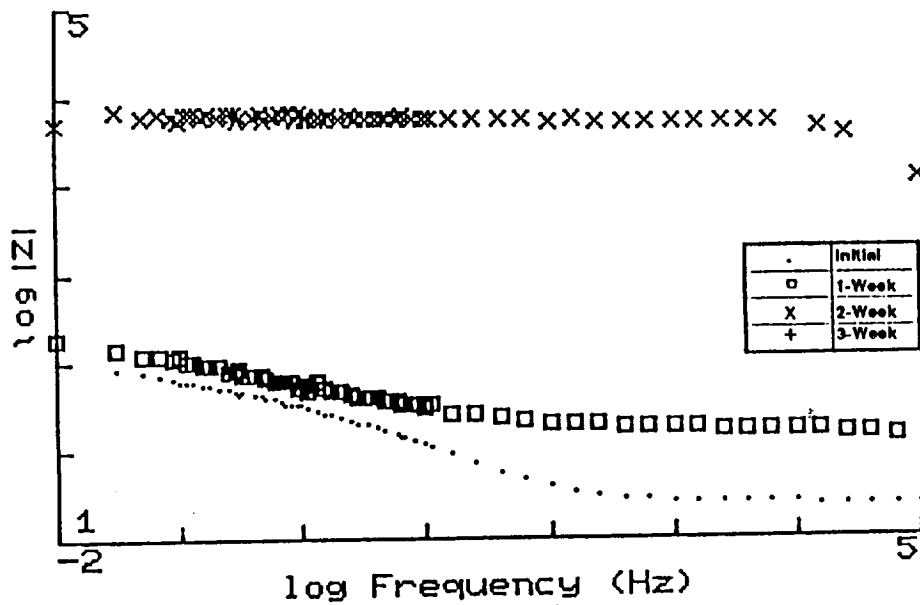


Figure 17. Bode Magnitude Plot for Glidden Glidzinc 5530

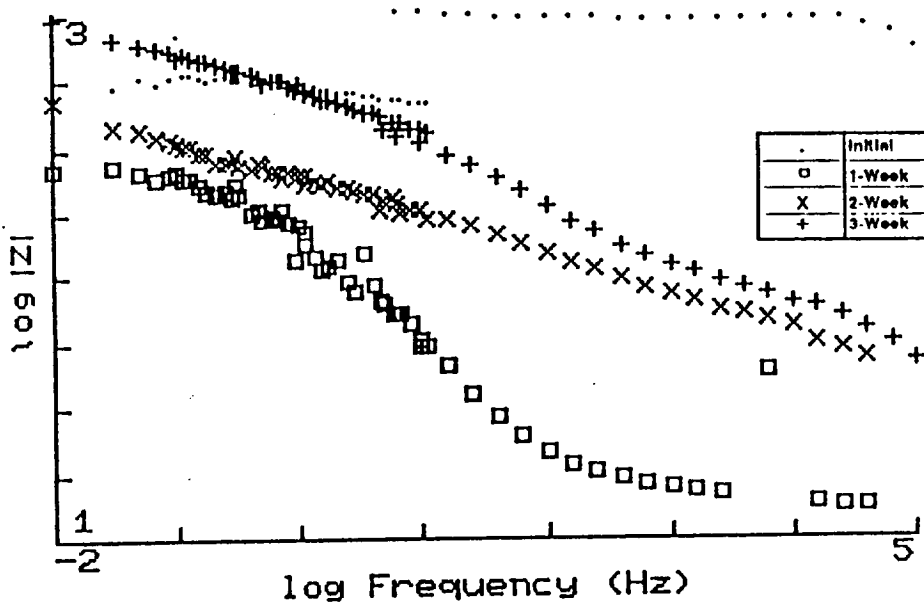


Figure 18. Bode Magnitude Plot for Byco SP-101

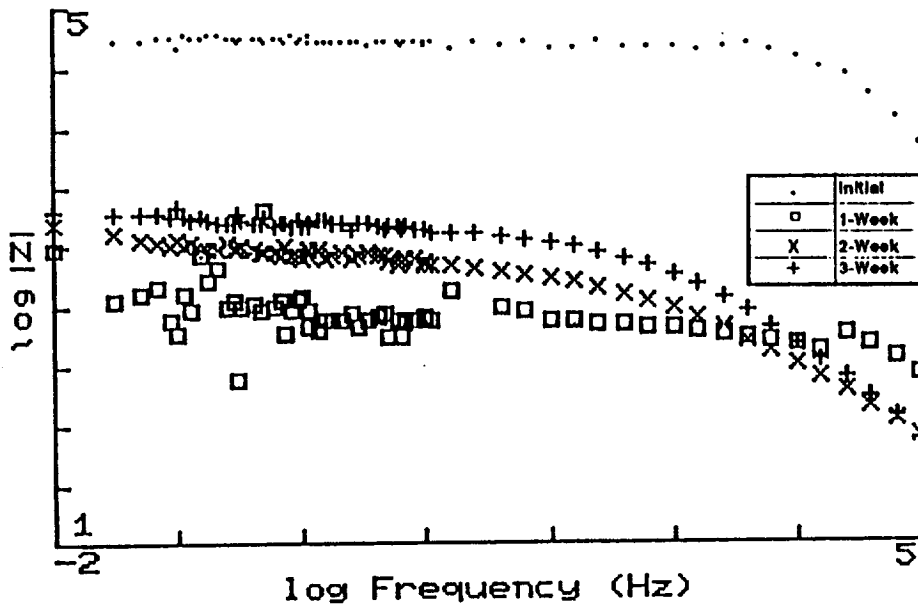


Figure 19. Bode Magnitude Plot for Tnemec 90E-75

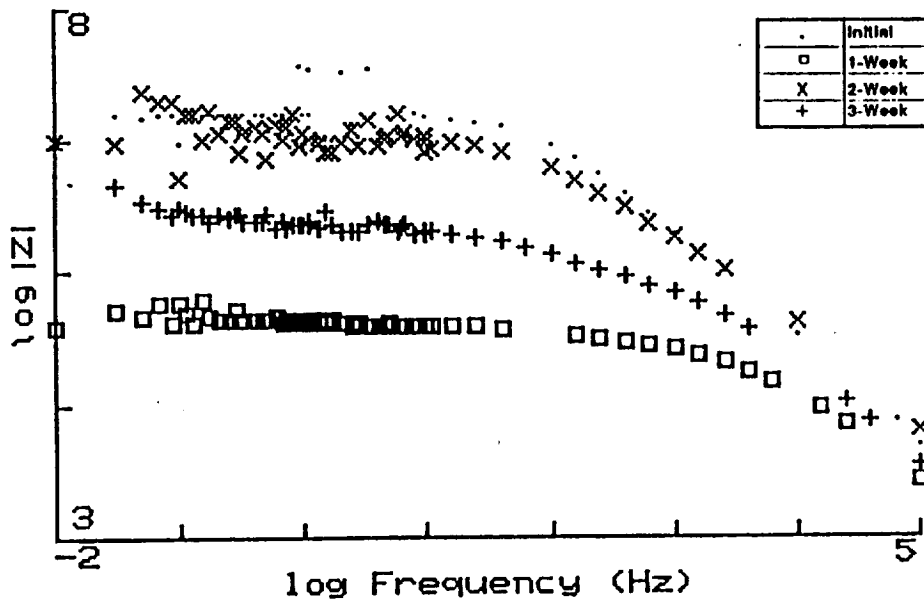


Figure 20. Bode Magnitude Plot for Devco Catha-Coat 302H

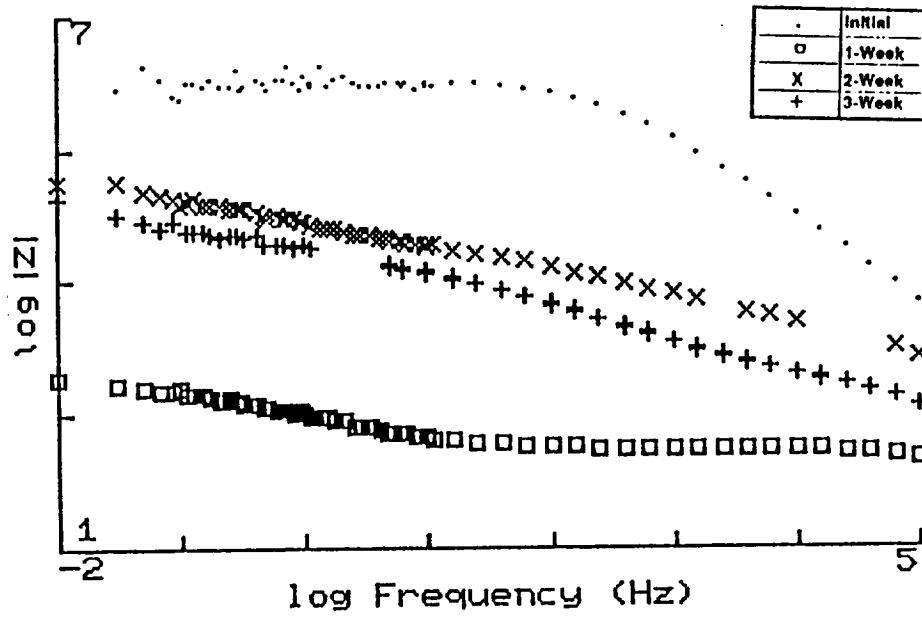


Figure 21. Bode Magnitude Plot for Glidden Glidzinc 5536

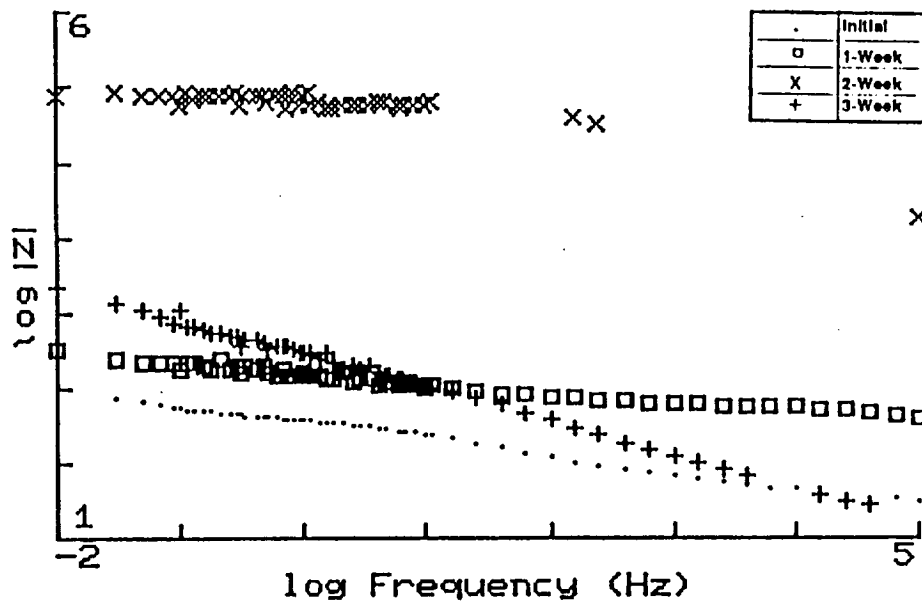


Figure 22. Bode Magnitude Plot for Koppers 701

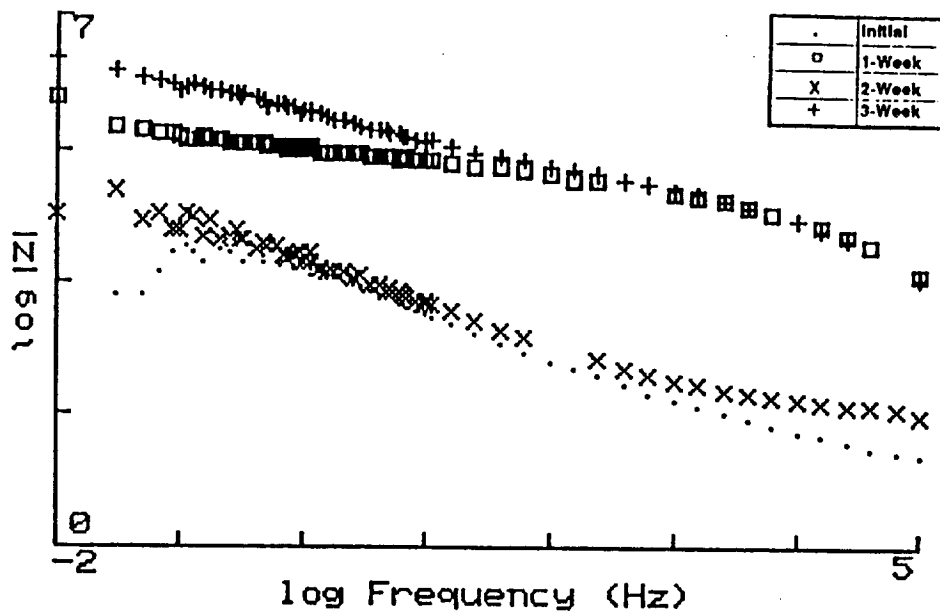


Figure 23. Bode Magnitude Plot for Ameron D-21-5

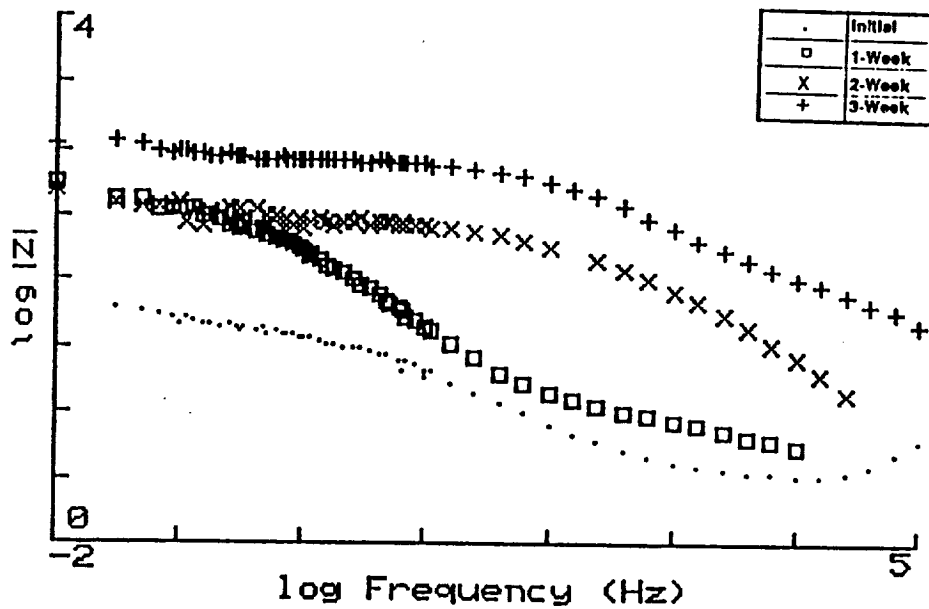


Figure 24. Bode Magnitude Plot for Coronado 935-152

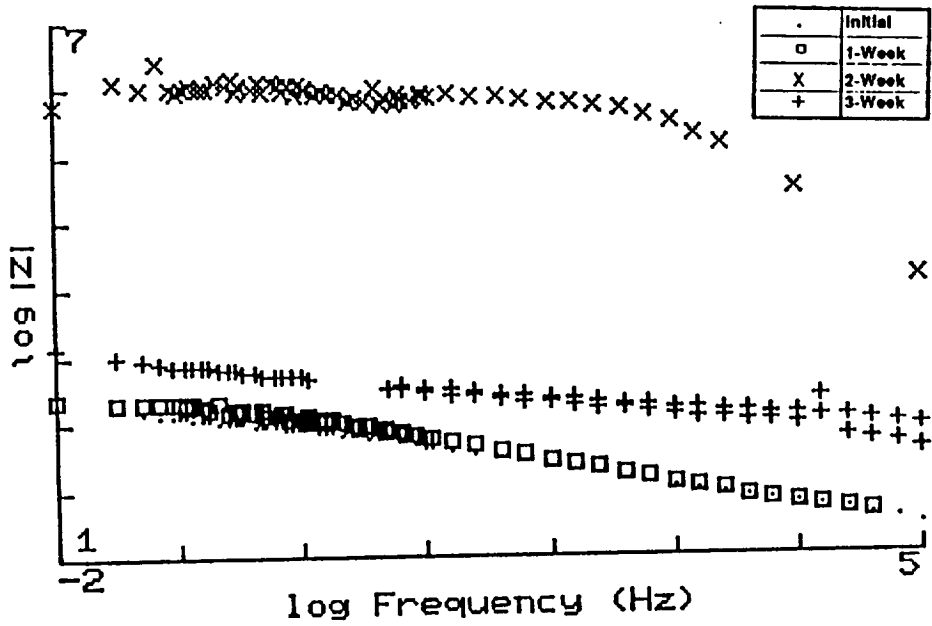


Figure 25. Bode Magnitude Plot for Subox Galvanox V

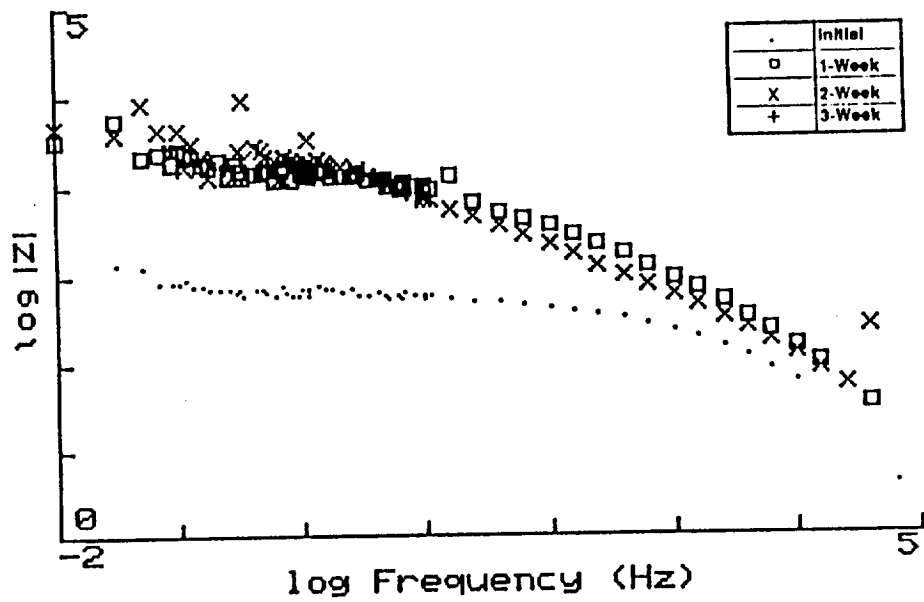


Figure 26. Bode Magnitude Plot for Galvanized steel

IV. REFERENCES

-
- ¹L. G. MacDowell, Evaluation of Protective Coating Systems for Carbon Steel Exposed to Simulated SRB Effluent After Eighteen Months of Seacoast Exposure, Document No. MTB-268-86B, NASA, Kennedy Space Center, Materials Testing Branch, February 16, 1988.
 - ²L. G. MacDowell, Status Report on Protective Coating Systems for Carbon Steel Exposed to Simulated SRB Effluent - 36 Month Exposure, Document No. MTB 268-86C, NASA, Kennedy Space Center, Materials Testing Branch, September 19, 1989.
 - ³L. G. MacDowell, Volatile Organic Content (VOC) Compliant Coating Systems for Carbon Steel Exposed to the STS Launch Environment - Application, Laboratory and 18 Month Exposure results, Study No. 93-2004, NASA, Kennedy Space Center, Materials Testing Branch, February 22, 1993.
 - ⁴L. G. MacDowell, L. M. Calle, Evaluation of High Performance Metal Alloys in the STS Launch Environment Using Electrochemical Impedance Spectroscopy, Document No. MTB-610-89A, NASA, Kennedy Space Center, Materials Testing Branch, August 16, 1990
 - ⁵C. Ontiveros, L. G. MacDowell, Electrochemical Impedance Spectroscopy for Evaluating Inorganic Zinc Rich Primers on Steel in the STS launch Environment, Interim Report, Document No. 91-4336, NASA, Kennedy Space Center, Materials Testing Branch, July 31, 1991.
 - ⁶L. G. MacDowell, Final Report on Protective Coating Systems for Carbon Steel Exposed to Simulated Solid Rocket Booster (SRB) Effluent -60 Mo9nth Exposure, Report No. 91-4821, NASA, Kennedy Space Center, Materials Testing Branch, October 18, 1991.

1994018761

N94-23234

1993 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA

54-66
197190
P 14

A FRAMEWORK FOR SFC/DC SYSTEM DATA ANALYSIS

PREPARED BY:	Dr. Martha A. Centeno
ACADEMIC RANK:	Assistant Professor
UNIVERSITY AND DEPARTMENT:	FAMU/FSU College of Engineering Department of Industrial Engineering
NASA/KSC	
DIVISION:	Shuttle Operations
BRANCH:	Vehicle Processing
NASA COLLEAGUE:	Amanda Mitskevich
DATE:	August 26, 1993
CONTRACT NUMBER:	University of Central Florida NASA-NGT-60002 Supplement: 11

A FRAMEWORK FOR SFC/DC SYSTEM DATA ANALYSIS

Martha A. Centeno¹

Athena Colucci

Department of Industrial Engineering
FAMU/FSU College of Engineering
Tallahassee, Florida 32316

ABSTRACT

The Shop Floor Control / Data Collection (SFC/DC) system is one of four major modules of the Integrated Work Control System being implemented at the NASA Kennedy Space Center. The SFC/DC system, which has been in operation approximately one and a half years, collects detailed information regarding the duration of tasks as well as work stoppages duration. This information needs to be effectively analyzed, so as to be able to manage the shuttle processing activities under the continuous improvement philosophy. We have studied the contents and form of the SFC/DC database, and based on the results of such study, we have designed a Shop floor Modeling, Analysis, and Reporting Tool (SMART). SMART is a collection of compatible tools (off-the-shelf as well as customized C-based tools) that work cohesively to prepare inputs for the analysis models and process the various models. The feasibility of SMART has been demonstrated through an initial prototype.

1. INTRODUCTION

This report is a summary of two extended reports that were written for the NASA TP-OAO office of the Kennedy Space Center. A multitude of valuable, yet confidential, information has been suppressed from this report.

This effort sought to design a cohesive framework for the effective analysis of the data being collected through the Shop Floor Control / Data Collection (SFC/DC) system. The design of such framework required a thorough understanding of the contents and form of the data in the SFC/DC database. Once this was established, a conceptual design and data exchange protocols for a Shop Floor Modeling, Analysis, and Reporting Tool (SMART) were devised. The SMART framework incorporates methods and tools to answer questions regarding delay duration such as:

- Do delays follow a pattern within a flow?
- Are delays correlated to the orbiter systems?
- Are delays correlated to the shuttle elements?
- Are delays correlated to the shuttle processing facility?
- What delays cause the most impact to the whole process?

Two protocols to provide inputs to the analysis models were considered: *directly from the*

¹ This effort has been funded through a NASA/ASEE fellowship and through a grant from the Council on Research and Creativity of the Florida State University.

database and from intermediate text files. At the moment of writing this report, NASA engineers had limited access to the various SFC relational tables; therefore, we chose the second protocol: *intermediate text files*. The procedures required to work with these models include instructions to prepare the text files and to transform them into the format of the analysis tool. STATGRAPHICS® is being recommended as the analysis tool based on its ease of use and its *macros* programming capabilities.

Section 2 of this report describes briefly what SFC is and how its database is populated. Section 3 describes the conceptual design for SMART. Section 4 gives a sample a proposed procedures to analyze SFC data. Section 5 summarizes the results of this effort and their significance to NASA and ASEE.

2. SFC/DC AND ITS DATABASE

SFC is one of the four clusters that make up the Integrated Work Control System (IWCS), which will integrate the shuttle processing databases at Kennedy Space Center (KSC). The IWCS framework will enable communication among the four clusters and add new data collection protocols (Figure 1).

The Shop Floor Control (SFC) module has been operational for one and a half years; however, at this stage, automatic links to the other 3 modules have not been implemented yet; except for a partial link to IOS (CASPR). SFC revolves around a DB/2 database with PFORMS acting as the database management system (DBMS). Tables in the SFC database are populated either from the IOS database or directly by the user. The main dynamic entity in the SFC and IOS database is a *task*; thus, the physical storage location and update privileges are driven by the status of the WAD. Figure 2 shows the various *states* that a WAD may experience throughout its life-cycle as well as the software systems (SFC or IOS) governing the WAD at each one of these statuses.

The majority of updates to WADs are done by technicians. Once all tasks in IOS that fall in the 11-day window are transferred to SFC, they are *available* to the technicians to work on. (table *orderopn*). When technicians are ready to work on a task, or when they want to change a task's status to hold, they clock in and provide pertinent information. This information does not go directly into *tsf01_delay* or *tsf02_task_worked*; rather, it goes to other SFC tables, such as ACTVEMPL from which the information for *tsf01_delay* and *tsf02_task_worked* is extracted once a week, as shown in Figure 3.

Tables in the SFC database are populated either from the IOS database or directly by the user. Figure 4 depicts the data exchange protocol that is being utilized to link IOS and SFC databases. This protocol is executed in batch mode, and it enables *almost real time* updates to both databases. It falls short of real time because the protocol is executed either once a day (from IOS to SFC) or twice a day (from SFC to IOS). IOS is not a DB/2 database; consequently, when transferring data to (and from) SFC, it is necessary to *extract* the data to a format neutral to both databases (ASCII or binary) and then *load* the data into the database unique format. This conversion is done automatically by a series of COBOL and REXX/SQL procedures. These procedures have been designed to preserve data integrity: they avoid duplicate records, prevent unauthorized updates, create tables for unexpected errors in the data, and so forth.

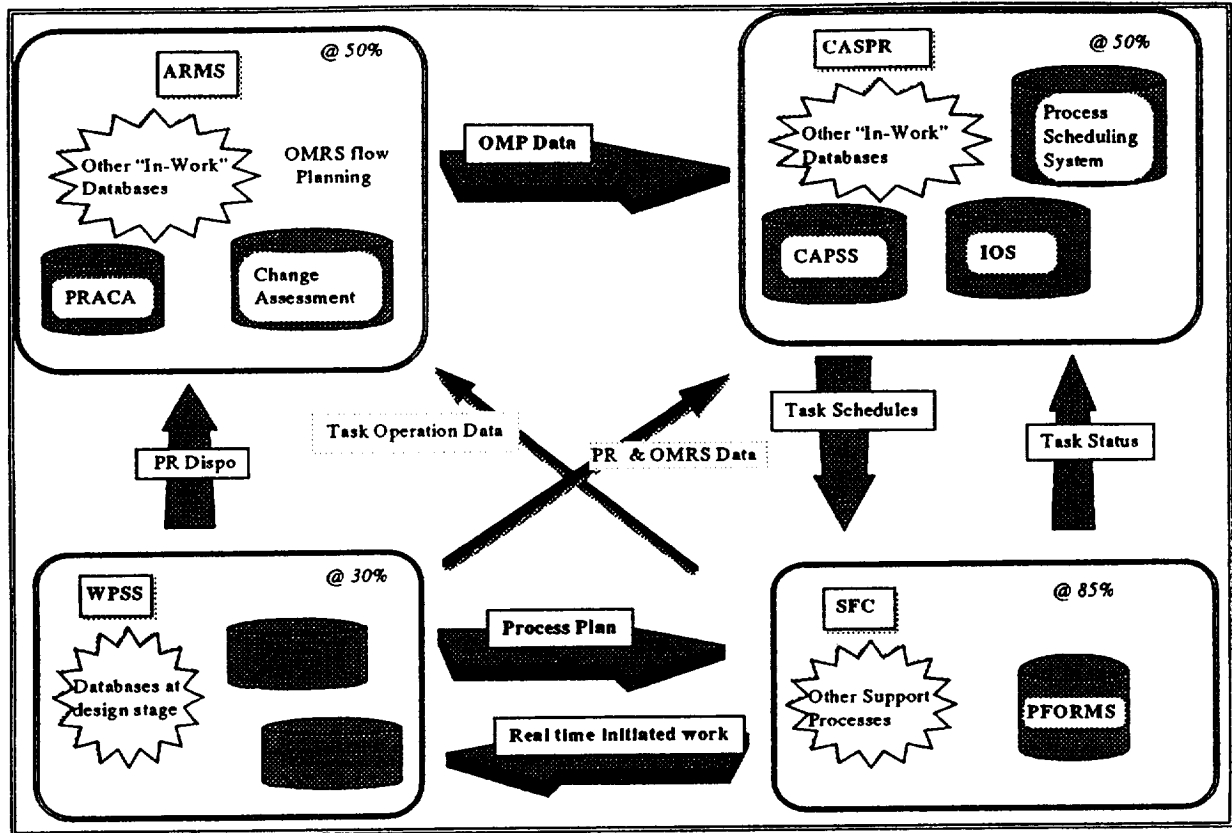


Figure 1: The IWCS Concept

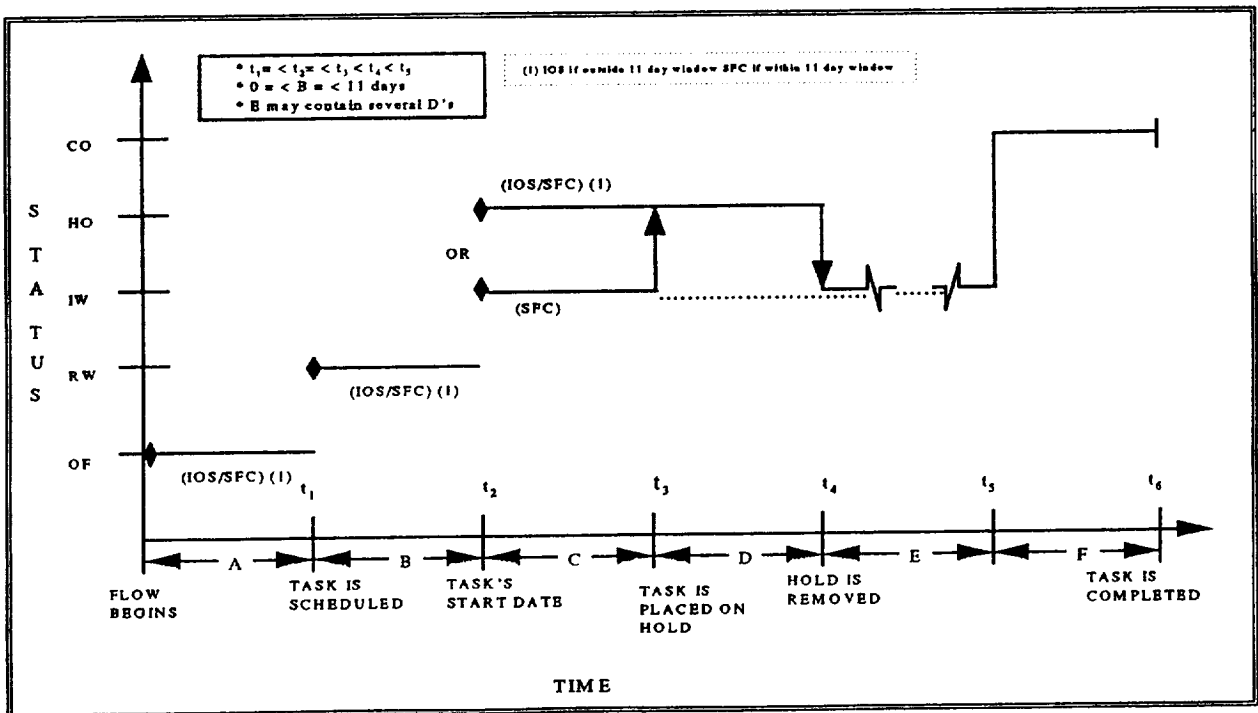


Figure 2: Generic Realization of a WAD

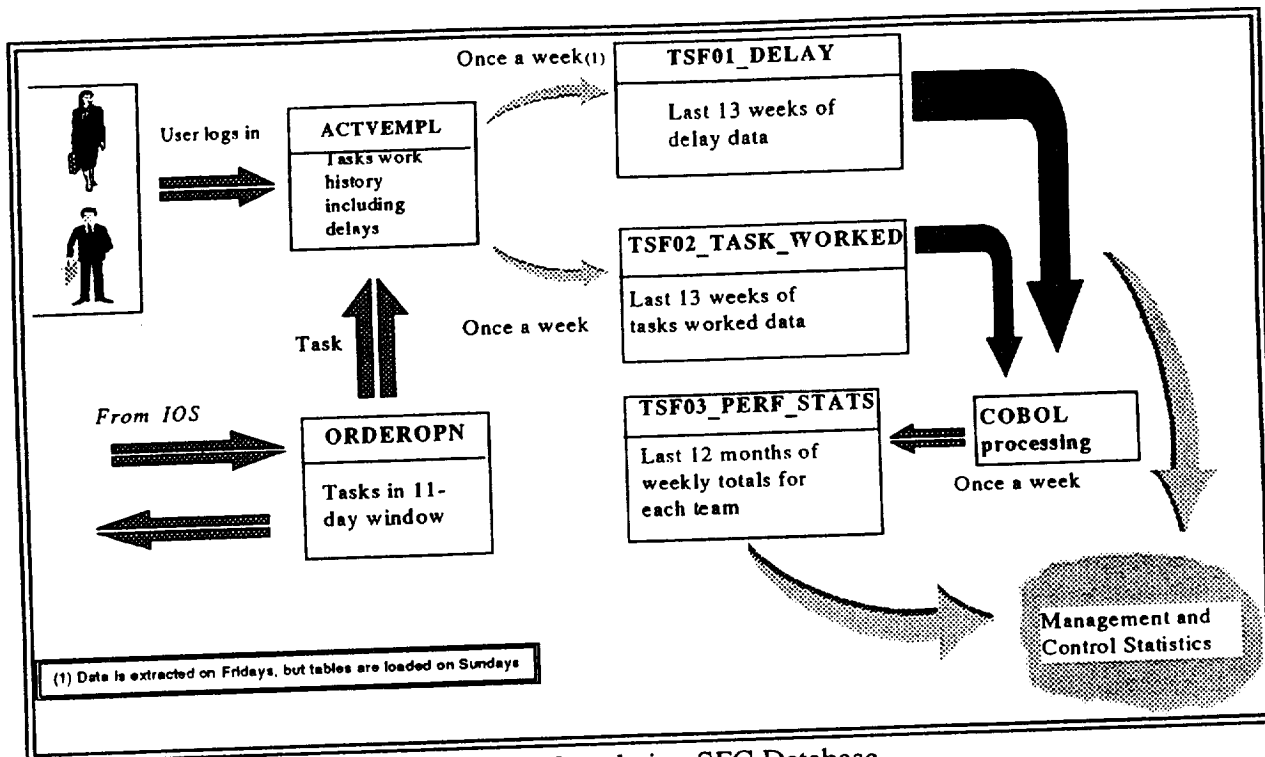


Figure 3: Populating SFC Database

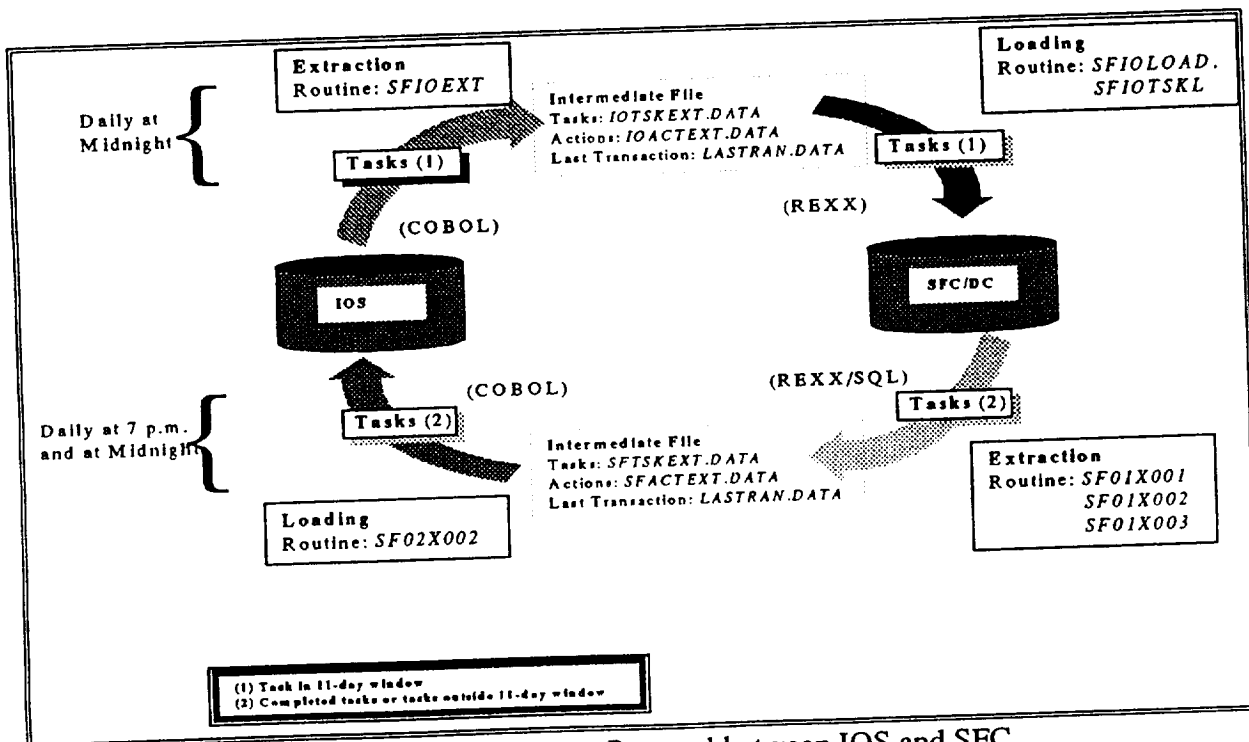


Figure 4: Data Exchange Protocol between IOS and SFC

The SFC database is truly a DB/2 database; therefore, it is accessible through DB/2 SQL commands and QMF SQL, i.e. it is not necessary to go through the PFORMS application to extract data out of it. Figure 5 shows the various ways in which NASA engineers may gain direct access to the SFC/DC database, with QMF being the most suitable via to achieve the goals of this on-going effort.

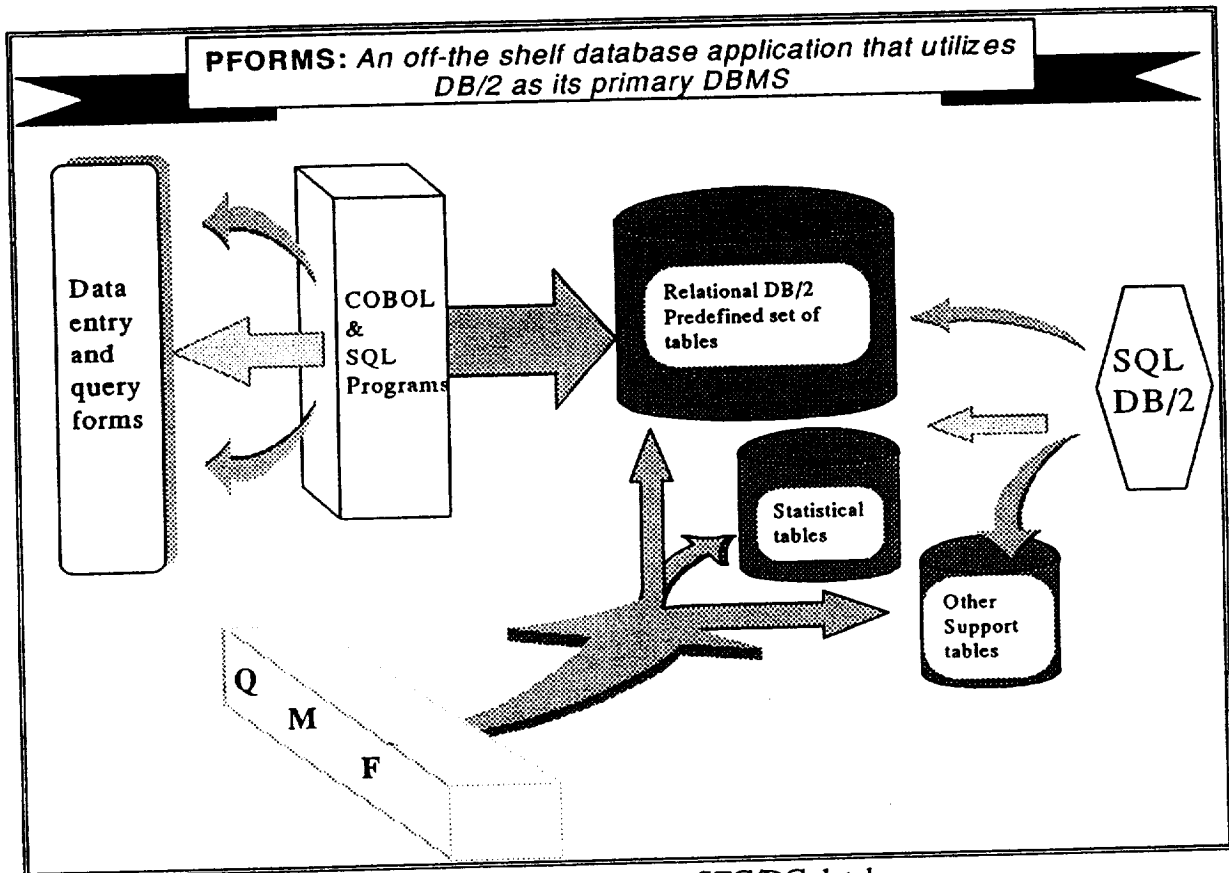


Figure 5: Gaining access to SFC/DC database

In a final report given to the TP-OAO office, we have given more detailed information regarding the populating of the SFC/DC database, as well as extensive documentation on the various tables in it. The reader is directed to contact TP-OAO office if interested in more details than those provided in this report.

3. THE SMART CONCEPT

The Shop floor Modeling, Analysis and Reporting Tool (SMART) is a framework designed around a relational database (Figure 6). The database contains statistical information regarding work time and delay duration. Under this framework, a link to the SFC/DC database (via text files) has been set. The framework is sensitive to the user of it. We have identified three possible users:

- * The *Statistical user* interacts with SMART to execute pre-defined analysis procedures. This user would request from the *designer user* modifications to SMART, so that new analysis models can become part of it.
- * The *Modeling user* interacts with SMART to extract raw data as well as statistical summaries from its database. This user would request from the *designer user* modifications to SMART, so that new modeling ideas can be readily tested.
- * The *Designer user* interacts with SMART through its SQL query language and other DBMS tools. This user would service the *statistical user* and *modeling user* needs regarding the modeling and analysis activities at the TP-OAO office.

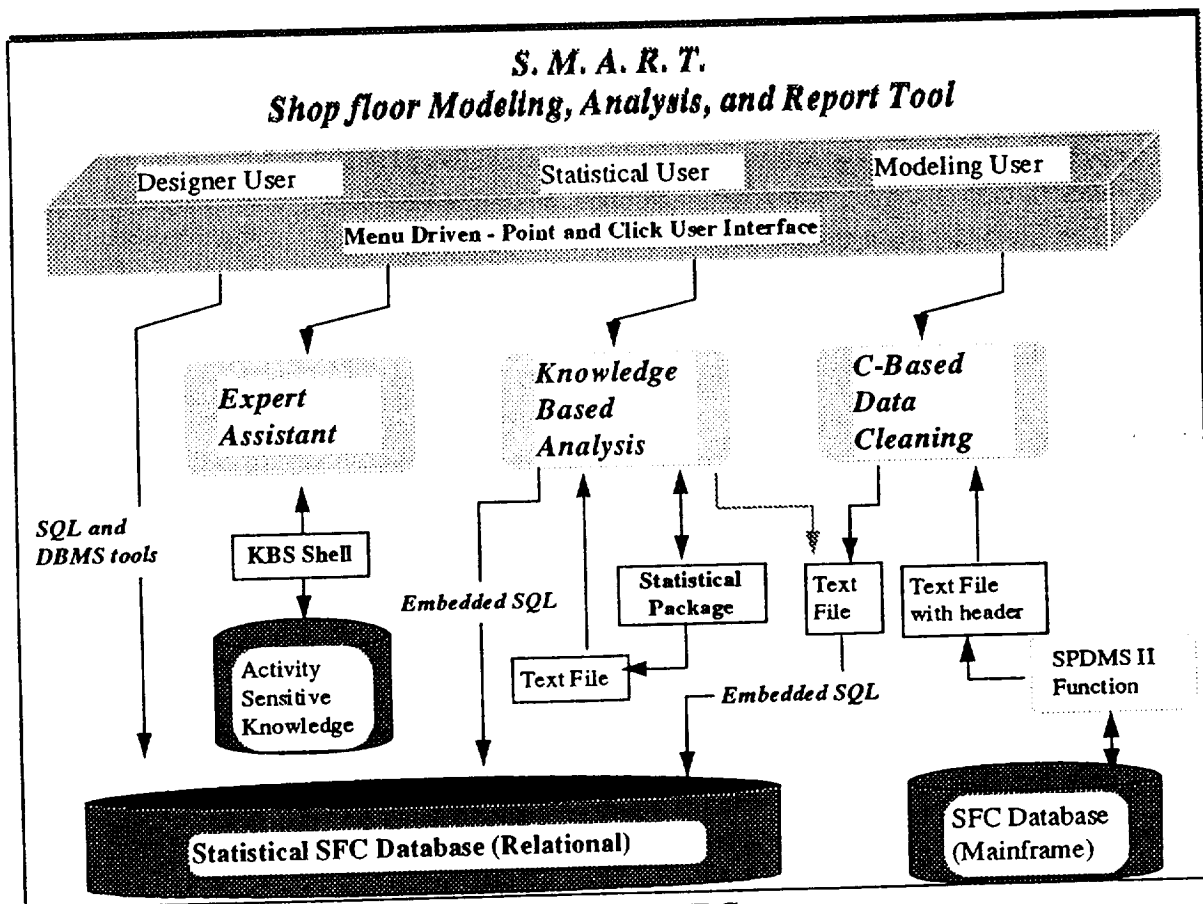


Figure 6: The SMART Concept

From Figure 6, we can see that there three major components to SMART, plus the access for the *designer user*. Each one of these components adds capabilities to SMART as follows:

- * SQL and DBMS Tools Access: It enables the creation of tables as new analysis

requirements are defined. It enhances accessibility of data from the analysis point of view. It enables the creation of "canned" reports using the **report** tool of the DBMS. It also enables the development of data entry forms as well as of specific applications (as needed) using the **forms** and **menu builder** tools.

- * *Expert Assistant*: It is an knowledge based expert system that operates in three different modes. It can be a tutor regarding the various statistical analysis, or it can be an assistant to the NASA engineer in choosing appropriate analysis methods, or it can provide user-sensitive and procedure sensitive help for the entire SMART framework.
- * *Knowledge Based Analysis*: This component computes descriptive statistics and estimates on delay duration, conducts various correlation analyses, conduct various hypotheses testing on the similarity of facilities, computes various joint probability functions for delay duration, and it enables the collection of historic behavior for possible use in other modeling activities.
- * *C-based Data Cleaning* It enables a semi automated link between the statistical database and current SFC database by preparing extraction procedures and then cleaning the resulting files. Once the files are cleaned, it utilizes embedded SQL to insert the data into the statistical tables. It also enables the creation of input files based on the format of the analysis and modeling tool to use.

The feasibility of SMART has been shown through a small prototype. This prototype includes a menu driven interface, through which a cleaning procedure can be executed. The data extraction and data cleaning prototype is partly operational for analyses on *total delayed time for a given flow*, *weekly delay duration*, and *probability functions for delay duration*. In the next section, we give one of such analysis procedures.

4. A SAMPLE PROCEDURE

The procedures developed are based on the premises that data in the SFC database is reliable, and that necessary software routines will be developed to achieve the objectives of the data extraction specifications.

Some variables definitions in this section utilized standard database notation which calls for the prefixing of *fields* with the *table* name. For example, in the definition below, ***actvempl*** is the name of the table that contains the field *sdate* (start date of work), *stime*, *actcdate*, and *actctime*. The + sign indicates that the two fields will be concatenated. Boldface font is being used as a reminder of the notation.

$$W_{ijk} = \mathbf{actvempl.sdate+actvempl.stime} - \mathbf{actvempl.actcdate+actvempl.actctime}$$

All the field names used in the definitions correspond to the actual names in the physical SFC table. The meaning of each field can be found in the final report given to the TP-OAO office.

Most procedures described herein are composed of 4 major steps:

1. Transfer of summary information from SFC data tables to statistical tables residing in the mainframe. This transfer is done automatically by issuing SQL statements with the INSERT/SELECT combination.
2. Extraction of summary information from statistical tables to text files, so as to generate reports and/or uses STATGRAPHICS for analysis. This is done by executing SQL statement SELECT and utilizing QMF extraction capabilities in conjunction with OS/2 data transfer protocol.
3. Cleaning of data text files. When extracted the files have some header information that must be removed before STATGRAPHICS can use the file. This cleaning is done by running a C program.
4. Analysis of data using STATGRAPHICS, which includes the generation of appropriate graphs and summaries.

Some of the queries given in this report have been tested using the QMF version of SQL; others have not been tested; however, we believe that no major changes are needed to these queries to make them work as desired.

The nature to of the QMF version of SQL has forced us to extract fields that may seem redundant from the analysis point of view, but these *fields* and *grouping/column functions* are necessary to make QMF SQL work.

Total delay time for a flow

OBJECTIVES: To compute (for an entire flow) 1) the total delayed time, 2) the total delayed time per task, 3) the total delayed time per delay category

DEFINITIONS: A *delay* refers to the amount of time (hours) that technicians could not work on a task due to the lack of the appropriate resources (materials and personnel).

$$D_{ijk} = tsf01_delay.delay_duration$$

$$TDS_{ij} = \sum_{k=1}^{ND_{ij}} D_{ijk}$$

$$TDT_i = \sum_{j=1}^{NS_i} TDS_{ij} = \sum_{j=1}^{NS_i} \sum_{k=1}^{ND_{ij}} D_{ijk}$$

$$\text{TDF} = \sum_{i=1}^{\text{NT}} \text{TDT}_i = \sum_{i=1}^{\text{NT}} \sum_{j=1}^{\text{NS}_i} \sum_{k=1}^{\text{ND}_{ij}} \mathbf{D}_{ijk}$$

$$\text{TDCF}_l = \sum_{i=1}^{\text{NT}} \sum_{j=1}^{\text{NS}_i} \sum_{k=1}^{\text{ND}_{ij}} \mathbf{D}_{ijk} \quad \text{for each } \mathbf{D}_{ijk} \text{ with delay code } = l$$

where:

- \mathbf{D}_{ijk} = k th delay time duration (hours) for the i th task at the j th shop department, regardless of code type.
- TDS_{ij} = Total delay time for the i th task at the j th shop department
- TDT_i = Total delay time for the i th task
- TDF = Total delay time for the flow
- TDCF_l = Total delay time for code l in the entire flow
- ND_{ij} = number of delays, regardless of code type, that the i th task experienced at the j th shop department. $\text{ND}_{ij} = 1, 2, 3, \dots$
- NS_i = number of shops that worked on the i th task. $\text{NS}_i = 1, 2, 3, \dots$
- NT = number of executed tasks in the flow. $\text{NT} = 1, 2, 3, \dots$

WHEN TO EXECUTE PROCEDURE: The **TDF**, **TDF**, and **TDT** should each be computed at the end of a flow. The **TDF** report must be generated and trend analyses must follow, so as to identify any adjustments necessary in forthcoming flows.

DATA EXTRACTION and ADDITION:

1. A text file (*tdf.dat*) which contains **sts_num**, **orbiter_no**, **NT**, and **TDF** (query 4.2.7)
2. A text file (*tdt.dat*) which contains actual entries for each task in a flow (query 4.2.8)
3. A text file (*tdcf.dat*) which contains total delay time for delay code for the given flow.(query 4.2.9)
4. 1 record to *total_delay_flow* table (query 4.2.1)
5. **NT** records to *total_delay_task_flow* table (query 4.2.3)
6. Up to 84 records to *total_delay_code_flow* table (query 4.2.5)

PROCEDURE:

1. Ask user for **sts_no**, **orbiter_no**
2. Create an ASCII file (*before.sql*) with the following SQL commands. The lines numbered 4.1.i are not part of *before.sql*; they are comments in this report to explain what each SQL command is supposed to achieve.

4.2.1 *Command to determine the total delay time in the given flow. It produces 1 record with 3 fields (out of 4 in the target schema), and it inserts the*

record into the table total_delay_flow

```
INSERT INTO total_delay_flow (sts_no, tdf, num_of_delays)
SELECT   sts_no, orbiter_no, sum(delay_duration), count(*)
FROM     historic_tsf01_delay
WHERE    sts_no = (id for a flow, e.g. sts-51)
```

- 4.2.2 *Command to update the record created by query 4.2.1. This query fills the 4th field in the target schema.*

```
UPDATE total_delay_flow
SET   orbiter_no = (orbiter use in this sts_no. Given by the user)
WHERE sts_no = (id for a flow, e.g. sts-51)
```

- 4.2.3 *Command to find out the total delay time per task in the given flow. It produces one record per task executed in the given flow. Each resulting record has 6 fields (out of seven in the target schema). These records are inserted into the total_delay_task_flow*

```
INSERT INTO total_delay_task_flow (sts_no, task_no, wad, wad_type,
tdt, num_delays, avg_del, std_del, max_del, min_del)
SELECT   sts_no, task_no, max(wad), max(wad_type),
sum(delay_duration), count(*), avg(delay_duration),
std(delay_duration), max(delay_duration),
min(delay_duration)
FROM     historic_tsf01_delay
WHERE    sts_no = (id for a flow, e.g. sts-51)
GROUP BY sts_no, task_no
```

- 4.2.4 *Command to update the records created by query 4.2.3. This query fills the 7th field in the target schema.*

```
UPDATE total_delay_task_flow
SET   orbiter_no = (orbiter use in this sts_no. Given by the user)
WHERE sts_no = (id for a flow, e.g. sts-51)
```

- 4.2.5 *Command to find the total delay time per delay code in the given flow. This query creates up to 84 records since there are 84 different codes. Each record has 4 fields (out of 5 in the target schema). The records are inserted into the total_delay_code_flow table*

```
INSERT INTO total_delay_code_flow (sts_no, delay_cat, tdt,
occurrence, avg_del, std_del, max_del, min_del)
SELECT   sts_no, delay_cat, sum(delay_duration), count(*),
avg(delay_duration), std(delay_duration),
```

```

                                max(delay_duration), min(delay_duration)
FROM      tsf01_delay
WHERE     sts_no = (id for a flow, e.g. sts-51)
GROUP BY  sts_no, delay_cat

```

- 4.2.6 Command to update the records created by query 4.2.5. This query fills the 5th field in the target schema.

```

UPDATE total_delay_code_flow
SET    orbiter_no = (orbiter use in this sts_no. Given by the user)
WHERE  sts_no = (id for a flow, e.g. sts-51)

```

3. Execute SQL commands in *before.sql* (only if the process is automated)
4. Create ASCII files (names given at each query) with the outputs of the following SQL commands, using the QMF interface in conjunction with OS/2 data transfer protocol. Transfer text files to appropriate sub directory (where C program for cleaning data resides).

- 4.2.7 Query to extract delay time summaries for each flow (*tdf.dat*)

```

SELECT  sts_no, orbiter_no, tdf, num_of_wads
FROM    total_delay_flow

```

- 4.2.8 Query to extract delay time summaries for each task in the given flow (*tdt.dat*)

```

SELECT  tdt, wad, wad_type, num_delays
FROM    total_delay_task
WHERE   sts_no = (id for a flow, e.g. sts-51)
ORDER BY wad

```

- 4.2.9 Query to extract delay time summaries for each delay code in the given flow (*tdcf.dat*)

```

SELECT  delay_cat, tdcf, occurrence
FROM    total_delay_code_flow
WHERE   sts_no = (id for a flow, e.g. sts-51)
ORDER BY tdcf DESC

```

5. Open text files *tdf.dat*, *tdt.dat*, read only mode.
6. Create a new report file: *dtfststs##.out*
7. Count the number of records (NT) in *tdt.dat* and *tdcf.dat* (NDC)
8. Allocate arrays of size NT for *tdt*, *orbiter*, *wad*, *wad_type*, *num_delays*, of size NDC for code and *tdc*
9. Set *min_twt = 1000000* and *max_twt = -10000*.

10. *loop1*: i = 1 to NT
 - 10.1 Read from *tdt.dat*: *orbiter(i)*, *wad(i)*, *wad_type(i)*, *tdt(i)*
 - 10.2 If *min_tdt* > *twl(i)* Then
 - min_tdt* = *tdt(i)*
 - min_wad* = *wad(i)*
 endif
 - 10.3 If *max_tdt* < *tdt(i)* Then
 - max_tdt* = *tdt(i)*
 - max_wad* = *wad(i)*
 endif*end loop1*

11. *loop2*: i=1 to NDC
 - 11.1 Read from *tdcf.dat*: *code(i)*, *tdc(i)**end loop2*

12. Locate in *tdf.dat* the value of TDF for the given flow
13. Print report as laid out in Figure 4-4 and Figure 4-5.
14. Write a message interpreting the meaning of the values in the report
15. Run STATGRAPHICS and import text file (*tdf.dat*), by using the following instructions:
 - 15.1 click on utils, click on procedures macros, click on run macro
 - 15.2 choose *SFCITDF1* and press F6 key.

5. SUMMARY

The results reported here include the following:

1. A better understanding of the accessibility of the SFC database. We investigated issues pertaining to the physical storage of the SFC database as well as to the data base management system (DBMS) administering the SFC database. The fact that DB/2 is the DBMS used for SFC opens the possibility of accessing SFC 1) directly via SQL commands, 2) directly via a C-based program, specially designed and developed for SFC, or 3) indirectly via LSOC engineers. Each approach has its advantages and disadvantages, and we have discussed them in the TP-OAO report.
2. Specific data extraction protocols are being recommended. These protocols address data extraction from existing tables in the SFC database onto new tables as well as from the new tables onto ASCII files, so as to enable the interface with the statistical tool.

3. Specific recommendations on direct access to SFC by a NASA engineer are being made. These recommendations have been made to facilitate the communication of NASA's needs to LSOC software engineers.
4. Design of SMART, a framework that will enable NASA engineers to manipulate SFC data, analyze it, and do other modeling activities, without being concerned with the knowledge of SQL. This framework utilizes existing software to carry out the actual statistical analyses, and it includes a customized expert system to interpret the results before they are reported to NASA.
5. Specific procedures have been developed to conduct a variety of statistical analyses on the SFC data. These procedures give as much detail as necessary to enable the engineer to conduct analysis even if SMART is not implemented. Some of these procedures have been partially tested.
6. Specific subsets of SMART have been identified as candidates for an applied research proposal.
7. A worthwhile, unique experience for both Ms. Colucci and Dr. Centeno. Among other things that make this experience worthwhile are the facts that Ms. Colucci has seen first hand the application of some of the IE techniques she has been learning at school, and Dr. Centeno has accumulated new examples that can be used in courses such as Operations Research, Simulation, and Statistics. Similarly, NASA engineers have become more familiar with databases, SFC, and statistical analysis in general.

REFERENCES

- Cheu, D. & B. Linden** (1990), SQL Language Reference Manual, V6.0, ORACLE Corporation, 500 Oracle Parkway, Redwood Shores, CA 94065. (800) 633-0598.
- Coppola, A.** (1988), ORACLE Tools for MS-DOS and OS/2: Installation User's Guide, V6.0, ORACLE Corporation, 500 Oracle Parkway, Redwood Shores, CA 94065. (800) 633-0598.
- Lapin, L.L.** (1990), Probability and Statistics for Modern Engineering, Second Edition, PWS-Kent Publishing Company, Boston.
- Scheffe, H.** (1970), The Analysis of Variance, 6th printing, Wiley & Sons, Inc. Canada

1994018762
N94-23235

1993 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA

ES-37
197191
p. 20

MASS PROPERTIES MEASUREMENT SYSTEM DYNAMICS

PREPARED BY:	Dr. Keith L. Doty
ACADEMIC RANK:	Professor
UNIVERSITY AND DEPARTMENT:	University of Florida Department of Electrical Engineering
NASA/KSC	
DIVISION:	Mechanical Engineering
BRANCH:	Special Projects
NASA COLLEAGUE:	Bill Jones Todd Graham
DATE:	August 7, 1993
CONTRACT NUMBER:	University of Central Florida NASA-NGT-60002 Supplement: 11

TABLE OF CONTENTS

ACKNOWLEDGMENTS.....	3
ABSTRACT.....	3
1. INTRODUCTION	4
2. MPMS KINEMATICS.....	5
3. MPMS END FRAME JACOBIAN.....	6
4. MPMS DYNAMICS.....	7
Potential Energy Terms	7
Measuring MPMS Link Mass and Link Center-of-Gravity	8
MPMS Kinetic Energy Terms.....	9
5. DYNAMICS OF THE UNKNOWN MASS	13
6. MEASURING THE UNKNOWN MASS	14
7. PROPOSED DESIGN FOR MPMS.....	17
8. CONCLUSION.....	20
REFERENCES.....	20

ACKNOWLEDGMENTS

The author gratefully acknowledges the support of the University of Central Florida faculty and staff along with the I-NET and NASA staffs whose diligence and kindness make the NASA Faculty Fellowship program at the Kennedy Space Center so efficient and productive. Particular thanks to Bill Jones and Todd Graham of NASA; Kedron Wolcott and Rich Bennett of I-NET; and Ray Hosler and Kari Stiles of the University of Central Florida. A special thanks to Willis Crumpler of NASA who proposed the original concept of the MPMS. Finally, I would like to thank NASA Faculty Fellows Robert Byers and Narasimha Malladi for stimulating discussions.

ABSTRACT

The MPMS mechanism possess two revolute degrees-of-freedom and allows the user to measure the mass, center of gravity, and the inertia tensor of an unknown mass. This paper develops the dynamics of the Mass Properties Measurement System (MPMS) from the Lagrangian approach to illustrate the dependency of the motion on the unknown parameters.

1. INTRODUCTION

The Mass Properties Measurement System (MPMS), illustrated in Figure 1.1, consists of a serial kinematic mechanism with two intersecting revolute axes, z_1 and z_2 , that intersect with fixed angle α_1 . The joint angles and rates denoted by θ_1 , θ_2 and $\dot{\theta}_1$, $\dot{\theta}_2$, respectively, turn about the respective joint axes.

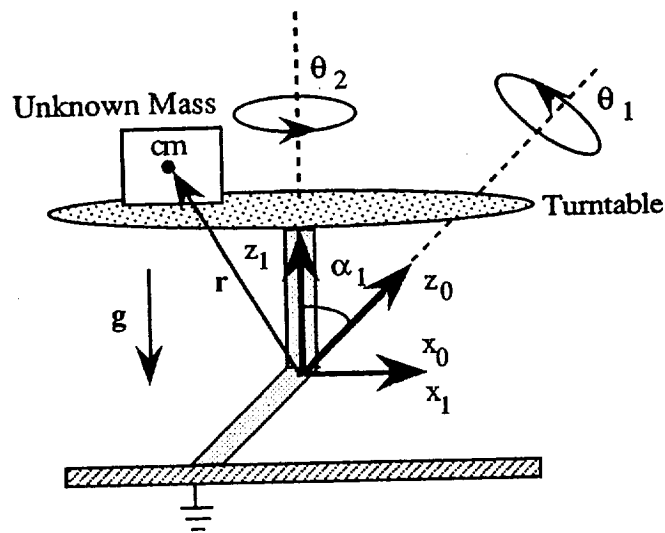


Figure 1.1 Mass Properties Measurement System (MPMS).

The function of the MPMS is to measure the mass, center-of-gravity and second-order mass moments of an unknown mass placed on the turntable. The mass and center-of-gravity can be determined from static measurements. The inertia tensor must be determined from the dynamics.

The structure of the Mass Properties Measurement System (MPMS) lends itself to a straightforward dynamics analysis using the Lagrangian approach. The dynamics analysis assumes α_1 as a parameter.

The symbolic software package *Mathematica* was used to produce and verify the equations presented here. In the analysis to follow, the standard Denevit-Hartenberg kinematic parameters of the MPMS were employed.

2. MPMS KINEMATICS

Table 2.1 lists the Denevit-Hartenberg kinematic parameters for the MPMS mechanism.

Table 2.1 : Kinematic Parameters for the MPMS Mechanism

Joint	d	θ	a	α	
1	r	0	θ_1	0	α_1
2	r	0	θ_2	0	0°

From the DH-parameters of the MPMS mechanism listed in Table 2.1, the two link transforms compute to

$$L_1 = \begin{pmatrix} {}^0R_1 & \mathbf{0} \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad L_2 = \begin{pmatrix} {}^1R_2 & \mathbf{0} \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (2-1a)$$

$$\text{where } {}^0R_1 = \begin{pmatrix} c_1 & -\tau_1 s_1 & \sigma_1 s_1 \\ s_1 & \tau_1 c_1 & -\sigma_1 c_1 \\ 0 & \sigma_1 & \tau_1 \end{pmatrix} \quad {}^1R_2 = \begin{pmatrix} c_2 & -s_2 & 0 \\ s_2 & c_2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (2-1b)$$

and $c_i := \cos(\theta_i)$, $s_i := \sin(\theta_i)$, $\tau_i := \cos(\alpha_i)$ and $\sigma_i := \sin(\alpha_i)$. The forward kinematics transform of the MPMS equals

$${}^0T_2 = L_1 L_2 \quad (2-2)$$

which computes to

$${}^0T_2 = \begin{pmatrix} {}^0R_2 & \mathbf{0} \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (2-3)$$

where

$${}^0\mathbf{R}_2 = \begin{pmatrix} c_1 c_2 - \tau_1 s_1 s_2 & -c_1 s_2 - \tau_1 s_1 c_2 & \sigma_1 s_1 \\ s_1 c_2 + \tau_1 c_1 s_2 & -s_1 s_2 + \tau_1 c_1 c_2 & -\sigma_1 c_1 \\ \sigma_1 s_2 & \sigma_1 c_2 & \tau_1 \end{pmatrix} \quad (2-4)$$

The forward analysis presented here serves as reference. The rotation part indicates how to change frame F_2 vector representations into frame F_0 representations. The dynamics analysis presented later will make use of the forward kinematics ${}^0\mathbf{T}_2$.

3. MPMS END FRAME JACOBIAN

The Jacobian of the MPMS relates the joint-rates $\dot{\mathbf{q}} = [\dot{\theta}_1 \quad \dot{\theta}_2]^\tau$ to the frame-velocity $\mathbf{V} = [\mathbf{v}^\tau \quad \boldsymbol{\omega}^\tau]^\tau$ of the end-frame,

$$\mathbf{V} = \mathbf{J} \dot{\mathbf{q}}. \quad (3-1)$$

The Jacobian of the MPMS computes [1] [2] to

$${}^{2,0}\mathbf{J}_{2,2} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ \sigma_1 s_2 & 0 \\ \sigma_1 c_2 & 0 \\ \tau_1 & 1 \end{pmatrix} \quad (3-2)$$

The leading superscript 2 means that this Jacobian is expressed in frame F_2 while the 0 indicates the motion the end-frame, designated by the second subscript, is relative to the base frame F_0 of the MPMS. The first subscript indicates the frame origin at which the linear velocity is measured.

For convenience, we write

$${}^{2,0}\mathbf{J}_{2,2} = \begin{pmatrix} {}^{2,0}\mathbf{J}_{v2,2} \\ {}^{2,0}\mathbf{J}_{\omega,2} \end{pmatrix}, \text{ where } {}^{2,0}\mathbf{J}_{v2,2} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \text{ and } {}^{2,0}\mathbf{J}_{\omega,2} = \begin{pmatrix} \sigma_1 s_2 & 0 \\ \sigma_1 c_2 & 0 \\ \tau_1 & 1 \end{pmatrix} \quad (3-3)$$

4. MPMS DYNAMICS

The Lagrangian approach to dynamics of the mechanism requires the calculation of the kinetic and potential energy of the moving masses.

Potential Energy Terms

Assume the center-of-mass vectors for the first and second links equal r_1^* and r_2^* . These vectors are constants in their own frames,

$${}^1r_1 = \begin{pmatrix} r_{x1} \\ r_{y1} \\ r_{z1} \end{pmatrix} \quad \text{and} \quad {}^2r_2 = \begin{pmatrix} r_{x2} \\ r_{y2} \\ r_{z2} \end{pmatrix} \quad (4-1)$$

The gravitational field vector in Figure 1.1 equals

$${}^0g = g_c \begin{pmatrix} 0 \\ \sigma_1 \\ -\tau_1 \end{pmatrix} \quad (4-2)$$

and the potential energy of a mass m at position p equals

$$P = -m g^\tau p \quad (4-3)$$

hence, the potential energies P_1 and P_2 of the first and second link equal

$$P_1 = -m_1 g^\tau r_1^* = -m_1 {}^0g^\tau {}^0R_1 {}^1r_1^* \quad (4-4a)$$

and

$$P_2 = -m_2 g^\tau r_2^* = -m_2 {}^0g^\tau {}^0R_2 {}^2r_2^* \quad (4-4b)$$

The torques associated with changes in the potential energy of the mechanism links equal $\tau_{pe1} = \frac{\partial P_1}{\partial \theta_1} + \frac{\partial P_2}{\partial \theta_1}$ and $\tau_{pe2} = \frac{\partial P_1}{\partial \theta_2} + \frac{\partial P_2}{\partial \theta_2} = \frac{\partial P_2}{\partial \theta_2}$. These torques compute to

$$\tau_{pe1} = m_1 \sigma_1 g_c \{ r_{x1} c_1 + r_{y1} \tau_1 s_1 + r_{z1} \sigma_1 s_1 \} + m_2 \sigma_1 g_c \{ r_{x2} (c_1 c_2 - \tau_1 s_1 s_2) + r_{y2} (-\tau_1 c_2 s_1 - c_1 s_2) + r_{z2} \sigma_1 s_1 \} \quad (4-4a)$$

$$\tau_{pe2} = m_2 \sigma_1 g_c \{ r_{x2} (-c_2 \tau_1 + \tau_1 c_1 c_2 - s_1 s_2) + r_{y2} (\tau_1 s_2 - c_2 s_1 - \tau_1 c_1 s_2) \} \quad (4-4b)$$

Equations (4-4) will provide the means for measuring the mass and center-of-gravity for each link.

Measuring MPMS Link Mass and Link Center-of-Gravity

By measuring the balancing torque on joint one and setting the joint angles at different angles one can determine the required information about the first-order mass moments of each link. This information will allow us to subtract the gravity torque terms due to the mass of the MPMS when we desire to measure the mass properties of the unknown mass.

Experiment 4.1 $\theta_1 = 0, \theta_2 = 0,$

$$\tau_{x1} := m_1 \sigma_1 g_c r_{x1} + m_2 \sigma_1 g_c r_{x2} \quad (4-5)$$

Experiment 4.2 $\theta_1 = 0, \theta_2 = \pi/2,$

$$\tau_{x2} := m_1 \sigma_1 g_c r_{x1} - m_2 \sigma_1 g_c r_{y2} \quad (4-6)$$

From these two measurements one computes

$$m_2 r_{y2} = \frac{\tau_{x1} - \tau_{x2}}{2 \sigma_1 g_c} \quad \text{and} \quad m_2 r_{x1} = \frac{\tau_{x1} + \tau_{x2}}{2 \sigma_1 g_c} \quad (4-7)$$

Experiment 4.3 $\theta_1 = \pi/2, \theta_2 = 0,$

$$\tau_{x3} := m_1 \sigma_1 g_c (r_{y1} \tau_1 + r_{z1} \sigma_1) + m_2 \sigma_1 g_c (-r_{y2} \tau_1 + r_{z2} \sigma_1) \quad (4-8)$$

Experiment 4.4 $\theta_1 = \pi/2, \theta_2 = \pi/2,$

$$\tau_{x4} := m_1 \sigma_1 g_c (r_{y1} \tau_1 + r_{z1} \sigma_1) + m_2 \sigma_1 g_c (-\tau_1 r_{x2} + r_{z2} \sigma_1) \quad (4-9)$$

From *Experiments 3 and 4* one can find

$$m_2 r_{x2} = \frac{\tau_{x3} - \tau_{x4}}{\sigma_1 \tau_1 g_c} + m_2 r_{y2} \quad (4-10)$$

Relation (4-10) does not apply when $\tau_1 = 0$, i.e., when $\alpha_1 = \pi/2$.

From (4-7) and (4-10) we can develop a linear relation in the other unknowns,

$$\begin{aligned} \tau'_{pe1} := \frac{1}{s_1} \left\{ \frac{\tau_{pe1}}{\sigma_1 g_c} - m_1 r_{x1} c_1 - m_2 r_{x2} (c_1 c_2 - \tau_1 s_1 s_2) \right. \\ \left. + m_2 r_{y2} (-\tau_1 c_2 s_1 - c_1 s_2) \right\} = m_1 \{ r_{y1} \tau_1 + r_{z1} \sigma_1 \} + m_2 r_{z2} \sigma_1 \quad (4-11) \end{aligned}$$

Since the unknowns in (4-11) have fixed coefficients, those unknowns cannot be resolved further.

In summary, (4-7), (4-10) and (4-11) provide the necessary information for determining the MPMS gravity terms. The actual mass values and center-of-gravity terms do not need to be determined completely.

MPMS Kinetic Energy Terms

The total kinetic energy K of the MPMS motion equals $K = K_1 + K_2$, where K_i , $i = 1, 2$, equals the kinetic energy of link L_i defined with respect to the origin define in Figure 1.1. The joint torques τ_{ke1} and τ_{ke2} required to generate the motion, assuming a conservative system, equal, according to Lagrange,

$$\tau_{kei} = \frac{d}{dt} \left[\frac{\partial K}{\partial \dot{\theta}_i} \right] - \frac{\partial K}{\partial \theta_i}, \quad i = 1, 2. \quad (4-12)$$

The torque terms associated with each kinetic energy

$$\tau_{keji} = \frac{d}{dt} \left[\frac{\partial K_j}{\partial \dot{\theta}_i} \right] - \frac{\partial K_j}{\partial \theta_i}, \quad j = 1, 2; i = 1, 2. \quad (4-13)$$

sum to produce the total torque for that joint,

$$\tau_{kei} = \tau_{ke1i} + \tau_{ke2i} \quad (4-14)$$

Since the origin does not translate, the kinetic energy for both links of the MPMS equals the rotational kinetic energy,

$$K_i = \frac{1}{2} \omega_i \langle \cdot \rangle I_i \omega_i, \quad i = 1, 2, \quad (4-15)$$

where ω_i equals the angular velocity and I_i equals the standard inertia matrix for link L_i . In frame F_i the matrix I_i has constant terms.

From (3-1) and (3-3),

$${}^{1,0}\omega_1 = \dot{\theta}_1 \begin{pmatrix} 0 \\ \sigma_1 \\ \tau_1 \end{pmatrix}, \quad {}^{2,0}\omega_2 = {}^{2,0}J_{\omega,2} \begin{pmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \end{pmatrix} \quad (4-16)$$

The kinetic energy of the first link

$$K_1 = \frac{1}{2} {}^{1,0}\omega_1 \langle \cdot \rangle {}^1I_1 {}^{1,0}\omega_1 \quad (4-17)$$

does not depend on either of the joint variables nor $\dot{\theta}_2$, hence,

$$\tau_{ke11} := \frac{d}{dt} \left[\frac{\partial K_1}{\partial \dot{\theta}_1} \right] - \frac{\partial K_1}{\partial \theta_1} = \ddot{\theta}_1 (0 \quad \sigma_1 \quad \tau_1)^\tau {}^1I_1 \begin{pmatrix} 0 \\ \sigma_1 \\ \tau_1 \end{pmatrix} \quad (4-18)$$

$$\tau_{ke12} := \frac{d}{dt} \left[\frac{\partial K_1}{\partial \dot{\theta}_2} \right] - \frac{\partial K_1}{\partial \theta_2} = 0 \quad (4-19)$$

The kinetic energy of the second link equals

$$K_2 = \frac{1}{2} \begin{pmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \end{pmatrix} \langle \cdot \rangle {}^{2,0}J_{\omega,2}^\tau {}^2I_2 {}^{2,0}J_{\omega,2} \begin{pmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \end{pmatrix}, \quad (4-20)$$

and does not depend upon θ_1 . Hence, the torque contribution by joint one to the motion of the second link equals

$$\begin{aligned} \tau_{ke21} &:= \frac{d}{dt} \left[\frac{\partial K_2}{\partial \dot{\theta}_1} \right] - \frac{\partial K_2}{\partial \theta_1} = \frac{d}{dt} \left[\frac{\partial K_2}{\partial \dot{\theta}_1} \right] = {}^{2,0} \mathbf{J}_{\omega 2}^{\tau} {}^2 \mathbf{I}_2 {}^{2,0} \mathbf{J}_{\omega,2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \langle \cdot \rangle \begin{pmatrix} \ddot{\theta}_1 \\ \ddot{\theta}_2 \end{pmatrix} + \\ & {}^{2,0} \dot{\mathbf{J}}_{\omega 2}^{\tau} {}^2 \mathbf{I}_2 {}^{2,0} \mathbf{J}_{\omega,2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \langle \cdot \rangle \begin{pmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \end{pmatrix} + {}^{2,0} \mathbf{J}_{\omega 2}^{\tau} {}^2 \mathbf{I}_2 {}^{2,0} \dot{\mathbf{J}}_{\omega,2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \langle \cdot \rangle \begin{pmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \end{pmatrix} \end{aligned} \quad (4-21)$$

Similarly, the torque contribution by joint two to the motion of link two equals

$$\begin{aligned} \tau_{ke22} &:= \frac{d}{dt} \left[\frac{\partial K_2}{\partial \dot{\theta}_2} \right] - \frac{\partial K_2}{\partial \theta_2} = {}^{2,0} \mathbf{J}_{\omega 2}^{\tau} {}^2 \mathbf{I}_2 {}^{2,0} \mathbf{J}_{\omega,2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \langle \cdot \rangle \begin{pmatrix} \ddot{\theta}_1 \\ \ddot{\theta}_2 \end{pmatrix} + \\ & {}^{2,0} \dot{\mathbf{J}}_{\omega 2}^{\tau} {}^2 \mathbf{I}_2 {}^{2,0} \mathbf{J}_{\omega,2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \langle \cdot \rangle \begin{pmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \end{pmatrix} + {}^{2,0} \mathbf{J}_{\omega 2}^{\tau} {}^2 \mathbf{I}_2 {}^{2,0} \dot{\mathbf{J}}_{\omega,2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \langle \cdot \rangle \begin{pmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \end{pmatrix} \\ & - \begin{pmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \end{pmatrix} \langle \cdot \rangle \left\{ \frac{\partial {}^{2,0} \mathbf{J}_{\omega 2}^{\tau}}{\partial \theta_2} {}^2 \mathbf{I}_2 {}^{2,0} \mathbf{J}_{\omega,2} + {}^{2,0} \mathbf{J}_{\omega 2}^{\tau} {}^2 \mathbf{I}_2 \frac{\partial {}^{2,0} \mathbf{J}_{\omega,2}}{\partial \theta_2} \right\} \begin{pmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \end{pmatrix} \end{aligned} \quad (4-22)$$

From (3-3) compute

$${}^{2,0} \dot{\mathbf{J}}_{\omega,2} = \sigma_1 \dot{\theta}_2 \begin{pmatrix} c_2 & 0 \\ -s_2 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \frac{\partial {}^{2,0} \mathbf{J}_{\omega,2}}{\partial \theta_2} = \sigma_1 \begin{pmatrix} c_2 & 0 \\ -s_2 & 0 \\ 0 & 0 \end{pmatrix} \quad (4-23)$$

and deduce

$${}^{2,0} \dot{\mathbf{J}}_{\omega 2}^{\tau} {}^2 \mathbf{I}_2 {}^{2,0} \mathbf{J}_{\omega,2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = {}^{2,0} \mathbf{J}_{\omega 2}^{\tau} {}^2 \mathbf{I}_2 {}^{2,0} \dot{\mathbf{J}}_{\omega,2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad (4-24)$$

$${}^{2,0} \mathbf{J}_{\omega 2}^{\tau} {}^2 \mathbf{I}_2 {}^{2,0} \mathbf{J}_{\omega,2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sigma_1 (s_2 I_{xz2} + c_2 I_{yz2}) \\ I_{zz2} \end{pmatrix}, \quad (4-25)$$

Relations (4-24) and (4-25) allows us to simplify (4-22), and, coupled with (4-19) yields the joint-two, kinetic energy derived torque τ_{ke2} ,

$$\begin{aligned} \tau_{ke2} = \tau_{ke22} + \tau_{ke12} = & {}^{2,0}\mathbf{J}_{\omega 2}^{\tau} {}^2\mathbf{I}_2 {}^{2,0}\mathbf{J}_{\omega,2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \langle \bullet \rangle \begin{pmatrix} \ddot{\theta}_1 \\ \ddot{\theta}_2 \end{pmatrix} + \\ - \begin{pmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \end{pmatrix} \langle \bullet \rangle & \left\{ \frac{\partial {}^{2,0}\mathbf{J}_{\omega 2}^{\tau}}{\partial \theta_2} {}^2\mathbf{I}_2 {}^{2,0}\mathbf{J}_{\omega,2} + {}^{2,0}\mathbf{J}_{\omega 2}^{\tau} {}^2\mathbf{I}_2 \frac{\partial {}^{2,0}\mathbf{J}_{\omega,2}}{\partial \theta_2} \right\} \begin{pmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \end{pmatrix} \end{aligned} \quad (4-26)$$

The joint-one, kinetic energy derived torque τ_{ke1} equals the sum of (4-18) and (4-21),

$$\begin{aligned} \tau_{ke1} = \tau_{ke11} + \tau_{ke21} = & \ddot{\theta}_1 \mathbf{I}_{T1} + {}^{2,0}\mathbf{J}_{\omega 2}^{\tau} {}^2\mathbf{I}_2 {}^{2,0}\mathbf{J}_{\omega,2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \langle \bullet \rangle \begin{pmatrix} \ddot{\theta}_1 \\ \ddot{\theta}_2 \end{pmatrix} \\ + \left\{ {}^{2,0}\dot{\mathbf{J}}_{\omega 2}^{\tau} {}^2\mathbf{I}_2 {}^{2,0}\mathbf{J}_{\omega,2} + & {}^{2,0}\mathbf{J}_{\omega 2}^{\tau} {}^2\mathbf{I}_2 {}^{2,0}\dot{\mathbf{J}}_{\omega,2} \right\} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \langle \bullet \rangle \begin{pmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \end{pmatrix} \end{aligned} \quad (4-27)$$

where

$$\mathbf{I}_{T1} := (0 \quad \sigma_1 \quad \tau_1)^{\tau} {}^1\mathbf{I}_1 \begin{pmatrix} 0 \\ \sigma_1 \\ \tau_1 \end{pmatrix} \quad (4-28)$$

cannot be resolved further by any joint measurements

Using *Mathematica* to expand (4-27) and (4-26) yields

$$\begin{aligned}
 \tau_{ke1} = & \ddot{\theta}_1 I_{T1} + I_{zz} \tau_1 (\ddot{\theta}_2 + \ddot{\theta}_1 \tau_1) + \sigma_1 \{ I_{yy} c_2 \sigma_1 (\ddot{\theta}_1 c_2 - 2 \dot{\theta}_1 \dot{\theta}_2 s_2) \\
 & + I_{xx} \sigma_1 s_2 (2 \dot{\theta}_1 \dot{\theta}_2 c_2 + \ddot{\theta}_1 s_2) \\
 & + I_{yz} (-\ddot{\theta}_2 c_2 - 2 \ddot{\theta}_1 c_2 \tau_1 + \dot{\theta}_2^2 s_2 + 2 \dot{\theta}_1 \dot{\theta}_2 \tau_1 s_2) \\
 & + I_{xz} (-\dot{\theta}_2^2 c_2 - 2 \dot{\theta}_1 \dot{\theta}_2 c_2 \tau_1 - \ddot{\theta}_2 s_2 - 2 \ddot{\theta}_1 \tau_1 s_2) \\
 & - I_{xy} \sigma_1 (2 \dot{\theta}_1 \dot{\theta}_2 c_{2\theta 2} + \ddot{\theta}_1 s_{2\theta 2}) \} \quad (4-29)
 \end{aligned}$$

and

$$\begin{aligned}
 \tau_{ke2} = & I_{zz} (\ddot{\theta}_2 + \ddot{\theta}_1 \tau_1) + \sigma_1 \{ I_{xy} \dot{\theta}_1^2 c_{2\theta 2} \sigma_1 + I_{xz} (\dot{\theta}_1^2 c_2 \tau_1 - \ddot{\theta}_1 s_2) \\
 & + I_{yz} (-\ddot{\theta}_1 c_2 - \dot{\theta}_1^2 \tau_1 s_2) - \frac{1}{2} I_{xx} \dot{\theta}_1^2 \sigma_1 s_{2\theta 2} + \frac{1}{2} I_{yy} \dot{\theta}_1^2 \sigma_1 s_{2\theta 2} \} \quad (4-30)
 \end{aligned}$$

Equations (4-29) and (4-30) yield the essential relations for determining the unknown inertia matrix I_2 and the term I_{T1} from torque, velocity and acceleration measurements.

5. DYNAMICS OF THE UNKNOWN MASS

The functional objective of the MPMS is to measure the unknown mass properties of an object placed on the table of link two (Figure 1.1). The unknown mass is rigidly attached to link two during the measurements. Conceptually, this makes the unknown object a part of the second link, hence, the combined inertia tensor I_2' equals

$$I_2' = I_2 + I_u \quad (5-1)$$

where I_u is the inertia matrix of the unknown mass described about the common origin of frame F_0 and F_1 . The dynamics analysis in Section 4 applies to this problem with I_2 replaced by I_2' . Thus, the torques derived from the kinetic

energy of the unknown mass alone have the form of $\tau_{ke1} - \ddot{\theta}_1 I_{T1}$ in (4-29) and τ_{ke2} in (4-30) with the inertia terms replaced by those of the unknown mass,

$$\begin{aligned} \tau_{keu1} := & I_{zzu} \tau_1 (\ddot{\theta}_2 + \ddot{\theta}_1 \tau_1) + \sigma_1 \{ I_{yyu} c_2 \sigma_1 (\ddot{\theta}_1 c_2 - 2 \dot{\theta}_1 \dot{\theta}_2 s_2) \\ & + I_{xxu} \sigma_1 s_2 (2 \dot{\theta}_1 \dot{\theta}_2 c_2 + \ddot{\theta}_1 s_2) \\ & + I_{yzu} (-\ddot{\theta}_2 c_2 - 2 \ddot{\theta}_1 c_2 \tau_1 + \dot{\theta}_2^2 s_2 + 2 \dot{\theta}_1 \dot{\theta}_2 \tau_1 s_2) \\ & + I_{xzu} (-\dot{\theta}_2^2 c_2 - 2 \dot{\theta}_1 \dot{\theta}_2 c_2 \tau_1 - \ddot{\theta}_2 s_2 - 2 \ddot{\theta}_1 \tau_1 s_2) \\ & - I_{xyu} \sigma_1 (2 \dot{\theta}_1 \dot{\theta}_2 c_{2\theta 2} + \ddot{\theta}_1 s_{2\theta 2}) \} \end{aligned} \quad (5-2)$$

and

$$\begin{aligned} \tau_{ke2} = & I_{zzu} (\ddot{\theta}_2 + \ddot{\theta}_1 \tau_1) + \sigma_1 \{ I_{xyu} \dot{\theta}_1^2 c_{2\theta 2} \sigma_1 + I_{xzu} (\dot{\theta}_1^2 c_2 \tau_1 - \ddot{\theta}_1 s_2) \\ & + I_{yzu} (-\ddot{\theta}_1 c_2 - \dot{\theta}_1^2 \tau_1 s_2) - \frac{1}{2} I_{xxu} \dot{\theta}_1^2 \sigma_1 s_{2\theta 2} + \frac{1}{2} I_{yyu} \dot{\theta}_1^2 \sigma_1 s_{2\theta 2} \} \end{aligned} \quad (5-3)$$

Equations (5-2) and (5-3) yield the essential relations for determining the unknown inertia matrix I_u from torque, velocity and acceleration measurements.

6. MEASURING THE UNKNOWN MASS

Since the torque equations are essentially the same for the inertia terms of link two of the MPMS and the unknown mass, only the experimental technique for the latter will be developed here.

The first question to resolve:

Will measurements of the torque, angular position, velocity and acceleration of joint one provide sufficient data to compute all the inertia parameters of the unknown mass?

The answer will to this question is *No*. To verify this claim, note that, potentially, all six independent parameters in I_u might be determined from (5-2) given that

Experiment 6.1: $\dot{\theta}_2 = 0, \ddot{\theta}_2 = 0,$

$$\tau_{keul} := \ddot{\theta}_1 \{ I_{zzu} \tau_1^2 + \sigma_1^2 \{ I_{yyu} c_2^2 + I_{xxu} s_2^2 - I_{xyu} s_{2\theta_2} \} - 2 \sigma_1 \tau_1 \{ I_{yzu} c_2 + I_{xzu} s_2 \} \}$$

which, in inner product form, equals,

$$\tau_{keul} = \ddot{\theta}_1 \mathbf{k} \langle \bullet \rangle \mathbf{I}_v \quad (6-1)$$

where \mathbf{k} is the *coefficient vector* and \mathbf{I}_v a six-vector of the independent parameters in I_u :

$$\mathbf{k} := \begin{pmatrix} \tau_1^2 \\ \sigma_1^2 c_2^2 \\ \sigma_1^2 s_2^2 \\ -\sigma_1^2 s_{2\theta_2} \\ -2 \sigma_1 \tau_1 c_2 \\ -2 \sigma_1 \tau_1 s_2 \end{pmatrix} \text{ and } \mathbf{I}_v := \begin{pmatrix} I_{zzu} \\ I_{yyu} \\ I_{xxu} \\ I_{xyu} \\ I_{yzu} \\ I_{xzu} \end{pmatrix} \quad (6-2)$$

If a vector can be found orthogonal to the coefficient vector for all choices of MPMS configurations, then, six independent coefficient vectors \mathbf{k} cannot be obtained and, therefore, six independent equations in the six unknowns in \mathbf{I}_v cannot be obtained. The vector $\mathbf{x} := [\sigma_1^2 \quad -\tau_1^2 \quad -\tau_1^2 \quad 0 \quad 0 \quad 0]^T \neq \mathbf{0}$ is such a vector: $\mathbf{k} \langle \bullet \rangle \mathbf{x} = 0$ for all possible \mathbf{k} . This implies no non-singular measurement matrix \mathbf{M} can be formed from the set of possible \mathbf{k} that will yield six independent equations. Since $\mathbf{M} \mathbf{x} = \mathbf{0}$ for any \mathbf{M} whose rows are constructed from different \mathbf{k} 's, the matrix \mathbf{M} possesses a non-zero vector in its null space and, therefore, must be singular. Note that this observation holds for any choice of angle α_1 between the two joint axes of the mechanism.

Observe that \mathbf{x} cannot be solution to any inertia problem because it forces at least one of the diagonal terms of I_u negative, an impossible situation physically.

Substitute $\mathbf{x} = \mathbf{I}_v$, this is not a physically possible \mathbf{I}_v , into the general expression for τ_{keu1} and find that $\tau_{keu1} = I_{zzu} \sigma_1 \tau_1 \ddot{\theta}_2$ which proves that, minimally, one must measure $\ddot{\theta}_2$ on the second joint.

Experiment 6.2 $\dot{\theta}_1 = 0, \ddot{\theta}_1 = 0,$

$$\tau_{keu1} := I_{zzu} \tau_1 \ddot{\theta}_2 + \sigma_1 I_{yzu} (-\ddot{\theta}_2 c_2 + \dot{\theta}_2^2 s_2) + \sigma_1 I_{xzu} (-\dot{\theta}_2^2 c_2 - \ddot{\theta}_2 s_2) \quad (6-3)$$

hence,

$$\tau_{keu1} = \mathbf{k} \langle \cdot \rangle \mathbf{I}_s \quad (6-4)$$

where \mathbf{k} is the *coefficient vector* and \mathbf{I}_s a three-vector :

$$\mathbf{k} := \begin{pmatrix} \tau_1 \ddot{\theta}_2 \\ \sigma_1 (-\ddot{\theta}_2 c_2 + \dot{\theta}_2^2 s_2) \\ \sigma_1 (-\dot{\theta}_2^2 c_2 - \ddot{\theta}_2 s_2) \end{pmatrix} \text{ and } \mathbf{I}_s := \begin{pmatrix} I_{zzu} \\ I_{yzu} \\ I_{xzu} \end{pmatrix} \quad (6-5)$$

By measuring the torque on joint one and the angular position, velocity and acceleration on joint two, one can measure three independent equations (6-4) to compute $I_{zzu}, I_{yzu}, I_{xzu}$. For example, if $\theta_2 = \frac{\pi t^2}{4}, \dot{\theta}_2 = \frac{\pi t}{2}, \ddot{\theta}_2 = \frac{\pi}{2}, 0 \leq t \leq \sqrt{2}$, and $\alpha_1 = \tan^{-1}[\sqrt{2}]$, the measurement values of (6-3) for $t=0, t=1, t = \sqrt{2}$ lead to the measurement matrix

$$\mathbf{M} = \begin{pmatrix} \mathbf{k}^\tau(t=0) \\ \mathbf{k}^\tau(t=1) \\ \mathbf{k}^\tau(t=\sqrt{2}) \end{pmatrix} = \begin{pmatrix} 0.9069 & 0 & -1.28255 \\ 0.9069 & -2.33145 & 0.517655 \\ 0.9069 & -1.28255 & 4.02925 \end{pmatrix}$$

whose determinant equals $\det[M] = -9.13734$, proving that M non-singular and that $M\mathbf{I}_s = \begin{pmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \end{pmatrix}$ can be solved for \mathbf{I}_s . With \mathbf{I}_s known, the remaining inertia parameters can be determined from the measurements in Experiment 6.1.

Conclusion: If one measures torques and angular acceleration on joint one, then the angular position, velocity and acceleration on joint two must be measured (or calculated) in order to obtain all the inertia parameters of the unknown mass.

7. PROPOSED DESIGN FOR MPMS

Based upon the analysis in this paper, we note that

1. All the mass properties of the MPMS and the unknown mass itself cannot be determined by measurements performed only on joint axis one quantities.

Discussions with Kedron Wolcott on the construction of the current prototype indicates

2. Construction of the bearings for the MPMS for a twist $0 < \alpha_1 < 90^\circ$ present significant cost and design penalties.

A talk with Richard Bennett indicates that

3. The current design occupies a large volume due to the twist $\alpha_1 \approx 54^\circ$ between the two axes.
4. The current design is not easily scalable to handle larger loads.

Based upon the four considerations above, the author proposes that the two joint axes of the MPMS be oriented at right angles, $\alpha_1 = 90^\circ$, as shown in Figure 7.1. This design eliminates the problems mentioned in items 2,3, and 4, but will require a torque sensor on the second joint. The following analysis demonstrates the latter observation.

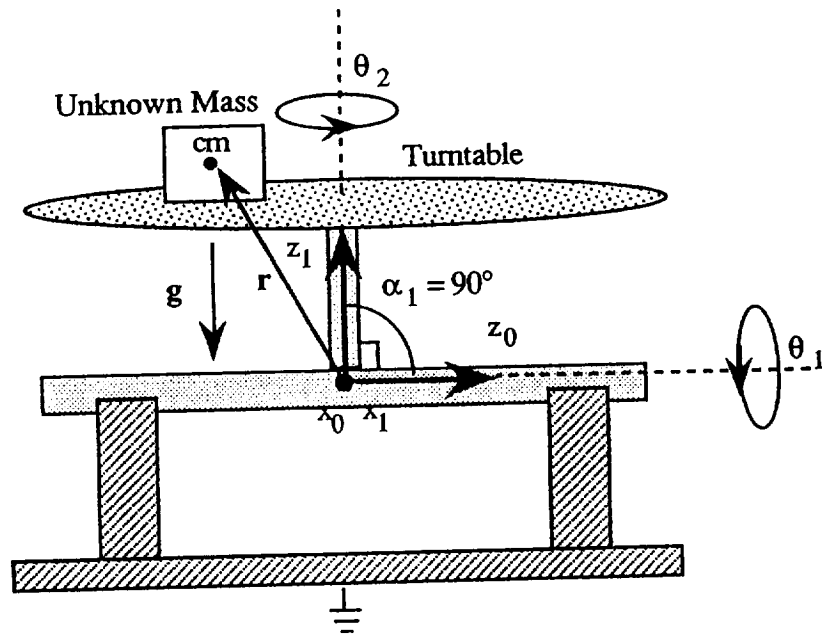


Figure 7.1 Proposed MPMS.

Evaluate (5-2) and (5-3) for $\alpha_1 = 90^\circ$,

$$\begin{aligned} \tau_{keu1} := & I_{yyu} c_2 (\ddot{\theta}_1 c_2 - 2 \dot{\theta}_1 \dot{\theta}_2 s_2) + I_{xxu} s_2 (2 \dot{\theta}_1 \dot{\theta}_2 c_2 + \ddot{\theta}_1 s_2) \\ & + I_{yzu} (-\ddot{\theta}_2 c_2 + \dot{\theta}_2^2 s_2) - I_{xzu} (\dot{\theta}_2^2 c_2 + \ddot{\theta}_2 s_2) - I_{xyu} (2 \dot{\theta}_1 \dot{\theta}_2 c_2 \theta_2 + \ddot{\theta}_1 s_2 \theta_2) \end{aligned} \quad (7-1)$$

and

$$\begin{aligned} \tau_{ke2} = & I_{zzu} \ddot{\theta}_2 + I_{xyu} \dot{\theta}_1^2 c_2 \theta_2 - I_{xzu} \ddot{\theta}_1 s_2 - I_{yzu} \ddot{\theta}_1 c_2 \\ & - \frac{1}{2} I_{xxu} \dot{\theta}_1^2 s_2 \theta_2 + \frac{1}{2} I_{yyu} \dot{\theta}_1^2 s_2 \theta_2 \end{aligned} \quad (7-2)$$

Observe that I_{zzu} cannot be determined from (7-1). However, if one measures τ_{ke2} and fixes $\dot{\theta}_1 = 0$, $\ddot{\theta}_1 = 0$, then (7-2) yields

$$I_{zzu} = \frac{\tau_{ke2}}{\ddot{\theta}_2} \quad \text{when } \dot{\theta}_1 = 0, \ddot{\theta}_1 = 0. \quad (7-3)$$

To measure the other inertia parameters set $\dot{\theta}_2 = 0, \ddot{\theta}_2 = 0$ in (7-1) and (7-2) to obtain

$$\tau_{keu1} := (I_{yyu} c_2^2 + I_{xxu} s_2^2 - I_{xyu} s_2 \theta_2) \ddot{\theta}_1 \quad (7-4)$$

$$(I_{xyu} c_2 \theta_2 - \frac{1}{2} I_{xxu} s_2 \theta_2 + \frac{1}{2} I_{yyu} s_2 \theta_2) \dot{\theta}_1^2 - \tau_{ke2} = (I_{xzu} s_2 + I_{yzu} c_2) \ddot{\theta}_1 \quad (7-5)$$

The procedure would be to use (7-4) to obtain $I_{yyu}, I_{xxu}, I_{xyu}$ and then substitute these values in (7-5) to obtain the remaining inertia terms I_{xzu}, I_{yzu} .

Experiments 7.1 and 7.2 entail measuring both torques, $\dot{\theta}_1$ and $\ddot{\theta}_1$ with the second joint axis fixed, respectively, at $\theta_2 = 0, \pi/2$ and $\pi/4$. For the last angle, measurement of the torque on the second joint is not required.

Experiment 7.1 $\dot{\theta}_2 = 0, \ddot{\theta}_2 = 0$

$$1. \theta_2 = 0, \quad I_{yyu} = \frac{\tau_{ke1}}{\ddot{\theta}_1} \quad (7-6a)$$

$$2. \theta_2 = \pi/2, \quad I_{xxu} = \frac{\tau_{ke1}}{\ddot{\theta}_1} \quad (7-6b)$$

$$3. \theta_2 = \pi/4, \quad I_{xyu} = \frac{I_{yyu} c_2^2 + I_{xxu} s_2^2 - \tau_{ke1}}{\ddot{\theta}_1} \quad (7-6c)$$

Experiment 7.2 $\dot{\theta}_2 = 0, \ddot{\theta}_2 = 0$

$$1. \theta_2 = 0, I_{yzu} = \frac{I_{xyu} \dot{\theta}_1^2 - \tau_{ke2}}{\ddot{\theta}_1} \quad (7-6d)$$

$$2. \theta_2 = \pi/2, I_{xzu} = \frac{-I_{xyu} \dot{\theta}_1^2 - \tau_{ke2}}{\ddot{\theta}_1} \quad (7-6e)$$

Equations (7-3) and (7-6) determines the inertia tensor of the unknown mass . Further, the same measurements performed on the MPMS when it is unloaded will allow one to compute the inertia tensor for the turntable and link.

8. CONCLUSION

Theoretical analysis proves the inertia parameters of an unknown mass can be determined from the joint torques, positions, angular velocities and angular accelerations of the Mass Properties Measurement System (MPMS). In particular, the existing system requires measurement of the torque and angular acceleration on joint one and the angular position, velocity and acceleration on joint two. A proposed system, where the twist angle between the MPMS turning axes equals 90° , permits computing the inertia terms in a simple manner with the additional requirement of measuring the second joint torque for one set of measurements. The new design offers some significant advantages. It packages more compactly, allows simple mechanical scaling, and is easier and less costly to construct.

The next report will discuss actual physical measurements with the existing system to determine the precision with which the MPMS can measure the mass properties of an unknown mass.

REFERENCES

1. Doty, Keith L, 1992, EEL 6667 Class Notes.
2. Doty, Keith L., 1989, "Tabulation of the Midframe Jacobian," *Int. J. of Robotics Research*, 6:3.

1994018763

N94-23236

1993 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA

S6-37
197192
P. 24

MASS PROPERTIES MEASUREMENT SYSTEM:
DYNAMICS AND STATICS MEASUREMENTS

PREPARED BY:	Dr. Keith L. Doty
ACADEMIC RANK:	Professor
UNIVERSITY AND DEPARTMENT:	University of Florida Department of Electrical Engineering
NASA/KSC	
DIVISION:	Mechanical Engineering
BRANCH:	Special Projects
NASA COLLEAGUE:	Bill Jones Todd Graham
DATE:	August 7, 1993
CONTRACT NUMBER:	University of Central Florida NASA-NGT-60002 Supplement: 11

TABLE OF CONTENTS

ABSTRACT.....	3
ACKNOWLEDGMENTS.....	3
1. INTRODUCTION	4
MPMS Dynamics.....	5
MPMS Prototype.....	5
2. COMPUTATION OF THE CENTER-OF-GRAVITY.....	6
MPMS Statics Coefficient Matrix	7
Condition Number of the MPMS Statics Coefficient Matrix	9
Selecting the Twist & Tilt to Optimize Condition Number ..	10
MPMS Configurations with Condition Number of One.....	12
3. STATICS MEASUREMENTS.....	15
Tilt Axis Torsion.....	16
Center-of-Gravity Measurements	17
4. DYNAMICS MEASUREMENTS	18
5. OBSERVATIONS OF MPMS PERFORMANCE.....	22
6. RECOMMENDATIONS.....	23
7. CONCLUSIONS.....	23
REFERENCES.....	24

ABSTRACT

This report presents and interprets experimental data obtained from the Mass Properties Measurement System (MPMS). Statics measurements yield the center-of-gravity of an unknown mass and dynamics measurements yield its inertia matrix. Observations of the MPMS performance has lead us to specific design criteria and an understanding of MPMS limitations.

ACKNOWLEDGMENTS

Willis Crumpler originated the basic concept of the Mass Properties Measurement System (MPMS) and Kedron R. Wolcott has led the theoretical analysis and data acquisition. R.J. McNulty contributed to its design and realization while Ron Remus guided its construction.

The authors gratefully acknowledge the support of the University of Central Florida faculty and staff along with the I-NET and NASA staffs whose diligence and kindness make the NASA Faculty Fellowship program at the Kennedy Space Center so efficient and productive. Particular thanks to Bill Jones and Todd Graham of NASA and James Spencer and Rich Bennett of I-NET.

1. INTRODUCTION

The Mass Properties Measurement System (MPMS), illustrated in Figure 1.1, consists of a serial kinematic mechanism with two intersecting revolute axes, z_0 and z_1 , that intersect with fixed angle α_1 . The joint angles and rates denoted by θ_1, θ_2 and $\dot{\theta}_1, \dot{\theta}_2$, respectively, turn about the respective joint axes. The angle θ_1 rolls the table about z_0 and will be called *tilt*.

The function of the MPMS is to measure the mass, center-of-gravity and second-order mass moments of an unknown mass placed on the turntable. The mass and center-of-gravity can be determined from statics measurements. The inertia tensor must be determined from the dynamics. Reference[1] provides a derivation of the MPMS dynamics equations.

In all the experimental work reported here, we have fixed the angle θ_2 at some value during a measurement. In other words, $\dot{\theta}_2 = 0$ and $\ddot{\theta}_2 = 0$ for all measurements performed in this study.

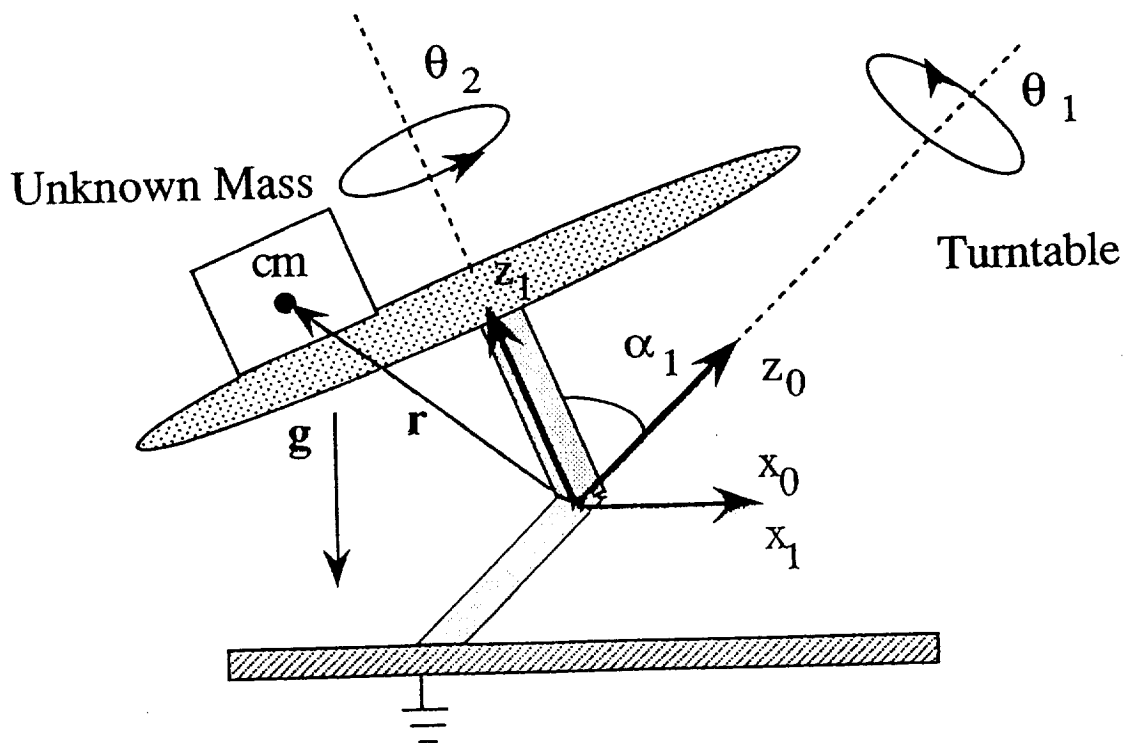


Figure 1.1 The Mass Properties Measurement System

MPMS Dynamics

The total torque τ_{total1} about the z_0 -axis equals

$$\tau_{total1} = \tau_{KE1} + \tau_{PE1} \quad (1-1)$$

According to (4-4a) in [1], with $\sigma_1 = \sin \alpha_1$, $\tau_1 = \cos \alpha_1$, $c_i = \cos \theta_i$, $s_i = \sin \theta_i$, $i = 1, 2$, $r_1 = 0$, $r := r_2$ and $m := m_2$ in that equation, the potential energy related torque equals

$$\tau_{PE1} = m \sigma_1 g_c \{ r_x (c_1 c_2 - \tau_1 s_1 s_2) + r_y (-\tau_1 c_2 s_1 - c_1 s_2) + r_z \sigma_1 s_1 \}. \quad (1-2)$$

τ_{KE1} equals the torque associated with the motion of the system (refer to (4-29) in [1]),

$$\begin{aligned} \tau_{ke1} = & \ddot{\theta}_1 I_{T1} + I_{zz} \tau_1 (\ddot{\theta}_2 + \ddot{\theta}_1 \tau_1) + \sigma_1 \{ I_{yy} c_2 \sigma_1 (\ddot{\theta}_1 c_2 - 2 \dot{\theta}_1 \dot{\theta}_2 s_2) \\ & + I_{xx} \sigma_1 s_2 (2 \dot{\theta}_1 \dot{\theta}_2 c_2 + \ddot{\theta}_1 s_2) \\ & + I_{yz} (-\ddot{\theta}_2 c_2 - 2 \ddot{\theta}_1 c_2 \tau_1 + \dot{\theta}_2^2 s_2 + 2 \dot{\theta}_1 \dot{\theta}_2 \tau_1 s_2) \\ & + I_{xz} (-\dot{\theta}_2^2 c_2 - 2 \dot{\theta}_1 \dot{\theta}_2 c_2 \tau_1 - \ddot{\theta}_2 s_2 - 2 \ddot{\theta}_1 \tau_1 s_2) \\ & - I_{xy} \sigma_1 (2 \dot{\theta}_1 \dot{\theta}_2 c_{2\theta 2} + \ddot{\theta}_1 s_{2\theta 2}) \}. \end{aligned} \quad (1-3)$$

The term I_{T1} describes the inertia of the first link and the other inertia terms, I_{xx} , I_{yy} , I_{zz} , I_{xy} , I_{xz} , I_{yz} describe the inertia of the second link *plus* the unknown mass. Ultimately, to determine the inertia tensor of the unknown mass will require knowledge of the inertia terms of the MPMS links.

The analytical relation (1-2) provides the means for computing the center-of-gravity moment $\tau_g = m g r$ while (1-3) allows one to compute the inertia matrix I_2 formed by the combined mass of link two and the unknown mass.

MPMS Prototype

The first axis of the MPMS prototype aligns with the diagonal of a unit cube (Figure 1.3), hence, the angle the first axis makes to the floor equals

$\text{atan2}(\sqrt{2}, 1) = 35.26^\circ$. The complement of this angle equals the twist $\alpha_1 = \text{atan2}(1, \sqrt{2}) \approx 54.74^\circ$. The body attached frame z-axis initially aligns with the axis of θ_1 . A rotation of the body by $\theta_2 = 120^\circ$ will align the body attached frame x-axis to the first axis of the MPMS. Similarly, $\theta_2 = 240^\circ$ will align the body attached frame y-axis to the first axis of the MPMS. This particular alignment permits determination of the diagonal entries of the inertia matrix by simply measuring the torque and position of the first axis. Computation of the off diagonal terms, however, requires θ_2 to be time varying and measured.

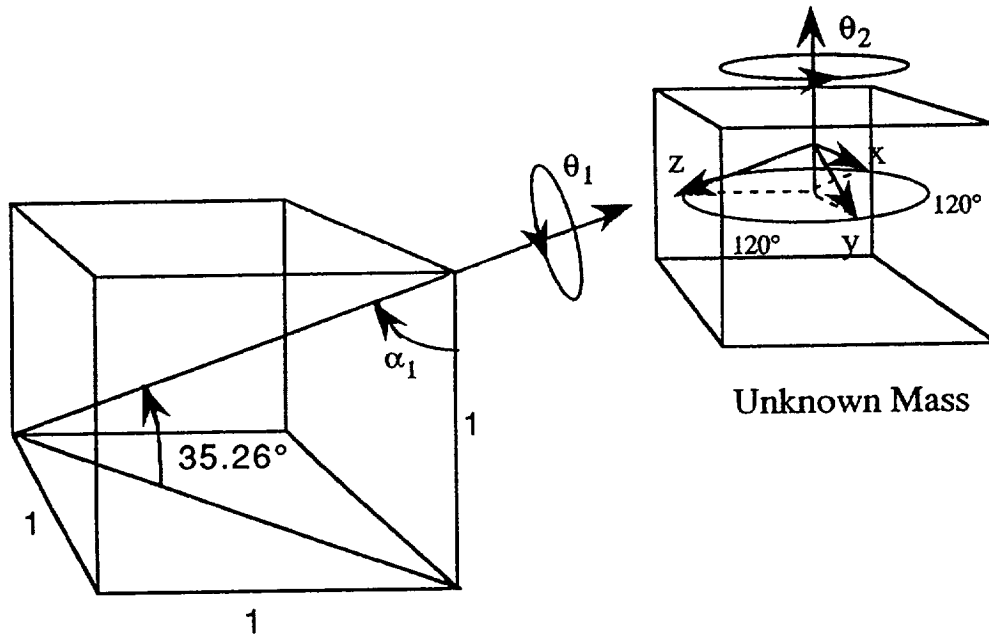


Figure 1.3 MPMS prototype with $\alpha_1 = \text{atan2}(\sqrt{2}, 1) \approx 54.74^\circ$.

2. COMPUTATION OF THE CENTER-OF-GRAVITY

To compute the center-of-gravity vector $\mathbf{r} = \frac{\tau_g}{m g}$, the total mass m and gravitational constant g must be known. In the experiments to follow, we treat $\sigma_1 m g r_x$, $\sigma_1 m g r_y$ and $\sigma_1 m g r_z$ as the unknown terms and fix both axes. Measuring the reaction torque on the first axis for different configurations we can determine the actual values of r_x , r_y and r_z when the solutions are divided by $\sigma_1 m g$. Details of this procedure follow.

MPMS Statics Coefficient Matrix

The coefficients of $\sigma_1 m g r_x$, $\sigma_1 m g r_y$ and $\sigma_1 m g r_z$ in (1-2) equal

$$c_x := c_1 c_2 - \tau_1 s_1 s_2, \quad c_y := -\tau_1 c_2 s_1 - c_1 s_2 \quad \text{and} \quad c_z := r_z \sigma_1 s_1, \quad (2-1)$$

respectively. At $\theta_2 = \mu, \mu + 120^\circ$ and $\mu + 240^\circ$, where μ equals an arbitrary, but fixed, angle, the coefficients form the coefficient matrix

$$C_f = \begin{pmatrix} c_\mu c_1 - \tau_1 s_\mu s_1 & & & \\ -c_1 (c_\mu + \sqrt{3} s_\mu)/2 - \tau_1 (\sqrt{3} c_\mu - s_\mu) s_1/2 & & & \\ c_1 (-c_\mu + \sqrt{3} s_\mu)/2 + \tau_1 (\sqrt{3} c_\mu + s_\mu) s_1/2 & & & \\ & -c_1 s_\mu - c_\mu \tau_1 s_1 & \sigma_1 s_1 & \\ & -c_1 (\sqrt{3} c_\mu - s_\mu)/2 + \tau_1 (c_\mu + \sqrt{3} s_\mu) s_1/2 & \sigma_1 s_1 & \\ & c_1 (\sqrt{3} c_\mu + s_\mu)/2 + \tau_1 (c_\mu - \sqrt{3} s_\mu) s_1/2 & \sigma_1 s_1 & \end{pmatrix} \quad (2-2)$$

A note about these choices for θ_2 . Physically, the angles $\theta_2 = 0, 120^\circ$ and 240° were chosen to align axis one with three orthogonal axis F_b selected in the unknown mass. These choices of θ_2 permit calculating the diagonal entries of the inertia matrix, expressed in F_b coordinates, from torque and position measurements on the first axis alone. Serendipitously, these angles allows us to find an MPMS configuration for which the condition number of C_f equals one. In fact, any three angles separated by 120° will allow the MPMS to be configured such that a condition number of one can be achieved. Recall that a matrix C_f with a condition number of one guarantees the optimum numerical stability when inverting C_f .

By measuring the torque T_1 at joint one for each of the three angles selected for θ_2 , we obtain the linear system of equations in the center-of-gravity unknowns,

$$\begin{pmatrix} T_{11} \\ T_{21} \\ T_{31} \end{pmatrix} = \sigma_1 m g C_f \begin{pmatrix} r_x \\ r_y \\ r_z \end{pmatrix} \quad (2-3)$$

Solving (2-3) produces

$$\begin{pmatrix} r_x \\ r_y \\ r_z \end{pmatrix} = \frac{1}{\sigma_1 m g} C_f^{-1} \begin{pmatrix} T_{11} \\ T_{21} \\ T_{31} \end{pmatrix} \quad (2-4)$$

Two central issues arise: 1) For what configurations is C_f invertible and 2) how well conditioned is C_f for those situations where it is invertible?

The coefficient matrix C_f is invertible as long as

$$\det[C_f] = \frac{-3^{3/2}}{16} (6 + 2 c_{2\alpha_1} - c_{2(\alpha_1 - \theta_1)} + 2 c_{2\theta_1} - c_{2(\alpha_1 + \theta_1)}) \sigma_1 s_1 \neq 0. \quad (2-5a)$$

Notice that the determinant does not depend upon the initial angle μ selected for θ_2 as long as the angles differ by 120° . The determinant of the coefficient matrix can also be written simply as

$$\det[C_f] = \frac{-3^{3/2}}{2} (1 - \sigma_1^2 s_1^2) \sigma_1 s_1, \quad (2-5b)$$

by using the not so obvious, but directly verifiable identity

$$6 + 2 c_{2\alpha_1} - c_{2(\alpha_1 - \theta_1)} + 2 c_{2\theta_1} - c_{2(\alpha_1 + \theta_1)} = 8 (1 - \sigma_1^2 s_1^2). \quad (2-6)$$

The condition $\sigma_1 s_1 \neq 0$ means that the angle θ_1 and the twist α_1 between the two axes cannot be zero or a multiple of π . These design constraints make physical sense. Positioning the turntable at three different angles with $\theta_1 = 0$ will not change the z-component of the center-of-mass, and, therefore, cannot provide information about r_z . The same observation applies to the case where the twist α_1 equals zero.

The term $(1 - \sigma_1^2 s_1^2) = 0$ when $\alpha_1 = \theta_1 = \pm \frac{1}{2} \pi$. The MPMS with this physical configuration can only measure r_z .

Design Criterion

The MPMS configuration and structure must satisfy $\sigma_1 s_1 \neq 1$ and $\sigma_1 s_1 \neq 0$. In particular, the MPMS must have a non-zero twist α_1 between the two axes.

In the course of experimentation, sometimes the computed solutions based on the data varied wildly. Analysis of the causes revealed that one error source was the ill-conditioning of the numerical inversion of C_f .

Condition Number of the MPMS Statics Coefficient Matrix

The condition number c_n for C_f equals

$$c_n = \sqrt{\left| \frac{\lambda_{\max}}{\lambda_{\min}} \right|} \quad (2-7)$$

where λ_{\max} and λ_{\min} are the maximum and minimum eigenvalues, respectively, of $C_f^T C_f$. We now calculate the condition number c_n of C_f as a function of the twist α_1 between the shafts and θ_1 , the angle of the first shaft. This will allow us to determine the angle settings for θ_1 which

Using (2-2), direct computation yields

$$C_f^T C_f = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix},$$

where $\lambda_1, \lambda_2, \lambda_3$ clearly equal eigenvalues of $C_f^T C_f$ and compute to

$$\lambda_1 = \frac{3}{16} (6 + 2 c_2 \alpha_1 - c_2 (\alpha_1 - \theta_1)) + 2 c_2 \theta_1 - c_2 (\alpha_1 + \theta_1), \quad (2-8a)$$

$$\lambda_2 = \frac{3}{16} (6 + 2 c_2 \alpha_1 - c_2 (\alpha_1 - \theta_1)) + 2 c_2 \theta_1 - c_2 (\alpha_1 + \theta_1), \quad (2-8b)$$

$$\lambda_3 = 3 \sigma_1^2 s_1^2, \quad (2-8c)$$

and do not depend upon the arbitrary angle μ added to 0° , 120° and 240° to θ_2 .

Define $\gamma = \sqrt{\left| \frac{\lambda_1}{\lambda_3} \right|}$. Applying (2-6) produces the simple expression

$$\gamma := \sqrt{\frac{1 - \sigma_1^2 s_1^2}{2 \sigma_1^2 s_1^2}} \quad (2-9)$$

Whenever $\lambda_3 \leq \lambda_1$, the condition number c_n of C_f equals γ . On the other hand, if $\lambda_1 < \lambda_3$, the condition number c_n of C_f equals $1/\gamma$,

$$\text{if } \gamma < 1, \text{ then } c_n = 1/\gamma \text{ else } c_n = \gamma. \quad (2-10)$$

Selecting the Twist and Tilt to Optimize Condition Number

A plot of c_n versus θ_1 appears in Figure 2.1 for $\alpha_1 = \text{atan2}(1, \sqrt{2}) \approx 54.74^\circ$, the angle employed in the prototype (Figure 1.3). Details of the W-shaped portion of the curve is illustrated in Figure 2.2.

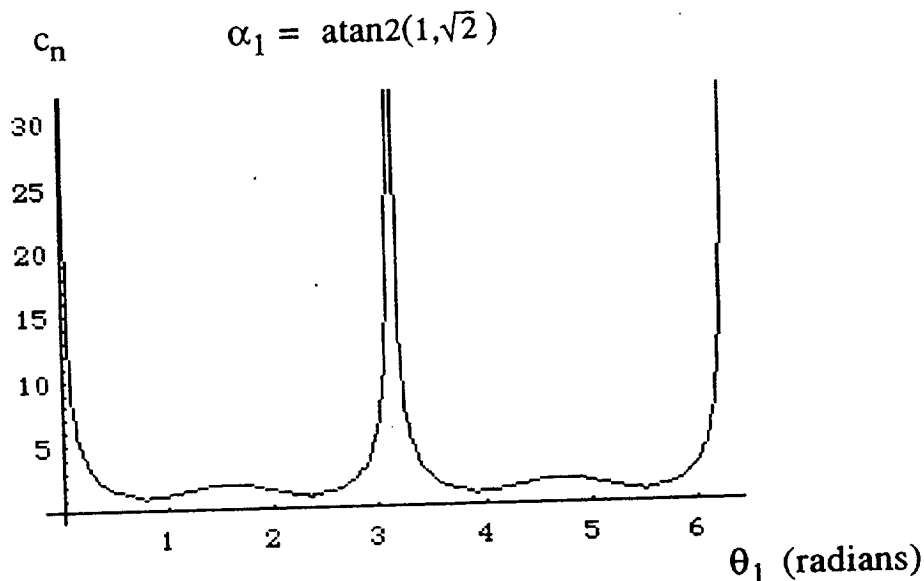


Figure 2.1 The condition number c_n of the coefficient matrix C_f vs. the angle θ_1 of the first shaft for $\alpha_1 = \text{atan2}(\sqrt{2}, 1)$.

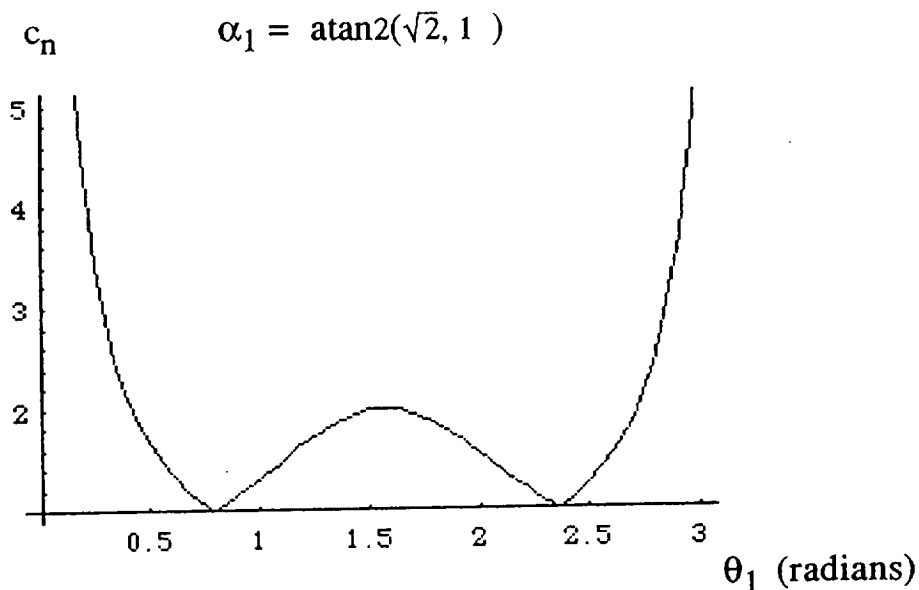


Figure 2.2 Details of the condition number $\alpha_1 = \text{atan2}(\sqrt{2}, 1)$.

The condition number $c_n = 1$, its optimum value, at that value of θ_1 for which $c_n - 1 = 0$. This occurs at $\theta_{1\text{opt}} = 0.7853981633974484 \cdot 180/\pi = 45^\circ$.

A plot of c_n versus θ_1 appears in Figure 2.3 for $\alpha_1 = \pi/2$. Details of one period of the curve is illustrated in Figure 2.4.

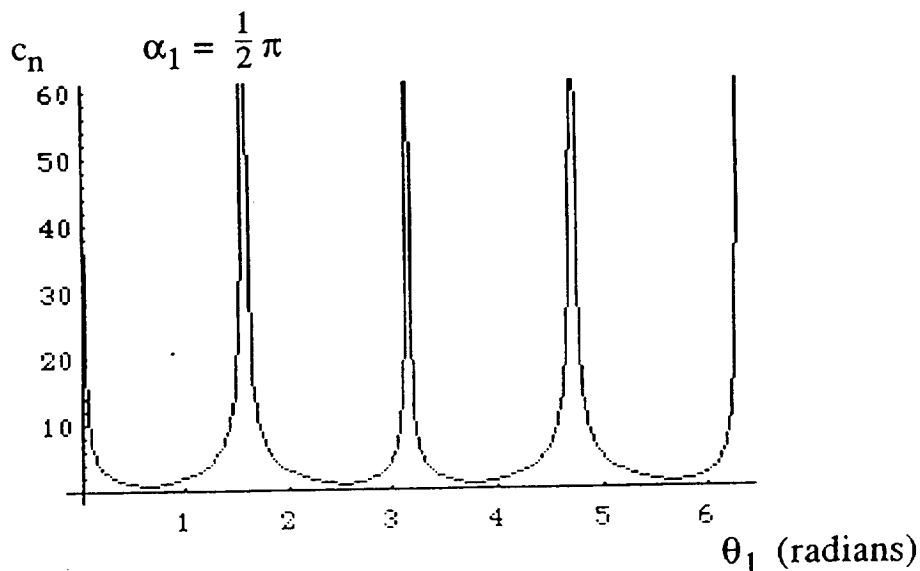


Figure 2.3 The condition number c_n of the coefficient matrix C_f vs the angle θ_1 of the first shaft for $\alpha_1 = \frac{1}{2} \pi$.

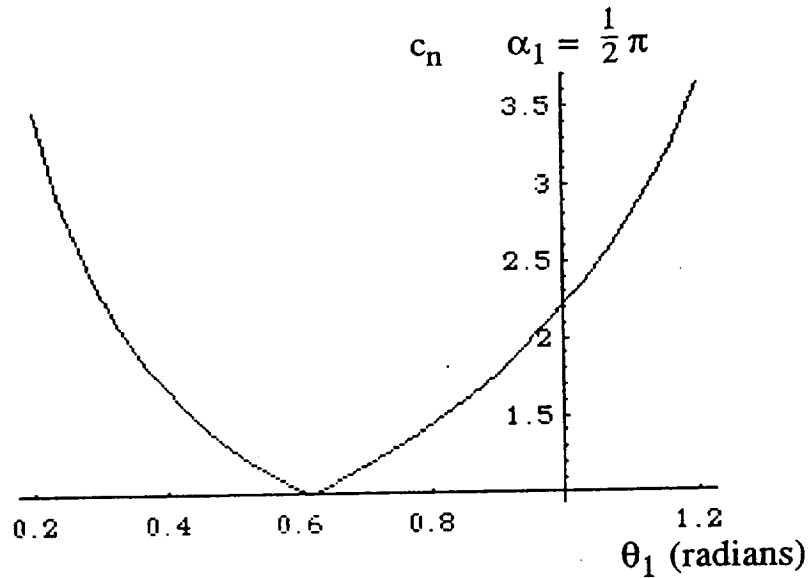


Figure 2.4 Details of the condition number for $\alpha_1 = \frac{1}{2} \pi$.

The condition number $c_n = 1$ at that value of θ_1 for which $c_n - 1 = 0$. This occurs at

$$\theta_{1opt} = 0.6154797086703873 \cdot 180/\pi = 35.2644 = \alpha_1 = \text{atan2}(\sqrt{2}, 1).$$

MPMS Configurations with Condition Number of One

A general plot of the condition number (Figure 2.5) illustrates the nature of the condition number when the three angles selected for θ_2 split the circle into three 120° arcs. The plot has been truncated at $c_n = 5$ to display the important characteristics of the function.

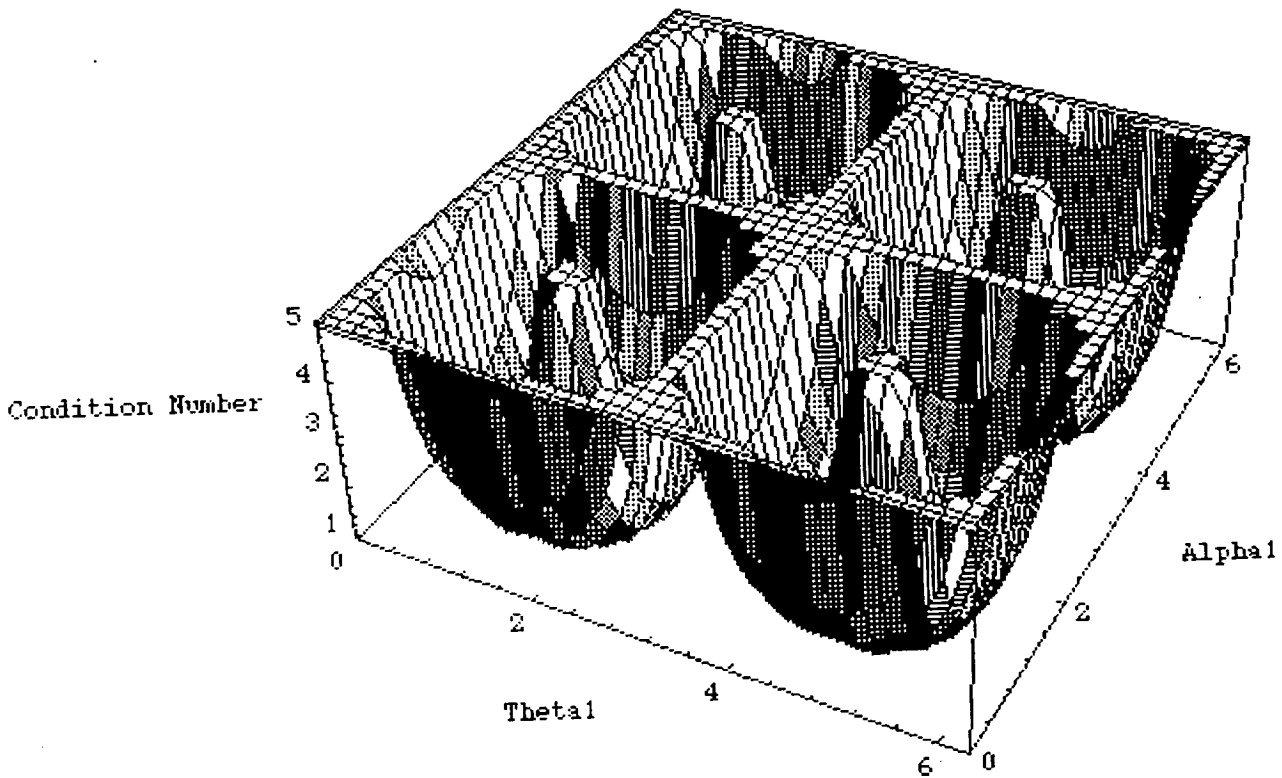


Figure 2.5 Plot of the condition number, clipped at $c_n=5$, as a function of θ_1 and the angle α_1 between the first and second axis.

The condition number plotted in Figure 2.5 applies for any three angles chosen for θ_2 which are separated by 120° . The function c_n is symmetric in θ_1 and α_1 and has period of π along each axis. Figure 2.6 illustrates a single quadrant of Figure 2.5 in more detail and Figure 2.7 depicts an overhead view.

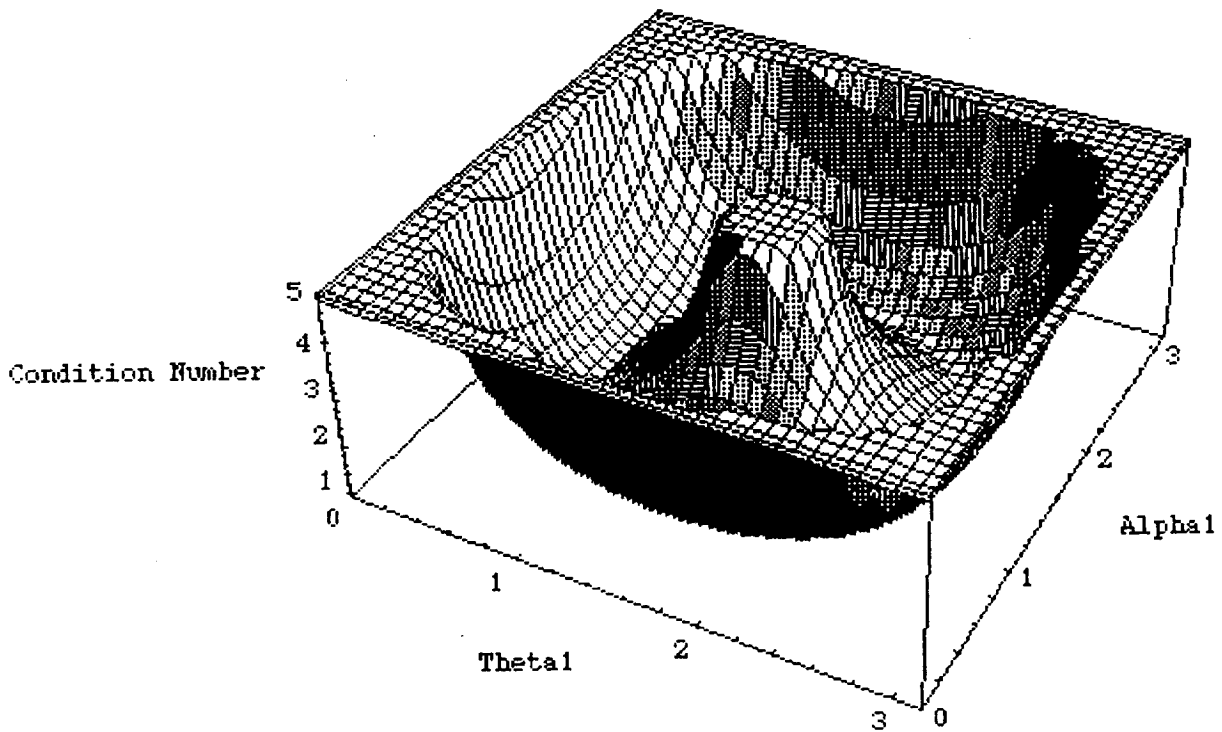


Figure 2.6 Plot of one quadrant of the condition number clipped at $c_n=5$.

When the condition number is one, all the eigenvalues of $C_f^T C_f$ must equal one. In particular, for $c_n = 1$, (2-9) implies

$$\lambda_3 = 3 \sigma_1^2 s_1^2 = 3 \sin^2(\alpha_1) \sin^2(\theta_1) = 1. \quad (2-11)$$

Furthermore, if $\lambda_3 = 1$, (2-9) and (2-8) easily proves that $\lambda_1 = \lambda_2 = 1$ and, hence, $c_n = 1$. Consequently, (2-11) provides a design equation for the MPMS. For example, the minimum tilt θ_1 possible for $c_n = 1$ occurs when $\sigma_1^2 = 1$, i.e., $\alpha_1 = \pm 90^\circ$. The minimum angle θ_1 must, therefore satisfy $s_1^2 = \frac{1}{3}$, yielding $\theta_1 = \pm 35.26^\circ$.

The crease at the bottom of the well in Figure 2.7 depicts the curve (2-11) on which the condition number of the coefficient matrix equals one.

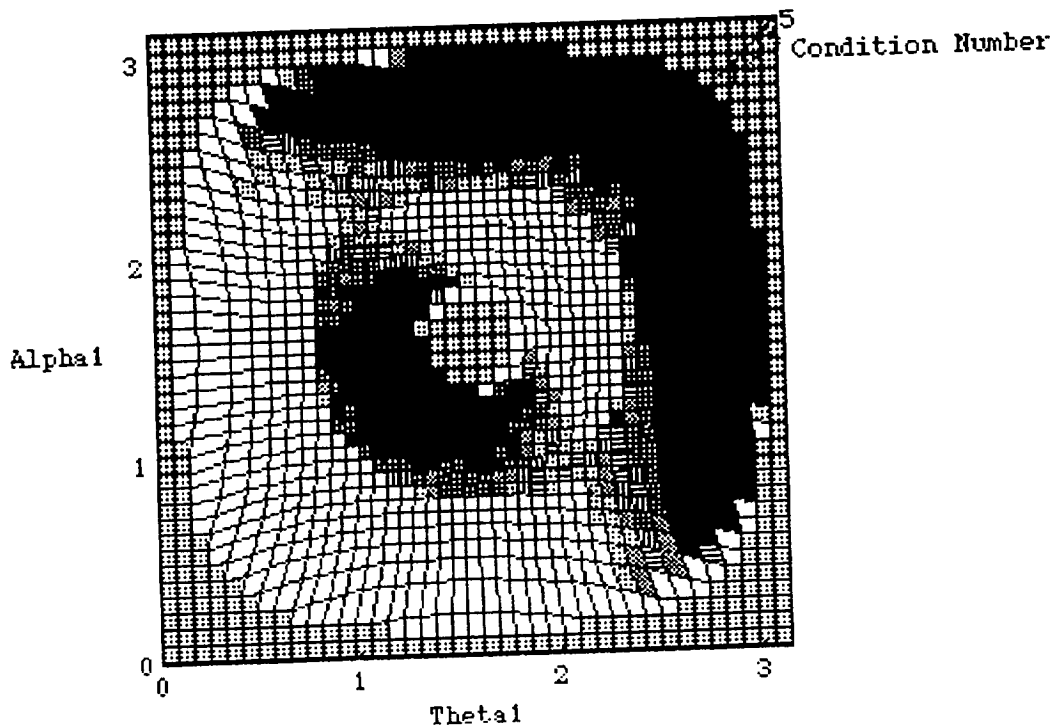


Figure 2.6 Top view of one quadrant of the condition number clipped at $c_n=5$.

3. STATICS MEASUREMENTS

Another source of error discovered in MPMS center-of-gravity measurements was the absolute accuracy of the torque sensor attached to z_0 . The torque sensors was rated 1% of full-scale rating of approximately 1900 in-lb. Measured torques of 40 in-lb could be off as much as 50%. Furthermore, if the condition number of the coefficient matrix were as little as 4, for example, with a twist $\alpha_1 = 54.73^\circ$ and a tilt of $\theta_1 = 25.24^\circ$, the error in the values for the center-of-gravity components could be as large as 100%. In fact, it was just such large errors that prompted us to look at the conditioning of the coefficient matrix and seek twist, tilt and turntable angles that would yield a condition number of one.

Statics Measurement Criterion

The three turntable angles (θ_2) should trisect the circle and the tilt angle θ_1 should be selected to satisfy (2-11) for a given α_1 .

Tilt Axis Torsion

The dynamics measurements later will require oscillating the tilt axis at some prescribed frequency. Any significant torsion spring action of the tilt axis would affect the dynamics measurement. Essentially, we wanted to determine the MPMS natural frequency to ascertain whether the tilt axis torsion would have to be modeled in the dynamics. The prototype MPMS, upon which all experiments were performed, possesses the twist angle $\alpha_1 = 54.73^\circ$.

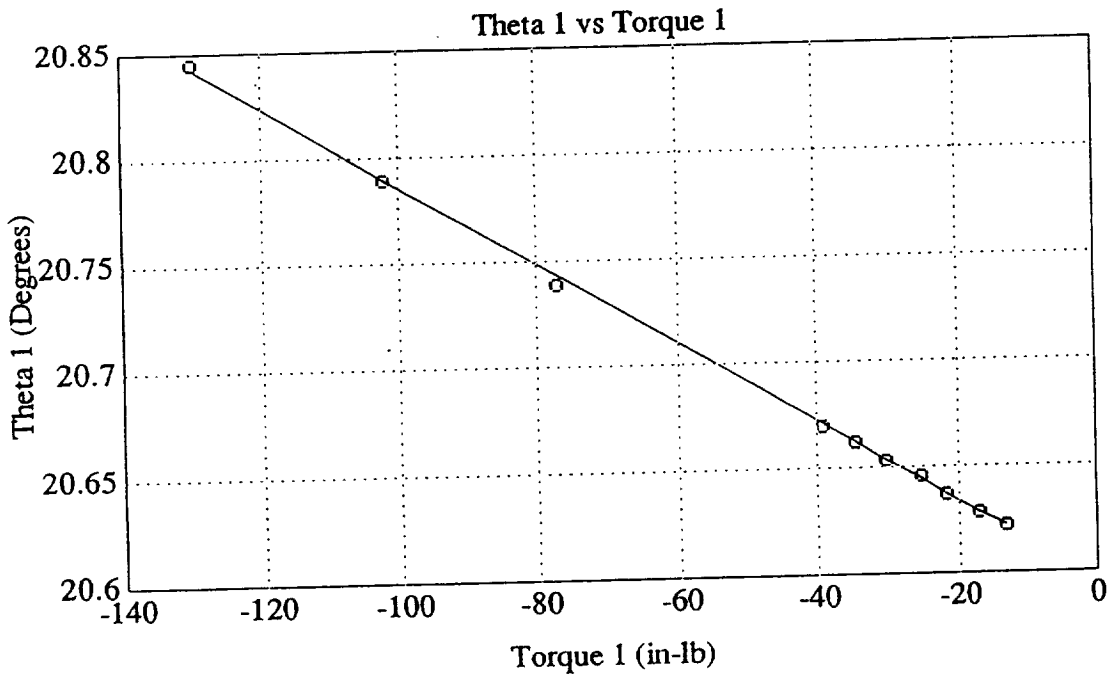


Figure 3.1 Compliance of the tilt axis: $\delta\theta_1 = (-0.0019^\circ/\text{in-lb}) \tau_1 + 20.5965^\circ$

Figure 3.1 depicts the relationship between the angular deflection $\delta\theta_1$ of the tilt axis as a function of the torque τ_1 exerted about that axis. The experimental compliance $c = 1/k$ of the tilt axis equals $c = 0.0019^\circ/\text{in-lb}$. Assuming frictionless operation, the natural frequency of the MPMS with the second joint fixed approximately equals 16 Hz for a 50 lb force, 4 in from the first axis,

$$f_n = \sqrt{\frac{k}{I_m}} \approx \sqrt{\frac{32.2*12}{0.0019*50*4^2}} \approx 16 \text{ Hz.} \quad (3-1)$$

In order to ignore the tilt axis torsion spring effect in the dynamics measurements, we should oscillate the table well below the resonant frequency of the second link and payload. A payload 9 times larger than the 50 lb force, however, would produce a resonant frequency of about 5.3 Hz. Such a low resonance would probably require modelling the shaft torsion for the dynamics measurements.

Center-of-Gravity Measurements

The minimum-norm, least-squares values for the first mass-moments of the turntable, plus an unknown mass placed on the turntable, yielded a first-mass moment vector of $m \ g \ r \approx [160 \quad -41 \quad 284]$ (in-lb). The vector $m \ g \ r$ was calculated from the average values of 3000 torque samples per experiment, taken from 30 experiments. The 30 experiments consisted of permutations of 10 measurements taken at each of the three MPMS configurations $\{ (\theta_1, \theta_2) \} = \{ (35^\circ, 0), (-35^\circ, 120^\circ), (-35^\circ, 240^\circ) \}$.

Since the torque sensor used in the prototype MPMS generated absolute errors of 19 in-lbs, small, first-order mass moments measurement were exceptionally error prone. Figure 3.2 illustrates this point quite clearly. The three curves indicate the error of each experiment about the solution $m \ g \ r \approx [160 \quad -41 \quad 284]$ (in-lb). The "blue" plot (the one that starts in the upper left-hand corner of the diagram) indicates the most error in measurement. Although 19 in-lb almost equal 50% of the measured y-axis first mass-moment of -41 in-lb, the huge number of samples over which averages were computed produced consistent results within $\pm 6\%$. The larger values of the x-axis and z-axis first mass-moments produced more accurate readings; within $\pm 1\%$ for $m \ g \ r_x$ (the "red" plot has a negative relative error for the first experiment) and within $\pm 0.5\%$ for $m \ g \ r_z$ ("green" plot). These results are consistent with the fact that the absolute error of 19 in-lb has much less effect on the larger torque measurements.

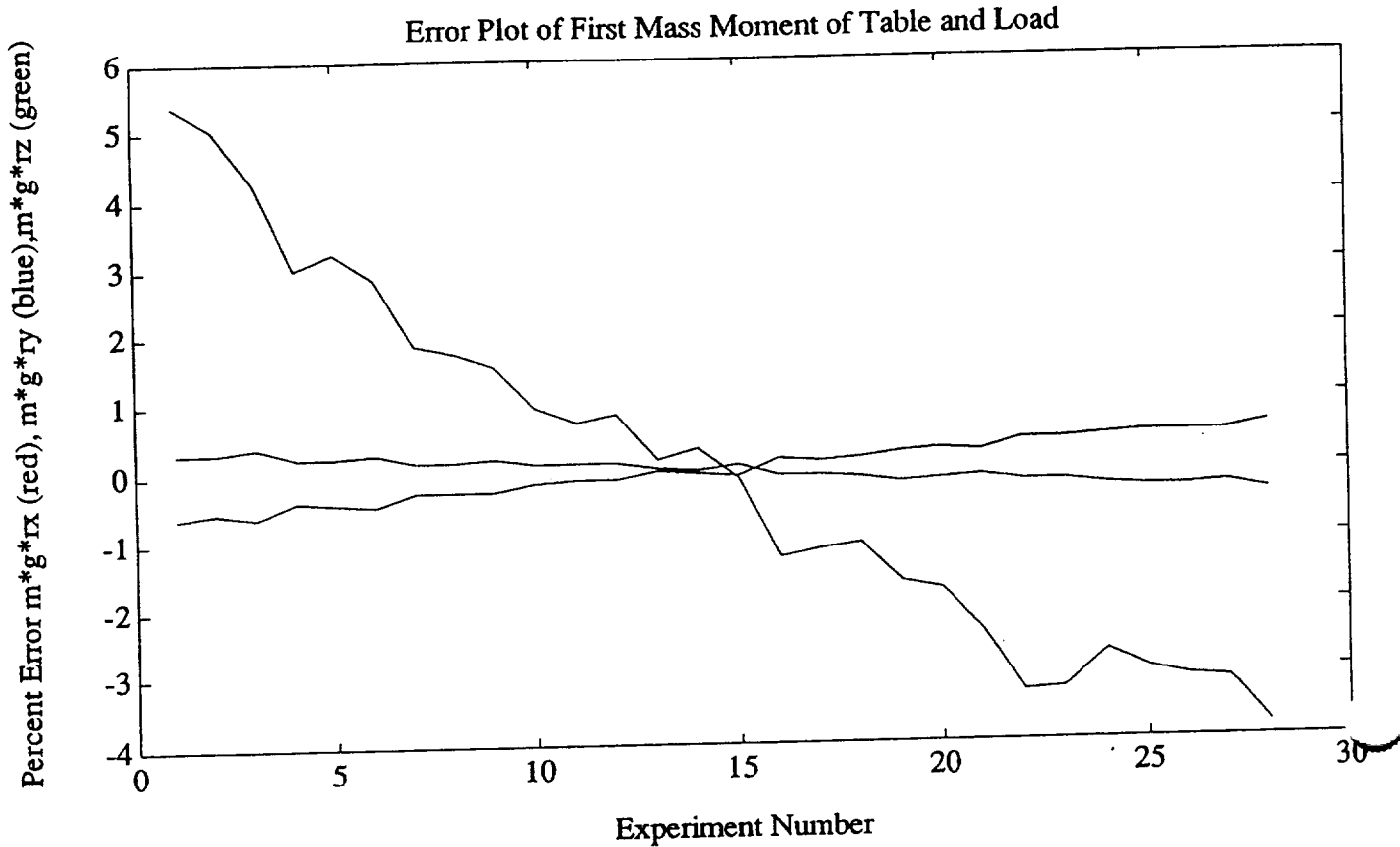


Figure 3.2 Percent error about mean of first order mass moment. At the far left the top plot is "blue", the middle plot "green" and the lower plot "red".

4. DYNAMICS MEASUREMENTS

The dynamics equation (1-3) for $\alpha_1 = 54.73^\circ$ ($\sigma_1 = \sqrt{\frac{2}{3}}$, $\tau_1 = \frac{1}{\sqrt{3}}$), $\ddot{\theta}_2 = 0$, $\dot{\theta}_2 = 0$, produces

$$\tau_{ke1} = \{I_{T1} + I_m\} \ddot{\theta}_1, \tag{4-1}$$

where

$$I_m := \frac{1}{3} (2 I_{xx} s_2^2 + 2 I_{yy} c_2^2 + I_{zz} - 2 I_{xy} s_{2\theta_2} - 2^{3/2} I_{xz} s_2 - 2^{3/2} I_{yz} c_2) \ddot{\theta}_1. \tag{4-2}$$

To test the MPMS prototype, we measured the angular position θ_1 and total torque $\tau_{total1} = \tau_{KE1} + \tau_{PE1}$ at joint one versus time in order to calculate the total inertia seen by the first joint,

$$\{I_{T1} + I_m\} = \frac{\tau_{total1} - \tau_{PE1}}{\ddot{\theta}_1} \quad (4-3)$$

The gravity term τ_{PE1} exceeds τ_{KE1} by a factor from 10 to 60, hence, the torque difference $\tau_{total1} - \tau_{PE1}$ will be quite small. Subtracting two nearly equal numbers produces large numerical errors. Initial experiments proved this to be a significant problem. To eliminate it we balanced the unknown mass so as to make the gravity torque $\tau_{PE1} \approx 0$. We then manually oscillated the second link and the unknown mass point at about 8.7 Hz about the $\tau_{PE1} = 0$ and measured the resultant torque and angular position.

We compute the angular acceleration $\ddot{\theta}_1$ from θ_1 by the central difference formula [1],

$$\ddot{\theta}_1(t_0) = \frac{-\theta_1(t_2) + 16\theta_1(t_1) - 30\theta_1(t_0) + 16\theta_1(t_{-1}) - \theta_1(t_{-2})}{12h^2} + O(h^4) \quad (4-4)$$

where $h = t_i - t_{i-1} = 4 \text{ ms}$ and $O(h^4)$ indicates errors on the order of h^4 .

Figures 4.1 and 4.2 depict typical position and torque signals, respectively, for a given experiment. The 12 in-lb absolute accuracy of the torque sensor used in the prototype resulted in large instantaneous errors when measuring the small inertia

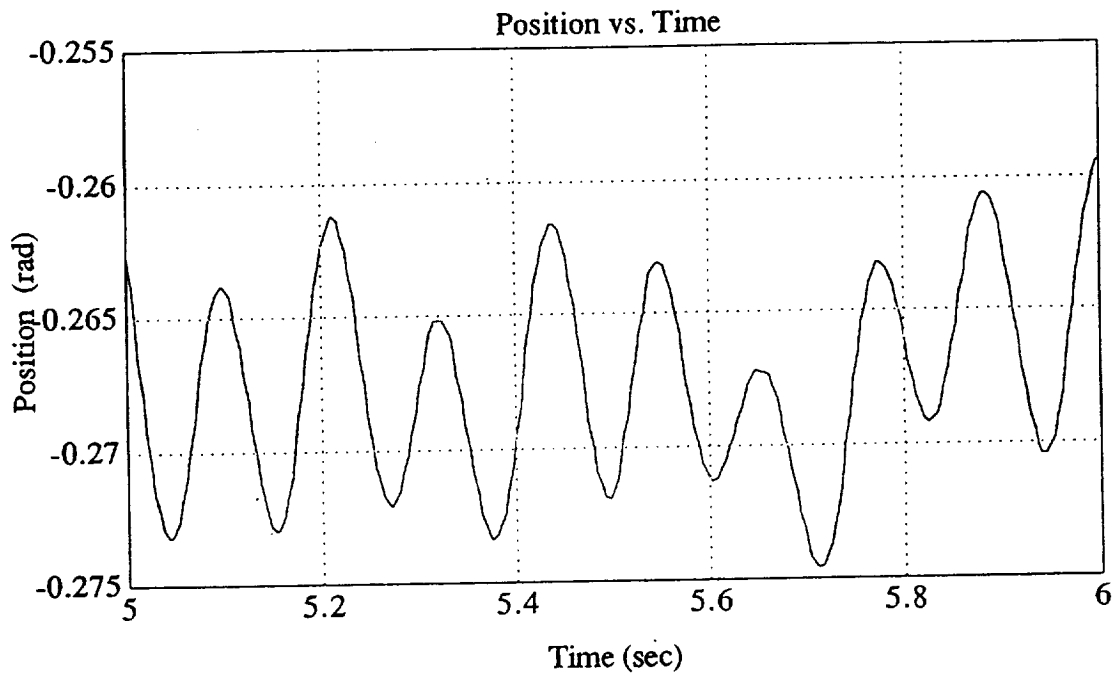


Figure 4.1 A plot of θ_1 vs time.

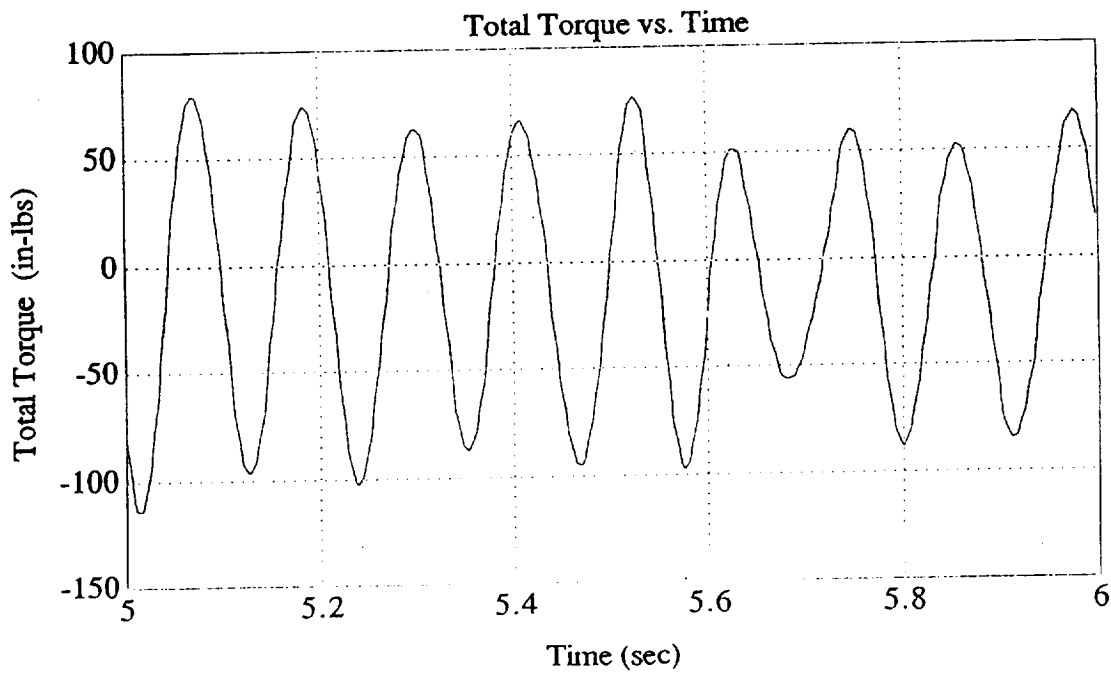


Figure 4.2 A plot of τ_{total1} vs time.

torques. The mean of the absolute value of the torque $\overline{|\tau_{kel}|}$ and position derived angular acceleration $\overline{|\ddot{\theta}_1|}$ were computed for each experiment consisting of 5000 points. The ratio of these values were then taken as the estimate for the measured inertia $\{I_{T1} + I_m\}$,

$$\{I_{T1} + I_m\} \approx \frac{\overline{|\tau_{kel}|}}{\overline{|\ddot{\theta}_1|}} \quad (4-5)$$

Figure 4.3 indicates the inertia measurement percent error for 10 different experiments with respect to their mean value. The error is bounded $\pm 10\%$ about a mean of 4.34 in-lb-s^2 . These large error bounds can plausibly be attributed to the torque sensor sensitivity. The measured torque values in Figure 4.2 were less than 10% full scale, hence, the torque sensor, which is rated 1% accuracy at full scale, produces approximately 10% error in the measurements.

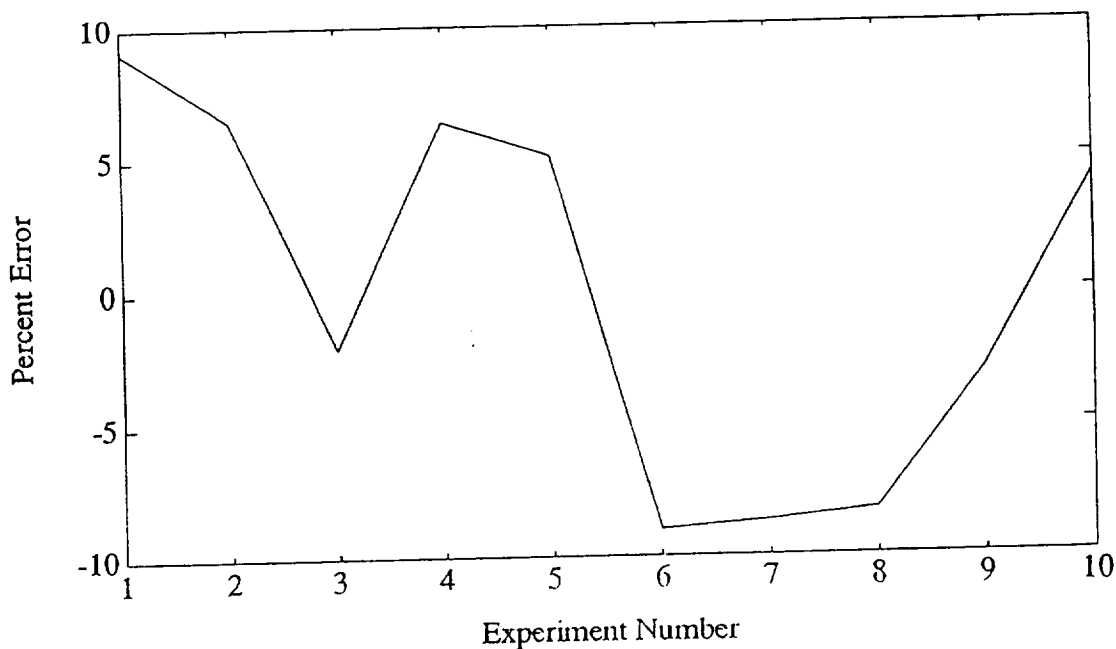


Figure 4.3 Percent Error in Inertia about Mean of 4.34 in-lb-s^2

5. OBSERVATIONS OF MPMS PERFORMANCE

With single torque sensor and position sensor, one can measure the center-of-gravity and the moments of inertia about three orthogonal axes. However, one must minimally move and measure the velocity and acceleration of the second axes in order to obtain the full inertia tensor. These facts were established in [1].

To obtain accurate center-of-gravity measurements one should select MPMS configurations where the condition number c_n of the coefficient matrix equals one. In other words, the choice of angles selected for measuring the center-of-gravity matter significantly. A poor choice of angles can multiply the measured error by a factor $1 \leq c_n < \infty$. This explained why, in some of our earlier experiments the errors were 600%.

Torque readings for both statics and dynamics measurements at the lower limit of the torque sensor resolution also had significant impact on the measurement reliability. The location of the center-of-gravity of the unknown mass affects the size of the torque readings. This suggests placing the mass center away from the z_1 -axis in order to drive the force sensor harder and produce more accurate results. Taking these details into consideration center-of-gravity measurements we made highly repeatable measurements ($\pm 0.4\%$) for a specific configuration and setup.

Gravity torque dominates dynamics measurements and extreme measures must be taken to minimize its effects. Dynamics measurements were made at a balance point of the turntable using "high" frequency oscillations (≈ 8 Hz). The combination of the smaller gravity terms about the balance point and the increased inertia effects for high speed oscillations enabled us to more accurately measure the much smaller inertial terms than would otherwise be possible. Higher frequency oscillations are precluded because of the MPMS resonance near 16 Hz. Dynamics measurements were repeatable ($\pm 3.0\%$) for a given configuration and setup and stayed within $\pm 10.0\%$, even though the measurements were near the lower resolution limit of the torque sensor.

Stiction and friction also appear to play an important role in inertia measurements on the MPMS. Stiction was particularly noticeable when balancing the turntable. These effects were not modeled in this effort, but should be taken into account in a later prototype.

6. RECOMMENDATIONS

In [1] it was recommended to make the twist $\alpha_1 = 90^\circ$. Further support for this design change arises in the condition number c_n analysis provided earlier. That analysis proved that the minimum tilt angle of 35.26° for $c_n = 1$ can only be achieved when $\alpha_1 = 90^\circ$. By comparison $\alpha_1 = 5.74^\circ$ in the current prototype and the tilt angle must be 45° in order for $c_n = 1$.

At this point the limitations of the current prototype appear to be well understood and further experimentation is contraindicated. If management decides to continue with the MPMS concept, the follow-on prototype will require some modifications to correct the discovered problems with the first model.

The prototype MPMS permitted only manual actuation of joint one. A second generation MPMS should :

1. Include automatic operation of both axes,
2. Match torque sensor with expected loads,
3. Utilize low friction bearings,
4. Add a position sensor on the second axis,
5. Incorporate a torque sensor on the second joint,
6. Explore concept with potential customers

7. CONCLUSIONS

Theoretical results verify the MPMS concept in principle, but experimental results indicate several engineering limitations in the current implementation. In particular, the torque sensor must be able to accurately measure a broad range of values. The gravity terms dominate the dynamics and this fact makes it difficult to obtain good inertial measurements.

To obtain the optimum performance from the MPMS when measuring center-of-gravity:

1. Place the unknown mass's center-of-gravity away from the turntable center.

2. Measure the center-of-gravity every 120° turn of the turntable.
3. Tilt the turntable to the optimum angle to produce a condition number of one for the coefficient matrix.

When measuring inertia :

4. Place the unknown mass so that the turntable axis passes through its center-of-gravity.
5. Oscillate the turntable well below the resonant frequency of the MPMS about the null gravity torque configuration .

Items 1 and 4 clearly compete with one another. Placement issues have not been addressed here and would need further exploration should the MPMS project proceed.

REFERENCES

1. Doty, Keith L,1993, "Mass Properties Measurement System Dynamics", Kennedy Space Center, NASA Faculty Fellow internal report.
2. Gerald, C.F. and Wheatley, P.O., "Applied Numerical Analysis", Addison Wesley,1984, pp.243.

1994018764
N94-22237

1993 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA

57-45
197193
p. 29

A PROTOTYPE EXPERT/INFORMATION SYSTEM FOR EXAMINING
ENVIRONMENTAL RISKS OF KSC ACTIVITIES

PREPARED BY:	Dr. Bernard A. Engel
ACADEMIC RANK:	Associate Professor
UNIVERSITY AND DEPARTMENT:	Purdue University Department of Agricultural Engineering
NASA/KSC	
DIVISION:	Biomedical Operations and Research
BRANCH:	Biological Research and Life Support
NASA COLLEAGUE:	John C. Sager
DATE:	August 6, 1993
CONTRACT NUMBER:	University of Central Florida NASA-NGT-60002 Supplement: 11

Abstract

Protection of the environment and natural resources at the Kennedy Space Center (KSC) is of great concern. An expert/information system to replace the paper-based KSC Environmental Checklist was developed. The computer-based system requests information only as required and supplies assistance as needed. The most comprehensive portion of the system provides information about endangered species habitat at KSC. This module uses geographic information system (GIS) data and tools, expert rules, color graphics, computer-based video, and hypertext to provide information.

Summary

A decision support system project has been initiated to develop software to assist with environmental and natural resource management at KSC. A module for this decision support system was developed based on the paper-based KSC Environmental Checklist. This checklist must be completed before most activities of significance can be conducted/completed at KSC. The paper-based checklist requires information and expertise that those completing the checklist often don't have.

A prototype expert/information system based on this checklist was developed. This system was developed using KnowledgePro Windows (an expert system and multimedia development tool) and PC ARC/View (a geographic information system (GIS) data display and analysis tool). The resulting system requires a PC with Microsoft Windows and PC ARC/View. The checklist was implemented such that users complete a computer-based checklist in a manner similar to completing the paper-based check list. Based on items checked, the system requests additional information only as required. Users of the system can request additional information to complete checklist items through hypertext links. The hypertext links provide a variety of information. For example, one of the checklist items asks whether endangered species habitat will be affected. Selection of this link accesses an extensive expert/information system that provides information about KSC wildlife habitat, KSC habitat locations, wildlife species, and wildlife habitat requirements. Much of the textual information in this section is based on a biodiversity publication being prepared by NASA contractor personnel. Hypertext links are used to provide access to this information. This system also makes extensive use of video clips, graphics, and GIS data.

The prototype system is in a usable form, but many additional features and capabilities should be added. With only slight modifications, the habitat and wildlife module could be used as an educational and training tool at KSC and many other locations including schools.

TABLE OF CONTENTS

I.	INTRODUCTION
II.	RELATED LITERATURE
III.	PROCEDURES
3.1	Expert/Information System Development
3.2	GIS Data Development and Documentation
3.3	GIS Data at KSC
3.4	Expert System Development Tool
3.5	Habitat Module
3.6	Video Capture
3.7	Graphics Capture
IV.	RESULTS AND DISCUSSION
4.1	Example Screens from ARC/View
4.2	Example Screens from the Expert/Information System
V.	SUMMARY AND CONCLUSIONS
VI.	RECOMMENDATIONS FOR CONTINUED WORK
APPENDIX A	CONVERSION OF ERDAS GIS DATA FOR USE IN ARC/VIEW
APPENDIX B	CLIPS FILE FOR ADDING TEXT LABELS TO INFO TABLE FOR WETLANDS LAYER
APPENDIX C	CONVERSION OF UNIX ARC/VIEW COMPATIBLE FILES FOR USE WITH THE MS WINDOWS ARC/VIEW
APPENDIX D	STRATEGY FOR DEVELOPMENT OF ENVIRONMENTAL AND NATURAL RESOURCES DECISION SUPPORT SYSTEMS AT KSC
VII.	REFERENCES

LIST OF ILLUSTRATIONS

Figure	Title
4.1	Portion of an ARC/View Data <i>View</i>
4.2	Data Table from an ARC/View Data <i>View</i>
4.3	Portion of the Expert System Environmental Checklist
4.4	List of KSC Habitat Types
4.5	Menu to Access Scrub Habitat Information
4.6	Species that Require and Use Scrub Habitat
4.7	Menu to Access Scrub Jay Information
4.8	Textual Scrub Jay Information

I. INTRODUCTION

The Kennedy Space Center (KSC) is not only home to many NASA activities but it is also home to 22 wildlife species listed as Endangered or Threatened on either the Federal or State lists. When NASA purchased land for the development of KSC in the early 1960's, not all of the land was needed for the space program. As a result, the U.S. Fish and Wildlife Service in cooperation with NASA established the Merritt Island National Wildlife Refuge in 1963. Thus, the effects of NASA activities on the environment and natural resources are of great concern.

A variety of monitoring and research activities are conducted by scientists at KSC to protect the KSC environment and natural resources and to better understand the relationships between these systems and NASA operations. These activities include: water quantity and quality monitoring and modeling (Dwor-nik, 1984; Heaney et al., 1984; Bennett, 1989; Dierberg and Jones, 1989), soil resources inventory development (Schmalzer and Hinkle, 1990a and 1991), climate monitoring (Madsen et al., 1989; Dreschel et al., 1990; Mailander, 1990), wildlife monitoring (Breininger and Schmalzer, 1990; Dreschel et al., 1991), and vegetation/habitat monitoring (Breininger, 1990; Schmalzer and Hinkle, 1990b; Provan-cha and Hall, 1991).

To assist in protection of the environment and natural resources at KSC, an environmental checklist must be completed prior to construction and many activities at KSC. The KSC Environmental Checklist was recently developed for KSC along with a document to provide assistance in completion of the check-list. Several of the checklist items that must be addressed require access to databases and information that reside in various formats and at various locations at KSC. In many cases, this information is not well documented and people requiring access don't know it exists. Other questions on the checklist can be difficult to answer. Examples of such checklists can be found in KHB 8800.6.

The objective of this project was to develop a prototype expert/information system to assist with the completion of the KSC environmental checklists. It is envisioned that this system would become part of a larger environmental/natural resources decision support system that is being developed for KSC.

II. RELATED LITERATURE

Environmental concerns and issues have and will continue to greatly shape the political agenda at local, state, and national levels. Environmental considerations have become important components of project planning as a result of the National Environmental Policy Act (NEPA) (Liroff, 1976) (Bear, 1989) and public pressures. Environmental considerations are often accounted for through environmental impact assessments.

Environmental Impact Assessment (EIA) requires the qualitative and quantitative prediction, analysis, and assessment of the impacts of human activities on the environment. A variety of methods and tools are used to perform EIAs. EIA techniques range from simple checklists and qualitative impact matrices to complex computer-based approaches that use simulation, optimization, GIS, and expert system techniques. The formats of these techniques range from narrative and qualitative descriptions to various attempts at quantification and formalization, from monetization to graphical methods. The procedures may involve experts, expert teams, panels, workshops, public hearings, and court proceedings (US EPA, 1992). The legal, procedural and institutional components of EIAs are important aspects but can vary greatly from project to project.

Some of the EIA techniques that have been successfully used include:

- Graphic overlay methods (McHarg, 1968; Dooley and Newkirk, 1976)
- USGS Matrix (Leopold et al., 1971)
- Network Analysis (Sorensen, 1971; Sorensen, 1972)
- Cross-impact Simulation (Kane, 1972)
- EES Environmental Evaluation System (Dee et al., 1973)
- HEP Habitat Evaluation Procedures (US Fish and Wildlife Service, 1976)
- Decision Analysis (Keeney and Raiffa, 1976)
- WRAM Water Resources Assessment (Solomon et al., 1977; Richardson et al., 1978)
- EQA Environmental Quality Assessment (Duke et al., 1977)
- METLAND Landscape Planning Model (Fabos et al., 1978)
- Goals Achievement Matrix (Hill, 1968)
- WES Wetland Evaluation System (Galloway, 1978)
- AEAM Adaptive Environmental Assessment (Holling, 1978)
- EQEP Environmental Quality Evaluation Procedure (Duke, 1979)
- CBA Cost-Benefit Analysis and related methods
- Mathematical modeling (Bonazountas et al., 1988; Forgensen, 1991)
- Interactive Systems Analysis and Decision Support (Fedra et al., 1987; Lein, 1989; Fedra 1991; Fedra et al., 1991)

Additional details concerning EIA techniques can be found in Clark et al. (1984), Beanlands and Duinker (1983), Tomlinson and Atkinson (1987), Bregman and Mackenthun (1992), Bisset (1988), and numerous

others.

Fedra (1992) describes several intelligent environmental information systems applications the International Institute for Applied Systems Analysis has developed. XSPILL is a dynamic analytical model to simulate the propagation of an accidental spill of a chemical in a river. XSPILL integrates GIS, a river simulation, a hypertext information system, and graphical visualization techniques. Another system uses several simulations, finite element techniques, GIS, visualization techniques, and expert systems to simulate the movement of groundwater contaminants under various scenarios. Fedra (1993) describes the expert systems that are used in the overall system in more detail. The system is used for the management of hazardous waste, site selection and risk assessment of landfills, and the design and evaluation of groundwater remediation measures.

Fedra (1991) describes a series of computer-based decision support systems to assist with environmental impact assessment. The decision support modules use expert systems, environmental assessment checklists in rule-based formats, databases, GIS, simulations, and visualization techniques. Extensive modules assess air quality, surface water quality, and groundwater quality.

Fedra et al. (1991) developed MEXSES, a rule-based expert system for environmental impact assessment at a screening level, starting at an early stage of project planning and design. MEXSES uses environmental assessment checklists in expert system form, databases, GIS, a hypertext explanation facility, and visualization techniques. The system was implemented on a UNIX workstation using X Windows. MEXSES was successfully applied to the Lower Mekong Basin of Thailand.

Negahban et al. (1993) developed a GIS-based decision support system to assist with regional environmental planning in the Lake Okeechobee, Florida basin. The primary contaminant considered by the system is phosphorus from agricultural areas of the basin. Kowalski and Bagley (1993) describe the US Army computer-based natural resources management system. Carlson et al. (1993) describe the Soil Conservation Service (SCS) computer-based systems that assist with erosion control and water quality protection. Garcia et al. (1993) developed a decision support system that uses GIS and habitat models to quantify the quality of wildlife habitat. Shih and Mills (1993) developed a spatial and ecological forest management system for upland central hardwood species of the US. Gough and Edwards-Jones (1993) developed a knowledge-based system to provide EIA training. Gurganus et al. (1993) developed a comprehensive educational/information transfer hypermedia program on environmental assessment. This program makes extensive use of graphics and hypertext.

Denno and Brail (1993) describe a decision support system that uses GIS, models, video and graphics to assist with transportation planning. Kindleberger and Meyers (1993) describe several applications in which multimedia techniques and GIS are being integrated to develop decision support systems and training systems. The Home Ownership Counseling System they developed incorporates audio, video, animation, text, and GIS data to assist potential first time home buyers.

Many of the systems and techniques described above are applicable to this project and the larger KSC environmental decision support system project.

III. PROCEDURES

3.1 Expert/Information System Development

The development of the environmental checklist expert/information system followed conventional expert system development techniques (Jones and Barrett, 1989; Waterman, 1986). The focus of the expert/information system was narrowed to the paper-based KSC Environmental Checklist. The project was further narrowed to focus on items in the checklist for which the KSC Ecological Program and the Remote Sensing and GIS Lab routinely provide information.

A conceptual design of the system was developed following discussions with those involved in the larger environmental decision support system project. It was decided to maintain the checklist format of the original KSC Environmental Checklist. Expert/information system modules would be accessed from individual items in the checklist.

Several sources of information were collected for use in the system development. Environmental checklists and environmental documents that are used at KSC were examined. Publications describing work conducted at KSC related to the environment and natural resources were reviewed. Literature describing EIA techniques and tools that are commonly used was reviewed as indicated in the Related Literature section.

The sections that follow describe the tools, information development and information collection procedures that were used. The system was developed on a 486 PC running Microsoft Windows. The system utilized expert system techniques, hypertext, GIS, computer-based video, and graphics.

3.2 GIS Data Development and Documentation

Spatial data play an important role in the management of the environment and natural resources. As described earlier, spatial data stored in GIS are often used in environmental impact assessment and other systems for environmental management. Answers to several of the questions on the KSC Environmental Checklist require the use of spatial data.

For this project, ARC/View (Environmental Systems Research Institute, 1992) was selected to display and manipulate the spatial data. ARC/View is a tool for displaying, browsing, and querying geographic or spatial data. ARC/View provides several significant advantages to this project. These advantages include:

- Easy to use. ARC/View is entirely icon driven. Within a few hours, someone with no experience with GIS tools can master the capabilities of ARC/View.
- Inexpensive. The cost of ARC/View for PCs and Macs is approximately \$400 and for UNIX workstations approximately \$900.
- Runs on multiple computers. ARC/View is available for PCs running Microsoft Windows, for Macs, and for UNIX workstations.
- Personnel familiar with ARC/View. Several NASA contractor personnel have experience with ARC/View.
- Supports multiple data formats. ARC/View is able to display spatial data that have various formats including ARC/Info and ERDAS.

The disadvantages of ARC/View to this project are listed below. However, several of these disadvantages will be eliminated with the release of a new version of ARC/View in late 1993.

- Limited data analysis capabilities. ARC/View is only able to display spatial data and generate simple tables of statistics. Additional data analysis capabilities will be included in the new version of ARC/View. However, additional capabilities will mean a more complex tool and thus additional training requirements/commitments.
- No data entry/modification capabilities. Data can't be entered or modified using ARC/View. Other GIS tools must be used to modify or add data sets. The new release of ARC/View will eliminate some of this problem. However, this again increases the knowledge of ARC/View that a user must have.

Despite these limitations, ARC/View should provide the capabilities required early in the project and the KSC environmental decision support project. If additional capabilities are required, the new release of ARC/View or ARC/Info would likely be viable options.

3.3 GIS Data at KSC

Extensive GIS data sets have been developed for KSC by the KSC Remote Sensing and GIS Lab to support ecological and habitat research and management. The data sets identified for use in this project include:

- Vegetation types
- Habitat types
- Soils
- Wetlands
- Groundwater recharge areas
- Scrub jay habitat
- KSC boundaries

Other GIS data sets have been developed by KSC Master Planning. The data sets identified for use in this project include:

- Roads
- Buildings and parking lots
- Water lines
- Sewer and storm drains
- Power lines
- Drainage ditches
- Railroads

Numerous other GIS data layers are available from the Remote Sensing and GIS Lab and from Master Planning. Portions of the remaining data sets will be useful to future decision support system

development efforts.

The above data sets listed above served as a starting point for development of the spatial data component of the system. However, before using these data in ARC/View, additional data labeling/documentation was required. Data from the KSC Remote Sensing and GIS Lab were stored in ERDAS software formats. Although ARC/View could display these data without modification, the tabular reports of statistics aren't available for this format. Statistics are only available within ARC/View for data in ARC/Info formats. The ERDAS data were converted to the ARC/Info format as described in Appendix A (Conversion of ERDAS GIS Data for Use in ARC/View). In addition to conversion of formats, the addition of labels was required. For example, the converted soil map contained only numerical labels for areas and thus soil series names were added to the Info data tables. This process is also described in Appendix A. Data from Master Planning were used without modification.

Once data layers were modified using the procedure described above, several *views* were constructed. Within ARC/View, a view is a collection of GIS data layers and information about how to display these data and the associated tabular data. A view defines the data that is to be displayed when it is opened, the data that is available for display, the order of data layers in the display menu, data layer legends, and other information. Several views were developed to support the different components of the expert system.

3.4 Expert System Development Tool

Expert system development tools are usually used to facilitate the development of expert systems (Engel et al., 1991). For this project, KnowledgePro Windows (Knowledge Garden, 1991) was selected as the expert system development tool. KnowledgePro Windows supports both object-oriented and rule-based programming approaches. The reasons for its selection include:

- Preferred computer for delivery. The preferred computer for delivery of the prototype system is a PC.
- Multimedia capabilities. KnowledgePro Windows supports PC multimedia extensions and provides several multimedia capabilities. Graphics and hypertext can be easily incorporated within applications using KnowledgePro commands.
- No runtime license. KnowledgePro Windows applications do not require the purchase of a runtime license.
- Interaction with other Microsoft Windows applications. KnowledgePro Windows runs under Microsoft Windows and thus interacts well with other Windows applications including PC ARC/View.

3.5 Habitat Module

One of the checklist questions asks whether endangered species habitat will be affected. To answer this question, an extensive expert system module was developed. A publication on biodiversity at KSC (currently under development) served as the starting point for this module. This document contains extensive descriptions of wildlife species at KSC, habitat descriptions, and wildlife species habitat requirements.

The major sections of the habitat module include:

- Habitat descriptions
- Habitat locations at KSC
- Species use of habitats
- Species descriptions

Existing GIS data layers were reclassified to provide a GIS layer showing the habitat locations described in the biodiversity publication. Numerous wildlife and habitat graphics were added to the module. The process is described below.

3.6 Video Capture

Video footage of wildlife, habitat, and facilities at KSC was taken using a VHS video camera. Short video clips ranging from 3 to 10 seconds were captured and compressed for use in the expert system using Intel Smart Video Recorder hardware and software (Intel, 1993). Once captured, these video clips can be played on a PC without the video board using Microsoft Video for Windows (Microsoft, 1992). The runtime version of Microsoft Video for Windows does not require a runtime license and is easily accessed from within KnowledgePro Windows applications.

The video was captured at a 160x120 pixel resolution at between 12 and 15 frames per second. This resolution and speed was found to be the most effective option, given computer speed and file storage requirements. A compressed video clip of 75 frames (provides 5 seconds of video at 15 frames/second) required approximately 425 Kbytes of storage. This storage requirement was obtained using the compression algorithm recommended to obtain the highest quality video clips. Other compression options are available but will degrade the quality of video clips. Video for Windows is capable of uncompressing and playing compressed files at the speeds at which they were captured.

3.7 Graphics Capture

Slides and photographs of habitat and wildlife were captured using a scanner for the photographs and a slide digitizing device for the slides. The digitized graphics were saved in the bmp (bitmap for Windows) format. Graphics in this format were easily incorporated in KnowledgePro Windows applications.

IV. RESULTS AND DISCUSSION

This section describes the *views* developed for ARC/View and shows example screens from these views. The computer-based environmental checklist is also described and example screens are provided. Keep in mind however, these systems are difficult to describe because of their dynamic nature and should be viewed on the computer to be fully appreciated.

4.1 Example Screens from ARC/View

Several *views* were developed in ARC/View using the GIS data layers described previously. These views can be accessed from the appropriate locations in the expert/information system or used in a stand-alone manner with ARC/View.

Figure 4.1 shows one of the views that was developed within ARC/View. The tools window (left side of the figure) allows users to query the database, measure features, find areas impacted within a given distance of a feature, and zoom within the data displayed. The center portion of the figure shows the habitat map for KSC. Using the tools menu, one can zoom to the location of interest and determine the type or types of habitat that would be affected and generate statistics of interest. The right side of the figure shows the data layers available within the view and the legend for the habitat layer. Other data layers and their legends can be displayed as desired.

Figure 4.2 shows a portion of the habitat layer from the previous figure. In this case, the user has zoomed to an area of interest and selected a map polygon. A table for the polygon is displayed showing the area, perimeter and habitat type. Other more complex tables of information are readily generated.

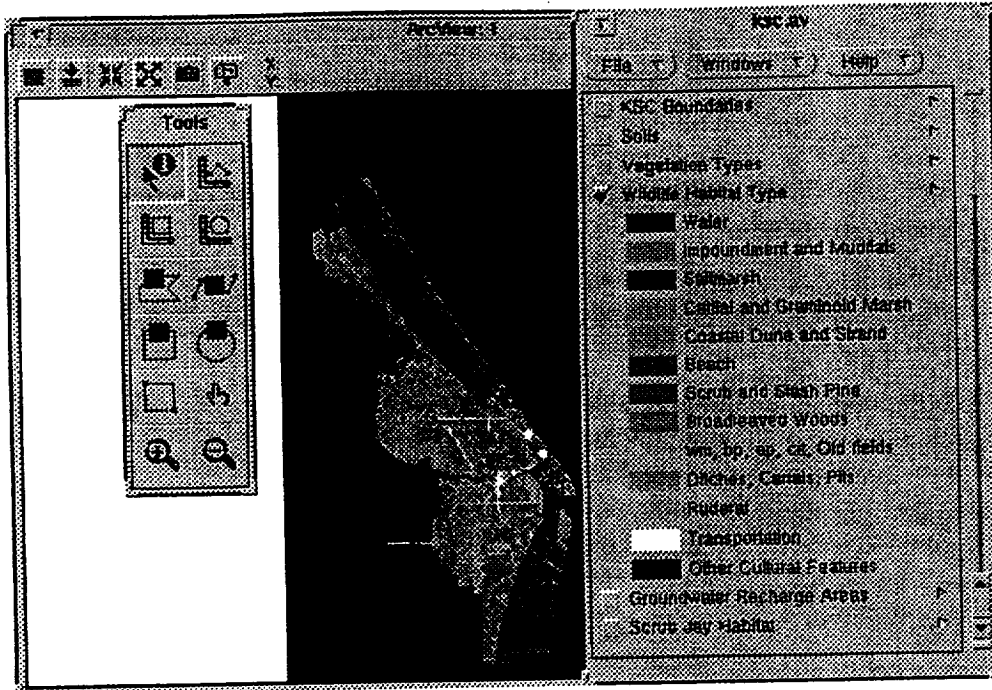


Figure 4.1 Portion of an ARC/View Data View

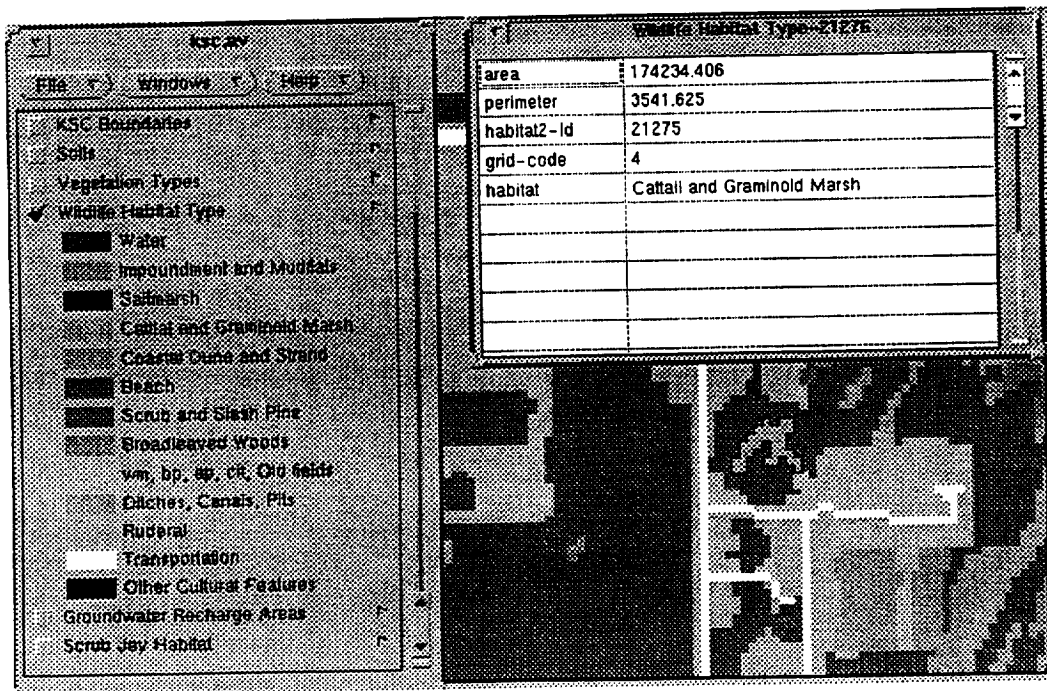


Figure 4.2 Data Table from an ARC/View Data View

4.2 Example Screens from the Expert/Information System

This section describes the operation of the expert/information system and provides example screens. Once the system is started, the user is asked to provide the same information as requested by the paper-based KSC Environmental Checklist. A portion of the computer-based checklist is shown in Figure 4.3. The mouse is used to check the appropriate items. Checking some items invokes questions that require additional information. For example, indicating that construction will be done in a flood plain invokes a question requesting whether the construction will be in a 100 year or 500 year floodplain.

Items within the system that are underlined are hypertext links. Selecting these items with the mouse provides additional information. For example, selecting *Construction in wetlands* provides access to a GIS layer showing wetland locations and types at KSC.

Selecting the *Endangered species habitat* link from the checklist provides access to the habitat and wildlife species expert/information system. Example screens are shown in Figures 4.4-4.8. After selecting the *Endangered species habitat* link, the user is presented a menu (Figure 4.4) requesting the type or types of habitat that will be affected. Selecting habitat types provides additional information about that habitat at KSC. Figure 4.5 shows the menu that is presented when the Scrub and Slash Pine habitat is selected from the previous menu. Figure 4.6 shows the information displayed when the *Species utilizing Scrub and Slash Pine* link is selected from Figure 4.5. This table lists the species that use this habitat and describe how they use it. Each of the species is linked to additional information about that species as shown in Figure 4.7 for scrub jays. Figure 4.8 shows a portion of the additional textual information available about scrub jays. The system user is returned one level (window) at a time after accessing additional information to allow additional links to be explored.

Does the construction, installation, removal, activation, or operation of the proposed project involve: (Fill out one page for each alternative considered; See instructions, when in doubt, check yes)

Alternative number and description:

- YES NO Discharge of any substance to the environment
- YES NO Land alteration, excavation, dewatering or removal of vegetation
- YES NO Construction in wetlands
- YES NO Construction in floodplain
- YES NO Generation of ionizing or non-ionizing radiation or use of any radiation source
- YES NO Asbestos-containing materials or facilities
- YES NO PCB-contaminated materials or equipment
- YES NO Generation of waste other than normal construction wastes
- YES NO Use or storage of Hazardous or Toxic Materials
- YES NO Aboveground or underground storage tanks
- YES NO Construction of high-voltage transmission lines (over 100kV)

Figure 4.3 Portion of the Expert System Environmental Checklist

What types of habitat are found within the area of interest?

- Estuarine Waters and Edge
- Impounded Open Water and Mud Flats
- Salt Marsh
- Cattail and Graminoid Marshes
- Coastal Dune and Strand
- Ocean Beach
- Scrub and Slash Pine
- Broadleaved Forests
- Miscellaneous Disturbed [e.g. citrus, Brazilian Pepper]
- Islands
- Ditches and Canals
- Ruderal Grass
- Continue

Figure 4.4 List of KSC Habitat Types

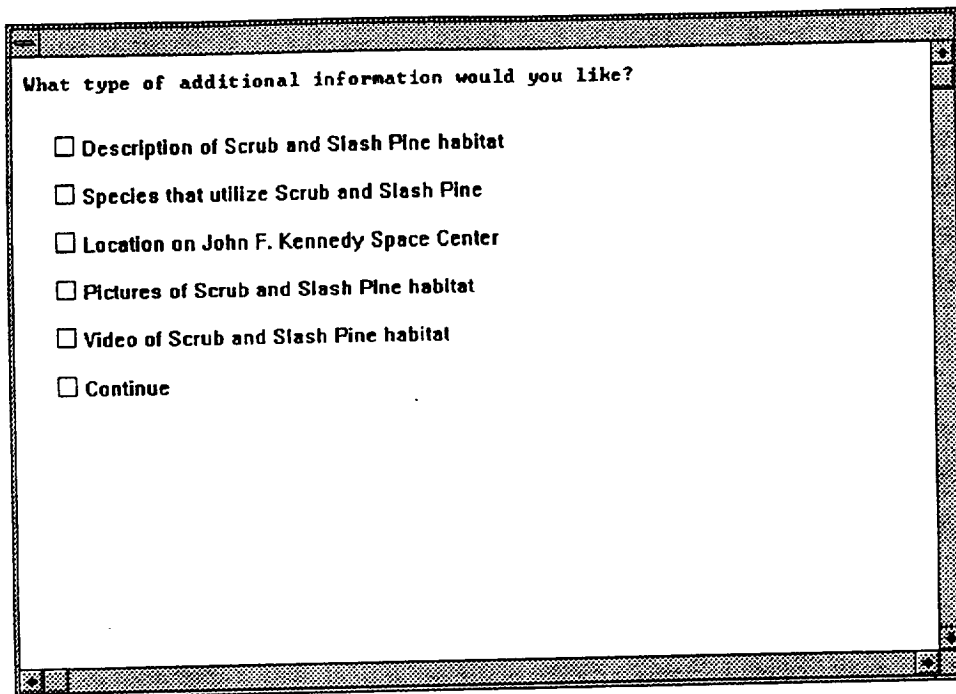


Figure 4.5 Menu to Access Scrub Habitat Information

Scrub and Slash Pine Species

The total number of priority taxon requiring this habitat is 15.
The total number of priority taxon using this habitat is 30.

Habitat Use	
<u>Florida Scrub Jay</u>	4
<u>Southern Bald Eagle</u>	B4F1
<u>Southeastern Beach Mouse</u>	2R
<u>Eastern Indigo Snake</u>	4
<u>Florida Pine Snake</u>	4
<u>E. Diamondback Rattlesnake</u>	4
<u>Arctic Peregrine Falcon</u>	F1
<u>Gopher Tortoise</u>	4
<u>Gull-billed Tern</u>	F1
<u>Bobcat</u>	4
<u>Kingsnake</u>	1
<u>Osprey</u>	B4
<u>Barn Owl</u>	F2
<u>Pileated Woodpecker</u>	3
<u>Eastern Coachwhip</u>	4
<u>Red-shouldered Hawk</u>	F2
<u>Cooper's Hawk</u>	3
<u>Northern Harrier</u>	F2
<u>Barred Owl</u>	F2
<u>Florida Mouse</u>	4
<u>Merlin</u>	F3
<u>Dusky Pigmy Rattlesnake</u>	4

Figure 4.6 Species that Require and Use Scrub Habitat

What type of additional Florida Scrub Jay information would you like?

- Seasonal abundance, population size, and measures of endangerment
- Protected status on John F. Kennedy Space Center
- Habitat location on John F. Kennedy Space Center
- Pictures of scrub habitat
- Video of scrub
- Picture of a Florida Scrub Jay
- Video of a Florida Scrub Jay
- General information
- Continue

Figure 4.7 Menu to Access Scrub Jay Information

Florida Scrub Jay information

Florida Scrub Jay (*Aphelocoma coerulescens coerulescens*)

The Florida Scrub Jay is referred to as a disjunct race of Scrub Jay which is widespread in the western U.S. and Mexico. Recent genetic studies have shown that the Florida Scrub Jay has become distinct enough to be considered a separate species and it should be officially recognized as such within the near future (J. Fitzpatrick and D. MacDonald, pers. comm.). Its population has declined by at least 50% due to habitat loss and degradation (Cox 1987) and many remaining populations are vulnerable due to low population size, habitat fragmentation, and degradation (Fitzpatrick et al. 1991).

The Florida Scrub Jay diet consists mainly of insects and small vertebrates throughout most of the year. Acorns are extremely important in winter when insect numbers are low. Nesting season ranges from late February to early June. Reproductive success varies from one year to the next at Archbold Biological Station in Highlands County (Woolfenden and Fitzpatrick 1984) and on KSC (Breininger and Smith unpublished data). Unlike the western Scrub Jay, the

Figure 4.8 Textual Scrub Jay Information

V. SUMMARY AND CONCLUSIONS

A decision support system project has been initiated to develop software to assist with environmental and natural resource management at KSC. A module for this decision support system was developed based on the paper-based KSC Environmental Checklist. This checklist must be completed before most activities of significance can be conducted/completed at KSC. The paper-based checklist requires information and expertise that those completing the checklist often don't have and can't easily obtain.

A prototype expert/information system based on this checklist was developed. This system was developed using KnowledgePro Windows (an expert system and multimedia development tool) and PC ARC/View (a geographic information system (GIS) data display and analysis tool). The resulting system requires a PC with Microsoft Windows and PC ARC/View. The checklist was implemented such that users complete a computer-based checklist in a manner similar to completing the paper-based check list. Based on items checked, the system requests additional information only as required. Users of the system can request additional information to complete checklist items through hypertext links. The hypertext links provide a variety of information. For example, one of the checklist items asks whether endangered species habitat will be affected. Selection of this link accesses an extensive expert/information system that provides information about KSC wildlife habitat, KSC habitat locations, wildlife species, and wildlife habitat requirements. Much of the textual information in this section is based on a biodiversity publication being prepared by NASA contractor personnel. Hypertext links are used to provide access to this information. This system also makes extensive use of video clips, graphics, and GIS data.

The prototype system is in a usable form, but many additional features and capabilities should be added. Suggestions for further development are included in the next section. With only slight modifications, the habitat and wildlife module could be used as an educational and training tool at KSC and many other locations including schools.

VI. RECOMMENDATIONS FOR CONTINUED WORK

This section describes areas of continued and future work. Suggestions for continued work on the KSC environmental and natural resources decision support system are given in Appendix D.

The expert system developed in this project is by no means complete. Several improvements to the habitat module should be made. Additional graphics and video of wildlife and habitat should be added. Graphics of habitat at various stages of development and for the different seasons of the year should be added. Additional hypertext links within the program should be added. Further testing of current links should also be conducted. At the present time, the GIS data is not linked to the expert system because the computer used for development did not have enough RAM to allow both KnowledgePro Windows and ARC/View to run concurrently. The habitat module could be slightly modified and used for educational purposes at KSC, schools, the Fish and Wildlife Service, the National Park Service, and other special interest groups.

The GIS data for ARC/View should be subdivided based on KSC regions. The time required to display the current data layers for all of KSC on the computer used for development (33 MHz 486 with 4 Mbytes RAM) is very slow. The habitat layer with slightly over 23,000 polygons requires approximately 16.5 minutes to be entirely displayed. The scrub jay habitat with approximately 7,800 polygons requires just over 4 minutes for display. Dividing KSC into areas and subdividing the GIS data based on these regions would improve the display time to an acceptable level. The appropriate regions could be accessed through a menu or graphically using a map of KSC.

Additional GIS data layers for ARC/View must be converted for use on the PC. Disk space available on the PC was not adequate to move all of the data desired. Appendix C provides instructions for conversion of the data layers.

Additional expert system modules could be developed and added to the checklist to assist with items not addressed in the current system. Other GIS data layers would be useful in answering some of the checklist items. For example, one of the questions asks whether the proposed action will be in the 100 or 500 year flood plains. Only paper copies of these maps are available at the present time.

Other environmental checklists are used at KSC for certain activities. These checklists could also be implemented as expert systems. Other expert system modules could be developed based on questions/problems encountered by KSC pollution officers. Development of these expert systems will also require additional GIS data layers.

Opportunities exist to develop more comprehensive decision support tools for KSC such as those described in the Related Literature section. Such tools would help with day-to-day operations and long term planning.

APPENDIX A. CONVERSION OF ERDAS GIS DATA FOR USE IN ARC/VIEW

ERDAS GIS files can be used directly by ARC/View. However, when used in this way, the data is treated as an image and thus statistics are not available within ARC/View. If statistics from ERDAS files are desired within ARC/View, a conversion process must be followed. If labeling of areas is desired, include the steps below that will allow labeling of polygons (steps 3-9). The necessary steps are described below.

1. Use ARC/Info to convert ERDAS data to ARC GRID format (erdasgrid <in_erdas_file> <out_grid>).
2. Use ARC/Info to convert the ARC GRID format to the ARC POLY format (gridpoly <in_grid> <out_cover>). This step is necessary since the statistics desired within ARC/View can't be obtained from GRID data.
3. Within Info, select the Info file that contains the appropriate GIS layer statistics (SEL <info_file>). The GIS layer files that must be modified are those that end with .PAT.
4. Export the information in selected file to an ascii file (EXPORT <filename> ASCII).
5. Cancel the association between the Info file and any external files (EXTERNAL)
6. Remove the data from the Info file (PURGE).
7. Modify the template of the Info file to include a new column or columns that will contain the labels that will be added (MODIFY). MODIFY is used to interactively add a column or columns. Do not remove existing columns.
8. An additional column or columns of data can be added to the file that was exported. A simple CLIPS program was used to add a column containing labels that described numeric polygon labels. The CLIPS code used for the wetlands data is provided in Appendix B (CLIPS File for Adding Text Labels to Info Table for Wetlands Layer). Other techniques could also be used to add labels. For example, soil series names were added as labels when developing the soils layer.
9. Import the external file into the modified template (IMPORT <ascii_file_name> ASCII).
10. The resulting information can now be accessed by ARC/View and the labels added will be displayed in ARC/View tables.

**APPENDIX B. CLIPS FILE FOR ADDING TEXT LABELS
TO INFO TABLE FOR WETLANDS LAYER**

```
(defrule xx
(declare (salience 5))
(initial-fact)
=>
(open "wetlands.exp" ex.file "r") ;; open the exported wetlands data file
(open "wet" ex2 "w")           ;; open a file to write new data in
)
```

```
(defrule xx2
;;; read original data and determine text label to add - then add new label
(initial-fact)
=>
(bind ?c 1)
(bind ?c2 0)
(while (< ?c2 24000)
(while (< ?c 6)
(bind ?x (read ex.file))
(bind ?c (+ ?c 1))
(fprintout ex2 ?x crlf)
)
(readline ex.file)
(if (= ?x 1) then (fprintout ex2 "
(if (= ?x 2) then (fprintout ex2 "
(if (= ?x 3) then (fprintout ex2 "
(if (= ?x 4) then (fprintout ex2 "
(if (= ?x 5) then (fprintout ex2 "
(if (= ?x 6) then (fprintout ex2 "
(if (= ?x 7) then (fprintout ex2 "
(if (= ?x 8) then (fprintout ex2 "
(if (= ?x 9) then (fprintout ex2 "
(if (= ?x 10) then (fprintout ex2 "
(if (or (< ?x 1) (> ?x 10)) then (fprintout ex2 "
(fprintout ex2 crlf)
(bind ?c 1)
(bind ?c2 (+ ?c2 1))
)
)
```

APPENDIX C. CONVERSION OF UNIX ARC/VIEW COMPATIBLE FILES FOR USE WITH THE MS WINDOWS ARC/VIEW

Before files that are used with the UNIX ARC/View can be used with the MS Windows version of ARC/View, these files must be converted. The conversion process is described below.

1. A file called `pname.c` is installed with ARC/View during the installation of the MS Windows version of ARC/View. This file must be moved to a machine running UNIX and compiled. The executable file that is produced should be called `pname`.
2. A UNIX shellsript file called `pcwksp` is also installed during the installation of the MS Windows version of ARC/View. This file should also be moved to a machine running UNIX. It may be necessary to modify the shell script to include the path to the `pname` executable file that was created in the previous step.
3. Use `pcwksp` to create a PC workspace from the UNIX workspace of GIS files (`pcwksp unix_workspace pc_workspace`).
4. Copy the PC workspace created in the previous step to the PC with the MS Windows version of ARC/View. Once copied, the MS Windows version of ARC/View should access the data in a manner similar to ARC/View on UNIX. The speed at which MS Windows ARC/View displays coverages may be significantly slower than under UNIX, especially with files containing large numbers of polygons.

**APPENDIX D. STRATEGY FOR DEVELOPMENT OF ENVIRONMENTAL
AND NATURAL RESOURCES DECISION SUPPORT SYSTEMS AT KSC**

1. Identify potential users of the decision support system.
2. Provide key spatial (GIS) data layers to a small group (5 or 6) potential users.

ARC/View should be used to provide access to the GIS data. ARC/View is a tool for viewing GIS data. It also provides minimal data manipulation capabilities.

The strengths of ARC/View include: 1) very easy to use - minimal training of those using it would be required (probably 2 one hour sessions); 2) cost - ARC/View for a PC is less than \$500; 3) NASA contractor personnel are familiar with ARC/View; 4) runs on several machines (PCs under Microsoft Windows, Macs, and UNIX machines)

Data layers that will be provided initially must be identified. These will likely include: soils, wetlands, vegetation, habitat, facilities, and roads. For this limited number of data layers, existing computer hard disks may be of suitable size. Additional hard disk space may be required or the data layers could be placed on CD ROMs.

Involving potential users of software at the initiation of a project is almost always a good idea. Involvement will increase their support for the project and acceptance of the final product.

3. Work with potential users of the decision support system to identify their needs and questions they must answer that require the use of spatial and other environmental and natural resource data.

Providing ARC/View and some GIS data layers to potential users (see above) will provide an opportunity to better determine what capabilities the decision support systems should have.

4. Once user needs are identified, decision support system modules can be identified to fit the needs of the users. Software modules can also be prioritized for development at this point.

The additional data (both GIS and non-spatial) that will be required must also be identified. The potential users can also assist with identification of additional data requirements.

5. Develop decision support system modules. Prototypes can be quickly delivered to potential users to gain feedback.
6. Provide final versions of decision support system modules to potential users.
7. Maintain and develop new databases.
8. Maintain software modules.

VII. REFERENCES

- Beanlands, G. E. and P. N. Duinker. 1983. An Ecological Framework for Environmental Impact Assessment in Canada. Institute for Resource and Environmental Studies, Dalhousie University, Halifax, Nova Scotia.
- Bear, D. 1989. NEPA at 19: a primer on an "old" law with solutions to new problems. *Environmental Law Reporter* 19 ERL 10060-10069.
- Bennett, D.B. 1989. Water Resources Analysis of a Multiobjective Drainage Network in the Indian River Lagoon Basin. Water Resources Research Center Publication No. 108. University of Florida, Gainesville, FL. 154p.
- Bisset, R. 1988. Developments in EIA methods. In: *Environmental Impact Assessment, Theory and Practice* (ed. P. Wathern). Unwin Hyman, Boston. 332 p.
- Bonazountas, M., A. Brecker, and R. G. Vranka. 1988. Mathematical environmental fate modeling. Ch. 5. In: I. Bodek, W. J. Lyman, W. F. Reehl and D. H. Rosenblatt (eds.), *Environmental Inorganic Chemistry*, Pergamon Press, New York.
- Bregman, J. I., and K. M. Mackenthun. 1992. *Environmental Impact Statements*. Lewis Publishers, Inc. Chelsea, Michigan 48118. 279 p.
- Breining, D.R. 1990. Avifauna of hammocks and swamps on John F. Kennedy Space Center. *Florida Field Naturalist* 18:21-44.
- Breining, D.R. and P.A. Schmalzer. 1990. Effects of fire and disturbance on plants and birds in a Florida oak/palmetto scrub community. *American Midland Naturalist* 123:64-74.
- Carlson, J.R., S.L. Ekblad, L.P. Aydelott, and P.E. Fuglestad. 1993. The conservation planning application system of the soil conservation system. In: *Application of Advanced Information Technologies: Effective Management of Natural Resources*. ASAE Publication 04-93. pp. 24-33.
- Clark, B.D., A. Gilad, R. Bisset, and P. Tomlinson. 1984. *Perspectives on Environmental Impact Assessment*. Reidel, Dordrecht, Netherlands.
- Dee, N. 1973. Environmental Evaluation for Water Resource Planning. *Water Resources Research* 9:523-535.
- Denno, N. and R.K. Brail. 1993. Decision support systems offer organizational efficiency for transportation planning. *GIS World*. June 1993. pp. 40-45.
- Dierberg, F.E. and K.L. Jones. 1989. Assessment of stormwater runoff for recycle in cooling towers at KSC, Florida. *Water Resources Bulletin* 25(1):43-47.

- Dooley, J.E. and Newkirk, R.W. 1976. Corridor Selection Method to Minimize the Impact of an Electrical Transmission Line. James F. McLaren Ltd. Toronto.
- Dreschel, T.W., B.C. Madsen, L.A. Maull, C.R. Hinkle, and W.M. Knott, III. 1990. Precipitation chemistry: Atmospheric loading to the surface waters of the Indian River Lagoon basin by rainfall. *Florida Scientist* 53(3):184-188.
- Dreschel, T.W., R.B. Smith, and D.R. Breininger. 1991. Florida scrub jay mortality on roadsides. *Florida Field Naturalist* 18:82-83.
- Duke, K.M. 1977. Environmental Quality Assessment in Multiobjective Planning. Batelle Columbus Laboratories. Columbus, Ohio.
- Duke, K.M. 1979. Environmental Quality Evaluation Procedure Implementing Principles and Standards for Planning Water Resources Programs. US Water Resources Council. Washington, DC.
- Dwornik, D.S. 1984. A Stormwater Collection and Reuse System for the Kennedy Space Center, Florida. Water Resources Research Center Publication No. 83. University of Florida, Gainesville, FL. 197p.
- Engel, B.A., D.D. Jones, J.R. Wright, and S. Benabdallah. 1991. Selection of an expert system development tool. *AI Applications in Natural Resources*. 5(1):15-22.
- Environmental Systems Research Institute. 1992. ArcView User's Guide. ESRI, Redlands, CA.
- Fabos, J.G. 1978. The METLAND Landscaping Planning Process: Composite Landscape Assessment, Alternative Plan Formulation and Plan Evaluation. Part 3 of the Metropolitan Landscape Planning Model. Research Bulletin No. 653. University of Massachusetts Agricultural Experiment Station. Amherst, Massachusetts.
- Fedra, K., M. Karhu, T. Rys, M. Skoc, M. Zebrowski, and W. Ziemia. 1987. Model-based Decision Support for Industry-Environmental Interactions: A Pesticide Industry Example. WP 87-97. International Institute for Applied Systems Analysis. Laxenburg, Austria.
- Fedra, K. 1991. A computer-based approach to environmental impact assessment. In: A.G. Colombo and G. Premazzi (eds.). *Proceedings of the Workshop on Indicators and Indices for Environmental Impact Assessment and Risk Analysis*. pp. 11-40.
- Fedra, K., L. Winkelbauer, and V.R. Pantulu. 1991. Expert systems for environmental screening: An application in the lower Mekong basin. International Institute for Applied Systems Analysis, RR-91-19. Laxenburg, Austria. 169p.
- Fedra, K. 1992. Intelligent environmental information systems. *Vortrage Wasserbau-Symposium Wintersemester 1991/92 "Okologie und Umweltvertraglichkeit"*. pp. 289-311.

- Fedra, K. 1993. Expert systems in water resources simulation and optimization. In: J.B. Marco et al. (eds.). *Stochastic Hydrology and its Use in Water Resources Systems Simulation and Optimization*. Kluwer Academic Publishers. The Netherlands. pp. 397-412.
- Galloway, C.E. 1978. *Assessing Man's Impact on Wetlands*. Report No. 78-136. University of North Carolina Water Resources Research Institute. Raleigh, North Carolina.
- Garcia, L.A., M. Armbruster, K. Buehler, and G. Williams. 1993. Spatially based decision support system developed for evaluation of wildlife habitat. In: *Application of Advanced Information Technologies: Effective Management of Natural Resources*. ASAE Publication 04-93. pp. 98-105.
- Gough, M. and G. Edwards-Jones. 1993. A knowledge management system for training planners to be aware of the environmental consequences of development projects. In: *Application of Advanced Information Technologies: Effective Management of Natural Resources*. ASAE Publication 04-93. pp. 342-350.
- Gurganus, J., M. Pearson, B.A. Engel, and D.D. Jones. 1993. *Environmental Assessment Resource Guide*. Agricultural Engineering Department, Purdue University, W. Lafayette, IN.
- Heaney, J.P., B. Koopman, D.S. Dwornik, and D.R. Saliwanchik. 1984. *Water Resources of the Kennedy Space Center*. Water Resources Research Center Publication No. 82. University of Florida, Gainesville, FL. 259p.
- Hill, M.A. 1968. A Goals-Achievement Matrix for Evaluating Alternative Plans. *Journal of the American Institute of Planners* 34:19-28.
- Holling, C.S. 1978. *Adaptive Environmental Assessment and Management*. Wiley, Chichester. 377p.
- Intel. 1993. *Intel Smart Video Recorder Installation Guide and Video Tips*. Intel Corporation, Hillsboro, OR.
- Jones, D.D. and J.R. Barrett. 1989. Building expert systems. In: *Knowledge Engineering in Agriculture*, J.R. Barrett and D.D. Jones (Eds.), ASAE Monograph No. 8, ASAE, St. Joseph, MI. pp. 21-31.
- Jorgensen, S. E. 1991. Environmental management modelling. In: *Introduction to Environmental Management* (eds., P. E. Hansen and S. E. Jorgensen). Elsevier, New York. 403 p.
- Kane, J. 1972. A Primer for a New Cross-Impact Language -KSIM. *Tech. Forecasting Soc. Change*. 4:129-142.
- Keeney, R.L. and II. Raiffa. 1976. *Decisions with Multiple Objectives: Preferences and Values Tradeoffs*. Wiley, New York, NY.
- Kindleberger, C.P. and T. Meyers. 1993. Local governments tap affordable multimedia

applications. GIS World. June 1993. pp. 46-50.

Knowledge Garden. 1991. KnowledgePro Windows User Manual. Knowledge Garden, Setauket, NY.

Kowalski, D.G. and C.F. Bagley. 1993. United States army natural resources information management. In: Application of Advanced Information Technologies: Effective Management of Natural Resources. ASAE Publication 04-93. pp. 1-8.

Lein, J. 1989. An expert system approach to environmental impact assessment. International Journal of Environmental Studies, 33:13-27.

Leopold, L.B., F.E. Clarke, B.B. Hanshaw, and J.R. Balsley. 1971. A Procedure for Evaluating Environmental Impact. US Geological Survey Circular 645. US Government Printing Office. Washington, DC.

Liroff, R.A. 1976. A National Policy for the Environment. NEPA and its Aftermath. Indiana University Press. Bloomington, Indiana. 273p.

Madsen, B.C., T.W. Dreschel, and C.R. Hinkle. 1989. Characterization and evaluation of acid rain in Central Florida from 1978 to 1987 - Ten year summary report. NASA Technical Memorandum TM 102149. Kennedy Space Center, FL.

Mailander, J.L. 1990. Climate of the Kennedy Space Center and vicinity. NASA Technical Memorandum 103498. Kennedy Space Center, FL. 62p.

McHarg, I. 1968. A Comprehensive Highway Route-Selection Method. Highway Research Record No. 246. Highway Research Board, Washington, DC.

Microsoft. 1992. Microsoft Video for Windows User's Guide. Microsoft Corporation.

Negahban, B., C. Fonyo, W. Boggess, J. Jones, K. Campbell, G. Kiker, E. Hamouda, E. Flaig, and H. Lal. 1993. A GIS-based decision support system for regional environmental planning. In: Application of Advanced Information Technologies: Effective Management of Natural Resources. ASAE Publication 04-93. pp. 169-178.

Provanha, J.A. and C.R. Hall. 1991. Observations of associations of sea grass beds and manatees in East Central Florida. Florida Scientist 54:87-98.

Richardson, S.E., W.J. Hansen, R.C. Solomon, and J.C. Jones. 1978. Preliminary Field Test of the Water Resources Assessment Methodology (WRAM): Tensas River, Louisiana. Paper No. Y-78-1. US Army Corps of Engineers. Vicksburg, Mississippi.

Schmalzer, P.A. and C.R. Hinkle. 1990a. Geology, geohydrology and soils of Kennedy Space Center: A review. NASA Technical Memorandum 103813. Kennedy Space Center, FL. 46p.

Schmalzer, P.A. and C.R. Hinkle. 1990b. Flora and threatened and endangered plants of John F. Kennedy Space Center, Florida. NASA Technical Memorandum 102791. Kennedy Space Center, FL. 68p.

Schmalzer, P.A. and C.R. Hinkle. 1991. Dynamics of vegetation and soils of oak/saw palmetto scrub after fire: Observation from permanent transects. NASA Technical Memorandum 103817. Kennedy Space Center, FL. 149p.

Shih, T. and W.L. Mills, Jr. 1993. FMDSS: A spatial and ecological forest management tool for upland hardwoods of the central United States. In: Application of Advanced Information Technologies: Effective Management of Natural Resources. ASAE Publication 04-93. pp. 199-193.

Solomon, R.C, B.K. Colbert, W.J. Hansen, S.E. Richardson, L. Canter, and E.C. Vlachos. 1977. Water Resources Assessment Methodology (WRAM) - Impact Assessment and Alternative Evaluation. Technical Report No. Y-77-1. US Army Corps of Engineers. Vicksburg, Mississippi.

Sorensen, J.C. 1971. A Framework for Identification and Control of Resource Degradation and Conflict in the Multiple Use of the Coastal Zone. Dept. of Landscape Architecture, University of California at Berkeley.

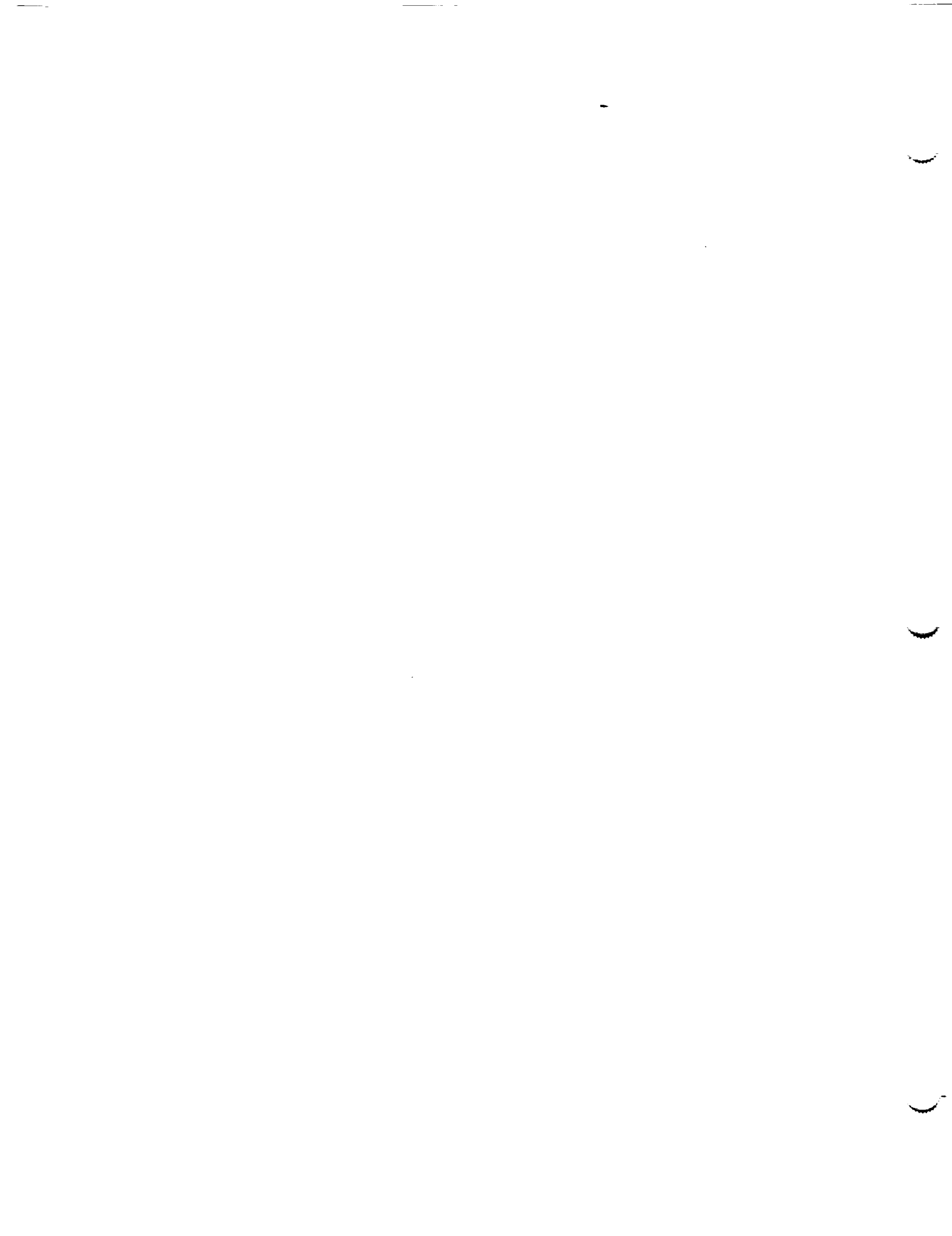
Sorensen, J.C. 1972. Some Procedures and Programs for Environmental Impact Assessment. In: R. B. Ditton and T.L. Goodale (eds.). Environmental Impact Analysis: Philosophy and Methods. University of Wisconsin Sea Grant Program. Madison, Wisconsin. pp. 97-106.

Tomlinson, P., and S. F. Atkinson. 1987. Environmental audits: A literature review. Environmental Monitoring and Assessment 8: 239-261.

U.S. Environmental Protection Agency. 1992. Framework for Ecological Risk Assessment. EPA/630/R-92/001.

US Fish and Wildlife Service. 1976. Habitat Evaluation Procedures. US Fish and Wildlife Service. Department of the Interior. Washington, DC.

Waterman, D. 1986. A guide to expert systems. Addison-Wesley Publishing Co., Reading, MA, 419p.



1994018765

N94-23238

1993 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA

58-47
197194
p. 30

LIGHTNING STUDIES USING LDAR AND LLP DATA

PREPARED BY:	Dr. Gregory S. Forbes
ACADEMIC RANK:	Associate Professor
UNIVERSITY AND DEPARTMENT:	The Pennsylvania State University Department of Meteorology
NASA/KSC	
DIVISION:	Communications and Instrumentation
BRANCH:	Instrumentation and Measurements
NASA COLLEAGUE:	Carl Lennon
DATE:	August 16, 1993
CONTRACT NUMBER:	University of Central Florida NASA-NGT-60002 Supplement: 11

ACKNOWLEDGEMENTS

Participation in the NASA/ASEE Summer Faculty Fellowship Program has been an enjoyable and professionally stimulating experience. Appreciation is extended to all those who support this program financially, and to the many individuals who made presentations or gave tours to the Faculty Fellows. The tours and presentations made the Faculty Fellows more aware of the scope and complexity of activities that take place at the Kennedy Space Center. Particular thanks are extended to Dr. E. Ramon Hosler, Co-Director of the NASA/ASEE Summer Faculty Fellowship Program, and to Kari L. Stiles, Administrative Assistant, who coordinated the program extremely well.

It was a great pleasure to work with Carl Lennon and others associated with the Lightning Detection and Ranging (LDAR) system in TE-CID-3. Carl Lennon is to be congratulated for the development of the LDAR system, which shows great capabilities in revealing the presence of electrical discharges within, and emanating from, thunderstorms. His support and encouragement of the applied research reported here is appreciated. The author benefitted from discussions with Launa Maier, who has vast knowledge of thunderstorms and their electrical processes. She helped acquire Lightning Location and Protection (LLP) data and helped the author establish computer networking within the NASA TE-CID-3 Weather Research Facility.

The author is very appreciative of the efforts of the NYMA employees who operate the LDAR system and archive its data, and for patiently dealing with many data requests and the increased computer load while the data were being processed. Steve Schaefer was particularly helpful in setting up access to the data-processing computer and in retrieving data for the studies described below. Preston Porter helped acquire real-time LLP data.

Greg Taylor and Mark Wheeler of the Applied Meteorology Unit (AMU; ENSCO) helped acquire Patrick AFB radar data for a case study using the McGill radar display system. Mike Maier of CSR and Gerry Talley of TE-CID-2 also helped provide LLP data.

The author extends his best regards to the meteorologists of the 45th Weather Squadron, detached out of Patrick AFB, USAF Space Command, who provide KSC with operational weather support from the Cape Canaveral Forecast Facility. It is hoped that the applied research begun this summer can help them in the difficult task of issuing lightning advisories.

Finally, the author acknowledges the efforts of all those involved in weather support of the space program. This includes Dr. Jack Ernst, Director of the Weather Support Office, NASA Headquarters (ME); John Madura, Chief, Weather Projects Office (TM-LLP-2); Dr. Frank Merceret, AMU Chief (TM-PCO-4); Jim Nicholson (DE-AST); the Spaceflight Meteorology Group of the National Oceanic and Atmospheric Administration at Johnson Space Center; atmospheric scientists at Marshall Space Flight Center; and many others.

ABSTRACT

This study intercompared lightning data from LDAR and LLP systems in order to learn more about the spatial relationships between thunderstorm electrical discharges aloft and lightning strikes to the surface. The ultimate goal of the study is to provide information that can be used to improve the process of real-time detection and warning of lightning by weather forecasters who issue lightning advisories. The Lightning Detection and Ranging (LDAR) system provides data on electrical discharges from thunderstorms that includes cloud-ground flashes as well as lightning aloft (within cloud, cloud-to-cloud, and sometimes emanating from cloud to clear air outside or above cloud). The Lightning Location and Protection (LLP) system detects primarily ground strikes from lightning. Thunderstorms typically produce LDAR signals aloft prior to the first ground strike, so that knowledge of preferred positions of ground strikes relative to the LDAR data pattern from a thunderstorm could allow advance estimates of enhanced ground strike threat. Studies described in the report examine the position of LLP-detected ground strikes relative to the LDAR data pattern from the thunderstorms. The report also describes other potential approaches to the use of LDAR data in the detection and forecasting of lightning ground strikes.

SUMMARY

The Lightning Detection and Ranging (LDAR) system shows great capabilities in detecting lightning within thunderstorms, and emanating from thunderstorms to ground or to clear air outside or above the clouds. Real-time displays of LDAR data -- monitored on a daily basis and compared to personal observations of cloud, thunder, and lightning conditions in the vicinity of KSC -- often revealed intriguing patterns of discharges. For example, in the late stages of thunderstorm systems, many long quasi-horizontal discharges spanned tens of kilometers along parallel paths which appeared to follow the top and bottom of the anvil cloud layer. LDAR can give insights into thunderstorm dynamics.

Thunderstorms typically produce LDAR signals aloft prior to the first ground strike, so that knowledge of preferred positions of ground strikes relative to the LDAR data pattern from a thunderstorm could allow advance estimates of enhanced ground strike threat. The average lead time in this study was 5.26 minutes. Only 19% of the storms showed a strike to ground within a minute of the first LDAR detection of the storm. Two remote sensing methods of detecting ground strikes were examined: the Lightning Location and Protection (LLP) system, which detects primarily ground strikes from lightning, and LDAR, using data points below 2.6 km.

Eighty-five percent of the LLP-detected ground strikes occurred within the boundaries of the LDAR data cluster associated with the thunderstorms, and 98% occurred within or less than 2 km beyond the storm LDAR data boundaries. Ground strikes were detected during 41% of the minutes while LDAR detected thunderstorms. Ground strikes tended to occur near the center of the LDAR-defined storm. Over 4000 LLP ground strikes and nearly two million LDAR data points were used in the study, which involved cases from June and July, 1993.

The report also describes other potential approaches to the use of LDAR data in the detection and forecasting of lightning ground strikes. These involved the clustering of LDAR data into volumes encompassing the lightning activity of a thunderstorm. For intercomparison, LDAR and LLP data were grouped together in samples of one-minute duration. The characteristics and evolutions of these "LDAR-defined storms" were examined in relation to concurrent and subsequent ground strikes. Three hundred ninety-one LDAR storms were defined in this manner, of which 169 (43%) produced ground strikes at some time during their existence. At forecast times more than 10 minutes, more than half of the ground strikes occur outside the boundaries of the current LDAR boundaries of the storm, indicating the development of new thunderstorm cells along the flanks of the existing storms. On days when thunderstorms were quasi-stationary, some preference was shown for future ground strikes along the southeast flank of the current LDAR pattern; when thunderstorms were moving, some preference was shown for future ground strikes along the rear (west-southwest) flank.

The LDAR characteristics of storms during minutes with ground strikes were compared to those without ground strikes. On average, when there were ground strikes the storms had many more LDAR events, much larger volume containing LDAR events, larger cross-sectional area containing LDAR events, a much larger vertical depth of LDAR events, a much greater number of LDAR events per unit volume, and a greater number of LDAR events per unit cross-sectional area.

Storms that produced ground strikes typically had more LDAR events than those which did not produce ground strikes. Sixty-two percent of the storms without ground strikes never had LDAR data rate greater than 50 data points per minute. Only 11% of the storms producing ground strikes had peak LDAR data generation rate that low. By contrast, 81% of the storms with ground strikes had a period of more than 150 LDAR data points per minute.

TABLE OF CONTENTS

Section	Title
I	INTRODUCTION
II	GENERATION OF THE DATA BASE
2.1	LDAR Data
2.2	LLP Data
2.3	LDAR-Defined Storms
2.4	Composite Data Base
2.4.1	Derived LDAR Parameters
III	INTERCOMPARISON OF LDAR AND LLP DATA
3.1	LDAR Lead Time
3.2	Percentage of Time LDAR Storms Produce Ground Strikes
3.3	LDAR Detection of Ground Strikes
3.4	Location of Ground Strikes Relative to LDAR Storms
3.5	Subsequent Ground Strikes Relative to Current LDAR Storm
3.6	Probability of Ground Strikes More Than 5 Nautical Miles Beyond LDAR Storm Edge
3.7	Cessation of LDAR Activity and End of Ground Strike Threat
3.8	Differences Between Storms That Produce Ground Strikes and Those That Never Produce Ground Strikes
3.9	LDAR Storm Differences During Minutes With and Without Ground Strikes
3.10	LDAR-Based Parameters in Relation to Variations in Ground Strike Rate
IV	CONCLUDING REMARKS

LIST OF ILLUSTRATIONS

- Figure 1-1. Horizontal Locations of LDAR Data Points that Occurred on the Map Domain
- Figure 2-1. LDAR Cubes Superimposed on LDAR Data Points Used in LDAR Storm Definition
- Figure 2-2. LDAR Storm Identification Numbers Resulting from the First Pass of the Storm Identification Scheme
- Figure 2-3. LDAR Storm Identification Numbers After the Second Pass of the Storm Identification Scheme
- Figure 2-4. Map of LDAR Storm Tracks
- Figure 2-5. Time Series of LDAR Storm X, Y, and Z Centroid Positions
- Figure 2-6. Time Series of Three LDAR-Derived Storm Parameters in Relation to LLP Ground Strikes
- Figure 2-7. Time Series of Three Additional LDAR-Derived Storm Parameters in Relation to LLP Ground Strikes
- Figure 3-1. Locations of Ground Strikes Relative to LDAR Storm Center, by Fractional Distance to Edge of LDAR Storm
- Figure 3-2. Locations of Ground Strikes Relative to LDAR Storm Center, by Fractional Positions East/West and North/South
- Figure 3-3. Locations of Ground Strikes Relative to LDAR Storm Center, by Distance East/West and North/South
- Figure 3-4. Locations of Ground Strikes 10 Minutes Later Relative to Current LDAR Storm Center, by Fractional Positions East/West and North/South

I. INTRODUCTION

The Kennedy Space Center (KSC) is located in one of the regions of the United States (and even the world) that encounters the most lightning strikes to ground per unit area (refs. 1,2,3). The possibility of lightning at the surface or aloft is, of course, a hazard that must be avoided during launches. On a daily basis, however, there are many operations at KSC which must be curtailed if there is a threat of a lightning strike to ground in the vicinity. The accuracy and timeliness of lightning advisories, therefore, has both safety and economic implications. The ultimate goal of the research described in this report is to provide information that can be used to improve the process of real-time detection and warning of lightning by weather forecasters who issue lightning advisories.

Prior to the development and implementation of remote sensing techniques for detection of lightning during the 1970s, weather forecasters had little ability to know in detail where lightning was occurring in regions beyond the range of eyesight. The locations of thunderstorms had to be estimated using radar -- based upon crude correlations between the intensity of precipitation and cloud electrification, or by recognizing characteristic shapes of thunderstorm clouds in satellite imagery. National networks of lightning remote sensing systems now exist (ref 4), that allow a direct knowledge of the location of lightning strikes to ground with accuracy of a few kilometers. Two special, denser networks of remote sensing equipment have been established to provide highly accurate information concerning lightning in the vicinity of KSC: the Lightning Location and Protection (LLP) system and the Lightning Detection and Ranging (LDAR) system. In addition, a Launch Pad Lightning Warning System (LPLWS) is operated to detect electric fields at the surface induced by thunderstorms or other atmospheric conditions, and a Catenary Wire Lightning Instrumentation System (CWLIS) detects electrical surges in wires at the launch pads when struck by lightning. Data from the LLP and LDAR systems were used in this study.

The LLP system (ref 5) detects lightning ground strikes through use of a network of magnetic direction finding antennae which sense electromagnetic disturbances in a broad band of frequencies triggered by lightning. Individual antennae detect a particular ground strike at different azimuth angles, and the location of the ground strike is essentially determined by finding the point of intersection of lines drawn from the antennae toward the source of the disturbance. The LLP system is approximately 90% efficient in detecting ground strikes near KSC, with position accuracy of about 1 km.

The LDAR system was developed by Carl Lennon and colleagues at KSC TE-CID-3 (ref 6). Its antennae detect lightning-induced disturbances at 66 MHz frequency. This system uses a time of arrival (TOA) approach, and extremely accurate timing through use of the Global Positioning System (GPS). The lightning-induced disturbance, travelling at the speed of electromagnetic propagation, arrives at different antennae at slightly different times. The three-dimensional position of the lightning source is determined by converting these time offsets into distance differences, and then performing a triangulation. The LDAR system began real-time operation in June, 1992.

The LDAR system can generate up to 10,000 data points per second, yielding numerous data points per lightning flash. Tests of the position accuracy of the LDAR data by Launa Maier have shown that within 10 km of the central antenna, 95% of the data points are accurate to better than 200m, and 50% are accurate to better than 100m. Figure 1-1 illustrates a sample plot of LDAR data during one minute, projected to their positions at the surface.

Because the LDAR system can detect thunderstorm electrical discharges aloft, it has the potential to allow anticipation and warning of imminent ground strikes.

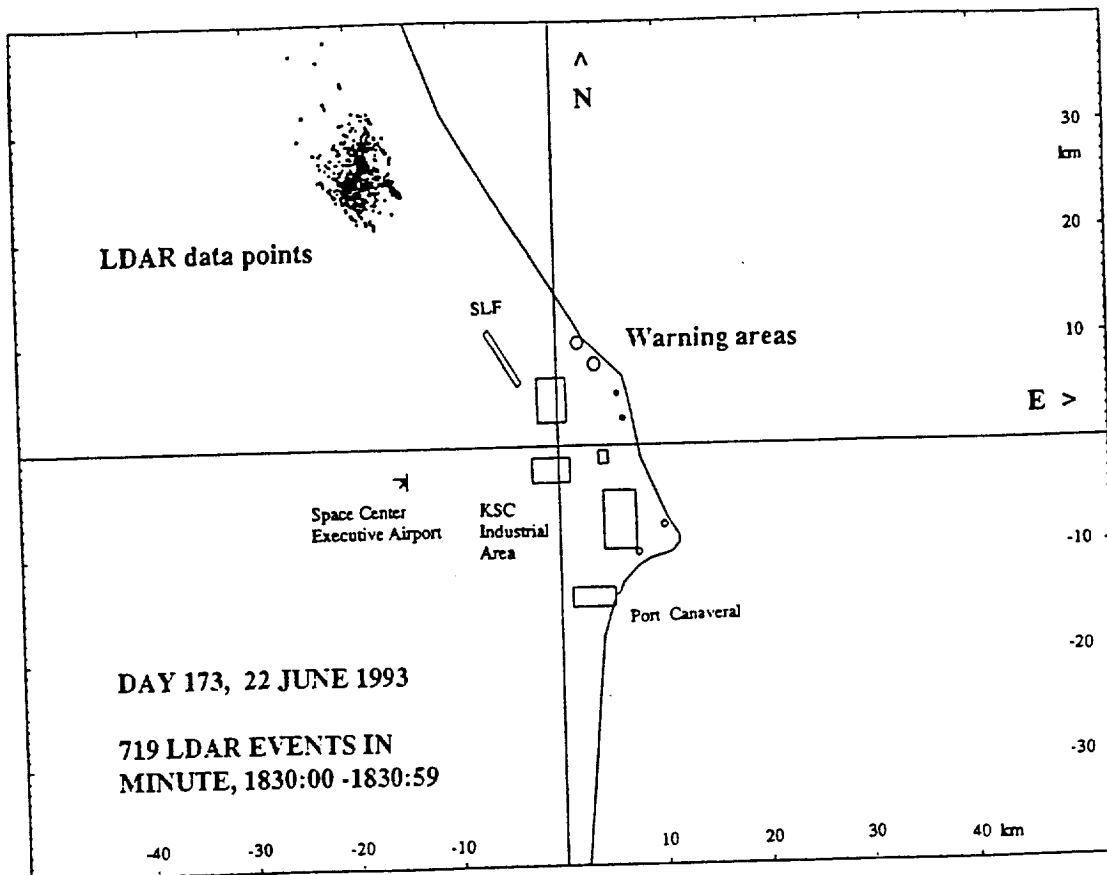


Figure 1-1. Horizontal locations of the LDAR data points that occurred on the map domain during the minute beginning at 1830 UTC on 22 June 1993 (Julian date 173), shown by dots. Circles and rectangles represent KSC and U.S. Air Force warning areas. Solid lines cross at the LDAR central site, about 1 km north of the KSC Headquarters Building. The map domain extends 52 km west and east of the central site and 40 km north and south.

For that potential to be realized in the lightning detection and warning process, however, additional knowledge must be acquired regarding the relationship between lightning aloft and ground strikes. This project contributes to that effort.

This study examines spatial and temporal relationships between lightning aloft detected by LDAR and ground strikes detected by the LLP and LDAR systems. Several questions are addressed. (1) What is the lead time between the first detection of LDAR events and ground strikes? (2) Are there preferred locations of occurrence of ground strikes relative to the pattern of discharges detected by LDAR? (3) Can future positions of ground strikes be anticipated through knowledge of the current LDAR event pattern? (4) Are there signatures in the LDAR data that can be used to determine when the threat of ground strikes has ended? (5) Are there signatures in the LDAR data that reveal which storms produce ground strikes and which do not?

II. GENERATION OF THE DATA BASE

2.1 LDAR DATA

The time series of signals received from the LDAR network antennae are used to generate a sequential listing of LDAR data containing the x, y, and z positions of each event and the time of occurrence. The horizontal locations are given as distances east (positive x) or west and north (positive y) or south of the LDAR central site (site 0). The original data, therefore, can be used to reconstruct details of individual flashes. Data are stored in files one hour in duration.

For purposes of this study, where spatial patterns were the primary issue, the original LDAR data were converted from a data sequence into a 4-dimensional array format. Time was originally reported by day, hour, minute, second, and microsecond, and position in meters and fractions of meters. Array data in this study were accumulated in one-minute segments, and cubes 1 km^3 in volume. Time accumulation was done by truncating the seconds and microseconds from the time, thereby referencing all data to the start of the minute. Horizontal accumulation was done by converting the original x, y, and z positions to integer kilometers. Finally, LDAR data points more than 52 km from the LDAR central site in the east-west direction or more than 40 km in the north-south direction were discarded. This yielded a converted data base, referred to hereafter as "LDAR cubes" or "LDAR bins", containing the number of LDAR data points per minute within each element of the array of LDAR 1 km^3 volumes. Like the original files, the LDAR cube files contain data for one hour.

The array horizontal dimensions of $-52:52$ by $-40:40$ (km) were chosen (1) to confine the study to a domain where LDAR data were highly accurate, yet (2) allow a domain sufficiently large that moving storms could be followed for at least one hour without moving out of the domain. The rectangular shape was chosen to match the aspect ratio of computer screens. The array vertical dimension was 0-20 (km), with accumulation done by integerizing the z position to the nearest kilometer. The use of 1 km vertical resolution kept the size of the data base reasonably small, but certain results below suggest that use of higher (~ 0.25 km) vertical resolution would have been preferable.

Figure 2-1 shows the LDAR cubes emanating from the LDAR data of Fig. 1-1. Some of the LDAR cubes overlap, suggesting the nucleus of an LDAR-defined storm classification scheme, discussed in section 2.3.

2.2 LLP DATA

The original LLP data were recorded sequentially in time and by latitude and longitude. For use in this study the time was referenced to the start of the minute, as with the LDAR data, and the horizontal position was converted to distance in kilometers east-west and north-south of the LDAR central site. LLP positions were stored as real numbers. Additional information available in the LLP data files, such as amplitude of the signal, number of return strokes per ground strike, and polarity of the lightning event, were not used in this study. Original LLP files contained all ground strikes during the day. Processed LLP data used in this study were output in hourly files to match the LDAR files.

2.3 LDAR-DEFINED STORMS

Weather forecasters use satellite imagery and weather radar to detect and track cloud and precipitation patterns, respectively. They are accustomed to identifying the patterns of satellite and radar imagery associated with rain showers

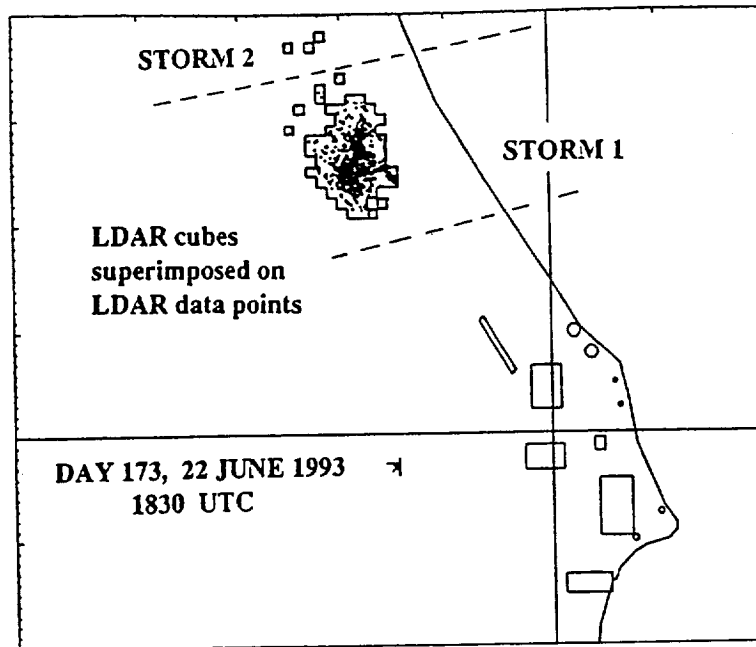


Figure 2-1. LDAR cubes superimposed on LDAR data points used in LDAR storm classification, as in Fig. 1-1, except only a portion of the domain is shown.

and thundershowers. Likewise, as can be seen from Fig. 2-1, the LDAR data points are clustered into packets affiliated with thunderstorms. It seemed natural, then, to cluster the LDAR cube data into groups, yielding "LDAR-defined storms" or "LDAR storms."

While humans can visually recognize patterns and clusterings of data readily, rules must be specified before comparable results can be obtained by computer. In the first pass of the computerized storm classification procedure, LDAR cubes from a particular minute are examined. Contiguous groupings of LDAR cubes, or near-contiguous groups within 3 km horizontal distance of each other, are clustered together to form LDAR-defined storms. The storm classification scheme proceeds by first finding the mean (x,y) position of all LDAR cubes during the minute and clustering the LDAR cubes nearest the domain mean into an LDAR storm. The storm classification scheme then continues until all LDAR cubes within the domain have been assigned a storm number. LDAR cubes separated by more than 3 km from an identified LDAR storm trigger the initiation of a new LDAR storm. For example, the closest of the three LDAR cubes farthest north on Fig. 2-1 is 4 km from the cluster of LDAR cubes to the south, so that these three will be classified as LDAR storm 2. All other LDAR cubes are within 3 km of a neighbor, and are clustered to form LDAR storm 1.

Figure 2-2 shows the result of the first pass of the LDAR storm classification scheme, from a different case. This is for minute 1800 UTC of Julian day 188, 7 July 1993. Five LDAR storms have been identified in this pass.

A second pass in the storm classification procedure eliminates "false storms" and renumbers the remaining LDAR storms consistently within the hourly file. LDAR storms were considered to be false storms if they had volume less than 4 km^3 . In terms of LDAR cubes, this means that at least four LDAR cubes had to contain at least one LDAR data point during the minute. This lenient rule would not discard a true lightning event, characterized by hundreds or thousands of LDAR data points, but typically eliminates spurious points such as those generated by electrically noisy automobiles, test signals at the central site, and the occasional false LDAR point passed by the LDAR quality control procedures.

Figure 2-3 shows the result of the second pass of the LDAR storm classification scheme. In this case, false LDAR storms 2 and 5 of Fig. 2-2 have been discarded due to sub-threshold volume dimensions, and the remaining three storms have been renumbered.

Consistent LDAR storm numbering from minute to minute becomes complicated if LDAR storm shape changes dramatically or when adjacent LDAR storms expand toward each other and merge. The storm classification procedure keeps track of the horizontal position of each storm by computing the centroid position. This is the weighted-mean position of the LDAR cubes comprising the storm, where weighting is done by the number of LDAR events in the cubes. The rule invoked in the storm classification scheme is that a new storm identification number is initiated if the centroid position of the storm being processed is more than 6 km from the position of any storm during the preceding minute. Otherwise, the storm with centroid position nearest to a storm in the preceding minute is assigned its identification number. Should a large LDAR storm split into several LDAR storms (separated, by definition, by more than 3 km), the new storm nearest the original storm centroid would take the parent storm number and the other storms resulting from the split would obtain new storm identification numbers.

Consistent LDAR storm numbering is also complicated by the fact that in early and late stages some LDAR storms flash intermittently. During individual minutes the LDAR storm may disappear. The second pass of the storm classification procedure keeps track of the previous positions and identification numbers of all LDAR storms during the past 10 minutes. Intermittent LDAR storms that "reappear" within 10 minutes and 6 km of their previous positions retain their original numbers. In the unlikely event that a storm was inactive for more than 10 minutes, but reappears, it would be reassigned a new identification number.

In summary, after a two-pass procedure, LDAR-defined storms are numbered consistently for each minute during the hourly file. Clusters of LDAR cubes are identified as new LDAR storms when they are separated by more than 3 km from other LDAR cube clusters (storms) and when their centroid is separated by more than 6 km from any previously identified LDAR storm. This can occur (1) when isolated storms suddenly appear; (2) when an existing LDAR storm expands so dramatically and asymmetrically that its centroid shifts by more than 6 km; (3) when existing LDAR storms merge into a single, larger storm; (4) when an existing LDAR storm splits into several smaller LDAR storms; (5) when a previous LDAR storm that has been inactive for more than 10 minutes reappears. The percentage breakdown of these possibilities has not yet been studied, but it is believed that (1), (2), and (3) dominate.

The storm classification procedure generated a new three-dimensional (x,y,t) array of data containing LDAR storm identification numbers. This array identifies during each minute the LDAR storm number associated with the horizontal position x,y . If no LDAR cube in the column above point x,y contained any LDAR events during the minute, the array value was set to zero.

In the intercomparison studies of section 3, LLP ground strikes were attributed to the storm nearest to the LLP location. The nearest storm was identified by

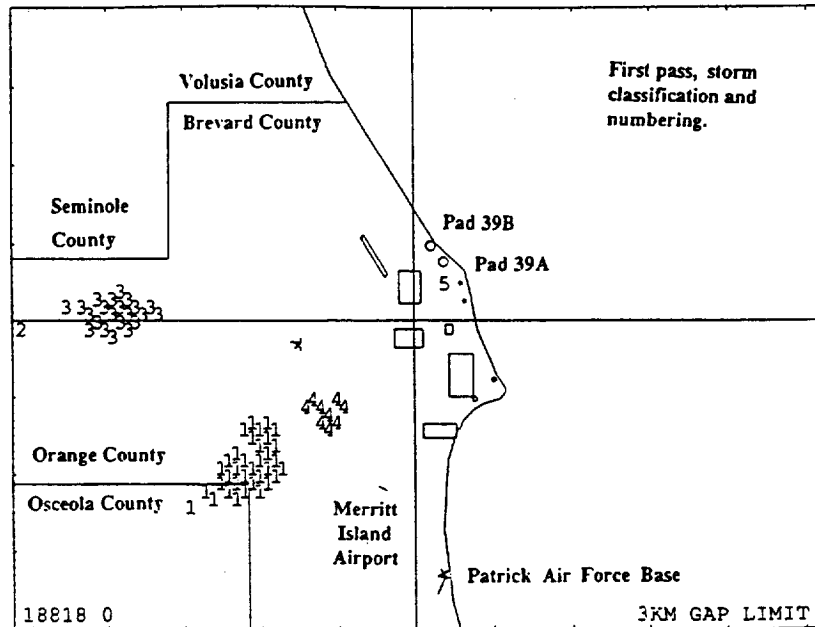


Figure 2-2. LDAR storm identification numbers resulting from the first pass of the storm identification scheme, from 1800 UTC, Day 188, 7 July 1993. A few additional landmarks have been shown.

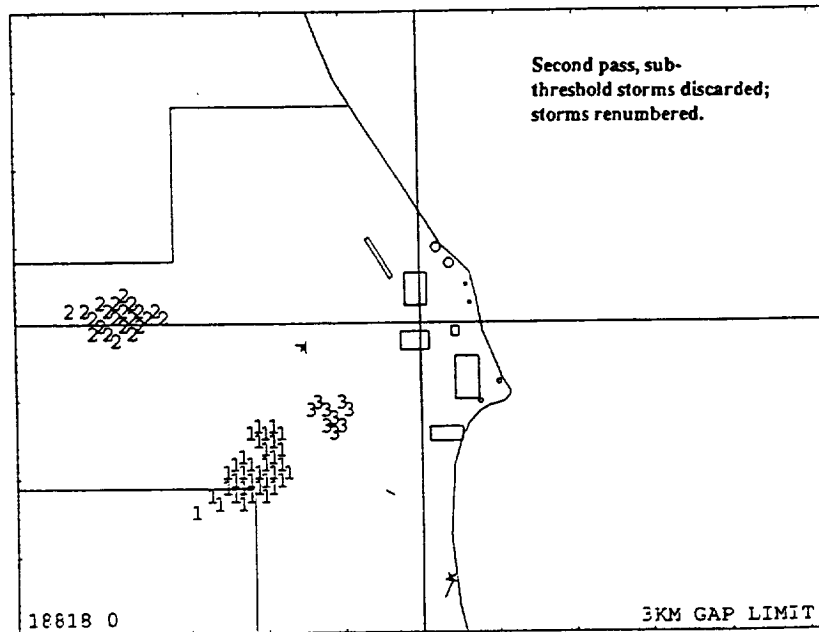


Figure 2-3. LDAR storm identification numbers after the second pass of the storm identification scheme. Identical to Fig. 2-2 except "false" storms 2 and 5 of Fig. 2-2 have been dropped and the remaining storms renumbered.

first finding the LDAR cube that contained LDAR events located nearest to the LLP data point, and then determining the LDAR storm identification number attributed to that cube.

2.4 COMPOSITE DATA BASE

The data base used in this study is the set of hours when both LDAR cube files and processed LLP files were available, and contained lightning events within the analysis domain. This consisted of 33 hours during June and July, 1993. The primary set of cases spanned the period from 22 June (day 173) to 21 July (day 201). A few additional hourly LDAR files were examined when no LLP events occurred within the map domain, but results from these have not been included in the results reported below.

The composite data base contained 1,853,170 LDAR data points and 4179 LLP ground strikes. The two-phase storm classification procedure described above identified 319 LDAR storms in this data set, of which 169 (43%) were affiliated with an LLP event during one or more minutes of their existence.

2.4.1 DERIVED LDAR PARAMETERS. During the LDAR storm classification procedure and subsequent data processing a number of additional parameters were derived from the LDAR storm and LDAR cube data base information. These included:

- (1) LDAR Storm Events -- total number of LDAR data points within the LDAR storm volume;
- (2) LDAR Storm Volume -- sum of the number of LDAR cubes (total km³) containing LDAR events;
- (3) LDAR Storm Area -- sum of the number of LDAR columns which contained LDAR events in some LDAR cube (essentially the LDAR storm cross-sectional area in km²);
- (4) LDAR Volume Density -- number of LDAR data points per km³, obtained by dividing the number of LDAR Storm Events by the LDAR Storm Volume;
- (5) LDAR Area Density -- number of LDAR data points per km², obtained by dividing the number of LDAR Storm Events by the LDAR Storm Area;
- (6) LDAR Centroid X, Y, Z Positions -- weighted mean positions of the LDAR cubes comprising the LDAR storm, weighted by the number of LDAR data points within the cubes;
- (7) LDAR Storm Breadth XS and YS and Depth ZS-- standard deviations of the x, y, and z positions of the LDAR cube positions comprising the storm;
- (8) LDAR Storm Z95 -- height below which 95% of the LDAR data points occur in the storm;
- (9) LDAR Storm Velocity Components CX, CY -- speed of LDAR storm movement in the east-west (positive when toward E) and north-south (positive when toward N) directions, based upon a least-squares fit of all storm centroid positions during the hour;
- (10) LDAR Storm Duration -- number of minutes between first and last appearance of the LDAR storm;
- (11) LDAR Storm Minutes -- number of minutes during which an LDAR storm produced LDAR events (i.e., eliminating "inactive" minutes);
- (12) LDAR-Based Ground Strikes -- these were obtained by identifying LDAR cubes below a threshold altitude, determined empirically, as described in Section 3.3.

Figure 2-4 shows a map of LDAR storm tracks that occurred between 1500 and 1559 UTC on 29 June 1993. This map was produced after pass 1 of the storm classification scheme, and storms 5, 7, and 9 are discarded during pass 2. LDAR storm movement tends to be somewhat wobbly, but an overall movement toward the east-southeast can be seen for storms 3 and 4. Other storms had shorter duration. The wobbly character of LDAR storm movement is due partly to quasi-

horizontal lightning flashes into new sections of cloud, causing abrupt lateral shifts in the LDAR storm centroid from one minute to the next.

Figure 2-5 shows time series of the x, y, and z centroid values as a function of time for storm 4 of Fig. 2-4. Dotted lines depict values one standard deviation above and below the mean, giving a measure of storm width in east-west and north-south directions and storm depth. Considerable growth is shown in all dimensions prior to the period of most frequent ground strikes.

Figure 2-6 shows the time variations of three other LDAR-derived parameters in relation to ground strikes: LDAR storm events, LDAR storm volume, and LDAR storm area. These are again for storm 4 of Figs. 2-4 and 2-5.

Figure 2-7 shows the time variations of three additional LDAR-derived storm parameters in relation to ground strikes for storm 4: LDAR volume density, LDAR area density, and Z95, the height below which 95% of the LDAR storm data points occur. LDAR volume density and LDAR area density appear to correlate rather well in this case with number of ground strikes per minute, with the peaks in ground strike rate occurring about 4 minutes after the peaks in volume and area density. Variations in Z95 are much more suppressed than the other parameters, and have apparently been suppressed due to use of LDAR cubes 1 km in vertical dimension. This appears to have been too coarse, and perhaps the LDAR bins should have had vertical dimension of about 0.25 km.

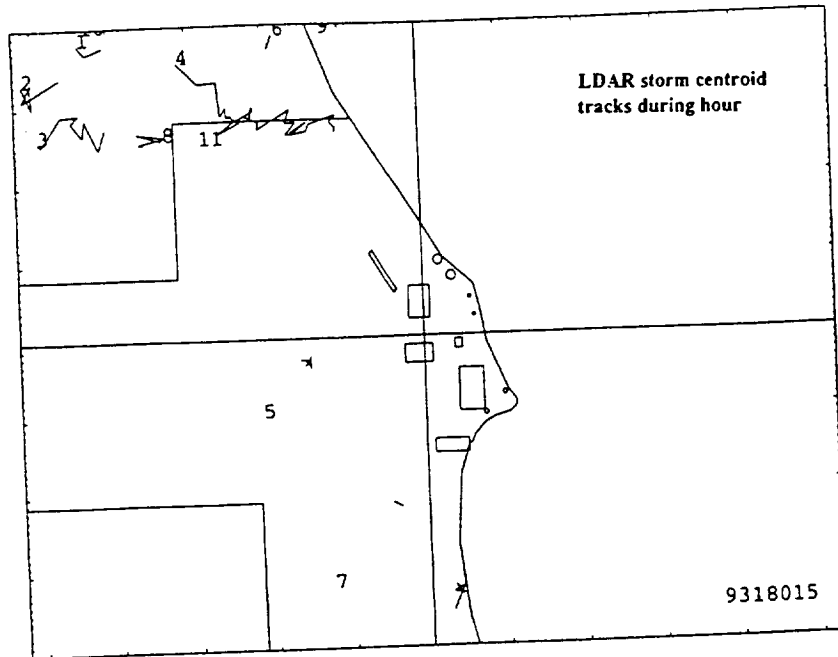


Figure 2-4. Map of LDAR storm tracks, from hour beginning 1500 UTC of day 180, 29 June 1993, after first pass of LDAR storm identification scheme.

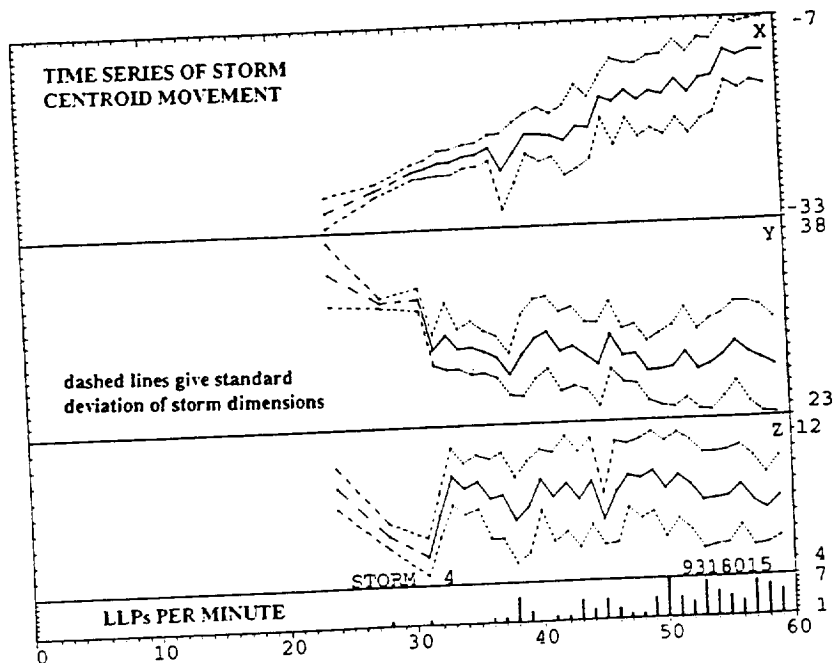


Figure 2-5. Time series of LDAR storm x, y, and z centroid positions, for storm 4 of Fig. 2-4. Number of LLP ground strikes per minute also shown.

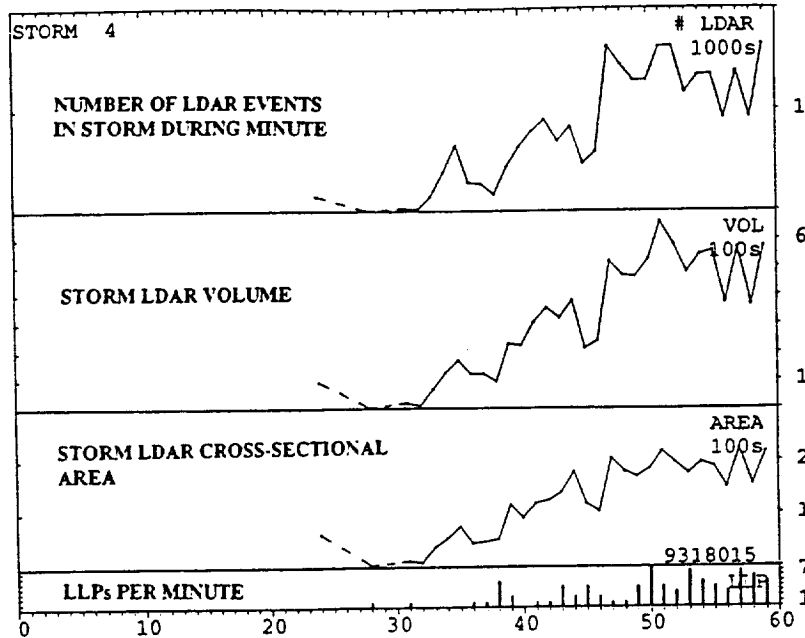


Figure 2-6. Time series of three LDAR-derived storm parameters in relation to LLP ground strikes: LDAR storm events, LDAR storm volume, LDAR storm area. For storm 4 of Figs. 2-4 and 2-5.

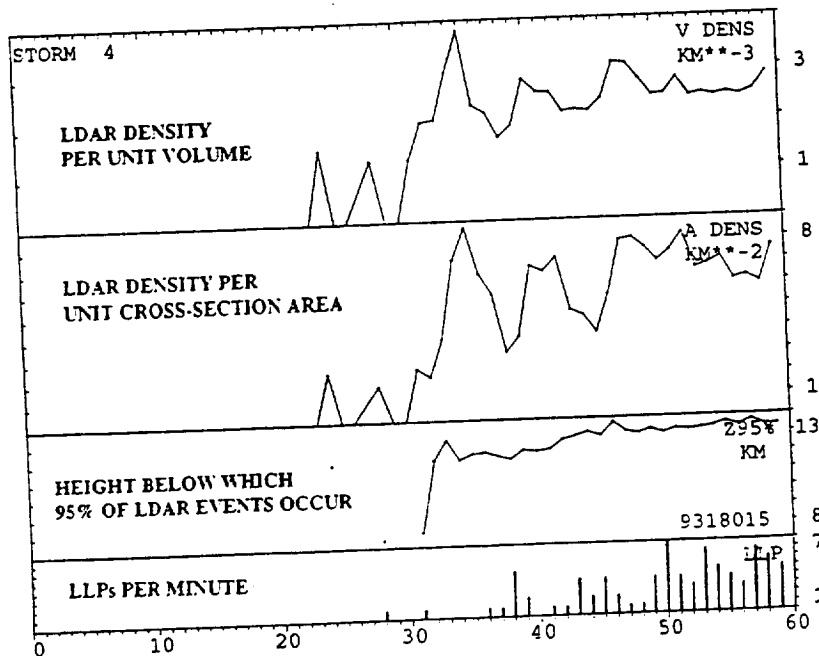


Figure 2-7. Time series of three additional LDAR-derived storm parameters in relation to LLP ground strikes: LDAR volume density, LDAR area density, and LDAR storm Z95. See text for explanations. For storm 4 of Figs. 2-4 to 2-6.

III. INTERCOMPARISON OF LDAR AND LLP DATA

3.1 LDAR LEAD TIME

Time ran out in the summer project before the author could conduct phase 3 of the storm classification project. This would have intercompared hourly files and renumbered storms consistently from one hourly file to the next. Because this phase of the project was not completed, studies of LDAR lead time used only a subset of the total data base. Eight hourly files were used, from days when these hours assuredly contained the initial formation of storms within the domain. Twenty-seven LDAR storms occurred during these hours. The average lead time between the first LDAR data defining a storm and the first LLP ground strike associated with the storm was 5.26 minutes. Only 3 of the 27 LDAR storms, 11%, had LLP ground strikes during the first minute of their existence.

Data were actually computed using all hourly files, and considering all LDAR storms in these files. Some of the first LDAR storm minutes determined in this manner will be from the first minute of the hour, though, and not really represent newly formed storms. Other first LDAR storm minutes will be from existing storms that moved into the domain. A qualitative inspection of the evolution of selected storms suggested that ground strikes were more frequent during the mature stages of storms than in the initial minutes or the late stages. Thus, LDAR storm "first minutes" in the full sample are expected to reveal a higher percentage producing ground strikes. This would also be manifested in a fictitiously low average lead time of 2.71 minutes. Of the 169 LDAR storms in the full sample, 19% of the "first" minutes had LLP ground strikes, but this number is fictitiously high.

3.2 PERCENTAGE OF TIME LDAR STORMS PRODUCE GROUND STRIKES

LDAR-defined storms were compared to ground strikes identified by LLP. Forty-one percent of the storm-minutes were associated with LLP ground strikes. Some LDAR storms never were associated with LLP ground strikes.

3.3 LDAR DETECTION OF GROUND STRIKES

LDAR cubes centered at 0, 1, 2, and 3 km were considered as potential indicators of lightning ground strikes. The highest LDAR events in these cubes are at altitudes slightly less than 0.5, 1.5, 2.5, and 3.5 km, respectively. Table 3-1 shows the number of LDAR-based column-minutes present in the sample using various height thresholds. A column-minute is defined as a minute when any LDAR cube in the column of 1 km² cross-sectional area contained LDAR data points at an altitude below the threshold. Very few LDAR data points are detected below 0.5 km.

TABLE 3-1. NUMBER OF COLUMN MINUTES WITH LDAR CUBES BELOW REFERENCE ALTITUDES

<u>Altitude</u>	<u>LDAR Column Minutes</u>
1.5 km	494
2.5 km	2840
3.5 km	7801

If it is assumed that a ground strike only occurs once in an LDAR column per minute, the numbers in Table 3-1 represent the number of LDAR-deduced ground strikes using the LDAR cube data base. To match the number of ground strikes detected using the LLP system when LDAR data was available, 3190, the best threshold altitude for LDAR data is determined by interpolation to be 2.62 km.

LLP ground strikes and LDAR-estimated ground strikes, using a threshold of 2.5 km for LDAR columns, were also intercompared on an LDAR storm basis. Both LLP and LDAR systems showed ground strikes associated with the storms during 798 minutes. Both LLP and LDAR systems showed that ground strikes were absent from storms during 1174 minutes. There were 309 storm-minutes when the LLP system indicated ground strikes were present and LDAR did not. There were 318 storm-minutes when LDAR indicated ground strikes were present and LLP did not.

Case studies (and past studies of LDAR detection efficiency) suggest that the efficiency of detection of near-ground strikes may decrease with increasing distance from the LDAR central site. Thus, it would probably be most accurate to determine a threshold altitude that was a function of distance from the LDAR central site.

3.4 LOCATION OF GROUND STRIKES RELATIVE TO LDAR STORMS

Positions of ground strikes detected by LLP and LDAR systems were composited with respect to the LDAR storm centers. This was done in three ways. One compositing method counted the number of ground strikes in one-kilometer distance increments with respect to the LDAR storm center. The mean distance to the LDAR storm edge was computed along a radial from the LDAR storm centroid through the ground strike location. The drawback to this method is that individual storm widths varied greatly, making it difficult to determine the percentage of ground strikes which occurred within the LDAR storm boundary. Ground strikes were also composited by distance east/west and north/south of the LDAR storm center.

A second, "fractional" compositing method involved computing the distance to the ground strike location and the distance to the LDAR storm edge along the radial between the LDAR storm center and the ground strike location. The ground strike locations inside the LDAR storm boundary were composited by fractional distance to the storm edge, in categories 0 to 10. Category 0 included ground strikes from the LDAR storm center outward to a distance less than 5% of the way to the storm edge. Category 10 included ground strikes at a distance 95% of the distance to the storm edge outward to 5% beyond the storm edge (or to just less than 0.5 km beyond storm edge, whichever was greater). Category 5, for example, included ground strikes from 45% to just less than 55% of the distance to the storm edge. Ground strikes were also composited by fractional distance east/west and north/south of LDAR storm center. When ground strikes were outside the LDAR storm edge in this second method, they were composited in one-kilometer increments as in method 1. Category 11 included ground strikes 0.5 to just less than 1.5 km beyond LDAR storm edge and category 20 representing ground strikes 9.5 km or more LDAR beyond storm edge.

A third compositing method was used with moving LDAR storms. Here it was anticipated that there might be a preference for ground strikes in the quadrant toward which the LDAR storm was moving, for example. Thus, the third compositing method used a storm-relative frame of reference: forward/rear and left/right with respect to storm movement. Left and right orientations are determined by looking from behind the LDAR storm toward the direction it is moving. Ground strikes were then composited by distances and by fractional positions forward/rear and left/right of the LDAR storm center with respect to the storm movement vector.

Figure 3-1 shows the locations of ground strikes relative to the LDAR storm during the same minute period. This display is based upon the second compositing method above; i.e., by fractional distance to the LDAR storm edge, which is located at category position 10 along the x axis. Figure 3-1a depicts ground strikes detected using the LLP system; Figure 3-1b depicts LDAR ground strikes in terms of LDAR columns below 2.5 km. Eighty-five percent of the LLP ground strikes occur within the boundaries of the LDAR storm, and 98% occur either inside or within 2km beyond the edge of the LDAR storm. By definition, 100% of the LDAR-deduced ground strikes occur within the boundaries of the LDAR storm.

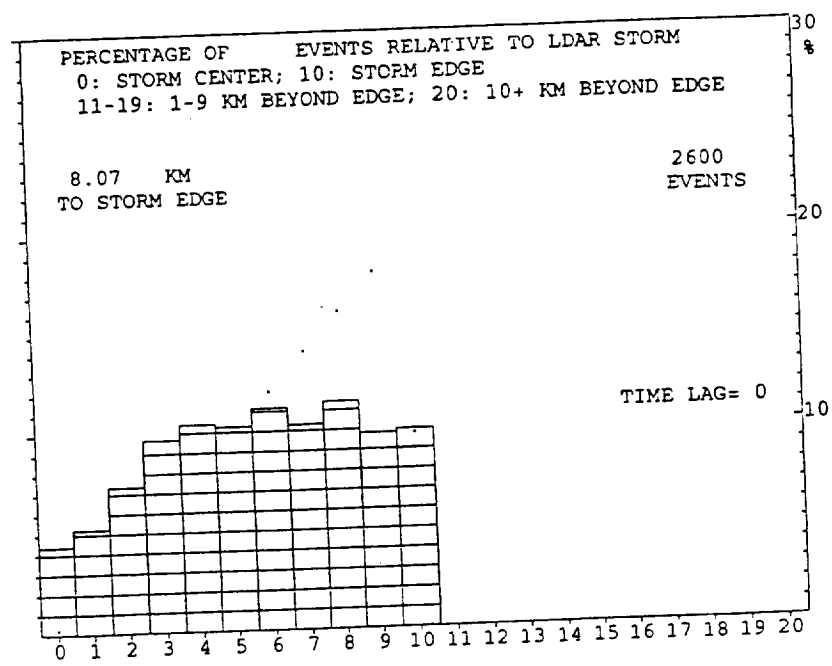
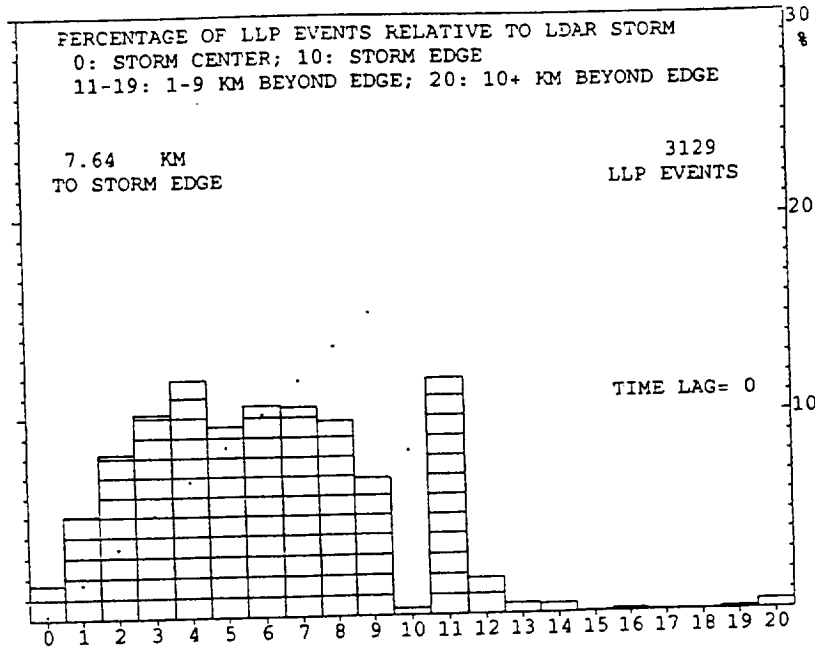


Figure 3-1. Locations of ground strikes relative to LDAR storm center, by fractional distance to edge of LDAR storm, which is at position 10.
 a. (Top) LLP ground strikes.
 b. (Bottom) LDAR-derived ground strikes, using LDAR data cubes below 2.5 km.

Care must be taken in interpreting Fig. 3-1. The fractional categories represent ring-shaped areas concentric about the LDAR storm center. The area within the ring increases with distance from the LDAR storm center. Thus, the area represented by small-numbered fractional categories near the LDAR storm center is small, and a small percentage of the events would be expected to occur there. For this reason, also shown on Fig. 3-1 are dots. The dots represent the percentage of ground strikes expected in the fractional categories if ground strikes were distributed in equal concentration (number per unit area) within the LDAR storm boundaries. Graphed percentages above the dots, then, indicate observed ground strike concentrations greater than average. Graphed percentages below the dots indicate observed ground strike concentrations less than average. It can be seen that the inner half of the composite LDAR storm contained greater ground strike concentrations than the outer portion.

Figure 3-2 shows the locations of ground strikes relative to the LDAR storm during the same minute period, broken down into east/west and north/south positions. Again, the dots represent expected percentages if ground strikes concentration was uniform inside the LDAR storm. The ground strikes are essentially distributed evenly on each side of the LDAR storm, with greater concentrations than average in the inner half of the composite storm.

Figure 3-3 shows the locations of ground strikes relative to the LDAR storm during the same minute period, but composited by distance from the LDAR storm center (using compositing method 1). Here the tendency for ground strikes near the LDAR storm center shows up very strongly.

3.5 SUBSEQUENT GROUND STRIKES RELATIVE TO CURRENT LDAR STORM

Weather forecasters must issue lightning advisories when ground strikes are expected within 5 nautical miles (9.25 km) of a warning site. Because they need to issue the advisories in advance of the first ground strike, it is desirable to examine the relationship of future ground strikes to the current LDAR position. Such intercomparisons have been done with time "lags" of 1, 3, 5, 10, 15, and 20 minutes.

In performing these compositings, it was anticipated that the results might be different for quasi-stationary storms and moving storms. It was also recognized that in the case of moving storms a storm-relative frame of reference would probably be most appropriate. In the latter approach, compositing method 3 of section 3.4, the movement of the LDAR storm was considered, so that the ground strike was composited with respect to the LDAR storm displaced to its future location at the time of the ground strike. The LDAR storm was presumed to remain the same size and shape during the period. In actuality, the LDAR storm-relative position of the ground strikes was obtained by displacing the ground strike points by a vector $-C dt$, where C is the movement vector of the LDAR storm and dt is the "time lag" (ground strike time minus current LDAR storm display time).

Table 3-2 shows the results of the time-dependent compositings for LLP ground strikes with respect to LDAR storms. For reference to the above discussion, time 0 is shown, representing intercomparison of ground strikes to LDAR storm within the same minute. The columns affiliated with "All Storms" represent an intercomparison of all storms, and the storms are presumed to remain at their current positions in their current shapes and sizes. In the columns affiliated with "Quasi-stationary" storms, only storms whose speed could be computed and with speed of movement components less than 2.5 kts (4.6 km/h) in both the east/west and north/south directions have been composited. As in the case of "All Storms", compositing has been done with respect to the current position of the LDAR storm. There were 435 storm-minutes of this type at time 0, decreasing to 324 by time = 20 minutes (an artifact of only using hourly files). In the columns affiliated with "Moving" storms, only LDAR storms having speed of movement greater than 2.5 kts (4.6 km/h) and less than 49 kts (90.7 km/h) were composited. The setting of an upper limit discarded a few storms where the computed LDAR storm velocity was

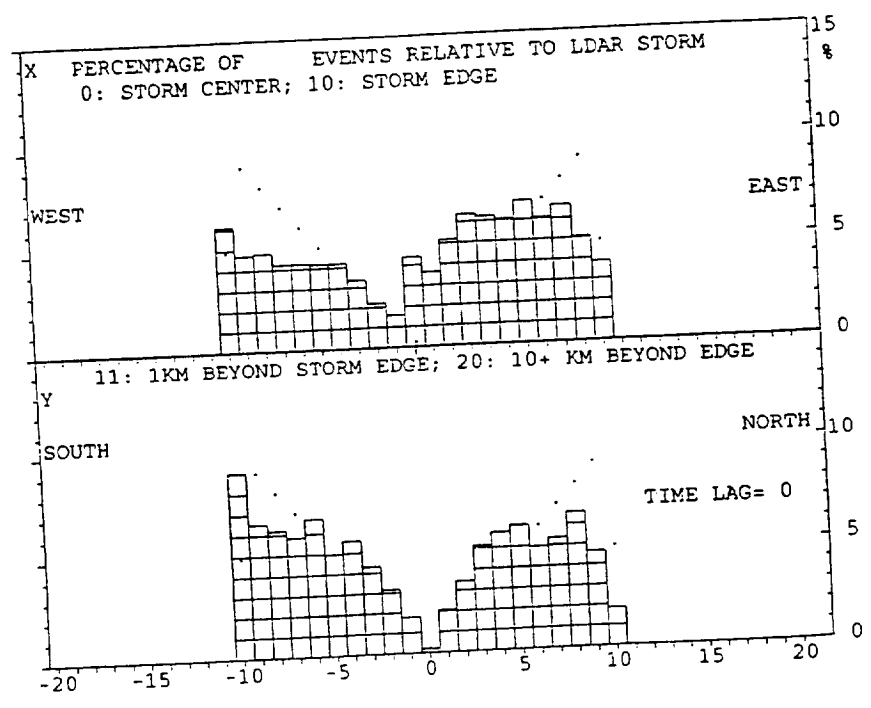
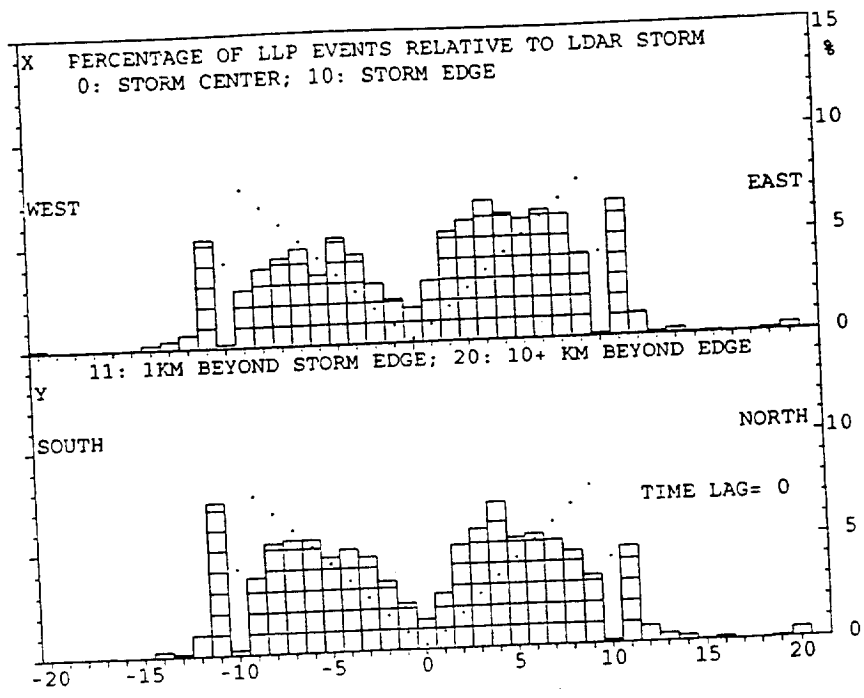


Figure 3-2. Locations of ground strikes relative to LDAR storm center, by fractional positions east/west and north/south. Edge of LDAR storm is at position 10. a. (Top) LLP ground strikes. b. (Bottom) LDAR-derived ground strikes, using LDAR data cubes below 2.5 km.

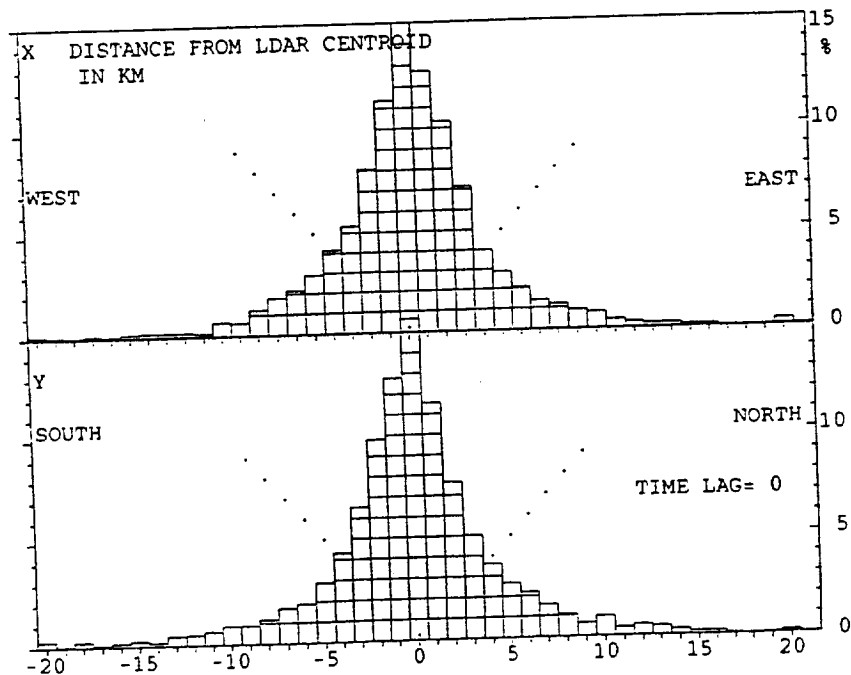
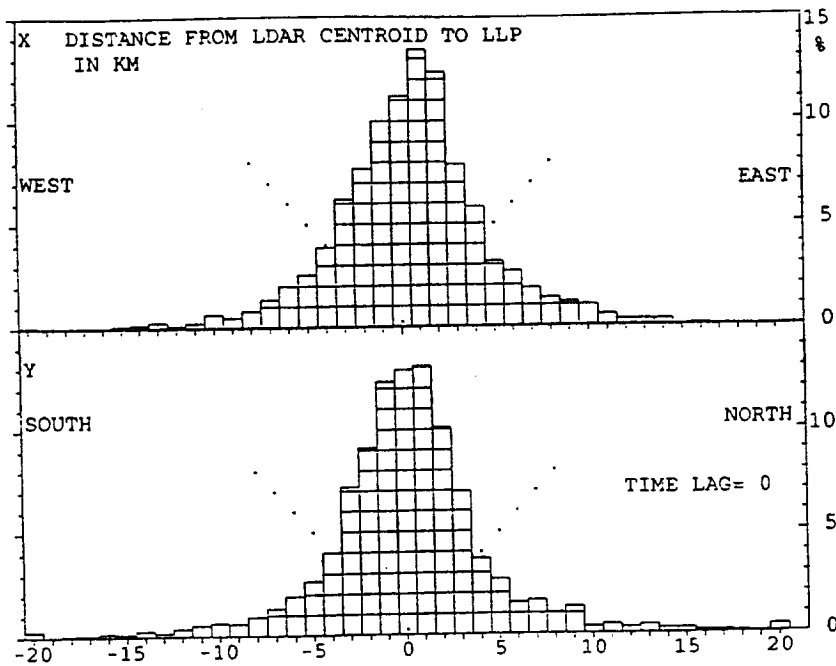


Figure 3-3. Locations of ground strikes relative to LDAR storm center, by distance east/west and north south. a. (Top) LLP ground strikes. b. (Bottom) LDAR-derived ground strikes, using LDAR data cubes below 2.5 km.

probably spurious. There were 2483 moving LDAR storm-minutes at time 0, decreasing to 1553 by time = 20 minutes. Compositing is in an LDAR storm-relative frame of reference for "Moving" storms.

The values in Table 3-2 indicate that there is about a 1% increase in percentages at time 0 for quasi-stationary and moving storms. This is because a few storms have been discarded for which no LDAR storm velocity could be determined. These tended to be short-lived storms, probably of dubious validity.

TABLE 3-2. COMPOSITING OF FUTURE GROUND STRIKES WITH RESPECT TO THE CURRENT LDAR STORM BOUNDARIES, BY CATEGORIES OF STORM MOVEMENT, IN PERCENT

TIME	ALL STORMS			QUASI-STAT.			MOVING		
	in	<1	<2	in	<1	<2	in	<1	<2
0	85	97	98	86	98	99	86	97	99
1	75	86	90	80	89	93	77	87	90
3	71	84	88	79	88	93	74	84	89
5	64	77	84	71	81	86	67	79	84
10	44	59	67	51	65	68	50	61	67
15	31	42	50	41	54	63	36	47	53
20	20	29	35	30	41	50	26	36	39

Notes: "in" means inside LDAR storm boundaries
 <1 means within 1 km beyond LDAR storm edge
 <2 means within 2 km beyond LDAR storm edge

The values in Table 3-2 seem to indicate the merit of using compiling according to LDAR storm movement. Percentages under "All Storms" are with respect to the initial position of the LDAR storm while, in reality, more than 75% of the storms were moving. Partly for this reason, the percentage of future ground strikes falling inside the current LDAR storm position decreases to 20% by 20 minutes. For storms that truly were quasi-stationary, the percentage is 30% at 20 minutes, a 10% "improvement", and is 50% versus 35% for ground strikes within 2 km beyond LDAR storm edge. Increases of up to 6% are seen in the percentage of ground strikes inside the LDAR storm due to use of the LDAR storm-relative frame of reference with moving storms. The increase is smaller for ground strikes within 2 km beyond storm edge.

One way to summarize Table 3-2 is to note that more than half of the ground strikes occur outside the boundaries of the current LDAR storm after 10 minutes. Half of the ground strikes occur more than 2 km beyond the boundaries of the current LDAR storm after 20 minutes for quasi-stationary storms, and after 15 minutes for moving storms.

The probable explanation of the increase in percentage of ground strikes occurring outside the LDAR storm with time is that thunderstorms and their lightning ground strike patterns evolve on rather short time scales. Thunderstorms and their lightning patterns tend to grow with time, and that growth has not been anticipated or factored into the compositings of Table 3-2. Some of the "external" ground strikes of Table 3-2 are also due to the development of new storms along the flanks of the existing storm.

The data base can be used to determine if future ground strikes occur preferentially on one flank of the the LDAR storm. Figure 3-4 shows the composite locations of ground strikes after 10 minutes, with respect to quasi-stationary storms (Fig. 3-4a) and with respect to moving storms in a storm-relative frame of reference (Fig. 3-4b). There is a slight preference for future ground strikes on the south and east flanks of the current quasi-stationary LDAR storms. There is a slight preference for future ground strikes on the rear flank of moving LDAR storms. In the mean the rear flank was the west-southwest flank in this study, though there was much variation in direction of storm movement.

3.6 PROBABILITY OF GROUND STRIKES MORE THAN 5 NAUTICAL MILES BEYOND LDAR STORM EDGE

Weather forecasters must issue lightning advisories when lightning is anticipated within 5 nautical miles (9.25 km) of a warning area. Table 3-3 shows the frequency of LLP ground strikes that occur more than 5 n.mi. beyond the edge of the LDAR storm boundary. In the case of moving storms at future times, the position is relative to the projected position of the current LDAR storm boundaries. Probabilities for "All storms" and "Quasi-Stationary" storms are with respect to the current positions of the LDAR storm boundaries.

TABLE 3-3. PROBABILITY OF GROUND STRIKES MORE THAN 5 N.MI. BEYOND EDGE OF LDAR STORM

<u>TIME</u>	<u>ALL STORMS</u>	<u>QUASI-STAT.</u>	<u>MOVING</u>
0	0.5%	0.0%	0.4%
1	3.0	2.5	2.7
3	2.6	2.3	2.5
5	3.4	4.0	3.7
10	7.5	8.5	7.3
15	14.8	7.9	15.5
20	30.0	16.1	30.8

The probabilities shown in Table 3-3 are not point probabilities, but apply to large areas. A conservative estimate of the probability that a particular area of size 1 km² will be struck can be obtained from the values in Table 3-3 by dividing by 61.3. This factor has been derived by using a mean LDAR storm edge position 7.64 km from the LDAR storm center, and very conservatively assuming that all ground strikes occurred between 9.25 and 10.25 km beyond LDAR storm edge. Thus, the probability of a particular square kilometer area being hit when beyond 5 n.mi. from the LDAR storm edge is less than 0.008% during the current minute and not more than 0.5% after 20 minutes. Depending upon what risk is acceptable, forecasters can probably make effective use of the current LDAR storm boundary (and not need to add an extra margin beyond it) in issuing advisories for lightning within a 5 n.mi. radius of warning sites.

3.7 CESSATION OF LDAR ACTIVITY AND END OF GROUND STRIKE THREAT

Weather forecasters must issue statements indicating that the threat of ground strikes has ceased. One potential tool in this task is the disappearance of the LDAR storm. However, some LDAR storms are intermittent--especially in their late stages--and may reappear after being inactive for a few minutes. Do these storms still pose a ground strike threat? Table 3-4 indicates the probability of an LDAR storm still yielding a ground strike after being inactive (producing no LDAR events) for the indicated number of minutes.

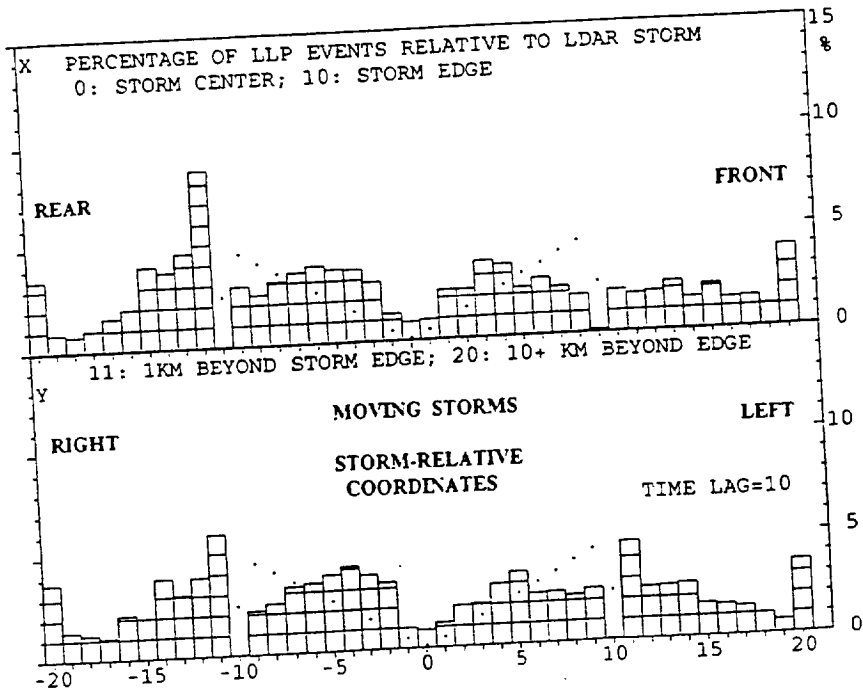
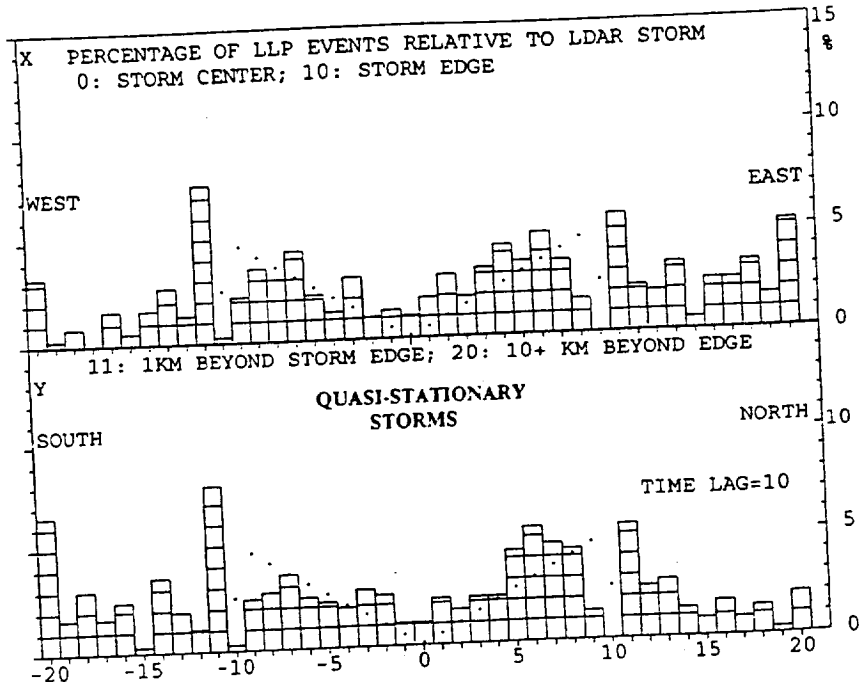


Figure 3-4. Locations of ground strikes 10 minutes later, relative to current LDAR storm center, by fractional positions east/west and north/south.
 a. (TOP) Quasi-stationary storms; compositing relative to current position.
 b. (Bottom) Moving storms; compositing relative to projected position of current LDAR storm after 10 minutes, in storm-relative coordinates.

**TABLE 3.4 PROBABILITY OF AN LDAR STORM THAT HAS BEEN INACTIVE
YIELDING A FUTURE GROUND STRIKE,
AS A FUNCTION OF THE PERIOD OF INACTIVITY**

<u>INACTIVE PERIOD</u>	<u>PROBABILITY</u>
1 min	46%
> 1	22
> 2	13
> 3	7
> 4	5
> 5	3
> 6	2
> 7	1
> 8	0.5

3.8 DIFFERENCES BETWEEN STORMS THAT PRODUCE GROUND STRIKES AND THOSE THAT NEVER PRODUCE GROUND STRIKES

LDAR storms that never produced ground strikes had short duration. Eighty-nine percent had duration of 5 minutes or less. Only 27% of the LDAR storms which produced LLP ground strikes had durations that brief. Many (55%) of the LDAR storms which never produced ground strikes were 1-minute events, representing odd discharge patterns into areas not previously affected. This caused an appreciable (> 6 km), momentary shift in storm centroid and caused the objective storm classification scheme to trigger a new storm identification number. During the next minute, when the odd discharge was no longer present, the LDAR storm centroid returned to near its previous location, and the previous storm identification number was resumed.

Only 10% of the LDAR storms having 1 or 2-minute duration were associated with LLP ground strikes. Only 6% of the LDAR storms having 1 or 2-minute duration were associated with LLP ground strikes if they contained less than 100 LDAR events per minute. From a weather forecasting perspective, the implication is that sudden large bursts of low-density LDAR activity that extend beyond the previous LDAR storm boundary do not normally pose a ground strike threat.

The maximum number of LDAR events is less in LDAR storms that never produce an LLP ground strike. Sixty-two percent of such storms never contained more than 50 LDAR events per minute. Only 11% of the LDAR storms associated at some time with LLP ground strikes peaked at such a low LDAR data rate. By contrast, 81% of the LDAR storms associated with LLP ground strikes contained more than 150 LDAR events per minute at some point during their existence.

3.9 LDAR STORM DIFFERENCES DURING MINUTES WITH AND WITHOUT GROUND STRIKES

It was indicated previously that only 41% of the LDAR storm minutes are associated with LLP ground strikes. Is the LDAR storm character different during minutes with LLP ground strikes? For brevity of wording, LDAR storm minutes associated with LLP ground strikes will be referred to below as "LLP minutes"; those without ground strikes will be referred to as "non-LLP minutes".

LLP minutes had many more LDAR events per storm. The median values were 500 for LLP minutes and 165 for non-LLP minutes. Of course, some of these LDAR events could be produced by the ground strike itself.

LLP minutes had a much bigger LDAR storm volume. The median values were 200 km³ for LLP minutes and 50 km³ for non-LLP minutes. Again, some of this difference could be due to an expansion of LDAR storm volume due to the ground strike extending the storm volume downward to the surface. However, this hypothesis is

not fully supported by LDAR storm height data. There was little difference in the median heights of the LDAR storm centroid during LLP and non-LLP minutes. Likewise, there was little difference in the median z95 heights, suggesting that LDAR storm top was not significantly different (at least with the vertical resolution used in this study). LLP minutes, however, did have larger storm depth, as measured by standard deviation of z. Using twice the standard deviation, median values were 4.15 versus 2.75 km for LLP and non-LLP minutes, respectively. Some downward extension of the LDAR storm may be revealed here.

LLP minutes had larger LDAR storm cross-sectional area. The median values were 80 km² for LLP minutes and 36 km² for non-LLP minutes.

Despite having larger LDAR storm volume and cross-sectional areas, LLP-minutes also had larger LDAR volume densities and LDAR area densities than non-LLP minutes. Median values for LDAR volume density were 3.1 and 2.2 LDAR events per km³, respectively. Median values for LDAR area density were 6.6 and 3.7 LDAR events per km², respectively.

3.10 LDAR-BASED PARAMETERS IN RELATION TO VARIATIONS IN GROUND STRIKE RATE

Because section 3.9 showed many differences between LDAR storm characteristics during minutes with and without LLP ground strikes, and because some storms (such as in Fig. 2-7) revealed that LDAR-derived parameters were correlated to number of LLP ground strikes, the correlation between LDAR-derived storm parameters and ground strike rate was examined in more detail. Temporal variations of eight LDAR-derived parameters were examined in relation to temporal variations in numbers of LLP ground strikes associated with the LDAR storms. LDAR storms were only included in these correlation studies if they were associated with LLP ground strikes at some time during their existence. The eight LDAR-derived storm parameters examined were: (1) NLDAR, number of LDAR data points within the LDAR storm volume; (2) VOL, LDAR storm volume; (3) AREA, LDAR storm area; (4) VOLDENS, LDAR volume density; (5) AREADENS, LDAR area density; (6) ZMEAN, LDAR storm centroid height; (7) ZSD, standard deviation of height of LDAR cubes within the LDAR storm volume, a measure of LDAR storm depth; (8) Z95, the height below which 95% of the LDAR data occurs within the LDAR storm.

Table 3-5 shows correlations between number of LLP ground strikes per minute associated with the LDAR storm and LDAR storm parameter values. The correlations are tabulated for simultaneous relationships (time lag 0), and for LLP ground strikes which lag the LDAR storm parameter values by the indicated numbers of minutes. The five LDAR-derived parameters with highest correlations are listed for each time lag, in order of decreasing correlation.

In the correlation studies, ZSD, a measure of LDAR storm depth, had the highest correlation for 9 of the 11 lag times. LDAR storm volume had the highest correlation for the other two lag times, 1 and 5 minutes. The number of LDAR data points within the LDAR storm had the third best correlation overall, having the second highest correlation for 3 lag times. Mean LDAR storm height ranked second for 2 lag times. Z95 and LDAR storm area each had third highest correlation for some lag time. Despite impressive relationships for some storms, LDAR volume and area densities never ranked better than sixth for any lag time when all LDAR storms that produced LLP ground strikes were considered.

Unfortunately, the correlations in Table 3-5 were highest for time lag of zero and generally decreased with increasing lag time. The exception was for lag time of two minutes, when the correlation was higher than at 1 minute. Probably the contribution of LDAR data points to the LDAR storm characteristics by the ground strike itself is causing the highest correlation for lag time of zero.

It must also be noted that even the highest correlation by an individual LDAR storm parameter is 0.592. This means that variations in LDAR storm depth from minute to minute would explain only 35% of the variations in number of LLP ground

TABLE 3-5. CORRELATION BETWEEN LDAR-DERIVED STORM PARAMETERS AND LLP GROUND STRIKE RATE, AS A FUNCTION OF LAG TIME BETWEEN LDAR PARAMETER AND LLP GROUND STRIKE

<u>LAG</u>	<u>RANK 1</u>	<u>RANK 2</u>	<u>RANK 3</u>	<u>RANK 4</u>	<u>RANK 5</u>
0	ZSD 0.592	VOL 0.579	AREA 0.542	NLDAR 0.521	ZMEAN 0.495
1	VOL 0.492	ZSD 0.492	NLDAR 0.457	AREA 0.441	ZMEAN 0.422
2	ZSD 0.497	VOL 0.486	NLDAR 0.460	AREA 0.441	ZMEAN 0.439
3	ZSD 0.478	VOL 0.466	NLDAR 0.446	ZMEAN 0.433	AREA 0.424
4	ZSD 0.450	VOL 0.433	ZMEAN 0.415	NLDAR 0.414	AREA 0.396
5	VOL 0.433	ZSD 0.429	NLDAR 0.422	ZMEAN 0.402	AREA 0.392
6	ZSD 0.396	NLDAR 0.384	VOL 0.383	ZMEAN 0.377	Z95 0.359
7	ZSD 0.376	ZMEAN 0.361	Z95 0.345	VOL 0.344	NLDAR 0.344
8	ZSD 0.355	NLDAR 0.348	VOL 0.347	ZMEAN 0.340	Z95 0.325
9	ZSD 0.337	NLDAR 0.323	VOL 0.321	ZMEAN 0.315	Z95 0.303
10	ZSD 0.300	ZMEAN 0.291	Z95 0.278	NLDAR 0.265	VOL 0.259

NOTE: LAG is LLP ground strike time minus LDAR storm parameter time, in minutes

strikes per minute using a regression approach. However, experience with multiple linear regression shows that this would be a good start if additional LDAR storm parameters were not highly intercorrelated. If the temporal variations of other LDAR storm parameters were poorly correlated to ZSD variations, then they would make additional contributions in a multiple regression equation aimed at explaining the variations in number of LLP ground strikes per minute from LDAR storms. Time has not permitted further pursuit of this topic.

IV. CONCLUDING REMARKS

The ultimate goal of this study has been to begin a process that may help operational weather forecasters give better and more timely lightning advisories. LDAR seems to provide information that can be useful to the forecasters.

In section I, five questions were posed that operational weather forecasters might ask regarding use of the LDAR data. The answers to those questions are reviewed here.

What is the lead time between the first detection of LDAR events and ground strikes? The average lead time is 5.26 minutes. Only 11% of the LDAR storms produced LLP ground strikes during the first minute of their existence (Section 3.1).

Are there preferred locations of occurrence of ground strikes relative to the pattern of discharges detected by LDAR? Eighty-five percent of the LLP ground strikes occur within the boundaries of the LDAR-defined storm, and 98% occur either within the LDAR storm boundaries or within 2 km beyond them. The highest concentration of ground strikes occurs near the center of the LDAR storm (Section 3.4).

Can future positions of ground strikes be anticipated through knowledge of the current LDAR event pattern? The percentage of ground strikes falling within the boundaries of the current LDAR storms decreases with time. Half of the ground strikes occur more than 2 km beyond the boundaries of the current LDAR storm (in the case of quasi-stationary storms) or the projected position of the current LDAR storm (in the case of moving storms) after 15 minutes for moving storms and after 20 minutes for quasi-stationary storms. There was some preference for future ground strikes to occur on the south and east flanks of quasi-stationary storms and on the rear (west-southwest) flank of moving storms in this study (Sections 3.5, 3.6).

Are there signatures in the LDAR data that can be used to determine when the threat of ground strikes has ended? In latter stages of LDAR storms, they often become intermittent, often disappearing for several minutes on the LDAR display. Only 5% of LDAR storms have future ground strike if they have disappeared for more than 4 minutes; only 0.5% if they have disappeared for more than 8 minutes (Section 3.7). Other signatures probably exist, but have not been quantified.

Are there signatures in the LDAR data that reveal which storms produce ground strikes and which do not? Sudden large bursts of low-density LDAR activity that extend beyond the previous LDAR storm boundary do not normally pose a ground strike threat. There are a number of LDAR-derived parameters that show significant differences, on average, between storms that produce ground strikes and those that do not (Sections 3.8-3.10). Further research is needed before these parameters can be used reliably.

LDAR shows great promise as a tool for better understanding the dynamics of thunderstorm electrification processes, especially when the LDAR data is used in conjunction with Doppler radar, surface mesonet, and wind profiler network data that can reveal the location of convergence zones. These convergence zones, when related to the flanks of LDAR- and radar-defined storms, can suggest the preferred flanks upon which new thunderstorm cell growth will be preferred and where future ground strikes may be concentrated. Intercomparison of field mill data should also prove informative. Expert system and neural network studies of lightning prediction may be ways to assimilate the various types of information.

REFERENCES

1. Maier, M.W., A.G. Boulanger, and R.I. Sax: An initial assessment of flash density and peak current characteristics of lightning flashes to ground in South Florida. NUREG/CR-1024, U.S. Nuclear Regulatory Commission, 1979.
2. MacGorman, D.R., M.W. Maier, W.D. Rust: Lightning strike density for the contiguous United States from thunderstorm duration records. NUREG/CR-3759, U.S. Nuclear Regulatory Commission, 1984.
3. Maier, L.M., E.P. Krider, and M.W. Maier: Average diurnal variation of summer lightning over the Florida Peninsula. *Mon. Weather Rev.*, 112(6), June, 1984, pp. 1134-40.
4. Orville, R.E., R.W. Henderson, and L.F. Bosart: An East Coast lightning detection network. *Bull. Amer. Meteorological Soc.*, 64, 1983, pp. 1029-37.
5. Krider, E.P., A.E. Pifer and D.L. Vance: Lightning direction-finding system for forest fire detection. *Bull. Amer. Meteorological Soc.*, 61, 1980, pp. 980-86.
6. Lennon, C. and L. Maier: Lightning mapping system. NASA CP-3106, Vol. II, 1991 International Aerospace and Ground Conference on Lightning and Static Electricity, 1991, pp. 89-1 to 89-10.

1994018766
N94-28230

1993 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA

59-33
197195
p. 33

REDUCED COST ALTERNATIVES TO PREMISE WIRING
USING ATM & MICROCELLULAR TECHNOLOGIES

PREPARED BY:	Dr. Raghvendra R. Gejji
ACADEMIC RANK:	Assistant Professor
UNIVERSITY AND DEPARTMENT:	Western Michigan University Department of Electrical Engineering
NASA/KSC	
DIVISION:	Electronic Systems
BRANCH:	Communication
NASA COLLEAGUE:	John Schnitzius
DATE:	July 16, 1993
CONTRACT NUMBER:	University of Central Florida NASA-NGT-60002 Supplement: 11

ACKNOWLEDGMENTS

Valuable ideas for this project were contributed by:

John Schnitzius - NASA colleague
Jerry Barnes, DL-ESS-13
Charles Arnett, INI-18
Dave Miller, INI-18
Gregg Atkins, INI-18
Armistead Fredlock, TE-CID-2A
Jeanne O'Bryan, CS-GSD-21B
Denise Rousseau, TE-CID-2B
Scott Thurston, TV-ETD-11
Fred McKenzie, DL-ESS-11

For organizing and supporting my activities under NASA-ASEE-SFRFP, thanks are due to:

E. Ramon Hosler, Director
Kari Stiles, Administrator
Tom Barron, Program Director
Jerry Barnes, Section Chief
Perry Rogers, Branch Chief

For assistance in learning the equipment and operation, a debt of gratitude is due to:

Don McCoy, INI-18
Felix Soto, DL-ESS-13

ABSTRACT

The cost of premises wiring keeps increasing due to personnel moves, new equipment, capacity upgrades etc. It would be desirable to have a wireless interface from the workstations to the fixed network, so as to minimize the wiring changes needed. New technologies such as microcellular personal communication systems are promising to bring down the cost of wireless communication. Another promising technology is Code Division Multiple Access (CDMA), which could dramatically increase the bandwidth available for wireless connections. In addition, Asynchronous Transfer Mode (ATM) technology is emerging as a technique for integrated management of voice, data and video traffic on a single network. The focus of this investigation will be to assess the future utility of these new technologies for reducing the premise wiring cost at KSC. One of the issues to be studied is the cost comparison of "old" versus "new", especially as time and technology progress. An additional issue for closer study is a feasible time-line for progress in technological capability.

SUMMARY

New technologies such as personal communication systems (PCS) and asynchronous transfer mode (ATM) hold the promise to usher in a new era of mobile multimedia communication. This could have significant impact at KSC in operations and in premise wiring cost. Switching networks have evolved from circuit switching to TDMA to packet switching. ATM is the next step in the evolution of switching systems, with the promise of carrying multimedia service on a single network. Cellular communication service is evolving from an analog FM system to a digital system based on TDMA. CDMA technology has also been proposed as an alternative for digital cellular. At present, the cellular business is driven by the voice market. But the lessons learned from the switching business indicate that cellular networks will evolve toward packet based technology e.g. ALOHA, for adapting to multimedia service based on ATM backbone network. The technology for building a multimedia terminal is already here. The challenge for industry is to design and standardize a packet based cellular network that can act as a wireless front end to the backbone ATM network. This may take about 8 years yet.

TABLE OF CONTENTS

<u>Section</u>	<u>Title</u>
I.	INTRODUCTION
1.1	THE PROMISE OF PCS TECHNOLOGY
1.2	THE PROMISE OF ATM TECHNOLOGY
1.3	KSC APPLICATIONS POTENTIAL
II.	ATM AND PCS TECHNOLOGY BACKGROUND
2.1	DEVELOPMENT BACKGROUND OF ATM TECHNOLOGY
2.2	ISSUES AND OPTIONS IN MOBILE COMMUNICATION TECHNOLOGY
2.2.1	DIGITAL CELLULAR
2.2.2	POWER-BANDWIDTH TRADEOFF IN DIGITAL CELLULAR
2.2.3	QPSK, CODING AND TDMA
2.2.4	CDMA, POWER CONTROL, VOICE ACTIVATION, CELL SECTORIZATION
2.2.5	ALOHA AND PRMA
III.	DISCUSSION
3.1	ISSUES IN MOBILE MULTIMEDIA
3.1.1	GRANULARITY, BANDWIDTH AND CAPACITY
3.1.2	ACCESS METHOD AND ATM ADAPTATION
3.1.3	COST
3.1.4	TECHNOLOGY DEVELOPMENT TIMELINE
3.1.5	RF SPECTRUM USAGE AT KSC
IV.	CONCLUSIONS
V.	REFERENCES
	FIGURES

LIST OF ILLUSTRATIONS

Figure Title

1. Promise of PCS technology.
2. Promise of ATM technology.
3. KSC applications potential.
4. Circuit switching.
5. TDMA switching.
6. Packet switching.
7. Cellular system concepts.
8. Example of 3- and 4-channel reuse plans.
9. Cellular system performance.
10. Capacity enhancing techniques for digital cellular.
11. Comparative features of TDMA and CDMA.

LIST OF ABBREVIATIONS

ADPCM	--	Adaptive differential pulse code modulation
AMPS	--	Advanced mobile phone system
ATM	--	Asynchronous transfer mode
CDMA	--	Code division multiple access
CIR	--	Carrier to interference ratio
CRC	--	Cyclic redundancy check
CSMA	--	Carrier sense multiple access
DCT	--	Discrete cosine transform
DS	--	Direct Sequence
FDDI	--	Fiber distributed data interface
FDMA	--	Frequency division multiple access
FEC	--	Forward error correction
FH	--	Frequency hopping
FM	--	Frequency modulation
GMSK	--	Gaussian minimum shift keying
GSM	--	Groupe spatiale mobile
KSC	--	Kennedy space center
LAN	--	Local area network
MTSO	--	Mobile telephone switching office
NTSC	--	National telecommunication standards committee
OISD	--	Operational intercom system digital
OMI	--	Operation and maintenance instruction
OTV	--	Operational television
PCS	--	Personal communication system
PN	--	Pseudo noise
PRMA	--	Packet reservation multiple access
PRN	--	Packet radio network
QPSK	--	Quaternary phase shift keying
RF	--	Radio frequency
SPDMS	--	Shuttle processing data management system
SSMA	--	Spread spectrum multiple access
TDMA	--	Time division multiple access

I. INTRODUCTION

The cost of premises wiring keeps increasing due to personnel moves, new equipment, capacity upgrades etc. It would be desirable to have a wireless interface from the workstations to the fixed network, so as to minimize the wiring changes needed. New technologies such as microcellular personal communication systems are promising to bring down the cost of wireless communication. Another promising technology is Code Division Multiple Access (CDMA), which could dramatically increase the bandwidth available for wireless connections. In addition, Asynchronous Transfer Mode (ATM) technology is emerging as a technique for integrated management of voice, data and video traffic on a single network. The focus of this investigation will be to assess the future utility of these new technologies for reducing the premise wiring cost at KSC. One of the issues to be studied is the cost comparison of "old" versus "new", especially as time and technology progress. An additional issue for closer study is a feasible time-line for progress in technological capability.

In Section 1 we will outline the motivation for study of mobile multimedia technology at KSC. In Section 2 we will outline the current status, issues and trends in PCS and ATM technologies, especially as they relate to scenarios for development of mobile multimedia technology. Section 3 will assess the technological options of Sec. 2 against the requirements posed by the applications in Sec. 1. A reader wishing to get an overview without too much detail may wish to skip Sec. 2 and proceed directly to Sec. 3. Final conclusions are presented in Sec. 4.

1.1 THE PROMISE OF PCS TECHNOLOGY

Personal communication system (PCS) technology is a predicted outgrowth of current generation cellular communication systems. In the 1970s, mobile telephone service was expensive and available to the rare few. By the mid 80s, cellular technology made it possible to expand service to a large number of users in automobiles. By 1990, personal communications were emerging in the form of lightweight mobile telephones that could be used by automobile passengers as well as pedestrians. Yet, further improvements in mobile communications are needed. Present technology is mainly geared for voice communications. There is a need to develop data and fax services in the near future, and video services further down the road. Another critical issue is improving the management of bandwidth in order to allow closer to a universal service penetration. These issues are under active investigation [1]. Indoor mobile communication needs further development, in order to allow tether-less access in the home and office environments.

When fully developed, PCS technology promises to deliver an array of communication services to a portable hand-held unit at any desired location. Fig. 1 summarizes some of the features of future PCS technology. Visions of Dick Tracy and Star Trek will be fully realized. The PCS terminal will be lightweight, operable indoors and outdoors, so the same terminal will allow completely tetherless and mobile access to the network from home, office, automobile, or even out of town. Global roaming will allow the same terminal to be usable anywhere. Just imagine, your teenage daughter can start talking with her friend as soon as she wakes up, and need not

hang up until she is ready to go to bed at night. Custom network services will be delivered on the basis of user identity, not terminal identity or location. This means custom services such as speed dialing, custom directory, call screening, billing authorizations, etc., will travel with you when you go to a different city. Each person will have only one telephone number, whether they are at home, office, automobile, hotel or fishing. The future PCS terminal will be more than just a voice telephone. It will offer voice, video, and data services all in a single terminal. This means that in addition to talking, you can use the PCS terminal for accessing your computer, or for that matter, any authorized computer, receive fax and e-mail messages, make video phone calls, and access video libraries. This issue is further discussed in the next subsection.

1.2 THE PROMISE OF ATM TECHNOLOGY

Asynchronous transfer mode (ATM) technology is a packet switching technology. The purpose of all switching techniques is to route data streams to specified network addresses. In traditional circuit switching, e.g. in the telephone network, the address, i.e. the dialed digits, are used to set up a dedicated wireline connection for that particular call. The message, i.e. the conversation, is carried on this dedicated wireline connection. In packet switching, the message bits and address are combined into a packet, and forwarded from node to node. Although packet switching has the overhead of including the address in every packet, it has been found to be more efficient than circuit switching for bursty data, such as is generated by a computer terminal. Circuit switching, on the other hand, is more suited for fixed rate of data generation, such as voice.

ATM technology [2],[3], has been proposed as the technology of choice for mixed data types. Fig. 2 summarizes the potential benefits of ATM technology. ATM is a packet switching technology with fixed-size packets of 53 bytes, called cells. The adoption of ATM technology promises to allow mixed data such as voice and video to be carried over the same network, instead of on separate telephone and LAN networks, as is done today. This will allow the deployment of multimedia terminals, incorporating voice, video, and data, to communicate with each other over a single switched network. By and large, video today is carried over non-switched broadcast networks such as TV stations, and cable TV. The possibility of switched video and switched multimedia service will allow services such as dial-up personal videophone, dial-up videoconference to any location, dial-up video libraries, and electronic town hall.

1.3 KSC APPLICATIONS POTENTIAL

Let us consider the potential impact of combining PCS and ATM technologies. As we noted above, PCS technology is targeted for universal access to mobile communications, and ATM technology is targeted for universal multimedia communications. Together, these two technologies will usher in a new era of mobile multimedia communications. The concept of mobile multimedia communications will make possible bold new services and applications not yet conceived. Some of the ways this could impact operations at KSC are considered (See Fig. 3).

One possible application is described by the title of the project. The possibility of universal

wireless access to a switched ATM network will allow the deployment of tetherless multimedia workstations. In particular, intercomputer communication will incorporate a wireless front end. This means that a particular workstation may be placed anywhere, and will receive the appropriate level of bandwidth service, regardless of location. Cable plant changes would not be required following personnel moves. With today's technology, the hardwired cable network is designed to handle a predicted pattern of bandwidth demand. If the demand pattern changes following personnel moves, say for example if a CAD engineer interchanges office locations with a secretary, the cable plant has to be rewired to meet the new location of high demand. At the same time, the bandwidth resource previously deployed for a high demand location, will be wasted when the demand at that location goes down.

Availability of wireless multimedia can significantly impact the information flow required for shuttle and payload operations. Currently the information is stored in the form of OMI's on a database system such as SPDMS. At the site of the operations, the information is carried on paper or by voice on OISD. The availability of wireless multimedia, will allow video instructions for operations, fast wireless scheduling updates, and videoconference among operation personnel for problem resolution. With two-way wireless support for multimedia service, it becomes possible to deliver rapid video verification of operation closeouts, as well as live video monitoring of critical operations by remotely located expert personnel, e.g. firing room.

II. ATM AND PCS TECHNOLOGY BACKGROUND

In this section we will review some background material on the development of ATM and PCS technology, emphasizing some key features that will be important in their application to mobile multimedia technology. In the next section, we will evaluate the various technological options against the criteria posed by the applications outlined in Sec. 1. Section 2 may be skipped by readers not interested in many of the details.

2.1 DEVELOPMENT BACKGROUND OF ATM TECHNOLOGY

Traditionally, the telephone network has been circuit switched. For our purpose, what this means is that the destination address, is first presented to the network in the form of dialed digits. The network processes this information in a centrally located intelligence unit to set up an end-to-end link between source and destination. The message, i.e. voice signal, is then applied to this dedicated link and appears at the desired destination by virtue of circuit connectivity. This principle is summarized in Fig. 4.

In addition to so-called space switching described above, modern telephone networks also employ time-switching. Time switching depends on time division multiplexing (TDMA). In the context of the preceding discussion, we can visualize a TDMA network as shown in Fig. 5. Here, the address is processed by a multiplexor and a demultiplexor located at either end of the link. The multiplexor takes a number of message streams and places them on the link in time sequence as dictated by the address. Thus the address is translated into a fixed position in the time sequence for the corresponding message stream. The important feature to note is that the processing of the address is now distributed between two intelligence units (multiplexor and

demultiplexor) on either end of the link. We see that as technology progresses, the trend is away from a central intelligence unit to a distribution of the intelligence and processing tasks.

Next, packet switching is considered. The conceptual operation of a packet switching network is shown in Fig. 6. In packet switching, the address and message are bundled together to form a packet, much as the message contained in a letter is bundled in an envelope with the address on it. The packet is forwarded from node to node in the network. At each node, the address is examined and a decision is made about which should be the next link for forwarding the message toward its ultimate destination. Thus, packet switching is an example of distributed processing, whereby the intelligence is distributed throughout the network.

The various technological options are thus seen as methods for managing access to the communication link. At a more fundamental level, it is the bandwidth of the link that the users are really demanding. The fundamental parameters of network design are the physical deployment of links, their bandwidth capacities, and the network intelligence to manage access to the links in a way that is fair to everyone. In traditional circuit switching, the physical access to the link is closely controlled by the network intelligence. But once the physical connection for link access is set up, communication can proceed with little help from the network. On the other hand, in ethernet, an example of packet switching, all nodes have physical access to the link, and a great deal of processing is required in each and every node to ensure orderly usage of the link, and delivery of the message at only its proper destination. It is in this context of links, physical access and intelligence location that we will set our problem.

It is well-known that circuit switching is efficient for fixed-rate data, such as voice, where the source generates bits at a steady rate over long periods of time. But circuit switching is not efficient for bursty data such as from a computer terminal, which typically produces long periods of idle punctuated by short voluminous bursts of data bits. Packet switching provides more efficient service for bursty data. For this reason, today we have a circuit switched telephone network for voice, but packet switched LAN's such as ethernet and FDDI for computer generated data communication. Meanwhile video or TV signals are carried over broadcast networks such as network TV or cable TV.

The next step in technology evolution comes from a desire to have a single network for all types of signal streams: data, voice and video. We noted above that voice is efficiently carried over circuit switched network, and data over packet switched network. What type of network should be designed for carrying a mix of traffic? This fundamental problem arises because of the wide disparity in the bandwidth demand of the different types of traffic. Data requires much less bandwidth than voice, which in turn requires much less bandwidth than video. A circuit switched network, including TDMA, assigns a fixed amount of bandwidth to all users. This means all users must be assigned as much bandwidth as those with the highest demand, leading to a great deal of wasted bandwidth resource. Circuit switching ends up assigning more than their fair share of bandwidth to users with low bandwidth demands. On the other hand, a packet switched network favors users with high bandwidth demand. Since packet switching is a random access technology, anyone can access the bandwidth resource. The low demand users will quickly release the resource for others to use. Not so the high demand users. Because of their greater needs, having

once gained control of the link, they will retain control for longer periods of time, thus effectively locking out users with low demand.

In the ATM specification, the packet size is fixed as a resolution of the above problem. Keeping the packet size relatively small forces the high bandwidth user to give up the link more often, thus giving a chance for other users to access it. At the same time, too small a packet size reduces efficiency for the high bandwidth users, since the overhead cost of the address in the packet is incurred more frequently. The ATM specified packet size of 53 bytes is a compromise between the needs of users with low and high bandwidth needs.

2.2 ISSUES AND OPTIONS IN MOBILE COMMUNICATION TECHNOLOGY

Cellular phone communication for automobiles has been available for about a decade. A summary of cellular system concepts is given in Fig. 7. The cellular concept is to divide the geographical service area into circular regions called cells. All the mobile users in a particular cell are serviced by a single base station located in the center of the cell. All the base stations are linked through a switched network. The exchange servicing the network connecting the base stations is called the MTSO, or mobile telephone switching office. The links from the base stations to the MTSO can be land lines or high capacity point-to-point dedicated microwave links. The MTSO is responsible for overall supervision and control of the cellular network.

A fundamental concept in cellular system planning is the number of channel sets, usually 4 or 7. All available channels are divided into a fixed number of channel sets. Each cell site or base station is assigned one channel set, such that neighboring cells do not have the same channel set. An example frequency reuse plan with 3 or 4 channel sets is shown in Fig. 8. Since there is only a finite number of channel sets, there will be other cells in the system using the identical channel set. There is a fixed mathematical relationship between the number of channel sets and the smallest distance (as measured in terms of the cell radius) between cell sites using the same channel set. Clearly, the smaller the number of channel sets, the shorter the distance before the same channel set will need to be reused. An important capacity measure in a cellular system is the number of users/cell per MHz of allocated bandwidth. For example let us consider a cellular system with 24 MHz of bandwidth. Each mobile user is assigned a channel or frequency-pair for 2-way communication with his base station. Say the bandwidth required for analog FM voice communication is 30 kHz in each direction, for a total of 60 kHz. Thus there are a total of $24 \text{ Mhz}/60 \text{ kHz} = 400$ distinct channels available. If the number of channel sets is 7, then each cell can support 57 users. The number of users/cell/MHz is 2.4. If the number of channel sets is reduced to 4, then each cell may support 100 users. Thus, fewer channel sets lead to higher service capacity.

Due to frequency reuse, for a user assigned to a given channel, there are other mobiles using the same channel, located in other cells with the same channel set. These mobiles will cause interference with the transmission from the user of interest. The amount of interference is larger if the cochannel users are located closer. If this interference is intolerable, the separation distance between cochannel users must be increased. The only way this can be done is to increase the number of channel sets, with a corresponding reduction in the number of users/cell that can be

supported. This is the basic tradeoff seen in cellular system planning. The number of channel sets is a compromise between the desire to have low interference and the desire to have as many users/cell as possible. This is shown in Fig. 9.

When cellular phones were first introduced, the price was high and the demand low. Only a few large cells were required to service the demand. As prices came down, and the demand grew, it became necessary to introduce more and more smaller cells. This has led to the growth of the so-called microcellular systems technology. Microcellular systems are characterized by a large number of small cells, some as small as a city block. In the next generation PCS technology, the cells may be as small as individual floors in buildings. Each cell will be served by separate base stations. One issue under active investigation is the effect of small cells on hand-offs. Hand-off is the network action by which a mobile using a frequency in one cell, is automatically assigned a new frequency in the adjacent cell, as it moves from one cell to the next.

2.2.1 DIGITAL CELLULAR

Although cellular phone service has seen phenomenal growth in the last few years, there are a number of issues that need resolution. The most important of these is how to increase service penetration among the population, so cellular service begins to approach the universal accessibility envisioned in the PCS scenario. In general, the most obvious ways to increase service capacity in cellular systems are to convince the FCC to allocate more spectrum, or to build more, smaller cells. Initially, cellular service was based on analog FM modulation. Beginning in the 1990's, the industry has started a shift to digital modulation to increase service capacity without getting new frequency allocations from the FCC. Instead of modulating the carrier directly with the analog voice signal, the voice signal is first sampled and digitized, and the resulting bit stream is used to modulate the carrier. Digital modulation can be binary or M-ary. In binary modulation, only two waveforms are transmitted, representing binary 1 or 0. In M-ary modulation, the bits are grouped, and a group of $\log_2 M$ bits are used to select one of M different transmission waveforms. Use of M-ary modulation reduces the channel bandwidth requirement by a factor of $\log_2 M$. The two most widely used modulation options are GMSK, which is used in the European GSM system, and QPSK, which is used in the Japanese and North American digital cellular standards.

Digital modulation allows a number of advantages related to capacity and quality. Digitized voice can be processed by compression techniques to bring down the bit-rate to around 8 to 10 kb/s. This means that up to 3 users may be served in the 30 kHz bandwidth used by analog FM. The bandwidth requirement may be even further reduced by using bandwidth-efficient M-ary modulation. (However, use of M-ary modulation introduces the additional issue of power-bandwidth tradeoff, discussed later in this report.) Additionally, digitized voice may be protected against channel errors by FEC coding. This helps protect against errors in transmission which are unacceptably likely to occur in an unprotected channel, due to reflections from buildings etc. In fact, both the European and North American digital cellular standards do incorporate Viterbi coding. However, use of FEC does require additional bandwidth, as discussed below.

2.2.2 POWER-BANDWIDTH TRADEOFF IN DIGITAL CELLULAR

The power-bandwidth tradeoff is fundamental to all of the technological options being considered to increase the capacity in digital cellular systems. The power bandwidth tradeoff is a result of Shannon's equation. The maximum rate of information transfer, in bits/second, over a channel of bandwidth B is

$$R = B \log_2 \left(1 + \frac{S}{N} \right)$$

where S/N is the signal-to-noise ratio at the receiver. For our purpose, we will use S/N interchangeably with C/I , the ratio of the received power of the signal carrying desired information, to the received interference power. C/I is sometimes referred to as CIR, the carrier to interference ratio.

For bandwidth efficient communication, to make it possible to serve the largest possible number of users, we would like to keep the channel bandwidth low. From Shannon's equation we see that for a given information transfer rate, we can reduce the channel bandwidth if we can increase the SNR. But we saw above, that in cellular systems, increase in the CIR comes from increasing the distance to the cochannel cells, which in turn reduces capacity. The precise effect of the power bandwidth tradeoff in digital cellular is still a subject of ongoing research [4]. The various technology options for enhanced capacity in digital cellular and the implications of the power bandwidth tradeoff are summarized in Fig. 10.

2.2.3 QPSK, CODING AND TDMA

As was stated above, QPSK has been adopted as the standard modulation scheme for next generation North American digital cellular. QPSK is 4-ary modulation, and therefore provides for a doubling of bit-rate within the same bandwidth. Thus it should be possible to transmit 60 kb/s over the standard 30 kHz analog cellular (AMPS) channel. However, from our discussion of the power bandwidth tradeoff above, it should not immediately be concluded that this alone will lead to increase in the system's service capacity.

The TDMA digital cellular standard also calls for coding. There are two forms of coding. The first is voice or source coding, and the second is channel coding. Source coding extracts the essential features of the information and represents them in fewer bits than when the analog signal is sampled and digitized. Source coding applies only to voice or video signals. When sampled and digitized, a standard voice signal requires 64 kb/s. Voice coding techniques such as ADPCM, LPC etc. can be used to reduce the bit-rate. The TDMA standard calls for use of LPC to reduce the bit-rate of voice signals to 10 kb/s. By the way, NTSC full-motion color video signal requires 76 Mb/s when sampled and digitized. Digital data compression techniques such as ADPCM, DCT, wavelet technique, etc. can be used to reduce the bit-rate of video signals.

FEC or channel coding is used to protect against errors in the transmitted signal induced by channel imperfections. Parity bits and CRC bits are two common forms of channel coding that

allow for error detection but not error correction. More advanced types of channel coding allow for a limited amount of automatic error correction. There are two types of FEC coding: block and convolution. Convolution coding with Viterbi decoding is a widely used FEC technique. Use of FEC increases the number of bits to be transmitted. It would appear that this is undesirable due to increase in the required channel bandwidth. However, the principle of power-bandwidth tradeoff mitigates the bandwidth expansion by a corresponding reduction in the required CIR. Thus channel coding can be a benefit. The digital cellular standard calls for convolution coded FEC.

The new digital cellular standard uses TDMA. The users must take turns to use the channel in an orderly synchronized way. Nominally, the total channel capacity is 60 kb/s, and the bit-rate of raw digitized voice is 64 kb/s. When the effects of source coding, channel coding, the need for guard time between users, and the reservation of time slots for network control is all taken into account, it is seen that three digital users can fit into the 30 kHz bandwidth of one analog user.

2.2.4 CDMA, POWER CONTROL, VOICE ACTIVATION, CELL SECTORIZATION

CDMA has been proposed as an alternative to TDMA for digital cellular [1, p. 472]. The main benefit claimed for CDMA is that it will provide for 20 times the capacity of the analog cellular (AMPS) system. As noted above, the TDMA architecture provides for a tripling of capacity.

CDMA is based on the spread spectrum technology that the military has used for many years. There are two kinds of spread spectrum: Direct Sequence and Frequency Hopping. Direct sequence is more popular because it is easier to implement in practice. In DS-SS each user bit is converted into a specific long sequence of bits. The resulting string is transmitted using a simple modulation such as BPSK. The user specific sequence of bits is called the PN code. The length of the sequence is called the processing gain.

When CDMA is used, the transmitted bit-rate, and hence the channel bandwidth is increased by a factor equal to the processing gain. Because of the increased bandwidth, by the principle of power-bandwidth tradeoff, CDMA is able to operate with correspondingly reduced CIR. For example if the user bit rate is 10 kb/s, which is typical for compressed speech, a processing gain of 100 would lead to a transmission of 1 Mb/s. At the same time the required CIR can be reduced, say from 3 to 0.03. It is the reduction in required CIR that leads to a capacity benefit in cellular systems.

The CDMA system is designed to operate in an environment of many interfering users. All users use the same frequency band. Each user in a CDMA system is assigned a unique PN code. The receiver is tuned to a particular code. It uses a process called matched filtering to select the user of interest and reject other interfering transmissions.

In addition to the reduction in CIR, additional capacity enhancing features are used in the proposed CDMA cellular system. These are power control, voice activation, and cell

sectorization. Transmitter power control is used so that transmissions from users close to the cell site do not overwhelm the weaker received signals from users farther away. Voice activation is based on the statistical observation that a typical voice signal is active only 38% of the time. Shutting off the transmitter whenever there is no speech signal present, reduces interference to the other users. Sectorization means that the circular cell is divided into three independent 120° sectors by the use of directional antennas. In essence, cell sectorization is like having three smaller cells, except that the hand-off between sectors can be handled locally, without interference by the MTSO.

It can be shown that out of the factor of 20 capacity benefit of CDMA, about a factor of 10 can be attributed to cell sectorization and voice activation, and only about a factor of 2 to the lower CIR. In general, it is difficult to improve capacity by applying the power-bandwidth tradeoff alone. In CDMA systems, the increase in channel bandwidth offsets some of the capacity gains obtained from the CIR reduction. On the other hand, in M-ary modulation, the increase in required CIR offsets some of the capacity gains obtained from the bandwidth reduction. The comparative features of TDMA and CDMA are shown in Fig. 11.

2.2.5 ALOHA AND PRMA

ALOHA is the protocol developed at the university of Hawaii for random access packet data communication [1, p.231]. It is inherently suited for bursty data sources such as computer terminals. It avoids the wasted bandwidth associated with the uniform division of bandwidth associated with such techniques as FDMA, TDMA, CDMA, etc.

Essentially, in ALOHA protocol, a terminal transmits a packet as soon as it is ready. The transmission succeeds if no other terminal transmits simultaneously. Because success is a random variable, this is also referred to as statistical multiplexing. A successful transmission is verified by the destination terminal transmitting an acknowledgment to the source. If a transmission fails, the packet must be retransmitted. The maximum utilization of the channel in ALOHA is 18%. Even so, this is an improvement over TDMA when the data sources are very bursty. An improvement of ALOHA is slotted ALOHA in which the time-scale is divided into uniform periods of length equal to one packet duration, called slots. Every transmission is constrained to start at the beginning of a slot. The maximum utilization in slotted ALOHA is 36%. A further improvement is carrier sense multiple access (CSMA), in which each terminal monitors the channel for any ongoing transmission, and defers its transmission until the current transmission is complete. CSMA is used in ethernet. Each improvement in channel utilization comes at the cost of increased complexity in the receiver. In general, packet communication networks use some sort of statistical multiplexing. This is also true for ATM. Statistical multiplexing performs well when the total network load is light, but the channel utilization drops dramatically when the network traffic is high, due to repeated collisions and retransmissions.

There has not been a great deal of attention devoted to packet communication in the study of cellular systems. This may be due to the fact that cellular service started with voice. In fact, even the voice market is yet to mature fully. The most noteworthy work in the cellular packet communication is the development of PRMA (packet reservation multiple access) [5]. Cellular

packet communication is still a subject of ongoing research [6]. Outside the cellular context, packet radio networks (PRN) have been extensively studied [7] for their application to DARPA survivable radio network architecture.

We will restrict our attention to PRMA here. PRMA is a combination of slotted ALOHA and TDMA. Each PRMA frame consists of a fixed number of slots. Each slot can hold a single user packet of 44 bytes of payload and 4 bytes of address. Users contend for a slot by randomly accessing a vacant slot. If the base station successfully decodes a user's transmission in a particular slot, that slot is reserved for that particular user in succeeding frames, much like TDMA. The base station continuously broadcasts information about reserved and accessible slots in addition to down link data for users already connected. Each packet is also acknowledged by the base station. A user maintains control of his slot as long as he has data to send. When he has no data to send, he relinquishes control, and other users may contend for that slot. Due to statistical multiplexing, the number of users that can be supported can be larger than the number of slots. The exact number will depend on the burstiness of the data streams. For voice traffic, assuming voice activated transmission, there can be 1.6 users per slot [6].

III. DISCUSSION

In this section we will consider the technology options presented in Sec. 2 and evaluate them against criteria posed by the applications of Sec. 1 to develop a scenario for the development of mobile multimedia technology.

It is assumed that there will be a large number, possibly thousands, of wireless ports scattered throughout the service area, indoors as well as outdoors. Users are free to access these radio links anywhere, and demand variable amounts of bandwidth service, i.e. voice, data, or video. The goal is to service a shifting pattern of bandwidth demand in a way that the cost is proportional to the aggregate demand rather than the product of the peak demand times the number of ports. In other words, although each port should be capable of servicing the peak bandwidth demand, say for video, the actual network resource needed to serve the port should depend only on the instantaneous bandwidth demand, be it voice, video or data. ATM networks have this feature.

3.1 ISSUES IN MOBILE MULTIMEDIA

Following is a potential list of issues for discussion:

- 1) Granularity. How close should the ports be located?
- 2) Access method. TDMA, CDMA or ALOHA based.
- 3) Modulation and coding method.
- 4) Intercell network.

- 5) ATM adaptation. Whether packetization should take place in the terminal or in the base station.
- 6) Service capacity per port
- 7) Operation flexibility. How will a shifting demand pattern be serviced?
- 8) Spectrum. How much bandwidth will be needed, and where will it come from?
- 9) Cost
- 10) Technology development timeline

3.1.1 GRANULARITY, BANDWIDTH AND CAPACITY

The answers to these three issues are intimately linked. The number of users per cell that can be supported is fixed by the spectrum available, and technology choice e.g. CDMA, PRMA etc. Of course, users can have variable bandwidth demand, so by the word user we mean a fictitious user with average bandwidth demand. Once the number of users per cell is fixed, the number of cells required is fixed by the total demand in the service area. If there is enough bandwidth to support 40 mixed service users, we may designate one port per wing in an office environment. If the number of users per cell drops to 10 we may need a base station in every room. The smaller the cell size, the less the bandwidth as well as power required. Thus smaller cells have a practical advantage. In theory, smaller cells would require closer management by the network as the demand moves from cell to cell.

As demand moves around, it is possible that there will be a demand imbalance from cell to cell. To address this imbalance, it is proposed here that the intercell network be an ATM switch. The traditional MTSO and cell planning which treats each cell equally and assigns equal bandwidth to every cell are not desirable.

3.1.2 ACCESS METHOD AND ATM ADAPTATION

This issue should be clear to the reader by now. A review of the path taken by switching technology and the trends seen in cellular technology suggests that many of the mistakes of switching are being repeated by cellular. Switching technology developed with only voice in mind, and had to go through a radical rethinking when data applications proliferated. The same trend is now seen in cellular. The big debate in cellular is whether TDMA or CDMA is better for capacity. It is proposed here that both are methods for dividing capacity equally among users, and are suited for voice, not data. Mixed media service will require a packet based technology. It could be PRMA, or some other technology yet to come. The choice of packet technology for cellular use is still an active research area. Some lessons learned for packet radio applications in DARPA may come in useful for this.

The above discussion also addresses the issue of ATM adaptation. Since the PCS scenario

envision a variety of terminal with varying bandwidth demands, it is essential that the access to the radio port be packet based. Thus packetization should take place in the terminal, not the port.

3.1.3 COST

All the technology needed for a multimedia terminal is available today. It would most likely combine a notebook PC, a miniature video camera, a packet network interface, similar to an ethernet card, and a radio such as in a cellular phone. All of these items can be purchased today for a total cost of \$4000. The cost is likely to come down as the product is integrated and mass produced. Based on the price trends seen in PCs, we can estimate that the cost would be around \$1000 in about 5 years.

3.1.4 TECHNOLOGY DEVELOPMENT TIMELINE

The real challenge in mobile multimedia technology will be to develop the network. This will have to go through a lengthy and complicated process. First, the architecture has to be developed further, and then standardized. This could involve a fairly long process of negotiation between vendors, regulating agencies, and standardization bodies. Based on the experience with digital cellular standard, we can guess that this will take about 5 to 7 years. After this, the network will have to be deployed. Based on the experience with recent products such as PCs and cellular phones this is a more rapid process and may take anywhere from 3 to 5 years. Therefore, we should start to see prototypes in about 5 years, and widespread deployment in 8 years.

3.1.5 RF SPECTRUM USAGE AT KSC

An important issue is how much bandwidth will become available for mobile multimedia technology. This is very much affected by how the RF spectrum is currently being used at KSC, and the possibility and cost of interference with critical operations. For this reason a short review of RF usage at KSC is included.

The primary operational communication systems, viz. OISD and OTV are not based on RF. Mission critical communication is often carried on land links in order to increase reliability. Where wireless techniques are conceived for mission critical operations, suitable error protection such as CRC or FEC should be included for reliability. Wireless communications are primarily used for communication with the orbiter itself. This includes telemetry, voice, data and video. The communication is carried in the S-band (2.2 GHz) and Ku-band (15 GHz) using 80 Mhz of bandwidth. The Ku-band is used when the vehicle is in orbit, while the S-band is used during ascent and descent. In base operations, RF is used for emergency communications such as police and fire rescue. These communications are carried out in the 150 Mhz, 400 Mhz and 800 Mhz bands, each 20 Mhz wide. KSC must share the frequencies reserved for government use with other government agencies in the area.

As technology progresses, new methods to utilize the spectrum more efficiently become available. To implement these requires a consensus with entities holding existing claims to the spectrum. Finding spectrum for new applications is not so much an issue of taking it away from

existing stakeholders, as it is of finding new ways to share it for maximum benefit to the community as a whole.

IV. CONCLUSIONS

A review of ATM and PCS technologies shows that mobile multimedia technology could offer benefits at KSC in operations as well as in ground support. Current trends in cellular system technology are driven by the voice market rather than multimedia. Significant effort will be required in the development of network technology to support mobile multimedia. A shift toward packet based cellular communication will make mobile multimedia possible. Prototypes costing around \$1000 are possible in about five years, and widespread deployment in about 8 years.

V. REFERENCES

- [1] Norman Abramson, ed., *Multiple access communications: foundations for emerging technologies*, Piscataway, N.J.: IEEE Press, 1993.
- [2] Chris Dhas, K. Konangi, and M. Sreetharan, eds., *Broadband switching: architectures, protocols, design, and analysis*, Los Alamitos, Calif.: IEEE Computer Society Press, 1991.
- [3] Amit Bhargava, ed., *Integrated broadband networks*, Boston, Mass.: Artech House, 1991.
- [4] R. R. Gejji, "Channel efficiency in digital cellular capacity," *Microwave Journal*, vol. 35, no. 8, pp. 22-27. Also in *Proc. 42nd IEEE Vehicular Technology Conference*, Denver, May 1992, pp. 1005-1007.
- [5] D.J. Goodman, "Cellular packet communications," *IEEE Trans. Commun.*, vol. 38, August 1990, pp. 1272-1280.
- [6] R.A. Ziegler and G.P. Pollini, "Experimental implementation of the PRMA protocol for wireless communication," *Proc. IEEE 43rd Vehicular Technology Conference*, Secaucus, N.J., May 1993, pp. 909-912.
- [7] B.M. Leiner, D.L. Nielson, and F.A. Tobagi, eds., Special issue on packet radio, *Proceedings of the IEEE*, vol. 75, no. 1, January 1987.

PROMISE OF PCS TECHNOLOGY

PCS: PERSONAL COMMUNICATION SYSTEM

- Same phone anywhere -- home, office, car.
- Tetherless access
- Global roaming
- Access to custom terminal environment from any point in the network
- Multimedia services -- voice, fax, data, video
- Lightweight, portable terminal

Fig. 1. Promise of PCS Technology.

PROMISE OF ATM TECHNOLOGY

ATM: ASYNCHRONOUS TRANSFER MODE

- Packet switching based technology
- Mixed services -- voice, video, data, on single network
- Multimedia services
- Dial-up video library
- Teleconference and video intercom
- Electronic town hall

Fig. 2. Promise of ATM Technology.

KSC APPLICATIONS POTENTIAL

- Tetherless workstations. No cable plant changes. More flexible deployment of bandwidth.
- Two-way wireless multimedia support for operations.
- Video instructions and scheduling
- Fast wireless video closeouts
- Video conference to/from mobile locations

Fig. 3. KSC Applications Potential.

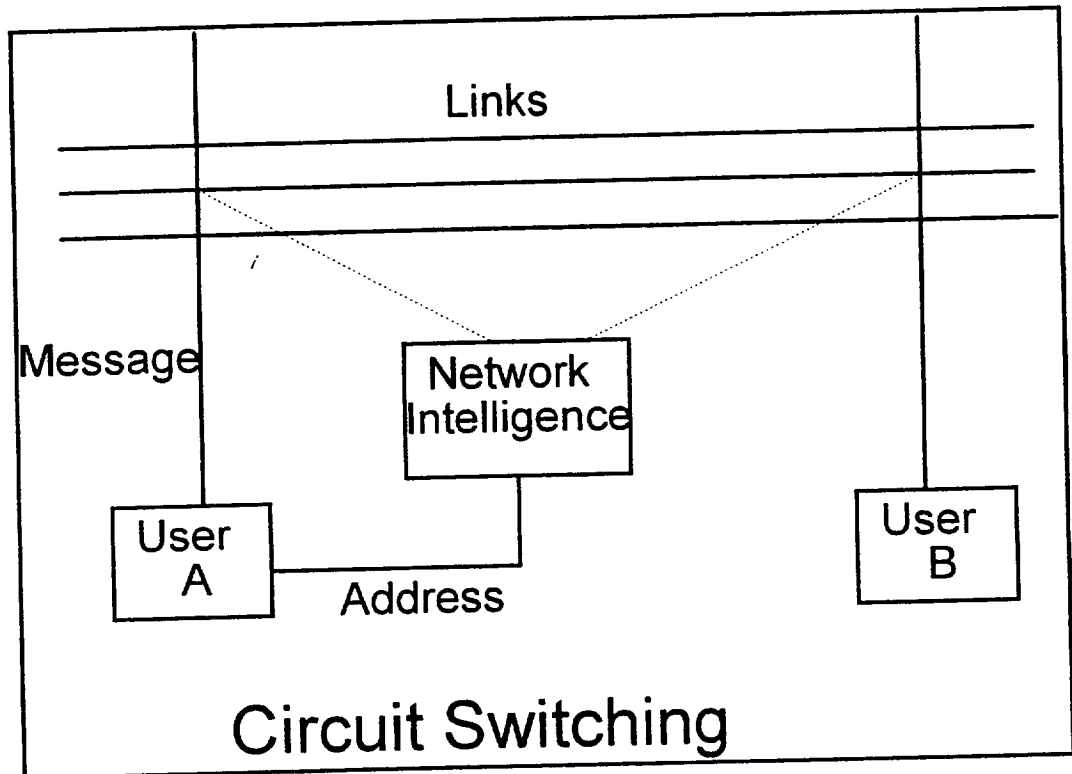


Fig. 4. Circuit Switching.

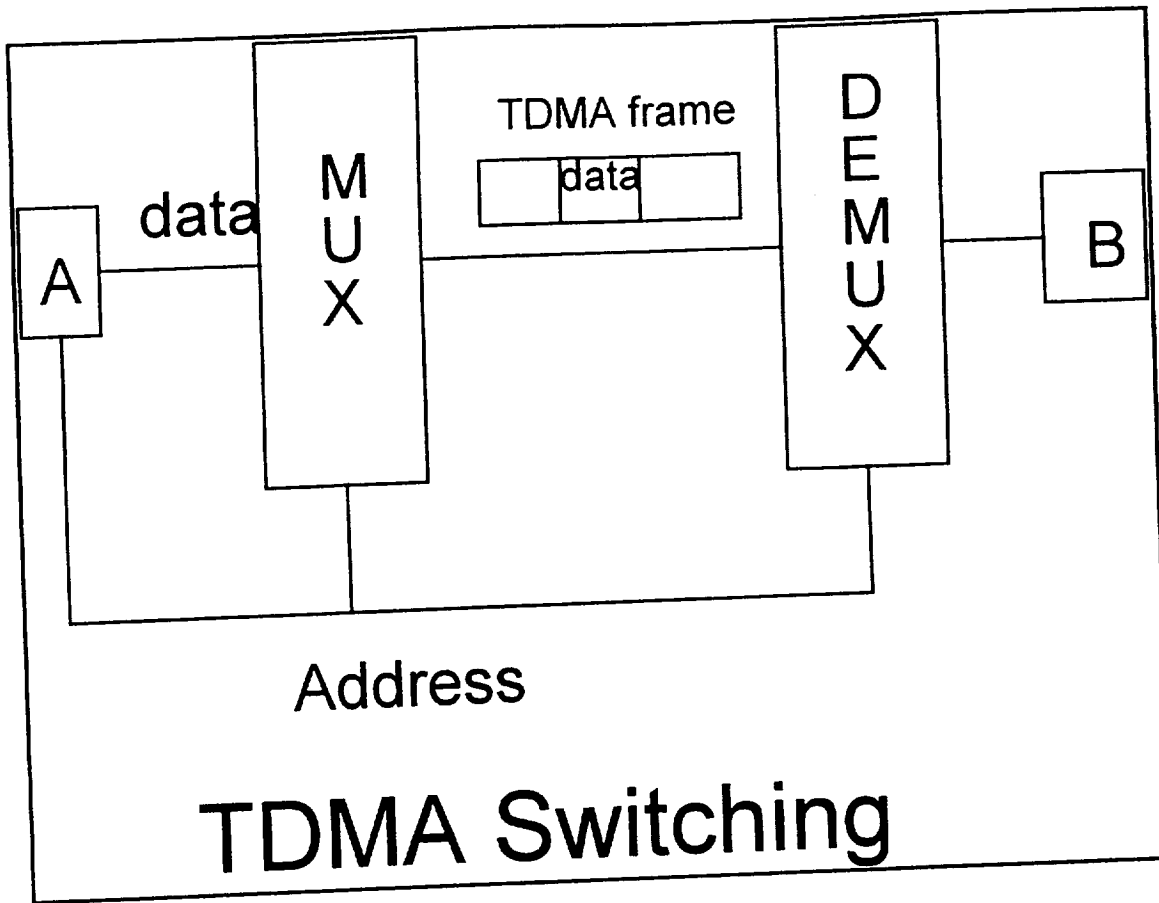


Fig. 5. TDMA Switching.

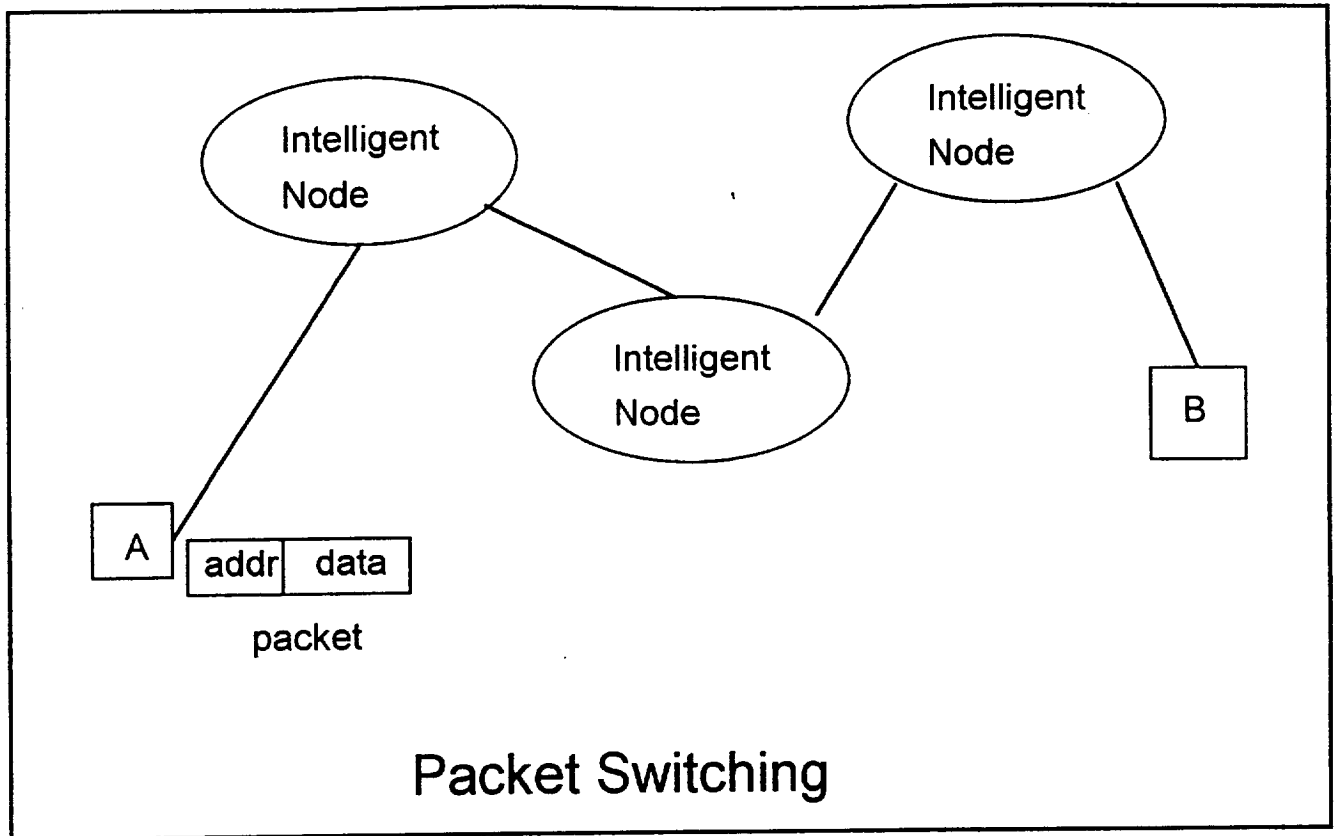
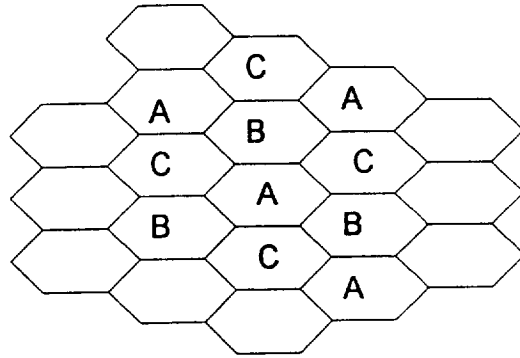


Fig. 6. Packet Switching.

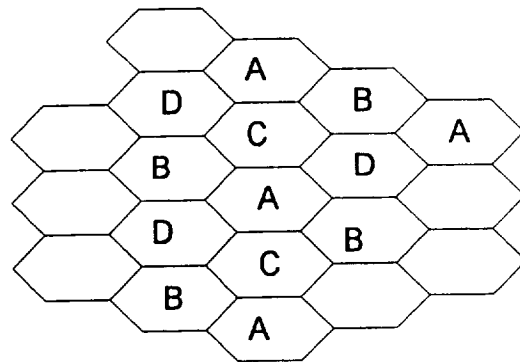
CELLULAR SYSTEM CONCEPTS

- **Basic concept:** Divide service area into circular cells. Assign channels to cells; neighbor cells do not have same. Reuse of channel in other cells.
- **Channel :** User's set of assigned frequencies for 2-way comm. within the cell
- **Cochannel cell:** Where the same channel is being reused
- **Reuse distance:** Smallest distance between cochannel cells
- **Cochannel interference**
- **Capacity grows as inverse of reuse distance**
- **Interference also grows same way**
- **Key problem:** How to increase capacity but not interference.
- **Hand-off :** Automatic assignment of new channel as mobile moves from one cell to another
- **Base station or cell site :** Radio port to land-based network, one for each cell.
- **MTSO :** Mobile telephone switching office. Switch for intercell communication.

Fig. 7. Cellular system concepts.



THREE CHANNEL REUSE PATTERN



FOUR CHANNEL REUSE PATTERN

Fig. 8. Example of 3- and 4- channel reuse patterns.

CELLULAR SYSTEM PERFORMANCE

- Performance Measure : number of users per cell.
- Lee's formula:

$$m = \frac{B_t / B_c}{\sqrt{\frac{2C}{3I}}}$$

where m = users per cell

B_t = Total system bandwidth

B_c = Channel bandwidth

C/I = Worst case CIR (Carrier to interference ratio).

Varies directly as reuse distance.

Fig. 9. Cellular system performance.

CAPACITY ENHANCING TECHNIQUES

Method	BW	CIR (Reuse dist)
Voice compression	dec	none
Error correction	inc	dec
M-ary mod. (QPSK)	dec	inc
CDMA	inc	dec
Sectorization	none	dec
Voice activation	none	dec

Fig. 10. Capacity enhancing techniques for digital cellular.

PROPOSED DIGITAL CELLULAR STANDARDS

TDMA

- QPSK modulation
- Source coding (LPC)
- Channel coding (convolution)
- Triple AMPS capacity

CDMA

- Spread spectrum technique
- Cell sectorization
- Voice activation
- 20 times AMPS capacity

Fig. 11. Comparative features of TDMA and CDMA.

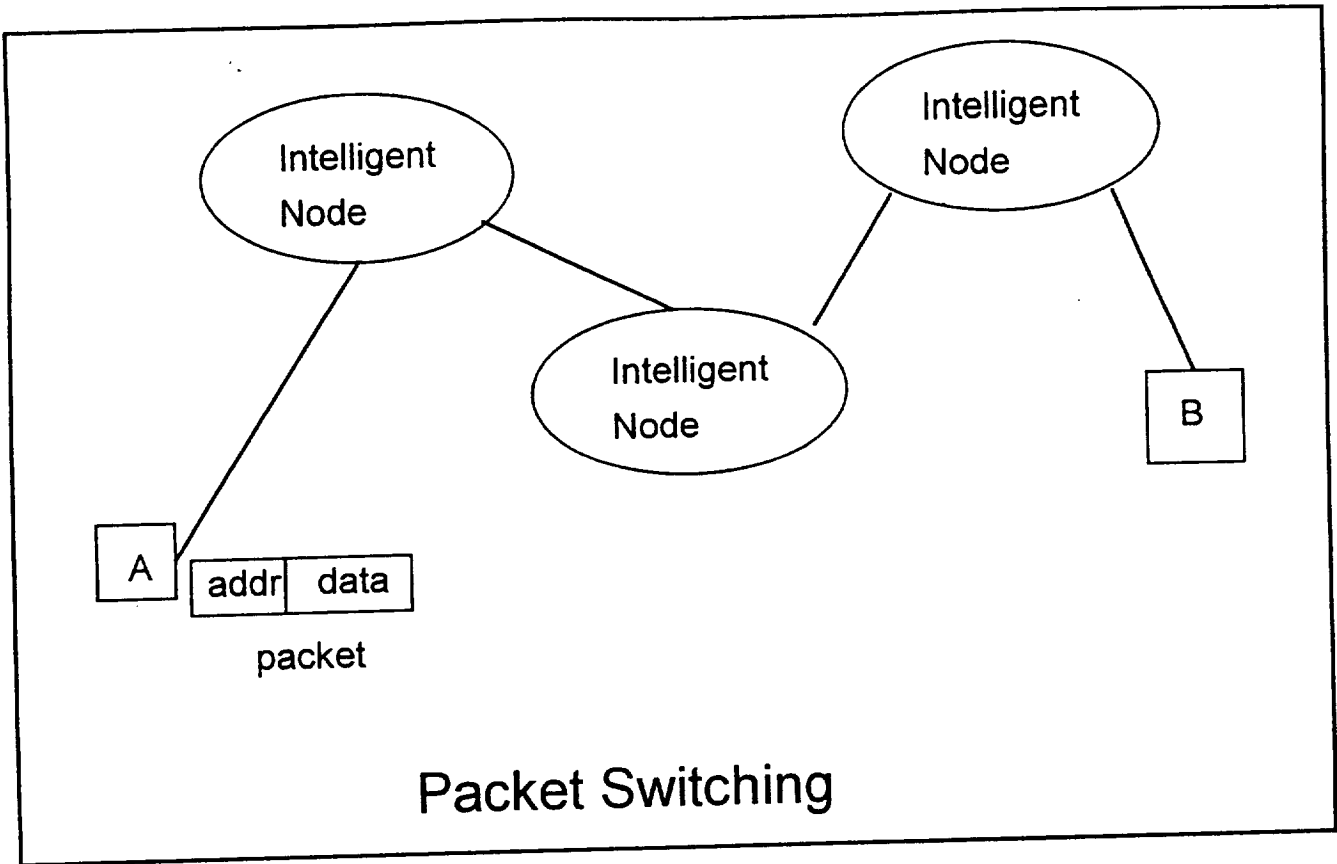
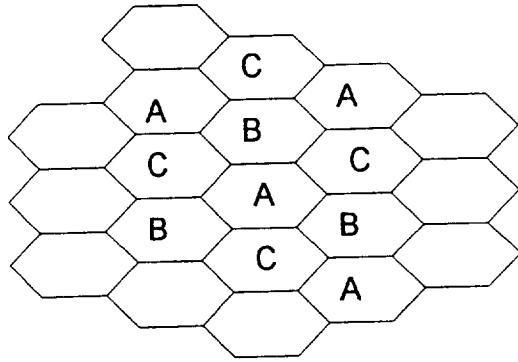


Fig. 6. Packet Switching.

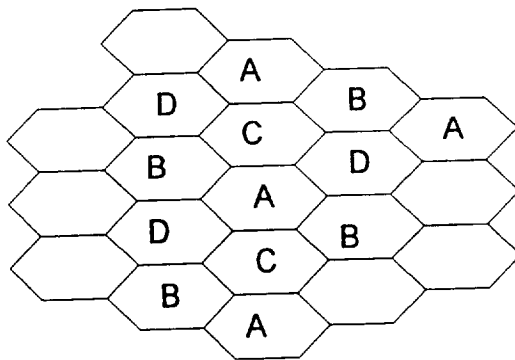
CELLULAR SYSTEM CONCEPTS

- **Basic concept:** Divide service area into circular cells. Assign channels to cells; neighbor cells do not have same. Reuse of channel in other cells.
- **Channel :** User's set of assigned frequencies for 2-way comm. within the cell
- **Cochannel cell:** Where the same channel is being reused
- **Reuse distance:** Smallest distance between cochannel cells
- **Cochannel interference**
- **Capacity grows as inverse of reuse distance**
- **Interference also grows same way**
- **Key problem:** How to increase capacity but not interference.
- **Hand-off :** Automatic assignment of new channel as mobile moves from one cell to another
- **Base station or cell site :** Radio port to land-based network, one for each cell.
- **MTSO :** Mobile telephone switching office. Switch for intercell communication.

Fig. 7. Cellular system concepts.



THREE CHANNEL REUSE PATTERN



FOUR CHANNEL REUSE PATTERN

Fig. 8. Example of 3- and 4- channel reuse patterns.

CELLULAR SYSTEM PERFORMANCE

- **Performance Measure** : number of users per cell.
- **Lee's formula:**

$$m = \frac{B_t / B_c}{\sqrt{\frac{2 C}{3 I}}}$$

where m = users per cell

B_t = Total system bandwidth

B_c = Channel bandwidth

C/I = Worst case CIR (Carrier to interference ratio).

Varies directly as reuse distance.

Fig. 9. Cellular system performance.

CAPACITY ENHANCING TECHNIQUES

Method	BW	CIR (Reuse dist)
Voice compression	dec	none
Error correction	inc	dec
M-ary mod. (QPSK)	dec	inc
CDMA	inc	dec
Sectorization	none	dec
Voice activation	none	dec

Fig. 10. Capacity enhancing techniques for digital cellular.

PROPOSED DIGITAL CELLULAR STANDARDS

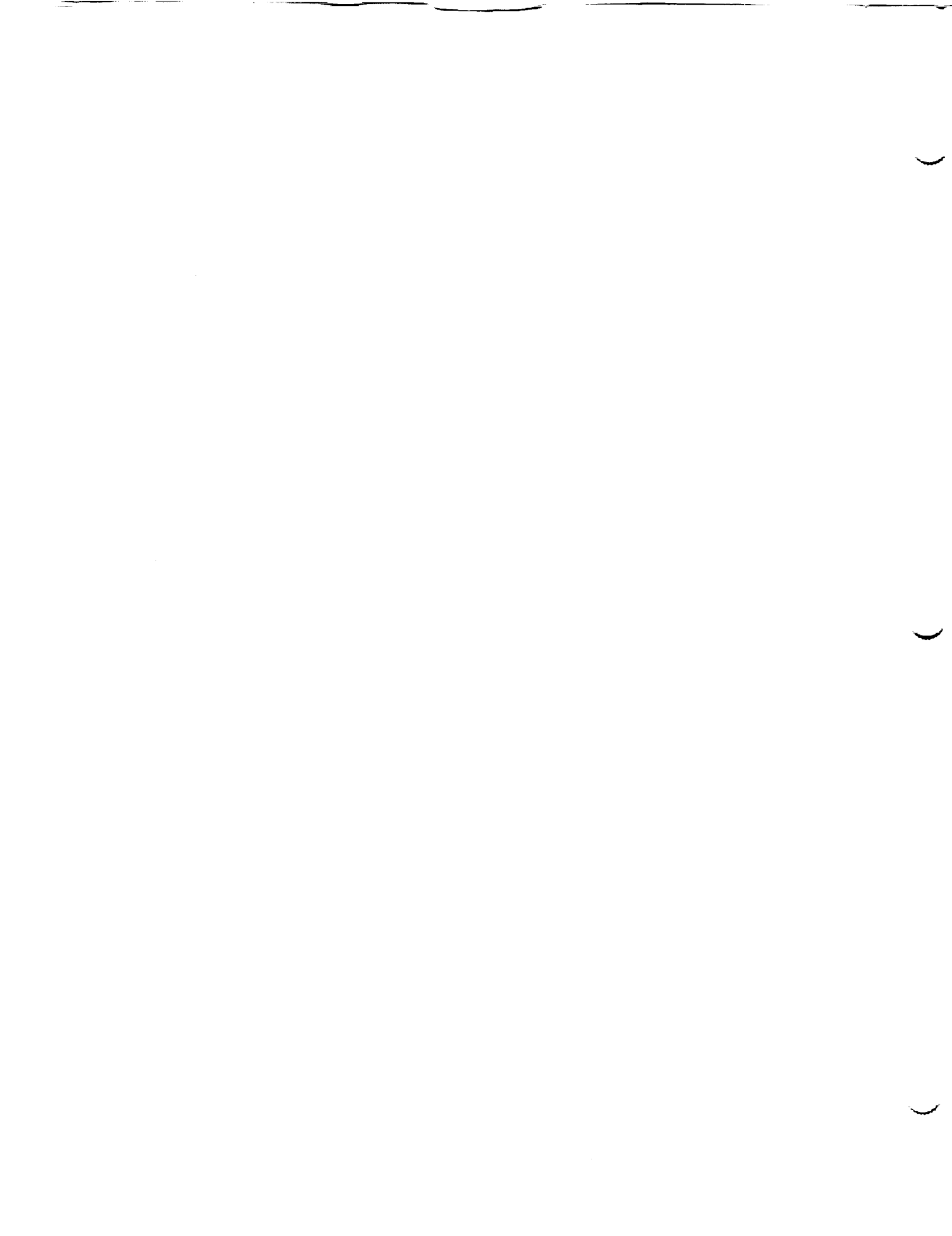
TDMA

- **QPSK modulation**
- **Source coding (LPC)**
- **Channel coding (convolution)**
- **Triple AMPS capacity**

CDMA

- **Spread spectrum technique**
- **Cell sectorization**
- **Voice activation**
- **20 times AMPS capacity**

Fig. 11. Comparative features of TDMA and CDMA.



1994018767
N94-28240

1993 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA

510-61
197196
p. 39

COST BENEFITS OF ADVANCED SOFTWARE:
A REVIEW OF METHODOLOGY USED AT KENNEDY SPACE CENTER

PREPARED BY:	Dr. Prafulla N. Joglekar
ACADEMIC RANK:	Professor
UNIVERSITY AND DEPARTMENT:	La Salle University Management Department
NASA/KSC	
DIVISION:	Shuttle Project Engineering
BRANCH:	Process Integration
NASA COLLEAGUE:	Arthur E. Beller
DATE:	August 6, 1993
CONTRACT NUMBER:	University of Central Florida NASA-NGT-60002 Supplement: 11

Acknowledgments

I am thankful to the NASA/ASEE fellowship program for this research opportunity. It has been exciting to witness space exploration from this proximity. It was intellectually very satisfying to be exposed to research in so many different disciplines related to NASA's mission. At the same time, it was very interesting to observe the management of the operations involved in getting an Orbiter ready for the next launch. It was most fascinating to study the development of future information systems to support that management.

I feel privileged to be able to propose a methodology for systematic and rational choices among information systems investment alternatives. Although I was not successful in demonstrating the actual use of my methodology, I have come to realize that, given the existence of the vicious circle I have discovered in this study, that was an unrealistically ambitious goal. On the other hand, I believe that I have enriched my own understanding of the methodology and designed a process for its implementation that is superior to any other process I have seen in the CBA literature.

I appreciate the generosity, and candidness of the many NASA and Lockheed employees I interviewed this summer. I have very freely used many of their ideas in this report without explicitly crediting them for the same. But I do want to thank them for the same. Special thanks are due to the authors of the four studies I was allowed to use as attachments to my report.

Ray Hosler (UCF), and Kari Stiles (UCF) provided valuable and very effective administrative support, as well as educational and entertainment opportunities. Gene Thurston's (TV-PEO-1) constant encouragement and guidance allowed me to overcome the occasional frustrations I encountered. Finally, this summer could not have been as enjoyable as it has been, without the help of Art Beller (TV-PEO-12) who is an outstanding mentor, thinker, facilitator, and by now, a close friend. I express my heartfelt gratitude to each one of these special individuals.

COST-BENEFITS OF ADVANCED SOFTWARE: A REVIEW OF METHODOLOGY USED AT KSC

**By
Prafulla Joglekar**

Abstract

To assist rational investments in advanced software, a formal, explicit, and multi-perspective cost-benefit analysis methodology is proposed. The methodology can be implemented through a six-stage process which is described and explained. The current practice of cost-benefit analysis at the Kennedy Space Center is reviewed in the light of this methodology. The review finds that there is a vicious circle operating. Unsound methods lead to unreliable cost-benefit estimates. Unreliable estimates convince management that cost-benefit studies should not be taken seriously. Then, given external demands for cost-benefit estimates, management encourages software engineers to somehow come up with the numbers for their projects. Lacking the expertise needed to do a proper study, courageous software engineers with vested interests use ad hoc and unsound methods to generate some estimates. In turn, these estimates are unreliable, and the vicious circle continues. The proposed methodology should help Kennedy Space Center to break out of this vicious circle.

COST-BENEFITS OF ADVANCED SOFTWARE: A REVIEW OF METHODOLOGY USED AT KSC

**By
Prafulla Joglekar**

Executive Summary

Advanced software (ASW) investment decisions are multi-stage, varied, complex, risky, and controversial. Therefore, we need a systematic methodology to assist rational ASW investment decisions. I propose a formal, explicit, and multi-perspective cost-benefit analysis (CBA) methodology for this purpose. I outline a number of rich concepts and principles of this methodology, and recommend a six-stage process for its implementation. In the light of this methodology, my review of the current practice of CBAs at KSC finds that the practice is seriously deficient.

The basic cause underlying these deficiencies is that we are caught in a vicious circle described by the following paragraph:

At present, CBA studies fail to capture all the relevant concerns. They measure only selected costs and benefits using questionable assumptions and unsound methods. As a result, the estimated costs and benefits are highly unreliable. Consequently, management looks at CBAs not as decision-making tools, but as mere exercises in generating numbers for external justification of decisions already made. Thus, management does not take CBA studies seriously, and simply leaves the conduct of CBAs up to the initiative of the software engineers involved in specific projects, without any provision for additional resources and expertise needed for these studies. Lacking resources, and the necessary expertise in economic analysis, but with vested interests in justifying their projects, courageous software engineers use creative, but ad hoc and unsound methods to conduct their CBAs. The resulting cost-benefit estimates are highly unreliable, and certainly not worthy of use in any rational decision-making. Thus, management's view that CBAs are to be used merely as exercises in generating numbers for external justification is reinforced, and so on. The vicious circle continues!

I recommend that at KSC, we should try urgently to break out of this vicious circle. The methodology I have proposed provides one exit point to break out of this circle. The other exit point is a change in management's perception of what a good methodology can do, and its willingness to provide adequate resources and appropriate expertise to the conduct of CBAs.

Table of Contents

<u>Section</u>	<u>Title</u>	<u>Page</u>
1.	INTRODUCTION	1
2.	THE PROPOSED METHODOLOGY	4
2.1	Richness of the methodology	4
2.2	A Clarification of Some Common Misperceptions	5
2.3	A Process for Implementation	6
2.4	Implementation Requirements and Advantages	11
3.	A REVIEW OF THE CURRENT PRACTICE AT KSC	12
3.1	The Vicious Circle	13
4.	CONCLUSION AND RECOMMENDATIONS	16
	REFERENCES	17
ATTACHMENT A	EXCERPTS FROM A CBA OF SCAN REPLATFORMING	A-1
ATTACHMENT B	CBA FOR THE DEPLOYMENT OF KATE	B-1
ATTACHMENT C	EXCERPTS FROM A CBA OF GPSS	C-1
ATTACHMENT D	RUBICON COST ANALYSIS STUDY	D-1

Abbreviations and Acronyms List

ASW	Advanced Software
CBA	Cost Benefit Analysis
CEA	Cost Effectiveness Analysis
CUA	Cost Utility Analysis
DTA	Decision Tree Analysis
GPSS	Ground Processing Scheduling System
G&A	General and Administrative overhead costs
KATE	Knowledge-based Autonomous Test Engineer
KSC	Kennedy Space Center
LRU	Link Replaceable Unit
RP	Rapid Prototyping
RUBICON	Reasoning Based on Intelligent Computer Operations and Networking
SCAN	Shuttle Connector Analysis Network
TA	Technology Assessment

COST - BENEFITS OF ADVANCED SOFTWARE: A REVIEW OF METHODOLOGY USED AT KSC

1. Introduction

Advanced software (ASW) projects are exciting. They keep us at the cutting-edge of technology; they help us develop and challenge the best minds in software development; they promise to capture the knowledge and expertise of the brightest and the most experienced personnel in the space program; they promise to minimize the chance of a human error while maximizing the chance of rapid trouble shooting in a launch count-down; and in general, they have the potential to help improve the effectiveness and efficiency of the operations at Kennedy Space Center (KSC). With the national focus on US competitiveness, we are also looking forward to ASW projects that promise commercial spin-offs.

As exciting as these promised benefits are, software development alone does not ensure actual realization of those benefits. Often many other necessary conditions must be obtained. For example, software such as Knowledge-based Autonomous Test Engineer (KATE) and Reasoning Based on Intelligent Computer Operations and Networking (RUBICON) will not enable us to actually reduce the manpower at the Firing Room consoles until management is willing to deviate from the traditions and practices that have clearly worked in the past, but that may be inferior and costly in the future compared to the use of these ASW.

On the other hand, advanced software development is not necessary to obtain certain improvements in operations efficiency. One well-known problem of today's computer systems is that their true potential is seriously under-utilized. For example, we are nowhere near realizing the reductions in hard-copy costs that are possible with the electronic communication capabilities already in place. Thus, detractors of ASW often suggest that what we need is not more investment in ASW, but more investment in the training and in the management of a change in people's attitudes and habits necessary for a fuller exploitation of the existing technology. Of course, proponents of ASW counter that exploiting even a small fraction of the potentially huge benefits of an ASW project may be well worth the costs of its development. Clearly, we need to identify the optimal mix of resources to spend on ensuring fuller use of existing technology and on developing new ASW.

In addition, there are a variety of interesting and challenging issues to resolve in ASW investment decisions. Given many ideas for ASW projects and limited resources at hand, we must decide which ideas to pursue and at what level of funding. By their very nature, ASW projects take many years to complete and carry the risks of technical, schedule, or operational failure. Thus, investment decisions pertaining to an ASW project are not simple one-shot, yes-or-no type decisions, but multi-stage decisions requiring a reassessment and redesign of the project at various stages in its life cycle. Below are a

few examples of the many interesting and challenging issues one has to deal with when making ASW investment decisions.

Some projects, such as the replatforming of the Shuttle Connector Analysis Network (SCAN), seem unavoidable given the obsolescence of the current platform. Yet, replatforming opens several possibilities for enhancements to current SCAN capabilities (e.g., LRU trace-through, Automated retest, Wire trace diagnostics, etc.), and total project costs depend upon the enhancements we decide to seek. We would be foolish not to exploit some of these opportunities for enhancements. However, the larger the set of enhancements we seek, the greater would be the project complexity and the consequent risk of failure. Thus, the real issue to be decided here seems to be what specific enhancements to seek and what not to.

Some projects, such as the Ground Processing Scheduling System (GPSS), seem to deserve continued funding on the basis of their past and measurable successes. However, the issue here may be who should fund it from this point on, and at what level? If GPSS's benefits are clearly demonstrable and the costs of its further development will be lower than its future benefits, is it time to spin it off as a commercial venture? Under this approach, a private firm will have to fund GPSS's further development and share in the rewards of its future success. Thus, a larger portion of Code C budget may be available to fund other ASW projects which may be too risky for a private (and risk-averse) entrepreneur but quite acceptable to a (risk-neutral) government. On the other hand, because of the many complicated legal and political issues involved, attempts to commercialize GPSS too soon could actually slow down its development and implementation.

Other ASW projects such as KATE, and RUBICON seem to deserve continued funding because they are based on truly visionary technologies. The issue here is whether these ASW projects represent a situation of "a solution looking for a problem to solve," and whether given our desire for being at the cutting-edge of technology, funding of visionary technologies is justified in and for itself.

Another issue pertaining to KATE and RUBICON seems to be the threshold level of funding needed to keep these projects at a reasonably productive pace. For some projects, no funding at all may be better than some funding below the threshold level. One concern is that with the speed at which some ASW projects are proceeding, there may be cheaper and better commercial products on the market long before our development is complete. Considering that possibility, the question is: Are we simply providing taxpayer-funded software development experience to the contractor?

When funding an ASW project (See Attachment A), we seem to budget for the time software engineers would spend on that project. In reality, the project uses many other resources in the organization. Computer hardware, and office supplies are the obvious examples of these. In addition, there are many *hidden* costs (hidden until we recognize them). For example, to the extent that ASW projects attempt to capture corporate knowledge and expertise, they require substantial time and cooperation from

various experts. Unless these experts' time is explicitly budgeted for the ASW project, project schedule and success may depend on the goodwill of the experts, and may even risk neglect of the experts' normal duties which may be launch-critical today. Unless all relevant costs of an ASW project are uncovered, added-up, and compared with the project's likely benefits, one does not know whether that ASW development would be a wise idea.

At the same time, it should be realized that if the experts are not convinced of the value of the project, or think that their jobs will be at risk once their expertise is captured, software engineers will not succeed in capturing their expertise. In other words, successful implementation of an ASW project often requires that each one of the many stakeholders of the project should find it cost-beneficial from his/her own perspective.

In short, ASW investment decisions are multi-stage, varied, complex, and risky, and their success depends on the cooperation of multiple stakeholders. It is no surprise that while there are a few success stories, there are many more instances of project failures, long delays, and wasted resources. Thus, most ASW investment decisions seem to be controversial. It is therefore imperative that we develop a systematic methodology to assist rational ASW investment decisions.

In Section 2, I propose a cost-benefit analysis (CBA) methodology to assist these decisions. I had hoped to demonstrate the use of this methodology in a couple of actual decision situations. Unfortunately, at KSC the concept of what a CBA methodology can do, and where to apply it, seems to be very different than mine. At KSC, CBAs are used to justify past decisions, or our preferred choices, to some external constituency. CBAs are not seen as an assistance to decision-making. Indeed, ASW projects that are facing serious decision points seem to avoid a systematic CBA. As a result, I did not really get a chance to demonstrate the use of my methodology. On the other hand, as is clear from the discussion in Section 2, I did have the opportunity to study several instances of the current practice of CBA at KSC. Attachments A through D present the relevant excerpts from the CBAs I studied. In section 3, I review the current practice as a whole and contrast it with my methodology. Section 4 provides my conclusions and recommendations.

2. The Proposed Methodology

Rational decision-makers always assess the costs, benefits, and risks of the alternative choices they have. However, this assessment is often informal, implicit, and only from a single (the decision-maker's) point of view. I recommend that at KSC the assessment of ASW investment alternatives be **formal, explicit, and multi-perspective**. Organizational decision-makers clearly recognize the need for a formal process of assessment. An explicit assessment forces us to articulate all underlying assumptions and verify their validity. An explicit process is also easier to study, improve over time, and pass on from one generation of decision-makers to the next. Many researchers suggest that a cost-benefit assessment be "objective." I believe that costs and benefits of an ASW lie in the "eye of the beholder." In other words, assessments, by their very nature, depend upon one's point of view, and hence are subjective. Instead of attempting to avoid this subjectivity, I recommend that the assessment be from the point of view of each one of the major stakeholders of an ASW investment. As I have suggested before, such a multi-perspective assessment improves our chances of obtaining full cooperation from all the stakeholders, and hence the chances of project success.

Rational decisions based on such a formal, explicit (therefore well documented), and multi-perspective assessment need no further efforts to justify them to our superiors or to the general public.

2.1 Richness of the Methodology

Formal CBAs have been done for over ninety years now, ever since the 1902 Harbor Act required that Army Corps of Engineers could build only those water projects that could be shown to generate more money than they consumed. Given the language of the Harbor Act, the foci of early CBA were on

- (i) *justifying* a decision already made, and
- (ii) quantifying all costs and benefits in *dollar terms*.

In many organizations, these foci continue to prevail even today. However, over the years, as CBAs are done in a wide variety of organizations analyzing a wide variety of decision situations, the CBA methodology has evolved considerably. In a previous publication [1], I have reviewed this evolution, and clarified a number of common misunderstandings about what a CBA methodology is, and is not.

Briefly, by now, we recognize that although a CBA can be used to justify a decision already made, *its most cost-effective use lies in arriving at the right decision*. We know that not all cost and benefits can be measured in dollar terms, if they can be measured at all. We have developed a variety of techniques such as cost-effectiveness analysis (CEA), cost-utility analysis (CUA), and technology assessment (TA) to accommodate variables that defy measurement and valuation in dollar terms. More importantly, we recognize that rational decisions can be made without forcing a quantification of the non-quantifiable, or a prediction of the unpredictable. I see these insights and techniques as an integral part of what I call "the CBA methodology."

The most fundamental principle of the CBA methodology is to account for (not necessarily quantify) **all incremental costs and benefits** resulting from a decision alternative. To enable us to do this task properly, the methodology provides a number of rich concepts and principles. For example, it describes the many different types of costs and benefits we may encounter, including: direct and indirect; tangible and intangible; fixed and variable; controllable and non-controllable; one-time and recurrent; etc. The methodology emphasizes the need to account for the *opportunity cost* of an action. The principle is to count the net benefits we would have reaped had we taken the best alternative action instead of a given action, as a cost of the given action.

The methodology tells us to pay attention to the cause-effect as well as the multi-producer-single-product relationships as may be present, and to attribute benefits and costs to the causes or the producers, as appropriate. It incorporates concepts and tools to adjust for the associated risks and uncertainties. In analyzing a multi-year stream of costs and benefits, the methodology provides us with techniques for converting these multi-year flows to comparable and consistent units, so that we do not "confuse apples for oranges". In short, the methodology is very rich and insightful.

2.2 A Clarification of Some Common Misperceptions

Unfortunately, in the information systems literature, some scholars have incorrectly equated CBA methodology with such financial techniques as internal rate of return or present value calculations. While accounting for the time value of money is an important principle of CBA methodology, the methodology is much broader in its scope than the narrow techniques it may use in specific analytical situations. I want to emphasize that I am recommending a **methodology**, not a single technique.

A methodology includes not only a toolkit, but also an understanding of the situations where each tool is most appropriate to use. Self-examination and improvement are integral parts of a methodology. Thus, answers to questions such as "Is cost-benefit analysis beneficial? Is cost-effectiveness analysis effective?" are legitimate parts of the methodology [2]. We recognize that some times, the benefits of conducting a formal and explicit CBA are not worth the time and costs required. The proposed methodology welcomes a formal, explicit, and rational decision not to pursue a CBA in such situations.

The methodology also requires that the scope and the level of detail of a CBA study be consistent with the magnitude of the likely costs of a wrong choice in an ASW investment decision, and with the time available for decision-making. A CBA study that costs \$10,000, when the largest possible difference between the net benefits of the best and the worst choice is only \$5,000, does not make any sense. Similarly, a study that takes a year to complete will not assist a decision that must be made within a month. Thus, in my view, a common fear, namely that a CBA will cost too much and take too long, is simply a misperception of the methodology.

One widely-held belief is that a CBA is useful only when a project is initially approved or disapproved, and it has no role to play in subsequent decisions about annual funding levels, etc., particularly so, if an original CBA was not conducted at the time of initial project approval. Once the methodology proposed here is in place, there will be no reason to assume that a CBA with properly defined scope and level of detail cannot assist the current year's funding decision pertaining to an on-going project, whether an initial CBA exists or not.

Of course, when an initial CBA does exist, the analysis in subsequent years is considerably easier. This is so because under my methodology, the initial CBA for an ASW project, incorporating Rapid Prototyping (RP) and anticipating a three year development cycle, would include a decision tree analysis (DTA) of the year-by-year alternative possible milestones of accomplishments and subsequent choices. Such a DTA spells out precisely what to do, once we know which one of the various possible milestones actually occurred during the previous year.

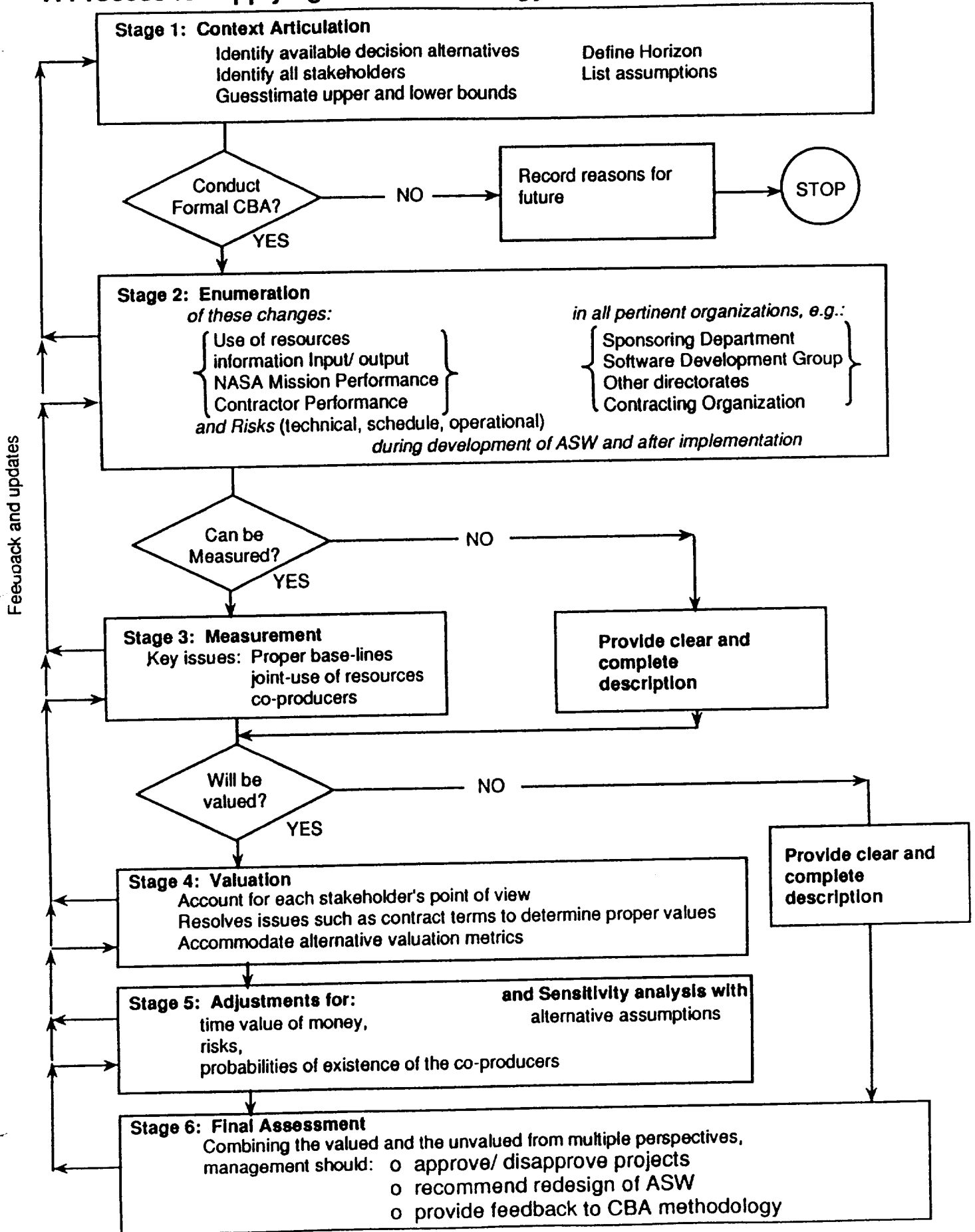
Perhaps the most pervasive misconception of the CBA methodology is that it accounts only for the "economic" costs and benefits, and ignores the many non-economic values we seek. With that misconception, some people even suggest that a CBA has no role to play in any government agency, let alone NASA, since government agencies exist precisely because market forces fail to provide for certain non-economic societal needs. I have shown elsewhere that economists in general, and CBA methodologists in particular, have always concerned themselves with the capture of the non-economic values [1]. The methodology I am proposing insists that all values, economic and non-economic, be captured, and captured explicitly. When this methodology is implemented, perhaps its greatest contribution may lie in the clarification of the real values at KSC, in such trade-offs as between obtaining assured launch success using existing (and proven) technology and developing ASW for more efficient and effective launch operations in the future.

2.3 A Process for Implementation

With this overall framework in mind, I propose that at KSC, we use the six-stage process depicted in Figure 1 for assessing various ASW investment alternatives.

Stage 1 requires that the decision context of a CBA study be articulated explicitly. That is, we must identify the decision alternatives to be evaluated in as specific terms as possible. For example, in the SCAN replatforming project (See Attachment A), evaluating the costs and benefits of the total replatforming effort does not help any decision, since in face of the obsolescence of the current platform, replatforming must be done. What we need is an assessment of the incremental costs and benefits of each enhancement sought while replatforming. We must still assess the costs and benefits of the basic (no enhancements) replatforming effort, but only to set the baseline from which the incremental costs, benefits, and risks of an enhancement can be assessed.

Figure 1 A Process for Applying CBA Methodology to ASW Investment Decisions



In addition, in this **Context Articulation Stage**, we identify all the major stakeholders of an ASW project, define the horizon (one year, or five years, etc.) over which benefits and costs will be assessed, guesstimate the upper and lower bounds on the costs and benefits of each alternative, and make decisions on which alternatives will be the subject formal CBA studies, and from which stakeholders' points of view. In other words, we make a judgment on which CBA studies would be cost-beneficial.

It is important to define a *reasonably long but limited* horizon. For example, it does not help any decision we can make today, if we assess the costs and benefits KATE assuming final completion and implementation of the total KATE vision, which is estimated to need \$27M in software engineers' time alone. At the current funding level of \$300K, it will take ninety years to realize that vision! (See Attachment B).

In the Context Articulation Stage, we should also begin to compile a list of assumptions underlying our study. In subsequent stages, we should be diligent in updating this list, as necessary.

Stage 2 requires the **enumeration** (or listing) of **all the categories of changes** resulting from an investment in an ASW alternative, both during the development of the ASW and after it is operational, but without going beyond the defined horizon. These changes may be in:

- (i) the use of resources including hardware, facilities, labor (both software engineers' time, and supporting experts' time), etc.,
- (ii) information input and output including quantity, quality, speed and timing,
- (iii) NASA's mission performance including on-schedule and safe launches, maximum productive use of available resources, being at the cutting edge of technology and providing commercial spin-offs, etc., and
- (iv) Contractor performance including profitability, productivity, etc.

We want to enumerate these changes not only in the sponsoring department (e.g., a vehicle flow manager in the case of GPSS), and the software development group, but also in the various non-sponsoring but potentially affected directorates and contractors. As suggested before, this may be important in obtaining the necessary cooperation from the experts in various affected organizations, without risking a neglect of their normal duties.

In addition to the above changes, we should also enumerate the technical, schedule and operational risks associated with an ASW project. Also, we should not forget to update the list of assumptions we began to compile in Stage 1. Indeed, as depicted by the feedback arrows in Figure 1, I visualize the six stages of this process to be overlapping, earlier stages requiring feedback and updating from later stages, and vice versa.

In short, Stage 2 ensures that we account for all costs, benefits, and risks of an ASW project, and their timings, within the defined horizon. It also ensures that immeasurable costs, benefits, and risks remain as prominent in our analysis as the

measurable ones. After all, making a decision (in Stage 6) inevitably involves a trade-off between the measured and the unmeasured.

Once the relevant changes are enumerated, it is important to identify those that defy any measurement (e.g., the quality of information), describe them as clearly and completely as possible, and determine if they are still amenable to valuation (perhaps through such approaches as the user's willingness to pay).

When feasible, measurement that occurs in Stage 3 is an important preliminary to valuation. However, even in the case of the measurable, such as the reduction in scheduling meeting durations attributable to GPSS (See Attachment C), we must have a proper historical *base-line* measurement, and the ability to project that base-line into two futures, one with GPSS implemented, and one without. Just because in the pre-GPSS environment, we used to schedule a meeting for an hour, and we used to invite 106 people to this meeting, does not mean we can use 106 hours per day as the base-line. We must examine as to how long these meetings actually used to last, and how many people used to actually attend.

If nothing else, Stage 3 tells us what data we must begin to collect, so as to track the performance improvements brought about by an ASW. In projecting the without-ASW future, it is important to not assume a simple status quo from the history. We must examine as to what other forces may be influencing the base-line. For example, experience in scheduling past Orbiter flows may also help reduce the scheduling meeting durations necessary for future flows.

Similarly, a reduction in weekend overtime, claimed as a benefit of GPSS (See Attachment C) may also be the result of a simple management policy to not approve certain types of overtime work regardless of what it does to the launch schedule, and the result of improved logistics and operations technologies in OPF. What is important is to *isolate and measure the incremental contribution* of GPSS to this reduction in overtime.

It is important in the measurement stage to identify the many co-producers (i.e., necessary conditions) a proposed ASW may need in producing a benefit. For example, to realize the savings in Firing Room manpower afforded by KATE or RUBICON (See Attachments B and D), a co-producer is the necessary cultural and attitudinal change in LCC management. When such co-producers are identified, one must estimate their probabilities of existence during each year of the defined horizon, and then in Stage 5, make the necessary adjustments to the measured or valued annual benefits, by multiplying the benefits with these probabilities. Thus, if the likelihood of a cultural change is zero, the expected benefits of manpower reduction due to KATE and RUBICON will be zero.

Another issue in the measurement of ASW project benefits is whether several projects are claiming the same benefits. For example, both KATE and RUBICON may be claiming the same reductions in the Firing Room manpower.

On the cost-estimation side, a similarly complicating issue is one of the joint use of same resources (e.g., the same computer and communications hardware) by many different projects. We need to develop a systematic method for identifying the incremental changes in these resources brought about by each ASW project.

Costs are often assumed to be easier to measure than benefits. However, in identifying exactly what costs are *incremental*, there are many issues that need to be resolved particularly in the contract management environment at KSC. If contractor compensation is based on head-count, will not the savings in direct labor on one task (brought about by an ASW) be simply "absorbed" (at least, in terms of their accounting) by some other tasks? If demonstrated savings will be accomplished only in future years through prudent contract negotiation, such a contract negotiation should be identified as a co-producer of those savings.

In Stage 3, the idea is to measure the changes in resources in their *physical units*, e.g., labor hours, CPU hours, etc. Then in Stage 4, we attempt an explicit valuation of these resource changes. Of course, we may deliberately exclude some of the resource changes from this valuation. For example, as long as the replatformed SCAN meets the desired maximum access time requirements, we may not place an explicit value on the system's actual access time. On the other hand, certain changes that could not be measured (such as better quality of information) could now be explicitly valued at least in subjective terms by the users of that information. This is possible as long as we do not insist on valuing everything in dollar terms. Thus, at least until Stage 6, some changes may be valued in dollars while others are valued on a "user satisfaction scale" of 1 to 10, etc.

Separation of valuation from measurement is critical in the multi-perspective analysis I am proposing. It allows us to recognize that different stakeholders value a given change in resources very differently. For example, from a cost-plus-fixed-fee contractor's point of view a cost saving has no positive or negative value. For an empire-building manager, the reduction in the manpower under his supervision has a negative value. If a fixed G&A pool will be collected by the contractor by the end of the year, regardless of the direct labor hours involved, should not G&A be left out of the rate NASA uses to value each labor hour saved? The proper labor rates to use in Attachments B, C, and D can be arrived at, only when issues of this sort are resolved.

For many other resources such as computer hardware or office facilities, market prices are commonly seen as an "objective" source of value. However, economists point out that market prices are not value-free; they derive from a particular income distribution and from existing institutional and legal arrangements. As such, at times it is necessary to adjust market prices to reflect specific stakeholders values. For certain benefits, such as the improved quality of decisions supported by an ASW, market prices may not be available and valuation must be imputed from the relevant stakeholder's beliefs, attitudes, and preferences.

Clearly, a number of assumptions are required in this valuation stage, and we must not forget to update our list of explicit assumptions. Sometimes, during valuation we realize that somethings we had originally decided not to measure can and need to be measured. Thus, there may be a feedback from this stage to Stage 3.

In Stage 5, the explicit values must be adjusted for the timing and uncertainty of their occurrence. It is in this **Adjustment Stage** that we must also adjust for the probabilities of existence of the co-producers of our benefits. These adjustments often require assumptions regarding discount rates and the various stakeholders' risk preferences. Thus, once again, we must update our list of assumptions. Finally, in this stage we must also conduct a sensitivity (i.e., what-if) analysis considering alternative values for the various assumptions, e. g., alternative discount rates, alternative timings of occurrence of particular events.

At the conclusion of Stage 5, the analyst's task is complete. In Stage 6, the decision-maker(s) must consider the valued and the unvalued together from each stakeholders point of view to arrive at the final assessment of an ASW alternative. Sometimes this **Final Assessment Stage** may provide a clear decision regarding the funding of the project, and sometimes it may lead to a redesign of the ASW project under consideration to make it more attractive to one or more stakeholders. In the latter case, we may have to repeat the entire process beginning with Stage 1.

2.4 Implementation Requirements and Advantages

From the many analytical issues I have identified, it should be clear that the conduct of this methodology cannot be left to the software engineers of an ASW project. The methodology must be guided by a person who is knowledgeable in the underlying philosophical, economic, and financial principles. This person would need the advice and cooperation of people familiar with contract terms and accounting systems, in addition to the advice and cooperation of the major stakeholders of an ASW project. The first time we apply this methodology, these requirements may seem prohibitively expensive and time consuming. However, once the first full study is complete, the methodology will be easy to apply to other ASW projects since a number of complicated measurement and valuation issues may be already resolved.

I think that an investment in this methodology will pay back many times over through better decision-making at KSC. As suggested in the foregoing discussion, the use of this methodology will also provide the following additional by-products:

- (i) No additional efforts needed to justify the decisions to external bodies,
- (ii) Better product designs of the ASW under consideration,
- (iii) Greater cooperation and commitment to the ASW project from the multiple stakeholders,
- (iv) Greater chance of on-schedule and successful development and implementation,
- and (v) Knowing the co-producers of our ASW's benefits may help us work on improving the probabilities of existence of those co-producers.

3. A Review of the Current CBA Practice at KSC

Before I say anything else, I must say that I appreciate the willingness of the authors of the CBAs in attachments A through D to subject their studies to a methodological review. Given that they had no background or training in the relevant philosophical and economic issues, I admire their creativity and courage in authoring these studies. I mean no harm or insult to these authors when I point out the conceptual errors in their methods. I particularly admire them for recognizing, on their own, that most of their numbers were simply wild guesses, and that the margin of error in their estimates was perhaps very large. I am most encouraged to find that these authors are highly interested in obtaining the necessary background, and in developing a better methodology for the future.

In Section 2, I have already commented on many specific conceptual issues in the studies represented in Attachments A to D. I will be happy to provide additional detailed comments and suggestions to the authors, if they so desire. However, here I want to review the overall practice of CBAs at KSC. In the light of my proposed methodology, we can observe many deficiencies in the current practice. However, two important deficiencies seem to be the root causes of the rest of them.

First, CBAs are not done to actively assist the decisions at hand. Instead, they seem to be produced for public relations (i.e., justification of past decisions), or documentation requirements (in the justification of a preferred decision). In project review meetings I observed, CBAs were often introduced casually with phrases such as "now let us see where we are going with our numbers." In other words, they are given little credibility, and practically no scrutiny.

Indeed, at KSC, I have observed instances where managers facing complex problems deliberately avoided CBAs. I believe that this practice is based on the many misperceptions of what a CBA is, and how it can assist decision-making, discussed earlier. I hope this report helps correct that misperception. At the same time, as I will explain in a minute, given the current state of CBA practice at KSC, these managers were fully justified in avoiding CBAs.

Second, the conduct of CBAs is left to the initiative of software engineers who have little background, training, or assistance in the pertinent methodology. Thus, each study seems ad hoc, developing its own methods and concepts. Indeed one engineer suggested that it was KSC's standard operating procedure "to build a brand new road every time we want to go to Orlando!"

Each one of the available studies seems to violate one or more of the fundamental principles of the CBA methodology. None of the studies I examined tried to capture *all* the costs and benefits, as is required by the methodology. None of them made *all* of their underlying assumptions explicit, or estimate probabilities that the explicit assumptions will be valid. Most studies did not seem to use proper base-lines or proper projection methods in the measurement of their costs and benefits. They failed to separate

measurement from valuation, and to address the many issues of valuation from the perspective of the multiple stakeholders. Even the more commonly understood practices of the CBA methodology, such as adjusting for time value of money in a multi-year stream of costs and benefits, were not used in the CBAs at KSC.

In short, the current practice is seriously deficient.

Speaking as a professor, I am sorry, but I must assign an F grade to this practice. At the same time, I must add that despite this team grade, most individuals who are involved in the current practice of CBAs get unqualified A grades. These individuals have been doing their parts sincerely and to the best of their abilities. They have also been very cooperative and candid with me and open to my ideas. As will be clear below, the deficiencies of the current practice are not the fault of any individual.

3.1 The Vicious Circle

As I think about the two root causes of deficiencies together, I have come to realize that we are caught in a vicious circle which can be described as below:

- Available CBA studies measure only selected (not all) changes brought about by the development and implementation of a given ASW. At times, they force quantification of the non-quantifiable, or prediction of the unpredictable. The baselines used in the measurement are often incorrect. Measurement (in physical units) is not separated from valuation. Valuation is from a single (as against each stakeholder's separate) point of view. Values are not adjusted for their probabilities or timing of occurrence. Sensitivity analysis is not done. In short, many principles of the CBA methodology are violated.

As a result,

- The focus of the CBA studies is primarily on the quantifiable. Very important but non-measurable costs, benefits, and risks are left out. The margin of errors in the quantified estimates is very large. The real values of the Agency mission, the values of senior managers, the values of the contractors, etc., are not captured by the analysis.

Then,

- Because CBA s do not capture and address the real values and issues, and because the studies' estimates are unreliable, Management looks at CBAs not as decision-making tools, but as mere exercises in generating numbers for external justification of decisions already made.

Thus,

- Management allocates few resources, and leaves the conduct of CBAs up to the initiative of the software engineers involved in specific projects.

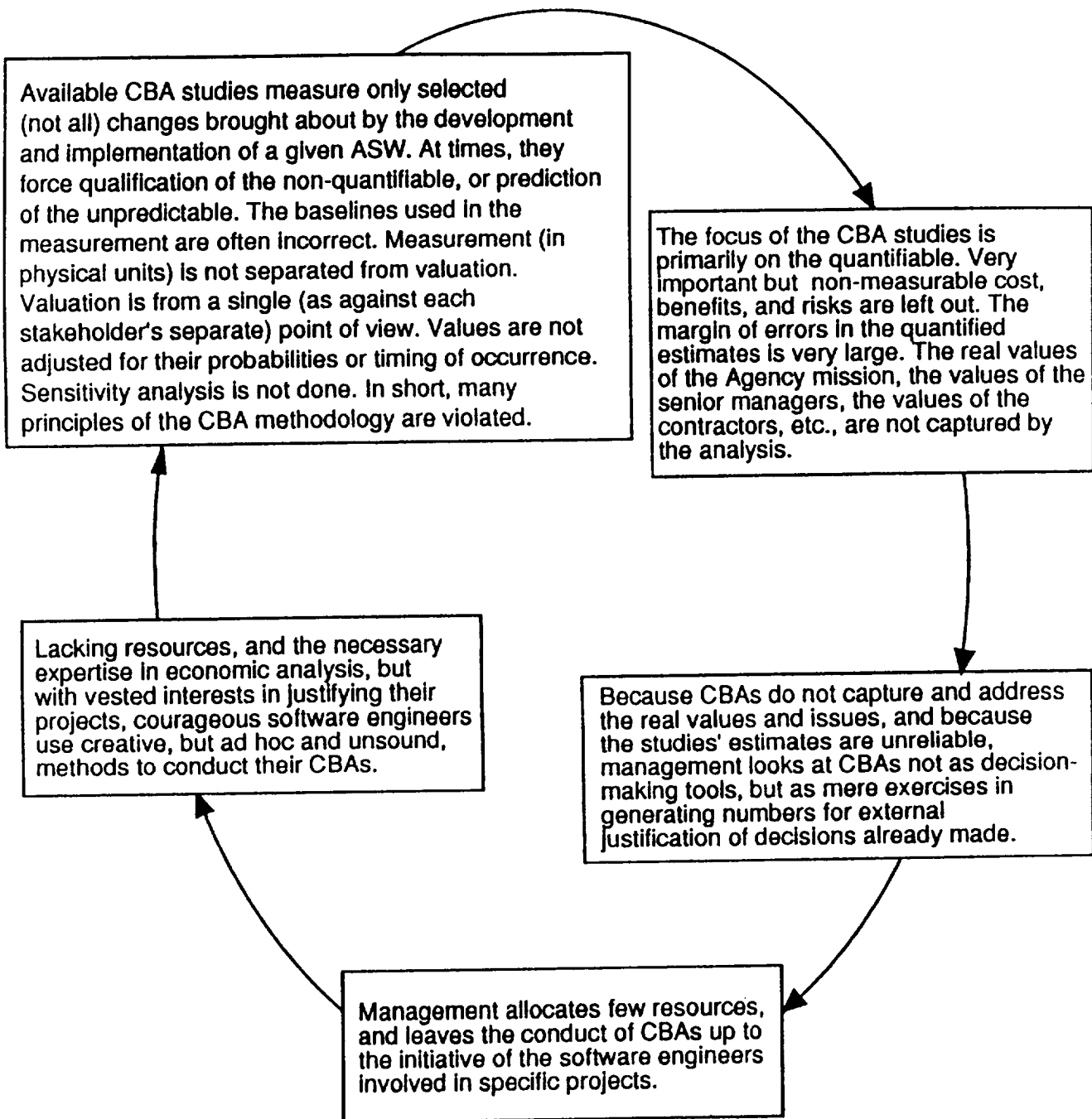
Next,

- **Lacking resources, and the necessary expertise in economic analysis, but with vested interests in justifying their projects, courageous software engineers use creative, but ad hoc and unsound, methods to conduct their CBAs.**

But this results exactly in the situation described in the starting bullet of this process, and the vicious circle continues!

Figure 2 depicts this vicious circle graphically.

Figure 2.
The Vicious Circle



4. Conclusion and Recommendations

I have argued that ASW investment decisions are multi-stage, varied, complex, risky, and controversial. Therefore, we need a systematic methodology to assist rational ASW investment decisions. I proposed a formal, explicit, and multi-perspective cost-benefit analysis (CBA) methodology for this purpose. I outlined a number of rich concepts and principles of this methodology, and described a six-stage process for its implementation. In the light of this methodology, we reviewed the current practice of CBAs at KSC.

Although I have concluded that current practice is seriously deficient, I believe that most NASA employees already knew that, and many are looking forward to improving that practice. I think my principal contribution is the identification of the vicious circle we are in, and consequently, my primary recommendation is:

Break out of that vicious circle.

The methodology I have proposed provides one exit point to break out of this circle. The other exit point is a change in management's perception of what a good methodology can do, and its willingness to provide adequate resources and appropriate expertise to the conduct of CBAs.

References

1. Toraskar, Kranti, and Joglekar, Prafulla, "Applying Cost-Benefit Analysis (CBA) Methodology for Information Technology Investment Decisions," in R. D. Banker, R. J. Kauffman, and M. A. Mahmood (Eds.): *New Perspectives on the Strategic and Economic Value of Information Technology Investments*, Idea Publishing Group, Middletown PA., December, 1992.
2. Gross, A. M., "Is Cost-benefit Analysis Beneficial? Is Cost-effectiveness analysis effective?" *The Heller School for Advanced Studies in Social Welfare*, Brandeis University, 1976. Distributed by the National Technical Information Service (NITS).

Excerpts from a CBA of SCAN Replatforming

5. Costs and Benefits

5.1 Benefits

The benefits for the replatforming of SCAN are difficult to quantify because they are primarily associated with mandated changes to LSDN. The planned migration toward the OSF operating system will make a number of the current SCAN software components (i.e. D3M, Dialog) obsolete which means that replatforming is required. However, with the necessity for replatforming comes the opportunity for a number of significant improvements to the system that will be the main benefits to the user community. The main benefits to be realized by the new system are as follows.

No required resync; engineers will be able to perform useful work immediately after accessing the system. Also, no garbage collection and no checkpoint operations as required by the LISP language.

A single on-line database which will make all data available to the users at all times and elimination of the need to reconstruct a new KB for each mission. Elimination of KB builds will also mean that system engineers will have access to modified circuitry as soon as EO modifications are entered.

Elimination of unused reports and replacement with reports that are more in line with the needs of the user Community.

5.2 Cost Breakdowns

Initial estimates indicate that SCAN replatforming will be an expensive undertaking. There are a number of key issues which must be resolved and the volume of functionality make the replatforming a non-trivial task. The major element of cost will be associated with software development, although some costs will be accrued by implementation as well as revised running costs. Cost estimates are defined in the following sections using CASE Methods estimation tools in conjunction with the best data available at this time.

5.2.1 Development Cost Estimates

The development cost of replatforming SCAN are best understood in terms of the four primary development stages described by CASE Methods. These stages are Analysis, Design, Build and Documentation. Each of these stages are described below with manpower estimates based on the best information available at this time.

The Analysis stage will verify the findings from the Strategy Stage and expand them into sufficient detail to ensure system accuracy, feasibility and a sound foundation for design.

The Design Stage will take the detailed requirements from the Analysis Stage and find the best way to fulfill them and achieve agreed service levels, given the technical environment and previous decisions on required levels of automation.

The Build stage will code and test programs, using appropriate tools. These depend on the technical environment and types of programs involved.

The Documentation Stage will deliver user manuals and operations hand-over documentation, which must be sufficient to support the system testing tasks in the concurrent build stage.

The current manpower estimates associated with the described tasks are as follows (calendar weeks equals total man-weeks divided by 4.2; current manning level).

See APPENDIX A.3 Manpower Analysis Summary, for details.

Analysis Stage	Total man-days: <u>475</u>
	Total man-weeks: <u>95</u>
	Calendar weeks: <u>23</u>
Design Stage	Total man-days: <u>696</u>
	Total man-weeks: <u>139</u>
	Calendar weeks: <u>33</u>
Build Stage	Total man-days: <u>715</u>
	Total man-weeks: <u>143</u>
	Calendar weeks: <u>34</u>
Documentation Stage	Total man-days: <u>18</u>
	Total man-weeks: <u>4</u>
	Calendar weeks: <u>1</u>

The primary impact of this analysis is an indication that the replatforming may not be achievable per the original schedule at the current manning level. Current manning levels indicate that a February 1994 completion date is a more reasonable estimate. To meet the planned schedule of a July 1993 completion date would require increased manpower as follows:

Stage	Required Manpower
Analysis	5.6
Design	4.6
Build-Doc	8.2

See Appendix A.4, Proposed Development Schedule and Milestones, for scheduling implications of the estimated manpower requirements.

5.2.2 Implementation Estimates

Implementation estimates will depend on hardware decisions which have not been made at this time. Implementation costs will also be associated with the transition methodology discussed in section 4.3.3.

5.2.3 Running Cost Estimates

Running cost estimates should be reduced significantly because of the elimination of the Knowledge Base build tasks currently performed by Data Bank. Running costs will also be modified by the maintenance level required on any new hardware. Running costs may be increased by additional backup requirements imposed on LSDN, which may be required to reliably support the planned RDBMS.

While all run cost data is not available at this time, it is expected that the replatformed system will present a net decrease in operational running costs.

5.3 Cost/Benefit Analysis

The analysis of costs versus benefits is virtually impossible in the context of SCAN platforming, for a number of reasons. The Replatforming is mandated by software obsolescence, the benefits to be gained are not easily quantifiable and some of the costs are unknown at this time. The best that can be stated is that the replatforming will be expensive as indicated by the estimates in section 5.2.1. To keep these costs in perspective, it must be remembered that it has taken more than five years for the SCAN system to achieve its current level of functionality, including effort associated with approximately 300 Problem Reports. The costs included in these estimates are a recognition that SCAN is still a complex system and they also represent a commitment to field a replatformed system, coded correctly the first time and ready for enhancements that will continue the reduce the System Engineering workload.

Costs Benefits Analysis for the Deployment of the Knowledge Based Autonomous Test Engineer (KATE)

INTRODUCTION

KATE is a tool for health monitoring of electromechanical and fluid systems which is capable of detecting, isolating and diagnosing failure causes. Capabilities exist for automatic correction of failure is , but is currently not deployed for use with Shuttle. Current work includes development of a C++ version of KATE for use by the Firing Room Integration console engineer's Vehicle Health Management System (VHMS).

What follows is a cost benefits analysis for the complete deployment of KATE into the Firing Room environment for both launch and day to day operations. As with all forecasts, several caveats and assumptions were incorporated into this analysis and are stated below. Since the development and deployment of intelligent process control and monitoring systems are in their infancy at KSC much of the cost data needed for a detailed cost benefits analysis either does not exist or is in a format/environment in which it is difficult to obtain. With that stated, this analysis for KATE must be viewed as being, in it's essence, a qualitative and speculative study of costs and benefits of the KATE system. In areas where it was felt sufficient cost data existed quantitative Dollar values were extrapolated from this data. For those areas where insufficient cost data existed a quantitative estimate was derived from plausibility arguments. The derivation of all the cost benefit values have been left outside of the main body of this text and reside as appendices at the end of this report.

ASSUMPTIONS

1. This analysis does not take into account the validation costs for KATE deployment. this is due to the fact that at the present time no decision has been made concerning the methodology to be used in validating non-GOAL, Firing Room resident applications.
2. This analysis does not take into account any transition costs other than estimates for training costs included into assumption # 5.
3. Since at the current time only one Shuttle system has been implemented in KATE (KATE-LOX) a linear extrapolation of costs is estimated for an expansion in the economy of scale from one application to a firing room wide system.
4. A measure of modelling complexity for a Particular Shuttle system is defined as the number of Function Designators (FDs) associated with that system.
5. A measure of model development time is given as 8 hr/FD. This assumption is based on experience and assumes a KATE neophyte programmer (i.e. this includes the time necessary to learn and model in KATE).
6. The labor rate is defined as 40.00 \$/hr, (approximate LSOC rate).

SUMMARY OF ANALYSIS

IMMEDIATE BENEFITS (< 3 yrs)

1. KATE can draw conclusions on system behavior, both current and future, based on analysis of measurement input. Upon operational activation, for the integration console alone savings of @ \$ 150,000 per vehicle per flow may be realized due to three less support engineers being required for daily monitoring operations (non-active testing). This figure is based on 3 on-station engineers working 3 shifts per day at 60 days per flow. Total cost savings, based on six flows per year, would equate to \$ 900,000 per year.

POTENTIAL BENEFITS (>3 yrs)

2. KATE represents one analytical tool which can be used for multiple subsystems by changing the knowledge base used for reasoning and analysis. Cost savings are in reduced system development time since only the knowledge base for each class of system needs to be developed. Life cycle cost savings are in sustaining engineering, since the same reasoning software is used for all subsystems.

2.1 Estimated Cost Expenditures

The costs in developing a total KATE Firing Room system (including knowledge bases for each Shuttle system as well as the KATE shell) is estimated to be @ \$ 30 Million.

Note that the capabilities currently residing in GOAL pertain only to control and monitoring capabilities, whereas KATE would have these with the addition of diagnostic capabilities.

2.2 Estimated Cost Savings

Sustaining engineering costs for a fully deployed KATE system, on a per year basis are estimated to be @ \$ 5.67 Million.

Note that this sustaining engineering cost estimate is approximately less than the estimated costs for sustaining the current complement of GOAL software (see appendix). The KATE sustaining engineering cost savings for maintaining ground software that may be realized are estimated to be @ \$ 2.33 Million .

Further cost savings are realized by allowing a reduction in the amount of Shuttle system engineering labor required to perform day-to-day Shuttle monitoring and maintenance operations. Following other cost benefit studies done concerning advisory systems (see appendix B) a 13 % reduction in man-power may be realizable without impairing safety. This savings equates to approximately a savings of \$ 5.62 Million per year.

The total costs savings that may be realized on a yearly basis is estimated to be @ \$ 7.95 Million per year.

Other Benefits

3. The same KATE/system version can be used for real time data analysis or as a simulation tool for training and off-line system evaluation. Costs savings are realized by deleting the necessity of maintaining separate software for operations and training activities as is currently done.
4. Reduction in the size of launch team for all other consoles based on the same rationale as used in #1 \$ #2.2 above.

ORIGINAL PAGE IS
OF POOR QUALITY

APPENDIX A

CONSOLE	FD	8 hr/fd	RATE	COST/CONSOLE
C2	8000	64,000 hrs	@ 40.00 \$/hr	\$ 2.56 M
C3	5600	44,800	—	1.79 M
C4	5600	44,800	—	1.79 M
C5	2400	19,200	—	0.77 M
C6	6500	52,000	—	2.08 M
C7	5100	40,800	—	1.63 M
C8	5900	47,200	—	1.89 M
C9	6400	51,200	—	2.05 M
C10	4200	33,600	—	1.72 M
INTG	2000	16,000	—	0.64 M
C11	7000	112,000 ¹	—	4.48 M
C12	10,000	160,000 ¹	—	6.40 M
TOTAL				\$ 27.80 M

ESTIMATED COSTS FOR KNOWLEDGE BASE DEVELOPMENT

1. This estimate was doubled to take into account the uncertainty in developing KATE applications involving high-speed digital systems.

CONSOLE	FD	SIZE
		80,000 LINES
C2	8000 X 10 LINES C++ FD	56,000
C3	5600	56,000
C4	5600	24,000
C5	2400	65,000
C6	6500	51,000
C7	5100	59,000
C8	5900	64,000
C9	6400	42,000
C10	4200	20,000
INTG	2000	70,000
C11	7000	100,000
C12	10,000	
TOTAL	68,700	4,081 K

KATE with control : Estimates based on experience with the KATE-ALO system suggest a 15 % increase in the amount of code needed to realize control procedures. :

$4,081 \text{ K} + 15 \% = 4,693 \text{ K lines}$

ESTIMATED SIZE OF TOTAL KATE C++ APPLICATION

SUSTAINING ENGINEERING

GOAL ESTIMATES

7 MILLION LINES OF GOAL CODE (⊙)

100 S/W ENGINEERS (⊙)

LSOC LABOR RATE = 40 \$/hr

ONE MAN YEAR = 2000 hr

SUSTAINING ENGINEERING COSTS ESTIMATE = 2000 hrs X 40 \$/hr X 100 = \$ 8,000,000 \$/Yr

~~8,000,000 \$~~
7,000,000 lines = 1.143 \$/line

KATE ESTIMATE

4,963,000 lines X 1.143 \$/line = 5,670,000 \$/Yr

POTENTIAL NET SAVINGS

8,000,000 \$/YR - 5,670,000 \$/YR = 2,330,000 \$/YR

Excerpts from a CBA of GPSS

COST SAVINGS

1. The daily scheduling meetings were considerably shortened. If the daily scheduling meeting would have been held using the non-AI schedule, approximately 83 hours of meeting time would have been expended for the entire flow. However, because of the use of the AI based scheduler, only 42.03 hours were spent, yielding a savings of about 40.23 hours. Based on a rate of \$34/hr for each of the 106 engineers that attend the meetings, the use of the GPSS schedule resulted in a savings of about \$144,988 (See attachment 1) for the engineers' time along with a savings of about \$12,000 for the person who used to physically "lay tape" for the paper schedules.
 2. Reduction/elimination of weekend overtime. Weekend overtime is expensive but often necessary to prevent even more costly delays. TPS (Thermal Protection Systems) technicians utilized the GPSS scheduler to predict required overtime during this flow based on vehicle configuration. Normally during a 16 week flow TPS spent \$481,950 on overtime. By utilizing the GPSS scheduler \$110,129 was actually spent for technician support. This results in a cost savings of \$371,821 (See attachment 2). The savings was accrued because GPSS was able to predict weekend overtime and reduce or eliminate in some instances due to better forecasting of conflicts.
- Total cost savings for STS-50, OV-102 are estimated at \$528,809.

RUBICON COST ANALYSIS STUDY

I. TASK DESCRIPTION

Evaluate the RUBICON concept from all aspects to determine if the continued development and subsequent implementation into CCMS 2 as well as the office environment will provide a cost pay back for shuttle operations.

II. ASSUMPTIONS

A. RUBICON development will continue in the direction described below.

- o The DLES displays and code will be used as the RUBICON DPS system monitor. LCC resolution, MDT, DEU dump analyzer and other analysis applications will run under the system monitor.

The following items are planned for implementation in FY93:

- o DLES must be modified to allow multiple applications to run on the same machine and then will become RUBICON.
- o The MDT CLIPS portion will be incorporated to run under this new RUBICON system.
- o Where feasible, the software developed by Rockwell will be utilized (ex. DPA keystrokes).
- o RUBICON must be converted to run under MOTIF.

Future plans/capabilities will depend on the direction system engineering decides to meet suit testing needs.

- B. The management issues for allowing vehicle monitoring from a remote location (outside the Firing Room) will be worked.
- C. The transmission of data from all 4 vehicles on a single network will be completed and validated. This is currently scheduled to be complete around the mid-1993 timeframe.
- D. The ethernet network used for the transmission of vehicle data and workstation to workstation communication will be fully maintained and supported.

RUBICON COST ANALYSIS STUDY

III. POSITIVE COST SAVINGS

A. Optimization of manpower resources

Description: RUBICON can be used either in the office or in one PR to monitor all 4 vehicles during times of vehicle power-up MONITOR ONLY support. It is estimated that DPS is in a monitor only mode approximately 85% of the time the vehicle is powered up.

Est Savings: \$274,000 (4 eng x \$33/hr x 40 hr/wk x 52 wks). DPS currently supports vehicle power up periods 3 shifts a day with 4 engineers on first shift, 4 on second shift and 2 on third shift for a total of up to 10 people a day. RUBICON can be utilized as a method for survivability by DPS hardware engineers to continue vehicle support when shuttle budget cutbacks impact the number of system engineers in the group. The number of engineers supporting vehicle testing could be reduced to 2 people per shift for a total of 6 people a day without impacting vehicle testing.

B. Avoid opening unnecessary IPR's

Description: Anomalous conditions that have been seen before and documented on an IPR can happen again at a later date. With no IPR/PR history data readily available, IPR's can be opened only to be researched long enough to find out that the problem was seen before and is an explained condition. The engineer must then close the IPR as an explained condition.

Est Savings: \$18,000/yr
There has been an average of approximately 16 explained condition IPR's opened per year (averaged over the last 12 years). It is estimated that out of those 16 IPR's that RUBICON could have potentially avoided approximately 9. These problems were either addressed on previous IPR's or the data in the RUBICON database could have helped understand the problem and avoid opening an IPR. The estimated cost to open and close an IPR is \$2,000 (This figure does not include the time required to investigate the problem).

C. Potential avoidance of an unnecessary launch scrub

Description: If the time remaining in the launch window is short, by providing a quick, precise explanation or work-around procedure for an LCC violation (one that can

ORIGINAL COPY IS
OF POOR QUALITY

be rationalized as being OK to launch given certain conditions are met) that would allow the count to resume quickly you could preclude an unnecessary launch scrub.

Est Savings: A minimum \$1 million savings would be realized.

D. Tradeoff between GOAL versus RUBICON maintenance for DPS System display monitoring.

Description: Once RUBICON is incorporated into C12 application software in CCMS 2, approximately 12 GOAL display programs will no longer be required. Thus, the maintenance manpower required on the GOAL software can be redirected to maintain the RUBICON system monitor software.

Est Savings: No cost savings but no additional cost incurred.

E. Improved training for new hires

Description: When new hires are brought on board they currently must go through extensive training. With the record and playback and debugger capabilities, engineers can be trained using actual vehicle data to get an understanding of how the DPS system works. Failures can also be inserted to test reactions to problems. This capability is supported totally separate from the CCMS and thus avoiding scheduling conflicts and reducing the impact on other systems.

On the software development side, most computer related degrees require C as a programming language. New hires can be more productive in a much shorter timeframe.

Est Savings: Savings is hard to determine but has the potential be a significant amount.

NI

F. Office/Firing Room tool to reduce the time it takes to find historical information.

Description: IPR/PR historical data and the PMW must often be researched to support anything from general management questions/concerns to troubleshooting problems. By searching a quick access database for the required information you can reduce the manpower required to provide the necessary information.

Est Savings: \$13,000/year
Estimated 15 system engineers spending .5 hours per

week searching documents that reside in the RUBICON database. The time required to search the PNW on a specific topic for example can be reduced by as much as 20 minutes.

0. "Path finder" for new approaches and capabilities for CCMS 2.

Description: New approaches to monitoring and analysis can be tested and evaluated. If more efficient techniques can be developed and proven in this environment they could be utilized faster and easier in the CCMS 2 environment. Conversely, techniques that are tested but do not work can be noted and the same mistakes could be avoided in the CCMS 2 environment.

Est Savings: CCMS 2 will require a massive learning curve. A side benefit of this and other similar projects will be a faster understanding of new approaches to system health monitoring.

1. "Path finder" for new approaches and capabilities for future launch vehicles.

Description: New approaches to monitoring and analysis can be tested and evaluated. If more efficient techniques can be developed and proven in this environment they could be utilized faster and easier in a future launch vehicle environment. Conversely, techniques that are tested but do not work can be noted thus repeating the same mistakes could be avoided in the new environment.

Est Savings: Future launch vehicles will require a significant learning curve. A side benefit of this and other similar projects will be a faster understanding of new approaches to system health monitoring utilizing "state of the art" hardware/software. Portable software modules can be reused with some minor modifications, greatly reducing software development time/costs.

GENERAL ISSUE OF POOR QUALITY

RUBICON COST ANALYSIS STUDY

IV. NEGATIVE COST SAVINGS

A. Dual maintenance of GOAL displays and RUBICON System Monitor.

Description: During the timeframe from October 1993 until CCMS 2 is operational, there will be maintenance required on both the GOAL software as well as the System monitor portion of RUBICON.

Cost Impact: \$4,500/year
Will require approximately 50% manhour increase on mandatory design center change packages. There were 26 mandatory change drivers (that impacted the 12 GOAL display programs that DLES can replace) over the past 2 years for a total of approximately 550 manhours. This figure does not include the LCC monitor.

B. Additional maintenance required to maintain CLIPS rules.

Description: An additional .5 engineer would be required to maintain the expert system portion. (Note: This includes maintenance on the LCC, MDT and any other CLIPS module.)

Cost Impact: \$35,000/year
Cost is more than absorbed in the reduction of the number of system engineers required.

C. Additional maintenance required for database.

Description: There will be a small increase in manpower required to maintain the database. However, automation routines will make this task a simple procedure. Database routines can be run while other tasks are performed.

Cost Impact: \$2,100/year (8 flows/year x 8 hr/flow x \$33/hr)
It will take an estimated 8 hours per flow to maintain the database. With documents on line there would be less need to manually update the paper version of the documents. A reduction in document distribution can also be realized as well as reducing the amount of paper used.

D. Additional work required to set up additional CM tracking procedures.

Description: The initial development of rules/guidelines will be a minimal impact and is already in work. This would be a one time impact that could (like all other processes) require periodic modification.

Cost Impact: No RUBICON cost impact. The operational impact is unknown but should not require a significant amount of additional work. (Note: This task will be done for other systems that are being developed anyway.)

E. Office hardware maintenance costs.

Description: The UNIX machines utilized to run RUBICON outside of the CCMS-2 environment would require maintenance from time to time.

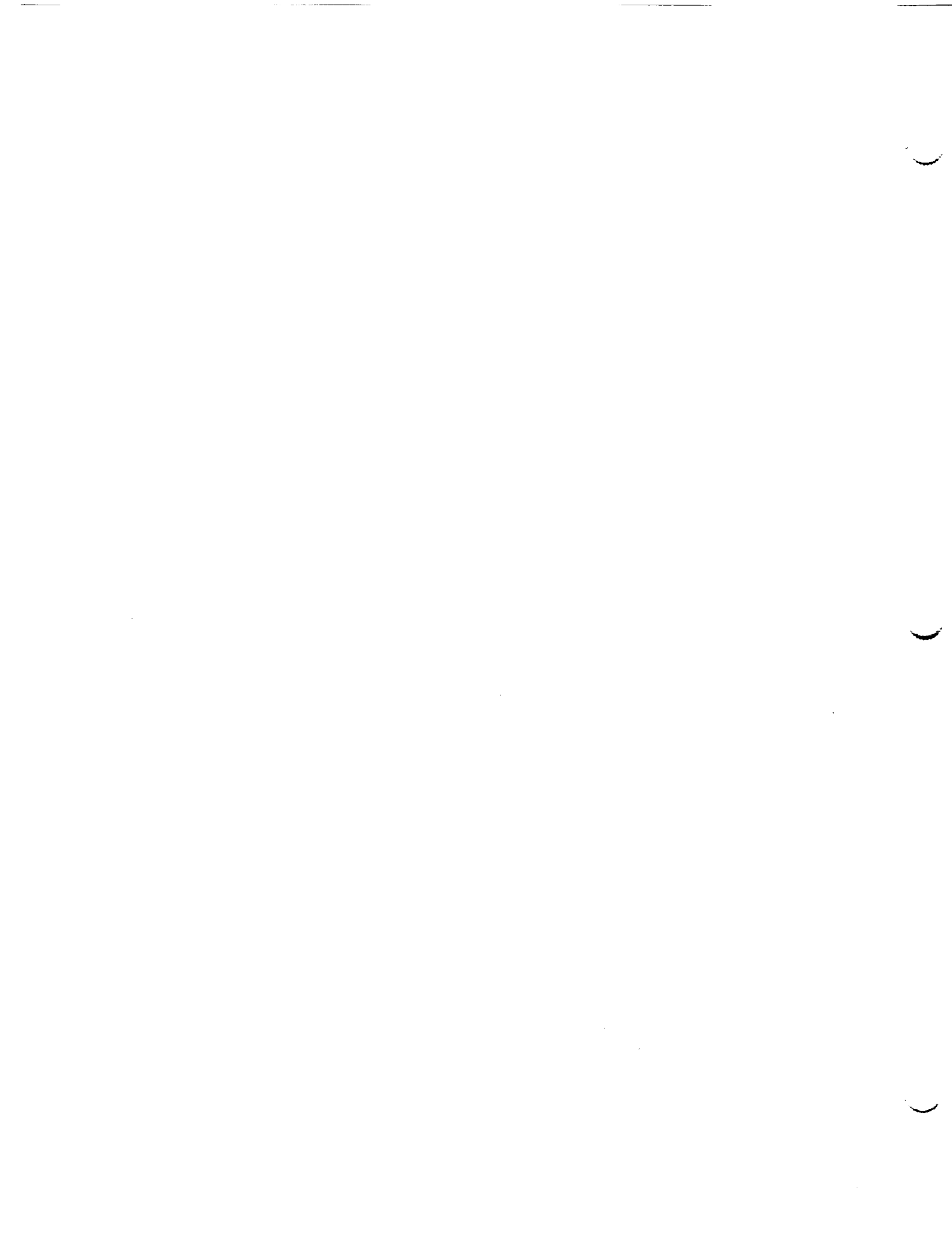
Cost Impact: Exact cost is not known at this time. The actual dollar amount will be insignificant since there is a large scale maintenance contract on the Apollo's. To date, no repair costs have been incurred (this covers a period of approximately 2 years). Plans are in work to fold the maintenance of these workstations under the LSDN maintenance plan. This will provide quick turnaround on hardware problems. Exact cost is not known at this time.

F. CLIPS validation.

Description: The validation of the CLIPS portion of RUBICON is under review. No data exists on this at this time.

Cost Impact: Exact cost is not known at this time.

ORIGINAL PAGE IS
OF POOR QUALITY



1994018768
N94-23241

1993 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA

511-51
197197
p. 21

CHEMICAL CHARACTERIZATION OF SOME AEROBIC LIQUIDS IN "CELSS"

PREPARED BY:	Dr. Brooks C. Madsen
ACADEMIC RANK:	Professor
UNIVERSITY AND DEPARTMENT:	University of Central Florida Chemistry Department
NASA/KSC	
DIVISION:	Biomedical Operations and Research
BRANCH:	Biological Research and Life Support
NASA COLLEAGUE:	John C. Sager
DATE:	July 16, 1993
CONTRACT NUMBER:	University of Central Florida NASA-NGT-60002 Supplement: 11

ACKNOWLEDGEMENTS

Interaction with several NASA and Bionetics Corporation personnel who are part of the Biological Research and Life Sciences Office was vital to successes achieved during the ten week period. John Sager, Bill Knott and Ross Hinkle were very supportive at the administrative level and made it possible for me to spend a second rewarding and refreshing summer. Conversations and discussions with Jay Garland, Dick Strayer, Barry Finger and Ray Wheeler were useful from a technical standpoint.

Special thanks are given to Barbara Peterson who for a second summer willingly relinquished a corner of the Analytical Laboratory to me. Opportunities to interact with former students are always special and the addition of Jennifer Batten to Barb's laboratory has helped to strengthen the high quality of work produced by this group.

ABSTRACT

Untreated aqueous soybean and wheat leachate and aerobically treated wheat leachate prepared from crop residues that are produced as a component of the Controlled Ecological Life Support System program designed to support long duration space missions have been compared and general chemical characterization has been accomplished. Solid phase extraction and high performance liquid chromatography were used to accomplish comparisons based on chromatographic and ultraviolet absorption properties of the components that are present. Specific compounds were not identified, however, general composition related to the initial presence of phenol-like compounds and their disappearance during aerobic treatment was explored.

SUMMARY

Aqueous leachate samples prepared from soybean and wheat crop residues that are produced as a component of the Controlled Ecological Life Support System program designed to support long duration space missions have been compared and general chemical characterization has been accomplished. Solid phase extraction and high performance liquid chromatography were used to accomplish comparisons based on chromatographic and ultraviolet absorption properties of the components that are present. Specific compounds were not identified, however, several compounds known to be present in these plant residues were used for comparison purposes. General composition related to the presence of phenol-like compounds was explored. The raw leachates and aerobically treated leachates are highly colored and aqueous fractions of the leachates exhibited chromatographic properties similar to flavonoid and tannin-like materials. Aerobic treatment of leachate samples removes or converts most of the discrete phenol-like constituents, however, flavonoid or tannin-like constituents remain after aerobic treatment. An exhaustive characterization is not a feasible undertaking because of the complexity associated with leachate composition.

TABLE OF CONTENTS

Section	Title
I.	INTRODUCTION
II.	PROCEDURES
2.1	Chemicals
2.2	Instrumentation
2.3	Leachates
2.4	Solid Phase Extraction
III.	RESULTS AND DISCUSSION
3.1	Representative Compounds
3.2	Leachate Solid Phase Extraction and Chromatography
3.3	Aerobically Treated Liquid
IV.	SUMMARY COMMENTS
V.	REFERENCES

LIST OF ILLUSTRATIONS

Figure	Title
1	Structures of Some Representative Compounds.
2	HPLC Chromatograms of some Representative Compound Mixtures Measured at 280 nm.
3	Ultraviolet Absorption Spectra for Selected Representative Compounds.
4	Chromatograms of a Wheat Leachate SPE Fraction Obtained at Several Wavelengths.
5	Chromatograms for Comparison of Fractions from SPE -C ₁₈ Treatment of a Soybean Leachate Sample Measured at 280 nm.
6	Chromatograms for Comparison of Subfraction from a Soybean Leachate Fraction (see Figure A, fE _B) Partitioned on SPE -SAX and Measured at 280 nm.
7	Chromatograms for Comparison of SPE -CH Fractions of Soybean Leachate measured at 280 nm.
8	Chromatograms for Comparison of SPE -SAX Fractions of Wheat Leachate Measured at 280 nm.
9	Chromatograms of SPE Fractions of 9.6 ml Wheat Leachate Obtained from 1000 mg Altex -C ₁₈ Measured at 280 nm.
10	Chromatograms for -SAX subfractions of Fraction f2 Wheat Leachate (see Figure F).
11	Chromatograms for -SAX subfractions of Fraction fE _{0.3c} Wheat Leachate (see Figure F).
12	Comparison Chromatograms of SPE -C ₁₈ Fractions of Wheat Leachate Obtained During and After Bioreactor Feeding.
13	Chromatograms of Tannic Acid Sample Fractions Partitioned on -SAX.

LIST OF TABLES

Table	Title
1	Conditions Used for Solid Phase Extraction of Soybean and Wheat Leachates.
2	Representative Compounds Used in Chromatographic Studies.
3	Solid Phase Extraction Partitioning Efficiency of Selected Representative Compounds on Four Different Extraction Cartridges.

I. Introduction

The Controlled Ecological Life Support Systems (CELSS) program has as a main objective the construction of a life support system that will provide food, water and oxygen for long duration space missions and will allow for recycling of some of the wastes produced (1-4). An ongoing project involves growing plants under hydroponic conditions. The plants of current focus include wheat, soybeans, lettuce and potatoes. It is desirable to recycle non-edible portions of the plants to minimize waste product buildup and to minimize the need for additional materials to support the plant growth cycle.

One current problem involves the characterization of crop residues, specifically the root, stem and chaff portions of these plants. The water soluble leachate of these dried residues contain both inorganic and organic constituents. The inorganic constituents have been characterized and can be recycled as fertilizer in the growth chambers (5). It is not unusual for a natural product material to contain in excess of 100 different chemical species. Specific fractions might include carbohydrates, organic acids, polyphenols, tannins and humic materials. A variety of analytical techniques are typically employed to derive solutions to complex chemical problems of this nature.

This report describes an extension of a previous study (6) performed to obtain a basic understanding of the chemical makeup of the organic composition of water leachate solutions derived from crop residues. A specific task was to investigate the composition of the liquid fraction taken from an aerobic reactor that is designed to decrease the total carbon load of the leachate (Strayer and Finger, private communication 1993). A brown color persists in the bioreactor and increases in intensity as the leachate is used to replenish the hydroponic solution. The yellow-brown color associated with aqueous systems that contain dissolved organic matter is typically associated with the presence of tannic and humic acids. Color can also be associated with the presence of flavonoid compounds. Anthocyanins, flavonols and flavones constitute this flavonoid group of compounds. The anthocyanins are intensely colored and account for the brilliant colors associated with many flowers. Substituted flavones and flavonols tend to exhibit yellow-brown colors that deepen as concentration increases. Preliminary investigations have shown that the organic composition of the untreated leachate can be subdivided into broad categories with complex chemical makeup (6). Potentially useful techniques for characterization of the leachate include liquid chromatography, spectrophotometry, gas chromatography and mass spectrometry. The utility of High Performance Liquid Chromatography (HPLC) in the chemical characterization of plant products can be demonstrated by studies such as reported by Lee, et al (7) where separation, identification and quantification of sugars, non-volatile acids, flavones, flavone glycosides, carotenoids and anthocyanidins were

8

reported. A photodiode array detector was used to characterize spectral properties of each class of compounds. Several recent reports use HPLC and UV spectrophotometry to characterize the flavonoid and/or phenolic composition of various plants (8-15). Methods of analysis for tannins have been the subject of a recent review (16) and HPLC methods for tannin determination have been published (17-19). The chemical makeup of tannins is extremely complex. Broad groups of water soluble compounds include hexahydroxydiphenic acids, gallic acid esters with molecular weights less than 3000, proanthocyanidins with molecular weights less than 20,000 and additional compounds with molecular weights near 200 are also present.

In the previous study (6) leachate fractions and subfractions from potato, soybean and wheat were compared by HPLC to determine if major similarities or differences exist. The present study has focused on wheat leachate both before and after aerobic treatment and suggests that phytotoxic compounds of simple phenolic structure have been removed by the aerobic treatment, however a tannin-like or flavonoid-like mixture persists that accounts for the brown color.

II. Experimental

2.1 CHEMICALS

Acid and phenol derivative samples were purchased from Aldrich Chemicals with the exception of Naringin, Coumarin 7-hydroxycoumarin and 4-methyl-7-hydroxycoumarin and Morin which were purchased from J. T. Baker and Quercetin which was purchased from Eastman Chemicals. Methanol and acetonitrile were HPLC grade purchased from Fisher or J. T. Baker. Water was obtained from a Hydrodyne Ultrapure/Picopure purification system which provided 18 megohm water filtered through a 0.45 μ filter. All HPLC mobile phase solvents were filtered through Nylon 66 0.45 μ filters prior to use. All aqueous based mobile phases were prepared fresh at three to four day intervals. Standard solutions of the various analytes were prepared in methanol at concentrations of 200-1500 ppm. All analyte solutions at these concentrations were stable and were stored at 4 °C as a precaution.

2.2 INSTRUMENTATION

Routine HPLC was performed using a Perkin-Elmer HPLC that included a PE Series 4 LC Microprocessor Controlled Solvent Delivery System, Rheodyne 7125-075 rotary injection valve with 50 μ l sample loop, reverse phase column and LC 85B Variable Wavelength Spectrophotometric Detector. Chromatograms were obtained using a LCI-100 Laboratory Computing Integrator. Typical injection volumes of 25 μ l were used in most applications and were delivered to the injector using a 100 μ l syringe. Full scale absorbance of 5.0 units (5.0 AUFS) was provided from the LC 85B. An attenuation of 1024 on the Computing Integrator is required to display the entire 5.0 unit absorbance scale on the integrator printout. A typical attenuation setting of 16(0.078 AUFS) was determined to be acceptable for the display of most chromatograms in this study. When an attenuation other than 16 was used, that attenuation is reported relative to the attenuation of 16.

Several different mobile phase compositions were evaluated, however the most extensively used mobile phase involved a gradient elution based on 0.3% aqueous phosphoric acid (solvent A) and 0.3% phosphoric acid in methanol (solvent B). Flow was 1.0 ml/min. with solvent A for 5 min., linear gradient to 50:50 A:B over 30 min., linear gradient to 100% B over 20 min. and 100% B for up to 30 min. A 250 x 4.6 mm Altex Econocil C₁₈ column with C₁₈ guard cartridge was used for all chromatograms.

10

Ultraviolet absorption spectra were obtained using either a Perkin-Elmer 552 UV-Visible spectrophotometer with Hitachi Model 52 X-Y recorder for output or a Beckmann Model DU-64 microprocessor controlled UV-Visible spectrophotometer. UV spectra were obtained as aqueous methanol dilutions.

2.3 LEACHATES

Three different leachate samples were investigated. A soybean leachate sample dated 3/26/93 was obtained from Dr. Jay Garland. Most studies centered on wheat leachate. The biomass identified as BWT-931, level 3 was used to prepare fresh leachate as previously described (6). These samples were stored frozen until used. Several samples of aerobically treated wheat leachate were obtained from bioreactor ISAB-2. After daily removal of approximately 1 liter of bioreactor slurry, the addition of 40g dry biomass and sufficient water to maintain the bioreactor volume at 8 liters was performed. Bioreactor liquids were obtained as subsamples of the slurry that had been removed or were drawn directly from the bioreactor.

2.4 SOLID PHASE EXTRACTION (SPE)

Most leachate sample were partitioned into fractions prior to HPLC investigation. Partitioning was accomplished using Sep-Pak Bond-Elute-SAX cartridges (100 mg sorbent bed) from Analytichem International, 300 mg sorbent bed $-C_{18}$ PrepSep cartridges from Fisher Scientific ($-C_{18}$ PS), 1000 mg sorbent bed $-C_{18}$ and $-CH$ or 100 mg sorbent bed $-C_{18}$, $-C_8$, $-C_2$ from Altex. Cartridges were conditioned prior to use by washing with multiple aliquots of methanol and multiple aliquots of water. Flow was controlled at less than 1.0 ml/min by vacuum applied to the cartridge. Accurately measured aliquots of leachate sample and eluent were used to selectively partition components and elute leachate fractions from the Sep-Pak. Leachates were filtered through 0.45μ Nylon 66 filters before SPE. Treatment of samples by SPE resulted in several eluent fractions depending on the specific scheme employed. Table 1 summarizes the specific fraction labelling scheme that was used. The SPE procedures were accomplished by eluting the SPE cartridges with the following solvents: Water which resulted in fraction f2, 20:65:15 (v/v) methanol:water:acetonitrile for fraction fE, 20:65:15:0.3 methanol:water:acetonitrile:phosphoric acid for fraction fE₃, and 20:65:15:1.0 methanol:water:acetonitrile:phosphoric acid for fraction fE₁₀. In several situations multiple portions of a single eluent were collected. When this occurred the fraction designation was expanded, e.g. f2A, F2B, etc. to represent collection of multiple water fractions. All leachate samples and SPE fractions were stored frozen at $-4\text{ }^{\circ}\text{C}$ and were warmed to room temperature prior to use.

III. RESULTS AND DISCUSSION

3.1 REPRESENTATIVE COMPOUNDS

Preliminary experimentation focused on the HPLC and SPE behavior of several commercially available compounds that are commonly found in various plant species. Structures for some of these compounds are presented in Figure 1 and expands the selection from that previously studied (6) to included several flavonoid compounds. Chromatograms that demonstrate the separations possible for mixtures of related compounds are presented in Figure 2. HPLC retention times obtained for several known compounds are presented in Table 2. These observations suggest that the separation of fairly complex mixtures of these types of compounds is feasible and practical. However, some of these compounds elute with nearly identical retention time. Other chromatographic conditions that utilize different columns and/or mobile phases may improve the quality of separation.

The SPE behavior of twelve different compounds is summarized in Table 3. The SPE partitioning behavior of individual compounds is drastically different on different SPE stationary phases. For example, the chlorogenic acid partition efficiency when water is used as eluent varies from 95% on $-C_2$ to 0% on $-C_{18}PS$.

3.2 LEACHATE SOLID PHASE EXTRACTION AND CHROMATOGRAPHY

Leachate composition is very complex (6), therefore the use of SPE to obtain simpler fractions prior to HPLC is essential. The use of various SPE reverse phase cartridges allows for fractionation by removal of nonpolar or moderately polar components from the aqueous sample upon passage through the cartridge. Removal is based on preferential solubility of each considered component in water or the organic functional group associated with the sorbent bed of the SPE cartridge. The SPE $-SAX$ cartridge will interact to remove negatively charged ionic species from aqueous solution based upon ion exchange interaction.

Conditions for SPE extraction of various leachate samples and aerobic reactor liquid phase are presented in Table 1. These schemes typically generate four to twelve fractions. The number of fractions is determined by sorbent bed size, 100-1000 mg, and the amount of sample to be partitioned. The SPE sequence number refers to the actual number assigned to the experimental sequence and the Figure Reference Number refers to the figure in this report that presents the HPLC chromatogram for the designated fraction (Figures 4 - 12). In some cases the SPE fractions were subjected to another SPE using a second sorbent, typically $-SAX$ (Figures 6, 10 and 11). The chromatograms displayed in Figures 5 and 7 demonstrate the influence of SPE sorbent selection upon

sample fractionation. Considerable differences (see Figures 5 and 7) are observed when using 1000 mg cartridges of $-C_{18}$ and $-CN$ sorbents. The $-CN$ sorbent was used exclusively in the previous study (6). Elution volumes were dictated by sample and sorbent bed size and visible observation of colored band accumulation and movement through the sorbent bed as a particular eluent was applied. For example in SPE Sequence 9/Figure 10 the apparent irregular size of eluent aliquots used was dictated by observation of yellow-brown band movement through the sorbent bed. The success of this approach can be found when the chromatograms for the various fractions are evaluated (see Figure 10). Figures 5 - 7 represent fractions and subfractions from soybean leachate and demonstrate the complexity of the samples and the utility of using a combination of SPE/HPLC for characterization. Figures 8 - 11 provide parallel information regarding untreated wheat leachate. The comparison with chromatograms for aerobically treated wheat bioreactor liquid (Figure 12) becomes dramatic.

The ability of organic molecules to absorb UV radiation is a function of chemical structure and bonding considerations. Each compound will absorb over a somewhat characteristic range of wavelengths. Chromatogram appearance can therefore be influenced by selection of monitoring wavelength. Chromatograms displayed in Figure 4 demonstrate this effect. Wheat leachate fractions were chromatographed and monitored at four different wavelengths. The changing amplitude of the peaks appearing at approximately 15.3 minutes, at approximately 19.6 minutes and the band that appears over the 25 - 45 minute interval demonstrates how selection of wavelength can influence chromatogram appearance and provide basic information regarding composition. At the same time it can be possible to derive some information about compound identity from peak amplitude changes that are observed. For example, the peaks at 17.3 and 19.5 minutes which do not appear in the chromatogram obtained at 320 nm rules out flavonoid type compounds. The UV absorption spectra for flavonoids exhibit relatively strong absorption in the 300 to 400 nm range.

The UV spectra (Figure 3) compliment those presented previously and include compounds that were not available for earlier study (6). For example, spectrum CfA is from a cinnamic acid derivative, Spectra A, B and CRY are from polyhydroxy flavones, spectrum NGN is from a polyhydroxyflavanone and R is from a flavone-rutenocide. Other spectra displayed in Figure 3 are for compounds commonly found in various plant species.

The use of UV spectral properties for identification is limited in a positive sense because these spectra are devoid of spectral fine-structure that is present in other techniques such as infrared and nuclear magnetic resonance spectroscopy and mass spectrometry. It is an ideal choice as a detection technique for HPLC because instrumentation is relative simple and inexpensive

and it allows detection to take place at very low concentration levels.

3.3 AEROBICALLY TREATED LIQUID

Before leachate is used in a plant growth chamber it is treated aerobically to reduce the total carbon content. This treatment results in the conversion of carbohydrates and sugars to carbon dioxide which is removed from the system. The organic content of the bioreactor liquid is decreased considerably. A persistent brown color characterizes the untreated and the treated leachate. Various samples of bioreactor liquid, designated ISAB-2, SPE Sequence 15/Figure 12 were subjected to SPE and HPLC of the SPE fractions. The chromatograms that were obtained from bioreactor samples obtained 30 min. after start of feeding, at 20 min. after completion of feeding and at 21 hours after completion of feeding demonstrates the effectiveness of the treatment in modifying composition. Chromatograms show a total lack of the individual resolvable peaks that are present in untreated wheat leachate SPE fractions that are obtained by the same procedure, see Figure 9. The observed behavior was verified based on Sequence 17 where samples were taken 2 hours before feeding, at 30 min. into the feeding and 20 min. after completion of feeding. It appears that simple phenolic and polyphenolic compounds are consumed by the aerobic digestion, however compounds that are responsible for the brown solution color are not totally removed. The chromatograms from fractions f₂ and fE₃, (Figure 12) suggest the presence of two groups of compounds that provide the broad band peaks in the corresponding chromatograms. Both fractions contain brown colored material but because the retention times associated with the bands are different some chemical difference in structure is suggested. In addition the fraction f₂ chromatograms in Figure 12 have peaks that elute near the column void volume near 3 minutes. UV spectra of collected eluent from various HPLC runs yield spectra with absorption maxima near 280 nm for these rapidly eluting peaks. UV spectra for f_{1C} and fE_{3B} fractions are similar with absorption band shoulders from 260-300 nm and 270-320 nm, respectively. In addition absorption band tailing occurs beyond 400 nm and 450 nm, respectively. These spectra appear to be quite similar to flavonoids such as Biochanin A, Crysin, Naringenin and Rutin but they are not similar to a tannic acid sample which exhibits an absorption maximum below 280 nm and little absorption beyond 310 nm. Two commercially available tannic acid samples were used. Both exhibited non-specific partitioning on -C₁₈. Multiplex/bands were present in all fractions from f₁ to fE₁₀ and were not similar to the fraction chromatograms obtained from the bioreactor liquid samples. It is possible that a thorough spectrophotometric study of these fractions will provide additional insight regarding composition.

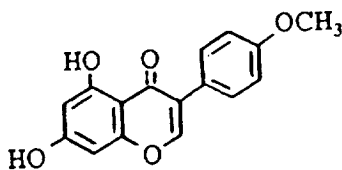
IV. SUMMARY COMMENTS

The use of solid phase extraction and high performance liquid chromatography for preliminary characterization of general similarities and differences in the chemical composition of aqueous leachates from crop residues and bioreactor liquids in the Controlled Ecological Life Support System program has been accomplished. The composition is complex. Wheat leachate has been studied both before and after aerobic treatment to reduce to total carbon load. The presence of highly colored anionic materials, as well as other colored and colorless compounds that can be fractionated from the samples by solid phase extraction provides the basis for an involved study to determine composition. High performance liquid chromatography using ultraviolet absorption as the detection tool was used to further characterize composition. It is apparent that a modest number of individual compounds of unique structure and other groups of compounds with similar structure constitute the organic composition of the untreated leachates. No specific compounds were identified, however, it is possible to determine the presence of flavonoid-like and phenolic-like classes of compounds in untreated leachate and in aerobically treated leachate. It has been demonstrated that several specific phenolic-like compounds that are produced by many crop plants are not present in measurable quantities in the leachate samples that were available. However, it is possible that related compounds are responsible for the chromatographic behavior observed. Aerobic treatment produces a liquid that is quite different from untreated leachate. An extension of this work should include the identification of some of these compounds. Once identified, controlled experiments that relate to persistence and accumulation of these compounds in the treated leachate enriched hydroponic solutions used in the growth chambers will be possible. An exhaustive characterization is not possible because of the complexity associated with the leachate composition. In fact it is probably not necessary to fully characterize the composition as long as a general understanding of the type of compounds that are present is documented and studies undertaken to understand the role played by these compounds in the hydroponic solutions.

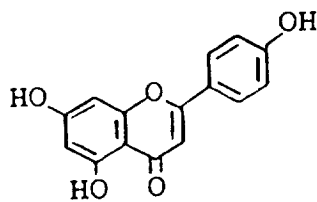
V. REFERENCES

1. Knott, W. M. The Breadboard Project: A Functioning CELSS Plant Growth System. *Adv. Space Res.* Vol. 12. No. 5. 1992. pp. 45-52.
2. Koller, A. M. CELSS Breadboard Facility Project Plan. Biomedical Operations and Research Office. National Aeronautics and Space Administration, Kennedy Space Center. March 1986.
3. Mackowiak, C. L., L. P. Owens-Hall, C. R. Hinkle, and R. P. Prince. Continuous Hydroponic Wheat Production Using a Recirculating System. NASA Tech. Memorandum TM 102784 September 1990.
4. Prince, R. P., and W. M. Knott. CELSS Breadboard Project at the Kennedy Space Center. In: D. W. Ming and D. L. Henninger (eds.), *Lunar Base Agriculture: Soils for Plant Growth.* Amer. Soc. Agron., Madison, WI 1989.
5. Garland, J. L., and C. L. Mackowiak. Utilization of the Water Soluble Fraction of Wheat Straw as a Plant Nutrient Source. NASA Tech. Memorandum TM 103497 June 1990.
6. Madsen, B. C., "Chemical Characterization of some Aqueous Leachates from Crop Residues in CELSS." pp.229-259/260. NASA/ASEE Summer Faculty Fellowship Program. KSC/UCF. Contract Report NASA CR-191004. 1992.
7. Lee, H. S., R. D. Carter, S. M. Barros, D. J. Dezman, and W. S. Castle. Chemical Characterization by Liquid Chromatography of Moro Blood Orange Juice. *J. Food Composition and Analysis.* Vol. 3. 1990. pp. 9-19.8.
8. Geibel, M., D. Treutter and N Meier. Characterization of sour cherries by HPLC analysis of the Bark-Flavonoids Combined with Multivariate Statistics. *Euphytica.* Vol. 45. 1990. pp. 229-235.
9. Koster, J., A. Zuzok and W. Barz. High-Performance Liquid Chromatography of Isoflavones and Phytoalexins from *Cicer Arietinum*. *J. Chromatog.* Vol 270. 1983. pp. 392-395.
10. Onyeneho, S. N. and N. S. Hettiarachy. Antioxidant Activity of Duram Wheat Bran. *J. Agric. Food Chem.* Vol. 40. 1992. pp. 1496-1500.

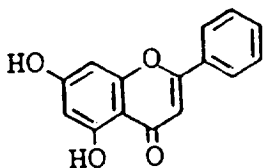
11. Graham, T. L. Flavonoid and Isoflavonoid Distribution in Developing Soybean Seedling Tissues and in Seed and Root Exudates. *Plant Physiol.* Vol. 95. 1991. pp. 594-603.
12. Banwart, W. L., P. M. Porter, T. C. Granato and J. J. Hassett. HPLC Separation and Wavelength Area Ratios of more than 50 Phenolic Acids and Flavonoids. *J. Chem. Ecology.* Vol. 11(3). pp. 383-395.
13. Eldridge, A. C. High-Performance Liquid Chromatography Separation of Soybean Isoflavones and their glucosides. *J. Chromatog.* Vol. 234. 1982. pp. 494-496.
14. Pietta, P. G., P. L. Mauri, LC. Gardana and A. Bruno. High-Performance Liquid Chromatography with Diode-Array Ultraviolet Detection of Methoxylated Flavones in Orthosiphon Leaves. *J. Chromatog.* Vol. 547. 1991. pp. 439-442.
15. Koupai-Abyazani, M. R., J. McCallum and B. A. Bohm. Identification of the Constituent Flavanoid Units in Sainfoin Proanthocyanidins by Reverse Phase High-Performance Liquid Chromatography. *J. Chromatog.* Vol. 594. 1992. pp. 117-123.
16. Okuda, Takuo, T. Yoshida, and T. Hatano. New Methods of Analyzing Tannins. *J. Natural Products.* Vol. 52(1). 1989. pp. 1-31.
17. Zhu, J., J. Ng and L. J. Filippich. Determination of Tannic Acid and its Phenolic Metabolites in Biological Fluids by High-Performance Liquid Chromatography. *J. Chromatog.* Vol. 577. 1992. pp. 77-85.
18. Turley, D. J., M. T. Kelly and M. R. Smyth. High-Performance Liquid Chromatographic Method for the Comparison of Tanning Capacity of Tannic Acid Batches used in the Manufacture of Pregnancy Testing Kits. *J. Chromatog.* Vol. 513. 1990. pp. 263-269.
19. Hatano, T., K. Reiko, M. Yoshizake and T. Okuda. Seasonal Changes in the Tannins of Liquidambar Formosana Reflecting their Biogenesis. *Phytochem.* Vol. 25(12). 1986. pp. 2787-2789.



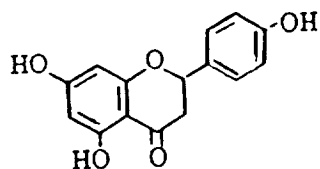
Biochanin A (F.W. 284.27)



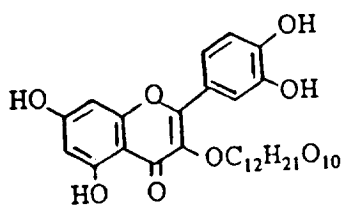
Apigenin (F.W. 270.24)



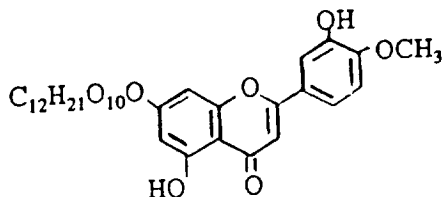
Chrysin (F.W. 254.24)



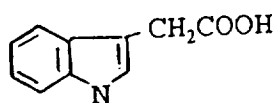
Naringenin (F.W. 272.26)



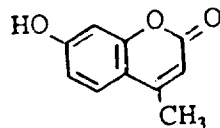
Rutin (F.W. 664.57)



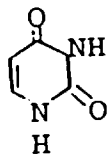
3',5,7-trihydroxy-4'-methoxy flavone-7-rutinoside (F.W. 608.56)



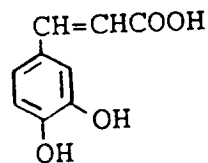
Indole-3-Acetic Acid (F.W. 175.19)



7-Hydroxy-4-Methyl Coumarin (F.W. 176.17)



Uracil (F.W. 112.09)



Caffeic Acid (F.W. 180.16)

Figure 1. Structures of Some Representative Compounds.

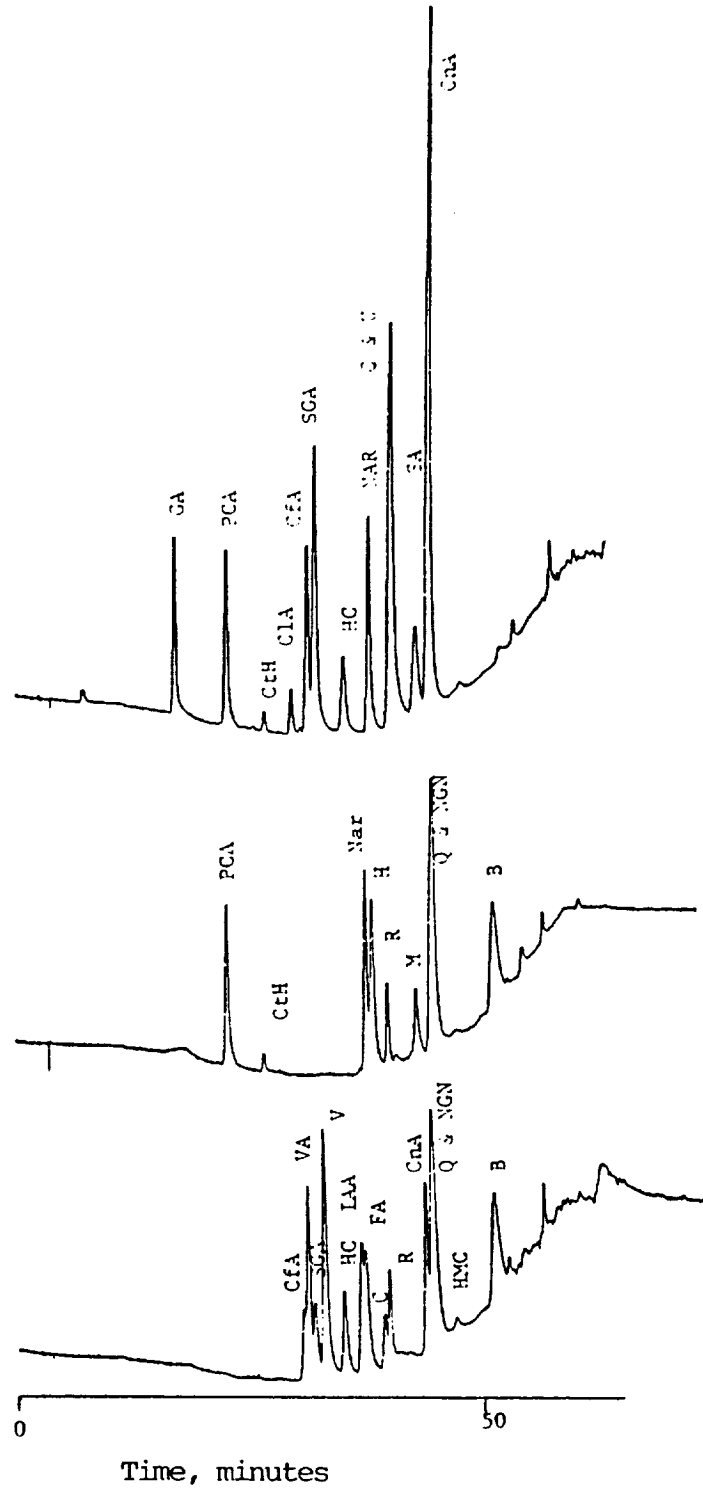
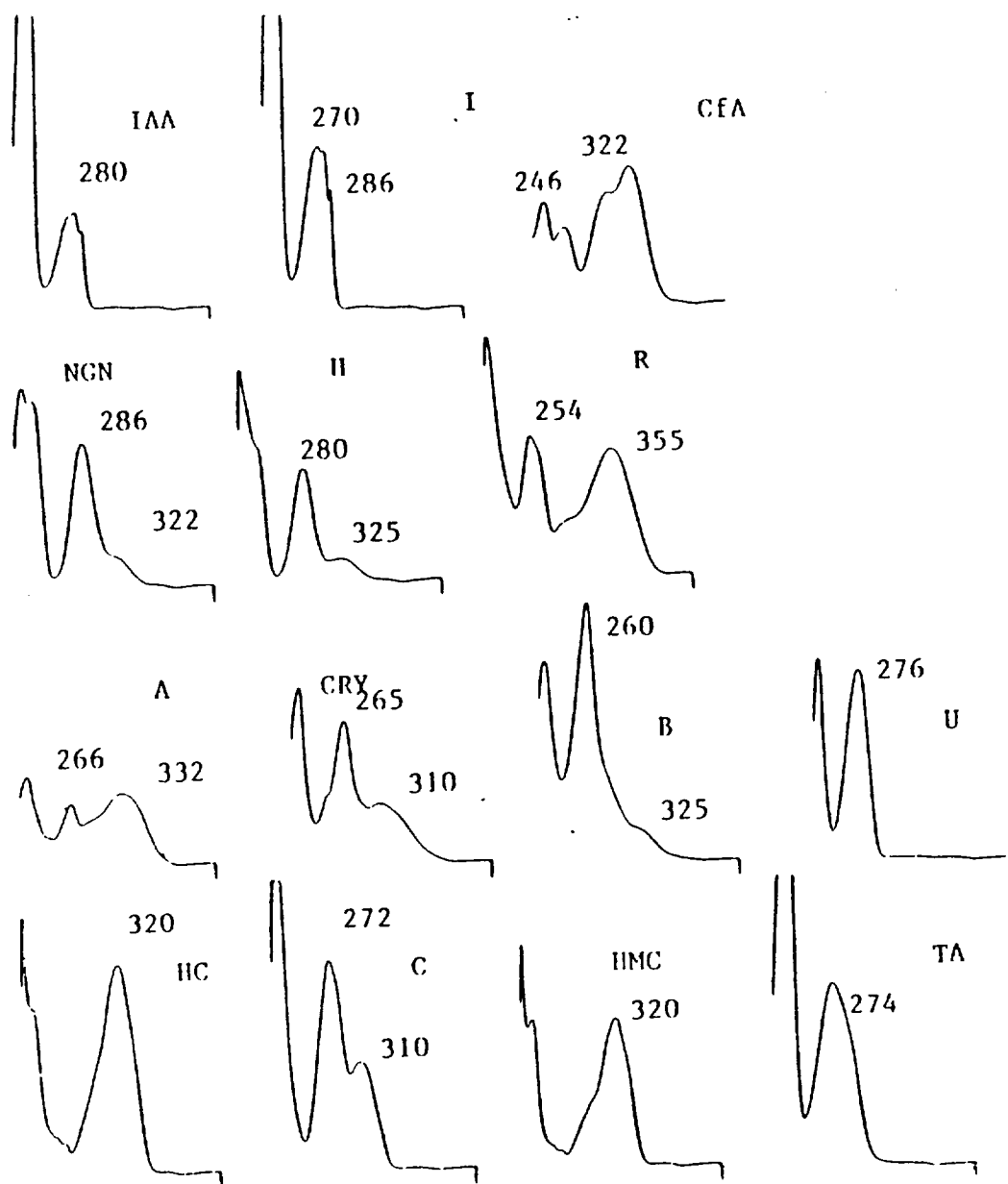


Figure 2. HPLC Chromatograms of some Representative Compound Mixtures Measured at 280 nm. See Table 2 for identification of peaks. See text for chromatographic conditions.



Compound	$\mu\text{g/ml}$	Compound	$\mu\text{g/ml}$
IAA	7.9	B	2.8
I	1.9	U	8.2
CfA	3.6	HC	7.1
NGN	7.5	C	8.0
H	12.0	HMC	4.8
R	13.9	TA	8.4
A	2.7		
CRY	3.6		

x-axis represents 200nm-450nm.

y-axis 1.0cm represents 0.25 Absorbance.

Figure 3. Ultraviolet Absorption Spectra for Selected Representative Compounds.

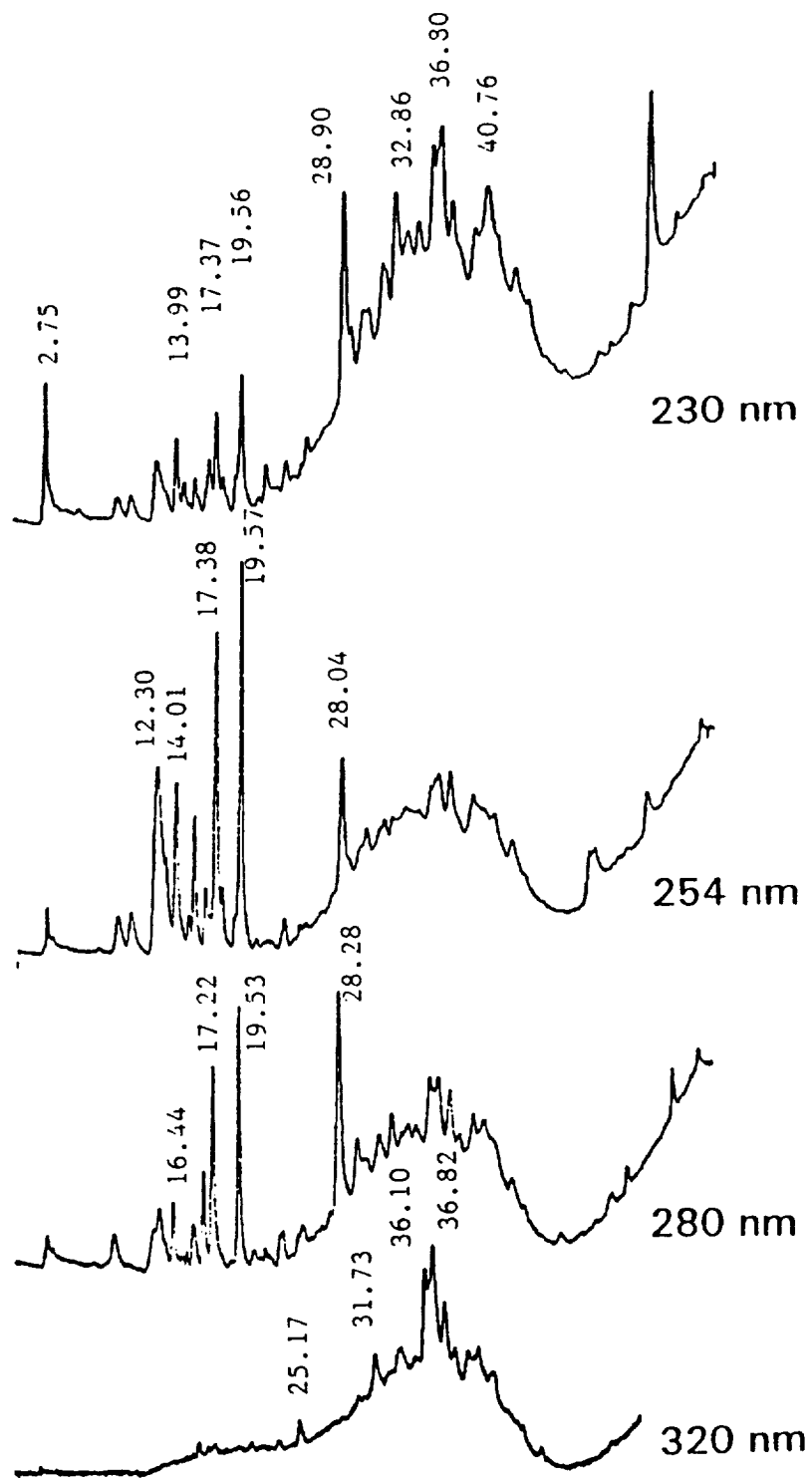


Figure 4. Chromatograms of a Wheat Leachate SPE Fraction Obtained at Several Wavelengths. See Figure 9 for specific identification of sample, SEQ. 9, fE_{3A}.

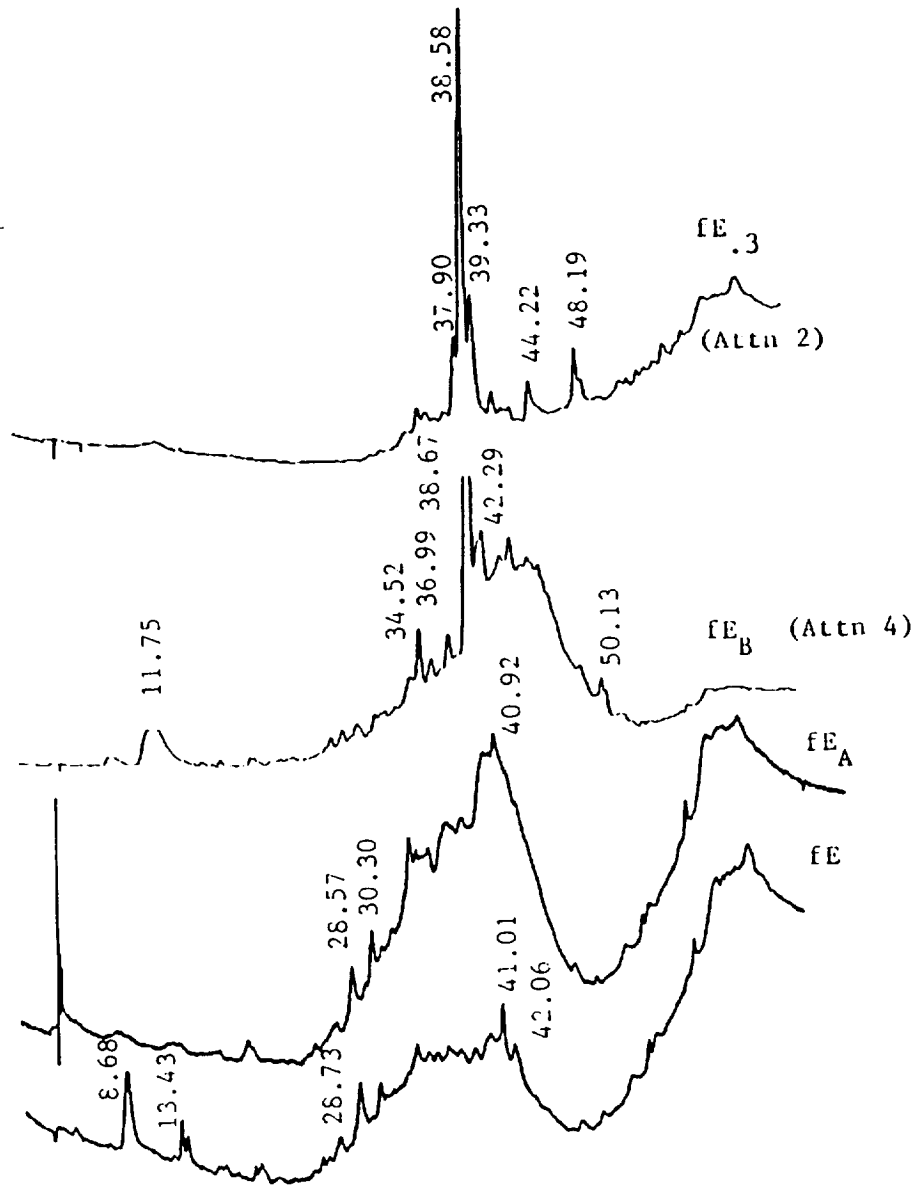


Figure 5. Chromatograms for Comparison of Fractions from SPE -C₁₈ Treatment of a Soybean Leachate Sample Measured at 280 nm.

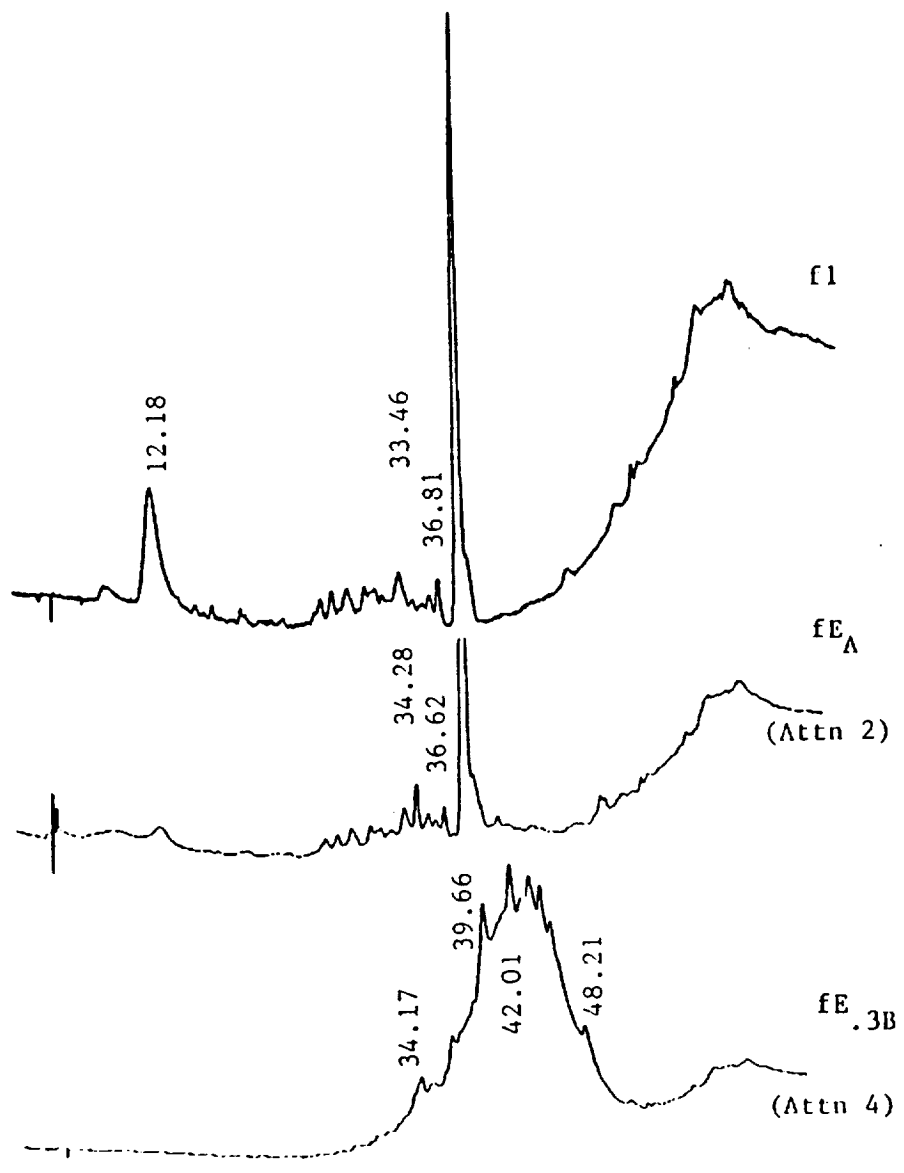


Figure 6. Chromatograms for Comparison of a Subfraction from a Soybean Leachate Fraction (see Figure 5, fE_B) Partitioned on SPE -SAX and Measured at 280 nm.

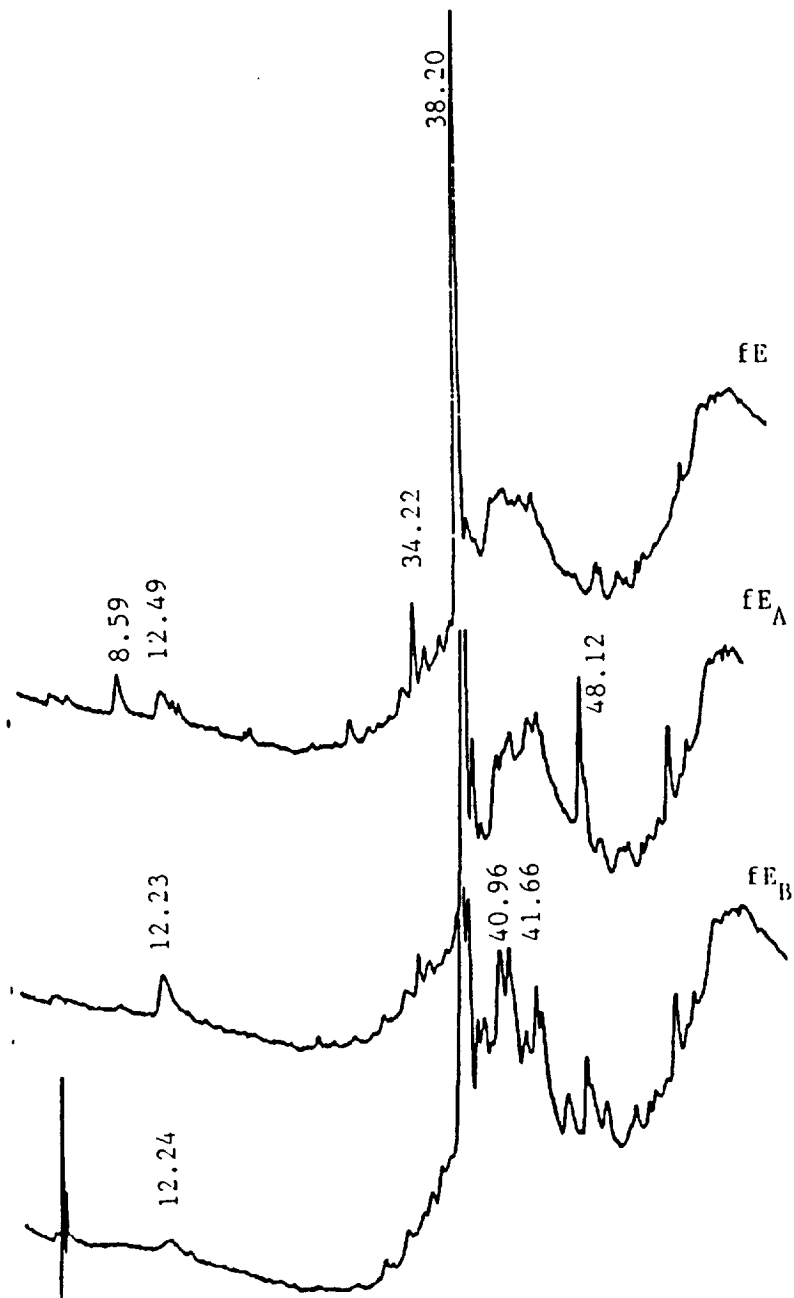


Figure 7. Chromatograms for Comparison of SPE -CH Fractions of Soybean Leachate measured at 280 nm.

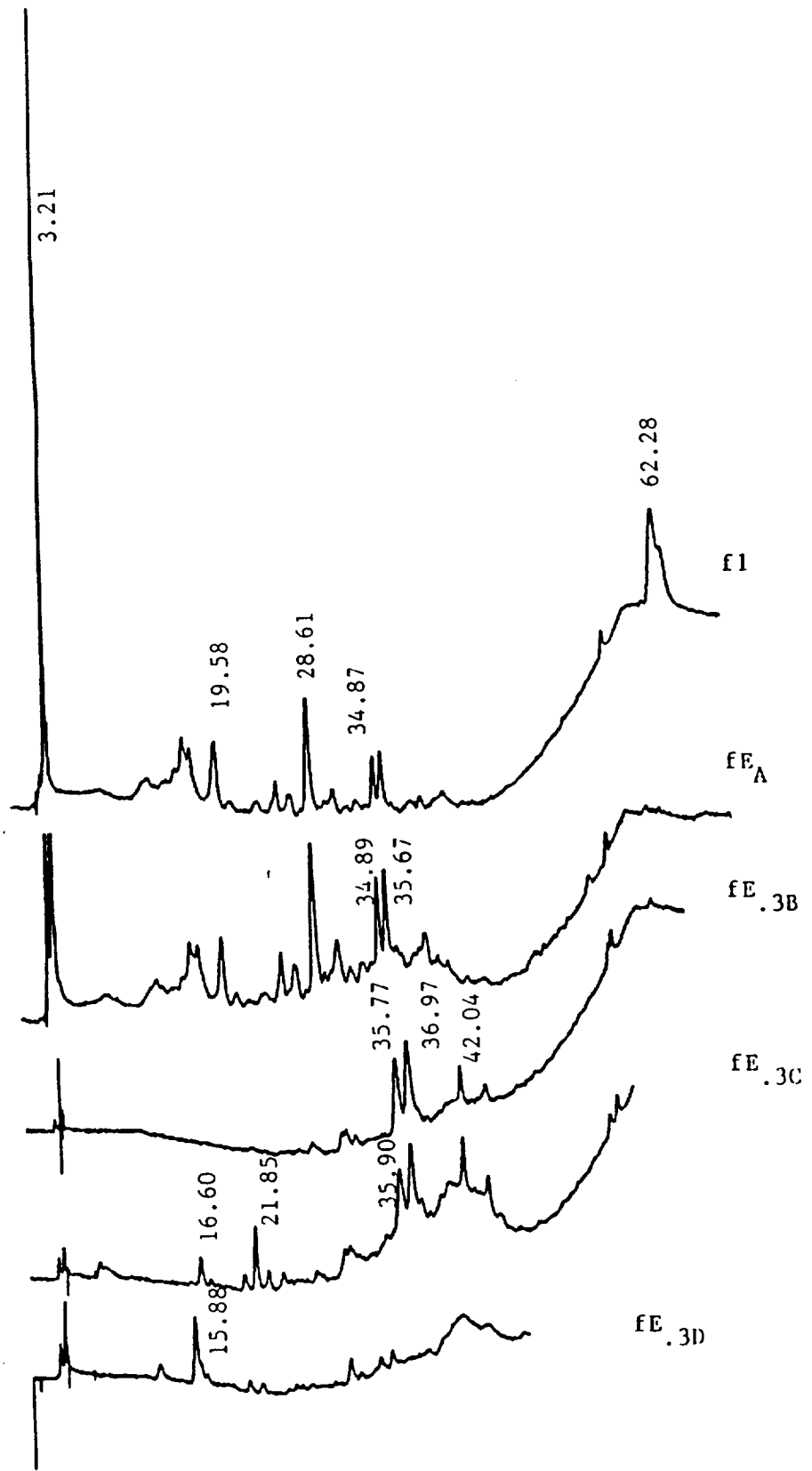


Figure 8. Chromatograms for Comparison of SPE -SAX Fractions of Wheat Leachate Measured at 280 nm.

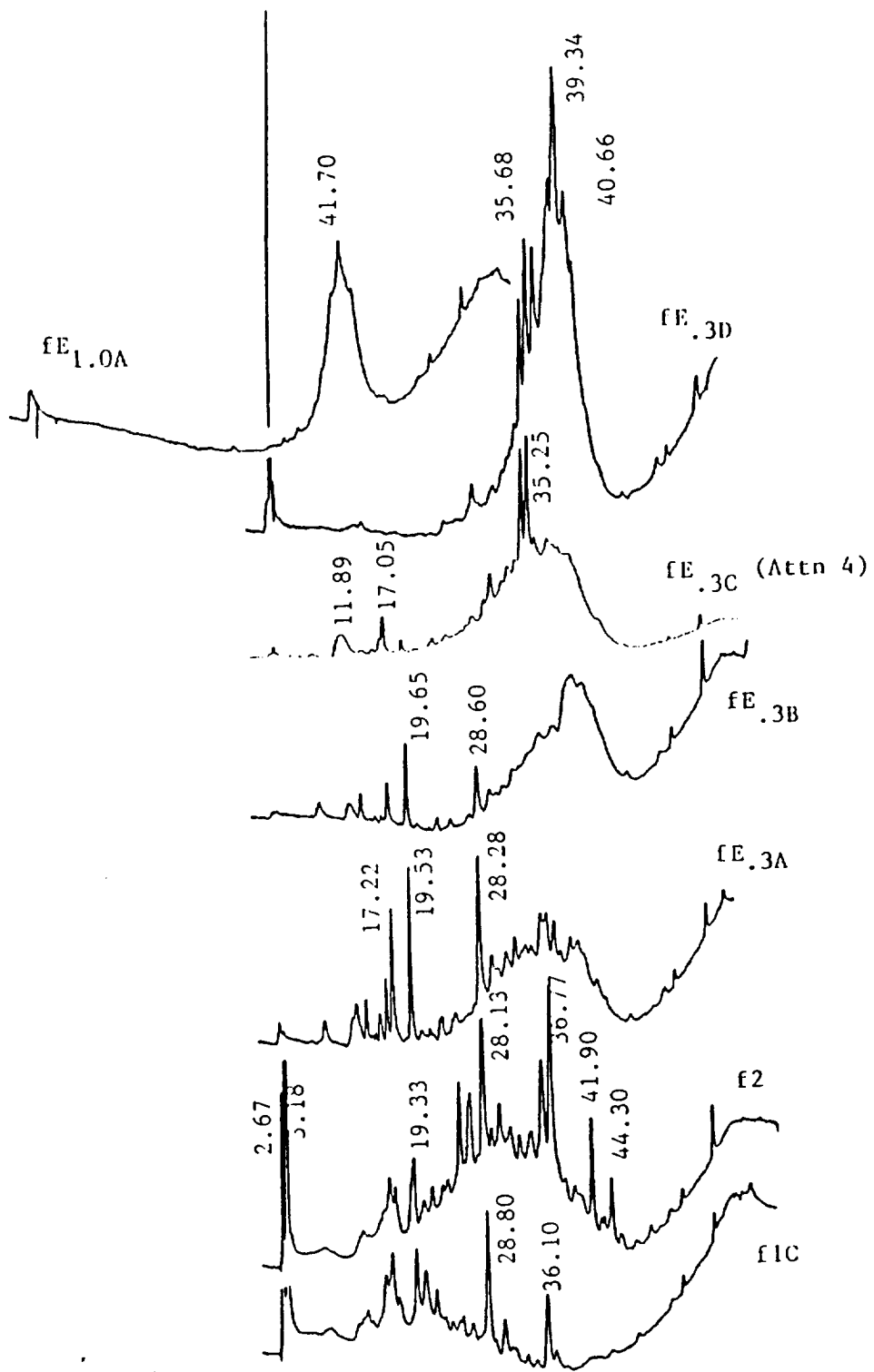


Figure 9. Chromatograms of SPE Fractions of 9.6 ml Wheat Leachate Obtained from 1000 mg Altex -C₁₈ Measured at 280 nm.

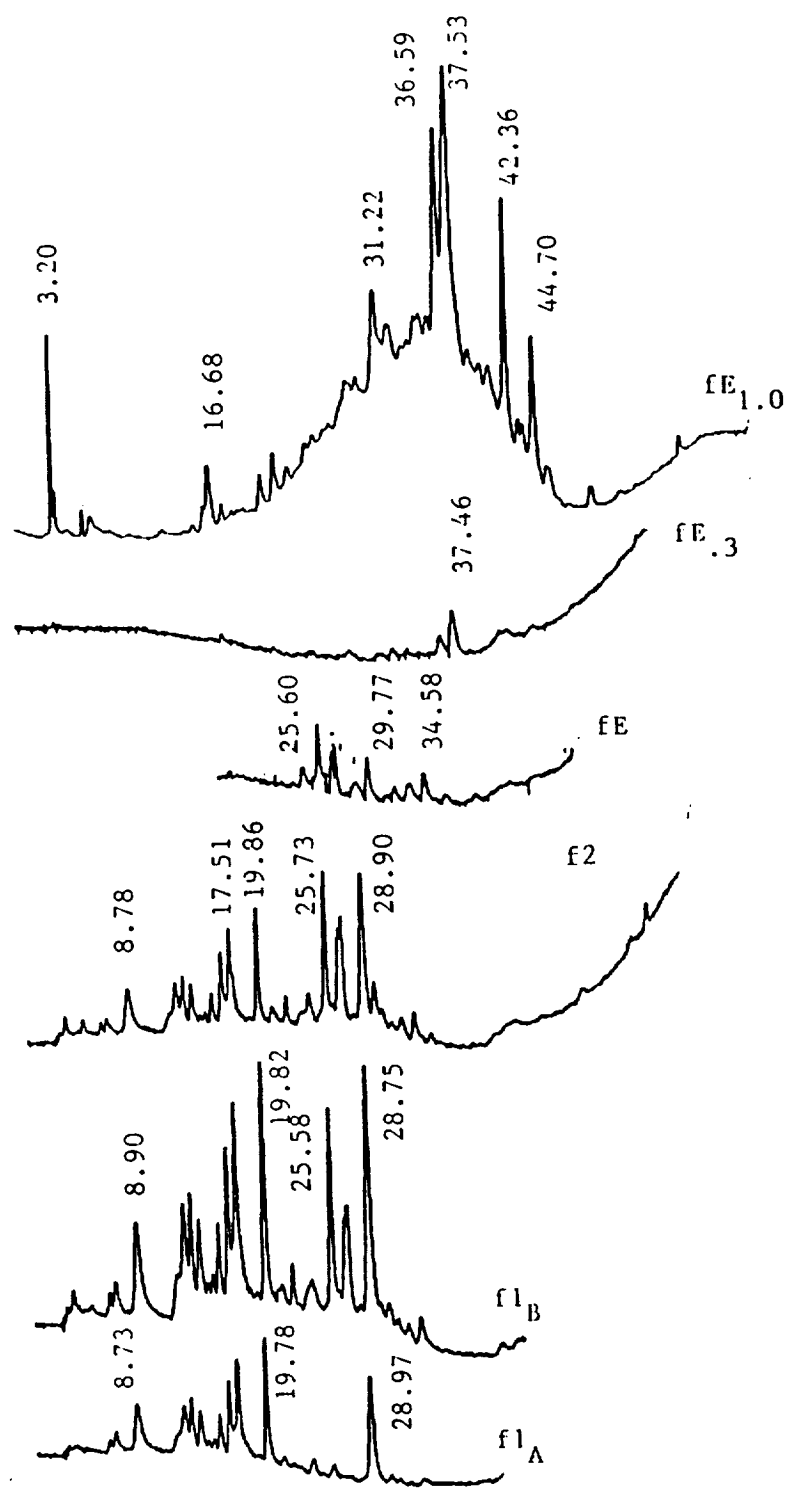


Figure 10. Chromatograms for -SAX subfractions of Fraction f2 Wheat Leachate (see Figure 9).

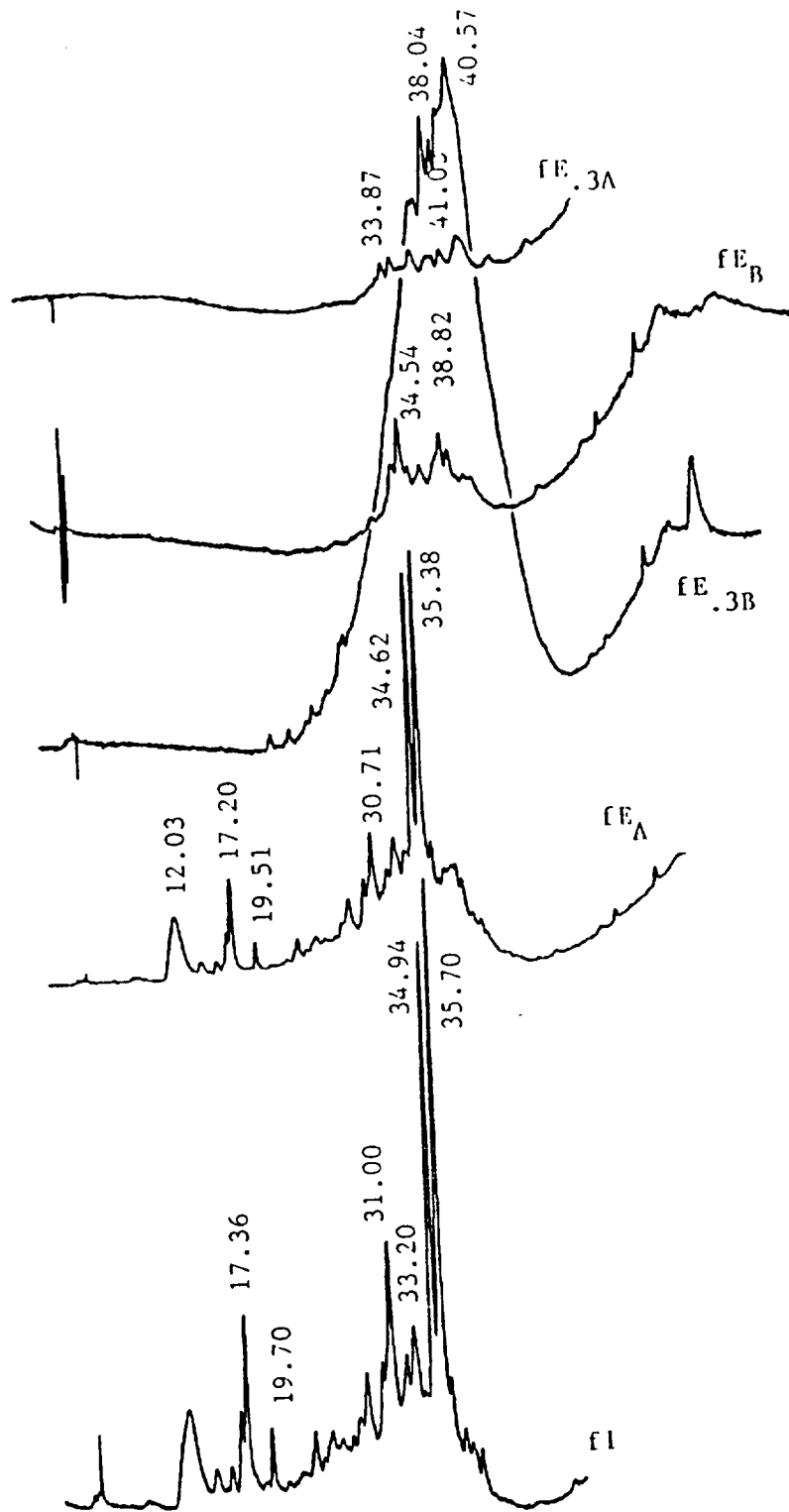


Figure 11. Chromatograms for -SAX subfractions of Fraction $fE_{0.3C}$ Wheat Leachate (see Figure 9).

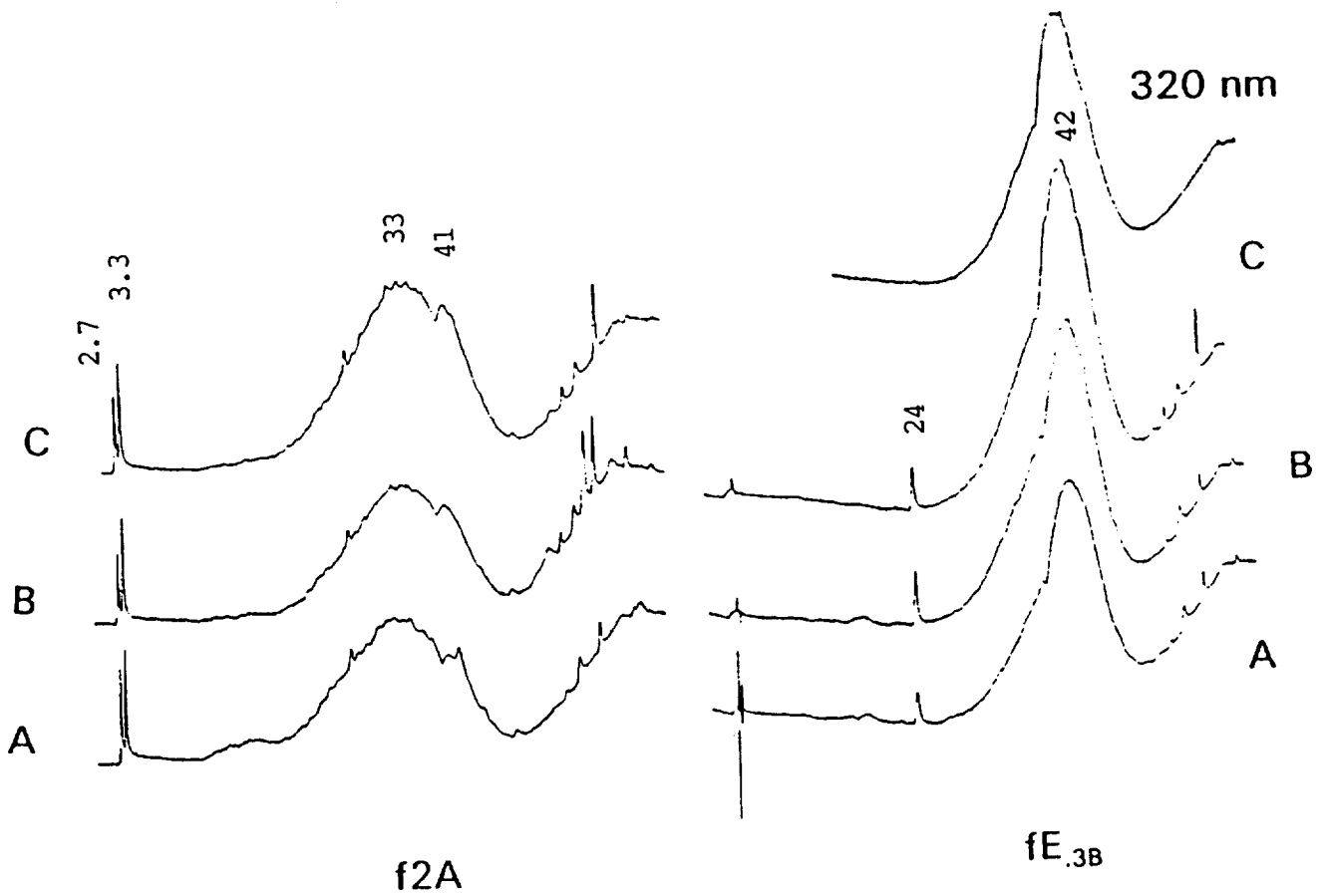


Figure 12. Comparison Chromatograms of SPE -C₁₈ Fractions of Wheat Leachate Obtained During and After Bioreactor Feeding. Chromatograms are obtained at 280 nm unless noted.

- A. 30 Minutes after start of feeding.
- B. 20 Minutes after completion of feeding.
- C. 21 Hrs. after completion of feeding, feed duration 2 hrs.

Table 1. Conditions Used for Solid Phase Extraction of Soybean and Wheat Leachates.

Sample, size	SPE Cartridge	Fraction Scheme and Designation	SPE Seq. / Fig. Reference
Soybean, 4.8ml	1000mg -C18	1.6ml f2 1.0, 2x0.5ml fE 1.0ml fE ₃	7 / 5
fE _B (seq. 7) 0.28ml	100mg -SAX	2x0.4ml fE 2x0.3ml fE ₃	7A/ 6
fE _A (seq. 7) 0.28ml	100mg -SAX	see Seq. 7A	7B/
Soybean, 4.8ml	1000mg -CH	see Seq. 7	7 / 7
Wheat, 0.8ml	100mg -SAX	2x0.4ml fE 4x0.4ml fE ₃	8 / 8
Wheat, 9.6ml	1000mg -C18	1.6ml f2; 0.8, 0.6, 1.5, 1.1 ml fE ₃ ; 2x1.6ml fE _{1.0}	9 / 9
fE _{3B} (Seq. 9) 0.28ml	100mg -SAX	2x0.4ml fE ₃ 2x0.4ml fE _{1.0}	9B /
fE _{3C} (Seq. 9) 0.35ml	100mg -SAX	see Seq. 9B	9C / 11
f2(Seq. 9), 0.8ml	100mg -SAX	see Seq. 9B and 2x1.6ml fE _{1.0}	9W / 10
Wheat, 0.5ml pH 2	100mg -C18	0.5ml f2, fE ₃ , fE _{1.0}	10 /
Wheat, 0.5ml pH 8.5	100mg -C18	0.5ml f1/2; see Seq. 10	11 /
ISAB-2 1 ml, day 14	-C18(1)	0.5ml f2, fE, fE ₃ and fE _{1.0}	12 /
ISAB-2 1 ml, day 14	-SAX(2)	see Seq. 12 -C18	
ISAB-2 1 ml, days 25 and 28	-C18(1,3) -SAX(2,4)	1.0ml f2, 0.5ml 0.5ml fE, fE ₃ , fE _{1.0}	14 /
ISAB-2 day 52	-C18	2x1.2ml f2; 0.8, 1.2, 2x0.8ml fE ₃ ; 3x0.8ml fE _{1.0}	15 / 12
Tannic Acid 1 ml, 5000 ppm	-SAX	0.4ml f2; 0.8ml fE; 0.4ml fE ₃ ; 0.6ml fE _{1.0}	16 / 13
ISAB-2 1 ml, day 59	100mg -C18	0.5ml f1/2; 1.0ml fE; 0.5 ml fE ₃	17 /

50

Table 2. Representative Compounds Used
in Chromatographic Studies

Compound	Abbrev.	HPLC Retention Time, min.
Biochanin A	B	51.7 - 51.8
Caffeic Acid	CfA	31.9 - 32.5
Catechin Hydrate	CtH	26.8 - 27.0
Chlorogenic Acid	ClA	29.6 - 29.9
Cinnamic Acid	CnA	44.1 - 44.4
Coumaric Acid	CA	36.3 - 36.4
Coumarin	C	39.7 - 40.7
Ethyl Vanillin	EV	37.7 - 37.9
Ferulic Acid	FA	37.1 - 37.4
Gallic Acid	GA	16.7 - 17.2
Hesperidin	H	38.7 - 39.0
7-Hydroxycoumarin	HC	35.2 - 35.6
7-Hydroxy-4-methylcoumarin	HMC	47.3 - 47.6
p-HydroxyBenzoic Acid	pHBA	28.1 - 28.7
Indole	I	42.0 - 42.6
Indole-3-Acetic Acid	IAA	37.1
Morin	M	42.9 - 43.6
Naringin	Nar	37.9 - 38.3
Naringenin	NGN	44.9 - 45.2
Protocatechuic Acid	PCA	22.7 - 23.2
Quercetin	Q	45.0 - 45.3
Rutin	R	40.2
Salicylic Acid	SA	41.2 - 41.9
Syringic Acid	SGA	32.1 - 32.9
Uracil	U	39.4
Vanillic Acid	VA	30.8 - 31.3
Vanillin	V	32.9 - 33.4

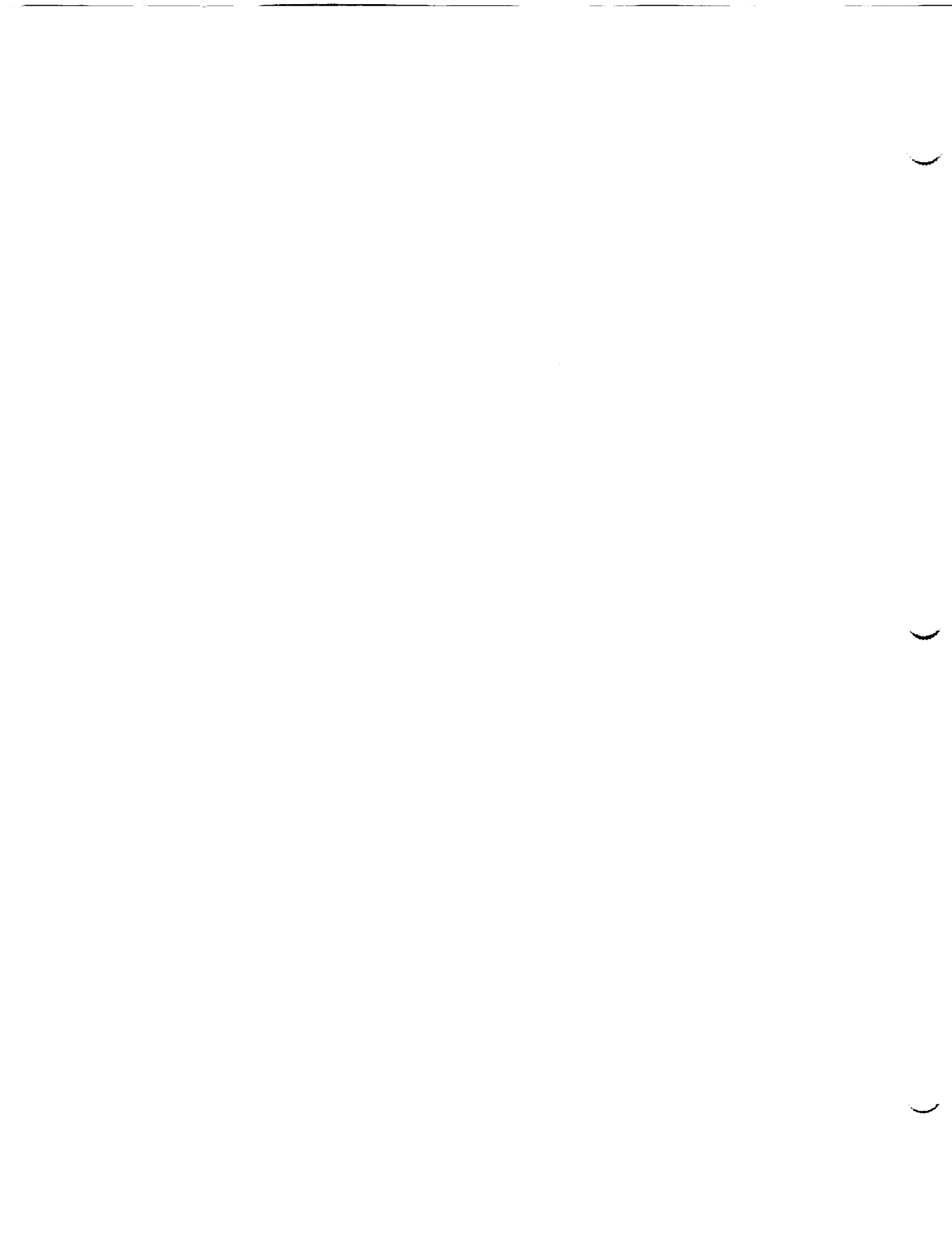
Table 3. Solid Phase Extraction Partitioning Efficiency of Selected Representative Compounds on Four Different Extraction Cartridges

Compound	μg^*	100mg -C ₁₈			300mg -C ₁₈ PS		
		f1/2	fE	fE ₃	f1/2	fE	fE ₃
Gallic Acid	5.2	100			100		
Protocatechuic Acid	10.4	100			85	15	
Catechin Hydrate	2.1	67	33			100	
Chlorogenic Acid	2.2	50	50			100	
Caffeic Acid	5.2	62	31	7		95	5
Syringic Acid	10.0	54	42	4		97	3
7-Hydroxycoumarin	5.3	22	78			40	60
Naringin	11.7	2	91	7		33	67
Coumarin**	12.0	5	95			9	91
Uracil**	2.5						
Indole	5.6	81	13	6			
Cinnamic Acid	5.2	68	23	9		20	80

Compound	μg^*	100mg -C ₂			100mg -C ₃		
		f1/2	fE	fE ₃	f1/2	fE	fE ₃
Gallic Acid	5.2	100			100		
Protocatechuic Acid	10.4	100			100		
Catechin Hydrate	2.1		100		86	14	
Chlorogenic Acid	2.2	95	5		87	13	
Caffeic Acid	5.2	72	28		92	8	
Syringic Acid	10.0	64	33	3	93	7	
7-Hydroxycoumarin	5.3		100		56	44	
Naringin	11.7		93	7	4	88	8
Coumarin**	12.0		98	2	15	84	1
Uracil**	2.5			2.5			
Indole	5.6		100		92	8	
Cinnamic Acid	5.2	41	58	1	79	20	1

See text for SPE cartridge descriptions.
 1.0 ml of 12 component standard added.
 f1/2 1.0 ml water except 2.0 ml water for C₁₈PS.
 fE and fE₃ 1.0 ml each.

** Compounds not resolved in chromatogram. Partitioning efficiency reported represents a combination of recoveries for the two compounds.



1994018769
N94-22242

1993 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA

92-37
197192
p. 26

RACK INSERTION END EFFECTOR (RIEE) AUTOMATION

PREPARED BY:	Dr. Narasimha S. Malladi
ACADEMIC RANK:	Associate Professor
UNIVERSITY AND DEPARTMENT:	Tuskegee University Department of Mechanical Engineering
NASA/KSC	
DIVISION:	Mechanical Engineering
BRANCH:	Special Projects
NASA COLLEAGUE:	Bill Jones Todd Graham Allan Littlefield Tom Woods
DATE:	August 11, 1993
CONTRACT NUMBER:	University of Central Florida NASA-NGT-60002 Supplement: 11

ACKNOWLEDGMENTS

The author gratefully acknowledges the support of Dr. Ramon Hosler and Kari Stiles of the University of Central Florida and the staffs of NASA and I-NET whose diligence and friendliness, made the NASA-ASEE Faculty Fellowship program at the Kennedy Space Center productive and enjoyable. Particular thanks to Willis Crumpler, Albert J. Diggs, Bill Jones, Eduardo Lopez, Todd Graham and Allan Littlefield of NASA and Dr. Jeffery Hooker, James E. Spencer and the enthusiastic Band of Engineers of the Robotics Group of I-NET. I enjoyed the help of the co-op student Tom Woods and the fellowship of all the other NASA fellows and in particular Dr. Robert Byers and Dr. Keith Doty. Dr. Doty has provided me with valuable academic and spiritual fellowship throughout the stay. For all this, may God bless all.

ABSTRACT

NASA is developing a mechanism to manipulate and insert Racks into the Space Station Logistic modules. The mechanism consists of a base with three motorized degrees of freedom, a 3 section motorized boom that goes from 15 to 44 feet in length, and a Rack Insertion End Effector (RIEE) with 5 hand wheels for precise alignment. The robotics section has been tasked with the automation of the RIEE unit.

In this report, for the automation of the RIEE unit, application of the Perceptics Vision System has been conceptually developed to determine the position and orientation of the RIEE relative to the logistic module and a MathCad program is written to display the needed displacements for precise alignment and final insertion of the Rack.

The uniqueness of this report is that the whole report is in fact a MathCad program including text, derivations and executable equations with example inputs and outputs.

TABLE OF CONTENTS

1. INTRODUCTION
2. THE RIEE MECHANISM
3. DETERMINATION OF THE COORDINATES OF THE TARGETS IN THE CAMERA FRAME
4. DETERMINATION OF COORDINATE SYSTEM ATTACHED TO THE RACK INTERFACE PLATE ROTATION AXIS
5. SUPPORT COORDINATES IN BOOM ROBOT COORDINATE SYSTEM
6. DISPLACEMENTS IN THE BOOM ROBOT AND THE RIEE MECHANISM
7. CONCLUSIONS AND RECOMMENDATIONS
8. REFERENCES

RACK INSERTION END-EFFECTOR (RIEE) AUTOMATION

Narasimha S. Malladi, 1993 Summer Faculty Fellow, Tuskegee University

1. INTRODUCTION

Towards the automation of the RACK INSERTION END-EFFECTOR (RIEE) Mechanism, Application of the Perceptics Vision System has been conceptually developed to determine the position and orientation of the RIEE relative to the logistic module and a MathCad program is written to compute and display the displacements required to

- 1) Align the RIEE relative to the logistic module to prepare for final insertion of the Rack, and
- 2) Translate the Rack Rotation axis and rotate the Rack about it for the Final Insertion.

The above displacements are used for the actuation of motors to be incorporated in the Improved RIEE Mechanism.

The following sections describe concepts, derivations of formulae and the MathCad program with sample input and output. With the help of this report a scaled prototype of the Improved RIEE Mechanism with an integrated Vision System can be quickly developed and tested.

2. THE RIEE MECHANISM

The RIEE is a space mechanism with 5 degrees of freedom. It has 12 links and 15 joints. It is used to align the axis of rotation of the Rack and perform the final insertion of the Rack into the logistic module. The boom, carrying the RIEE, has 4 degrees of freedom. Fig.1 and Fig.2 show the RIEE in retracted and installed positions as per drawing set # 82K03931.

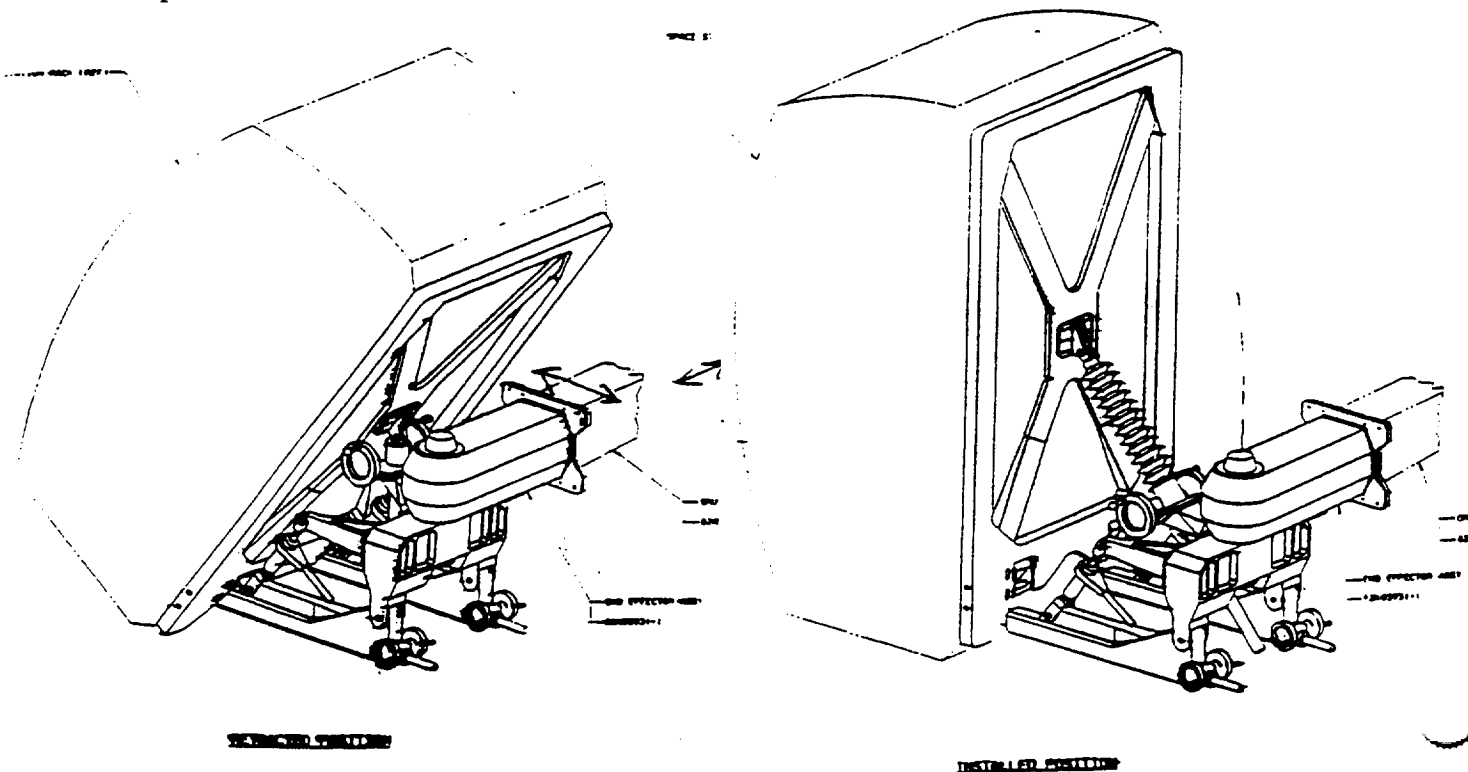


Fig. 1

Fig. 2

3. DETERMINATION OF THE COORDINATES OF THE TARGETS IN THE CAMERA FRAME

To insert the rack, we have to know the positions of the rack attach points on the module, shown as lower support points (S) in the Fig. 3 below. In the following analysis, we derive the expressions for the coordinates of the left and right lower support points by getting the images of the spherical targets suspended vertically from the supports employing a set of two cameras for each target. The figure shows the relationship of the cameras to the targets in the y-z plane of the boom coordinate system carrying the RIEE mechanism. The RIEE mechanism has 2 end effector points on the left and right sides shown by letter E, and a top end effector point P. Different views of this mechanism can be studied in the set of drawings # 89K03931. The points shown in the figure carry prefixes "R" for "Right" and "L" for "Left" in the analysis that follows.

RIEE MECHANISM SCHEMATIC

(Not to scale)

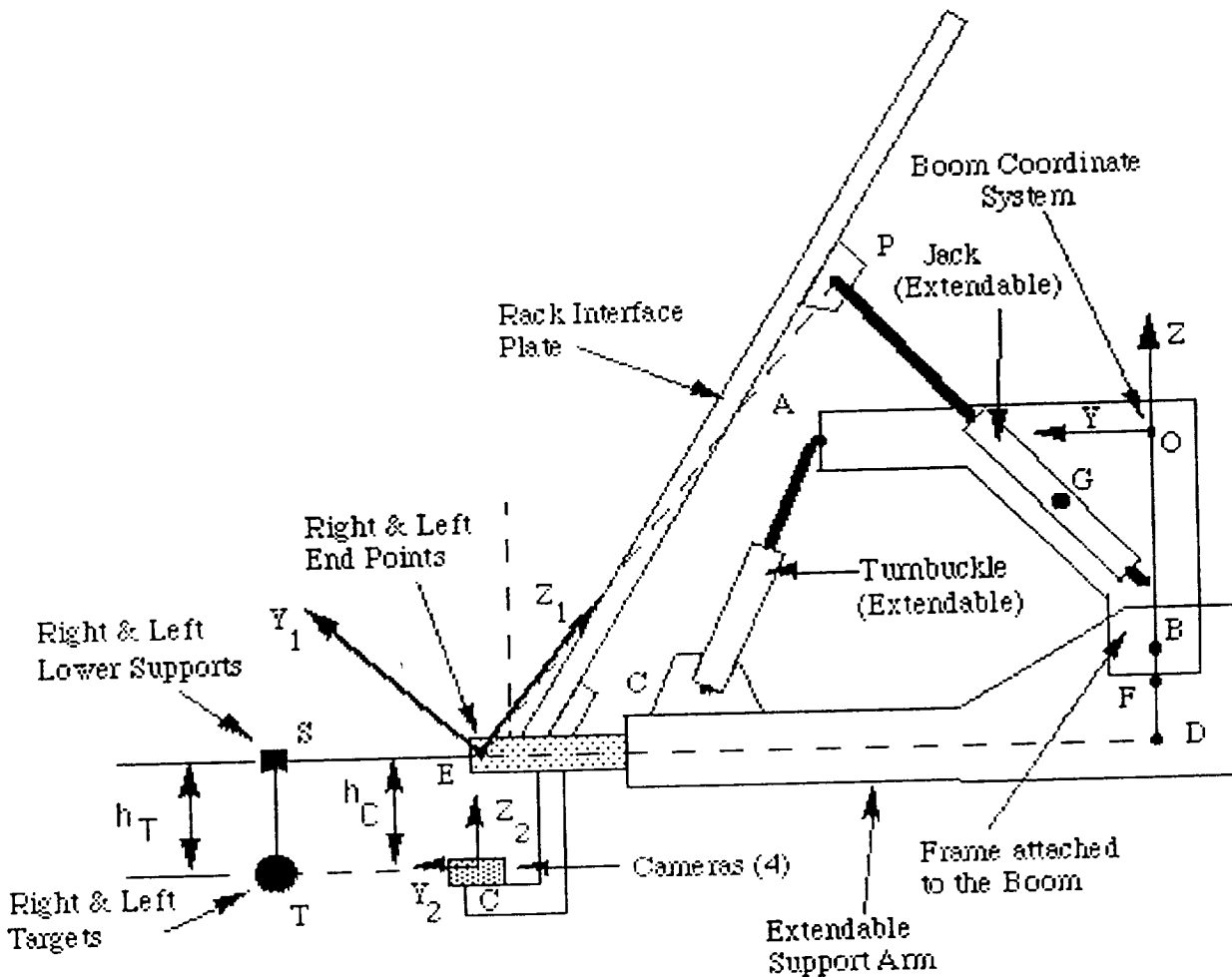


Figure 3

TARGET AND CAMERA ARRANGEMENT SCHEMATIC

(Not to scale)

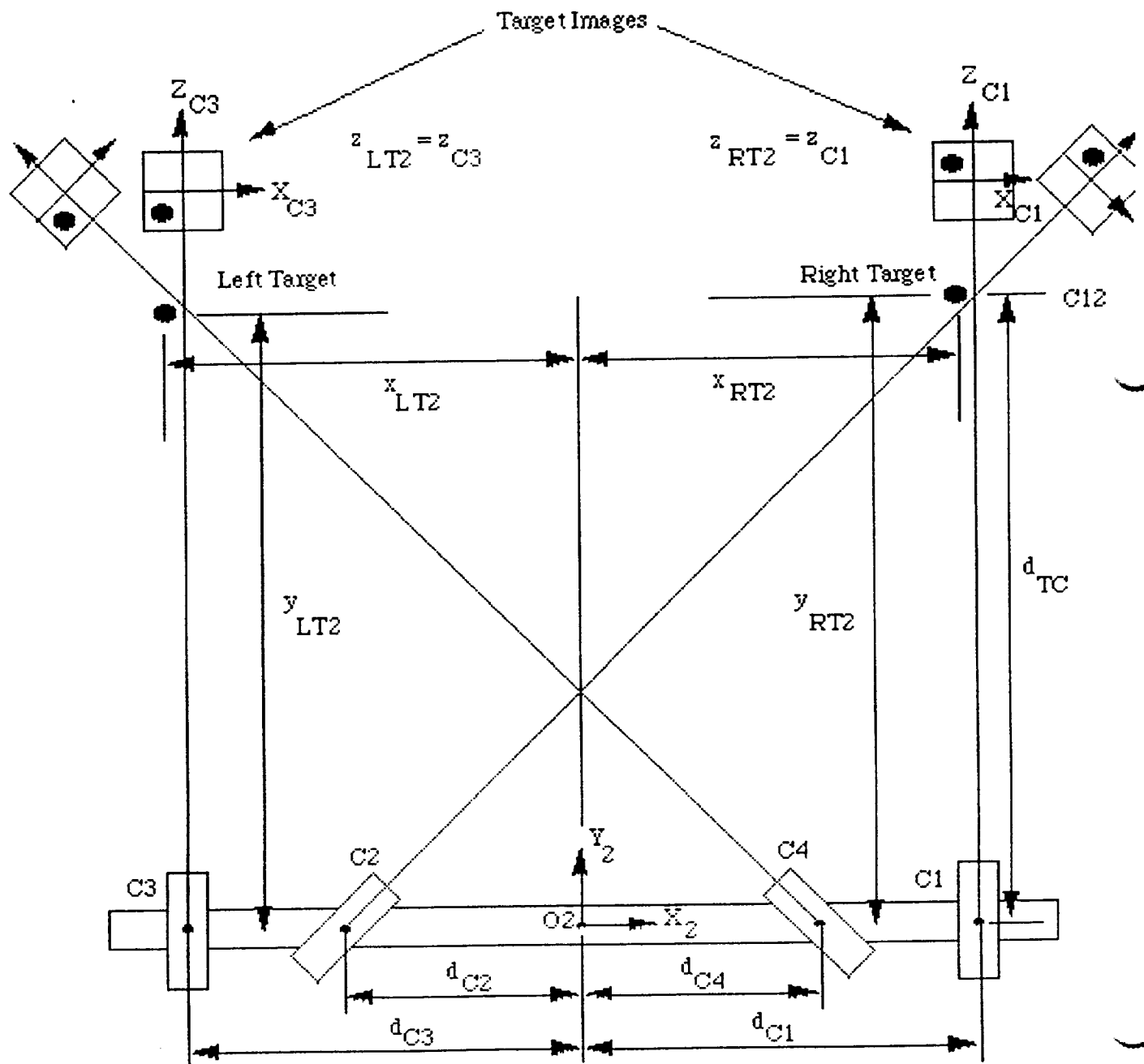


Figure 4

In Figure 4 the frame that carries the cameras of the Vision System is called the Camera frame. A coordinate system $O_2X_2Y_2Z_2$ is established at O_2 . X_2Y_2 plane is the plane containing the axes of the all 4 cameras. It should be made sure that all the camera axes are co-planar with proper calibration. The camera frame is called frame 2.

In Figure 4, camera 1 and camera 2 are used to get the coordinates of the right target. The coordinates obtained by camera 1 are x_{C1} and z_{C1} which are perpendicular to the axis of the camera. Let the distance of right target from the camera 1 be y_{C1} which cannot be measured. The coordinates obtained by camera 2 are x_{C2} and z_{C2} . Let the distance from the camera 2 to the right target by y_{C2} , which cannot be measured. In the following derivation we determine y_{C1} by eliminating y_{C2} .

$z_{C1} := 1$	z coordinate of Right Target image centroid from Camera 1
$x_{C1} := -2$	x coordinate of Right Target image centroid from Camera 1
$x_{C2} := -.707$	x coordinate of Right Target image centroid from Camera 2
$d_{TC} := 20$	Nominal Distance of the Target from the line of Cameras
	Distance of Target from Camera frame
$d_{C1} := 10$	Distance of Camera 1 from origin of frame 2 (camera frame)
$d_{C2} := 10$	Distance of Camera 2 from origin of frame 2 (camera frame)
$d_{C12} := d_{C1} + d_{C2}$	Distance between camera 1 and camera 2
$\alpha_{C12} := \text{atan}\left(\frac{d_{C12}}{d_{TC}}\right)$	Angle between axes of camera 1 and camera 2

The coordinates of the right target in frame #2 situated as shown in Fig. 3 are given by:

Using the output of camera 1

$$x_{RT2} = x_{C1} + d_{C1} \quad [1]$$

$$y_{RT2} = y_{C1} \quad [3]$$

$$z_{RT2} = z_{C1} \quad [5]$$

Using the output of camera 2

$$x_{RT2} = x_{C2} \cdot \cos(\alpha_{12}) + y_{C2} \cdot \sin(\alpha_{12}) - d_{C2} \quad [2]$$

$$y_{RT2} = -x_{C2} \cdot \sin(\alpha_{12}) + y_{C2} \cdot \cos(\alpha_{12}) \quad [4]$$

$$z_{RT2} = z_{C2} \quad [6]$$

Equating the right hand sides of equations [1] and [2], and rearranging, we have:

$$y_{C2} = \frac{x_{C1} - x_{C2} \cdot \cos(\alpha_{C12}) + d_{C1} + d_{C2}}{\sin(\alpha_{C12})}$$

or

$$y_{C2} = \frac{(x_{C1} - x_{C2} \cdot \cos(\alpha_{C12})) + d_{C12}}{\sin(\alpha_{C12})}$$

Substituting the above expression for y_{C2} into equation [4], we have

$$y_{RT2} = -x_{C2} \cdot \sin(\alpha_{C12}) + \frac{(x_{C1} - x_{C2} \cdot \cos(\alpha_{C12})) + d_{C12}}{\sin(\alpha_{C12})} \cdot \cos(\alpha_{C12})$$

$$y_{RT2} = \frac{-x_{C2} \cdot \sin(\alpha_{C12})^2 + x_{C1} \cdot \cos(\alpha_{C12}) - x_{C2} \cdot \cos(\alpha_{C12})^2 + d_{C12} \cdot \cos(\alpha_{C12})}{\sin(\alpha_{C12})}$$

$$y_{RT2} = \frac{-x_{C2} + x_{C1} \cdot \cos(\alpha_{C12}) + d_{C12} \cdot \cos(\alpha_{C12})}{\sin(\alpha_{C12})}$$

$$y_{RT2} = \frac{(d_{C12} + x_{C1}) \cdot \cos(\alpha_{C12}) - x_{C2}}{\sin(\alpha_{C12})}$$

Coordinates of Right Target in frame 2 are:

$$\begin{pmatrix} x_{RT2} \\ y_{RT2} \\ z_{RT2} \end{pmatrix} := \begin{bmatrix} x_{C1} + d_{C1} \\ \frac{(d_{C12} + x_{C1}) \cdot \cos(\alpha_{C12}) - x_{C2}}{\sin(\alpha_{C12})} \\ z_{C1} \end{bmatrix} \quad \begin{pmatrix} x_{RT2} \\ y_{RT2} \\ z_{RT2} \end{pmatrix} = \begin{pmatrix} 8 \\ 18.9998 \\ 1 \end{pmatrix}$$

In Figure 4, camera 3 and camera 4 are used to get the coordinates of the left target. The coordinates obtained by camera 3 are x_{C3} and z_{C3} which are perpendicular to the axis of the camera. Let the distance of left target from the camera 3 be y_{C3} which cannot be measured. The coordinates obtained by camera 4 are x_{C4} and z_{C4} . Let the distance from the camera 4 to the left target be y_{C4} , which cannot be measured. y_{C3} is determined by eliminating y_{C4} in the same way we did as for the right target. Following similar substitutions and simplifications, we get the coordinates of the left target in frame 2.

$z_{C3} := -1$	z coordinate of Left Target image centroid from Camera 3
$x_{C3} := -2$	x coordinate of Left Target image centroid from Camera 3
$x_{C4} := -2.121$	x coordinate of Left Target image centroid from Camera 4
$d_{C3} := 10$	Distance of Camera 3 from origin of frame 2 (camera frame)
$d_{C4} := 10$	Distance of Camera 4 from origin of frame 2 (camera frame)
$d_{C34} := d_{C3} + d_{C4}$	Distance between camera 3 and camera 4
$\alpha_{C34} := \text{atan}\left(\frac{d_{C34}}{d_{TC}}\right)$	Angle between the axes of camera 3 and camera 4

The coordinates of the left target in frame #2 situated as shown in Figure 4 are given by:

Using the output of camera 3

$$\begin{aligned} x_{LT2} &= x_{C3} - d_{C3} \\ y_{LT2} &= y_{C3} \\ z_{LT2} &= z_{C3} \end{aligned}$$

Using the output of camera 4

$$\begin{aligned} x_{LT2} &= x_{C4} \cdot \cos(\alpha_{C34}) - y_{C4} \cdot \sin(\alpha_{C34}) + d_{C4} \\ y_{LT2} &= x_{C4} \cdot \sin(\alpha_{C34}) + y_{C4} \cdot \cos(\alpha_{C34}) \\ z_{LT2} &= z_{C4} \end{aligned}$$

y_{C4} in the above is eliminated by following the procedure adopted for the right target.

Coordinates of Left Target in frame 2 are:

$$\begin{pmatrix} x_{LT2} \\ y_{LT2} \\ z_{LT2} \end{pmatrix} := \begin{pmatrix} x_{C3} - d_{C3} \\ \frac{(d_{C34} - x_{C3}) \cdot \cos(\alpha_{C34}) + x_{C4}}{\sin(\alpha_{C34})} \\ z_{C3} \end{pmatrix} \quad \begin{pmatrix} x_{LT2} \\ y_{LT2} \\ z_{LT2} \end{pmatrix} = \begin{pmatrix} -12 \\ 19.0005 \\ -1 \end{pmatrix}$$

4. DETERMINATION OF COORDINATE SYSTEM ATTACHED TO THE RACK INTERFACE PLATE ROTATION AXIS.

Derivation of the equations for determining the frame of reference 1 attached to the Rack Interface Plate rotation axis at a special point (CE) on the axis joining the left and right end points (LE and RE) of the end effector. This special point is the foot of the perpendicular from the top joint (P) on the rack onto the axis of the end effector endpoints. See Figure 3 and Figure 6 for details.

- $m := 10$ Distance between CE and CR
- $h := 30$ Distance between CE and P
- $g := \sqrt{m^2 + h^2}$ Distance between CR and P $g = 31.6228$
- $r_{GP} := 15$ Distance between G and P

G is a reference point on the boom where the end effector mechanism is connected, as shown in Figure 3 & Figure 6.

Coordinates of point G:

$$\begin{pmatrix} x_G \\ y_G \\ z_G \end{pmatrix} := \begin{pmatrix} 90 \\ 0 \\ 0 \end{pmatrix}$$

In the following, r_{AB} indicates the magnitude of vector AB, and x_{AB} , y_{AB} , and z_{AB} indicate the x,y,z coordinates of the vector AB. The points that join the vectors are shown in Fig. 5.

Further, the "R" and "L" in the subscripts always means "Right" and the "Left" respectively. For example, x_{RB} indicates the x coordinate of Right B (point B on the right side).

- $r_{RBD} := 5.25$ $r_{RBF} := 3.125$ $r_{RFC} := 23.25$
- $r_{LBD} := 5.25$ $r_{LBF} := 3.125$ $r_{LFC} := 23.25$

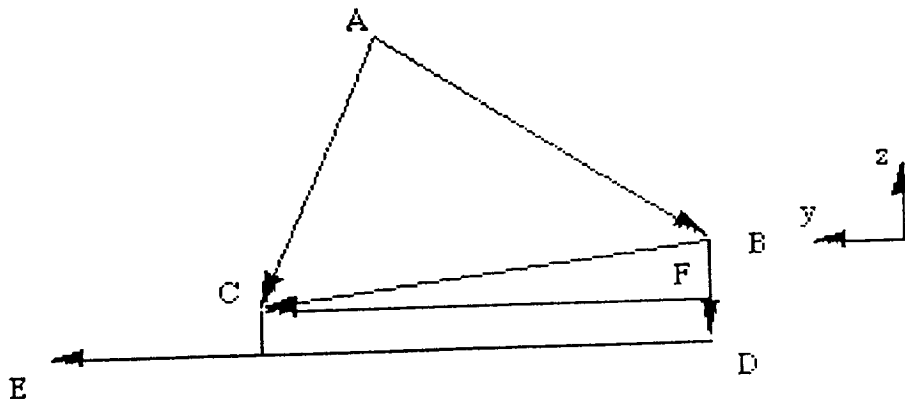


Figure 5

$$r_{RBC} := \sqrt{r_{RBF}^2 + r_{RFC}^2}$$

$$r_{RBC} = 23.4591$$

$$r_{LBC} := \sqrt{r_{LBF}^2 + r_{LFC}^2}$$

$$r_{LBC} = 23.4591$$

Coordinates of Right B

$$\begin{pmatrix} x_{RB} \\ y_{RB} \\ z_{RB} \end{pmatrix} := \begin{pmatrix} 100 \\ 0 \\ 0 \end{pmatrix}$$

Coordinates of Left B

$$\begin{pmatrix} x_{LB} \\ y_{LB} \\ z_{LB} \end{pmatrix} := \begin{pmatrix} 80 \\ 0 \\ 0 \end{pmatrix}$$

Components and magnitude of vectors AB

$$y_{RAB} := -17 \quad z_{RAB} := -11$$

$$y_{LAB} := -17 \quad z_{LAB} := -11$$

$$r_{RAB} := \sqrt{y_{RAB}^2 + z_{RAB}^2}$$

$$r_{RAB} = 20.2485$$

$$r_{LAB} := \sqrt{y_{LAB}^2 + z_{LAB}^2}$$

$$r_{LAB} = 20.2485$$

Nominal lengths of AC and DE

$$r_{RACnom} := 14.75 \quad r_{RDEnom} := 34.5$$

$$r_{LACnom} := 14.75 \quad r_{LDEnom} := 34.5$$

Increments in AC nominal and DE nominal (right and left)

$$\Delta_{RAC} := 0$$

$$\Delta_{LAC} := 0$$

$$\Delta_{RDE} := 0$$

$$\Delta_{LDE} := 0$$

$$r_{RAC} := r_{RACnom} + \Delta_{RAC}$$

$$r_{RAC} = 14.75$$

$$r_{LAC} := r_{LACnom} + \Delta_{LAC}$$

$$r_{LAC} = 14.75$$

$$r_{RDE} := r_{RDEnom} + \Delta_{RDE}$$

$$r_{RDE} = 34.5$$

$$r_{LDE} := r_{LDEnom} + \Delta_{LDE}$$

$$r_{LDE} = 34.5$$

The sets of points A,B and C on the right and left sides of the RIEE mechanism will each form a triangle. Here we derive the equations for finding the angle of the general (common to left and right) vector BC with known coordinates of points A and B, and the length AC in the y-z plane. Recalling Figure 5, we see:

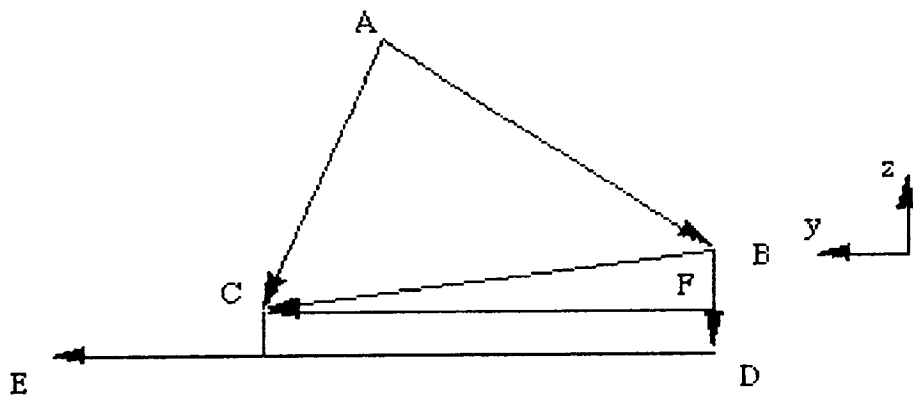


Figure 5

$$\text{Vector AC} = \text{Vector AB} + \text{Vector BC}$$

$$y_{AC} = y_{AB} + r_{BC} \cos(\theta_{BC}) \quad \text{and} \quad z_{AC} = z_{AB} + r_{BC} \sin(\theta_{BC})$$

Squaring and adding the above equations and simplifying, we have:

$$r_{AC}^2 = r_{AB}^2 + r_{BC}^2 + 2 \cdot y_{AB} \cdot r_{BC} \cos(\theta_{BC}) + 2 \cdot z_{AB} \cdot r_{BC} \sin(\theta_{BC})$$

Dividing by $2r_{BC}$ and further rearranging, results in:

$$y_{AB} \cos(\theta_{BC}) + z_{AB} \sin(\theta_{BC}) + \frac{r_{AB}^2 + r_{BC}^2 - r_{AC}^2}{2 \cdot r_{BC}} = 0$$

$$\text{Let } c_{BC} = \frac{r_{AB}^2 + r_{BC}^2 - r_{AC}^2}{2 \cdot r_{BC}}$$

$$\text{Hence } y_{AB} \cdot \cos(\theta_{BC}) + z_{AB} \cdot \sin(\theta_{BC}) + c_{BC} = 0$$

The above equation is of the form:

$$a \cdot \cos(\theta) + b \cdot \sin(\theta) + c = 0$$

Expanding the above equation using half angles, we have:

$$a \cdot \left(\cos\left(\frac{\theta}{2}\right)^2 - \sin\left(\frac{\theta}{2}\right)^2 \right) + 2 \cdot b \cdot \sin\left(\frac{\theta}{2}\right) \cdot \cos\left(\frac{\theta}{2}\right) + c \cdot \left(\cos\left(\frac{\theta}{2}\right)^2 + \sin\left(\frac{\theta}{2}\right)^2 \right) = 0$$

Dividing the above by $\cos\left(\frac{\theta}{2}\right)^2$ and letting $\tan\left(\frac{\theta}{2}\right) = t$ we have

$$a \cdot (1 - t^2) + 2 \cdot b \cdot t + c \cdot (1 + t^2) = 0 \quad \text{or} \quad (-a + c) \cdot t^2 + 2 \cdot b \cdot t + (a + c) = 0$$

$$\text{or} \quad (a - c) \cdot t^2 - 2 \cdot b \cdot t - (a + c) = 0$$

The above is a quadratic equation in "t" and has two solutions. The solution that is valid according to the given geometry is

$$t = \frac{2 \cdot b + \sqrt{4 \cdot b^2 + 4 \cdot (a^2 - c^2)}}{2 \cdot (a - c)} \quad \text{or} \quad t = \frac{b + \sqrt{a^2 + b^2 - c^2}}{(a - c)}$$

$$\text{Therefore } \theta = 2 \cdot \text{atan}(t)$$

The trigonometric equations to solve for the angle of vector BC in the triangles formed by points A, B and C on the right and left sides of the RIEE are:

$$y_{RAB} \cdot \cos(\theta_{RBC}) + z_{RAB} \cdot \sin(\theta_{RBC}) + c_{RBC} = 0 \quad \text{and}$$

$$y_{LAB} \cdot \cos(\theta_{LBC}) + z_{LAB} \cdot \sin(\theta_{LBC}) + c_{LBC} = 0$$

Where

$$c_{RBC} := \frac{r_{RAB}^2 + r_{RBC}^2 - r_{RAC}^2}{2 \cdot r_{RBC}}$$

$$c_{RBC} = 15.8311$$

$$c_{LBC} := \frac{r_{LAB}^2 + r_{LBC}^2 - r_{LAC}^2}{2 \cdot r_{LBC}}$$

$$c_{LBC} = 15.8311$$

$$\theta_{RBC} := 2 \cdot \operatorname{atan} \left(\frac{z_{RAB} + \sqrt{r_{RAB}^2 - c_{RBC}^2}}{y_{RAB} - c_{RBC}} \right)$$

$$\theta_{RBC} = -5.6652 \cdot \text{deg}$$

$$\theta_{LBC} := 2 \cdot \operatorname{atan} \left(\frac{z_{LAB} + \sqrt{r_{LAB}^2 - c_{LBC}^2}}{y_{LAB} - c_{LBC}} \right)$$

$$\theta_{LBC} = -5.6652 \cdot \text{deg}$$

The points B,C and F on the left and right sides form a right triangle and vector FC is parallel to vector DE, therefore, the following relationships are derived to compute the angles of DEs (left and right) from the positive y axis in the y-z plane.

$$\lambda_R := \operatorname{atan} \left(\frac{r_{RBF}}{r_{RFC}} \right)$$

$$\lambda_R = 7.6552 \cdot \text{deg}$$

$$\lambda_L := \operatorname{atan} \left(\frac{r_{LBF}}{r_{LFC}} \right)$$

$$\lambda_L = 7.6552 \cdot \text{deg}$$

$$\theta_{RDE} := \theta_{RBC} + \lambda_R$$

$$\theta_{RDE} = 1.99 \cdot \text{deg}$$

$$\theta_{LDE} := \theta_{LBC} + \lambda_L$$

$$\theta_{LDE} = 1.99 \cdot \text{deg}$$

Vector RE = Vector RB + Vector RBD + Vector RDE

$$\theta_{RBD} = \left(\theta_{RDE} - \frac{\pi}{2} \right)$$

Therefore, the coordinates of point RE are:

$$\begin{pmatrix} x_{RE} \\ y_{RE} \\ z_{RE} \end{pmatrix} := \begin{pmatrix} x_{RB} \\ y_{RB} + r_{RBD} \cdot \cos \left(\theta_{RDE} - \frac{\pi}{2} \right) + r_{RDE} \cdot \cos \left(\theta_{RDE} \right) \\ z_{RB} + r_{RBD} \cdot \sin \left(\theta_{RDE} - \frac{\pi}{2} \right) + r_{RDE} \cdot \sin \left(\theta_{RDE} \right) \end{pmatrix} \quad \begin{pmatrix} x_{RE} \\ y_{RE} \\ z_{RE} \end{pmatrix} = \begin{pmatrix} 100 \\ 34.6615 \\ -4.0488 \end{pmatrix}$$

Vector LE = Vector LB + Vector LBD + Vector LDE

$$\theta_{LBD} = \left(\theta_{LDE} - \frac{\pi}{2} \right)$$

Therefore, the coordinates of the point LE are:

$$\begin{pmatrix} x_{LE} \\ y_{LE} \\ z_{LE} \end{pmatrix} := \begin{pmatrix} x_{LB} \\ y_{LB} + r_{LBD} \cdot \cos\left(\theta_{LDE} - \frac{\pi}{2}\right) + r_{LDE} \cdot \cos(\theta_{LDE}) \\ z_{LB} + r_{LBD} \cdot \sin\left(\theta_{LDE} - \frac{\pi}{2}\right) + r_{LDE} \cdot \sin(\theta_{LDE}) \end{pmatrix} \quad \begin{pmatrix} x_{LE} \\ y_{LE} \\ z_{LE} \end{pmatrix} = \begin{pmatrix} 80 \\ 34.6615 \\ -4.0488 \end{pmatrix}$$

Components of the vector from Left end LE to right end RE are:

$$\begin{pmatrix} x_{LRE} \\ y_{LRE} \\ z_{LRE} \end{pmatrix} := \begin{pmatrix} x_{RE} \\ y_{RE} \\ z_{RE} \end{pmatrix} - \begin{pmatrix} x_{LE} \\ y_{LE} \\ z_{LE} \end{pmatrix} \quad \begin{pmatrix} x_{LRE} \\ y_{LRE} \\ z_{LRE} \end{pmatrix} = \begin{pmatrix} 20 \\ 0 \\ 0 \end{pmatrix}$$

$$r_{LRE} := \sqrt{x_{LRE}^2 + y_{LRE}^2 + z_{LRE}^2} \quad r_{LRE} = 20$$

The Components of Unit vector U1 along the line from left end LE to right end RE

$$\begin{pmatrix} x_{U1} \\ y_{U1} \\ z_{U1} \end{pmatrix} := \begin{pmatrix} x_{LRE} \\ y_{LRE} \\ z_{LRE} \end{pmatrix} \cdot \frac{1}{r_{LRE}} \quad \begin{pmatrix} x_{U1} \\ y_{U1} \\ z_{U1} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

Point CE is foot of the perpendicular from the top joint P on Rack Inserting Frame, to the line joining LE and RE. From the geometry it is m units from RE along the negative direction of the unit vector U1.

The Coordinates of the point CE are:

$$\begin{pmatrix} x_{CE} \\ y_{CE} \\ z_{CE} \end{pmatrix} := \begin{pmatrix} x_{RE} \\ y_{RE} \\ z_{RE} \end{pmatrix} - m \cdot \begin{pmatrix} x_{U1} \\ y_{U1} \\ z_{U1} \end{pmatrix} \quad \begin{pmatrix} x_{CE} \\ y_{CE} \\ z_{CE} \end{pmatrix} = \begin{pmatrix} 90 \\ 34.6615 \\ -4.0488 \end{pmatrix}$$

Our intention now is to establish a Coordinate system attached to the Rack Interface Plate with Unit vectors U1, V1 & W1 along its x,y & z axes and Point CE as the origin.. U1 is already defined above. V1 will be defined along the line from CE to P.

Our current objective is to determine the coordinates of the point P. We will find the coordinates by writing the equations defining its distance from three points in the RIEE mechanism. These three points are the reference point G, the right end point RE, and the point CE as defined above, and shown below.

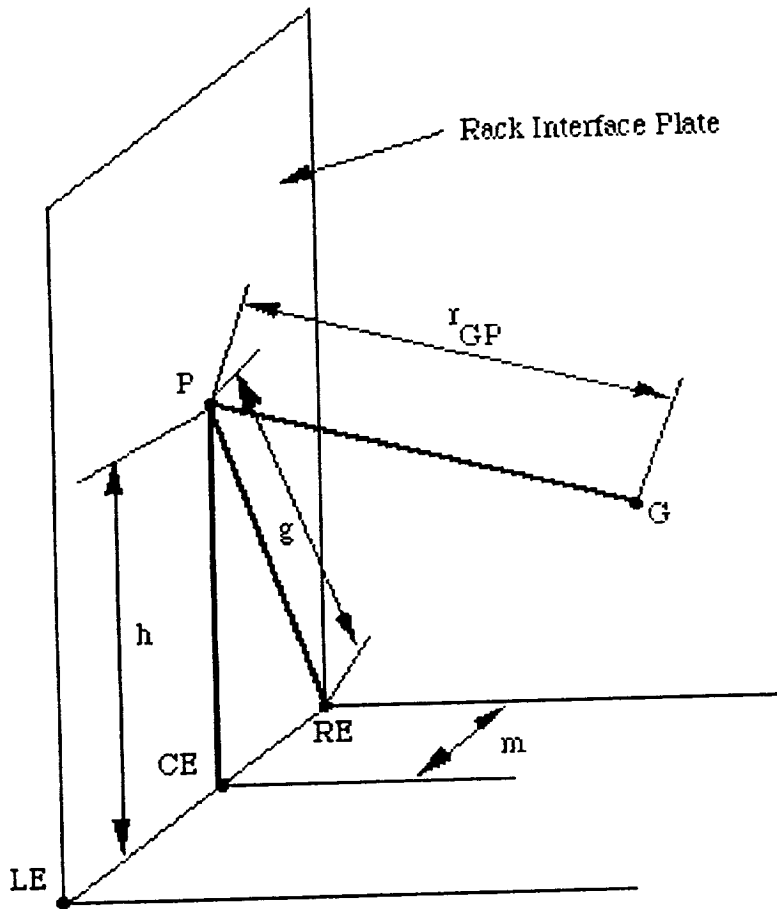


Figure 6

The equations are:

$$(x_P - x_G)^2 + (y_P - y_G)^2 + (z_P - z_G)^2 = r_{GP}^2 \quad [7]$$

$$(x_P - x_{CE})^2 + (y_P - y_{CE})^2 + (z_P - z_{CE})^2 = h^2 \quad [8]$$

$$(x_P - x_{RE})^2 + (y_P - y_{RE})^2 + (z_P - z_{RE})^2 = g^2 \quad [9]$$

Subtracting equations [8] and [9] from equation [7], we get the following equations:

$$a_1 \cdot x_P + b_1 \cdot y_P + c_1 \cdot z_P + G_2 - x_{CE}^2 - y_{CE}^2 - z_{CE}^2 = r_{GP}^2 - h^2 \quad [10]$$

$$a_2 \cdot x_P + b_2 \cdot y_P + c_2 \cdot z_P + G_2 - x_{RE}^2 - y_{RE}^2 - z_{RE}^2 = r_{GP}^2 - g^2 \quad [11]$$

where

$$\begin{aligned}
 a_1 &:= 2 \cdot (x_{CE} - x_G) & a_1 &= 0 \\
 a_2 &:= 2 \cdot (x_{RE} - x_G) & a_2 &= 20 \\
 b_1 &:= 2 \cdot (y_{CE} - y_G) & b_1 &= 69.323 \\
 b_2 &:= 2 \cdot (y_{RE} - y_G) & b_2 &= 69.323 \\
 c_1 &:= 2 \cdot (z_{CE} - z_G) & c_1 &= -8.0977 \\
 c_2 &:= 2 \cdot (z_{RE} - z_G) & c_2 &= -8.0977 \\
 G_2 &:= x_G^2 + y_G^2 + z_G^2 & G_2 &= 8.1 \cdot 10^3
 \end{aligned}$$

Equations [10] and [11] can be rearranged as:

$$a_1 \cdot x_P + c_1 \cdot z_P = -b_1 \cdot y_P + k_1 \quad [12]$$

$$a_2 \cdot x_P + c_2 \cdot z_P = -b_2 \cdot y_P + k_2 \quad [13]$$

where

$$\begin{aligned}
 k_1 &:= r_{GP}^2 - h^2 + x_{CE}^2 + y_{CE}^2 + z_{CE}^2 - G_2 & k_1 &= 542.8125 \\
 k_2 &:= r_{GP}^2 - g^2 + x_{RE}^2 + y_{RE}^2 + z_{RE}^2 - G_2 & k_2 &= 2.3428 \cdot 10^3
 \end{aligned}$$

Equations [12] and [13] can be arranged in the following matrix form:

$$\begin{pmatrix} a_1 & c_1 \\ a_2 & c_2 \end{pmatrix} \cdot \begin{pmatrix} x_P \\ z_P \end{pmatrix} = \begin{pmatrix} -b_1 \cdot y_P + k_1 \\ -b_2 \cdot y_P + k_2 \end{pmatrix} \quad [14]$$

$$D_{12} := \left| \begin{pmatrix} a_1 & c_1 \\ a_2 & c_2 \end{pmatrix} \right| \quad D_{12} = 161.9539$$

Using Cramer's Rule, we can solve for x_P and z_P

$$x_P = \frac{\begin{pmatrix} -b_1 \cdot y_P + k_1 & c_1 \\ -b_2 \cdot y_P + k_2 & c_2 \end{pmatrix}}{D_{12}} \quad z_P = \frac{\begin{pmatrix} a_1 & -b_1 \cdot y_P + k_1 \\ a_2 & -b_2 \cdot y_P + k_2 \end{pmatrix}}{D_{12}}$$

After simplifying, we have

$$x_P = b_3 \cdot y_P + k_3 \quad [15]$$

$$z_P = b_4 \cdot y_P + k_4 \quad [16]$$

where

$$b_3 := \frac{-b_1 \cdot c_2 + b_2 \cdot c_1}{D_{12}}$$

$$b_4 := \frac{-a_1 \cdot b_2 + a_2 \cdot b_1}{D_{12}}$$

$$k_3 := \frac{k_1 \cdot c_2 - k_2 \cdot c_1}{D_{12}}$$

$$k_4 := \frac{a_1 \cdot k_2 - a_2 \cdot k_1}{D_{12}}$$

$$b_3 = 0$$

$$b_4 = 8.5608$$

$$k_3 = 90$$

$$k_4 = -67.033$$

Recalling equation [7] and substituting for x_P and z_P from equations [15] and [16], we have:

$$(x_P - x_G)^2 + (y_P - y_G)^2 + (z_P - z_G)^2 = r_{GP}^2 \quad [17]$$

$$(b_3 \cdot y_P + k_3 - x_G)^2 + (y_P - y_G)^2 + (b_4 \cdot y_P + k_4 - z_G)^2 = r_{GP}^2 \quad [18]$$

Expanding equation [17] and rearranging, we get:

$$A \cdot y_P^2 + B \cdot y_P + C = 0$$

where

$$A := b_3^2 + 1 + b_4^2$$

$$A = 74.2878$$

$$B := 2 \cdot [b_3 \cdot (k_3 - x_G) - y_G + b_4 \cdot (k_4 - z_G)]$$

$$B = -1.1477 \cdot 10^3$$

$$C := (k_3 - x_G)^2 + y_G^2 + (k_4 - z_G)^2 - r_{GP}^2$$

$$C = 4.2684 \cdot 10^3$$

Equation [18] is a quadratic in "y_P" and has 2 solutions. The solution that is relevant to the geometry of the RIEE mechanism is:

$$y_P := \left(\frac{-B + \sqrt{B^2 - 4 \cdot A \cdot C}}{2 \cdot A} \right)$$

$$y_P = 9.2129$$

Substituting y_P back into equations [15] and [16], the coordinates of P are given by:

$$\begin{pmatrix} x_P \\ y_P \\ z_P \end{pmatrix} := \begin{pmatrix} b_3 \cdot y_P + k_3 \\ y_P \\ b_4 \cdot y_P + k_4 \end{pmatrix} \quad \begin{pmatrix} x_P \\ y_P \\ z_P \end{pmatrix} = \begin{pmatrix} 90 \\ 9.2129 \\ 11.8373 \end{pmatrix}$$

The z-axis given by unit vector W1 is along the line joining CE and P. Its components are given by:

$$\begin{pmatrix} x_{W1} \\ y_{W1} \\ z_{W1} \end{pmatrix} := \begin{pmatrix} x_P - x_{CE} \\ y_P - y_{CE} \\ z_P - z_{CE} \end{pmatrix} \cdot \frac{1}{h} \quad \begin{pmatrix} x_{W1} \\ y_{W1} \\ z_{W1} \end{pmatrix} = \begin{pmatrix} 0 \\ -0.8483 \\ 0.5295 \end{pmatrix}$$

The y-axis given by unit vector V1 is given by the cross product of vector W1 and vector U1. Its components are given by:

$$\begin{pmatrix} x_{V1} \\ y_{V1} \\ z_{V1} \end{pmatrix} := \begin{pmatrix} x_{W1} \\ y_{W1} \\ z_{W1} \end{pmatrix} \times \begin{pmatrix} x_{U1} \\ y_{U1} \\ z_{U1} \end{pmatrix} \quad \begin{pmatrix} x_{V1} \\ y_{V1} \\ z_{V1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0.5295 \\ 0.8483 \end{pmatrix}$$

Thus we have completely specified the coordinate system located at CE.

We now find the relationship between this coordinate system and a coordinate system located on the camera frame.

The x-axis of the camera frame defined by U2 is taken same as the U1, and the z-axis is defined by the unit vector W2, which is at an angle β in the clockwise direction from the unit vector W1 about the x-axis.

The origin of the camera frame is situated at a distance of d_{21} along the negative direction of the unit vector W2 from CE. Therefore:

$$\beta := \text{atan} \left(\frac{y_{CE} - y_P}{z_P - z_{CE}} \right) \quad \beta = 58.0257 \cdot \text{deg}$$

$$h_C := 12$$

The components of the unit vectors of the camera frame (2) as seen from rack insertion frame (1) are given by:

$$\begin{pmatrix} x_{U21} \\ y_{U21} \\ z_{U21} \end{pmatrix} := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \begin{pmatrix} x_{V21} \\ y_{V21} \\ z_{V21} \end{pmatrix} := \begin{pmatrix} 0 \\ \cos(\beta) \\ -\sin(\beta) \end{pmatrix} \quad \begin{pmatrix} x_{W21} \\ y_{W21} \\ z_{W21} \end{pmatrix} := \begin{pmatrix} 0 \\ \sin(\beta) \\ \cos(\beta) \end{pmatrix}$$

The coordinates of the origin of the camera frame O_2 as seen from Rack Interface frame (1) are given by:

$$\begin{pmatrix} x_{O21} \\ y_{O21} \\ z_{O21} \end{pmatrix} := \begin{pmatrix} 0 \\ -h_C \sin(\beta) \\ -h_C \cos(\beta) \end{pmatrix}$$

The transformation of coordinates of the right and left targets from the camera frame to the rack insertion frame is given by the following matrix equations:

Note that in the 4×4 transformation matrix on the right hand side:
 the first column contains the components of the unit vector U_2 as seen in frame 1,
 the second column contains the components of the unit vector V_2 as seen in frame 1,
 the third column contains the components of the unit vector W_2 as seen in frame 1, and
 the fourth column contains the coordinates of the origin of the camera frame as seen in frame 1.

$$\begin{bmatrix} x_{RT1} \\ y_{RT1} \\ z_{RT1} \\ S_{21} \end{bmatrix} := \begin{pmatrix} x_{U21} & x_{V21} & x_{W21} & x_{O21} \\ y_{U21} & y_{V21} & y_{W21} & y_{O21} \\ z_{U21} & z_{V21} & z_{W21} & z_{O21} \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_{RT2} \\ y_{RT2} \\ z_{RT2} \\ 1 \end{pmatrix} \quad \begin{bmatrix} x_{RT1} \\ y_{RT1} \\ z_{RT1} \\ S_{21} \end{bmatrix} = \begin{bmatrix} 8 \\ 0.73 \\ -21.9422 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} x_{LT1} \\ y_{LT1} \\ z_{LT1} \\ S_{21} \end{bmatrix} := \begin{pmatrix} x_{U21} & x_{V21} & x_{W21} & x_{O21} \\ y_{U21} & y_{V21} & y_{W21} & y_{O21} \\ z_{U21} & z_{V21} & z_{W21} & z_{O21} \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_{LT2} \\ y_{LT2} \\ z_{LT2} \\ 1 \end{pmatrix} \quad \begin{bmatrix} x_{LT1} \\ y_{LT1} \\ z_{LT1} \\ S_{21} \end{bmatrix} = \begin{bmatrix} -12 \\ -0.9662 \\ -23.0018 \\ 1 \end{bmatrix}$$

5. SUPPORT COORDINATES IN BOOM COORDINATE SYSTEM:

$h_T := 12$ Height of the supports above the targets as in Figure 3. Support Coordinates are:

$$\begin{bmatrix} x_{RS} \\ y_{RS} \\ z_{RS} \\ S_1 \end{bmatrix} := \begin{pmatrix} x_{U1} & x_{V1} & x_{W1} & x_{CE} \\ y_{U1} & y_{V1} & y_{W1} & y_{CE} \\ z_{U1} & z_{V1} & z_{W1} & z_{CE} + h_T \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_{RT1} \\ y_{RT1} \\ z_{RT1} \\ 1 \end{pmatrix} \quad \begin{bmatrix} x_{RS} \\ y_{RS} \\ z_{RS} \\ S_1 \end{bmatrix} = \begin{bmatrix} 98 \\ 53.6613 \\ -3.0488 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} x_{LS} \\ y_{LS} \\ z_{LS} \\ S_1 \end{bmatrix} := \begin{pmatrix} x_{U1} & x_{V1} & x_{W1} & x_{CE} \\ y_{U1} & y_{V1} & y_{W1} & y_{CE} \\ z_{U1} & z_{V1} & z_{W1} & z_{CE} + h_T \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_{LT1} \\ y_{LT1} \\ z_{LT1} \\ 1 \end{pmatrix} \quad \begin{bmatrix} x_{LS} \\ y_{LS} \\ z_{LS} \\ S_1 \end{bmatrix} = \begin{bmatrix} 78 \\ 53.6619 \\ -5.0488 \\ 1 \end{bmatrix}$$

6. DISPLACEMENTS IN THE BOOM ROBOT AND THE RIEE MECHANISM.

The center point CS on the line joining the left and right lower support points is given by

$$\begin{pmatrix} x_{CS} \\ y_{CS} \\ z_{CS} \end{pmatrix} := \begin{pmatrix} x_{RS} + x_{LS} \\ y_{RS} + y_{LS} \\ z_{RS} + z_{LS} \end{pmatrix} \cdot \frac{1}{2} \quad \begin{pmatrix} x_{CS} \\ y_{CS} \\ z_{CS} \end{pmatrix} = \begin{pmatrix} 88 \\ 53.6616 \\ -4.0488 \end{pmatrix}$$

Having found the coordinates of the supports and the midpoint, we move the boom such that the point CE gets the displacement in the x and z direction of the boom keeping the y-distance from the targets undisturbed. This displacement is shown exaggerated in Figure 7 below.

$$\begin{pmatrix} x_{CES} \\ y_{CES} \\ z_{CES} \end{pmatrix} := \begin{pmatrix} x_{CS} \\ y_{CS} \\ z_{CS} \end{pmatrix} - \begin{pmatrix} x_{CE} \\ y_{CE} \\ z_{CE} \end{pmatrix} \quad \begin{pmatrix} x_{CES} \\ y_{CES} \\ z_{CES} \end{pmatrix} = \begin{pmatrix} -2 \\ 19.0002 \\ -4.4409 \cdot 10^{-15} \end{pmatrix}$$

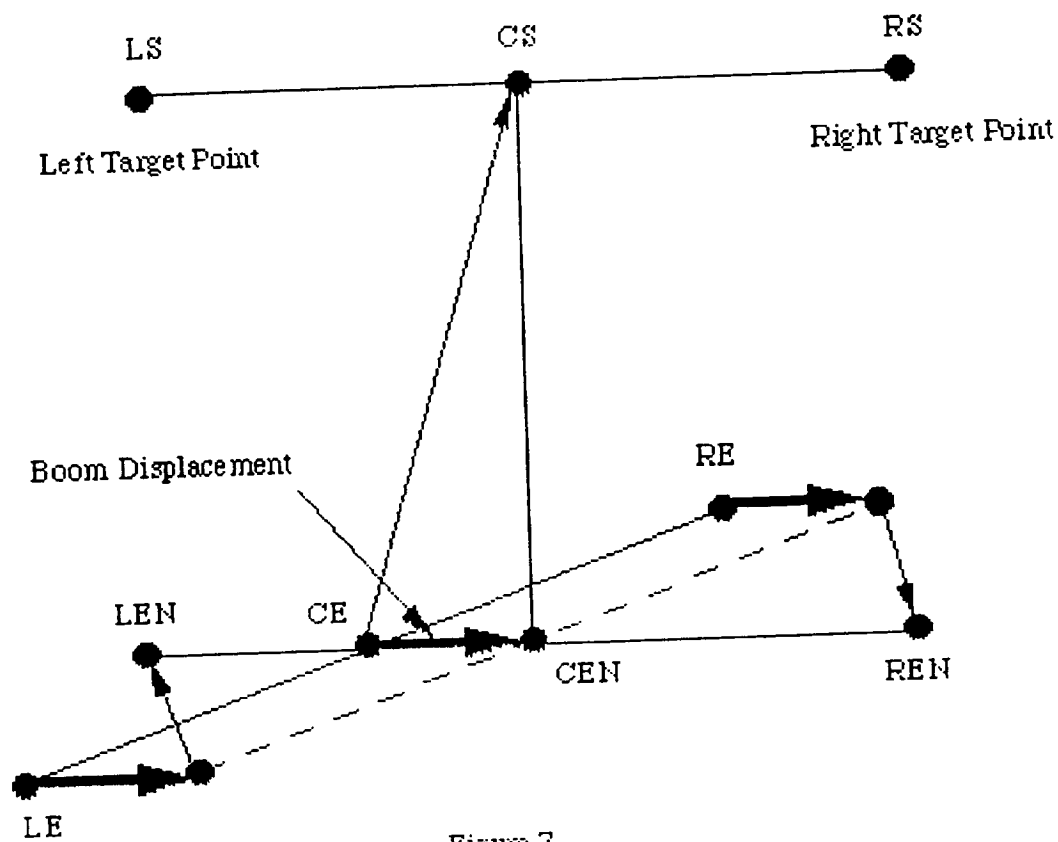


Figure 7

$$\begin{pmatrix} x_{Boom} \\ y_{Boom} \\ z_{Boom} \end{pmatrix} := \begin{pmatrix} x_{CES} \\ 0 \\ z_{CES} \end{pmatrix} \quad \begin{pmatrix} x_{Boom} \\ y_{Boom} \\ z_{Boom} \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \\ -4.4409 \cdot 10^{-15} \end{pmatrix}$$

The new coordinates of CE after the boom displacement

$$\begin{pmatrix} x_{\text{CEN}} \\ y_{\text{CEN}} \\ z_{\text{CEN}} \end{pmatrix} := \begin{pmatrix} x_{\text{CE}} \\ y_{\text{CE}} \\ z_{\text{CE}} \end{pmatrix} + \begin{pmatrix} x_{\text{Boom}} \\ y_{\text{Boom}} \\ z_{\text{Boom}} \end{pmatrix}$$

$$\begin{pmatrix} x_{\text{CEN}} \\ y_{\text{CEN}} \\ z_{\text{CEN}} \end{pmatrix} = \begin{pmatrix} 88 \\ 34.6615 \\ -4.0488 \end{pmatrix}$$

The new coordinates of Right and Left "B" are:

$$\begin{pmatrix} x_{\text{RBN}} \\ y_{\text{RBN}} \\ z_{\text{RBN}} \end{pmatrix} := \begin{pmatrix} x_{\text{RB}} \\ y_{\text{RB}} \\ z_{\text{RB}} \end{pmatrix} + \begin{pmatrix} x_{\text{Boom}} \\ y_{\text{Boom}} \\ z_{\text{Boom}} \end{pmatrix}$$

$$\begin{pmatrix} x_{\text{RBN}} \\ y_{\text{RBN}} \\ z_{\text{RBN}} \end{pmatrix} = \begin{pmatrix} 98 \\ 0 \\ -4.4409 \cdot 10^{-15} \end{pmatrix}$$

$$\begin{pmatrix} x_{\text{LBN}} \\ y_{\text{LBN}} \\ z_{\text{LBN}} \end{pmatrix} := \begin{pmatrix} x_{\text{LB}} \\ y_{\text{LB}} \\ z_{\text{LB}} \end{pmatrix} + \begin{pmatrix} x_{\text{Boom}} \\ y_{\text{Boom}} \\ z_{\text{Boom}} \end{pmatrix}$$

$$\begin{pmatrix} x_{\text{LBN}} \\ y_{\text{LBN}} \\ z_{\text{LBN}} \end{pmatrix} = \begin{pmatrix} 78 \\ 0 \\ -4.4409 \cdot 10^{-15} \end{pmatrix}$$

The new coordinates of Right "E" and components of new right BE derived from Figure 5 are:

$$\begin{pmatrix} y_{\text{REN}} \\ z_{\text{REN}} \end{pmatrix} := \begin{pmatrix} y_{\text{CEN}} \\ z_{\text{RS}} \end{pmatrix}$$

$$\begin{pmatrix} y_{\text{REN}} \\ z_{\text{REN}} \end{pmatrix} = \begin{pmatrix} 34.6615 \\ -3.0488 \end{pmatrix}$$

$$\begin{pmatrix} y_{\text{RBE}} \\ z_{\text{RBE}} \end{pmatrix} := \begin{pmatrix} y_{\text{RE}} - y_{\text{RB}} \\ z_{\text{RE}} - z_{\text{RB}} \end{pmatrix}$$

$$\begin{pmatrix} y_{\text{RBE}} \\ z_{\text{RBE}} \end{pmatrix} = \begin{pmatrix} 34.6615 \\ -4.0488 \end{pmatrix}$$

$$\begin{pmatrix} y_{\text{RBEN}} \\ z_{\text{RBEN}} \end{pmatrix} := \begin{pmatrix} y_{\text{REN}} \\ z_{\text{REN}} \end{pmatrix} - \begin{pmatrix} y_{\text{RBN}} \\ z_{\text{RBN}} \end{pmatrix}$$

$$\begin{pmatrix} y_{\text{RBEN}} \\ z_{\text{RBEN}} \end{pmatrix} = \begin{pmatrix} 34.6615 \\ -3.0488 \end{pmatrix}$$

The new coordinates of Left "E" and components of left BE derived from Figure 5 are:

$$\begin{pmatrix} y_{\text{LEN}} \\ z_{\text{LEN}} \end{pmatrix} := \begin{pmatrix} y_{\text{CEN}} \\ z_{\text{LS}} \end{pmatrix}$$

$$\begin{pmatrix} y_{\text{LEN}} \\ z_{\text{LEN}} \end{pmatrix} = \begin{pmatrix} 34.6615 \\ -5.0488 \end{pmatrix}$$

$$\begin{pmatrix} y_{\text{LBE}} \\ z_{\text{LBE}} \end{pmatrix} := \begin{pmatrix} y_{\text{LE}} - y_{\text{LB}} \\ z_{\text{LE}} - z_{\text{LB}} \end{pmatrix}$$

$$\begin{pmatrix} y_{\text{LBE}} \\ z_{\text{LBE}} \end{pmatrix} = \begin{pmatrix} 34.6615 \\ -4.0488 \end{pmatrix}$$

$$\begin{pmatrix} y_{LBEN} \\ z_{LBEN} \end{pmatrix} := \begin{pmatrix} y_{LEN} \\ z_{LEN} \end{pmatrix} - \begin{pmatrix} y_{LBN} \\ z_{LBN} \end{pmatrix} \qquad \begin{pmatrix} y_{LBEN} \\ z_{LBEN} \end{pmatrix} = \begin{pmatrix} 34.6615 \\ -5.0488 \end{pmatrix}$$

The lengths of the new BE and the new DE on the right and left sides are:

$$r_{RBEN} := \sqrt{y_{RBEN}^2 + z_{RBEN}^2} \qquad r_{RBEN} = 34.7953$$

$$r_{LBEN} := \sqrt{y_{LBEN}^2 + z_{LBEN}^2} \qquad r_{LBEN} = 35.0273$$

$$r_{RDEN} := \sqrt{r_{RBEN}^2 - r_{RBD}^2} \qquad r_{RDEN} = 34.397$$

$$r_{LDEN} := \sqrt{r_{LBEN}^2 - r_{LBD}^2} \qquad r_{LDEN} = 34.6316$$

The new angles of the vector DE on the right and left side are:

$$\theta_{RDEN} := 2 \cdot \operatorname{atan} \left(\frac{r_{RBD} + z_{RBEN}}{r_{RDEN} + y_{RBEN}} \right) \qquad \theta_{RDEN} = 3.6512 \cdot \text{deg}$$

$$\theta_{LDEN} := 2 \cdot \operatorname{atan} \left(\frac{r_{LBD} + z_{LBEN}}{r_{LDEN} + y_{LBEN}} \right) \qquad \theta_{LDEN} = 0.3326 \cdot \text{deg}$$

The angles for the new BC on the right and left are:

$$\theta_{RBCN} := \theta_{RDEN} - \lambda_R \qquad \theta_{RBCN} = -4.0039 \cdot \text{deg}$$

$$\theta_{LBCN} := \theta_{LDEN} - \lambda_L \qquad \theta_{LBCN} = -7.3225 \cdot \text{deg}$$

The new BC components are:

$$\begin{pmatrix} y_{RBCN} \\ z_{RBCN} \end{pmatrix} := \begin{pmatrix} r_{RBC} \cdot \cos(\theta_{RBCN}) \\ r_{RBC} \cdot \sin(\theta_{RBCN}) \end{pmatrix} \qquad \begin{pmatrix} y_{RBCN} \\ z_{RBCN} \end{pmatrix} = \begin{pmatrix} 23.4018 \\ -1.638 \end{pmatrix}$$

$$\begin{pmatrix} y_{LBCN} \\ z_{LBCN} \end{pmatrix} := \begin{pmatrix} r_{LBC} \cdot \cos(\theta_{LBCN}) \\ r_{LBC} \cdot \sin(\theta_{LBCN}) \end{pmatrix} \qquad \begin{pmatrix} y_{LBCN} \\ z_{LBCN} \end{pmatrix} = \begin{pmatrix} 23.2678 \\ -2.99 \end{pmatrix}$$

The new AC components are:

$$\begin{pmatrix} y_{RACN} \\ z_{RACN} \end{pmatrix} := \begin{pmatrix} y_{RAB} \\ z_{RAB} \end{pmatrix} + \begin{pmatrix} y_{RBCN} \\ z_{RBCN} \end{pmatrix}$$

$$\begin{pmatrix} y_{RACN} \\ z_{RACN} \end{pmatrix} = \begin{pmatrix} 6.4018 \\ -12.638 \end{pmatrix}$$

$$\begin{pmatrix} y_{LACN} \\ z_{LACN} \end{pmatrix} := \begin{pmatrix} y_{LAB} \\ z_{LAB} \end{pmatrix} + \begin{pmatrix} y_{LBCN} \\ z_{LBCN} \end{pmatrix}$$

$$\begin{pmatrix} y_{LACN} \\ z_{LACN} \end{pmatrix} = \begin{pmatrix} 6.2678 \\ -13.99 \end{pmatrix}$$

The new AC lengths are:

$$r_{RACN} := \sqrt{y_{RACN}^2 + z_{RACN}^2}$$

$$r_{RACN} = 14.167$$

$$r_{LACN} := \sqrt{y_{LACN}^2 + z_{LACN}^2}$$

$$r_{LACN} = 15.3298$$

The new increments are:

$$\Delta_{RDEN} := r_{RDEN} - r_{RDE}$$

$$\Delta_{RDEN} = -0.103$$

$$\Delta_{RACN} := r_{RACN} - r_{RAC}$$

$$\Delta_{RACN} = -0.583$$

$$\Delta_{LDEN} := r_{LDEN} - r_{LDE}$$

$$\Delta_{LDEN} = 0.1316$$

$$\Delta_{LACN} := r_{LACN} - r_{LAC}$$

$$\Delta_{LACN} = 0.5798$$

$$\eta := 0.1$$

The new rotations required are:

$$\Psi_{RDE} := \frac{\Delta_{RDEN}}{\eta} \cdot 2 \cdot \pi$$

$$\Psi_{RDE} = -370.8683 \cdot \text{deg}$$

$$\Psi_{RAC} := \frac{\Delta_{RACN}}{\eta} \cdot 2 \cdot \pi$$

$$\Psi_{RAC} = -2.0989 \cdot 10^3 \cdot \text{deg}$$

$$\Psi_{LDE} := \frac{\Delta_{LDEN}}{\eta} \cdot 2 \cdot \pi$$

$$\Psi_{LDE} = 473.7588 \cdot \text{deg}$$

$$\Psi_{LAC} := \frac{\Delta_{LACN}}{\eta} \cdot 2 \cdot \pi$$

$$\Psi_{LAC} = 2.0874 \cdot 10^3 \cdot \text{deg}$$

$$\begin{pmatrix} x_{LREN} \\ y_{LREN} \\ z_{LREN} \end{pmatrix} := \begin{pmatrix} x_{RBN} \\ y_{REN} \\ z_{REN} \end{pmatrix} - \begin{pmatrix} x_{LBN} \\ y_{LEN} \\ z_{LEN} \end{pmatrix}$$

$$\begin{pmatrix} x_{LREN} \\ y_{LREN} \\ z_{LREN} \end{pmatrix} = \begin{pmatrix} 20 \\ 0 \\ 2 \end{pmatrix}$$

$$r_{LREN} := \sqrt{x_{LREN}^2 + y_{LREN}^2 + z_{LREN}^2}$$

$$r_{LREN} = 20.0998$$

The new unit vector joining the left and right "E"s is:

$$\begin{pmatrix} x_{U1N} \\ y_{U1N} \\ z_{U1N} \end{pmatrix} := \begin{pmatrix} x_{LREN} \\ y_{LREN} \\ z_{LREN} \end{pmatrix} \cdot \frac{1}{r_{LREN}}$$

$$\begin{pmatrix} x_{U1N} \\ y_{U1N} \\ z_{U1N} \end{pmatrix} = \begin{pmatrix} 0.995 \\ 0 \\ 0.0995 \end{pmatrix}$$

$$r_{LRS} := \sqrt{(x_{RS} - x_{LS})^2 + (y_{RS} - y_{LS})^2 + (z_{RS} - z_{LS})^2}$$

$$r_{LRS} = 20.0998$$

The unit vector joining the left and right lower supports is:

$$\begin{pmatrix} x_{US} \\ y_{US} \\ z_{US} \end{pmatrix} := \begin{pmatrix} x_{RS} - x_{LS} \\ y_{RS} - y_{LS} \\ z_{RS} - z_{LS} \end{pmatrix} \cdot \frac{1}{r_{LRS}}$$

$$\begin{pmatrix} x_{US} \\ y_{US} \\ z_{US} \end{pmatrix} = \begin{pmatrix} 0.995 \\ -3.0052 \cdot 10^{-5} \\ 0.0995 \end{pmatrix}$$

Because these two unit vectors are now the same, they are parallel.

Let the final coordinates of Right and Left "E" be

$$\begin{pmatrix} y_{REF} \\ z_{REF} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} y_{LEF} \\ z_{LEF} \end{pmatrix}$$

In the forgoing equations, the above are used in place of "new" coordinates for RE and LE

$$\begin{pmatrix} y_{REN} \\ z_{REN} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} y_{LEN} \\ z_{LEN} \end{pmatrix}$$

and calculations are repeated to obtain the final displacements for the wheels of the Right and Left BE and AC.

7. CONCLUSIONS AND RECOMMENDATIONS:

In this report the procedure to obtain the necessary displacements to align the RIEE mechanism for the final insertion has been accomplished with example calculations. The MathCad software has allowed us to integrate the Text, Figures, Derivations and Calculations, saving time and space for a clear understanding of the procedures and calculations.

It is recommended that an experiment be set up to obtain the images of prototype targets and process the images to get the input data required for this software. The results of the software can then be verified be verified against the actual locations of the prototype targets.

8. REFERENCES:

- A. Drawing set #82K03931 and associated drawings

1994018770
N94-23243

1993 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA

197-201
197199
P 47

FTIR QUANTIFICATION OF INDUSTRIAL HYDRAULIC FLUIDS
IN PERCHLOROETHYLENE

PREPARED BY:	Dr. Narinder K. Mehta
ACADEMIC RANK:	Investigator
UNIVERSITY AND DEPARTMENT:	University of Puerto Rico - Mayaguez Department of Chemical Engineering
NASA/KSC	
DIVISION:	Material Science
BRANCH:	Microchemical Analysis
NASA COLLEAGUE:	Gale Allen
DATE:	August 5, 1993
CONTRACT NUMBER:	University of Central Florida NASA-NGT-60002 Supplement: 11

DEDICATION

In loving and living memory of my late
dear mother Joginder Vati,
whom I never knew.

ACKNOWLEDGMENT

I would like to thank NASA/ASEE Summer Faculty Fellowship Program, to Dr. E. Ramon Hosler of the University of Central Florida, to Ms. Gale Allen and Mr. William (Irby) Moore, Mr. Tom Barron and Mr. Frank Nesbit of the Kennedy Space Center (KSC) for providing me with the opportunity to work at KSC for another summer. Thanks to Ms. Kari Stiles for her willingness to help in all administrative matters.

Special thanks are extended to Ms. Gale Allen, KSC project coordinator and Mr. Jim Johns for suggesting the project, and for their gentleness, guidance and encouragement in all phases of the project.

There are many other professionals who assisted me during the course of this project. My sincere appreciation goes to Mr. Lee Underhill, Ms. Martha Williams, Dr. Orlando Melendez, Mr. Charlie Bassett and Ms. Valarie Frankline for their assistance throughout the project.

Special friendly thanks are extended to other professionals of the microchemical and failure analysis laboratories who made me feel welcome, and for their help in rush times.

The support of Dr. Jorge Ortiz, Dr. Felix Santiago, Dr. A. Rodriguez and Mr. Roberto Carrero of my home institute, the University of Puerto Rico is gratefully acknowledged.

ABSTRACT

The purpose of this summer research project was to investigate whether perchloroethylene can be used as a solvent for the quantitative analysis of industrial hydraulic fluids by infrared spectroscopy employing Beer's law. Standard calibration curves using carbon-hydrogen stretching (generic) and ester absorption peaks were prepared for a series of standard dilutions at low ppm levels of concentration of seven hydraulic fluids in perchloroethylene. The absorbance spectras were recorded with 1.5-10 mm fixed and variable path length sample cells made of potassium bromide. The results indicate that using ester infrared spectral peak, it is possible to detect about 20 ppm of the hydraulic fluid in perchloroethylene.

SUMMARY

1,1,2 trichloro-1,2,2 trifluoro ethane (CFC-113) commercially known as Freon-113 is the primary test solvent used for validating the cleaned hardware at the Kennedy Space Center (KSC). It is also used as a general laboratory solvent for analytical work in a variety of sample matrices at KSC. Due to the ozone depletion problem, the current United States policy calls for the phase out of Freons by 1995. A 1993 (May) preliminary report by Environmental Protection Agency (EPA) on the efforts to replace Freon-113 for the determination of oils and greases suggests the evaluation of the feasibility of utilizing infrared-based methods, and perchloroethylene (PCE) as an alternate validation and extraction solvent.

NASA's chlorofluorocarbon (CFC) replacement group at KSC has opted to pursue the recommendations made in the EPA's report to investigate infrared method and PCE since PCE has high solvation compatibility with organics and generally considered non-toxic (OSHA's LD50). Infrared method of analysis with dispersive infrared spectrometers has been used extensively for qualitative identification, and for quantification at high parts per million (ppm) level or at percent level concentration of the analyte. However, with the advent of the Fourier transform method, state-of-the-art spectrometers have become available for quantification at low ppm level concentration of the analyte. In this project, experimental data was obtained to assess the applicability of the infrared method to the detection of industrial hydraulic fluids in PCE at low ppm levels of concentration.

Standard calibration curves were prepared for a series of standard dilutions of hydraulic fluids in PCE. The infrared absorbance spectras were recorded with an FTIR instrument operated from a computer-based data collection and spectral analysis system, and using fixed and variable path length potassium bromide sample cells.

The results indicate that it is possible to detect about 20 ppm of the hydraulic fluid in PCE using a generic or any other absorbance spectral peak of the analyte. By optimizing certain variables which may affect the interpretability of the results obtained, the method may equally be applicable to the detection of hydraulic fluids below 20 ppm.

TABLE OF CONTENTS

<u>Section</u>	<u>Title</u>
I	INTRODUCTION
1.1	Objective
1.2	Background
II	MATERIALS AND METHODS
2.1	Instrumentation
2.1.1	FTIR Spectrometer
2.1.2	Sample Cells
2.2	Industrial Hydraulic Fluids
2.3	Perchloroethylene (Solvent)
2.4	Test Procedure
2.4.1	Preparation of Standards
2.4.2	Cleaning Process
2.4.2.1	Volumetric Flask Cleaning
2.4.2.2	Syringe Cleaning
2.4.2.3	Cell Cleaning
2.4.3	Sample Cell Filling
III	RESULTS AND DISCUSSIONS
3.1	Infrared Absorbance Spectrum
3.2	Calibration Curves
3.2.1	Mobile Jet II Fluid
3.2.2	Rapid Tap Cutting Fluid
3.2.3	Red Mil-H-5606 Fluid
3.2.4	Green # 1530 (Houghto) Fluid
3.2.5	Brown (ASTM # 3) # 1537 Fluid
3.2.6	Yellow Mil-H-83282 Fluid
3.2.7	PL-S Lube Oil # 1542 Fluid
3.2.8	Mixture of Hydraulic Fluids
IV	CONCLUSIONS
V	RECOMMENDATIONS

LIST OF ILLUSTRATION

<u>Figure</u>	<u>Title</u>
2-1	Block diagram of the FTIR system
3-1	Typical absorbance spectrum of a hydraulic fluid
FTIR Calibration Curves:	
Mobile Jet II	
3-2	Ester peak, 1.5 KBr cell
3-3	Ester peak, 1.5 KBr cell (two different days)
3-4	Ester and generic peaks, 1.5 KBr cell
3-5	Generic peak, 2.5 mm KRS-5 cell
3-6	Ester peak, 1.5 and 5 mm KBr cells
Rapid Tap Cutting Fluid	
3-7	Spectral peak, 1.5 mm KBr cell
3-8	Ester and generic peaks, 1.5 mm KBr cell
3-9	Generic peak, 2.5 mm KRS-5 cell
3-10	Ester peak, 1.5 and 5 mm KBr cells
Red Mil-H-5606 Fluid	
3-11	Generic peak, 1.5 mm KBr cell
3-12	Generic peak, 1.5 mm KBr cell (two different days)
3-13	Generic and ester peaks, 1.5 mm KBr cell
3-14	Generic peak, 2.5 mm KRS-5 cell
3-15	Ester peak, 5 mm Cell (two consecutive days)
Green # 1530 (Houghto) Fluid	
3-16	Spectral peak, 1.5 mm KBr cell
3-17	Ester peak, 1.5 and 10 mm KBr cells
Brown (ASTM # 3) # 1537 Fluid	
3-18	Generic peak, 1.5 and 10 mm KBr cells
Yellow Mil-H-83282 Fluid	
3-19	Ester peak, 1.5 and 10 mm KBr cells
PL-S Lube Oil (Gen Purpose) # 1542 Fluid	
3-20	Spectral peak, 1.5 and 10 mm KBr cells
3-21	Ester peak, 1.5 and 10 mm KBr cells
Mixture of Five Hydraulic Fluids	
3-22	Ester peak, 5 mm KBr cell

FTIR QUANTIFICATION OF INDUSTRIAL HYDRAULIC FLUIDS IN PERCHLOROETHYLENE

I-INTRODUCTION

1.1 Objective:

A feasibility study to investigate whether perchloroethylene can be used as a solvent for the quantitative analysis of industrial hydraulic fluids by Infrared Spectroscopy (FTIR).

1.2 Background:

1,1,2 trichloro-1,2,2 trifluoro ethane (CFC-113) known as Freon-113 is the primary test solvent used for validating the cleaned aerospace hardware at the Kennedy Space Center (KSC). For a safe space shuttle launch, it is of utmost importance that the hardware be verified clean of organic materials before use in an oxygen rich environment as that of the Orbiter.

Freon-113 is also used as a general laboratory solvent for analytical work in a variety of sample matrices at KSC. Material Science Laboratory's (MSL) standard procedure calls for the extraction of organic materials with Freon-113, and subsequent determination of the non-volatile residue (NVR) by gravimetric method.

Freons including CFC-113 are known to remain in the atmosphere long enough to migrate to altitudes due to their high volatility, and are known to be a threat to the earth's protective ozone layer. The Montreal protocol and the 1990 Clean Air Act calls for the accelerated reduction of the substances directly responsible for ozone depletion of the Antarctic ozone hole. The current United States Government policy calls for ending the production of these compounds, and phase out by 1995.

United States Environmental Protection Agency's (EPA) Clean Water Act (CWA), and Resource Conservation and Recovery Act (RCRA) requires an effluent discharge permit for all the industries connected to the wastewater treatment facilities. One of the parameters required for the permit is the analytical determination of oils and greases in the

effluents. EPA's standard methodology calls for the utilization of Freon-113 as the extraction solvent for the determination of oils and greases in the effluents.

A 1993 (May) preliminary report by EPA on the efforts to replace Freon-113 for the determination of oils and greases suggests the evaluation of infrared-based methods, and perchloroethylene (PCE, also known as tetrachloroethylene) as an alternate extraction solvent. This report also acknowledges that a lot of efforts are going on at the national and international level to find a substitute for Freon-113.

Chlorofluorocarbon (CFC) replacement group at KSC in collaboration with NASA's and other independent research centers is also working on to find a substitute for CFC-113 for the validation process of cleaned parts. Among the options, the MSL group decided to pursue the recommendations of the EPA's preliminary report of investigating the infrared method using PCE as a replacement solvent since PCE has high solvation compatibility with organics, generally considered non-toxic (OSHA's LD50), and has a high B.P. (121 oC) compared to 48 oC for Freon-113.

Infrared methods of analysis with dispersive infrared spectrometers have been used extensively as a general analytical technique for identifying materials or "fingerprinting", and for quantification of analytes at high parts per million (ppm) or at percent concentration levels.

Infrared detectors in dispersive spectrometers are not frequency sensitive. With the advent of the Fourier transform method, more sensitive, accurate, reliable and state-of-the-art Fourier transform infrared spectrometers (FTIR) have become available for the quantification work at low ppm levels since in FTIR, infrared frequencies are converted into audio frequencies, and the available detectors and electronics have the ability to track both frequency and intensity at the same time. This research project will focus basically to investigate the feasibility of extending the FTIR method for detecting low ppm concentration levels of industrial hydraulic fluids in PCE.

II-MATERIALS AND METHODS

2.1 Instrumentation:

2.1.1 FTIR Spectrometer:

The spectrometer used for this study was a Galaxy series 300 FTIR spectrometer manufactured by Mattson Instruments. The instrument employs a deuterium triglycine sulfate (DTGS) detector, a Michelson interferometer, and a standard high-intensity infrared source to cover the mid infrared range from 500 - 5,000 wavenumber (cm^{-1}). The Michelson interferometer, which consists of a cesium iodide beam splitter, a fixed mirror and a moving mirror, preserves both the intensity and frequency information at the same time. An internal He-Ne monochromatic laser monitors the position of the moving mirror within the scan. The spectrometer was operated from a computer-based data-collection and spectral analysis analytical package called FIRST (Fourier Infrared Software Tools). The FIRST analytical package automatically collects the indicated number of interferogram scans, coadds the scans, performs the fast Fourier transform, processes the data, and plots the resulting sample spectrum. The block diagram of the system is presented in Figure 2-1.

The results presented in this report are obtained using the setting of 120 interferogram scans* at 4cm^{-1} resolution with detector gain setting of 4. A background spectrum of the solvent (PCE) was always obtained before the start of collecting the sample spectra since with an appropriate background spectrum in memory, the spectrometer coupled with FIRST automatically calculates the absorbance spectrum as a ratio of the sample spectrum to the background spectrum.

2.1.2 Sample Cells

The most common method of sample preparation for infrared examination of liquid samples involves placing an undiluted sample (0.01-0.1 mm thickness) between a pair of transparent crystal windows. The recorded spectrum will be

*In FTIR, the scan is the mechanical displacement of the moving mirror assembly and their total reflects the sensitivity of the instrument (S/N ratio).

the most characteristic of the whole composition of the sample since no solvent or matrix interferences, and inter- and intra-molecular interactions will be present to contribute to the difficulties of band shifts and their identifications. However, in dilute solutions (less than 100 ppm range) one experiences all sorts of difficulties since spectras are recorded on thicker samples for obtaining useful analytical information. The greater the number of solvent molecules along the radiation path, the greater will be the interferences and absorption by solvent molecules. Therefore it is of utmost importance to optimize the sample thickness for obtaining useful spectral information.

Sealed- and variable path length liquid infrared cells (1.5-10 mm) having potassium bromide (KBr) and thallium bromide iodine (KRS-5) transparent windows were used for recording the absorbance spectras of the standard solutions of industrial hydraulic fluids in PCE at low ppm range of the analyte.

2.2 Industrial Hydraulic Fluids:

About 20 types of industrial hydraulic fluids and greases are commonly used at KSC. Eight of the industrial hydraulic fluids listed below were obtained from various sources at KSC, and were utilized for the preparation of the calibration curves. Most of them are ester-based hydrocarbon type fluids. During the course of this study, it was observed that one of the silicon-based fluid was not compatible with PCE. For this reason, this study is focussed on the preparation of standard curves for quantification for the seven hydraulic fluids.

<u>Sample I.D.</u>	<u>Name</u>
b	Red-Mil-H-5606
c	Rapid Tap Cutting Fluid
d	Dow Corning FS 1265 Fluid (this sample was found not to be compatible (Solubility) with PCE.
e	Mobile Jet II
f	Yellow Mil-H-83282
g	Green # 1530 (Houghto) Fluid
h	Brown (ASTM #3) 3 1537 Fluid
i	PL-S Lube Oil (General Purpose) # 1542

2.3 Perchloroethylene (solvent):

Perchloroethylene used for the preparation of all dilutions in this study was a Fisher brand reagent grade solvent having 0.5 percent of ethyl alcohol as a stabilizer. In the early part of the study, about 3/4 gallon of a bottle of the solvent was provided by a contractor group at KSC. Once the borrowed stock was consumed in the preparation of the dilutions, a newly purchased stock was used until the end of the study.

2.4 Test procedure:

2.4.1 Preparation of Standards:

Several series of standard dilutions of hydraulic fluids in PCE were prepared for recording the absorbance spectras. Dilutions lower than 100 ppm were prepared from a stock solution of each fluid in PCE. Since the objective of the project was to investigate the detection of the analyte in PCE, extreme care was taken not to cross-contaminate the standard dilutions. Auto pipettes of different sizes and brands available in the laboratory were used for dispensing small volumes of the solutions into the volumetric flasks for the preparation of standard dilutions. Tips for the pipettes were only used once for the transfer of solutions, and were disposed off properly. Dilutions of less than 100 ppm were prepared before the start of recording the absorbance spectrum, and were never stored for over 24 hours.

2.4.2 Cleaning Process:

The detection of an analyte at low ppm level of concentration requires that the whole solution handling process be standardized to avoid cross-contamination. It is not cost effective to use glassware and syringes in quantities for the preparation and transfer of large number of solutions. A simple and practical procedure was used for cleaning and filling of the sample cell with the prepared standard solutions.

2.4.2.1 Volumetric Flasks Cleaning:

25 ml volumetric flasks were used for the preparation of most of the standard dilutions. After each use, the flasks were washed with warm 0.5 percent Liquinox solution according

to the normal procedure used in the laboratory. After rinsing with deionized water, the flasks were dried in an oven maintained at 103 degrees Celcius. All the volumetric flasks were rinsed twice with PCE before the preparation of the standard dilutions.

2.4.2.2 Syringe Cleaning:

PCE for cleaning was transferred from a gallon bottle into two cleaned 50 ml beakers. The syringe used for filling the sample cell was cleaned twice with a 2 ml portion of the pure solvent from the first beaker. Both the times, the syringe was inverted and plunger moved to the outer-most end of the barrel for rinsing of the whole barrel with the solvent. It is not necessary to remove the barrel from the plunger. The plunger was moved up and down 3-5 times with paper tissue wipes at the tip of the syringe to soak the maximum amount of solvent on to the tissue wipes.

2.4.2.3 Cell Cleaning:

A cleaned syringe was filled with 1-2 ml of the solvent from the other beaker, and the solvent transferred gently into the sample cell. Depending on the size of the cell, the amount of the solvent to be transferred may vary. The teflon plugs were replaced and the cell was given a gentle swirling motion horizontally. The plugs were removed and the solvent was allowed to drain from the lower orifice (right orifice for circular adjustable cell) into a waste-beaker assigned for proper disposal. Paper tissue wipes were placed on top of the lower orifice of the cell to soak the maximum amount of the solvent onto the wipes without giving a drastic motion to the cell. The procedure was repeated with another portion of the solvent. It is probable that a trace amount of solvent may remain in the cell. It is assumed that this trace quantity of the solvent remained in the cell will be constant for all the standards in a series, and will have insignificant effect on the slope of the standard calibration curve.

2.4.3 Sample Cell Filling:

An appropriate amount of the standard solution was transferred into an aluminum dish from the 25 ml volumetric flask. 2-3 ml of this solution was drawn into a cleaned 5 ml Lauer syringe. This volume may vary depending on the

thickness of the sample cell. Using a very light hand pressure, the solution was injected into the lower orifice of the metal plate of the cell. It is important that the cell should never be filled from the upper orifice at the top of the front plate of the cell. After the cell was filled, teflon plugs were replaced with a twisting motion to ensure a good seal. All the times paper tissue wipes were used to avoid the contact of the transferring solution with the cell windows. The filled cell was physically observed for any entrapped bubbles of air before placing it in the cell holder of the instrument for recording the absorbance spectras.

III-RESULTS AND DISCUSSIONS

3.1 Infrared Absorbance Spectrum:

When radiation at a particular frequency strikes a molecule which has a vibrational frequency identical to the incident radiation, the molecule will absorb the radiant energy, and thus increasing the total energy of the molecule. If a pair or group of atoms is to absorb infrared radiation, it must undergo a change in dipole moment during the vibration. The changing dipole couples the vibration of the molecule with that of the incident radiation, and causes infrared absorption.

Identification of a certain principal spectral peak(s) of an analyte is necessary to be identified before calibration curves can be plotted for quantification. Special efforts were focussed to identify a particular generic peak in all the fluids which can be used for the preparation of a single calibration curve for detecting the total concentration of hydraulic fluids in a mixture. One percent dilutions of the hydraulic fluids were prepared in PCE, and their absorbance spectras were recorded for the identification of the most intense absorption peak(s). For quantitative work, the absorbance spectras were recorded at a sensitivity of 120 interferogram scans at 4 cm^{-1} resolution. Triplicate spectras were recorded for each of the standard dilution, and the average of the three absorbance values obtained for a particular spectral peak was used for the preparation of the standard curves.

Ester-based hydrocarbon type fluids used in this project have spectral bands around 2700 to 3000 cm^{-1} , and near 1700 and 1100 cm^{-1} . Because hydrogen is a light element, spectral bands due to stretching of C-H bonds occur at a fairly high frequency. The large broad band from 2700-3000 cm^{-1} (generic peak) is due to the stretching vibrations of the carbon-hydrogen bonds in the hydrocarbon chain backbone structure, and is very typical of organic compounds. Since it requires less energy to bend a bond than to stretch it, bands due to bond bending occur at low frequency. The peak near 1700 cm^{-1} is typical of an ester group, and also C=C stretch peaks appear in this area. A typical absorbance spectra of a hydraulic fluid is presented in Figure 3-1.

3.2 Calibration Curves:

The absorbance data is useful for quantitative analysis of analytes in dilute solutions where Beer's law is applicable, and is that a linear relationship exists between the measured net absorbance and the concentration of an analyte in a solution. The net absorbance of the spectral peaks obtained for seven hydraulic fluids at various levels of concentration and sample thicknesses are presented in Figures 3-2 to 3-22 in the form of calibration curves. The net absorbance data was obtained using the analytical software package called FIRST. Regression analysis was done on each set of standards using analyte concentration as an independent variable and absorbance units as a dependent variable. The validation of the model is presented in the form of R Squared. It ranges up to one with one being optimal.

3.2.1 Mobile Jet II:

Figure 3-2 is a calibration curve for Mobile Jet II Fluid in PCE using 1743 cm^{-1} ester peak of the spectras recorded with a 1.5 mm KBr variable cell. The graph demonstrates a linear relationship between absorbance and concentration of the analyte in the range of 50-500 ppm with R^2 of 0.99. The absorbance data of this series of standards (10-300 ppm only) along with the data obtained on standards prepared on a different date is presented in Figure 3-3. R^2 of the two sets of data was found to be optimal (0.98). Although variation in the absorbance units of the two series of standards were observed, the two curves were found to have similar slopes.

Figure 3-4 demonstrates the erratic scattering in the absorbance data obtained on the generic C-H stretching peak as compared to the ester peak of the recorded spectras with a 1.5 mm KBr cell. During the course of the work, it was observed that too much background noise was resulting in the area of the generic peaks of the spectras, and it was impossible for the detector to distinguish between small variations in the absorbance units in solutions with analyte concentration at low ppm levels. The presence of several peaks/bands around 3000 wavenumber in the absorbance spectrum obtained on a pure solvent reflected the presence of some contaminant in the solvent. As a check, a similar observation was made on a newly opened solvent bottle at a later stage. This observation suggests that the presence of the unwanted

organic compound(s) may be a part of this particular batch of solvent after all, and causing the background noise in the generic peak(s) area of the infrared spectra.

The absorbance data for the ester peak in Figure 3-4 reflect that it is possible to detect about 20 ppm of Mobile Jet II fluid in PCE R^2 equal to 0.98. It was reasoned that the detection of the analyte may be improved by increasing the absorbance units with a larger sample size. The only cell of higher path length available in the laboratory was a KRS-5 cell having a variable path length of 2.5 mm. A calibration curve (Figure 3-5) in the range of 1-10 ppm of the analyte for the generic peak was prepared using the spectras recorded with the KRS-5 cell. The graph demonstrates highly scattered data obtained using the longer path length KRS-5 cell. 5 mm and 10 mm KBr fixed cells were purchased from Perkin Elmer Corporation, and absorbance data was obtained on the ester peak (1743 cm^{-1}) using a 5 mm fixed path length sample cell. Figure 3-6 is a calibration curve of the net absorbance data obtained on an ester peak of the spectras recorded with 1.5 and 5 mm KBr cells. The slopes of the two lines were found to be reasonably similar with R^2 values of about 0.98. Efforts were also made to use the absorbance data of the generic peak (2930 cm^{-1}) for the preparation of the calibration curve. This resulted in failure due to the presence of several unresolved peaks in the generic peak area. The calibration curves (Figures 3-4 and 3-6) demonstrate that ester peak is stable and can be used for the quantification of Mobile Jet II in PCE. These graphs also demonstrate that about 20 ppm of the analyte in PCE can be detected under the experimental conditions.

3.2.2 Rapid Tap Cutting Fluid:

Figure 3-7 is a calibration curve for the Rapid Tap Cutting fluid in PCE using an infrared spectral peak of around 1086 cm^{-1} . This particular spectral peak was found to be more intense compared to the ester peak (1720 cm^{-1}) at high ppm levels of concentration of the analyte. Figure 3-7 demonstrates a linear relationship between the net absorbance units and the analyte concentration ($R^2 = 0.99$) in the range of 50-500 ppm with the spectras recorded using a 1.5 mm KBr cell.

The ester and the generic peak was investigated for the preparation of the calibration curves below 100 ppm of the analyte in PCE since the area around 1,100 wavenumber was found to be full of scattered noise at low concentration

levels. A series of standard dilutions of the analyte in PCE was made in the range of 20-100 ppm to investigate the possibility of detecting the analyte concentration below 50 ppm. The absorbance spectras were recorded with a 1.5 mm KBr cell. The net absorbance was calculated for the generic as well as for the ester peak, and was plotted as a function of the analyte concentration. Figure 3-8 demonstrates that we can detect about 20 ppm of the Rapid Tap Cutting fluid in PCE using the ester peak. R2 value of the absorbance data for the ester peak was found to be about 0.54. The lower R2 value may be attributed to a bad standard of 80 ppm of the analyte in PCE. The generic peak (2930 cm^{-1}) in the concentration range of 20-100 ppm with a 1.5 mm KBr cell (Figure 9) as well as in the concentration range of 1-10 ppm with a 2.5 mm KRS-5 cell (Figure 3-9) gave scattered absorbance data similar to the Mobile Jet II sample. This reflects that the generic peak is not recommendable to be used with this solvent for quantification work due to the presence of other organics in the solvent which produces background noise bands in the generic peak area.

A series of standards (20-100 ppm) used for the preparation of the calibration curve with a 1.5 mm KBr cell was also used for the 5 mm KBr cell study. Figure 3-10 is a calibration curve of the net absorbance data of the ester peak of the spectras recorded with 1.5 mm and 5 mm KBr cells. The absorbance data of the two graphs follows a similar pattern. This reflects that 80 ppm standard is indeed a bad standard causing R2 value to go down drastically. The data also suggest that the optimum sample path length is somewhere between 1.5-5 mm.

3.2.3 Red Mil-H-5606 Fluid:

A series of standard solutions of Red Mil-H-5606 in the range of 10-1,000 ppm in PCE was prepared to verify Beer's law. Figure 3-11 is a calibration curve of the analyte using the generic peak (2928 cm^{-1}) of the spectras recorded with a 1.5 mm KBr cell. The graph is found to be linear in the range tested except for some scattering of the data observed below 50 ppm level of the analyte. R2 was found to be equal to 0.99, and the graph suggests that it is possible to detect below 50 ppm of the analyte in PCE.

A new series of standards in the range of 20-100 ppm of the analyte was prepared in PCE for improving the detection of the analyte in PCE. The infrared absorbance spectras were recorded using 1.5 mm KBr cell. The absorbance data obtained

on the generic peak (2928 cm^{-1}) for this series is plotted in Figure 3-12 together with the generic peak absorbance data of Figure 3-11 between 20-100 ppm range (4 points) of the analyte. Scattering of the data for both the series was observed and R^2 was found to be 0.72 and 0.90 respectively. The graph suggests that it is possible to detect as low as 20 ppm of the analyte without difficulty.

Figures 3-13 is a calibration graph for the generic (2928 cm^{-1}) and the ester (1738 cm^{-1}) peak of the spectras recorded with a 1.5 mm KBr cell. R^2 for the ester peak was found to be 0.35 as compared to 0.72 for the generic peak. It reflects that there might be some other variables affecting the ester peak of Red Mil-H-5606 fluid in PCE at low ppm levels of concentration. In the range of 1-10 ppm concentration levels, scattering of the absorbance data was found to be similar to the Mobile Jet II and Rapid Tap Cutting fluid with the KRS-5 cell (Figure 3-14). Figure 3-15 is a calibration curve of the net absorbance data of the spectras recorded with a 5 mm KBr cell on two consecutive days on a series of standards with concentration of the analyte in the range of 20-100 ppm. The graph reflects that the absorbance data obtained on the ester peak with a larger sample size did not improve the detection limit of the Red Mil-H-5606 fluid in PCE. The absorbance data also reflect the instrumental drift on a low absorbance level (0.02 range) on a day to day basis.

3.2.4 Green # 1530 (Houghto) Fluid:

Figure 3-16 is a calibration graph for the Green #1530 Fluid in PCE in the concentration range of 20-1,000 ppm using a 1.5 mm KBr cell. The absorbance data of a more intense absorbance peak at 1157 wavenumber was used for the preparation of the calibration curve. The graph demonstrates a linear relationship for Beer's law with R^2 equal to 0.97. It also demonstrates that it is possible to detect about 20 ppm of the analyte in PCE.

10 mm KBr sealed cell was also used for recording the absorbance spectras for improving the detection of the analyte in PCE. It was tried to use the intense peak at 1157 wavenumber for measuring the absorbance units of the recorded spectras. However, the efforts were met with little success due to high background noise in the region of 1157 wavenumber peak. Figure 3-17 is a calibration curve of the net absorbance data obtained for the ester peak (1738 cm^{-1}) with 1.5 mm and 10 mm KBr cells. The absorbance data obtained for

the 10 mm KBr cell reflect that it is possible to detect about 20 ppm of the analyte in PCE, however scattering of the data was observed with the 10 mm KBr cell. With a 1.5 mm KBr cell, it was not possible to obtain a defined slope of the calibration graph using the ester absorbance spectral peak.

3.2.5. Brown (ASTM #3) # 1537:

Figure 3-18 is a calibration graph for the Brown # 1537 fluid using the generic peak (2854 cm^{-1}) net absorbance data of the recorded spectras obtained with a 1.5 mm KBr cell, and also the absorbance data of the ester peak (1739 cm^{-1}) of the spectras recorded with a 10 mm KBr cell. As expected, the absorbance units increased with the sample size. It is interesting to observe that the slope is identical for the graphs using two different absorption peaks. R2 for the absorbance data for all points obtained on the generic peak was found to be 0.97. The graph also reflects that about 20 ppm of the analyte can be detected in PCE.

3.2.6 Yellow Mil-H-83282:

The absorbance data obtained on the spectras recorded for a series of dilutions of Yellow Mil-H-83282 fluid in PCE with 1.5 mm and 10 mm KBr cells is presented in Figure 3-19. The slopes of the two graphs using the net absorbance data of the ester peak (1737 cm^{-1}) are similar with R2 values equal to 0.94 to 0.99 for 1.5 mm and 10 mm KBr cells respectively. The calibration curve demonstrates that it is possible to detect about 20 ppm of the analyte in PCE.

3.2.7 PL-S Lube Oil # 1542:

The absorbance data obtained on the standard dilutions for PL-S Lube Oil # 1542 in PCE using different sample sizes for recording the spectras are presented in Figures 3-20 and 3-21 in the form of the standard curves. The scattering of the absorbance data is observed for the generic and also for the ester peak using a 10 mm KBr cell. The slopes of the graphs using the generic and ester peaks with a 1.5 mm cell is found to be similar (Figures 3-20 and 3-21). The absorbance data obtained on the ester peak (Figure 3-21) reflect that it is possible to detect about 20 ppm of the analyte in PCE.

3.2.8 Mixture of Hydraulic Fluids:

All the calibration curves prepared above for the detection of the hydraulic fluids were of the solutions of the pure fluids in PCE. Most of the absorbance data obtained on individual fluids has demonstrated that it is appropriate to use the ester peak (1739 cm^{-1}) for the quantitative analysis of hydraulic fluids.

A series of dilutions of a mixture of five hydraulic fluids, Red Mil-H-5606, Rapid Tap Cutting fluid, Mobile Jet II, Yellow Mil-H-83282 and Brown ASTM # 3, was prepared in PCE to study the behavior of the ester absorbance peak in a mixture. The absorbance spectras were recorded with 1.5 mm and 5 mm KBr cells. The net absorbance data obtained with a 1.5 mm cell for the generic and for the ester peak were found to have an erratic behavior. Efforts were made to use the generic peak (2854 cm^{-1}) for the preparation of the calibration curve with a 5 mm KBr cell but the data demonstrated a highly erratic and similar behavior of 1.5 mm KBr cell due to high background interferences by the solvent in the region of generic peak(s) in the infrared absorbance spectrum.

The net absorbance data obtained on the ester spectral peak with a 5 mm KBr cell is presented in Figure 3-22. The figure demonstrates that it is possible to detect 20 ppm or less of the total hydraulic fluid using the ester peak and a 5 mm KBr cell. However, R2 value of the absorbance data was found to be equal to 0.73.

IV-CONCLUSIONS

1. The results obtained on the quantification of hydraulic fluids in perchloroethylene has demonstrated that it is possible to detect high ppm levels of the analyte easily with the infrared method using a generic or any other absorbance spectral peak of the analyte. R squared is found to be higher than 0.97 for the range (50-500 ppm) used for the preparation of the calibration curves.
2. The calibration curves prepared for quantification of the hydraulic fluids at low ppm levels of concentration (20-100 ppm) in PCE have demonstrated that it is possible to detect about 20 ppm of the analyte in PCE under the experimental conditions of the test runs.
3. The data has demonstrated that the generic spectral peak for most of the samples can not be used under the present experimental conditions for the detection of low ppm of the analyte due to background noise generated in the generic peak area by the solvent tested.
4. The optimal sample path length is found to be between 1.5-5 mm.
5. The results have also demonstrated that it is possible to extend the applicability of the infrared method to low ppm (2-20 ppm) levels of detection of the hydraulic fluids in PCE. However, one must control certain variables (sample size, contamination and background noise) which may affect the interpretability of the results obtained.

V-RECOMMENDATIONS

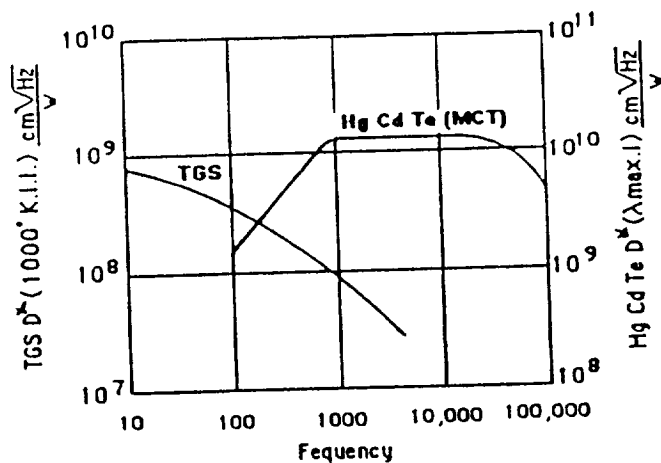
1. Infrared method, generally is not a quantitative method of analysis for the determination of analytes at low ppm levels of concentration. In order to extend its capability to low ppm level work, one must optimize certain variables which may directly affect the interpretability of the results.

- One of them is the sample size. It is found that the optimum length of the sample lies somewhere between 1.5 mm and 5 mm. It is suggested to carry out the experimental work to collect the data with a new variable path length KBr or BaF2 (preferred since it does not fog) cell.

- Another possible source of error in the detection of the analyte by the infrared method (or any other instrumental method) at low ppm levels may result from the use of contaminated supporting materials like auto pipettes. One of the auto-pipette (5ml range) used in the project had multi-users. Under the normal laboratory procedures, it is not possible to avoid contamination of the pipettes. It is suggested that pipettes and other supporting materials prone to contamination due to daily routine laboratory work should not be used for research projects.

- The infrared detector used in this research project was a deuterated triglycine sulfate with cesium iodide windows. It may be worth a try to use a more sensitive liquid nitrogen cooled mercury-cadmium-telluride detector which has a constant S/N ratio between 1,000-10,000 wavenumber for the quantification work at low ppm level.

D*
- is a function of audio frequency being detected, and also a measure of S/N ratio associated with the detector.



D* for TGS and HgCdTe Detectors

2. A careful evaluation of the recorded absorbance spectras revealed that sometimes the absorbance peak splits into two peaks or more depending on the characteristic bands of the compound. This evaluation on the splitting of the absorbance peaks reflect that their intensity (height) and position (wavenumber) varies with the concentration of the analyte especially at low ppm levels. It is suggested that peak area instead of peak height should be tried for the preparation of the calibration curves for the quantification of fluids at low ppm concentration levels of the analyte.
3. About 20 hydraulic fluids and greases are used at KSC. Some of them are silicon-based compounds which have limited solubility in PCE. It will not be possible to determine their concentration using infrared method. It is suggested to develop a model by running experiments on mixtures having fluids and greases by infrared method and also by the standard NVR method. The deviation in the results can be used to optimize a K factor in the following equation:

$$\text{Total Concentration} = K \times \text{Conc. (Infrared method)}$$

where K is a predetermined ratio of the NVR method concentration to the Infrared method concentration, and is always greater than one. Since there are only a limited known types of fluids and greases, it won't be very difficult to optimize the K factor.

4. For the removal of the non-volatile residue (NVR) material for the clean validation process, it is recommended to explore the possibility of using water at or near supercritical conditions (SCW), since at these conditions, water has properties as of a fluid and a gas thus increasing its solvation power. Experimental runs should be carried out to optimize the temperature and pressure conditions for the maximum removal of the NVR materials.

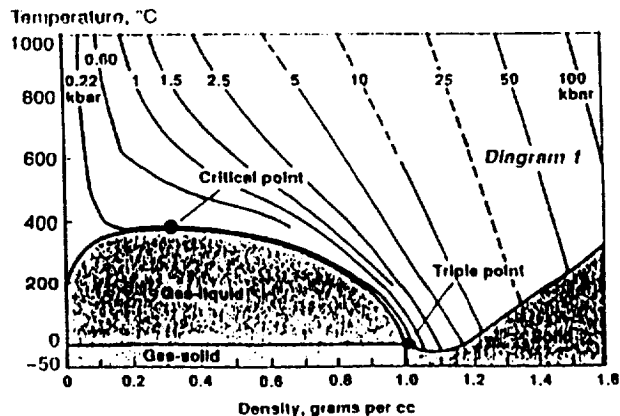
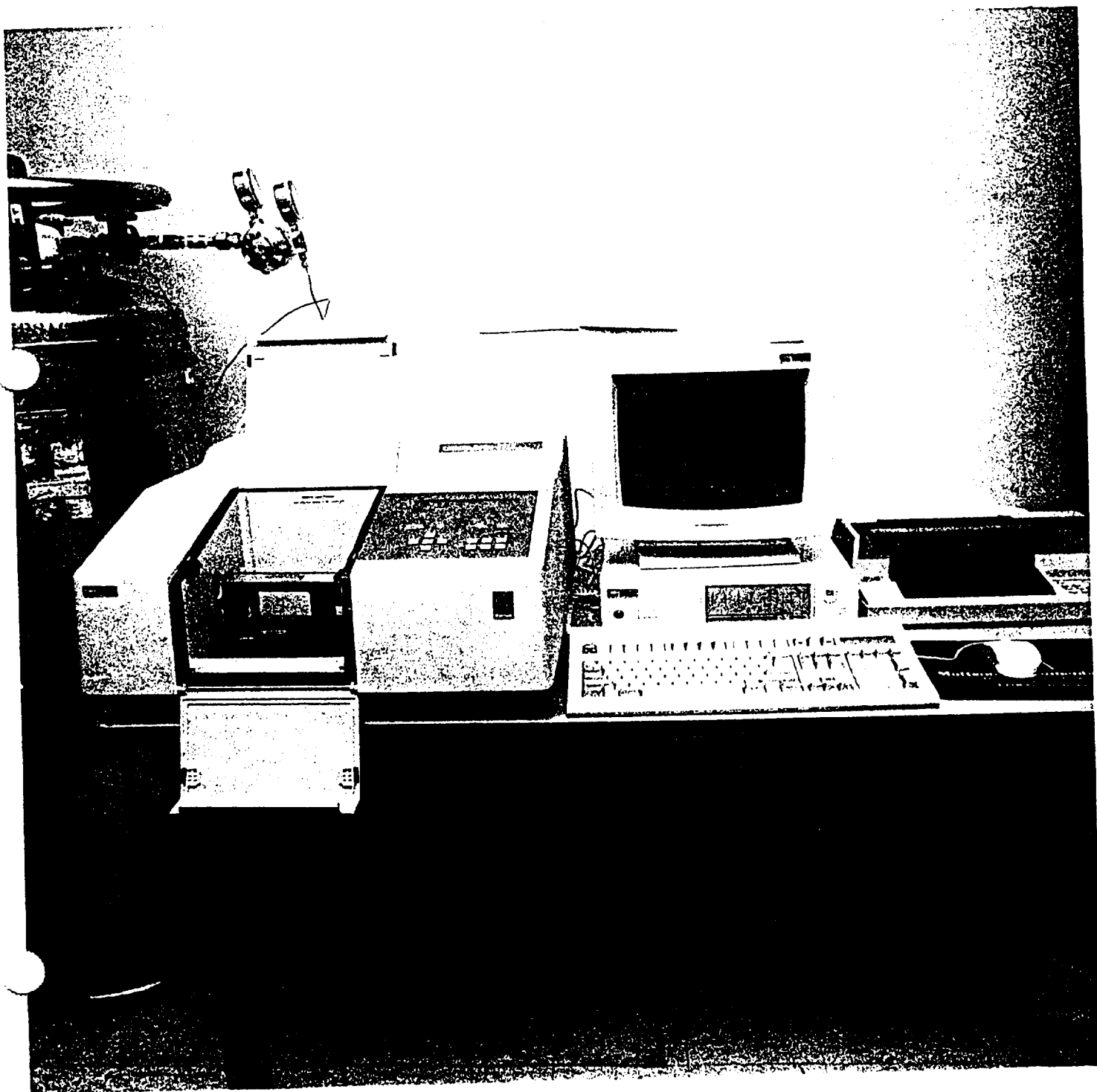
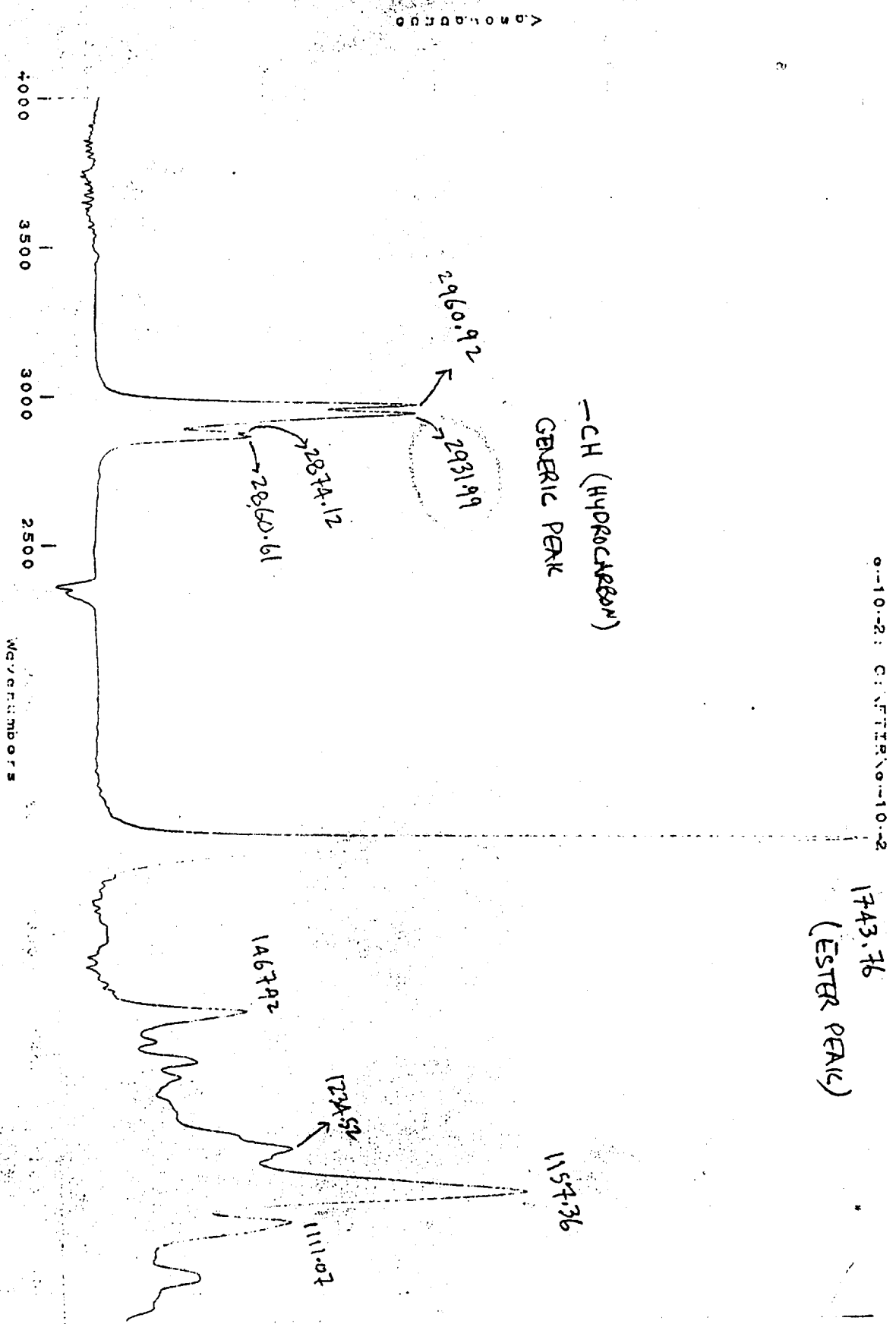


FIGURE 2-1



10% MOBILE JET II IN PCE

FIGURE 3-1



ORIGINAL PAGE IS OF POOR QUALITY

FIGURE 3-2

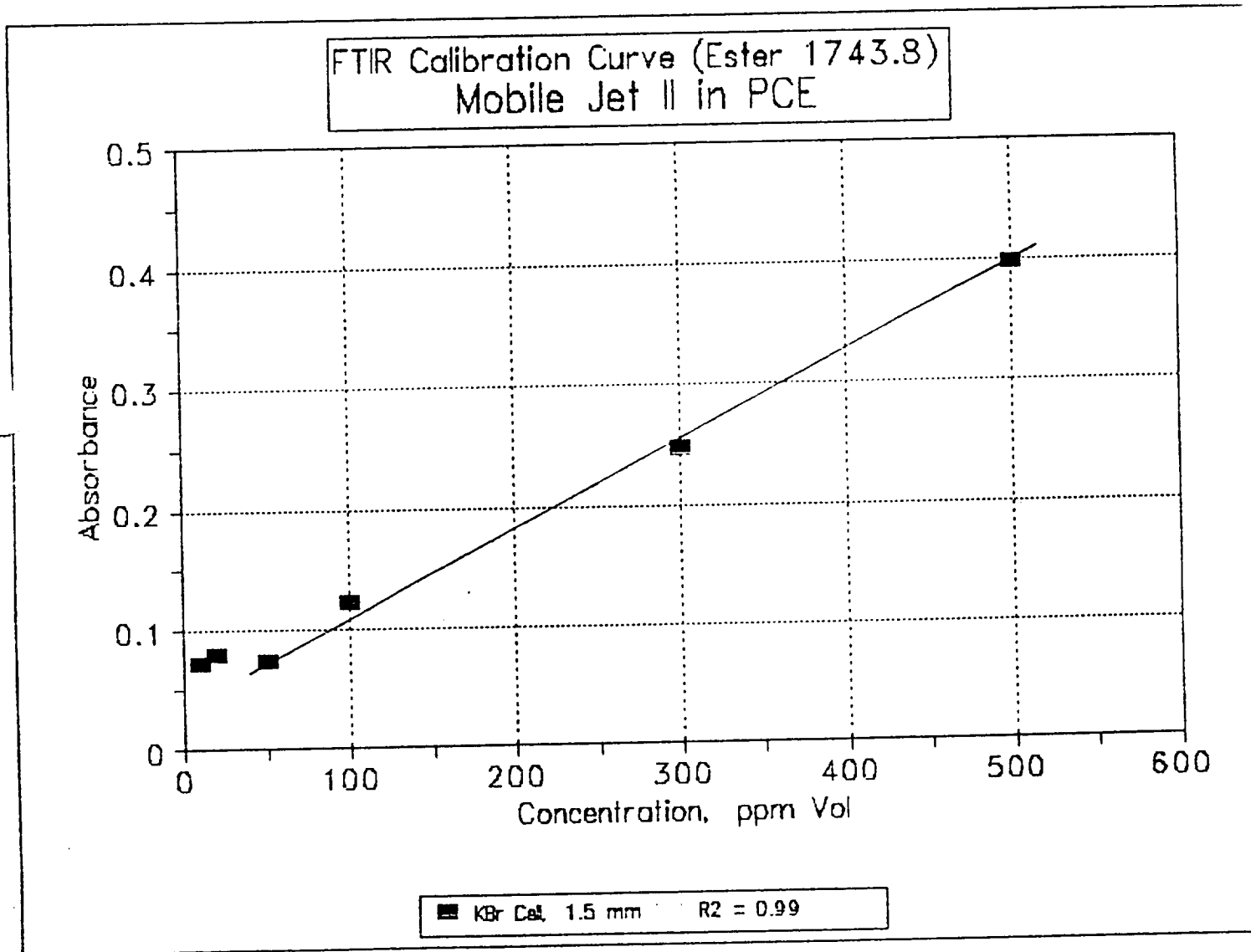


FIGURE 3-3

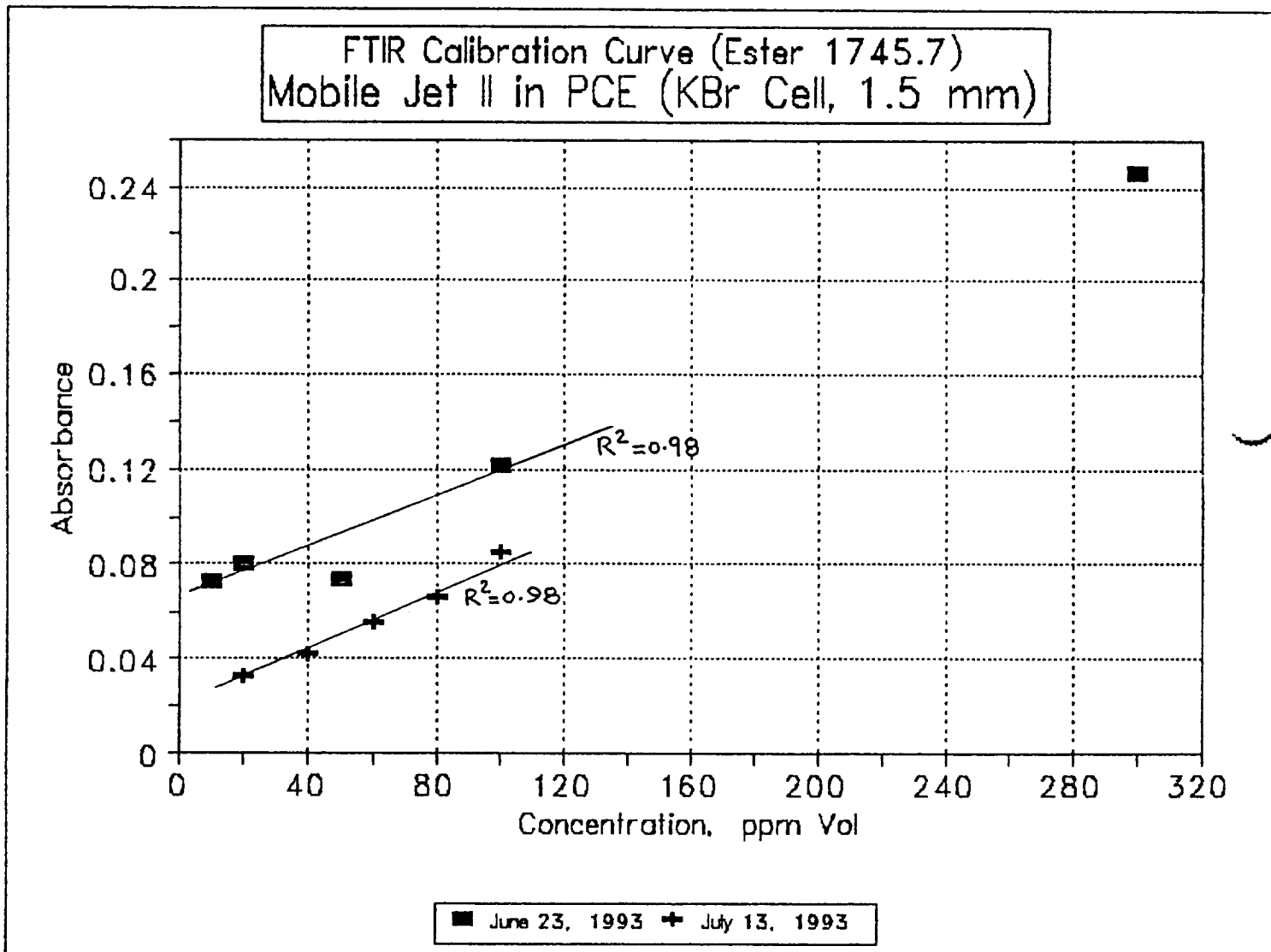


FIGURE 3-4

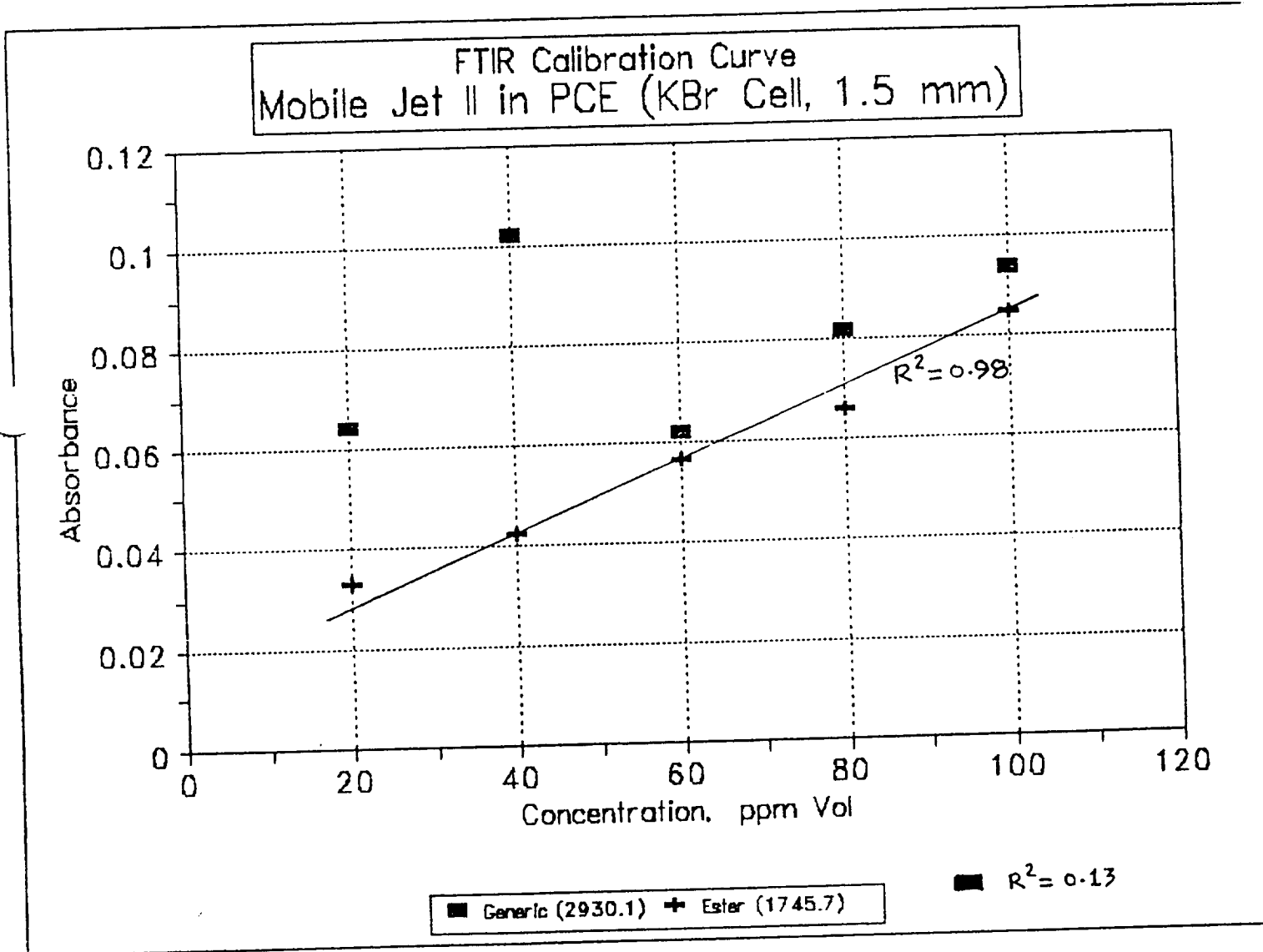


FIGURE 3-5

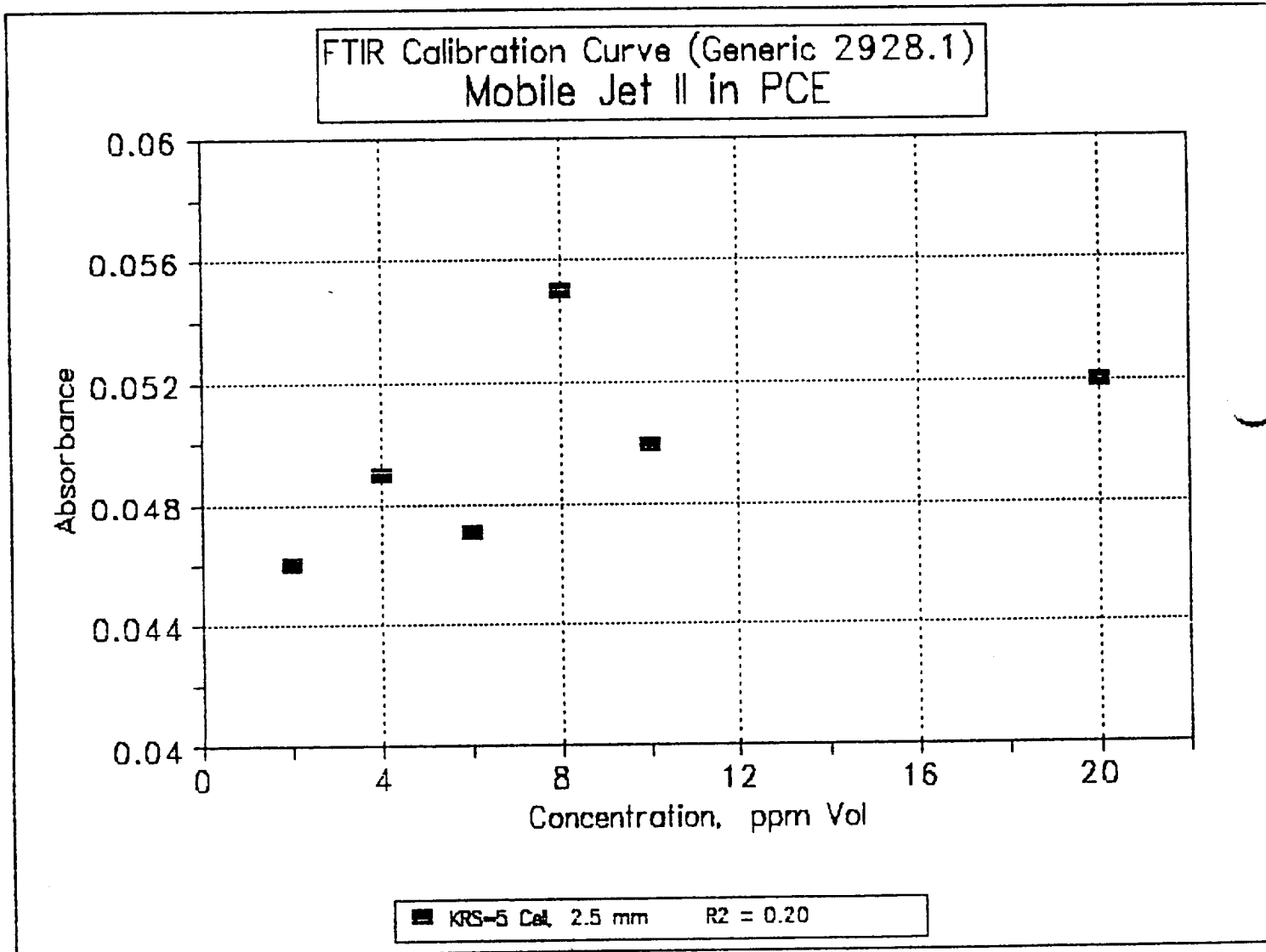


FIGURE 3-8

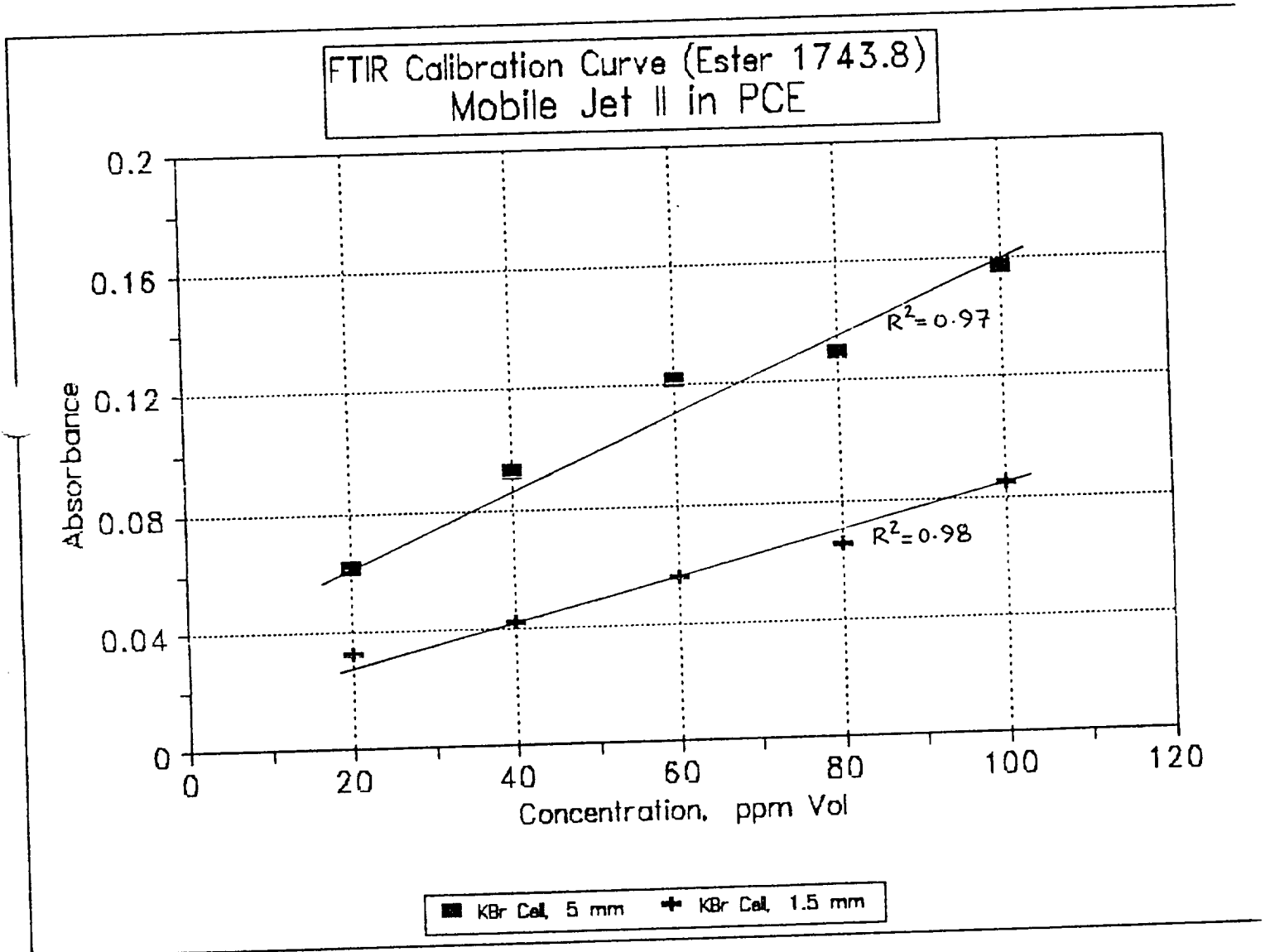


FIGURE 3-7

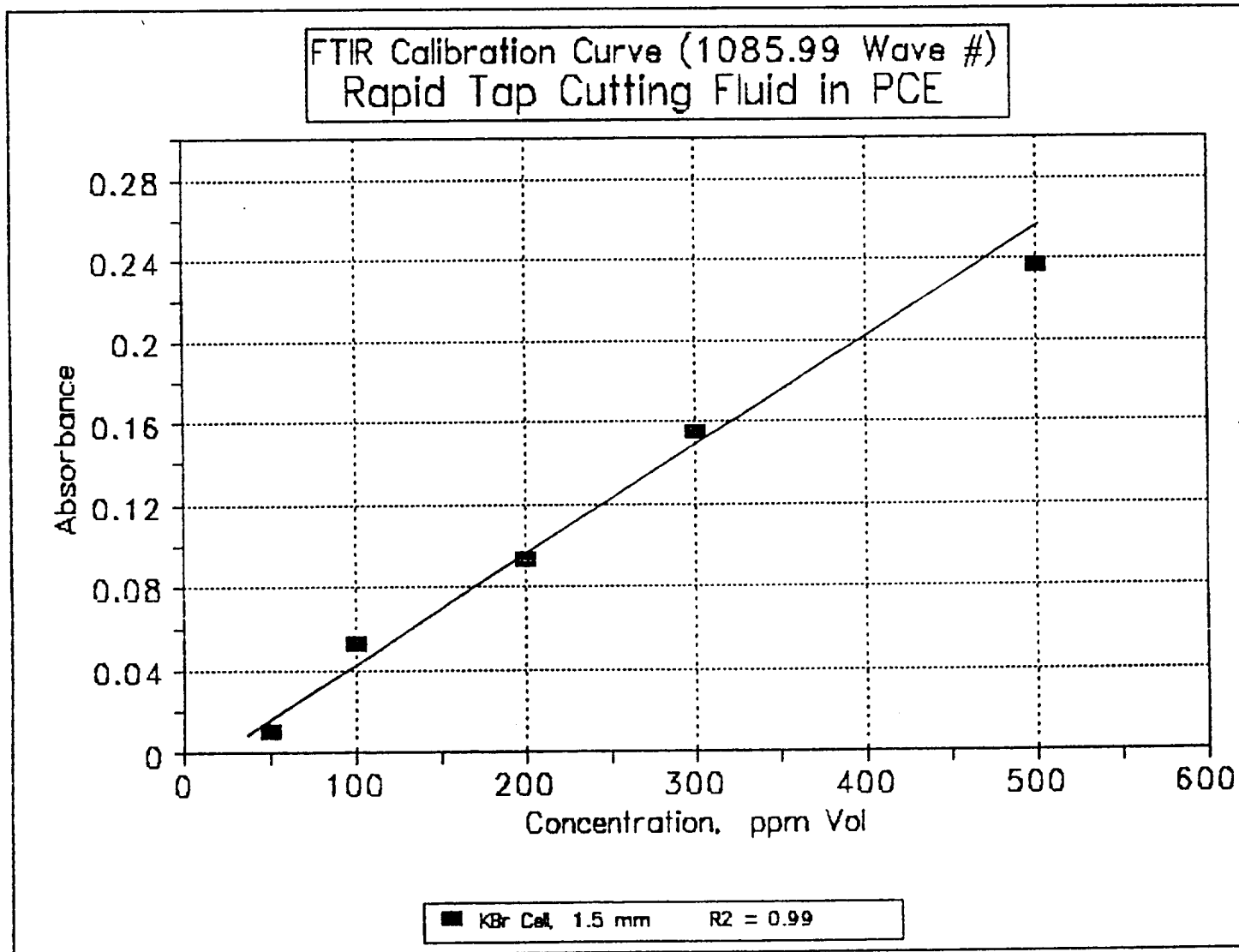


FIGURE 3-8

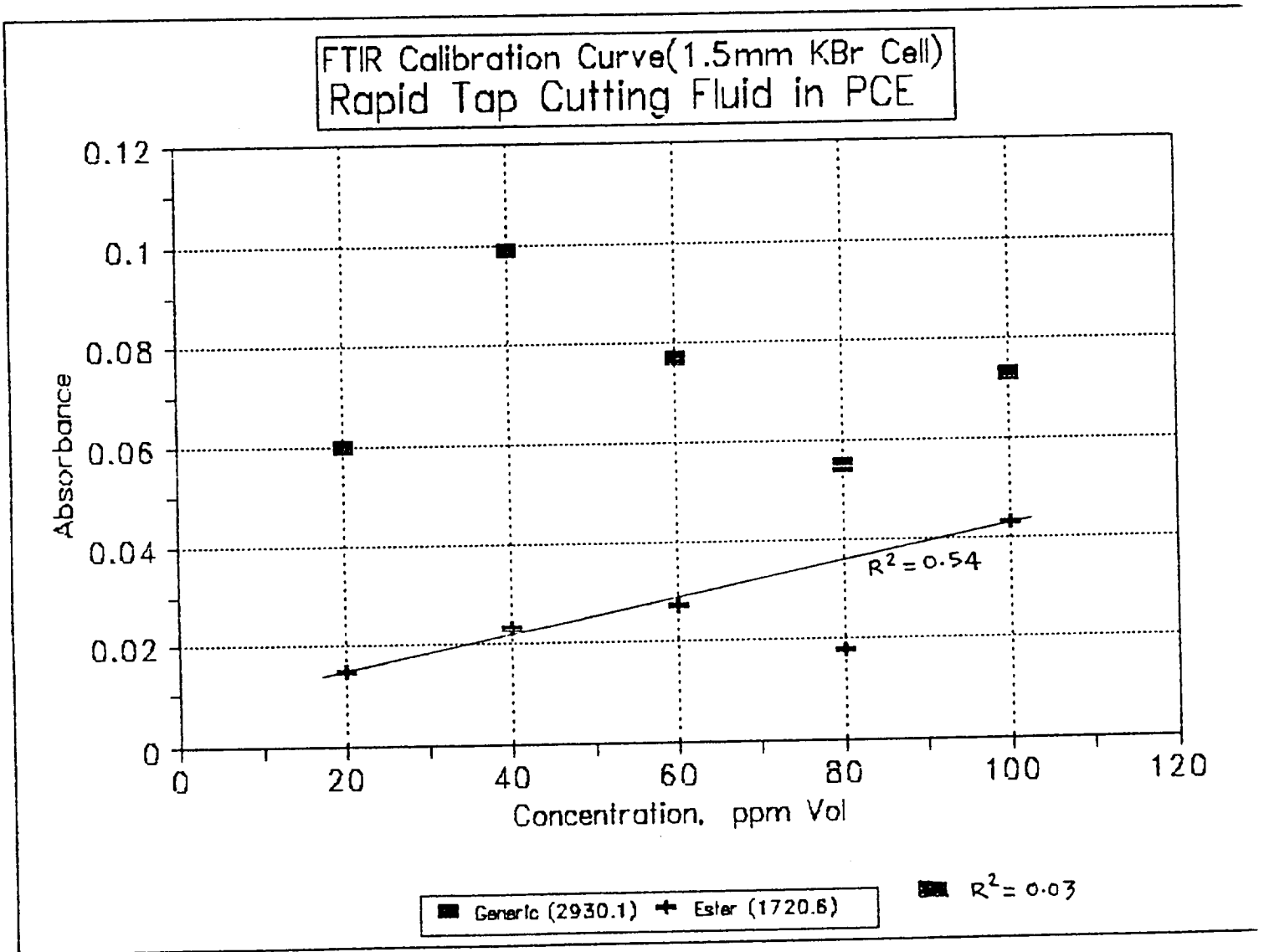


FIGURE 3-9

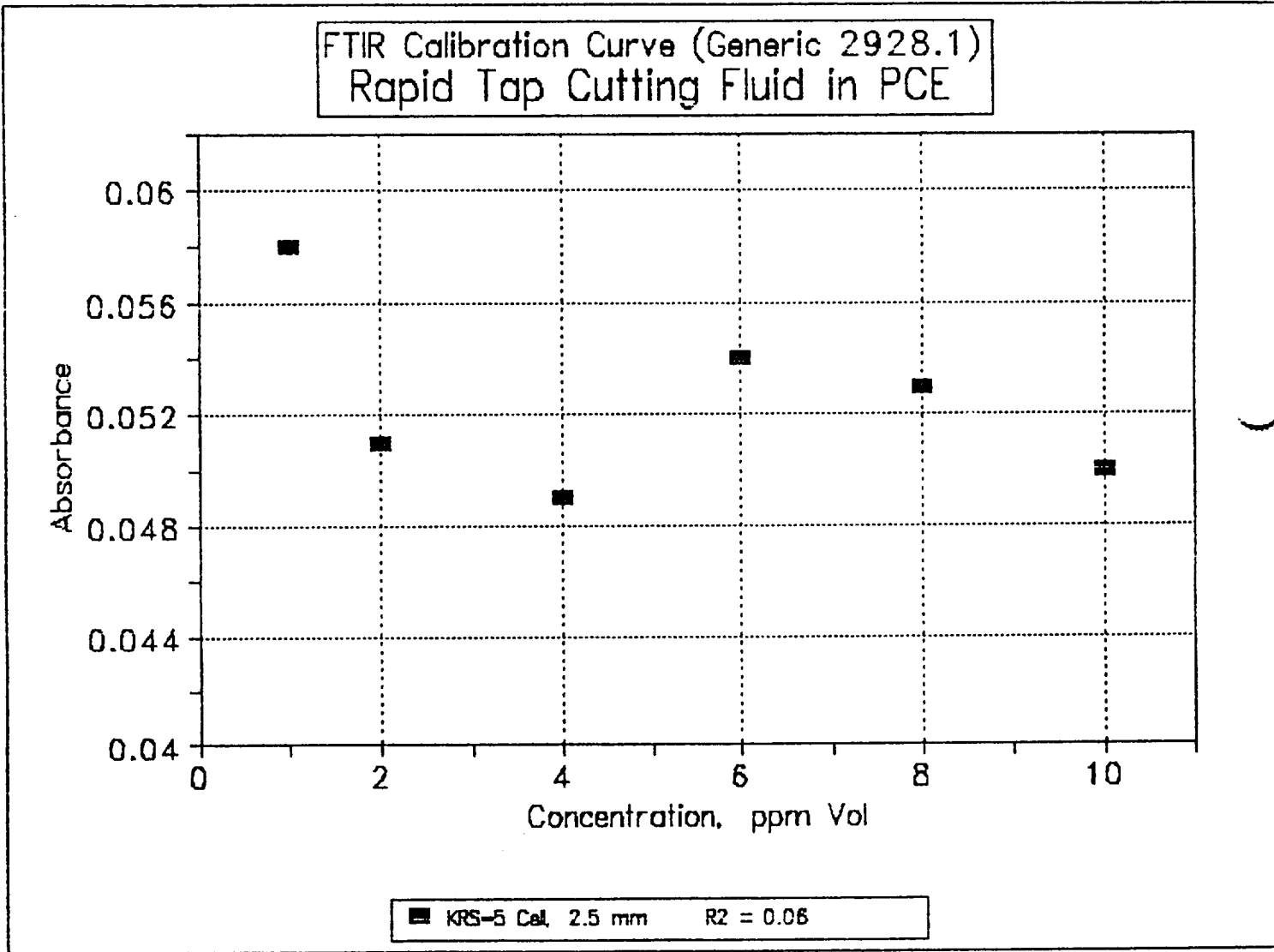


FIGURE 3-10

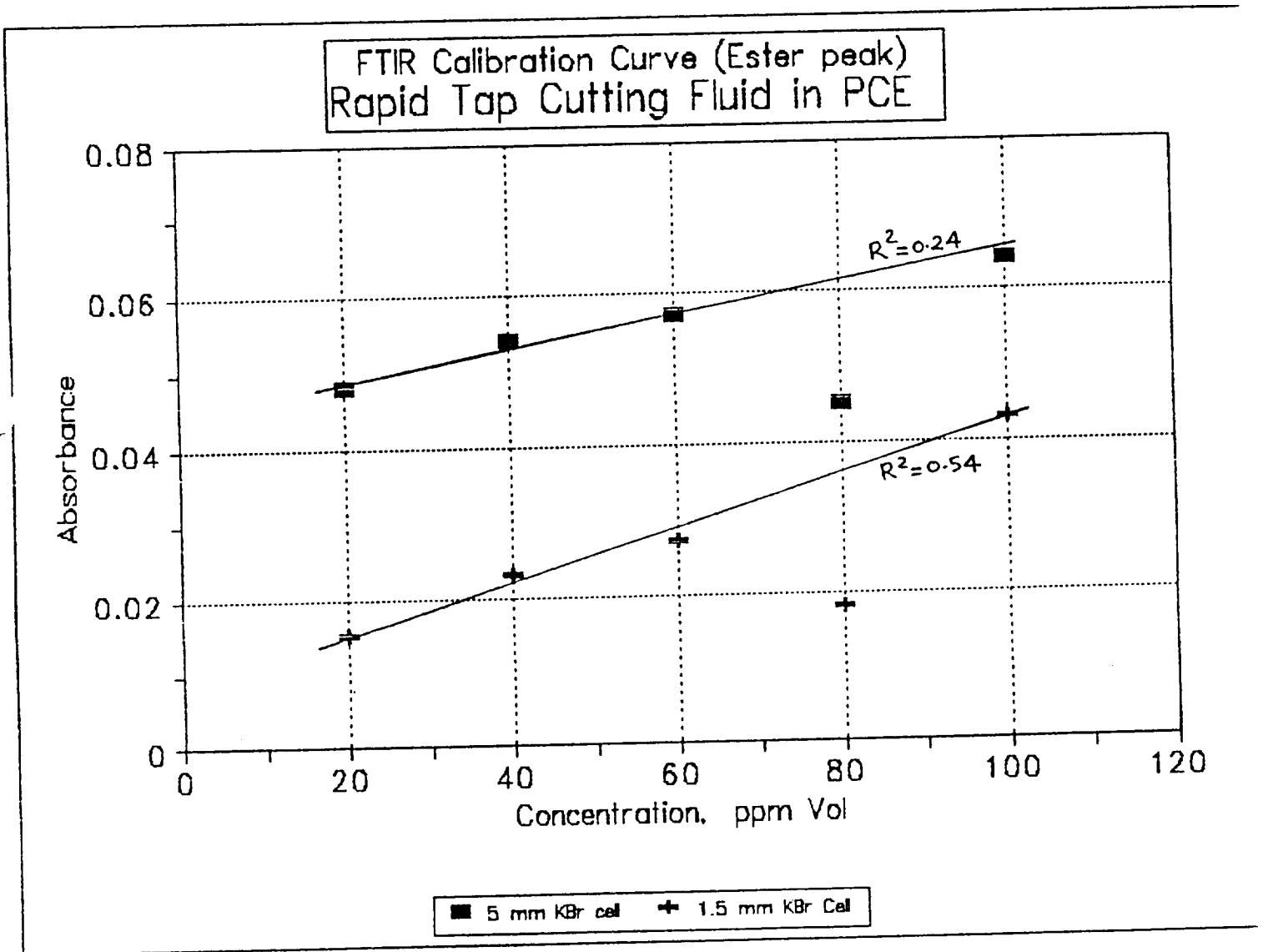


FIGURE 3-11

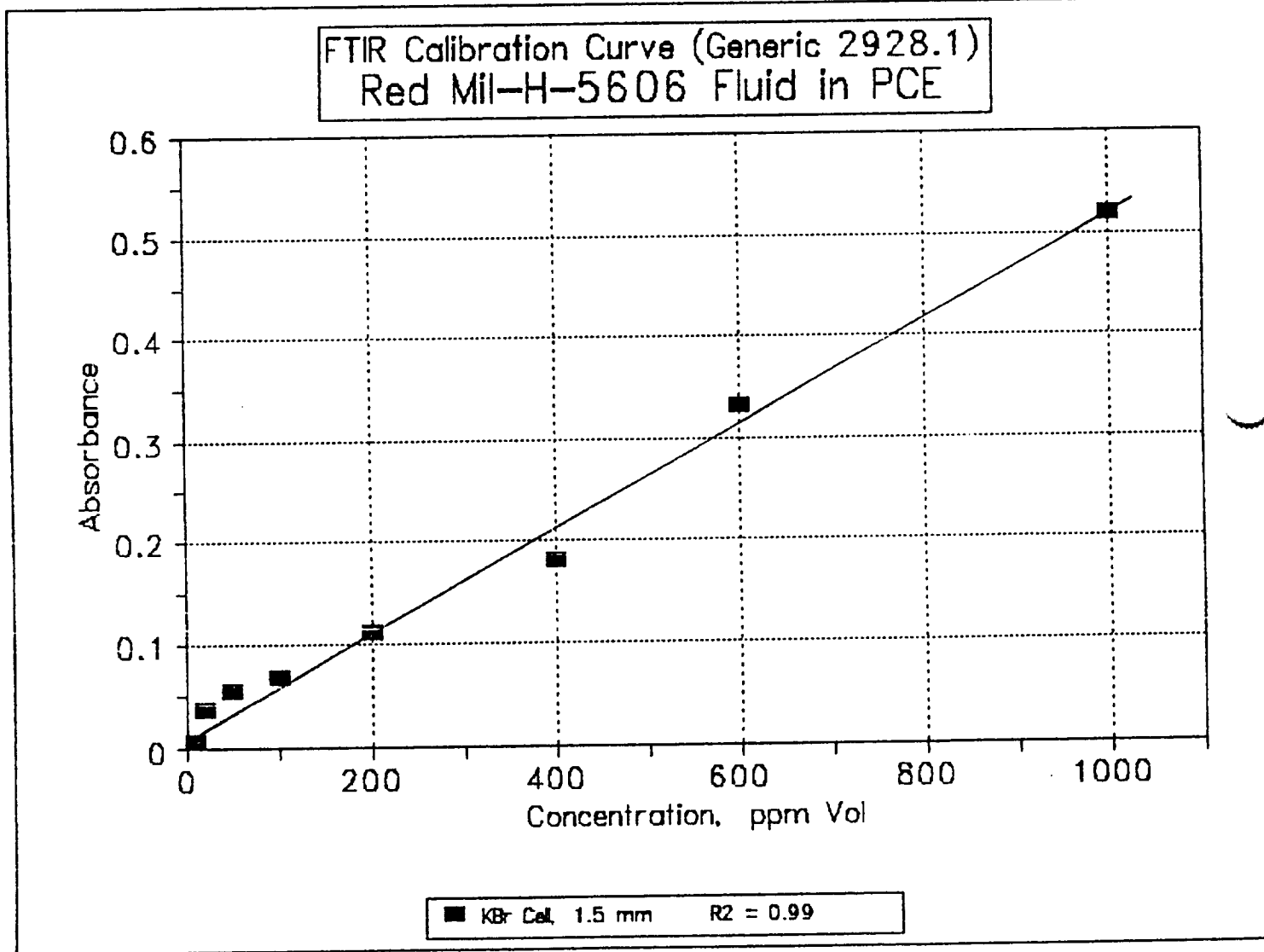


FIGURE 3-12

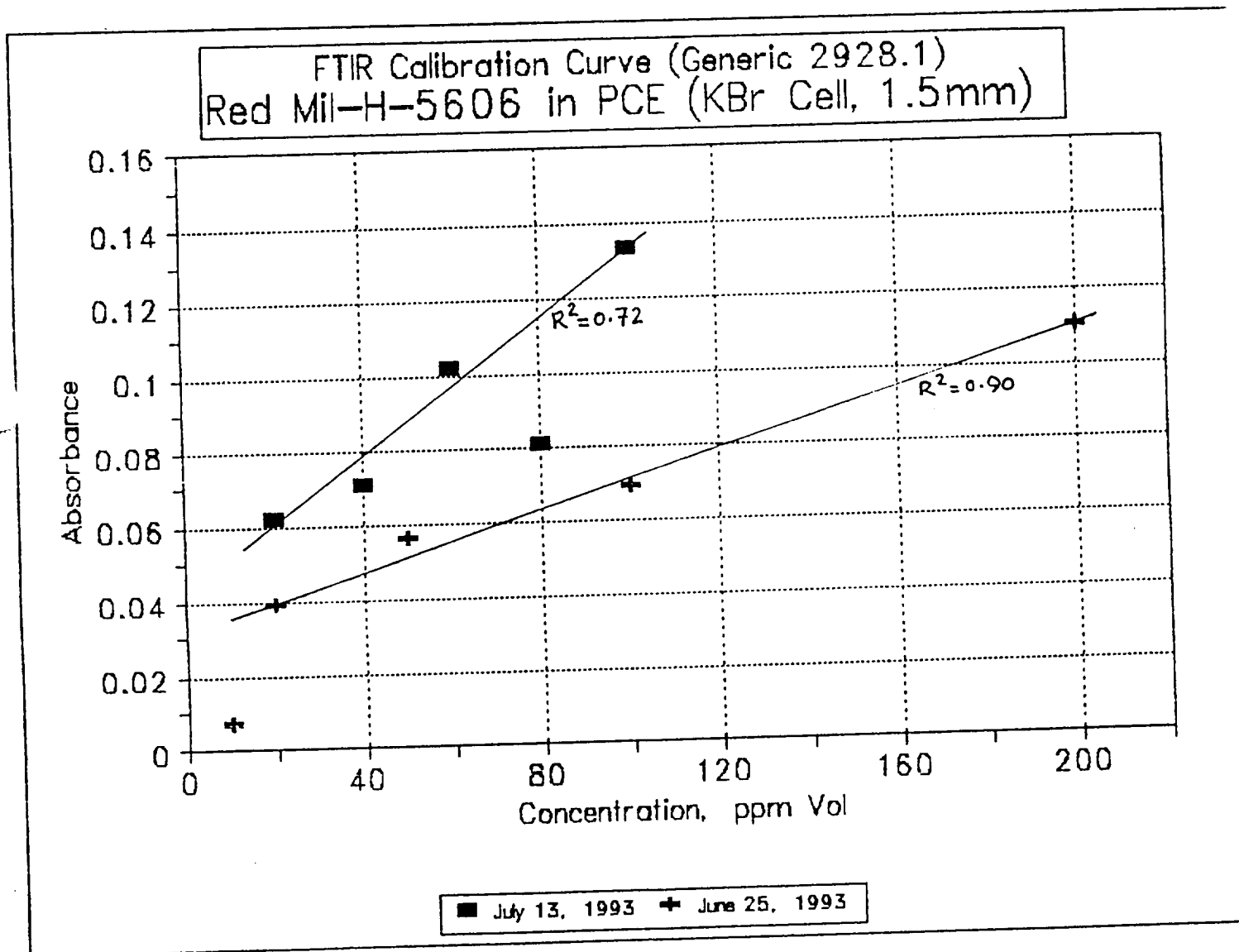


FIGURE 3-13

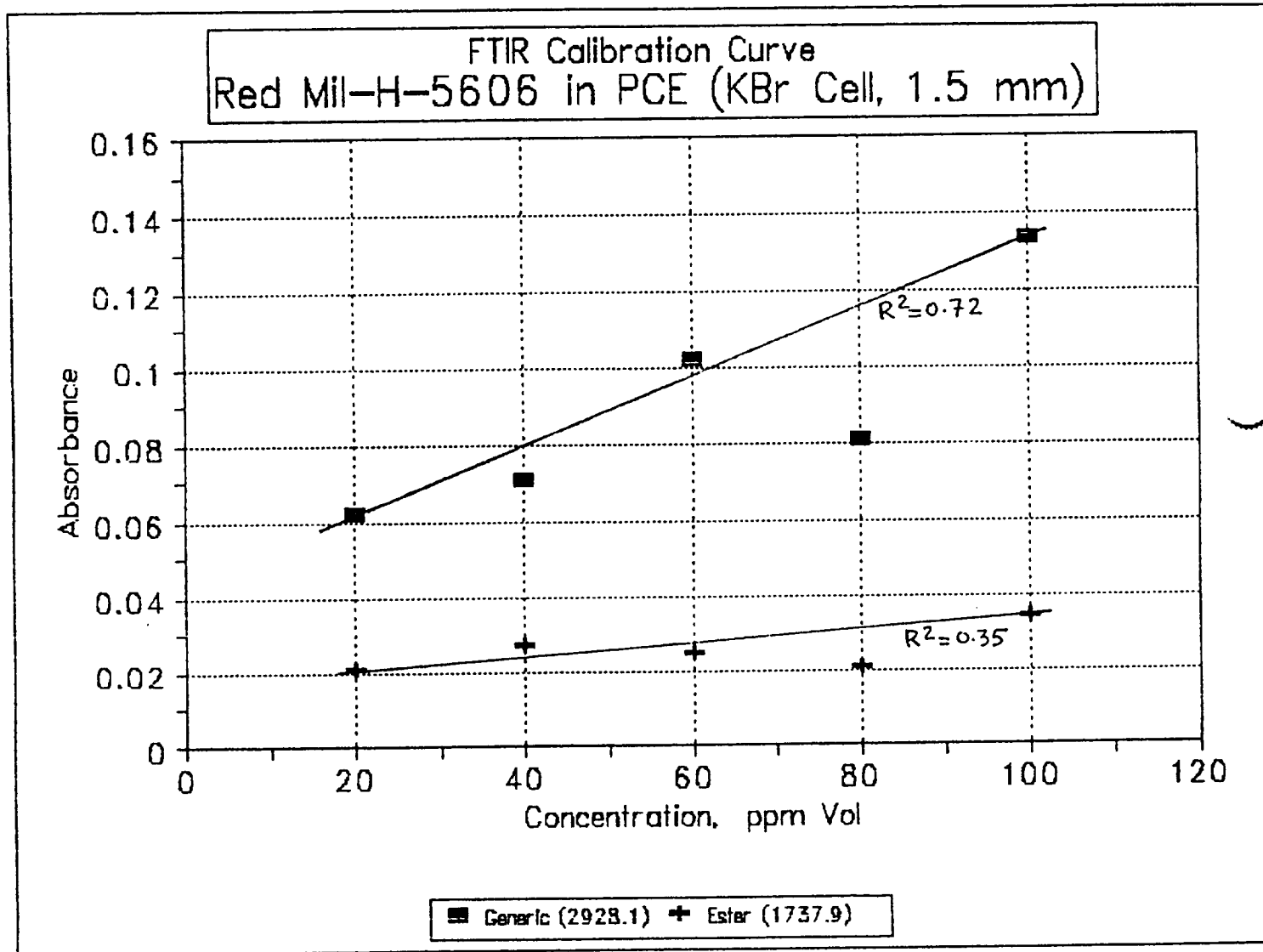


FIGURE 3-14

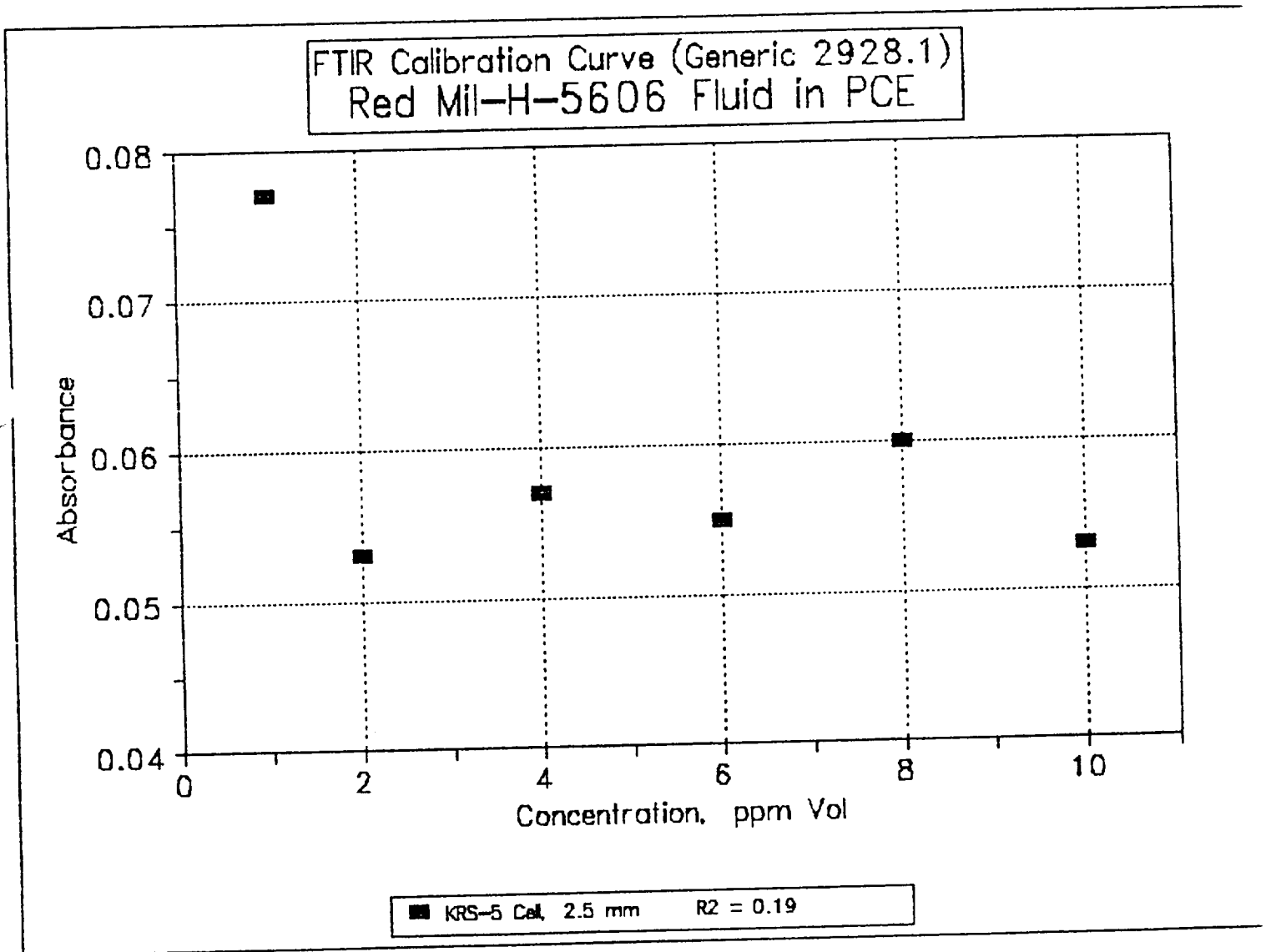


FIGURE 3-15

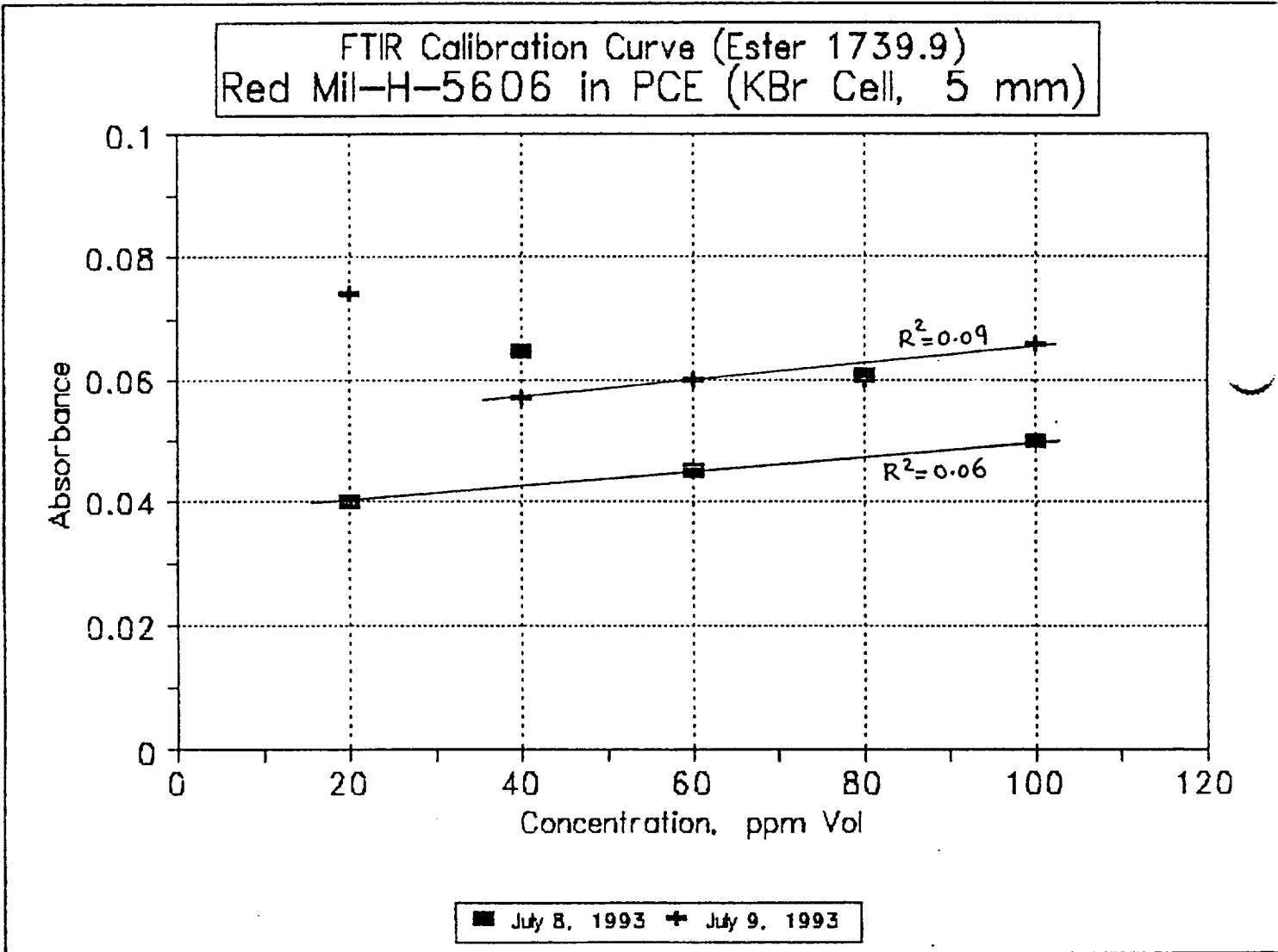


FIGURE 3-18

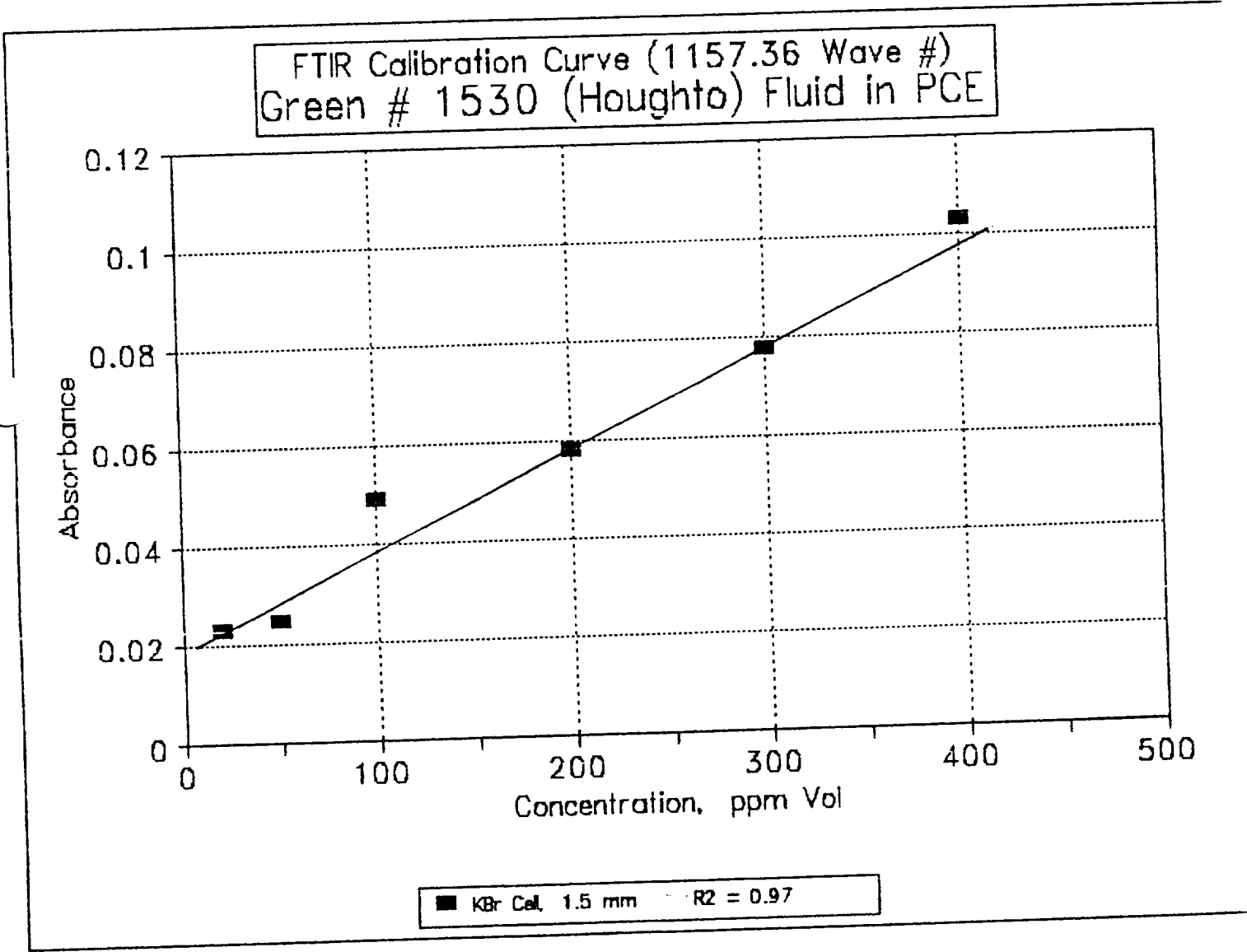


FIGURE 3-17

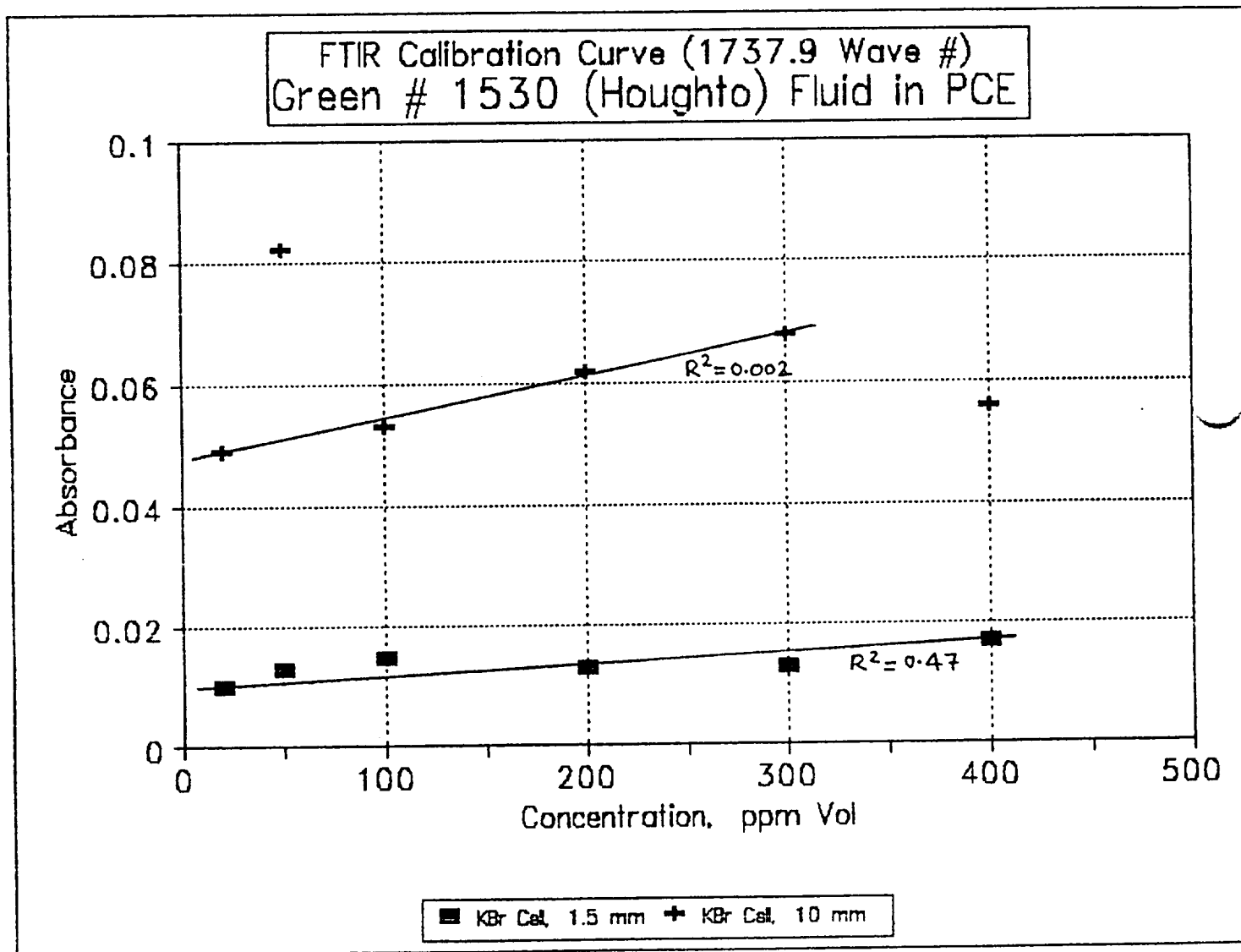


FIGURE 3-18

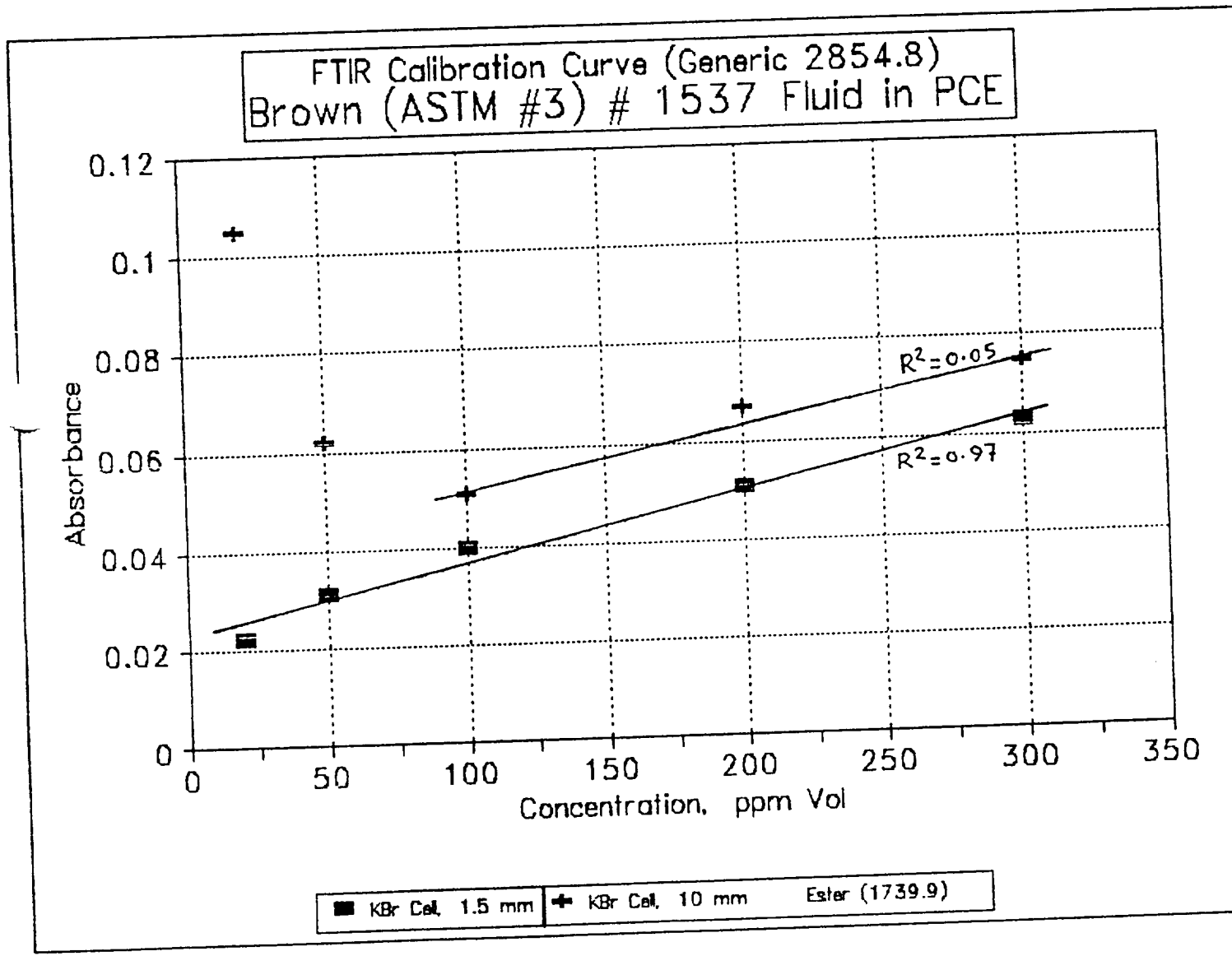


FIGURE 3-19

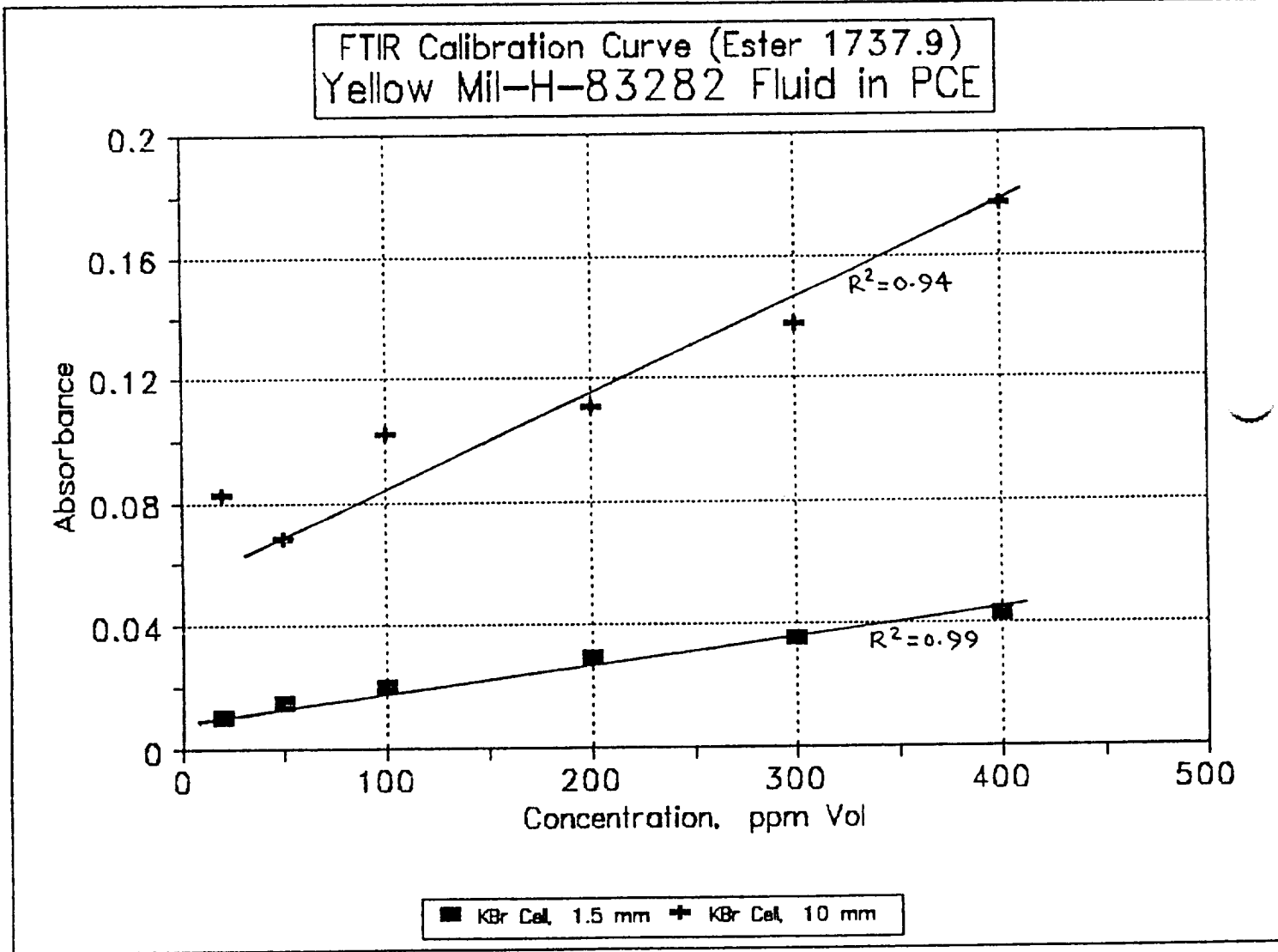


FIGURE 3-20

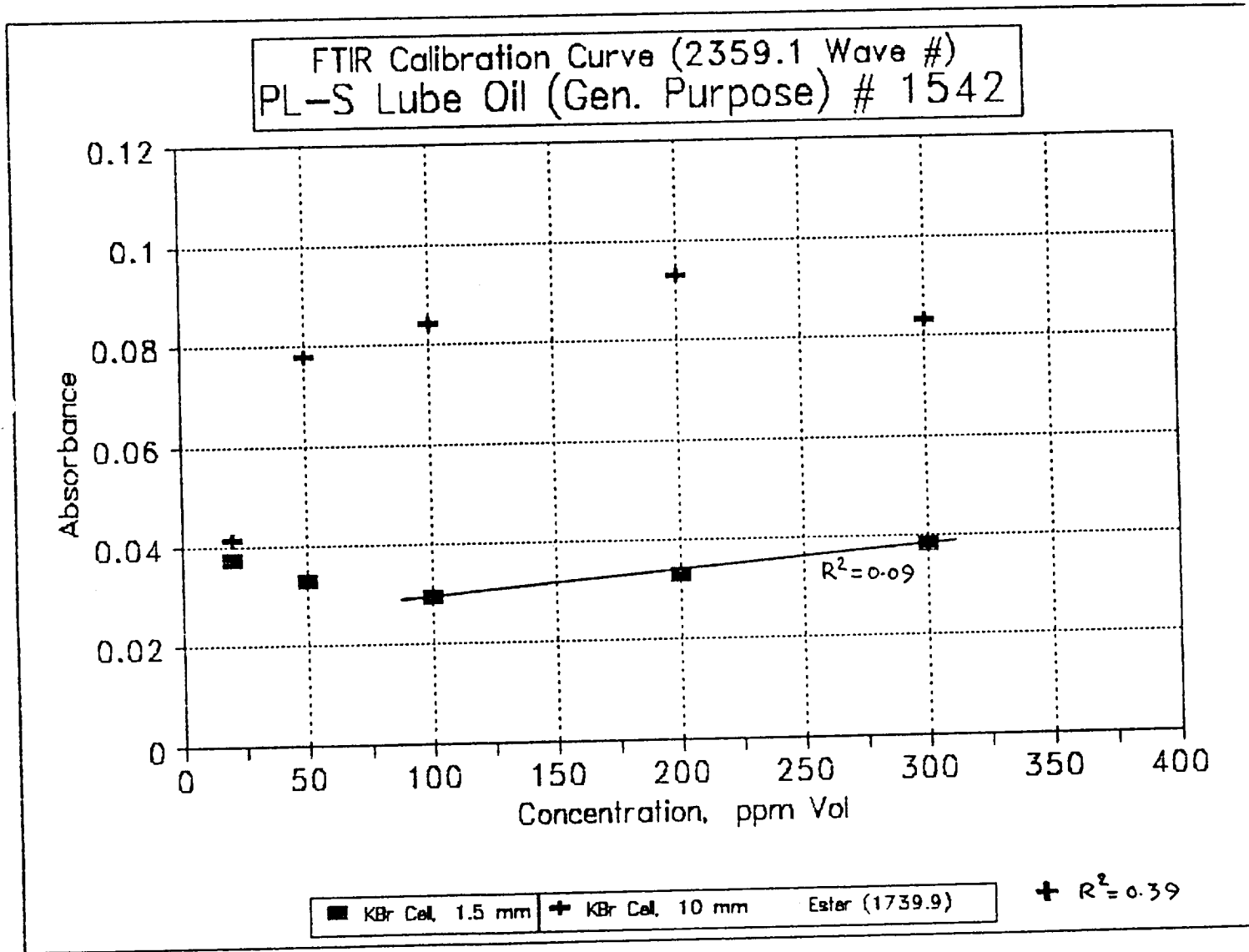


FIGURE 3-21

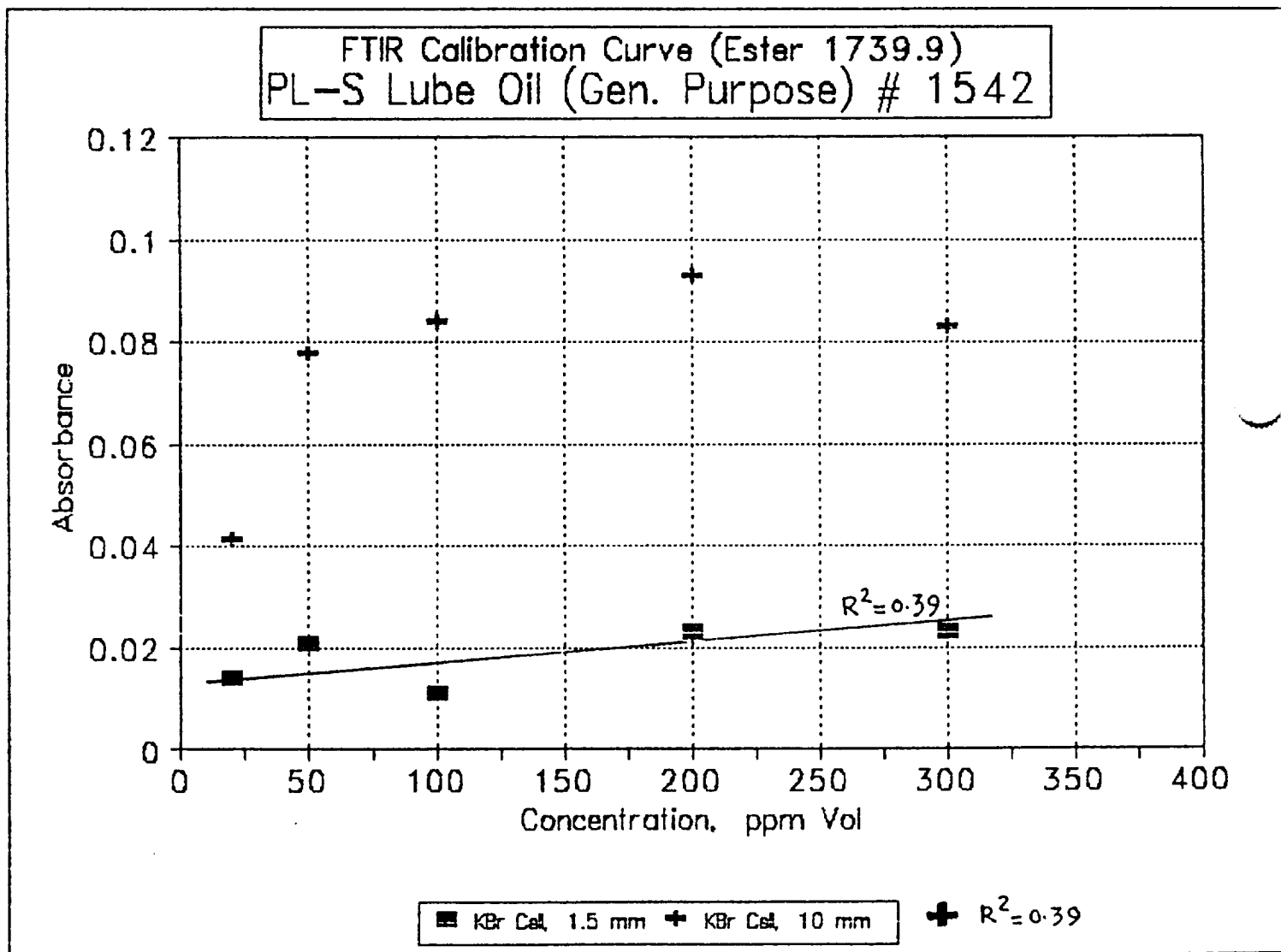
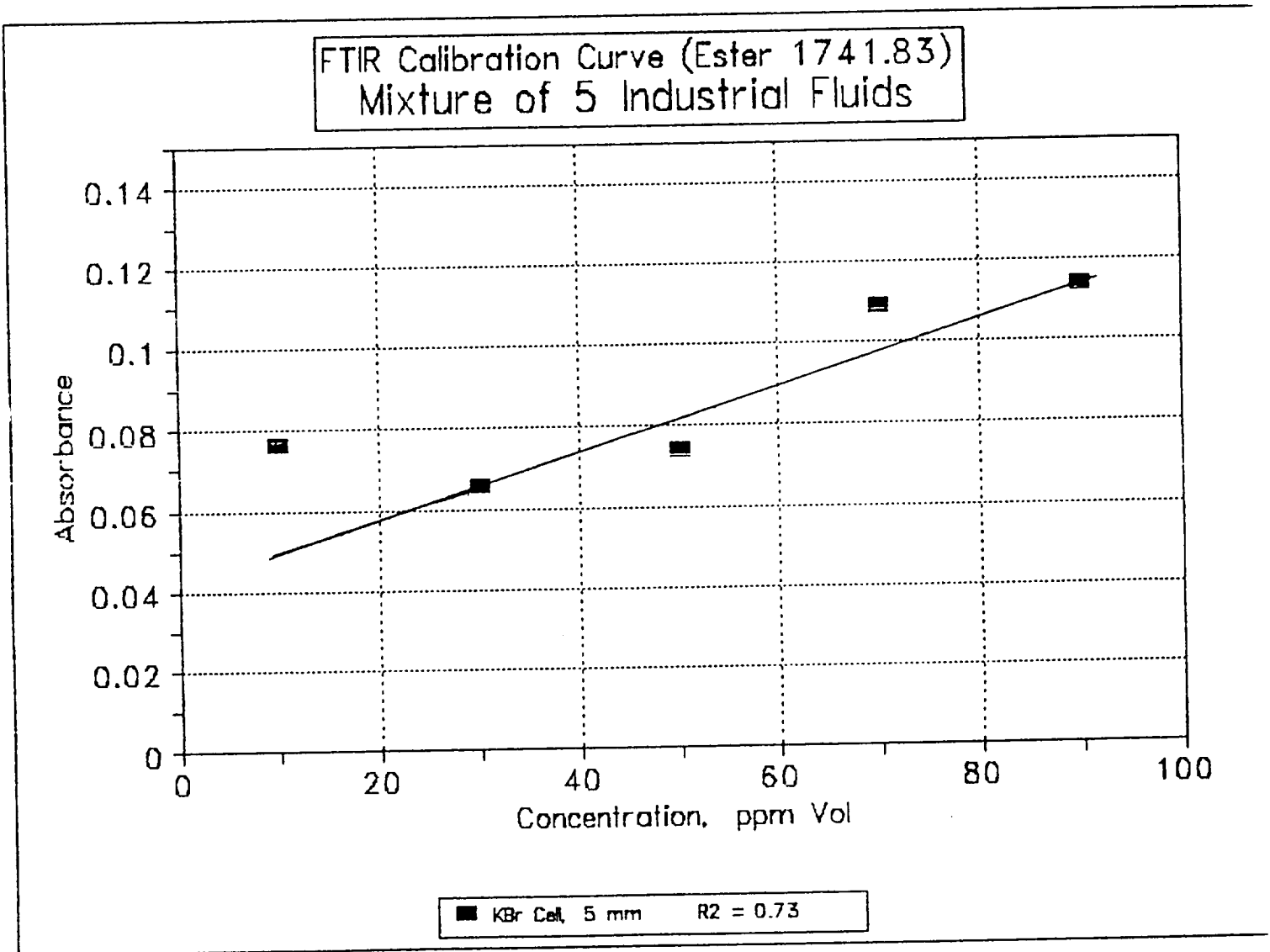
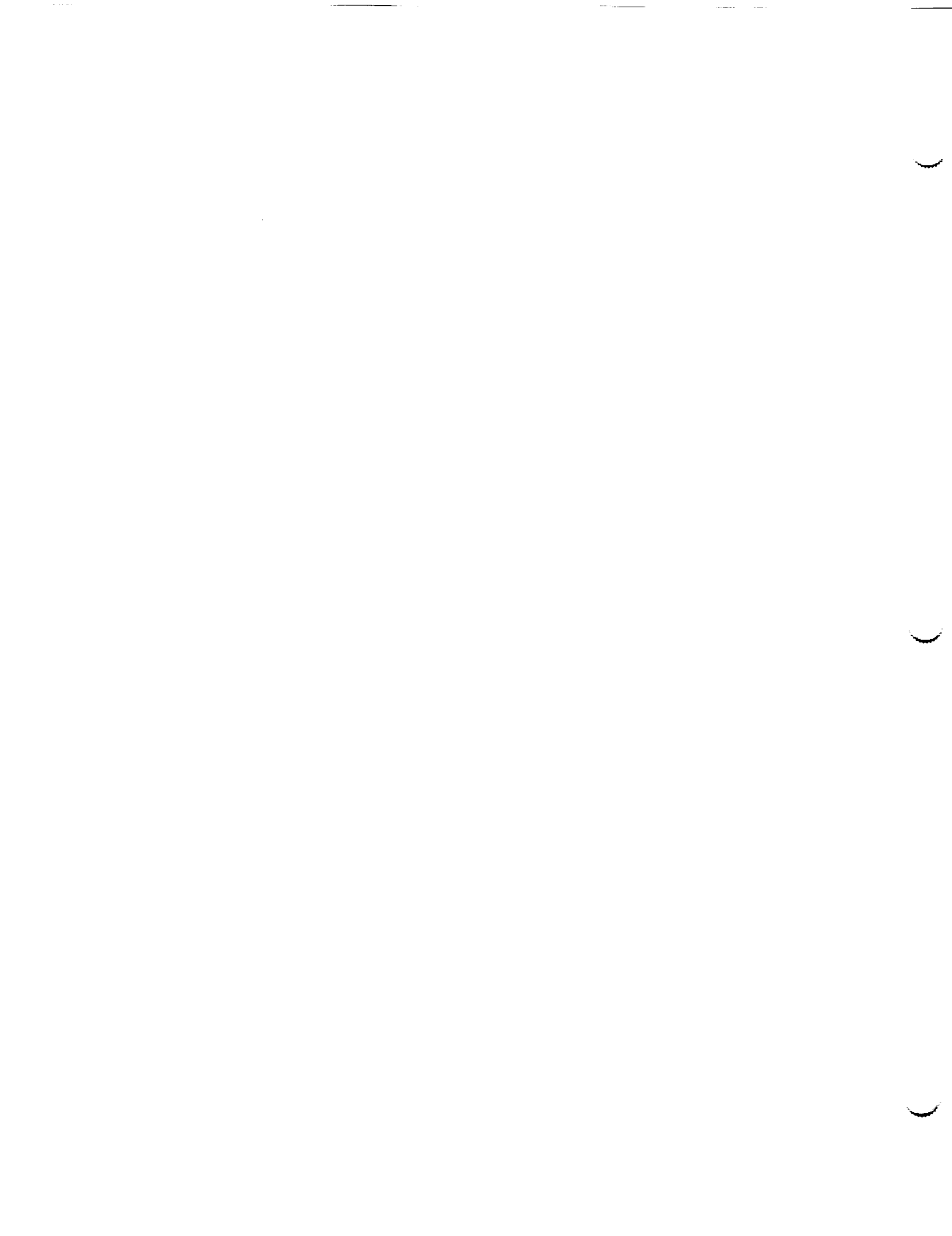


FIGURE 3-22





1994018770

N94-23244

1993 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA

197200
P. 98

STRESS CORROSION CRACKING PROPERTIES OF 15-5PH STEEL

PREPARED BY:	Dr. Ferdinand Rosa
ACADEMIC RANK:	Professor
UNIVERSITY AND DEPARTMENT:	University of Puerto Rico - Mayaguez Department of Mechanical Engineering
NASA/KSC	
DIVISION:	Material Science
BRANCH:	Failure Analysis and Materials Evaluation
NASA COLLEAGUE:	Rupert Lee
DATE:	August 5, 1993
CONTRACT NUMBER:	University of Central Florida NASA-NGT-60002 Supplement: 11

C-5

1

ACKNOWLEDGEMENTS

I would like to express my appreciation for being selected to participate in the NASA/ASEE Summer Fellowship Program. To Dr. Ramon Hosler [University of Central Florida] and to Mr. Tom Barron [NASA] for their capable administration of the program.

I also want to thank my NASA colleague Dr. Rupert Lee for his assistance and encouragement during this summer internship. To Peter Marciniak for his invaluable assistance whenever it was needed. My thanks go also to Ms. Kari Stiles for her dedication and invaluable help during my stay at Kennedy Space Center.

Last but not least I want to thank Mr. Scott Murray, Mr. William (Irby) Moore, Mr. Cole Bryant, and all the rest of the people in the Materials Section for making me feel at home during this summer.

ABSTRACT

Unexpected occurrences of failures, due to stress corrosion cracking (SCC) of structural components, indicate a need for improved characterization of materials and more advanced analytical procedures for reliably predicting structures performance.

Accordingly, the purpose of this study was to determine the stress corrosion susceptibility of 15-5PH steel over a wide range of applied strain rates in a highly corrosive environment. The selected environment for this investigation was a highly acidified sodium-chloride (NaCl) aqueous solution. The selected alloy for the study was a 15-5PH steel in the H900 condition. The Slow Strain Rate technique was selected to test the metals specimens.

SUMMARY

The catastrophic failure of some structural components, at Kennedy Space Center, due to environmentally assisted cracking has raised questions regarding the reliability of those structures. To that effect NASA has initiated a comprehensive program to identify materials which are immune to cracking under the above mentioned conditions, and recommend them for future applications.

The purpose of this study was to determine the behavior of some Precipitation Hardenable steels when exposed to a highly corrosive environment at different strain rates. The material selected for this study was a 15-5PH steel in the H900 condition. The environment selected consisted of a highly acidified Sodium-Chloride (3.5% NaCl) aqueous solution. The Slow Strain Rate technique was selected to test the metal's specimens. The test were programmed for strain rates between 10^{-3} to 10^{-5} inches per minutes.

The data obtained from these tests beside being useful for selection of materials on a sound engineering basis provides also for a better understanding of the Stress Corrosion Cracking phenomena.

TABLE OF CONTENTS

- 1.0 INTRODUCTION
 - 1.1 BACKGROUND
 - 1.2 STRESS CORROSION CRACKING MECHANISM
 - 1.3 LINEAR ELASTIC FRACTURE MECHANIC
 - 1.4 SLOW STRAIN RATE TECHNIQUES
- 2.0 MATERIALS AND PROCEDURES
 - 2.1 MATERIALS
 - 2.2 EQUIPMENT
 - 2.3 EXPERIMENTAL PROCEDURE
- 3.0 RESULTS AND DISCUSSION
- 4.0 CONCLUDING REMARKS
- 5.0 RECOMMENDATIONS
- 6.0 APPENDICES
 - 6.1 APPENDIX A
- 7.0 REFERENCES

LIST OF FIGURES

FIGURE	TITLE
1	Stress intensity factor vs specimen thickness typical curve
2	Crack growth rate vs stress intensity factor typical curve
3	Photomicrograph of 15-5 PH steel in the H900 condition showing precipitates distribution
4	Compact tension(CT)type, fracture toughness specimen
5	Load vs crack-mouth opening displacement curve
6	Tensile stress vs crack depth diagram illustrating the regions of no-crack, subcritical crack growth, and catastrophic failure

LIST OF TABLES

TABLE	TITLE
1	Nominal composition of 15-5PH steel
2	Heat treatment of 15-5PH steel
3	Mechanical properties of vacuum melted 15-5PH steel

1.0 INTRODUCTION

1.1 BACKGROUND

The solid rocket booster used to launch the space shuttle, use ammonium perchlorate as fuel and release approximately seventeen(17) tons of hydrochloric acid into the atmosphere in each launching. This acidification of the marine atmosphere has led to severe problems and premature failure of various structural components and critical equipment [1] at the launching facilities at Kennedy Space Center(KSC).

The fact that some structural components have failed catastrophically due to environmental cracking, has raised questions regarding the reliability of those structures. Even if catastrophic failures are rarely observed in practice, when they occur, they may be more costly in terms of human life and property damage than other types of failures. To that effect NASA has been working in a comprehensive program to identify materials which are not susceptible to cracking in the launch-pad's environment when stressed, and recommend them for future applications.

1.2 STRESS CORROSION CRACKING MECHANISM

Delayed failure of structural components subjected to an aggressive environment may occur under statically applied stresses well below the yield strength of the material. Failure under these conditions is caused by stress corrosion cracking(SCC) and has long been recognized as an important failure mechanism.

Although many tests have been developed to study this mode of failure the underlying mechanism for SCC are yet to be resolved [2,3,4] and quantitative design procedures against its occurrence are yet to be established. These difficulties are caused by the complex chemical, mechanical, and metallurgical interactions; the many variables that affect the behavior; the extensive data scatter [5,6]; and the poor correlation between laboratory test results and service experience.

Cracking of materials may be either intergranular or transgranular and may progress at velocities between 10^{-9} to 10^{-1} mm/s. Three broad categories of stress corrosion mechanism can be identified:

1. Pre-existing path mechanism - This mechanism relates the cracking susceptibility to the chemical activity of the grain boundaries (i.e. precipitates).
2. Strain assisted active path mechanism - This mechanism is related to the rupture of a protective film at the crack tip, followed by metal dissolution by the corrosive environment.
3. Absorption mechanism - This mechanism is based on the chemisorption of an environmental species on the crack tip which reduces the surface energy, and therefore reduces the local fracture strength of the metal lattice.

1.3 LINEAR ELASTIC FRACTURE MECHANIC

The application of Linear Elastic Fracture Mechanic (LEFM) concepts has met with considerable success in the study of SCC [7,8]. Because environmentally enhanced crack growth and stress intensity factor (K) can be used to characterize the mechanical component of the driving force in SCC.

The critical stress intensity factor or fracture toughness (K_{IC}) represents the inherent ability of a material to withstand a given stress-field intensity at the tip of a crack and to resist progressive tensile crack extension under plane strain conditions. Plane strain conditions requires that:

$$B = 2.5 (K_{IC} / S_{YS})^2, \text{ where:}$$

B = specimen thickness

S_{YS} = tensile strength

For materials that are susceptible to crack growth in a particular environment this threshold value is called K_{ISCC} and represent the value below which crack propagation does not occur for a given material-environment combination under plane strain conditions (See Fig.1).

Stress corrosion crack growth rates have been investigated in various material-environment combination, and the results suggest that the crack growth rate as a function of the stress intensity factor can be divided in three regions (See Fig.2). In region I the rate of crack growth is strongly dependent on the magnitude of the stress intensity factor, such that small changes in K results in large changes in crack growth rate. Region I also exhibit a stress intensity factor (K_{ISCC}) below which cracks do not propagate under sustained loads for a given material-environment system. In region II crack growth rate, for many systems, is moderately dependent on the magnitude of K and for some systems like high strength steels in gaseous hydrogen, crack growth rate is independent of K . The crack growth rates in region III increases rapidly with K as K approaches K_{IC} of the material.

1.4 SLOW STRAIN RATE TECHNIQUE

Before the early 1960's, constant load and constant strain testing on smooth and notched specimens of various configurations became very popular. However the 60's produced two accelerated-test techniques based on different mechanical approaches [9,10]. One of the techniques involved testing statically loaded mechanically precracked specimens using Linear Elastic Fracture Mechanics concepts. The second involved Slow Strain Rate testing of smooth specimens. These testing methods have often produced SCC in materials, where older techniques have failed to do so.

More recently Constant Strain Rate test have become widely accepted as quality control or screening technique quite apart from their usefulness in mechanistic studies. The prime justification for this technique is that it accelerates a

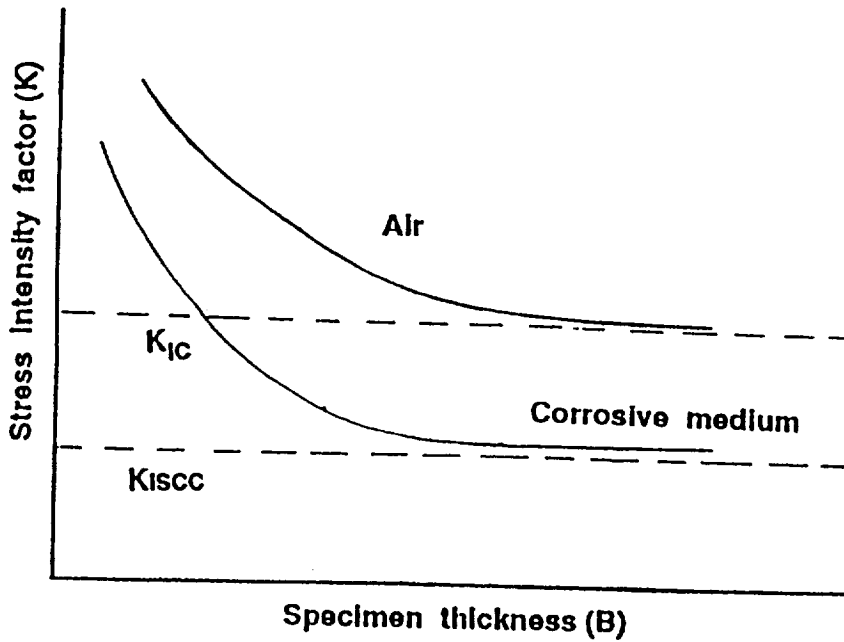


Fig. - 1 Stress Intensity Factor vs Specimen Thickness Curve

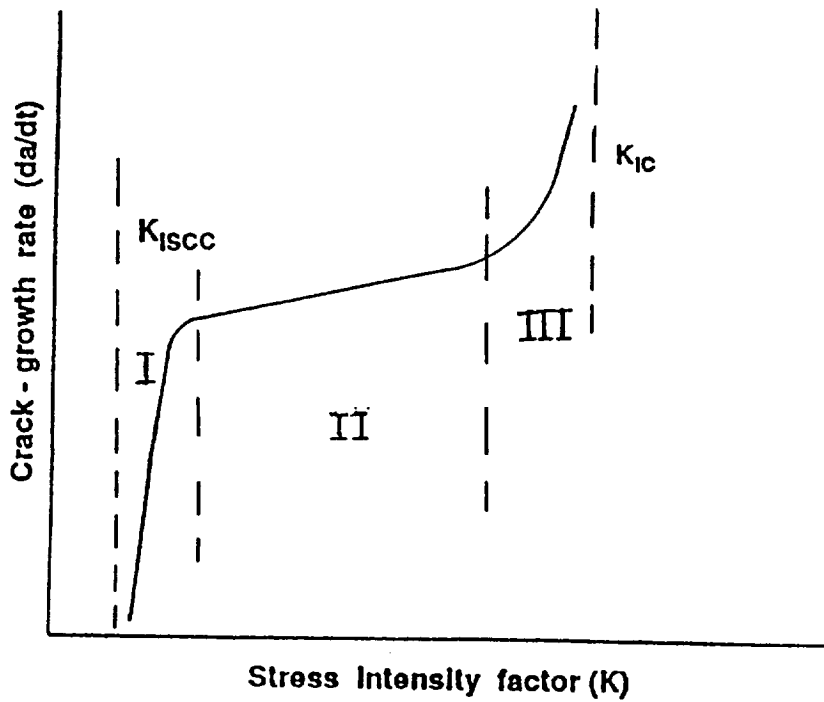


Fig. - 2 Typical Curve Of Crack-growth Rate vs Stress Intensity Factor

known rate determining step in the cracking mechanism of ductile alloys-aqueous environment system (i.e. oxide-rupture rate). It is not surprising then that good correlations are observed between SCC susceptibility rated by this technique and by more protracted methods involving static loads.

Strain rate is one of the most important single parameter in evaluating SCC susceptibility of any metal or alloy in a given environment. If strain rate is too high, fracture of the material will be mostly mechanic (ductile) because the corrosion process cannot keep pace with the straining process. On the other side if the strain rate is too low, SCC may be prevented due to repassivation of the exposed base metal, which may be too fast compared to the frequency of the film rupture event. It has been observed that strain rates in the range of 10^{-5} to 10^{-1} s^{-1} tend to promote SCC in most cases.

2.0 MATERIALS AND PROCEDURES

2.1 MATERIALS

The materials used in this investigation consist of 15-5PH steel in the H900 condition. The nominal chemical composition is shown in Table 1, and the thermal condition is shown in Table 2. Table 3 is a summary of the mechanical properties of the material. A photomicrograph showing the distribution of precipitates is shown in Figure 3.

Compact tension (CT) specimens were machined from the as-received material, such that the crack growth direction was perpendicular to the material's rolling direction. Specimens were cleaned to eliminate grease and other impurities from the machining operation and then they were immersed in a magnesium chloride solution to produce the starting crack required in this type of test. The specimen configuration is shown in Figure 4. The specimens were then loaded into the testing machine and strained until rupture occurred.

2.2 EQUIPMENT

Tests were conducted in a Satec's MATS II Universal Testing Machine equipped with "NuVision II" software package for automating the system. The minimum strain rate applied by the machine was 8×10^{-5} inches per minute. The crack-mouth opening was measured with a double cantilever beam type strain gage. A Laser based type extensometer developed at KSC was also used to measure the crack-mouth opening.

2.3 TESTING PROCEDURE

Following mounting in the testing machine, the specimens were initially tested in air at different strain rates. A curve of applied load vs crack-mouth opening similar to the one shown in Figure 5 was obtained. From that graph the crack growth rate and the stress intensity factor for the material could be evaluated.

Table - 1 Nominal Composition Of 15-5 PH Steel

Cr	Ni	Cu	Mn	Si	Cb+Ta	C
14.0- 15.0	3.5- 5.5	2.5- 4.5	1.0- Max	1.0- Max	0.15- 0.45	0.07- Max

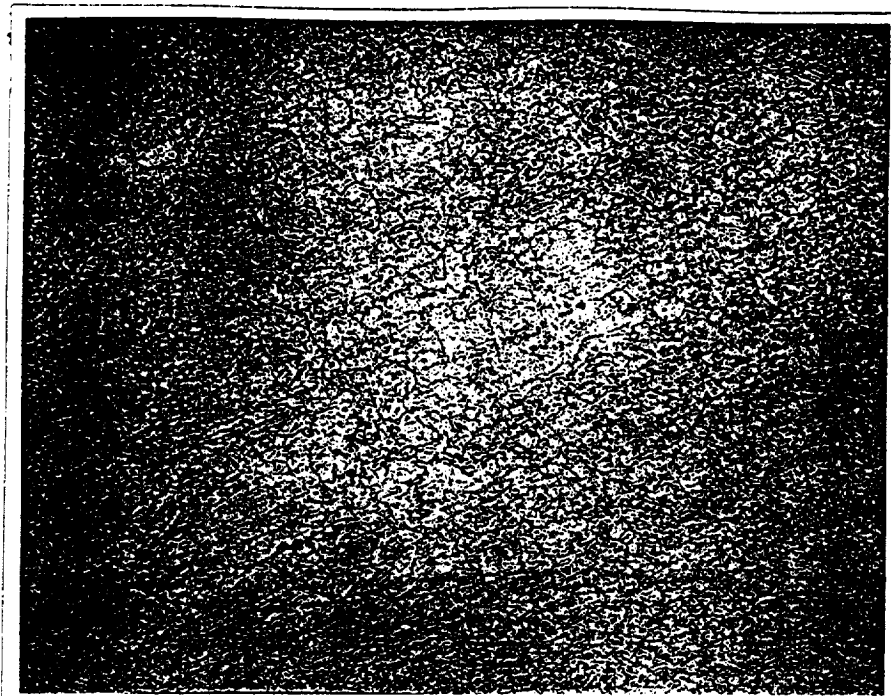
Table - 2 Heat Treatment Of 15-5 PH Steel

Temper	Heat Schedule
Solution Treated	1900 F for half (.5) hour Oil quench
H 900	800 F for one (1) hour Air cooled

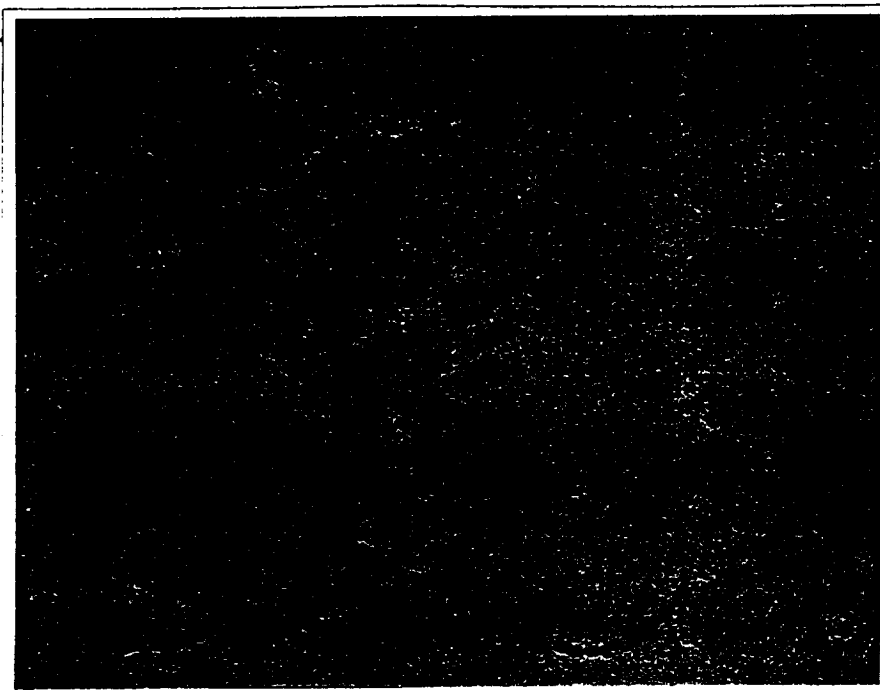
Table - 3 Mechanical Properties Of Vacuum Melted 15-5 PH Steel

Temper	Su	Sys	% Elong	% R.A.	Hardness
H 900	199.6	178.9	17	61	HB 401

Su = Tensile Strength (KSI)
 Sys = 0.2% offset yield strength (KSI)
 R.A. = Reduction in area
 HB = Brinell hardness

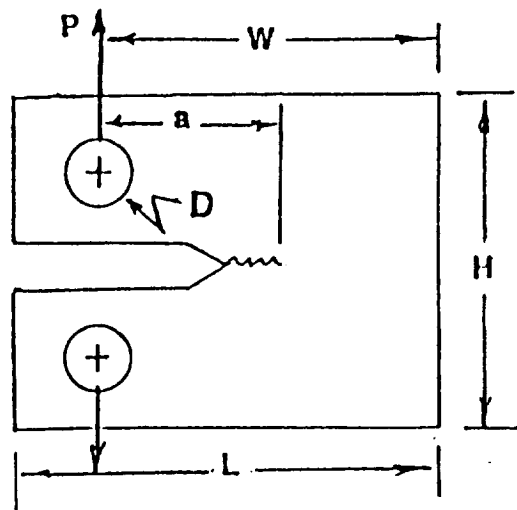


Mag.-100X



Mag.-500X

Fig.-3 Photomicrograph Of 15-5PH Steel In The H900 Condition Showing The Precipitate Distribution.



$H = 1.2W$
 $L = 1.25W$
 $D = 0.25W$
 $a = \text{crack length}$
 $P = \text{applied load}$

Fig. - 4 Compact - Tension (CT) Type Fracture Toughness Specimen

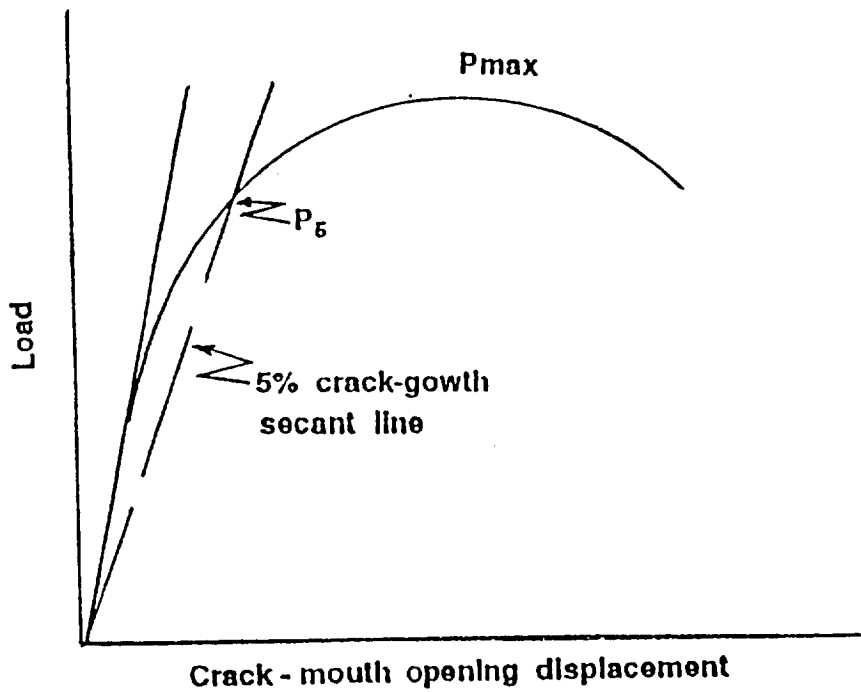


Fig. - 5 Load vs Crack-mouth Opening Displacement Curve

A similar set of tests in a corrosive solution of 3.5% NaCl acidified to 1N.HCl was also scheduled. From the data obtained from those tests the crack growth rate and the critical stress intensity factor under corrosive conditions were to be evaluated.

3.0 RESULTS AND DISCUSSION

3.1 GENERAL

The constant strain rate technique is a method to assess susceptibility of metals and alloys to SCC. It provides a rapid laboratory method to determine the SCC susceptibility of materials in environments in which other tests do not readily promote SCC. The results are positive in that failure occurs either in a ductile manner or prematurely in a brittle mode if SCC occurs.

Several properties are used to define and compare the severity of SCC of materials and aggressiveness of environments. Generally a measure of the time to failure, or reduction in ductility in a corrosive environment is compared to the behavior in an environment which does not promote SCC, for example air. Increased severity of SCC is indicated by shorter times to failure or reduced ductility, as measured by reduction in area or reduction in elongation. The presence or absence of SCC on the fractured specimen can be unequivocally determined only by metallographic examination. Results of those examinations can be presented quantitatively by comparing the number or length of secondary stress corrosion cracks.

The usefulness of Fracture Mechanics for defining SCC tendencies in high strength metals is derived from the ability to use the parameter K_{ISCC} for calculating the stress-flaw size combination necessary for the initiation of cracks growth. The value of K_{ISCC} is calculated from the load vs crack-mouth opening displacement curve. From that graph the stress intensity factor for the specific material-sample geometry-environment (K_Q) will be calculated according to:

$$K_Q = [P_Q / B\sqrt{W}] \cdot f(a/W) \quad \text{where:}$$

B= specimen thickness

W= specimen width

a= crack length, and

$$f(a/W) = [(2+a/W)(.886+464a/W-13.32a^2/W^2+14.72a^3/W^3-5.6a^4/W^4)]/[1-a/W]^{3/2}$$

$P_Q = P_5$ if the load at every point in the graph preceded P_5
(See Fig.5)

$P_Q = P_{max}$ if there is a maximum load preceding P_5 in the
graph (Fig.5)

The value of K_Q will be equivalent to K_{IC} or K_{ISCC} if:

$$P_{max}/P_5 \geq 1.10 \quad \text{and}$$

$$B \text{ and } a \geq 2.5(K_Q/S_{ys})$$

if the above two constraints are not satisfied then the test is not valid for determining K_{IC} and K_{ISCC} , and a new test have to be done using a thicker specimen, usually 1.5 times thicker. With values of K_{IC} and K_{ISCC} obtained from the slow-strain rate test and ultimate and yield strength of the material, a curve similar to that in Fig.6 can be obtain, that will permit the designer a better way of predicting the life of the structure for a given design stress-crack length combination.

The application of the Fracture Mechanic approach to design concepts relies on the definition of the boundary lines on the stress vs crack-depth diagram. Figure 6 shows the no-crack growth, subcritical-crack growth, and catastrophic failure regions. The boundaries are experimentally definable, with limits imposed by S_{ys} and S_U for smooth specimens and K_{IC} and K_{ISCC} for pre-cracked samples. Where S_{ys} is the material yield strength and S_U is the maximum tensile strength.

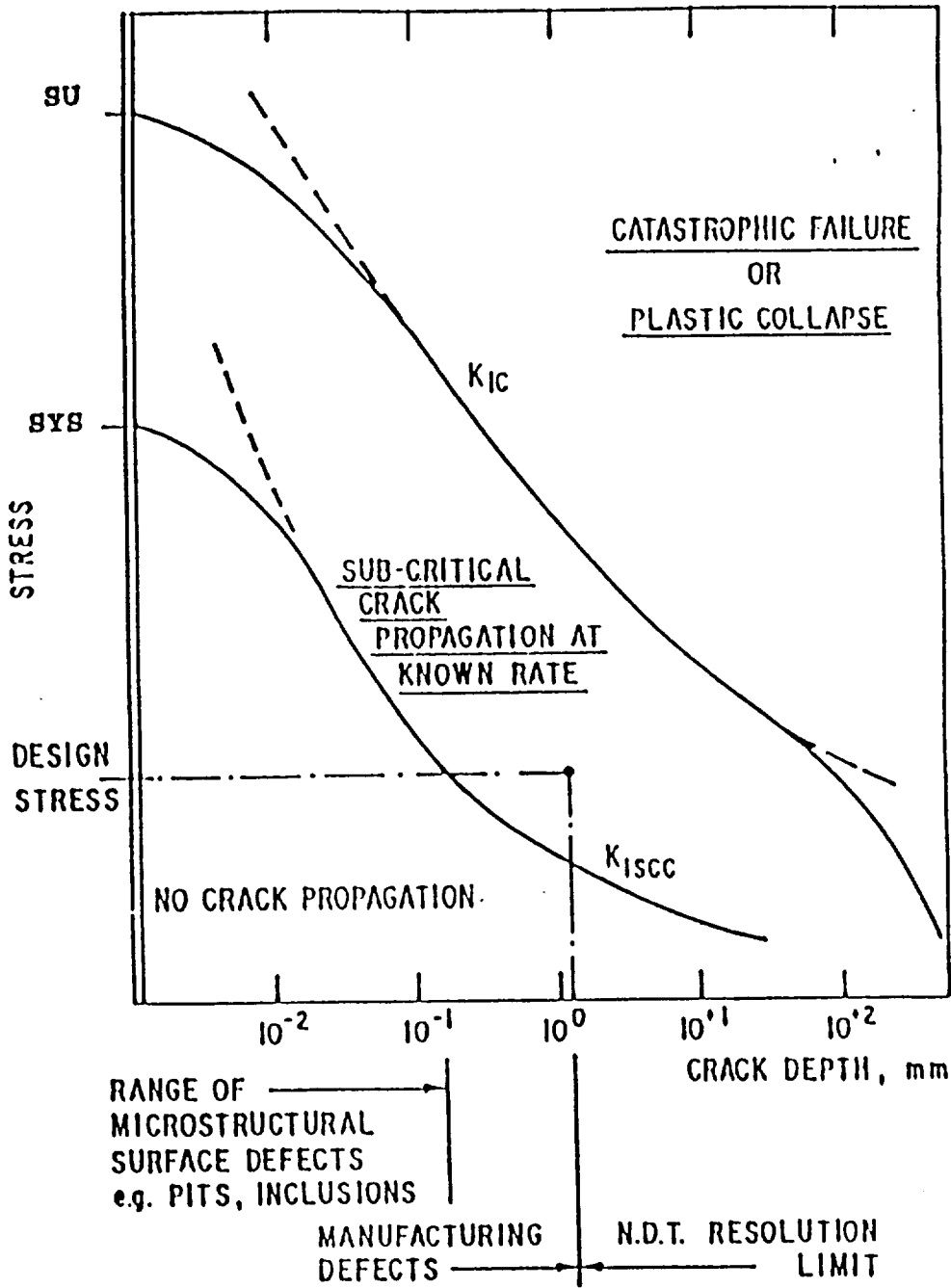


Fig.-6 Tensile Strength vs Crack Length Diagram Illustrating The Regions Of No-crack, Subcritical Crack Growth, And Catastrophic Failure.

Whether or not a component will crack from the time of commissioning will depend on the tensile stress-crack size combination. If the initial stress intensity associated with this combination is greater than K_{ISCC} , its lifetime or inspection periodicity may be determined by integration of the appropriate crack-velocity vs stress intensity curve.

4.0 CONCLUDING REMARKS

Constant strain rate techniques provide a rapid laboratory method to determine the susceptibility of metals to stress corrosion cracking. Absolute results are obtained because failure occurs either in a ductile manner or prematurely by a brittle mode when stress corrosion cracking is present. The absence of stress corrosion cracking assures that the material can be used safely under the specific condition examined. If stress corrosion cracking is observed, judgement is necessary, because the material still can provide sufficient useful life in service. The data obtained from those tests beside being useful for materials selection on a sound engineering basis provides also for a better understanding of the stress corrosion cracking mechanism.

5.0 RECOMMENDATIONS

Due to the short time available and some unexpected delays part of the testing program scheduled for the summer could not be finished. So it is suggested that this study be continued and expanded. The areas where expansion is possible include;

- 1- Testing at different temperatures,
- 2- Testing at other heat treatment conditions, and
- 3- Testing in real conditions (exposure to real atmosphere under static load).

A similar set of experiments can be also done using other types of Precipitation Hardenable steels.

6.0 APPENDICES

6.1 APPENDIX A

Manufacturer certification of properties and quality of
15-5 PH steel.



ARMCO STAINLESS & ALLOY PRODUCTS CERTIFICATION
A DIVISION OF ARMCO INC.

DATE: **JAN 15 1993**

FAVOR:

1 of 2

5123823	CUST. NO.	BILL TO	2110	286	SHIP TO
ANDERSON SHUMAKER CO		ANDERSON SHUMAKER CO			
824 S. CENTRAL AVE		824 S. CENTRAL AVE			
CHICAGO IL 60644		CHICAGO IL 60644			
PRODUCT NUMBER	TEST LOT ID	SALES ORDER NO.	PURCHASE ORDER NO.		
32280-18157	4083-7-B1262A01	23689	738		

SHIP LOT ID: 4083-7-B1262A01

MET LOT NO.: B1262A

15-5 FH SEMI FINISH BILLET STNLS XM-12 VAC CE FQ OA FOR FORGE

AMS 2300G * AMS 2315C EX ONE TEST SAMPLE 16" OK * AMS 5659G TYPE 1 EX
COND A HDNS WVD * ASME SA705 89ED TYPE XM-12 COND H900T ANAL & MECH PROP
CPBLTY ONLY * ASTM A564-89 TYPE XM-12 COND H900T ANAL & MECH PROP CPBLTY
ONLY * ASTM A705-89 TYPE XM-12 COND H900T ANAL & MECH PROP CPBLTY ONLY *
MIL-I-45208A * MIL-STD-2154 (SUPERCEDES MIL-I-8950) TYPE II (CONTACT) PER
ARMCO UTP 7 REV 4 *

LINE: 1	SIZE: 0.0000" RCS	SHAPE: 16' 0" TO 20' 0" MIN/MAX	WEIGHT: 4,420#
---------	-------------------	---------------------------------	----------------

TYPE ANAL: REGULAR	PARENT HEAT NO: 626233		
C: 0.050 MN: 0.540	F: 0.020 S: 0.004	SI: 0.380 CR: 14.91	
NI: 4.57 MO: 0.080	CU: 3.580 CB: 0.310	SN: 0.009 W: 0.003	
CO: 0.040 AL: 0.004	N: 0.034 TI: 0.004	Y: 0.000 TA: 0.008	
B: 0.00130	MULT: 5.	MULT X C: 0.250	
CB+TA LOWLIM: 0.250	CB+TA: 0.318		
TYPE ANAL: PRODUCT ANALYSIS	HEAT NO: 4083-7 *		
C: 0.047 MN: 0.350	F: 0.024 S: 0.004	SI: 0.390 CR: 14.86	
NI: 4.70 MO: 0.070	CU: 3.430 CB: 0.360	CO: 0.043 N: 0.024	
TA: 0.008	MULT X C: 0.235	CB+TA: 0.368	
MULT: 5.	HEAT NO: 4083-7 *		
TYPE ANAL: TOP	F: 0.024 S: 0.004	SI: 0.390 CR: 14.86	
C: 0.047 MN: 0.350	CU: 3.430 CB: 0.360	CO: 0.043 N: 0.024	
NI: 4.70 MO: 0.070	MULT X C: 0.235	CB+TA: 0.368	
TA: 0.008	HTL: 2-920199		
MULT: 5.	B HRDNS (P/F): OK	TEMP: 1300.	
AS SHIP TEST: OVERAGE =	MINS: 0.	QUENCH: AIR	
B HRDNS: 293.			
HOURS: 2.	B HRDNS (P/F): OK	TEMP: 900.	
CAPABILITY TEST: H900T	MINS: 0.	QUENCH: AIR	
B HRDNS: 401.	UTS KSI: 198.7	.2% YLD STR KSI: 180.9	
HOURS: 1.	% RED OF AREA: 20.60		
OFFSET %: 0.20			
% EL 2" OR EQ: 0.0			

PHYSICAL:

FERRITE PCT: <1	FERRITE (P/F): OK	MACRO 604 CL 1: A
MACRO 604 CL 2: A	MACRO 604 CL 3: A	MACRO 604 CL 4: A
MACRO 604 (P/F): OK	MFLUX FREQUENCY: 0.00	MFLUX SEVERITY: 0.00
MAGNAFLUX (P/F): OK	UT: OK	

METALLURGICAL
RELEASE *[Signature]* 1/20/93

ARMCO STAINLESS & ALLOY PRODUCTS CERTIFICATION
 DIVISION OF ARMCO INC.

DATE
 JAN 15 1993

PAGE NO.

2 of 2

5123823	CUST. NO.	BILL TO	2110	286	SHIP TO
ANDERSON SHUMAKER CO 824 S. CENTRAL AVE CHICAGO IL 60644		ANDERSON SHUMAKER CO 824 S. CENTRAL AVE CHICAGO IL 60644			
PRODUCT NUMBER	TEST LOT ID	SALES ORDER NO.	PURCHASE ORDER NO.		
32280-18157	4083-7-B1262A01	23689	735		

MATERIAL OVERAGED 1300 DEG F 2 HRS @TEMP 12 HRS T.T. AIR COOLED
 Material Ultrasonic tested per MIL-STD-2154 (SIZES 0"-10" INCL - CL A,
 OVER 10" - CL B) TYPE II (supercedes MIL-I-8950) per ARMCO UTP 7 REV 4
 (SIZES 0"-10" INCL - CL A, SIZES OVER 10" - CL B) (CONTACT) and was found
 Heat treated at TIME, TEMP and QUENCH indicated above.

Material electric furnace melted, AOD refined, VAC ARC remelted.
 Melted and manufactured in the U.S.A.

No welding or weld repair performed on this material.

ASAF certifies conformance to NRC 10CFR PT 21, and 10CFR PT 50 APP B.

This material was manufactured and tested in accordance with the noted specifications, and is in conformance with those specification requirements

ASAF certifies that this material is manufactured free from mercury,
 uranium, alpha source and low melting metal or alloy contamination.

This material was produced in accordance with the Quality Assurance Program, Quality Assurance Manual, Issue 5 Rev 0 dtd 6/1/92.

All testing procedures were conducted in accordance with the latest ASTM standards or applicable specifications.

The recording of false, fictitious or fraudulent statements or entries on this document may be punished as a felony under federal statutes including Federal Law, Title 18, Chapter 47. This certified test report has been delivered to a consignee of material purchased from ASAF. To avoid the possibility of its misuse on the redelivery of this report to a third party it must be recertified by and under the name of such consignee.

The chemical analyses and physical or mechanical test report are correct as contained in the records of the Corporation.

E. Famularo

E. FAMULARO, CERTIFICATION CLERK

METALLURGICAL
 RELEASE

Jan 15/1993

MSi Metallurgical Services Inc.
 1201 S. Ninth Avenue
 Maywood, Illinois 60153
 708-343-3444

Anderson Shumaker Company
 824 S. Central Ave.
 Chicago, IL 60644

Report No.: 9274-1
 Date: 2-5-93
 Order No.: 7401

Attention: Mr. Steven Tribble

SUBJECT

Tensile and hardness testing of one (1) test bar identified as 15-5 PHVAC, Condition H900, Heat #4083-7, Job #10553, AMS 5659.

TEST RESULTS*

Tensile Testing

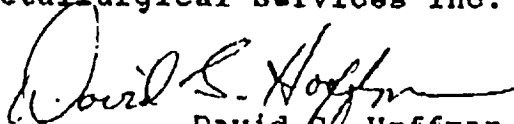
Tensile Strength, psi	199,600
Yield Strength, psi (.2% offset)	178,900
% Elongation in 2"	17
% Reduction of Area	61

Hardness Testing

Hardness, HB	401
--------------	-----

* Testing performed in accordance with ASTM E8-91 and E10-84.

Respectfully Submitted,
 Metallurgical Services Inc.


 David G. Hoffman
 Senior Metallurgical Engineer

DGH/mj

MSi Metallurgical Services Inc.

1201 S. Ninth Avenue
Maywood, Illinois 60153
708-343-3444

Anderson Shumaker Company
824 S. Central Ave.
Chicago, IL 60644

Report No.: 9274-2
Date: 2-5-93
Order No.: 7401

Attention: Mr. Steven Tribble

SUBJECT

Free ferrite examination of one (1) test bar identified as 15-5 PHVAC, Condition H900, Job #10553, Heat #4083-7, AMS 5659.

TEST RESULTS*

Free Ferrite Examination

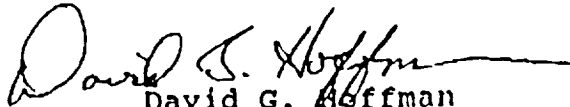
One (1) 15-5 PHVAC stainless steel sample was examined for percentage of free ferrite in accordance with Aerospace Material Specification (AMS) 2315A. The sample was sectioned perpendicular to the direction of rolling as identified in paragraph 3.1.6. The sample was metallographically prepared in accordance with ASTM E3-80 and examined in the etched condition at a magnification of 250X.

The microstructure was rated for percentage of free ferrite in accordance with the occupied squares method as outlined in paragraph 3.2.1. The worst field determined by metallographic examination was photographed at 250X and rated using a transparent grid overlay. A total of fifteen squares were used in the free ferrite calculation. The percentage of free ferrite was determined in accordance with paragraph 3.2.1.1. The results are as follows:

Percentage Free Ferrite

.92

Respectfully Submitted,
Metallurgical Services Inc.


David G. Hoffman
Senior Metallurgical Engineer

DGH/mj

7.0 REFERENCES

- 1-Kundalgurki, S.G.- "Effect Of Strain Rate And Temperature On The Stress Corrosion Cracking Tendency Of Engineering Alloys In Acidified Water", M.S. Thesis, Florida Atlantic University, Boca Raton, Fla., Dic.1992.
- 2-Ford, F.P.- "Stress Corrosion Cracking", Report No. 80CRD141, General Electric, Technical Information Series, Class I, Schenectady, N.Y., June 1980.
- 3-Swann, P.R.- "Stress Corrosion Failure", Scientific American, Feb.1966, Pag.121-135.
- 4-Parkins, R.N.- "Stress Corrosion Spectrum", Br. Corrosion J., Vol.7, Jan.1972, Pag.15-28.
- 5-Clarke Jr, W.C.-"Applications And Properties Of 15-5PH Steel" Source Book On Stainless Steel, A.S.M., Metals Park, Ohio, 1976.
- 6-Jaske, C.E.; Payer, J.H.; Balint, V.S.-"Corrosion Fatigue Of Metals In Marine Environments", MCIC 81-42, Batelle's Columbus Labs., July 1981.
- 7-Campbell, J.E.; Gerberich, W.W.; Underwood, J.H.-Editors- "Application Of Fracture Mechanics For Selection Of Metallic Structural Materials", A.S.M., Metals Park, Ohio, 1982.
- 8-Barsom, J.M.; Rolfe, S.T.-"Fracture & Fatigue Control In Structures - Applications Of Fracture Mechanics", 2nd Ed., Prentice Hall Inc., 1987.
- 9-Craig, H.L.; Editor- "Stress Corrosion - New approaches" A.S.T.M.-STP 610; Philadelphia, Pa., 1976.
- 10-Brown, B.F.; Beachem, C.D.-"A Study Of The Stress Factor In Corrosion Cracking By Use Of The Precracked Cantilever Beam Specimen", Corrosion, 5, 1965.

1994018772

N94-23245

1993 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA

75-25

197201

p. 30

INTERPRETATION OF DISCREPANCIES IN MASS SPECTROSCOPY DATA
OBTAINED FROM DIFFERENT EXPERIMENTAL CONFIGURATIONS

PREPARED BY:	Dr. John M. Russell
ACADEMIC RANK:	Associate Professor
UNIVERSITY AND DEPARTMENT:	Florida Institute of Technology Department of Mechanical and Aerospace Engineering
NASA/KSC	
DIVISION:	Electronic Systems
BRANCH:	Ground Support Equipment
NASA COLLEAGUE:	Frederick W. Adams Carolyn A. McCrary
DATE:	August 20, 1993
CONTRACT NUMBER:	University of Central Florida NASA-NGT-60002 Supplement: 11

ACKNOWLEDGEMENTS

I am indebted to my KSC colleagues RICK ADAMS and CAROLYN MCCRARY (DL ESS-24) for suggesting the problem addressed herein, for providing work space, for introducing me to other cognizant personnel, for arranging Temporary Area Access passes, and for continued interest, encouragement, and council during my summer stay.

I am indebted to Mr. LARRY LINGVAY and other personnel of the Hazardous Gas Detection Laboratory for furnishing me with a Macintosh SE/30 computer, for instructing me on the use of the Alcatel Helium Leak Detector, for carrying out experiments on the determination of the Permeation Ratio of that device, and for helping me to address numerous practical issues where my own expertise is wanting.

I am indebted to Mr. KEVIN KELLENBERGER (Main Propulsion Branch, Vehicle Engineering TV FSD-11) for calling my attention to a training course on the use of the Leybold Hereaus Mass Spectrometer Leak Detector, for obtaining permission from the course instructor for me to attend, for escorting me up to the port access hatch to the engine compartment on the shuttle Endeavor (where I was allowed to peer inside and observe technicians conducting probe-type leak tests), for helping me to obtain copies of *Operation and Maintenance Instruction OMI V1000.005* "[Main Propulsion System] Leak and Functional Test. (LPS) (LH2 Feed System)" and *Operations and Maintenance Requirements and Specifications Document "Orbiter OMRSD main propulsion system"* Volume 41, File III. I am further indebted to him for many useful conversations regarding the subject matter of my summer work.

I am indebted to Mr. RUSTIN VAN DYKE of Lockheed Training (LSO-155), for his able instruction in the course on the Leybold-Hereaus machine, for furnishing a copy of the Rockwell Specification on "Proof pressure and leak detection", for helping me to track down the personnel at the OPF shops who validate the Leybold Hereaus machine, and for many useful discussions on the general subject of this report.

I am indebted to Messrs. KELVIN R. POLK and J. PERRY GODFREY of the Mass Spectrometer Laboratory (LSO-417) for showing me their facilities in the Orbiter Processing Facility High Bay 3, for demonstrating how they determine the Permeation Ratio for the machine, for furnishing small pieces of reject hardware, for helping me to identify the Operation and Maintenance Instruction document that governs their routine tasks, and for a very useful conversation regarding the use of the Leybold Hereaus machine.

This report owes much to a document titled "Report of KSC-STG GHe Leak Check Independent Assessment" by DENNIS PETERSON, dated May 24, 1991. The PETERSON report identifies a 1968 Boeing study as the source of the so-called *Apollo factor* (for estimating the ratio of actual leak rate to the leak rate indicated by a leak detector equipped with a hand held sniffer probe). The PETERSON report points out that probe-type sampling systems with internally branched conduits have now replaced earlier ones in which the conduits were unbranched. These changes in instrumentation have induced changes in the conventions for the interpretation of meter readings during leakage inspections, which may, in turn, result in overestimation or underestimation of the rate of leakage from a test object. The PETERSON report argues that the possibility of underestimation is real.

I am indebted to the staff of the KSC Library for helping me to locate several obscure technical documents dating from the sixties.

Finally, I am indebted to Dr. E. RAYMON HOSLER and Mrs. KARI L. STILES for their able admission of the details of this year's Faculty Summer Fellow program.

ABSTRACT

Many helium-mass-spectrometer leak detectors at KSC employ sampling system that feature hand held sniffer probes. Authors of general leakage-testing literature recommend sniffer probes for leak location but not for quantitative leakage measurement. Their use in the latter application at KSC involves assumptions that may be subtle. The purpose of the research effort reported herein was to establish the significance of indicated leak rates displayed by sniffer-probe-equipped leak detectors and to determine whether the use of alternative hardware or testing procedures may reduce the uncertainty of leakage measurements made with them. The report classifies probe-type sampling systems for helium leak detectors according to their internal plumbing (direct or branched), presents a basic analysis of the fluid dynamics in the sampling system in the branched-conduit case, describes the usual test method for measuring the *internal* supply-to-sample flowrate ratio (*a.k.a* permeation ratio), and describes a concept for a sponge-tipped probe whose *external* supply-to-sample flowrate ratio promises to be lower than that of a simple-ended probe. One conclusion is that the main source of uncertainty in the use of probe-type sampling systems for leakage measurement is uncertainty in the *external* supply-to-sample flowrate ratio (In contrast, the present method for measuring the *internal* supply-to-sample flowrate ratio is quantitative and satisfactory). The implication is that probes of lower external supply-to-sample flowrate ratio must be developed before this uncertainty may be reduced significantly.

SUMMARY

As was stated in the Abstract above, the purpose of the research effort reported herein was to determine the significance of numbers displayed by helium-mass-spectrometer leak detectors equipped with hand held sniffer probes (which are normally of value in leak location but not in the measurement of leak rates) and to determine whether modifications of hardware or procedures may enable one to employ such devices for quantitative leakage measurement.

The report classifies probe-type sampling systems for helium leak detectors according to their internal plumbing, which may be direct or branched. In a direct conduit sampling system, gas flows from the probe tip directly to the analysis cell of the mass spectrometer and the user interface for the spectrometer displays information from which one may deduce the rate of transport of helium through that cell (or may, indeed, display that information directly). In a branched conduit sampling system, gas flows from the probe tip to a place (a mass separator cell) where the flow branches, a portion entering the analysis cell of the mass spectrometer via helium-permeable window and the rest entering the low pressure port of a vacuum pump. In a branched-conduit sampling system, the rate of transport of helium through the probe tip is typically a large multiple of the rate of transport of helium through the analysis cell. This large multiple (the *internal supply to sample flowrate ratio* or *permeation ratio*) is special to each particular branched conduit sampling system. The pressure in the mass separator cell is small relative to pressure in the atmosphere but large relative to the pressure in the analysis cell of the spectrometer. Under such circumstances, the fluid in the conduit is a continuum and obeys the ordinary equations of fluid dynamics.

The report then presents a basic analysis of the fluid dynamics in the sampling system in the branched-conduit case. The equations of motion of a gas admit a simple analytical solution that holds when $Re \ll 1$, where $R = \rho \bar{w} d / \mu$ is the REYNOLDS number based on mass density, ρ , and viscosity, μ , of the sampled gas, tube diameter, d , and cross sectionally averaged fluid speed, \bar{w} , and $\epsilon = d/\ell$ is a slenderness factor based on the tube diameter, d , and the tube length, ℓ .

The report then describes a standard test method for measuring the internal supply-to-sample flowrate ratio of a branched conduit sampling system. The probe is first exposed to atmospheric air and the rate of transport of gas into the probe tip is measured with a flowmeter. The probe is then fed gas of known helium concentration from a bottle containing a calibrated gas mixture. A valve between the calibrated gas bottle and the probe is adjusted until the flow rate of calibrated gas matches the flow rate of atmospheric air determined earlier. The rate of transport of helium into the probe tip is then the rate of transport of the ambient gas mixture times the helium concentration. Knowing the helium transport into the probe tip and the helium transport through the analysis cell (which may be deduced from information displayed by the user interface), the ratio of the former to the latter must equal the internal supply to sample flowrate ratio. The report illustrates the application of this test method in the case of an Alcatel leak detector, where the internal supply to sample flowrate ratio was found to be approximately 1100.

The report then describes a concept for a sponge-tipped probe to be used in conjunction with a branched conduit sampling system. The concept is meant to apply in the case when the rate of transport of sampled gas through the branched conduit is large in comparison with the rate of leakage of helium one is trying to measure. The aim is to lower the *external* supply to sample flowrate ratio to a value well below that of a simple open-ended probe. One places the sponge-tipped probe in contact with a suspect leak (much the way one would place the brush head of a vacuum cleaner over something one wants to vacuum up). The assumptions are meant to ensure that the rate of outward transport of helium by diffusion is small in comparison with the rate of inward transport by convection.

The main conclusion is that leak detectors equipped with sniffer probes are capable of determining the rate of transport of helium into the probe (even if the sampling system is internally branched), but that quantitative measurement of the leakage from a suspect leak site will become feasible only after probes of higher external supply-to-sample flowrate ratio are developed. The topics discussed in the report have implications in regard to the acceptability criteria of leakage from different parts of the shuttle orbiter and the report also discusses some of these implications.

TABLE OF CONTENTS

<u>Section</u>	<u>Title</u>	<u>Page</u>
	ACKNOWLEDGEMENTS	i
	ABSTRACT	ii
	SUMMARY	iii
	LIST OF TABLES	v
I	INTRODUCTION	1-1
1.1	Inspection for leaks. Documents. Motivation for the present project	1-1
1.2	On the expression of leak rates: dimensions	1-1
1.3	The analysis cell of a mass-spectrometer leak detector	1-3
1.4	Testing methods employing enclosures and sniffer probes	1-4
1.4.1	Enclosure testing for leak measurement	1-4
1.4.2	Sniffer-probe testing for leak location	1-5
1.4.3	Sniffer-probe testing for leak measurement	1-6
1.4.4	Difficulties in the interpretation of acceptability criteria	1-10
II	CAPILLARY FLOW	2-1
2.1	The equations of motion of a viscous gas on cylindrical coordinates	2-1
2.2	Simplifications	2-2
2.2.1	Neglect of the gravity terms	2-2
2.2.2	The neglect of time dependence	2-2
2.2.3	The assumption of axisymmetric flow without swirl	2-2
2.2.4	Use of the assumption that the capillary is slender	2-3
2.2.5	Use of an assumption involving the capillary Reynolds number	2-4
2.3	Deductions from the simplified equations	2-5
2.3.1	Solution for the axial velocity distribution	2-5
2.3.2	The rate of transport of fluid volume	2-5
2.3.3	Computation of the throughput $p\dot{\Delta}$. Use of an isothermal assumption. Streamwise variation of the pressure	2-5 2-6
2.3.4	The time for a fluid particle to travel the length of a capillary	2-6
2.4	Dimensions of the Quick Test conduit. A one-capillary model for predicting $p\dot{\Delta}$ and $t_1 - t_0$. Comparisons with observation	2-8
III	A CONCEPT FOR A MICROENCLOSURE, OR SPONGE-TIPPED PROBE	3-1
IV	RECOMMENDATIONS AND CONCLUSION	4-1
	REFERENCES	5-1

LIST OF TABLES

<u>Number</u>	<u>Title</u>	<u>Page</u>
1-1	Sniffer-probe evaluation results after SENEAR & BLAIR 1968 (Ref. 4, p 44)	1-7
1-2	External supply-to-sample flowrate ratio deduced from data in SENEAR & BLAIR 1968 (Ref.4, p 44).....	1-8
1-3	Leak-Point Matrix from Rockwell Specification MF0001-003 (Ref. 3, Appendix I, pp 13-15)	1-9

SECTION ONE

INTRODUCTION

1.1 INSPECTION FOR LEAKS. DOCUMENTS. MOTIVATION FOR THE PRESENT PROJECT

Uncontrolled leaks pose hazards to personnel and equipment and compromise the ability of spacecraft and ground support hardware to carry out their missions. Accordingly, shuttle operations personnel perform a variety of leak tests on the shuttle orbiter between flights.

Technicians involved in space shuttle operations at KSC perform routine tasks according to checklists. Such checklists belong to a generic class of documents called *Operations and Maintenance Instruction*, abbreviated OMI. Ref. 1, titled "MPS leak and functional test (LPS). (LH2 feed system)," is an example of an OMI governing leak checking of the Main Propulsion System, specifically the system for the transport of liquid hydrogen. Every OMI is subject to rules laid down in one or another *Operations and Maintenance Requirements and Specification Document*, (OMRSD). Ref. 2, titled "Orbiter OMRSD main propulsion system," (Rockwell International) is an example of an OMRSD, in this case the OMRSD that governs the procedures described in Ref. 1. A third category of governing document is a *Specification*. Ref. 3, titled "Proof pressure and leak detection—Aerospace plumbing systems and assemblies," (Rockwell International) is an example of a Specification. The Rockwell Specification (Ref. 3) expresses acceptance criteria for leak rates at various suspect leak sites on the orbiter and states what kinds of instrumentation are permissible to verify that the acceptance criteria are satisfied. This Specification allows for leakage measurement by means of hand-held sniffer probes. The OMI that pertains to leak testing of liquid hydrogen systems (Ref. 1) specifically names the mass spectrometer leak detector as the instrument to be used to perform various inspections.

A report titled "Report of KSC-STG GHe Leak Check Independent Assessment" by DENNIS PETERSON, dated May 24, 1991 (Ref. 4) points out that probe-type sampling systems with internally branched conduits have now replaced earlier ones in which the conduits were unbranched. These changes in instrumentation have induced changes in the conventions for the interpretation of meter readings during leakage inspections, which may, in turn, result in overestimation or underestimation of the rate of leakage from a test object. The PETERSON report argues that the possibility of underestimation is real.

The research reported herein was motivated by the need to address some of the questions raised in the PETERSON report (Ref. 4) and to consider possible changes in the design of probes and sampling systems that may eliminate some of the uncertainties of existing systems.

1.2 ON THE EXPRESSION OF LEAK RATES: DIMENSIONS

Before one can attribute differences in leak rate indications to differences in experimental configuration one must discuss what exactly one means by the term *leak rate*. A value of the leak rate is a quantity with certain dimensions and some discussion of these dimensions is in order. This subsection supplies preliminary information of this nature.

In fluid dynamics, a *streamline* is a mathematical curve in space that is tangent at all of its points to the local velocity direction. Given a velocity field and a geometric point in that field, one can always draw a streamline that passes through that point. If, alternatively, one is given a geometric *figure* in the form of a closed loop, all of whose points lie in a velocity field, then one can draw a streamline through every point on the perimeter of that loop. The family of streamlines so generated describes a tube-shaped surface and is called a *streamtube*. Since a streamtube is a surface generated by streamlines and streamlines are tangent to the velocity direction at all of their points, one concludes that a streamtube is also tangent to the local velocity direction at all of its points. It follows that a *streamtube is impermeable to fluid motion*—fluid never passes across it. Now every leak belongs to a particular streamtube. One may therefore apply all of the ideas

that fluid dynamicists have developed for general streamtubes to the particular case when the streamtube represents a leak.

Let S denote a geometric figure in the form of a surface that cuts across a streamtube at one particular station along it; let \mathbf{u} be the local fluid velocity vector; and let $d\mathbf{A}$ be the differential directed area element normal to S and directed away from the downstream side of it (the magnitude $|d\mathbf{A}|$ represents the geometric area of the element described by the vector $d\mathbf{A}$). Then the *volumetric rate of transport* of fluid across the infinitesimal directed area element $d\mathbf{A}$ per unit time is $\mathbf{u} \cdot d\mathbf{A}$. Let $\dot{\Delta}$ be the *resultant* volumetric rate of transport of fluid across the *whole* surface S . One may compute $\dot{\Delta}$ by summing the separate contributions from all of the infinitesimal subsurfaces of S and express the result in the form of the following double integral

$$\dot{\Delta} \equiv \iint_S (\mathbf{u} \cdot d\mathbf{A}). \quad (1.1)$$

The parameter $\dot{\Delta}$ has the dimensions of volume per unit time. One may, for example, express a value of $\dot{\Delta}$ in $(\text{cm})^3/\text{sec}$.

A *gram-mole* of a substance whose molecular mass is M is an amount of that substance whose mass in grams is numerically equal to M . Thus, a gram-mole of helium, whose molecular mass is 4 is an amount of helium whose mass equals 4 grams. Alternatively, one may say that M for helium has the value 4 grams/(gram-mole). Let \dot{n} be the *molar rate of transport* of fluid through a streamtube across a cross sectional surface S (one may, for example, express a value of $\dot{\Delta}$ in gm - moles/sec). Let ρ^* be the molar density of a gas (one may, for example, express a value of ρ^* in $\text{gm}/(\text{cm})^3$). The number of moles of gas per unit time that passes across the infinitesimal subsurface $d\mathbf{A}$ equals the product of the number of moles per unit volume, ρ^* , times the volume transport rate, $\mathbf{u} \cdot d\mathbf{A}$ across that subsurface, and therefore has the value $\rho^*(\mathbf{u} \cdot d\mathbf{A})$. One may compute \dot{n} by summing the separate contributions from all of the infinitesimal subsurfaces of S and express the result in the form of the following double integral

$$\dot{n} = \iint_S \rho^*(\mathbf{u} \cdot d\mathbf{A}). \quad (1.2)$$

The parameter \dot{n} has the dimensions of amount of substance per unit time. If the flow is steady, the equation of conservation of mass implies that the value of \dot{n} is the same for all stations along the streamtube. In this respect, the value of \dot{n} is an intrinsic property of the streamtube as a whole. This streamwise uniformity of \dot{n} is to be contrasted with the possible streamwise nonuniformity of the volumetric rate of transport $\dot{\Delta}$. Variations of $\dot{\Delta}$ may occur as a result of compressibility of the gas.

According to the ideal gas law $p = \rho^* \mathcal{R} T$, in which p is the pressure; T is the absolute temperature; and \mathcal{R} is the *universal gas constant*, whose value is

$$\mathcal{R} = 82.08 \frac{\text{atm} \cdot (\text{cm})^3}{\text{gm-mole} \cdot ^\circ\text{K}}. \quad (1.3)$$

The word "universal" in the phrase "universal gas constant" is meant to imply that \mathcal{R} is the same for all gases (at least those that can be modeled as ideal gases). If one multiplies (1.2) by $\mathcal{R} T$ and restricts attention to the case when T is uniform over S , one obtains

$$\dot{n} \mathcal{R} T = \mathcal{R} T \iint_S \rho^*(\mathbf{u} \cdot d\mathbf{A}) = \iint_S \rho^* \mathcal{R} T (\mathbf{u} \cdot d\mathbf{A}) = \iint_S p (\mathbf{u} \cdot d\mathbf{A}), \quad (1.4)$$

whose rightmost member has the dimensions of pressure times volume flow rate.

If the pressure is uniform over the cross section S , one may move p across the integral sign in (1.4) and obtain

$$\dot{n} \mathcal{R} T = p \iint_S (\mathbf{u} \cdot d\mathbf{A}) = p \dot{\Delta}, \quad (1.5)$$

in which the last equality follows from (1.1).

If the absolute temperature T does not differ greatly from laboratory conditions (say $T = 295^\circ\text{K} = 71.33^\circ\text{F}$), then one may carry out approximate calculations by assuming that \mathcal{RT} has the room-temperature value

$$\mathcal{RT} = 24,214 \frac{\text{atm}\cdot(\text{cm})^3}{\text{gm-mole}} \quad (\text{for } T = 295^\circ\text{K} = 71.33^\circ\text{F}). \quad (1.6)$$

If one accepts the value of \mathcal{RT} given in (1.6) then equation (1.5) furnishes a formula from which one may calculate \dot{n} from the value of the quantity $p\dot{\Delta}$ or $p\dot{\Delta}$ from the value of the quantity \dot{n} . In this respect, the two parameters \dot{n} and $p\dot{\Delta}$, though they have different dimensions, both measure the flux of gas through a streamtube and are both intrinsic properties of that streamtube.

The relation (1.5) between the molar rate of transport \dot{n} , the parameter $p\dot{\Delta}$, and the value of the conversion factor \mathcal{RT} given by (1.6) is independent of the choice of gas. If, for example, the absolute temperature has the value $T = 295^\circ\text{K}$ and the parameter $p\dot{\Delta}$ has the value $p\dot{\Delta} = 1 \times 10^{-6}$ atm cc/sec, then these data furnish enough information to calculate the molar rate of transport \dot{n} from (1.5) (in the present example, \dot{n} would have the value 4.13×10^{-11} gm-moles/sec). This value of \dot{n} is the same whether the gas is helium, hydrogen, nitrogen, or any other gas that can be modeled as an ideal gas. The observation that the parameter $p\dot{\Delta}$ has meaning as a measure of the rate of leakage and that this meaning is not specific to the choice of gas (or, for that matter, on the barometric pressure of the environment into which the gas discharges) helps account for why some leak detectors display leakage rate in units such as millibar-liters/sec or atm-(cm)³/sec (both of which are suitable units for the parameter $p\dot{\Delta}$).

The parameter $p\dot{\Delta}$ also has a interpretation in terms of energy concepts. Thus, when gas leaks from the inside to the outside of a test object via an orifice, some of the gas in the exterior region that had initially abutted the leak site is *displaced* by the leaked gas. Thus, as fluid leaks out of a test object, it performs work on the surrounding gas to displace it. The parameter $p\dot{\Delta}$ thus represents the rate at which the leaking fluid performs such displacement work per unit time. In mechanics, the rate at which work is performed per unit time is *power*, and the standard unit for power is the Watt (W). One concludes that $p\dot{\Delta}$ may be expressed in Watts. Now one atmosphere (atm) is a unit of pressure equal to 1.01325×10^5 Newtons per square meter ($\text{N}/(\text{m}^2)$), one Watt (W) is defined to be 1 N·m/sec, and one centimeter (cm) equals 10^{-2} meters (m). It follows that

$$1 \frac{\text{atm}\cdot(\text{cm})^3}{\text{sec}} = \frac{[1.01325 \times 10^5 \text{ N}/(\text{m}^2)](1 \times 10^{-2} \text{ m})^3}{\text{sec}} = 0.101325 \frac{\text{N}\cdot\text{m}}{\text{sec}} = 0.101325 \text{ W}. \quad (1.7)$$

Motivated by the foregoing considerations, I will refer to the parameter $p\dot{\Delta}$ as the *leakage rate, flowrate, or throughput* in this report. In this context, one may regard the familiar unit sccs (for *standard cubic centimeter per second*) as a shorthand for the more precise notation atm-(cm)³/sec or atm-cc/sec. Equation (1.7) shows that 1 sccs is equivalent to 101.325 milliwatts (exactly) regardless of the choice of gas.

1.3 THE ANALYSIS CELL OF A MASS-SPECTROMETER LEAK DETECTOR

A 1968 NASA Contractor Report titled *Leakage testing handbook*, prepared by J. W. MARR on behalf of General Electric Company for Jet Propulsion Laboratory, Pasadena, CA (Ref. 5), furnishes a general reference on the subject of leak testing. MARR's description of the action of the analysis cell of a mass spectrometer (*op. cit.*, p 150) suffices for the purposes of the present report:

In the sector field mass spectrometer... gas molecules entering the evacuated body of the instrument are ionized by a bombarding beam of electrons. The resultant positive ions are accelerated by the influence of a high voltage into a magnetic field whose direction is perpendicular to the plane of the mean path of the ions. Under the influence of this magnetic field, ions of different masses travel on arcs of different radii. Ions of a particular mass pass through a slit and fall on

a collecting electrode; from it they leak to ground through a high-value resistor. The ion current through the resistor is amplified and read on a meter. By varying the magnetic field or accelerating electric field, ions of any chosen mass may be collected.

One may surmise from the foregoing description of the principle of operation of the mass spectrometer that

$$i = e\eta\dot{n}, \quad (1.8)$$

in which: i is the electric current that passes through the receiver electrode belonging to a particular charge-to-mass ratio of molecule; e is the charge of one electron; \dot{n} is the total flux of molecules of a given charge-to-mass ratio through the instrument; η is the ionization efficiency, i.e. the ratio of the flux of current-bearing molecules to \dot{n} . The following calculation illustrates a self-consistency check of the units of measurement of the various factors in (1.8):

$$\frac{\text{Coulombs}}{\text{time}} = \frac{\text{Coulombs}}{\text{ion}} \cdot \frac{\text{ions}}{\text{molecule}} \cdot \frac{\text{molecules}}{\text{time}} \quad (1.9)$$

The basic presumption in the calibration of a mass-spectrometer leak detector is that $e\eta$ is constant, so $i \propto \dot{n}$. One also assumes that \dot{n} is proportional to the partial pressure of the helium on the upstream side of the inlet leak to the instrument.

To standardize a mass spectrometer leak detector, one must determine the instrument response to a known leak rate. A common method for producing a known leak rate is with the aid of a *calibrated leak*. According to MARR (*op. cit.*, p 112)

Calibrated leaks, leaks which deliver gas at a known rate, are sold by a number of vendors. . . Calibrated leaks may be divided into two distinct categories—the reservoir leaks, those which contain their own gas supply, and the non-reservoir leaks to which tracer gas is added during test

Among the kinds of reservoir leaks are *permeation leaks*. According to MARR (*op. cit.*, p 114), such leaks employ the principle [of] gas diffusion through a thin wall. . . Tracer gas [is] permeated from the area of high concentration to air or vacuum on the other side at a rate governed by the permeability of the thin membrane.

Tracer gas is the gas whose flux is measured by the analysis cell. Here, and elsewhere, I will assume that the tracer gas is helium. A reservoir-type leak based on the permeation principle may resemble a small bottle with a hand operated valve. The valve attaches to standard size vacuum couplings. In use, the operator attaches the standard leak to an input port on the leak detector and opens the valve. In some leak detectors, the meter reading on the electronic cabinet is expressed in the same units as the leak rate parameter $p\Delta$ (say atm·(cm)³/sec). The label on the reservoir-type calibrated leak indicates its nominal leakage. By comparing the instrument reading with the indication on the label, one adjusts the gain in the electronic interface circuitry so as to force agreement between the two numbers. On other leak detectors the meter reading may be in the same units as the electrical current in the ion beam, say Amperes. For such a device, the ratio of the instrument reading (in Amperes) to the known leak rate for the calibrated leak is interpreted as an *instrument sensitivity factor*. Regardless of how the instrument is standardized, the meter reading of a helium mass spectrometer leak detector is proportional to $(p\Delta)_{\text{He, Analysis cell}}$, namely *the rate of leakage of helium into the analysis cell of the mass spectrometer*.

1.4 TESTING METHODS EMPLOYING ENCLOSURES AND SNIFFER PROBES

1.4.1 ENCLOSURE TESTING FOR LEAK MEASUREMENT. According to one scheme for the quantitative measurement of leak rate, a fluid line or other test object is filled with gaseous helium under pressure. One measures the rate of discharge of helium from a leak by first enclosing the test object in some kind

of enclosure (e.g. a hood, boot, or bag), then flushing the enclosure with a suitable purge gas (the rate of transport of purge gas being known), and finally measuring the helium concentration of the effluent.

Let $(p\dot{\Delta})_{\text{purge}}$ be the rate of transport of purge gas into the enclosure. Let $(p\dot{\Delta})_{\text{He, leak}}$ be the rate of leakate of helium into the enclosure and let $(p\dot{\Delta})_{\text{eff}}$ be the rate of discharge of gas mixture in the effluent. If the absolute temperature of the gas is effectively uniform throughout the enclosure and the flow is stationary in time, the net inflow must equal the net outflow and one deduces that

$$(p\dot{\Delta})_{\text{purge}} + (p\dot{\Delta})_{\text{He, leak}} = (p\dot{\Delta})_{\text{eff}} \quad (1.10)$$

Let c denote the concentration of helium in the effluent. Then

$$c = \frac{(p\dot{\Delta})_{\text{He, leak}}}{(p\dot{\Delta})_{\text{eff}}} \quad (1.11)$$

If one eliminates $(p\dot{\Delta})_{\text{eff}}$ from (1.11) by means of (1.10), one obtains

$$c = \frac{(p\dot{\Delta})_{\text{He, leak}}}{(p\dot{\Delta})_{\text{purge}} + (p\dot{\Delta})_{\text{He, leak}}} \quad (1.12)$$

One may manipulate this equation algebraically to obtain a formula for $(p\dot{\Delta})_{\text{He, leak}}$, namely

$$(p\dot{\Delta})_{\text{He, leak}} = \left(\frac{c}{1-c} \right) (p\dot{\Delta})_{\text{purge}} \quad (1.13)$$

If one incorporates a flow metering device into the purge gas delivery system then $(p\dot{\Delta})_{\text{purge}}$ is known. If one then measures c by sampling the gas in the effluent and determining its helium concentration by mass spectrometry, then formula (1.13) furnishes a formula for the determination of $(p\dot{\Delta})_{\text{He, leak}}$. If c is very small compared to one, then one may approximate the expression (1.13) by $(p\dot{\Delta})_{\text{He, leak}} \approx c(p\dot{\Delta})_{\text{purge}}$. A 1987 AIAA Paper titled "Development of the Helium Signature Test for orbiter main propulsion system revalidation between flights" by VINCENT J. BILARDO, FRANCISCO IZQUIERDO, & R. SMITH (Ref. 6) describes an enclosure-type method now in use for the measurement of the aggregate helium leakage resulting from all individual leaks in the aft section of the shuttle orbiter. The Rockwell Specification on Proof pressure and leak detection (Ref. 3) specifies that the term *actual leak rate* refers to the leak rate measured by one or another of such enclosure-type tests (*op. cit.*, §6.2 p 13)

1.4.2 SNIFFER-PROBE TESTING FOR LEAK LOCATION. An enclosure-type test of a large or complicated test object may establish that a leak exists somewhere but may give no indication of how many leaks there are or where they are situated. A convenient method for locating such leaks involves the sampling of the air in the vicinity of a suspect leak site with a hand held *sniffer probe*. The sampled gas is then transported to a helium-mass-spectrometer leak detector where the rate of transport of helium into the leak detector is measured and displayed by the instrumentation. In such applications of sniffer probes, only a fraction of the helium that escapes from the leak in the original test object is collected by the sniffer probe, (the uncollected helium escapes to the surrounding atmosphere where it is dispersed). Here, and elsewhere, I will use the term *external supply-to-sample flowrate ratio* for the fraction

$$\frac{(p\dot{\Delta})_{\text{He, test object}}}{(p\dot{\Delta})_{\text{He, probe}}} \quad (1.14)$$

i.e. the ratio of the rate of release of helium from a leak in a test object to the rate of transport of helium into the sniffer probe. Methods for standardizing helium-mass-spectrometer leak detectors enable one to measure the rate of transport of helium into the sniffer probe quantitatively, but the measurement of the external supply-to-sample flowrate ratio is far more problematic. Notwithstanding the uncertainties in the external supply-to-sample flowrate ratio, the meter deflection that results when a sniffer probe is passed over a leak provides useful information, namely information on where the leak is located.

1.4.3 SNIFFER-PROBE TESTING FOR LEAK MEASUREMENT. A 1968 report titled "Measurement and correlation of helium and fluid leak rates" prepared by A. E. SENEAR & A. W. BLAIR on behalf of the Boeing Company for what was then the Manned Spacecraft Center, Houston (Ref. 7) reports results of extensive tests on artificially produced leaks. The aim of these tests was to separate the dependence of leakage rates on leak configuration (e.g. boundary shape and endpoint conditions for the pressure) from the dependence of leakage rates on the specific choice of leaked fluid (e.g. helium gas *versus* hypergolic liquids). In the course of this study, the Boeing investigators performed some tests that amount to a measurement of the external supply-to-sample flowrate ratio. Thus, SENEAR & BLAIR write (*op. cit.*, page 42)

The helium sniffer probe attached to a CEC helium leak detector with 5 feet of Tygon tubing and fitted with a 1/8" ID Tygon tip (per Grumman Specification LSP-1A-50121A) was used to test an AN scribed fitting leaking at a rate of 1×10^{-5} cc He/sec at 380 psig applied helium pressure. The absolute sensitivity of the leak detector at the end of the probe had previously been determined to be 3.4×10^{-10} atm cc/(sec division). When tested in the open laboratory and when holding the probe perpendicular to the axis of the leaking fitting a very large variation was obtained dependent upon precise positioning, with over an order of magnitude difference recorded 60° either side of the top of the fitting. Variations of about 50% at any one position were also noted. The maximum indicated leak rate detected during this experiment (265 divisions) was 9×10^{-8} cc/sec, by an operator with prior knowledge of the location of the leak.

Experiments were also attempted with discouraging results, using the fingers to hood the leak in order to improve sensitivity. These readings proved quite unstable and non-reproducible although an increase in sensitivity level was produced. Checking in the open air, if relatively still, seems to be the best, most reproducible method.

Table 8 [whose content is reproduced here as Table 1-1 on the next page] gives the results obtained by four different operators on four different leaks utilizing both the hooded-by-fingers and open still air probing techniques. These operators had no prior knowledge of the location of the leaks. Examination of the data will reveal that the sniffer probe technique is not a quantitative measurement. It is, to be sure, quite useful as a screening technique for locating relatively large leaks ($> 1 \times 10^{-6}$ cc/sec of helium).

SENEAR & BLAIR state the conclusion they draw from the foregoing results as follows (*op. cit.*, p 70)

The sniffer probe is of qualitative value only. The observed leak rates are approximately three orders of magnitude lower than the actual leak rate, and variations of 1300% are found between different operators, and different probe attitudes.

The ratio of the number in the leftmost column of Table 1-1 to the number in the rightmost column of the same row represents the ratio of actual to apparent leak rate. This ratio is typically on the order of one thousand. In these experiments, the conduit that transferred helium from the sniffer probe to the analyzer cell of the mass spectrometer was direct, *i.e.* all of the helium that entered the probe tip was counted in the computation of the apparent leak rate. One concludes that *in the Boeing data of Ref. 7 the ratio of actual to apparent leak rate (whose value was approximately one thousand) was equal to the external supply-to-sample flowrate ratio of the sampling system that the investigators employed.* Table 1-2 (on the second page after this one) lists the values of the external supply-to-sample flowrate ratio corresponding to the data in Table 1-1 (specifically, the data corresponding to the label "w/o hood").

The Rockwell Specification titled "Proof pressure and leak detection—Aerospace plumbing systems and assemblies" (Ref. 3) states what methods of measurement are acceptable for the determination of leakage rates. Method A of the Rockwell Specification, titled "Probe Technique" (*op. cit.* §3.5.1), indicates that sniffer probes may be used for leak measurement. An appendix to the Specification titled "Leak Point Matrix" defines the acceptable leakages from different parts of the shuttle orbiter. Table 1-3 (on the third page after this one) reproduces the content of the Rockwell Leak Point Matrix

The Leak Point Matrix has two columns of numbers headed "Actual He Leakage" and "Indicated He Leakage". For any one part of the orbiter (represented by a particular row in the Leak Point Matrix), the

Table 1-1. Sniffer-probe evaluation results
after Senear & Blair 1968 (Ref. 4, p 44)

Known Leak Rate (cc/sec)	Operator	Leak Indication (divisions)		L.D. Sensitivity [(cc/sec)/div]	Apparent Leak Rate (cc/sec)	
		w/hood	w/o hood		w/hood	w/o hood
1×10^{-5}	A	212	18	3.62×10^{-10}	7.67×10^{-8}	6.51×10^{-9}
	B	695	14	3.73×10^{-10}	2.59×10^{-7}	5.22×10^{-9}
	C	219	17	3.73×10^{-10}	8.16×10^{-8}	6.35×10^{-9}
	D	265	16	3.40×10^{-10}	9.01×10^{-8}	5.45×10^{-9}
5×10^{-6}	A	84	17	3.62×10^{-10}	3.04×10^{-8}	6.18×10^{-9}
	B	339	8	3.73×10^{-10}	1.26×10^{-7}	2.98×10^{-9}
	C	69	6	3.73×10^{-10}	2.67×10^{-8}	2.24×10^{-9}
	D	180	14	3.73×10^{-10}	6.70×10^{-8}	5.22×10^{-9}
1×10^{-6}	A	16	2	3.62×10^{-10}	5.8×10^{-9}	7.5×10^{-10}
	B	62	2	3.73×10^{-10}	2.31×10^{-8}	7.45×10^{-10}
	C	10	1	3.73×10^{-10}	3.73×10^{-9}	3.73×10^{-10}
	D	13	3	3.73×10^{-10}	4.85×10^{-9}	1.12×10^{-9}
5×10^{-7}	A	10	2	3.62×10^{-10}	3.62×10^{-9}	7.5×10^{-10}
	B	30	—†	3.73×10^{-10}	1.12×10^{-8}	—†
	C	3	—†	3.73×10^{-10}	1.12×10^{-9}	—†
	D	14‡	3‡	3.73×10^{-10}	$5.22 \times 10^{-9}‡$	$1.1 \times 10^{-9}‡$
† Could not possibly identify the number of divisions (if any) ‡ Questionable						

numbers in these two columns differ by a factor of exactly one thousand. The report titled "Report of KSC-STS GHe leak check Independent Assessment" by DENNIS PETERSON of KSC (Ref. 4) suggests that the Leak Point Matrix in the Rockwell Specification was developed at a time when sniffer-probe equipped leak detectors used by shuttle inspection personnel at KSC featured direct-conduit sampling systems similar to the ones used to generate the data in Tables 1-1 and 1-2. According to PETERSON, the factor one-thousand between the Actual and Indicated leak rates in the Rockwell Leak Point Matrix was developed during the Apollo era. Some personnel with whom I have spoken refer to this three-order-of-magnitude adjustment as the *Apollo factor*.

Now the sniffer-probe equipped leak detectors that feature direct conduit sampling systems of the sort used to generate the data in Table 1-1 have some disadvantages. Thus, the gas in conduit is at nearly the same pressure as that of the analyzer cell of the mass spectrometer. At such a low pressure, the mean free path of the gas molecules is large compared to the tube diameter. In order to transport a measurable amount of gas, the sniffer-probe hose must be thick and unwieldy. The response time is also slow when the hose is longer than a few meters.

Such practical considerations have motivated personnel in charge of leak testing the shuttle orbiter to consider alternative configurations of test equipment. In the late eighties, they abandoned sniffer-probe equipped leak detectors featuring direct-conduit sampling systems in favor of sniffer probe equipped leak detectors featuring systems of *branched conduit* type. In a branched conduit sampling system, the sniffed gas is allowed to flow at a significant fraction of atmospheric pressure through much of its passage from the probe end to the leak detector. In particular, the mean free path of the sampled gas is small compared to the inside diameter of the conduit (even when the latter is on the order of a millimeter or so) and the flow moves in accordance with the laws of viscous fluid motion. For such conduits, the response time may be no longer than eight seconds even when the conduit is one hundred feet long. To satisfy the condition of very

Table 1-2. External supply-to-sample flowrate ratio deduced from data in Senear & Blair 1968 (Ref. 4, p 44)

Known Leak Rate (cc/sec)	Operator	$\frac{(p\dot{\Delta})_{\text{He, test object}}}{(p\dot{\Delta})_{\text{He, probe}}}$
1×10^{-5}	A	1536
	B	1916
	C	1575
	D	1835
5×10^{-6}	A	809
	B	1678
	C	2232
	D	958
1×10^{-6}	A	1333
	B	1342
	C	2681
	D	893
5×10^{-7}	A	668
	B	—†
	C	—†
	D	455‡
† Could not possibly identify the number of divisions (if any)		
‡ Questionable		

low pressure in the analysis cell of the mass spectrometer, the fluid at the downstream end of the conduit branches. One branch is separated from the analysis cell by a helium permeable window (which is capable of reacting the difference in gas pressure between the analysis cell and the downstream end of the conduit) and the other branch is connected to the low pressure port of a vacuum pump (whose exhaust is discarded at atmospheric pressure). In the mean time, the only flow rate that the mass spectrometer measures is the rate of flow of helium into the analysis cell. Thus the branch of the the gas stream that passes through the the probe tip but does not pass through the analysis cell is *not* measured by the mass spectrometer (and so does not contribute to the leak rate value displayed by the instrumentation on the electronic cabinet of the leak detector).

Let $(p\dot{\Delta})_{\text{He, probe}}$ be the flux of helium into the probe and let $(p\dot{\Delta})_{\text{He, Analysis cell}}$ be the flux of helium into the analysis cell. Here, and elsewhere, I will refer to the ratio

$$\frac{(p\dot{\Delta})_{\text{He, probe}}}{(p\dot{\Delta})_{\text{He, Analysis cell}}} \quad (1.15)$$

as the internal supply-to-sample flowrate ratio. There is an Operation and Maintenance Instruction (OMI) that governs the procedures for validating the Leybold Hereaus Ultratest F mass spectrometer leak detector equipped with the optional Quick Test attachment (a sampling system of the internally branched type). The procedure described in this OMI (Ref. 8) involves the determination of a parameter (*the permeation ratio*), whose meaning is equivalent to the internal supply-to-sample flowrate ratio defined by (1.15) (see "cylinder samples check", sequence 02-010, p 26, *op. cit.*). Messrs. KELVIN R. POLK and J. PERRY GODFREY of the Mass Spectrometer Laboratory of the Orbiter Processing Facility Shops (LSO-417) perform the procedures

Table 1-3 Leak-Point Matrix from Rockwell Specification MF0001-003
(Ref. 3, Appendix I, pp 13-15)

Test Point	Actual He Leakage (scc/sec)	MSHP† Indicated‡ He Leakage (scc/sec)
<u>(OMS/RCS)</u>		
<u>[Orbital Maneuvering System/Reaction Control System]</u>		
All welded/brazed joints	1×10^{-4}	1×10^{-7}
All Mechanical joints	1×10^{-4}	1×10^{-7}
All He and propellant test points (capped)	1×10^{-4}	1×10^{-7}
All All vent ports (capped)	1×10^{-4}	1×10^{-7}
All He and propellant fill ports (capped)	1×10^{-4}	1×10^{-7}
<u>MPS [Main Propulsion System]</u>		
All cryogenic mechanical, welded and brazed joints	1×10^{-4}	1×10^{-7}
Tank door seals, flange seals, and swaged joints	5×10^{-3}	5×10^{-6}
Engine interface seal	1×10^{-4}	1×10^{-7}
Pneumatic joints less than 1/2 inch in diameter	1×10^{-4}	1×10^{-7}
Pneumatic joints greater than or equal to 1/2 inch in diameter	1×10^{-3}	1×10^{-6}
<u>EPS [Electrical Power System]—Cryo system</u>		
All welded and brazed joints in H ₂ and O ₂ systems	1×10^{-4}	1×10^{-7}
Quick Disconnects uncapped H ₂ and O ₂ system	NA	NA
Quick Disconnects capped H ₂ and O ₂ system	NA	NA
Valve packs in H ₂ system	1×10^{-4}	1×10^{-7}
Relief valve in H ₂ system	NA	NA
Cryo tanks metallurgical joints H ₂ system	1×10^{-4}	1×10^{-7}
All dynatube connections or other mechanical joints	1×10^{-4}	1×10^{-7}
<u>EPG [Electrical Power Generator]-Fuel cell system</u>		
All welded and brazed joints	1×10^{-4}	1×10^{-7}
All dynatube connections or other mechanical joints	1×10^{-4}	1×10^{-7}
<u>ECLSS [Environmental Control and Life Support System]</u>		
All welded and brazed or solder joints	1×10^{-4}	1×10^{-7}
All Quick Disconnects in ECLSS (capped or mated) (unless otherwise specified)	1×10^{-4}	1×10^{-7}
All dynatube connections or other mechanical joints	1×10^{-4}	1×10^{-7}
All test ports and instrumentation ports (capped)	1×10^{-4}	1×10^{-7}
<u>Window/Windshield Cavity Conditioning System</u>		
All welded/brazed joints	1×10^{-4}	1×10^{-7}
All mechanical joints	1×10^{-4}	1×10^{-7}
Quick Disconnects capped	1×10^{-4}	1×10^{-7}
Bulkhead seals	5×10^{-3}	5×10^{-6}
All test ports and instrumentation ports (capped)	1×10^{-4}	1×10^{-7}
† MSHP=Mass Spectrometer Hand Probe ‡ All leakage indication are "Single Point" leaks. These acceptance criteria, when using the Ultratest M, M2, or F system shall be corrected with the appropriate factors ... to provide the maximum indicated reading which may be accepted.		

described in this OMI. They inform me that the value of the permeation ratio for Quick Test systems they test is normally between 300 and 400, and that the value 400 is typical.

In the case when the sampling system is of branched conduit type, one may relate the actual-to-indicated leak rate ratio to the external and internal supply-to-sample flowrate ratios as follows

$$\frac{\underbrace{(p\dot{\Delta})_{\text{He, test object}}}_{\text{actual-to-indicated leakrate ratio}}}{\underbrace{(p\dot{\Delta})_{\text{He, analysis cell}}}_{\text{external supply-to-sample flowrate ratio}}} = \frac{(p\dot{\Delta})_{\text{He, test object}}}{(p\dot{\Delta})_{\text{He, probe}}} \times \frac{(p\dot{\Delta})_{\text{He, probe}}}{\underbrace{(p\dot{\Delta})_{\text{He, analysis cell}}}_{\text{internal supply-to-sample flowrate ratio}}} \quad (1.16)$$

1.4.4 DIFFICULTIES IN THE INTERPRETATION OF ACCEPTABILITY CRITERIA. With the aid of equation (1.16), one may bring into sharper focus the difficulties of interpreting the Rockwell Leak Point Matrix requirements (Table 1-3) in the case when the sniffer probe on the leak detector is used in conjunction with a sampling system of branched conduit type. To fix ideas, suppose that a probe operator inspects a test object with a Leybold-Hereaus Ultratest F leak detector equipped with a Quick Test sampling system. Now the Quick Test sampling system is of branched conduit type (as opposed to direct conduit type), so one may not assume automatically that the external supply-to-sample flowrate ratio is on the order 1,000 (as in the Boeing data reported in Table 1-2). If one makes this assumption anyway, and assumes further that the internal supply-to-sample flowrate ratio is 400, then a numerical computation of the actual-to-indicated leak rate ratio is, according to (1.16),

$$\underbrace{400,000}_{\text{actual-to-indicated leakrate ratio}} = \underbrace{1,000}_{\text{external supply-to-sample flowrate ratio}} \times \underbrace{400}_{\text{internal supply-to-sample flowrate ratio}} \quad (1.17)$$

Of course, the value 400,000 for the actual-to-indicated leak rate ratio is much larger than the value 1,000 indicated in the Rockwell Leak Point Matrix (Table 1-3). Two hypotheses for this apparent inconsistency come to mind:

Hypothesis I. *The Quick Test probe acts on the air in its neighborhood as a mass sink. Transport of helium toward the probe in consequence of this sink flow dominates over transport away from the leak source in consequence of molecular diffusion and lateral transport in consequence of ambient crossflow.* The value 1,000 for the external supply-to-sample flowrate ratio, which was assumed in the derivation of (1.17), was based on the 1968 Boeing data (Table 1-2). These data were taken with sampling systems of direct type. But the Quick Test system is of branched type. The sampling system used to collect the data in the Boeing study more nearly resembles the Leybold Hereaus *Standard Probe*, a sampling system of the direct conduit type which may be fit to the Leybold Hereaus Leak Detector as alternative to the Quick Test attachment. The rate at which this direct conduit system draws in atmospheric air is about 1×10^{-3} atm·(cm)³/sec. For such a device, one may conjecture that the mechanism that dominates the transfer of helium from the leak in the test object to the probe tip is molecular diffusion. Now molecular diffusion transfers helium outward from the source in all directions and the element of solid angle subtended by the probe tip is small compared with the total solid angle of a celestial sphere centered on the source. The effectiveness of such diffusive transfer is low at best and may be hampered further by ambient crossflow at the test site. Some such ambient crossflow will always be present owing to deliberate circulation of environmental air in the test area and buoyant convection from temperature nonuniformities of different objects in the test area (to name only two causes). One may surmise that the actual transfer of helium from the leak source to the tip of a Standard Probe with a draw rate of 1×10^{-3} atm·(cm)³/sec is an inefficient process which is easily

disrupted by unavoidable conditions in the test environment. The high value of the external supply-to-sample ratio observed by the Boeing investigators of Ref. 7 (*i.e.* the value 1,000) is, according to this argument, a reflection of the inefficiency of this diffusive transfer process.

Consider now the Quick Test system. The rate at which the Quick Test system draws in atmospheric air is about $1.5 \text{ atm}\cdot(\text{cm})^3/\text{sec}$. The flowrate of air through the Quick Test probe is therefore about 1,500 times the corresponding flowrate through the Standard Probe. The flowrate through the Quick Test probe is, moreover, much larger than the actual helium leakage of any marginally acceptable leak indicated in the Rockwell Leak Point Matrix (Table 1-3). One might conjecture that the Quick Test, unlike the Standard probe, exerts an action on the leaked gas analogous to the action of a vacuum cleaner. In particular, the flow of air exterior to the probe tip generated by the sucking of fluid into it is what fluid dynamicists call a *point sink*. The transfer of helium outward from the leak source due to molecular diffusion and lateral transfer due to ambient crossflow is counteracted to a greater or lesser extent by the transfer of helium toward the probe tip due to the sink flow. By making the sink flow large enough, one may arrange that the helium transfer in the neighborhood of the probe tip is *dominated* by the sink flow. One may surmise that the actual transfer from the leak source to the tip of a Quick Test Probe with a draw rate of $1.5 \text{ atm}\cdot(\text{cm})^3/\text{sec}$ is a more efficiently process than the transfer in the case of the Standard Probe (whose draw rate is three orders of magnitude smaller). A value of the external supply-to-sample ratio that is conspicuously lower than the one observed by the Boeing investigators of Ref. 7 would reflect the greater efficiency of the Quick Test system in transferring helium. If, for example, one assumes an external supply-to-sample ratio of 2.5 then, under Hypothesis I, a computation corresponding to (1.16) would be of the form

$$\underbrace{1,000}_{\text{actual-to-indicated leakrate ratio}} = \underbrace{2.5}_{\text{external supply-to-sample flowrate ratio}} \times \underbrace{400}_{\text{internal supply-to-sample flowrate ratio}} \quad (1.18)$$

Hypothesis II. When probe-type sampling systems are used, the term *Indicated Leak Rate* is ambiguous unless preceded by an appropriate adjective. Thus, *Uncorrected Indicated Leak Rate* refers to the flowrate of helium through the analysis cell and *Corrected Indicated Leak Rate* in the Rockwell Leak Point Matrix (Table 1-3) are to be interpreted as *Corrected Indicated Leak Rates*. Under Hypothesis II, one may relate the Corrected Indicated Leak Rate to the Uncorrected Indicated Leak rate by an identity of the form

$$\underbrace{(p\dot{\Delta})_{\text{He, probe}}}_{\text{Corrected Indicated Leak Rate}} = \underbrace{\frac{(p\dot{\Delta})_{\text{He, probe}}}{(p\dot{\Delta})_{\text{He, analysis cell}}}}_{\text{internal supply-to-sample flowrate ratio}} \times \underbrace{(p\dot{\Delta})_{\text{He, analysis cell}}}_{\text{Uncorrected Indicated Leak Rate}} \quad (1.19)$$

The Correction Factor, as used under Hypothesis II, equals the internal supply-to-sample flowrate ratio (*a.k.a.* the permeation ratio). One may compute the ratio of actual leak rate to corrected indicated leak rate as follows:

$$\frac{(p\dot{\Delta})_{\text{He, test object}}}{(p\dot{\Delta})_{\text{He, probe}}} = \frac{(p\dot{\Delta})_{\text{He, test object}}}{\frac{(p\dot{\Delta})_{\text{He, probe}}}{(p\dot{\Delta})_{\text{He, analysis cell}}} \times (p\dot{\Delta})_{\text{He, analysis cell}}}, \quad (1.20)$$

or

$$\underbrace{\frac{(p\dot{\Delta})_{\text{He, test object}}}{(p\dot{\Delta})_{\text{He, probe}}}}_{\text{Ratio of Actual Leak Rate to Corrected Indicated Leak Rate}} = \underbrace{\left(\frac{(p\dot{\Delta})_{\text{He, test object}}}{(p\dot{\Delta})_{\text{He, analysis cell}}} \right)}_{\text{end-to-end supply-to-sample flowrate ratio}} \div \underbrace{\left(\frac{(p\dot{\Delta})_{\text{He, probe}}}{(p\dot{\Delta})_{\text{He, analysis cell}}} \right)}_{\text{internal supply-to-sample flowrate ratio}} \quad (1.21)$$

which is just a rearrangement of equation (1.16) above. If one assumes an end-to-end supply-to-sample flowrate ratio of 400,000 and an internal supply-to-sample flowrate ratio of 400, then a computation corresponding to (1.21) is

$$\underbrace{1,000}_{\substack{\text{Ratio of Actual Leak} \\ \text{Rate to Corrected} \\ \text{Indicated Leak Rate}}} = \underbrace{400,000}_{\substack{\text{end-to-end} \\ \text{supply-to-sample} \\ \text{flowrate ratio}}} \div \underbrace{400}_{\substack{\text{internal} \\ \text{supply-to-sample} \\ \text{flowrate ratio}}}, \quad (1.22)$$

which is a rearrangement of (1.17).

Now Hypothesis II is the more conservative of the two hypothesis in that it is more likely to overestimate the rate of leakage from a test object. The information in the Rockwell Specification (Ref. 3) is not clear as to which of the above two hypotheses it requires. The relevant passages in the text of the specification read as follows:

3.2.1 ...The Leybold-Heraeus Ultratest M, M2, or F with Quick Test attachment shall have operating instructions attached that shall include the necessary ratio and entrainment information so that its indicated readings may be converted to actual leak figures.

3.2.2 The permeability ratio shall be determined using the volume method for obtaining the indicated leak rate and shall be 600 or less before the leak detector, with hose and probe, may be used (Ref. Calibration Procedure 344-2027 [titled "Calibration/Service Procedure for the Leybold Heraeus Mass Spectrometer Leak Detector"]).

One should also recall one of the footnotes to the Leak Point Matrix reproduced in Table 1-3 above. That footnote reads

All leakage indications are "Single Point" leaks. These acceptance criteria, when using the Ultratest M, M2, and F system shall be corrected with the appropriate factors (ref. 3.2.1) to provide the maximum indicated reading which may be accepted.

The quoted passage from §3.2.1 seems to require that the "operating instructions attached" to the Quick Test contain not merely the Permeation Ratio (*i.e.* internal supply-to-sample flowrate ratio), but rather the full end-to-end supply-to-sample flowrate ratio (including the poorly understood factor due to the *external* supply-to-sample flowrate ratio). In this respect, the Rockwell Specification seems to ask a great deal of the personnel responsible for preparing the machine for use. I have read "Calibration Procedure 344-2027" cited in §3.2.2 above* and am satisfied that the procedure for determining the Permeation Ratio is equivalent to the one now in use (*cf.* Operations and Maintenance Instruction OMI V6H31, Revision A-1 (Ref. 8)).

The footnote to the Leak Point Matrix just quoted above seems to suggest the use of some kind of correction factor to relate Indicated Leak Rate to Corrected Indicated Leak Rate. Parts of the specification that mention corrections for helium background effects also allude to such a correction factor. Thus, in §3.5.1.1 one reads

3.5.1.1 ...Normally/ideally the background should be low enough that the allowable leak (to be determined) is at least several times the MDL [Minimum Detectable Leak]. However, the MDL versus the allowable leak is the limiting factor. The background reading shall not exceed the Indicated Acceptance Criteria in Appendix I [the Leak Point Matrix, *cf.* Table 1-3 above], or for the Ultratest M, M21, or F the corrected Indicated Acceptance Criteria (see note 2 of Appendix I).

"Note 2" cited above reads as follows:

* I am indebted to Dr. WILLIAM VAN DUSEN of Rockwell (ZK36) for furnishing me with a copy of this document. I gather that Dr. VAN DUSEN, in turn, obtained his copy from Mr. RUSTIN VAN DYKE of Lockheed Training (LSO-155). Special thanks are therefore due to Mr. VAN DYKE.

- 2) When using the Ametek or Ultratest M, M2, or F, if the indicated leak rate reading does not exceed the corrected indicated acceptance criteria, a background reading need not be taken.

Here again is an allusion to a correction factor. The last excerpt that provides any clue about indicated leak rate is a definition in §3.5.1.4 (h), *viz*

- (h) The indicated leakage rate shall not exceed the leakage criteria of Appendix I [the Leak Point Matrix of Table 1-3] and shall be calculated using any mathematical tool that will produce the same results as the following formula:

$$\text{indicated leak rate} = \left(\frac{\text{leak detector}}{\text{sensitivity}} \right) \times (\text{total} - \text{background meter readings})$$

(DuPont, or Veeco machines [of direct conduit type]) [or]

$$\text{indicated leak rate} = \text{corrected} - \text{background meter readings}$$

(Ametek and Leybold-Heraeus Machine [of branched conduit type]) only if the total exceeds 1×10^{-7} [no units given in original].

Again, there is an allusion to a correction factor. Lacking clearer information, one is inclined to suppose that the intentions of the Rockwell specification are more accurately reflected by Hypothesis II (which involves the use of a correction factor and in which the external supply-to-sample ratio is assumed to be 1,000) than by Hypothesis I (which involves no correction factor and in which the external supply to-sample flowrate ratio is assumed to be approximately 2.5).

In the mean time, Operations and Maintenance Instruction V1009.005 (Ref. 1) covers the procedures used by probe operators inspecting parts of the hydrogen transport system for the shuttle orbiter. These instructions do not make any use of a correction factor. They appear, therefore, to be more consistent with Hypothesis I. There would seem to be a need to reconcile the requirements of References 1 and 3.

SECTION TWO

CAPILLARY FLOW

2.1 THE EQUATIONS OF MOTION OF A VISCOUS GAS IN CYLINDRICAL COORDINATES

Let $\{\hat{e}_r, \hat{e}_\theta, \hat{e}_z\}$ be a right handed orthogonal triad of unit vectors belonging to a cylindrical coordinate system with radial coordinate r , azimuth angle θ , and axial coordinate z . Let (u_r, u_θ, u_z) be the scalar components of the fluid velocity vector \mathbf{u} belonging to these coordinates. Let ρ be the mass density of the fluid (a quantity with the dimensions mass per unit volume). Then one may write the differential equation of conservation of mass in the form

$$\frac{\partial \rho}{\partial t} + \frac{1}{r} \frac{\partial(r\rho u_r)}{\partial r} + \frac{1}{r} \frac{\partial(\rho u_\theta)}{\partial \theta} + \frac{\partial(\rho u_z)}{\partial z} = 0. \quad (2.1)$$

The differential equation for the rate of change of translational momentum is a vector equation. In a three dimensional space, this vector equation has three scalar components. Let p denote the local instantaneous mechanical pressure in the fluid, and let (g_r, g_θ, g_z) be the scalar components of the local gravitational force per unit mass \mathbf{g} belonging to the cylindrical coordinate system introduced above. Let μ be the coefficient of shear viscosity in the fluid (a quantity with the dimensions of stress divided by strain rate, or, equivalently, force times time divided by area).

If one assumes that there are no volumetric forces exerted on the gas other than gravity and makes the usual assumption that the gas is linearly viscous, then the three scalar components of the momentum equation are as follows. The radial component of the momentum equation is

$$\rho \left(\frac{\partial u_r}{\partial t} + u_r \frac{\partial u_r}{\partial r} + \frac{u_\theta}{r} \frac{\partial u_r}{\partial \theta} + u_z \frac{\partial u_r}{\partial z} - \frac{(u_\theta)^2}{r} \right) = -\frac{\partial p}{\partial r} - \frac{2\mu}{r} \left(\frac{1}{r} \frac{\partial u_\theta}{\partial \theta} + \frac{u_r}{r} - \frac{\text{div } \mathbf{u}}{3} \right) + \frac{1}{r} \frac{\partial}{\partial r} \left[r 2\mu \left(\frac{\partial u_r}{\partial r} - \frac{\text{div } \mathbf{u}}{3} \right) \right] + \frac{1}{r} \frac{\partial}{\partial \theta} \left[\mu \left(\frac{1}{r} \frac{\partial u_r}{\partial \theta} - \frac{u_\theta}{r} + \frac{\partial u_\theta}{\partial r} \right) \right] + \frac{\partial}{\partial z} \left[\mu \left(\frac{\partial u_r}{\partial z} + \frac{\partial u_z}{\partial r} \right) \right] + \rho g_r, \quad (2.2)$$

the azimuthal component of the momentum equation is

$$\rho \left(\frac{\partial u_\theta}{\partial t} + u_r \frac{\partial u_\theta}{\partial r} + \frac{u_\theta}{r} \frac{\partial u_\theta}{\partial \theta} + u_z \frac{\partial u_\theta}{\partial z} + \frac{u_\theta u_r}{r} \right) = -\frac{1}{r} \frac{\partial p}{\partial \theta} + \frac{\mu}{r} \left(\frac{1}{r} \frac{\partial u_r}{\partial \theta} - \frac{u_\theta}{r} + \frac{\partial u_\theta}{\partial r} \right) + \frac{1}{r} \frac{\partial}{\partial r} \left[r \mu \left(\frac{1}{r} \frac{\partial u_r}{\partial \theta} - \frac{u_\theta}{r} + \frac{\partial u_\theta}{\partial r} \right) \right] + \frac{1}{r} \frac{\partial}{\partial \theta} \left[2\mu \left(\frac{1}{r} \frac{\partial u_\theta}{\partial \theta} + \frac{u_r}{r} - \frac{\text{div } \mathbf{u}}{3} \right) \right] + \frac{\partial}{\partial z} \left[\mu \left(\frac{\partial u_\theta}{\partial z} + \frac{1}{r} \frac{\partial u_z}{\partial \theta} \right) \right] + \rho g_\theta, \quad (2.3)$$

and the axial component of the momentum equation is

$$\rho \left(\frac{\partial u_z}{\partial t} + u_r \frac{\partial u_z}{\partial r} + \frac{u_\theta}{r} \frac{\partial u_z}{\partial \theta} + u_z \frac{\partial u_z}{\partial z} \right) = -\frac{\partial p}{\partial z} + \frac{1}{r} \frac{\partial}{\partial r} \left[r \mu \left(\frac{\partial u_r}{\partial z} + \frac{\partial u_z}{\partial r} \right) \right] + \frac{1}{r} \frac{\partial}{\partial \theta} \left[\mu \left(\frac{\partial u_\theta}{\partial z} + \frac{1}{r} \frac{\partial u_z}{\partial \theta} \right) \right] + \frac{\partial}{\partial z} \left[2\mu \left(\frac{\partial u_z}{\partial z} - \frac{\text{div } \mathbf{u}}{3} \right) \right] + \rho g_z, \quad (2.4)$$

in which $\text{div } \mathbf{u}$ is an abbreviation for the divergence of the velocity field, *i.e.*

$$\text{div } \mathbf{u} \equiv \frac{1}{r} \frac{\partial(r u_r)}{\partial r} + \frac{1}{r} \frac{\partial u_\theta}{\partial \theta} + \frac{\partial u_z}{\partial z}. \quad (2.5)$$

2.2 SIMPLIFICATIONS

2.2.1 NEGLECT OF THE GRAVITY TERMS. I will assume, here and elsewhere, that the gas in the capillary is ideal. Thus,

$$p = \rho RT, \quad (2.6)$$

in which the mass specific gas constant R is related to the universal gas constant \mathcal{R} by

$$r = \frac{\mathcal{R}}{M}, \quad (2.7)$$

in which M is the molecular mass. Now the elements of the list

$$\left(\frac{\partial p}{\partial r}, \frac{1}{r} \frac{\partial p}{\partial \theta}, \frac{\partial p}{\partial z} \right) \quad (2.8)$$

are the three scalar components of the pressure gradient vector ∇p and the elements of the list

$$(\rho g_r, \rho g_\theta, \rho g_z) \quad (2.9)$$

are the three scalar components of the gravitational force per unit mass $\rho \mathbf{g}$. In view of (2.6), we have

$$\frac{|\rho \mathbf{g}|}{|\nabla p|} = \frac{p}{RT} \frac{|\mathbf{g}|}{|\nabla p|}. \quad (2.10)$$

If ℓ is the length of the capillary, then an order of magnitude estimate of the quotient $p/|\nabla p|$ is

$$\frac{p}{|\nabla p|} \sim \ell. \quad (2.11)$$

The expression $RT/|\mathbf{g}|$ represents a length which is sometimes called the *scale height of the atmosphere*. To be specific, if an isothermal atmosphere is in static equilibrium under its own weight, then the scale height is the altitude increase within which the pressure decreases by one power of e . For air at $T = 296^\circ\text{K}$, the scale height is about 8.84 kilometers. If one denotes the scale height by H , then (2.10) and (2.11) imply that

$$\frac{|\rho \mathbf{g}|}{|\nabla p|} = \frac{p}{|\nabla p|} \frac{1}{H} \sim \frac{\ell}{H}. \quad (2.12)$$

I will assume, here and elsewhere, that $\ell \ll H$ and, accordingly, neglect those contributions to the momentum equations (2.2)–(2.4) due to $\rho \mathbf{g}$ (cf. the list (2.9)) in comparison with the contributions due to ∇p (cf. the list (2.8)).

2.2.2 THE NEGLECT OF TIME DEPENDENCE. I restrict the present analysis to flows with zero time dependence. Accordingly, all of the terms in the system (2.1)–(2.4) involving the time derivative operator $\partial(\)/\partial t$ will be equated to zero.

2.2.3 THE ASSUMPTION OF AXISYMMETRIC FLOW WITHOUT SWIRL. I restrict the present analysis to flows in which none of the flow variables depend upon the azimuth angle θ and the swirl component of the velocity, u_θ , is identically zero. Under the assumptions listed thus far, the system of equations (2.1)–(2.5) reduces to

$$\frac{1}{r} \frac{\partial(r \rho u_r)}{\partial r} + \frac{\partial(\rho u_z)}{\partial z} = 0. \quad (2.13)$$

$$\rho \left(u_r \frac{\partial u_r}{\partial r} + u_z \frac{\partial u_r}{\partial z} \right) = -\frac{\partial p}{\partial r} - \frac{2\mu}{r} \left(\frac{u_r}{r} - \frac{\text{div } \mathbf{u}}{3} \right) + \frac{1}{r} \frac{\partial}{\partial r} \left[r 2\mu \left(\frac{\partial u_r}{\partial r} - \frac{\text{div } \mathbf{u}}{3} \right) \right] + \frac{\partial}{\partial z} \left[\mu \left(\frac{\partial u_r}{\partial z} + \frac{\partial u_z}{\partial r} \right) \right], \quad (2.14)$$

$$0 = 0, \quad (2.15)$$

$$\rho \left(u_r \frac{\partial u_z}{\partial r} + u_z \frac{\partial u_z}{\partial z} \right) = -\frac{\partial p}{\partial z} + \frac{1}{r} \frac{\partial}{\partial r} \left[r \mu \left(\frac{\partial u_r}{\partial z} + \frac{\partial u_z}{\partial r} \right) \right] + \frac{\partial}{\partial z} \left[2\mu \left(\frac{\partial u_z}{\partial z} - \frac{\text{div } \mathbf{u}}{3} \right) \right], \quad (2.16)$$

and

$$\text{div } \mathbf{u} \equiv \frac{1}{r} \frac{\partial(r u_r)}{\partial r} + \frac{\partial u_z}{\partial z}, \quad (2.17)$$

respectively.

2.2.4 USE OF THE ASSUMPTION THAT THE CAPILLARY IS SLENDER. Let ℓ be the inside diameter of the capillary. If the capillary wall is impermeable, one may solve (2.13) for u_r as follows. If one regards ρ , u_r , & u_z as functions of the argument list (r, z) , then one may write (2.13) in the form

$$\frac{\partial}{\partial r} [r\rho(r, z)u_r(r, z)] = -r \frac{\partial}{\partial z} [\rho(r, z)u_z(r, z)] . \quad (2.18)$$

An equivalent equation written in different notation is

$$\frac{\partial}{\partial \xi} [\xi\rho(\xi, z)u_r(\xi, z)] = -\xi \frac{\partial}{\partial z} [\rho(\xi, z)u_z(\xi, z)] . \quad (2.19)$$

If one integrates (2.19) with respect to ξ from $\xi = d/2$ to $\xi = r$ and applies an impermeable-wall boundary condition $(u_r)_{r=d/2} = 0$, one obtains

$$r\rho(r, z)u_r(r, z) = - \int_{d/2}^r \xi \frac{\partial}{\partial z} [\rho(\xi, z)u_r(\xi, z)] d\xi , \quad (2.20)$$

or

$$u_r(r, z) = \frac{-1}{r\rho(r, z)} \int_{d/2}^r \xi \frac{\partial}{\partial z} [\rho(\xi, z)u_r(\xi, z)] d\xi . \quad (2.21)$$

Let U_r & U_z be velocity scales typical of the velocity components u_r and u_z , respectively. Then (2.21) implies that U_r is related to U_z by the following order-of-magnitude balance

$$U_r \sim U_z \frac{d}{\ell} , \quad (2.22)$$

or

$$U_r \sim U_z \epsilon , \quad (2.23)$$

in which

$$\epsilon \equiv d/\ell \quad (2.24)$$

is a *slenderness parameter* for the capillary. I will assume in what follows that $\epsilon \ll 1$.

Of the various viscous terms in the right member of the axial momentum equation (2.16), one may show that the term

$$\frac{1}{r} \frac{\partial}{\partial r} \left(r\mu \frac{\partial u_z}{\partial r} \right) \quad (2.25)$$

dominates over the others as follows. Let μ_0 be a viscosity scale typical of the viscosity of the gas in the capillary. Then

$$\frac{1}{r} \frac{\partial}{\partial r} \left(r\mu \frac{\partial u_z}{\partial r} \right) \sim \frac{\mu_0 U_z}{d^2} , \quad (2.26)$$

and

$$\frac{1}{r} \frac{\partial}{\partial r} \left(r\mu \frac{\partial u_r}{\partial r} \right) \sim \frac{\mu_0 U_r}{d\ell} \sim \frac{\mu_0 \epsilon U_z}{d\ell} = \frac{\mu_0 \epsilon^2 U_z}{d^2} , \quad (2.27)$$

in which the last equality follows from (2.24). Moreover,

$$\frac{\partial}{\partial z} \left(2\mu \frac{\partial u_z}{\partial z} \right) \sim \frac{\mu_0 U_z}{\ell^2} = \frac{\mu U_z}{d^2} \epsilon^2 . \quad (2.28)$$

Now

$$\text{div } \mathbf{u} = \frac{1}{r} \frac{\partial (ru_r)}{\partial r} + \frac{\partial u_z}{\partial z} \sim \frac{U_r}{d} + \frac{U_z}{\ell} \sim \frac{\epsilon U_z}{d} + \frac{U_z \epsilon}{d} ,$$

or

$$\operatorname{div} \mathbf{u} \sim \frac{U_z \epsilon}{d}. \quad (2.29)$$

Thus,

$$\frac{\partial}{\partial z} \left[2\mu \left(-\frac{\operatorname{div} \mathbf{u}}{3} \right) \right] \sim \frac{\mu_0 U_z \epsilon}{\ell d} = \frac{\mu_0 U_z}{d^2} \epsilon^2. \quad (2.30)$$

The viscous terms (2.27), (2.28), & (2.30) are therefore of the order ϵ^2 times the viscous term (2.26). If one neglects these higher order terms, the axial momentum equation (2.16) reduces to

$$\rho \left(u_r \frac{\partial u_z}{\partial r} + u_z \frac{\partial u_z}{\partial z} \right) = -\frac{\partial p}{\partial z} + \frac{1}{r} \frac{\partial}{\partial r} \left(r \mu \frac{\partial u_z}{\partial r} \right). \quad (2.31)$$

If one carries out estimates of the order-of-magnitude of the various terms in the *radial* momentum equation (2.14), one finds that the largest of the various terms is of the order ϵ times the largest of the viscous terms in the axial equation (2.16). By the same token, the ratio of the inertia terms in the radial equation to the inertia terms in the axial equation is also of the order ϵ . Thus, the terms that balance the radial pressure gradient in (2.14) are of the order ϵ times the terms that balance the axial pressure gradient in (2.16) and one concludes that $\partial p / \partial r$ is of the order ϵ times $\partial p / \partial z$. One is then led to treat p as a function of the axial coordinate z only and to write (2.31) in the form

$$\rho \left(u_r \frac{\partial u_z}{\partial r} + u_z \frac{\partial u_z}{\partial z} \right) = -\frac{dp}{dz} + \frac{1}{r} \frac{\partial}{\partial r} \left(r \mu \frac{\partial u_z}{\partial r} \right). \quad (2.32)$$

2.2.5 USE OF AN ASSUMPTION INVOLVING THE CAPILLARY REYNOLDS NUMBER. Let ρ_0 be a density scale typical of values of ρ in the capillary. One may estimate the terms in the left member of (2.31) as follows

$$\rho u_r \frac{\partial u_z}{\partial r} \sim \rho_0 \frac{U_r U_z}{d} \sim \frac{\rho_0 \epsilon U_z^2}{d}, \quad (2.33)$$

$$\rho u_z \frac{\partial u_z}{\partial z} \sim \rho_0 \frac{U_z^2}{\ell} \sim \rho_0 \frac{U_z^2}{d} \epsilon. \quad (2.34)$$

The ratio of the order of magnitude of the inertia terms in the left member of (2.32) to the order of magnitude of the viscous term in the right member is therefore

$$\frac{\left(\rho_0 \frac{U_z^2}{d} \epsilon \right)}{\left(\frac{\mu_0 U_z}{d^2} \right)} = \left(\frac{\rho_0 U_z d}{\mu_0} \right) \epsilon. \quad (2.35)$$

The expression $\rho_0 U_z d / \mu_0$ is the *Reynolds number* based on the capillary diameter and axial velocity scale. I will assume in what follows that

$$\left(\frac{\rho_0 U_z d}{\mu_0} \right) \epsilon \ll 1, \quad (2.36)$$

which allows one to approximate (2.32) by the simpler equation

$$\frac{dp}{dz} = \frac{1}{r} \frac{\partial}{\partial r} \left(r \mu \frac{\partial u_z}{\partial r} \right). \quad (2.37)$$

2.3 DEDUCTIONS FROM THE SIMPLIFIED EQUATIONS

2.3.1 SOLUTION FOR THE AXIAL VELOCITY DISTRIBUTION. If one assumes that the viscosity μ does not vary significantly across the cross section of the capillary, one may rewrite (2.37) in the form

$$\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial u_z}{\partial r} \right) = \frac{1}{\mu} \frac{dp}{dz}. \quad (2.38)$$

One may write the general solution of the differential equation (2.38) in the form

$$u_z = \frac{r^2}{4} \frac{1}{\mu} \frac{dp}{dz} + A \ln \left(\frac{r}{d/2} \right) + B, \quad (2.39)$$

in which the parameters μ , p , A , and B are all independent of r . If one imposes the condition that the velocity be finite on the axis $r = 0$, one deduces that $A = 0$. If, moreover, one applies the *no-slip boundary condition*

$$(u_z)_{r=d/2} = 0, \quad (2.40)$$

one may compute the value of B . The resulting formula for u_z is

$$u_z = \left[r^2 - \left(\frac{d}{2} \right)^2 \right] \frac{1}{4\mu} \frac{dp}{dz}. \quad (2.41)$$

Equation (2.41), or some equation of equivalent meaning, appears in nearly every textbook on fluid dynamics and is normally associated with the names HAGEN and POISEUILLE. All of the standard derivations of (2.41) with which I am familiar, however, assume that the fluid is incompressible. The present derivation, which makes no incompressibility assumption *per se* generalizes the result (2.41) to gas flow.

If one evaluates (2.47) on the centerline $r = 0$, one obtains

$$(u_z)_{\max} = - \left(\frac{d}{2} \right)^2 \frac{1}{4\mu} \frac{dp}{dz}. \quad (2.42)$$

If one orients the axes so that the positive z -axis points in the direction of the flow, then $(u_z)_{\max} > 0$ by construction. It follows from (2.42) that $dp/dz < 0$, *i.e.* that *pressure decreases in the direction of the flow* (as expected!).

2.3.2 THE RATE OF TRANSPORT OF FLUID VOLUME. In (1.1), let S be a circular cross section (normal to the axis of the capillary), in which $0 \leq r \leq d/2$ and $0 \leq \theta \leq 2\pi$. Let $d\mathbf{A} = r d\theta dr \hat{\mathbf{e}}_z$ and $\mathbf{u} = u_z \hat{\mathbf{e}}_z$. Then $\mathbf{u} \cdot d\mathbf{A} = u_z r d\theta dr$ and (1.1) becomes

$$\dot{\Delta} = \int_{\theta=0}^{2\pi} \int_{r=0}^{d/2} u_z r d\theta dr = \frac{1}{4\mu} \frac{dp}{dz} 2\pi \int_{r=0}^{d/2} \left[r^2 - \left(\frac{d}{2} \right)^2 \right] r dr = - \frac{\pi}{8\mu} \frac{dp}{dz} \left(\frac{d}{2} \right)^4. \quad (2.43)$$

2.3.3 COMPUTATION OF THE THROUGHPUT $\rho \dot{\Delta}$. USE OF AN ISOTHERMAL ASSUMPTION. STREAMWISE VARIATION OF THE PRESSURE. If one multiplies the two members of (2.43) by the mass density ρ , one obtains

$$\rho \dot{\Delta} = - \frac{\pi}{8\mu} \rho \frac{dp}{dz} \left(\frac{d}{2} \right)^4. \quad (2.44)$$

According to the ideal gas law $p = \rho RT$. One may therefore substitute $p/(RT)$ for ρ in the right member of (2.44) to obtain

$$\rho \dot{\Delta} = -\frac{\pi}{8\mu} \frac{p}{RT} \frac{dp}{dz} \left(\frac{d}{2}\right)^4 - \frac{\pi}{8\mu} \frac{1}{RT} \left(\frac{d}{2}\right)^4 \frac{d}{dz} \left(\frac{p^2}{2}\right). \quad (2.45)$$

Now $\rho \dot{\Delta}$ is the rate of transport of *mass* through the capillary and is therefore independent of the streamwise coordinate z . If one assumes, moreover, that the absolute temperature is effectively independent of z , then one may integrate (2.45) with respect to z to obtain

$$(\rho \dot{\Delta})(z - z_0) = -\frac{\pi}{8\mu} \frac{1}{RT} \left(\frac{d}{2}\right)^4 \left(\frac{p^2}{2} - \frac{p_0^2}{2}\right), \quad (2.46)$$

in which the constant of integration has been fixed by the condition

$$(p)_{z=z_0} = p_0. \quad (2.47)$$

If one multiplies (2.46) by $RT/(z - z_0)$ and recalls that $RT\rho = p$, one obtains

$$p \dot{\Delta} = -\frac{\pi}{16\mu} \left(\frac{d}{2}\right)^4 \frac{p^2 - p_0^2}{z - z_0}. \quad (2.48)$$

If one evaluates (2.48) at a station where

$$(p)_{z=z_1} = p_1, \quad (2.49)$$

then (2.48) implies that

$$p \dot{\Delta} = -\frac{\pi}{16\mu} \left(\frac{d}{2}\right)^4 \frac{p_1^2 - p_0^2}{z_1 - z_0}. \quad (2.50)$$

One may apply (2.50), for example, in the case when station zero is the inlet (at, say, atmospheric pressure) and station one is at the outlet (where the pressure is some given lower pressure). In this case, of course, $z_1 - z_0$ represents the length of the capillary, ℓ .

If one eliminates $p \dot{\Delta}$ between (2.48) and (2.50), one obtains, after simplification,

$$\frac{p^2 - p_0^2}{z - z_0} = \frac{p_1^2 - p_0^2}{z_1 - z_0}. \quad (2.51)$$

One may interpret (2.51) as an interpolation formula for the local pressure $p(z)$ at an arbitrary station z in the range $z_0 < z < z_1$. If one solves (2.51) for p , one obtains

$$p = \sqrt{p_0^2 + (p_1^2 - p_0^2) \left(\frac{z - z_0}{z_1 - z_0}\right)}. \quad (2.52)$$

which shows that in the case of gas flow, unlike the case of liquid flow, the pressure depends nonlinearly upon the axial coordinate.

2.3.3 THE TIME FOR A FLUID PARTICLE TO TRAVEL THE LENGTH OF A CAPILLARY. Consider a reference frame attached to a hypothetical material point which moves with the gas velocity \mathbf{u} . Since the material point moves with the fluid and the pressure decreases in the downstream direction, the pressure to which the material point is subject must be time dependent. If one denotes the time derivative of p relative

to the material point by Dp/Dt , one may show from the chain rule and the kinematic definitions of the velocity components in cylindrical coordinates that

$$\frac{Dp}{Dt} = \frac{\partial p}{\partial t} + u_r \frac{\partial p}{\partial r} + \frac{u_\theta}{r} \frac{\partial p}{\partial \theta} + u_z \frac{\partial p}{\partial z}. \quad (2.53)$$

The assumption that the flow is steady (cf. §2.2.2) implies that $\partial p/\partial t = 0$. The assumption that the flow is axisymmetric (cf. §2.2.3) implies that $\partial p/\partial \theta = 0$. The assumption that the flow lacks swirl (cf. §2.2.3) implies that $u_\theta = 0$. The assumption that the capillary is slender (cf. §2.2.4) implies that $u_r \ll u_z$ and $\partial p/\partial r \ll \partial p/\partial z$. Equation (2.53) thus reduces to

$$\frac{Dp}{Dt} = u_z \frac{dp}{dz} \quad (2.54)$$

under assumptions and idealizations already introduced. If one multiplies (2.54) by p^2 , one obtains

$$p^2 \frac{Dp}{Dt} = (pu_z) \left(p \frac{dp}{dz} \right). \quad (2.55)$$

Consider a cylindrical surface where u_z equals the cross sectionally averaged fluid velocity \bar{u}_z defined by

$$\bar{u}_z \equiv \frac{\dot{\Delta}}{\pi(d/2)^2}. \quad (2.56)$$

On a streamline where $u_z = \bar{u}_z$, (2.55) becomes

$$\left(p^2 \frac{Dp}{Dt} \right)_{u_z = \bar{u}_z} = \frac{p\dot{\Delta}}{\pi(d/2)^2} p \frac{dp}{dz}. \quad (2.57)$$

If one eliminates dp/dz from the right member by means of (2.43), one obtains

$$\left(p^2 \frac{Dp}{Dt} \right)_{u_z = \bar{u}_z} = \frac{p\dot{\Delta}}{\pi(d/2)^2} p \left(-\frac{8\mu}{\pi} \frac{\dot{\Delta}}{(d/2)^4} \right),$$

or

$$\left[\frac{D}{Dt} \left(\frac{p^3}{3} \right) \right]_{u_z = \bar{u}_z} = -\frac{8\mu}{\pi^2(d/2)^6} (p\dot{\Delta})^2. \quad (2.58)$$

The right member is independent of the axial coordinate z . It follows that the right member is independent of t in a reference frame moving with the velocity $u_z = \bar{u}_z$. One may thus integrate (2.58) with respect to time subject to the conditions

$$(p)_{t=t_0} = p_0, \quad (p)_{t=t_1} = p_1 \quad (2.59)$$

to obtain

$$\frac{p_1^3 - p_0^3}{3} = -\frac{8\mu}{\pi^2(d/2)^6} (p\dot{\Delta})^2 (t_1 - t_0). \quad (2.60)$$

The two equations (2.50) and (2.60) relate the inlet and the outlet pressures to three other fixed parameters for the capillary, namely the capillary diameter, d , the throughput, $p\dot{\Delta}$, and the transit time, $t_1 - t_0$. One may eliminate any one of these latter three parameters between (2.50) and (2.60) to obtain a formula relating the remaining two. If, for example, one eliminates the throughput $p\dot{\Delta}$ from (2.60) by means of (2.50), one obtains

$$\frac{p_1^3 - p_0^3}{3} = -\frac{8\mu}{\pi^2(d/2)^6} \left[-\frac{\pi}{16\mu} \left(\frac{d}{2} \right)^4 \frac{p_1^2 - p_0^2}{z_1 - z_0} \right]^2 (t_1 - t_0) = -\frac{8}{(2 \cdot 8)^2} \frac{1}{\mu} \left(\frac{d}{2} \right)^2 \frac{(p_1^2 - p_0^2)^2}{(z_1 - z_0)^2} (t_1 - t_0), \quad (2.61)$$

whence

$$t_1 - t_0 = -\frac{p_1^3 - p_0^3}{(p_1^2 - p_0^2)^2} \frac{32}{3} \mu \frac{(z_1 - z_0)^2}{(d/2)^2} \quad (2.62)$$

The result (2.62) expresses the time of transit $t_1 - t_0$ in terms of the inlet and outlet pressures, the fluid viscosity, and the ratio $d/(z_1 - z_0)$ (which is the slenderness parameter ϵ introduced in §2.2.4).

2.4 DIMENSIONS OF THE QUICK TEST CONDUIT. A ONE-CAPILLARY MODEL FOR PREDICTING $p\dot{\Delta}$ AND $t_1 - t_0$. COMPARISONS WITH OBSERVATION.

The Quick Test attachment to the Leybold Hereaus Ultratest F mass spectrometer leak detector may be operated with hoses of various internal diameters. Thus, a standard red hose has an internal diameter $d = d_r = 0.75$ mm, a standard green hose has an internal diameter $d = d_g = 0.90$ mm, and a standard blue hose has an internal diameter $d = d_b = 1.15$ mm*. During maintenance of the machine some combination of red and green hose is chosen by trial and error to ensure that the probe throughput $p\dot{\Delta}$ has the nominal value $1.5 \text{ atm}\cdot(\text{cm})^3/\text{sec}$. In a typical configuration, 25 feet of red hose is placed upstream of 75 feet of green hose. Upstream of the hoses is a metal capillary whose length is typically 8 cm and whose inside diameter is $d = d_c = 0.5$ mm.

Suppose one idealizes the problem by pretending that a single green hose of length 100 feet constitutes the whole conduit. Then in (2.50), one may take $z_1 - z_0 = \ell = 100 \text{ ft} = 3048 \text{ cm}$ and $d/2 = d_g/2 = 0.045 \text{ cm}$. For air at room temperature, say $T = 295^\circ\text{K} = 71.33^\circ\text{F}$, one may take $\mu = 1.8122 \times 10^{-4} \text{ dyne}\cdot\text{sec}/(\text{cm})^2$. If, moreover, one assumes that $(p_1/p_0)^2 \ll 1$, then (2.50) takes the approximate form

$$p\dot{\Delta} \approx \frac{\pi}{16\mu} \left(\frac{d_g}{2}\right)^4 \frac{p_0^2}{\ell}, \quad (2.63)$$

in which $p_0 = 1 \text{ atm} = 1.01325 \times 10^6 \text{ dyne}\cdot\text{sec}/(\text{cm})^2$. Thus,

$$p\dot{\Delta} \approx \frac{\pi(0.045 \text{ cm})^4 [(1.013 \times 10^6 \text{ dyne}/(\text{cm})^2)^2]}{16[1.8122 \times 10^{-4} \text{ dyne}\cdot\text{sec}/(\text{cm})^2](3048 \text{ cm})} = 1.496 \times 10^6 \frac{\text{dynes} (\text{cm})^3}{(\text{cm})^2 \text{ sec}}. \quad (2.64)$$

If one multiplies by

$$1 = \frac{1 \text{ atm}}{1.013 \times 10^6 \text{ dyne}/(\text{cm})^2}, \quad (2.65)$$

one obtains

$$p\dot{\Delta} \approx 1.477 \frac{\text{atm}\cdot(\text{cm})^3}{\text{sec}} \quad (\text{all green hose}). \quad (2.66)$$

This value agrees well with the target value $p\dot{\Delta} = 1.5 \text{ atm}\cdot(\text{cm})^3/\text{sec}$.

If one repeats the calculation of $p\dot{\Delta}$ but for an all red hose with $d = d_r = 0.75$ mm, instead of a green one (with $d = d_g = 0.90$ mm), one obtains

$$p\dot{\Delta} \approx 0.712 \frac{\text{atm}\cdot(\text{cm})^3}{\text{sec}} \quad (\text{all red hose}). \quad (2.67)$$

Thus a reduction of hose diameter from $d_g = 0.9$ mm to $d_r = 0.75$ mm roughly halves the throughput.

If one substitutes the same values of the parameters into formula (2.62) for the transit time $t_1 - t_0$, one obtains

$$t_1 - t_0 \approx \frac{32}{3} \frac{\mu \ell^2}{p_0 (d/2)^2}. \quad (2.68)$$

For the all green hose, (2.68) gives

$$t_1 - t_0 \approx \frac{32 [1.8122 \times 10^{-4} \text{ dyne}\cdot\text{sec}/(\text{cm})^2](3048 \text{ cm})^2}{3 [1.013 \times 10^6 \text{ dyne}/(\text{cm})^2](0.045 \text{ cm})^2} = 8.754 \text{ sec} \quad (\text{all green hose}). \quad (2.69)$$

The corresponding calculation for the all red hose gives

$$t_1 - t_0 = 12.61 \text{ sec} \quad (\text{all red hose}). \quad (2.70)$$

The response time for a typical Quick Test configuration is about 8 or 9 seconds. Once again, the idealization that the hose is all green agrees well with nominal performance data for the Quick Test.

* I am indebted to Messes KELVIN R. POLK and J. PERRY GODFREY of the Mass Spectrometer Lab of the OPF Shops (LSO-417) for furnishing me with this information.

SECTION THREE

A CONCEPT FOR A MICROENCLOSURE, OR SPONGE-TIPPED PROBE

The difficulty of obtaining repeatable measurements for the external supply-to-sample flowrate ratio is vexing. This problem becomes moot if one can construct an alternative probe design that simulates an enclosure, analogous to the brush head of a vacuum cleaner, rather than a probe end exposed to the free air in the test area.

I envisage a probe with a flat metal disk welded to the end (punctured at the center, so the probe capillary is not plugged). The diameter of the metal disk must be large compared to the capillary diameter. If the capillary diameter is 0.5 mm then the disk diameter may be as small as 0.5 cm. A porous sponge or foam material adheres to the disk and allows the user to place the probe flat against the test surface.

If the draw rate of air into the probe is large compared to the rate of discharge of helium from the leak one is trying to measure (as will be the case if the probe is attached to a typical branched conduit system), then there is reason to believe that external supply-to-sample flowrate ratio may be reduced to a level on the order of, say, 1.1 or less.

SECTION FOUR

RECOMMENDATIONS AND CONCLUSION

Recommendations:

1. Determine (by consultation with the authors, if necessary) whether the indicated-to-actual leak rate ratio in the Rockwell specification MF0001-003 should be interpreted as: (I) the end-to-end supply-to-sample flowrate ratio (*i.e.* the product of the internal and external supply-to-sample flowrate ratios); or, (II) the external supply-to-sample flowrate ratio only. If hypothesis (II) is correct then the Operations and Maintenance Instruction V1009.005 (and, hence, current practice at KSC) is in conflict with the Rockwell specification.
2. Initiate a program to develop a microenclosure probe, preferably a program that would lead to the fabrication and testing of a prototype device at KSC within the next twelve months.

Conclusion:

The observed flow rate $p\dot{\Delta}$ and the time $t_1 - t_0$ for a gas particle to travel the length of a capillary are predicted well by the theory of capillary flow of a gas developed in the written report for this project (provided it is applied under its stated assumptions).

REFERENCES

1. ANONYMOUS, MPS leak and functional test (LPS). (LH2 feed system). OMI Number V1009.005 June 21, 1993 (564 pages).
2. ANONYMOUS, Orbiter OMRSD main propulsion system, Space Shuttle Operations and Maintenance Requirements and Specifications Document, Volume 41, File III (Rockwell International), dated May 24, 1993 (679 pages).
3. ANONYMOUS, Proof pressure and leak detection-Aerospace plumbing systems and assemblies. Rockwell International specification (prepared by D.R. SUEME, & K.R. RAVAL). SPECIFICATION number MF0001-003 (Code Indent. No. 03953), Sept 27, 1992.
4. PETERSON, DENNIS M. Report of KSC-STG GHe leak check Independent Assessment. KSC SR&QA Systems Assurance Office May 24, 1991 (39 pages).
5. MARR, J. WILLIAM, *Leakage testing handbook*. Prepared by General Electric, Schenectady, N.Y. for Jet Propulsion Laboratory, Pasadena, CA. NASA Contractor Report CR-952, April 1968 (xx + 392 pages, with appendices).
6. BILARDO, VINCENT J., IZQUIERDO, FRANCISCO, & SMITH, R. KENNETH, Development of the Helium Signature Test for orbiter main propulsion system revalidation between flights. AIAA 25th Aerospace Sciences Meeting, Reno, NV, January 12-15 1987. AIAA Paper 87-0293 (9 pages).
7. SENEAR, A. E. & BLAIR, A. W., *Measurement and correlation of helium and fluid leak rates*. Prepared by Boeing for Manned Spacecraft Center, Houston, TX. Report of contract NASw-1650 (NASA CR 109977, Accession number N70-35954), July 1968 (189 pages).
8. ANONYMOUS, Preventive maintenance/validation procedure for the Leybold Ultratest F Helium Leak Detector. OMI number V6H31, Revision A-1, dated January 13, 1993 (34 pages).

1994018773
N94-28246

1993 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA

376-31
197202
p. 31

KSC MANAGEMENT TRAINING SYSTEM PROJECT

PREPARED BY:	Dr. Jose A. Sepulveda
ACADEMIC RANK:	Associate Professor
UNIVERSITY AND DEPARTMENT:	University of Central Florida Department of Industrial Engineering and Management Systems
NASA/KSC	
DIVISION:	Personnel Office
BRANCH:	Employee Development
NASA COLLEAGUE:	James Norman Carol Valdes
DATE:	August 3, 1993
CONTRACT NUMBER:	University of Central Florida NASA-NGT-60002 Supplement: 11

NASA/ASEE MANAGEMENT SCIENCE
FACULTY FELLOWSHIP PROJECT

KSC MANAGEMENT TRAINING SYSTEM PROJECT
FINAL REPORT- SUMMER 1993

TABLE OF CONTENTS

1.0 OBJECTIVES.....	2
2.0 THE INDIVIDUAL DEVELOPMENT PLAN SURVEYS FOR 1994.....	3
2.1 Commands Used to Query the Data Bases.....	3
2.1.1 Report #1 -- Courses by Type of Training by Organization.....	3
2.1.2 Report # 2 -- Type of Training by Organization.....	3
2.1.3 Report # 3 -- Summary by Organization.....	6
2.1.4 Report # 4 -- Summary by Type of Training.....	6
2.2 On Training Costs and Priorities.....	6
2.2.1 The Financial Impact of Training.....	6
2.2.2 The Priority of Training.....	10
2.3 Recommendations for Further Work.....	10
2.3.1 On the Financial Impact of Training.....	10
2.3.2 On the Priority of Training.....	10
3.0 THE EVALUATION OF TRAINING.....	11
3.1 Existing Evaluation Procedures.....	11
3.2 Objectives.....	11
3.3 Evaluation: A Literature Search.....	11
3.3.1 Sources.....	11
3.3.2 Approaches.....	13
3.3.3 Methods.....	13
3.3.4 Content Validity Ratio (CVR).....	13
3.3.5 References.....	14
3.4 The Evaluation Procedure.....	15
3.4.1 The Scanning Process.....	15
3.4.2 Evaluations.....	15
3.4.3 Reports.....	15
3.5 Expectations.....	20
3.6 Recommendations for Further Work.....	20
3.6.1 On Evaluation:.....	20
3.6.2 On the Training Management Information System:.....	20
3.6.3 On the Needs Assessment:.....	20
4.0 APPENDIX: Literature Review.....	21

NASA/ASEE MANAGEMENT SCIENCE
FACULTY FELLOWSHIP PROJECT

KSC MANAGEMENT TRAINING SYSTEM PROJECT
FINAL REPORT- SUMMER 1993

Faculty Fellow: José. A. Sepúlveda, Ph.D., P.E.
Associate Professor
Department of Industrial Engineering and Management Systems
University of Central Florida
Orlando, Florida 32816

1.0 OBJECTIVES

The stated objectives for the Summer of 1993 were:

1. Review the Individual Development Plan Surveys for 1994 in order to automate the analysis of the Needs Assessment effort.
2. Develop and implement evaluation methodologies to perform ongoing program-wide course-to-course assessment. This includes:
 - Propose a methodology to develop and implement objective, performance-based, assessment instruments for each training effort.
 - Mechanize course evaluation forms and develop software to facilitate the data gathering, analysis and reporting processes.
 - Implement the methodology, forms, and software in at least one training course or seminar selected among those normally offered in the summer at KSC.

Section 2 of this report addresses the work done in regard to the Individual Development Plan Surveys for 1994. Section 3 presents the methodology proposed to develop and implement objective, performance-based, assessment instruments for each training course offered at KSC.

2.0 THE INDIVIDUAL DEVELOPMENT PLAN SURVEYS FOR 1994

Section 2 of this report deals with the first objective, the Individual Development Plan Surveys for 1994. The information stored in the Personnel Training Developmental System (PTDS) was transferred to a Windows platform and processed using FoxPro for Windows. Section 2.1 presents the SQL commands used to query the data bases and produce the requested reports. As an example, the first page of each report is presented in the next pages. The full reports, already delivered to NASA, are considered part of this Final Report. Section 2.2 discusses issues associated to training cost and priority and presents recommendations for further research.

2.1 Commands Used to Query the Data Bases

2.1.1 Report #1 -- Courses by Type of Training by Organization

```
SELECT DISTINCT Report1.unit, Report1.tot, Report1.course_num,;
COUNT(Report1.course_num), SUM(Report1.cost), Courses.title,;
Traintyp.type, Units.unitname;
FROM Report1, Courses, Units, Traintyp;
WHERE Courses.course_num = Report1.course_num;
AND Units.unit = Report1.unit;
AND Traintyp.tot = Report1.tot;
AND Report1.unit = "AC";
GROUP BY Report1.course_num;
ORDER BY Report1.tot;
INTO CURSOR Report1
CREATE REPORT report1.frx FROM Report1 WIDTH 0 COLUMN NOOVERWRITE
REPORT FORM report1.frx TO PRINTER NOCONSOLE
```

2.1.2 Report # 2 -- Type of Training by Organization

```
SELECT DISTINCT Report1.unit, Report1.tot, COUNT(Report1.course_num),;
SUM(Report1.cost), Traintyp.type, Units.unitname;
FROM Report1, Courses, Units, Traintyp;
WHERE Courses.course_num = Report1.course_num;
AND Units.unit = Report1.unit;
AND Traintyp.tot = Report1.tot;
GROUP BY Report1.unit, Report1.tot;
ORDER BY Report1.unit, Report1.tot;
INTO CURSOR Report2
CREATE REPORT report2.frx FROM Report2 WIDTH 5120 COLUMN NOOVERWRITE
REPORT FORM report2.frx TO PRINTER NOCONSOLE NOEJECT
```

1994 Individual Development Plans

Organization: AC -- Comptroller

Type of Training: 10 -- Executive and Management

Course	Course Title	Number of Requests	Cost
		1	\$0.00
K1894	CONF SE EXEC PUB POL ISSUE	1	\$1,775.00
K2176	SCIENCE TECH AND PUBLIC POLICY	12	\$24,000.00
K534	NASA MANAGEMENT EDUCATION PROG	1	\$2,175.00
K545	EXEC DEVELOPMENT SEMINAR	5	\$4,465.00
K6140	RES MGT EDUC PROGRAM	1	\$1,350.00
K6166	MANAGEMENT DEV SEM	3	\$3,750.00
K6339	HUMAN ELEMENT (THE)	3	\$630.00
K6427	COMMUNICATION & INFLUENCE	1	\$670.00
K6715	MEP UPDATE:MGMT PRACTICES & ORGAN. CLIMATE	2	\$6,600.00
K6777	UNDERSTANDING BUSINESS POLICY & OPS.	4	\$0.00
K753	MANAGEMENT DEVELOPMENT SEMINAR	3	\$5,325.00
K8164	ADMINISTRATION OF PUBLIC POLICY	9	\$4,140.00
K8176	CROSSING DEPARTMENT LINES	3	\$0.00
K9716	TASK MANAGEMENT (NASA)	1	\$0.00
K9941	SEMINAR ON MANAGERIAL COMPETENCIES	2	\$0.00
K9942	EXECUTIVE PROJECT MANAGEMENT		
Subtotals for type 10 -- Executive and Management		52	\$54,880.00

1994 Individual Development Plans

Organization: AC -- Comptroller

Type of Training:	Number of Requests	
10 -- Executive and Management	52	\$54,880.00
20 -- Supervisory	68	\$37,784.00
3J -- Engineering - General	3	\$325.00
3M -- Engineering - Aerospace	2	\$2,550.00
3O -- Engineering - Electrical/Electronic	1	\$0.00
3T -- Engineering - Systems, Industrial	1	\$390.00
4A -- Public or Business Administration	46	\$13,136.44
4B -- Finance, Accounting, Auditing	89	\$33,332.00
4C -- Personnel	9	\$960.00
4D -- EEO, Race or Ethnic Studies	2	\$0.00
4E -- Procurement	97	\$24,871.00
4G -- Policy, Program, or Management Analysis	7	\$3,050.00
4N -- Security or Investigation	2	\$500.00
4U -- Program and Project Management	25	\$8,084.00
4Z -- Other Administrative	1	\$320.00
5A -- Systems, Safety and Area Access Training	4	\$0.00
5B -- Academic Course Work	101	\$22,797.88
6A -- Basic Clerical	8	\$634.00
6B -- Inter-personal and Office Management	11	\$1,009.00
6C -- Administrative Systems	10	\$750.00
6D -- Other Clerical	2	\$99.00
8A -- Orientation	1	\$0.00
8E -- Nets/Quality Circles	3	\$131.00
8F -- Productivity	1	\$386
8G -- Professional/Personal Development	13	\$1,418.00
8H -- Communication Skills	25	\$2,438.00
8Z -- Other General	1	\$595.00
9A -- Telecommunications, Networking	2	\$1,210.00
9B -- Programming and Programming Languages	26	\$1,255.00
9D -- Office Automation/AIM, Business	4	\$2,500.00
9E -- Software Management, Systems Software	2	\$1,127.00
9Z -- Other Computer Hardware and Software	2	\$0.00
TOTALS FOR AC -- Comptroller	621	216,532.32

2.1.3 Report # 3 -- Summary by Organization

```
SELECT DISTINCT Report1.unit, COUNT(Report1.course_num),;
SUM(Report1.cost), Units.unitname;
FROM Report1, Courses, Units, Traintyp;
WHERE Courses.course_num = Report1.course_num;
AND Units.unit = Report1.unit;
AND Traintyp.tot = Report1.tot;
GROUP BY Report1.unit;
ORDER BY Report1.unit;
INTO CURSOR Report3
CREATE REPORT report3.frx FROM Report3 WIDTH 640 COLUMN NOOVERWRITE
REPORT FORM report3.frx TO PRINTER NOCONSOLE NOEJECT
```

2.1.4 Report # 4 -- Summary by Type of Training

```
SELECT DISTINCT Report1.unit, COUNT(Report1.course_num),;
SUM(Report1.cost), Units.unitname;
FROM Report1, Courses, Units, Traintyp;
WHERE Courses.course_num = Report1.course_num;
AND Units.unit = Report1.unit;
AND Traintyp.tot = Report1.tot;
GROUP BY Report1.tot;
ORDER BY Report1.tot;
INTO CURSOR Report4
CREATE REPORT report4.frx FROM Report4 WIDTH 640 COLUMN NOOVERWRITE
REPORT FORM report4.frx TO PRINTER NOCONSOLE NOEJECT
```

2.2 On Training Costs and Priorities

2.2.1 The Financial Impact of Training

Training is expensive, however, lack of training could be even more expensive! Reliable estimates of training costs are needed for informed decision making by management concerning what training to offer and what to postpone. In PTDS, the COURSES.DBF's data base fields "TUITION", "BOOKS", and "OTHER COSTS" store information about each course's cost. Unfortunately, the information is not always available, and even when available, it is not always current, nor complete. There are a number of costs associated to a training effort, some of which are not included in the current data bases.¹

- *Direct*: Materials, instructor, tuition, cost for media, cost of employee travel and per diem, employee salary during training, opportunity cost of foregone production.
- *Indirect*: Administrative, office space, computers, simulators, depreciation of facilities.
- *Intangibles*: Costs associated with potential failure of personnel to perform a task or job; savings associated to finishing a job early due to good performance by everybody involved.

¹ The terms "direct" and "indirect" costs refer here to its accepted use in industrial engineering cost analysis studies. At KSC, "direct" cost may be understood as the cost borne by the trainee's organization, while "indirect" is any cost borne by a different organization.

1994 Individual Development Plans Summary by Organization

Organization	Number of Requests	Cost
-- No mail code provided	13	\$4,764.00
AC -- Comptroller	621	\$216,532.32
CC -- Chief Counsel	7	\$2,389.00
CD -- Center Director	20	\$23,616.00
CM -- Payload Management and Operations	40	\$18,154.00
CP -- Payload Projects Management	318	\$193,895.11
CS -- STS Payload Operations	2318	\$1,321,368.35
CV -- Expendable Vehicles	170	\$121,802.00
DE -- Engineering Development	136	\$59,804.00
DF -- Facilities Engineering	900	\$484,378.92
DL -- Electronic Engineering	1033	\$572,131.68
DM -- Mechanical Engineering	1225	\$459,930.46
EO -- Equal Opportunity Program	28	\$5,742.00
EX -- Executive Management	55	\$19,647.00
HM -- Human Resources and Management Systems	12	\$5,607.00
MD -- Biomedical Operations and Research	294	\$148,244.18
MK -- NASA Headquarters/Space Shuttle Operations	156	\$55,289.00
OP -- Procurement Office	710	\$301,214.91
PA -- Public Affairs Office	221	\$126,324.01
PM -- Personnel Office	733	\$103,086.00
PT -- Technology and Advanced Projects Office	154	\$81,938.00
RM -- Mission Assurance	490	\$294,545.00
RO -- Quality Assurance	2464	\$548,692.18
RQ -- Reliability and Quality Assurance	50	\$30,072.66
RT -- Safety and Reliability	659	\$226,156.10
SI -- Center Support Operations	1254	\$544,933.96
SS -- Space Station Project Office	262	\$159,045.00
TE -- Grounds Engineering	985	\$811,079.95
TL -- Shuttle Logistics Project Management	363	\$180,313.00
TM -- Shuttle Management and Operations	471	\$267,918.92
TP -- Shuttle Operations	842	\$338,121.95
TV -- Vehicle Engineering	3066	\$1,388,009.52
Totals for KSC	20070	\$9,114,747.17

1994 Individual Development Plans

Summary by Type of Training

Type of Training	Number of Requests	Cost
10 Executive and Management	1211	\$1,892,420
20 Supervisory	2691	\$1,655,615
3A Legal	31	\$17,805
3B Medical and Biomedical	33	\$4,620
3C Physical Sciences	19	\$6,655
3D Earth Sciences and Natural Resources	112	\$65,637
3F Human Factors	15	\$0
3H Mathematics and Statistics	5	\$178
3J Engineering - General	186	\$44,420
3K Engineering - Aeronautics	33	\$4,322
3M Engineering - Aerospace	698	\$391,765
3N Engineering - Chemical, Petroleum, or	146	\$56,735
3P Engineering - Civil, Architectural,	190	\$115,939
3Q Engineering - Electrical/Electronic	473	\$358,319
3R Engineering - Mechanics/Mechanical	165	\$97,324
3S Engineering - R & QA	246	\$128,603
3T Engineering - Systems, Industrial	55	\$72,268
3U Engineering - Safety	100	\$58,419
3W Other Legal, Medical and Scientific	46	\$5,565
4A Public or Business Administration	1224	\$138,947
4B Finance, Accounting, Auditing	249	\$75,729
4C Personnel	291	\$423,196
4D EEO, Race or Ethnic Studies	272	\$16,439
4E Procurement	745	\$188,475
4G Policy, Program, or Management Analysis	53	\$14,938
4J Quantitative Analysis or Operations	32	\$5,226
4K Safety Systems and Operation	34	\$12,407
4N Security or Investigation	84	\$12,470
4P Logistics	59	\$22,669
4Q Supply	20	\$3,345
4R General Safety and Health	288	\$59,557
4S Foreign Language	14	\$6,876
4U Program and Project Management	1065	\$836,679
4Y Sales, Marketing and Customer Service	22	\$5,545
4Z Other Administrative	86	\$27,610
5A Systems, Safety and Area Access Training	1325	\$0
5B Academic Course Work	2050	\$552,941
5C On-the-Job Training	24	\$0
6A Basic Clerical	109	\$9,086
6B Inter-personal and Office Management	354	\$35,828
6C Administrative Systems	242	\$30,297
6D Other Clerical	107	\$9,989
7B Engineering and Science Support/Journeyman	39	\$4,945
7Z Other Trade, Craft, Apprentice and	7	\$7,000
8A Orientation	24	\$0
8B Career Planning	243	\$6,836
8C Pre-Retirement	75	\$1,188
8E Nets/Quality Circles	472	\$29,117

File FIN_RPT.DOC, August 7, 1993, page 8

1994 Individual Development Plans

Summary by Type of Training

Type of Training	Number of Requests	Cost
8F Productivity	67	\$73,907
8G Professional/Personal Development	943	\$156,450
8H Communication Skills	1117	\$271,057
8Z Other General	85	\$68,633
9A Telecommunications, Networking	117	\$95,060
9B Programming and Programming Languages	764	\$710,244
9C Engineering Applications	146	\$106,807
9D Office Automation/AIM, Business	586	\$66,532
9E Software Management, Systems Software	71	\$40,547
9F Artificial intelligence, Expert Systems	6	\$1,875
9Z Other Computer Hardware and Software	104	\$9,690
TOTALS FOR KSC	20070	\$9,114,747

2.2.2 The Priority of Training

PTDS' IDP.DBF database file currently stores a "PRIORITY" field for each training request. Unfortunately, the field reflects the order in which the employee listed his/her requests (i.e., the first course listed is assigned priority 1, the second one has priority 2, etc.). Obviously this does not reflect the potential impact of the training for the Center, nor for the person making the request.

2.3 Recommendations for Further Work

2.3.1 On the Financial Impact of Training

Ensure that the training-related management information system (of which PTDS is currently the major component) includes the very best possible estimate of costs associated to each training course. The responsibility of keeping this data base up to date belongs to the Human Resources Development Branch. The cost data base should store, for each course, the training components (i.e, hours, units of resources, travel requirements, etc.), rather than the actual dollars spent. At report time the computer will calculate the dollar cost with actualized values per hour, unit of resource spent, etc. Whether the particular training component represents a direct cost to the trainee's organization or to other NASA organization should also be part of the training-related management information system.

2.3.2 On the Priority of Training

Include in the data base a field that reflects the potential importance that the requested training has for the Center and for the person making the request. This may be the result of an employee and supervisor review of the current requests. It would be time consuming, but it may be the best method to ensure that the user (NOT the Training Branch) defines the potential impact for the Center and the trainee.

As a possible indicator of priority of content and timing, the following scales were used in the Summer of 1992 in the Needs Assessment effort conducted by the author at KSC:

Potential Impact	Weight
Helps organizational performance	4
Helps individual's performance in current position	3
Enables person to perform more responsible position	2
Useful but not essential	1
There is no need for training in this subject matter	0

Priority on Timing	Weight
Needed within next three months	4
Needed within next six months	3
Needed within next year	2
Not an urgent need	1

3.0 THE EVALUATION OF TRAINING

3.1 Existing Evaluation Procedures

Evaluation of training courses is limited to KSC Form 13 (see next page). This form is processed manually and its major problems are that it is not linked to KSC's goals and objectives, there are no cost/benefits considered, the form is not used to assess performance improvement, and the results are not used for formal reporting.

3.2 Objectives

The evaluation of training at KSC is envisioned as having the following properties:

- Linked to KSC Mission, Goals and Objectives.
- Has a performance and values-based focus.
- Measures results in terms of performance improvement, financial impact, productivity and quality improvement, team building accomplishments, and less employee turnover.
- Resets priorities based on results.
- Tracks training and performance data.

The link to KSC's mission, goals and objectives and the performance and values-based focus will be ascertained through user involvement. It is the user (NOT the Human Resources Development Branch) who must define what is "required level of performance". The evaluation process will determine the training's impact on performance by asking the users: "What is the expected result of training?" and "What may happen if no training is given?"

3.3 Evaluation: A Literature Search

A literature search was performed to ensure that the proposed evaluation procedure agrees with standard practice. The following sections discuss the sources, most common approaches identified, typical methods to conduct the evaluation and an indicator of performance effectiveness, the "Content Validity Ratio". A detailed discussion of the search findings is provided in the appendix, along with the list of references used. Copies of all references were submitted to NASA in a separate binder and are considered part of this final report.

3.3.1 Sources

The literature search yielded 782 hits when the keep words "TRAINING", "EFFECTIVENESS", "ASSESSMENT", and "EVALUATION" were used. From those 782 titles, a total of 78 abstracts were selected and printed. The analysis of those abstracts resulted in 22 promising articles of which 14 were available at UCF's library (or through inter-library loans). Most articles are rather theoretical in nature but a few described specific applications in government agencies such as the Department of Agriculture (Plant Protection and Quarantine and Forest Service divisions), NASA Headquarters, the Department of the Navy's Navy Finance Center, and the Office of Personnel Management.

PARTICIPANT'S EVALUATION

SEMINAR :

PARTICIPANT'S NAME

ORGANIZATION

MAIL SYMBOL

TELEPHONE NUMBER

DATE(S)

SUPERVISOR'S NAME

PLEASE INDICATE YOUR EVALUATION OF THE SEMINAR IN TERMS OF THE FOLLOWING:

	EXCELLENT	GOOD	FAIR	POOR
1. Overall evaluation of seminar.				
2. Seminar content - amount of detail and extent of coverage.				
3. Level of presentation.				
4. Effectiveness of hand-out materials.				
5. Effectiveness of audio/visual aids.				
6. Instructor(s) knowledge of subject.				
7. Instructor(s) effectiveness.				
8. Application to your job.				

COMMENTS:

File FIN_RPT.DOC, August 3, 1993, page 12

SIGNATURE

TITLE

3.3.2 Approaches

Training evaluation usually takes one of four emphasis: (1) **Reaction**, where the evaluation is conducted at the end of the training exercise and focuses in the trainee's immediate feelings about the environment, instructor, and the material covered; (2) **Learning**, where the evaluation focuses on whether the student mastered specific knowledge or skills through exams or tests about the subject taught; (3) **Behavior**, where the evaluation consists on following the trainee over a period of time to ensure that the person's behavior changed as a result of the training; and (4) **Results**, where the emphasis of the evaluation is not the trainee's learning (or lack of it) or behavior but its impact on the organization's benefits accrued as a result of the training.

3.3.3 Methods

The most commonly used methods to gather evaluation data are surveys and questionnaires; interviews (individual or group); performance observations; and miscellaneous inquiries (from co-workers and supervisors). The Participant Action Plan Approach (PAPA, implemented by the Office of Personnel Management) asks the trainee at the end of the training session for a number of personal goals to achieve as a result of the training and comes back a year later to check whether those goals were attained. The main problems with observations, interviews, inquiries, and the PAPA method are that they may lack objectivity and require vast resources (time, personnel, and money) to conduct. For an on-going training evaluation system designed to include a large number of training programs, the best recommendation is a set of standardized surveys which printing, distribution, collection, analysis and reporting can be mechanized.

3.3.4 Content Validity Ratio (CVR)

The Content Validity Ratio is an indicator of the degree of acceptance of a given statement by a group of respondents. It ranges from -1 (absolute rejection) to +1 (unanimous approval). It is calculated as follows:

$$\text{CVR} = (\text{number who approve} - \text{number who reject}) / \text{number of respondents}$$

For example, if among 20 persons surveyed, 15 agree with the statement "I recommend that this training be taken by others in my organization", while 3 respondents disagree and 2 do not respond to the question, the CVR would be 0.67 [(15-3)/18 = 12/18 = 0.67]. Note that the indicator is non-linear.

3.3.5 References

- Brown, Mark G, "The Emperor's Still Naked: Why Management Remains in the Dark Ages," Performance and Instruction, v27 n9, Oct 1988, p1-4
- Cantor, Jeffrey A, "The Training Effectiveness Algorithm," Journal of Educational Technology Systems, v16 n3, 1988, p207-229
- Cantor, Jeffrey A, "A New Dimension in Military Instructional Development," Performance and Instruction, v27 n1, Jan 1988, p14-20
- Cantor, Jeffrey A, "Research and Development into a Comprehensive Media Selection Model," Journal of Instructional Psychology, v15 n3, Sep 1988, p118-131
- DeWine, Sue, "Evaluation of Organizational Communication Competency: The Development of the Communication Training Impact Questionnaire," Journal of Applied Communication Research, v15 n1-2, Spr-Fall 1987, p113-127
- Ford, J. Kevin, and Wroten, Steven P., "Introducing New Methods for Conducting Training Evaluation and for Linking Training Evaluation to Program Redesign," Personnel Psychology, v37 n4, Winter 1984, p651-65
- Geroy, Gary D., and Wright, Phillip C., "Evaluation Research: A Pragmatic, Program-Focused, Research Strategy for Decision-Makers," Performance Improvement Quarterly, v1 n3, 1988, p17-26.
- Hawley, John K., "A Practical Methodology for Determining Cost-Effective Instructional Programs," Performance and Instruction, v30 n5, May-Jun 1991, p17-23
- Korotkin, Arthur L., "A Taxonomic Approach to Integrating Job Analysis with Training Front-End Analysis," Performance Improvement Quarterly, v5 n3, 1992, p26-34.
- Kruger, Michael J., and Smith, Kay H., "Evaluating Management Training," Performance and Instruction, v26 n9-10, Nov-Dec 1987, p22-26
- Private Industry Council of Marin County, San Rafael, CA., "Training Program Evaluation Plan and Procedure," National Alliance of Business, Inc., Washington, D.C., Sponsoring Agency: Employment and Training Administration (DOL), Washington, D.C. Available from: Private Sector Initiative Program Clearinghouse, 1015 15th Street, N.W., Washington, D.C. 20005, Feb 1981, 38p.
- Salinger, Ruth, and Roberts, Cynthia, Evaluating the Impact of Training: A Collection of Federal Agency Evaluation Practices. Volume 2, Office of Personnel Management, Washington, D.C., Training Resources Management Division, 1984, 198p. Includes summaries of Volume 1.
- Stevens, Matthew D., and Hellweg, Susan A., "Toward a Standard of Communication Training Effectiveness Evaluation," Paper presented at the Annual Meeting of the Speech Communication Association (76th, Chicago, IL, November 1-4, 1990), 35p.
- Wreathall, John, and Connelly, Edward M., "Using Performance Indicators to Evaluate Training Effectiveness: Lessons Learned," Performance Improvement Quarterly, v5 n3, 1992, p35-43.

3.4 The Evaluation Procedure

The proposed evaluation procedure consists of a set of two evaluations and three reports. Copies of the evaluations (as proposed and as implemented) are included in the next pages.

3.4.1 The Scanning Process

An OpScan Model 5 from National Computer Systems (NCS) was purchased by KSC in early 1993 to mechanize the processing of all evaluation forms. The forms are designed with the help of DESIGN, a Windows based graphics package. DESIGN allows the user to define the questionnaire, to specify the types of responses and the response areas, and to print the questionnaire in forms supplied by NCS. Another NCS software, SURVEY, processes the designed forms and reads the scanned data into a file. A statistic (or a data base) package is then needed to analyze the data and produce reports. SPSS for Windows and FoxPro for Windows will be tried with the data collected in the first few training courses.

3.4.2 Evaluations

The proposed End-of-Course Evaluation is a "reaction" type evaluation. It has four sections. The first one, "Externals," deals with the training environment (location, handouts, etc.). The second, "Instruction," focuses in the instructor; the third, "Focus," on the contents; and the last, "Assessment," provides the trainee's feedback of the perceived value for the organization. An overall "Happiness Index" will be associated (based on a CVR type of analysis) to the trainees' reaction to each training course.

A Follow-Up Evaluation was proposed, to be delivered about six months after the training. Both the trainee and the supervisor were targeted. The form focused on the training's effect on performance, productivity, and satisfaction. It asked also for suggestions for improvement and recommendations for continuation of training. Although a standard form was proposed for both trainee and supervisor, the analysis of the trainee's responses would have concentrated in "behavior modification" issues, while the analysis of the supervisor's responses would have focused on the "organizational impact." An overall "Effectiveness Index" was to be associated (based on a CVR type of analysis) for the trainees' and supervisors' delayed assessment of the real impact of each training course.

In order to streamline the overall procedure, KSC's Human Resources Development Branch decided to replace the proposed Follow-Up Evaluation by a briefer form that targets only the supervisor. A copy of that form is also provided. A limited "Effectiveness Index" may be associated (based on a CVR type of analysis) to each training course, on the basis of the supervisor's delayed assessment.

3.4.3 Reports

The "Reaction" Report will be an End-of-Course Evaluation Report to be submitted to the Human Resources Development Branch. It will include information such as course ID, date, time, place, instructor, attendance, cost, student reactions, and an overall "happiness index" for the course.

The "Assessment of Impact" Report is the result of the follow-up evaluations. As this evaluation stands now, the report will essentially reflect the supervisor's assessment of the training impact on the organization's performance as reflected by his recommendation for future similar training for personnel under his supervision.

An End-of-Quarter Report is planned which will summarize the number of courses taught, the number of students, cost, and indicators of overall happiness and overall effectiveness.

**KENNEDY SPACE CENTER
TRAINING BRANCH
INITIAL EVALUATION OF TRAINING**

SSN: - -
MAIL CODE:

COURSE:	_____
INSTRUCTOR:	_____
COURSE DATE(S):	_____
SITE:	_____

INSTRUCTIONS: Indicate the degree to which you agree with the following statements using
SA=STRONGLY AGREE A=AGREE D=DISAGREE SD=STRONGLY DISAGREE NA=NOT APPLICABLE

EXTERNALS

1. The environment was suitable for learning.
2. There were no distractions during the session.
3. Handouts used at the session were effective.
4. Audio/visual aids used at the session were effective.

SA	A	D	SD	NA
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

INSTRUCTION

5. The instructor delivered the lessons clearly and effectively.
6. The instructor was knowledgeable of the material taught.
7. The instructor kept my interest throughout the session.
8. The instructor showed concern for the student's understanding of the material.
9. The hands-on component was delivered effectively.
10. Enough time was allotted to hands-on practice.

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

FOCUS

11. The material in this class is consistent with the actual and current requirements of my job.
12. The training focused on the specific tasks, knowledge and skills needed for acceptable job performance.
13. The knowledge and skills were taught to the appropriate level of proficiency.
14. Training was included that is NOT needed on the job.
15. Required skills and knowledge were NOT adequately covered.
16. The level of knowledge attained in training meets the level needed for acceptable job performance.

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

ASSESSMENT

17. Overall, I am pleased with this course.
18. This training will help me on my job.
19. This training course is right on target with KSC's needs.
20. I recommend that others from my organization attend this course.

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**KENNEDY SPACE CENTER
HUMAN RESOURCES DEVELOPMENT BRANCH**

INITIAL EVALUATION OF TRAINING

NAME: _____

COURSE: _____
 INSTRUCTOR: _____
 COURSE DATE(S): _____
 SITE: _____

INSTRUCTIONS: Indicate the degree to which you agree with the following statements using
SA=STRONGLY AGREE A=AGREE D=DISAGREE SD=STRONGLY DISAGREE NA=NOT APPLICABLE

PLEASE COMPLETELY FILL IN THE APPROPRIATE BUBBLES WITH A #2 PENCIL.

EXTERNALS

1. The environment was suitable for learning.
2. Handouts used at the session were effective.
3. Audio/visual aids used at the session were effective.

SA	A	D	SD	NA
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

INSTRUCTION

4. The instructor delivered the lessons clearly and effectively.
5. The instructor was knowledgeable of the material taught.
6. The hands-on component was delivered effectively.

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

FOCUS

7. The material in this class is consistent with the actual and current requirements of my job.
8. The level of knowledge attained in training will help improve my job performance.

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

ASSESSMENT

9. Overall, I am pleased with this course.
10. This training course is consistent with KSC's needs.
11. I would recommend that others from my organization attend this course.

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Suggestions for improvement or other comments:

TO: _____
(SUPERVISOR'S NAME)

MAILCODE: _____

**KENNEDY SPACE CENTER
HUMAN RESOURCES DEVELOPMENT BRANCH
FOLLOW-UP EVALUATION OF TRAINING**

PARTICIPANT'S NAME: _____

COURSE: _____
VENDOR: _____
COURSE DATE(S): _____

Approximately six months ago, the employee identified above attended the referenced training class. In order to assess the long range effects of this training, please indicate the degree to which you agree with the following statements using:

SA=STRONGLY AGREE A=AGREE D=DISAGREE SD=STRONGLY DISAGREE NA=NOT APPLICABLE

PLEASE COMPLETELY FILL IN THE APPROPRIATE BUBBLES WITH A #2 PENCIL.

1. This employee's job performance has improved as a direct result of this training.
2. I would recommend that others from my organization attend this course.

SA	A	D	SD	NA
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

TODAY'S DATE: _____

Suggestions for improvements or other comments concerning this training:

PLEASE RETURN THIS SURVEY TO **HM-PER-1** IN A U.S. GOVERNMENT MESSENGER ENVELOPE.

3.5 Expectations

The proposed evaluation procedure will (a) examine training requests from the perspective of KSC goals and objectives, (b) determine if performance improves as a result of investment in training, and (c) keep a continuous review of the training results. This will allow the Human Resources Development Branch to document and report to management and to departments (a) the time and resources needed to provide training, (b) the employee time and Center resources spent in training, and (c) an objective evaluation of the results obtained (aggregate values).

3.6 Recommendations for Further Work

3.6.1 On Evaluation:

The first recommendation is to implement the proposed evaluation procedure. It is suggested to try the proposed evaluation forms in at least two courses and to develop software to produce standard reports. Then, use the software to generate reports, and obtain feedback from the evaluators and from the report recipients.

3.6.2 On the Training Management Information System:

The second recommendation is to link the evaluation data to the PTDS (NTDS?) information system. The data should include (a) Cost data (NOT in dollars, but in hours, units of resources, etc.); (b) delivery data (who, where, how, when, how long, why, what (syllabus), number of attendants, instructor, student's initial and delayed evaluations, supervisor evaluation, comments); and (c) training evaluation data.

3.6.3 On the Needs Assessment:

The whole Training System should be consistent and dynamic. The evaluations should be used as a tool for constantly revising needs and priorities of what is being trained (Needs Analysis), how (Method/Mean Analysis), and how often. A procedure to this effect should be devised and implemented.

4.0 APPENDIX: Literature Review

The following pages were prepared by Joseph Espino, a SHARP (Summer High School Apprenticeship Research Program) student assigned to work under Dr. José A. Sepúlveda's mentoring during the Summer of 1993. In addition to preparing this literature search, Mr. Espino was instrumental in getting the NCS scanner in operation. He also wrote an operator manual for this hardware.

TRAINING EVALUATION - An Overview Of Existing Practices

Four categories of training evaluation are reaction, learning, behavior and results [Stevens & Hellweg, 1990].

Reaction is the trainee's liking of the training course. This response should be determined as soon as possible after the session is completed [Dewine, 1987]. Questions dealing with the atmosphere, physical environment, instructors and training media should be included on the first questionnaire. An example of a reaction based question is:

6. Rate the use of handouts in the training. _Excellent _Good _Fair _Poor

Learning is the evaluation of whether the trainee remembers (not necessarily uses) facts, principles and behaviors taught in class. This is the least relevant level of evaluation for an organization because increased knowledge does not necessarily mean the knowledge is used on the job [Stevens & Hellweg, 1990]. An example of a learning based question is:

2. Pentium class microprocessors must be super-cooled prior to use. _True _False

The **behavior** level is the most complex level of evaluation. This deals with the actual job improvement of the trainee. The behavior evaluation should come from trainee, its supervisor, subordinates, and peers. It should be performed at least three months after the training

[Stevens & Hellweg, 1990]. An example of a behavior based question is:

6. I am capable of speaking to large groups: Excellently Well Acceptably Poorly

Results are the actual effects that training has had on an organization. Items such as turnover, absences, sales, customer satisfaction, quantity and quality of performance, morale, cost reductions, grievance reductions, accident rates, suggestions, employee satisfaction and time use should be evaluated [Stevens & Hellweg, 1990].

8. Morale at the worksite has (since my training). Increased significantly
 Remained about the same
 Decreased significantly

The outcome of one of the four areas may affect other areas [Dewine, 1987]. The amount of time between training and a post training evaluation is not a standard. Ranges of suggested times are from two weeks to twenty-three months [Stevens & Hellweg, 1990]. An average time is from four to six months.

An alternate method of evaluation of training is content evaluation. **Content validity** is the evaluation of how the training content is related to the job. The training content must be identified. Items that are included are knowledge, skills, abilities, and other personal characteristics (KSAOs) that are needed for job performance. The content is then evaluated through a Content Validity Ratio (CVR) approach. People who have a thorough knowledge of the job rate each KSAO on its importance to the

job. The following formula is used to determine validity [Ford & Wroten, 1984].

$$\text{CVR} = \frac{\text{People stating KSAO is important} - \text{People stating KSAO is unimportant}}{\text{Total Number of People}}$$

Positive CVRs mean that greater than half of the people stated that the KSAO was important and negative CVRs mean that less than half of the people stated that the KSAO was important. A method for matching the CVR data and training needs would include the following steps. First divide the KSAOs into categories. Have your job experts rate the important KSAOs on a scale. This scale would determine how training intensity and resources should be spent on different areas of subject content [Ford & Wroten, 1984]. These methods would work well. However, they would require customized forms for each course and extra manpower because the procedure could not be automated.

Kruger & Smith (1987) recommended two different surveys: a skills/behavior survey and a health/stress survey. Both surveys use a system of examining statements and choosing to what degree you agree with the statements. The following is an example:

3. I am able to communicate with my manager. AGREE 9 8 7 6 5 4 3 2 1 DISAGREE

The **skills/behavior** survey deals with what was learned, how the trainee's performance has changed and how it has affected the organization. A survey is filled out by the trainee, its supervisor and some others. A sample

question on a skills/behavior survey might be:

6. I am capable of speaking to large groups: Excellently Well Acceptably Poorly

The **health/stress** survey identifies health and stress risks to the employee. It is only be filled out by the trainee. A sample question on a health/stress survey might be:

6. My knowledge of safety that is required for the job is: Proficient
 Adequate
 Not Adequate

Existing Training Evaluation Programs

The following are some training evaluation programs currently in place at different governmental and non-governmental agencies.

Department of Agriculture-Plant Protection and Quarantine: The Department of Agriculture's Plant Protection and Quarantine branch currently uses an evaluation program in which the evaluation is done by front-line supervisors. The supervisors are first trained in the skills needed to evaluate the trainees. The supervisors are then assigned to evaluate recently trained officers. The evaluators may use questionnaires, interviews, performance observations and miscellaneous inquiries to gather data. These are all used after the training. The trainees supply data concerning their performance through either questionnaires and interviews. The evaluators gather their own data through on-the-job performance evaluations and candid inquiries. The data is collected by a central organization and compiled to determine if the skills taught actually assisted the officers with their jobs. Advantages of this program include providing relevant data for evaluation and actively involving supervisors in the evaluation. Disadvantages include long length of time needed for process, requiring supervisors to do work solely on evaluation causes need to hire temporary replacements and expensive travel costs of

supervisors. A final disadvantage is that the evaluation is based on the specific job or supervisor's objectives and not those of the entire organization [Salinger & Roberts, 1984]

Department of Agriculture-Forest Service: The Forest Service currently uses a questionnaire-based evaluation system. Standard questionnaires are completed by the trainee, its supervisor and subordinates before the training session and four weeks after the session. Copies of the questionnaire are also completed by untrained personnel and their supervisors and subordinates in order to have a control population. Questions in the questionnaire are behavior based such as the following

To what extent do you think training:	GREATLY	NOT AT ALL
Will improve your ability to be a better supervisor?	9 8 7 6 5 4 3 2 1	
Will increase productivity?	9 8 7 6 5 4 3 2 1	
Is a good use of your time?	9 8 7 6 5 4 3 2 1	
Improve your communication skills on the job?	9 8 7 6 5 4 3 2 1	

Results are statistically tabulated. The results are then compiled and reported to the specific organizations. The questionnaires are behavior based and completely confidential. The tabulation stage tests the statistical significance of questions, trainee, supervisor and subordinate perceptions and comparison of trained results versus the control's results. Advantages of this program include accurate results due to use of a control group and small amounts of time and resources needed. The only significant disadvantage is the possibility of low return rate of questionnaires [Salinger & Roberts, 1984].

National Aeronautics & Space Administration: The evaluation program described below was for a career seminar sponsored by NASA Headquarters. The evaluation was directly built into the training program. The evaluation consisted of both pre and post training questionnaires, group interviews and individual interviews. The questionnaires contained general, open-ended questions dealing with current and future job goals. A sample question on a survey might be:

7. What have you done to achieve your goals in the past 30 days?

In the group interviews, the trainees spoke about their own job plans and how they were using what they learned.

Trainees could hold optional individual interviews with the instructor concerning their own progress. There were also informal evaluations during the course. The instructor was the only evaluator involved and there was no major computation of data in the evaluation. The instructor had to be available after the training session for many months in order to do the personal interviews. Advantages of the evaluation program included adaptability. Disadvantages include the fact that it is not very structured, forms need to be customized for different courses, constant need for the instructor, no organizational input and most of work has to be in groups which may be difficult to assemble [Salinger & Roberts, 1984].

Department of the Navy- Navy Finance Center: The Navy Finance Center evaluated one of its Interaction Management

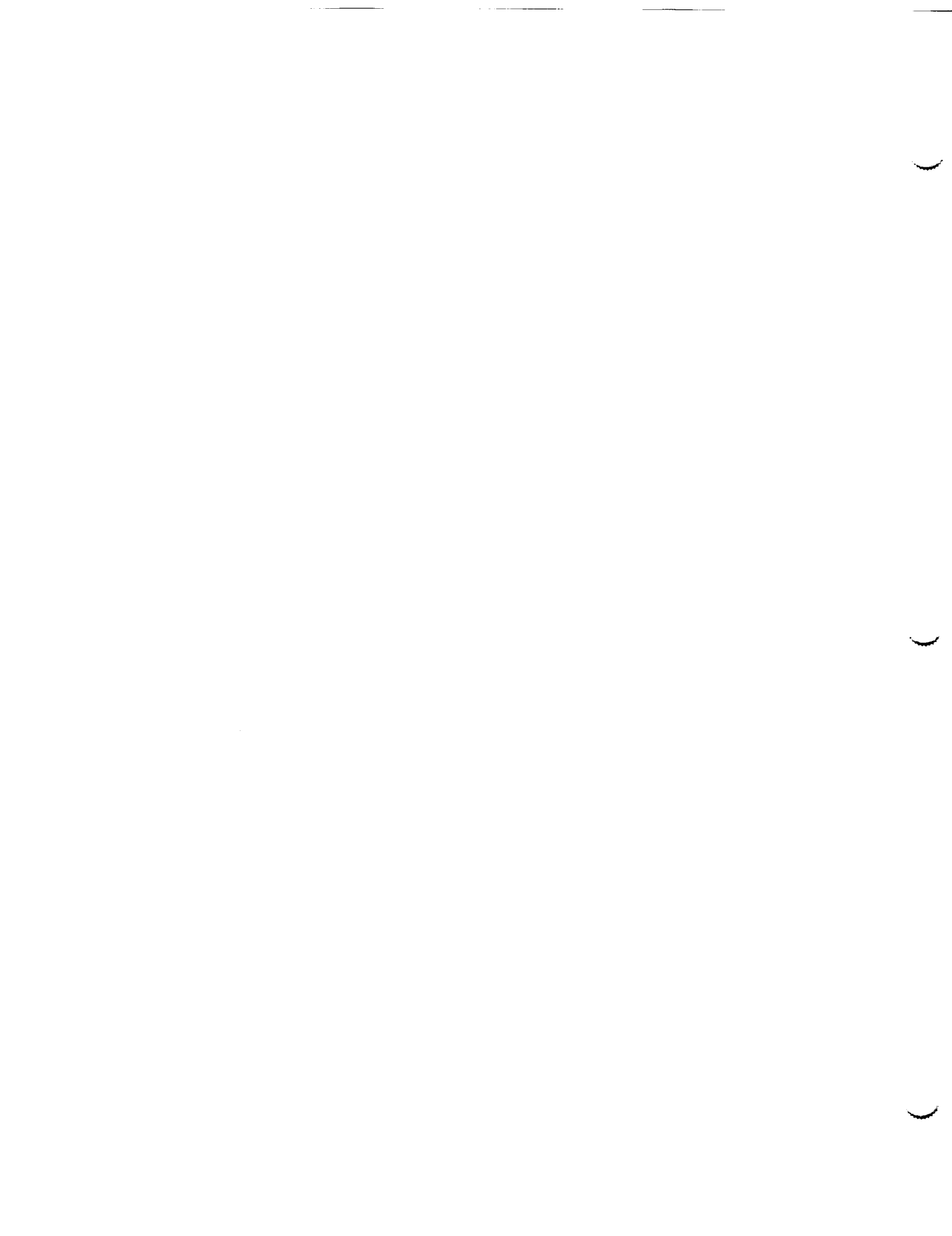
Courses through a system of questionnaires done in a group setting. The questionnaires were completely behavior based with questions such as:

When my supervisor speaks to me regarding <u>disciplinary action</u> , he/she ...	Always	Never
Uses my ideas when I offer them.	6 5 4 3 2 1	
Acknowledges my feelings about the situation.	6 5 4 3 2 1	

The questionnaires were completed by employees, their managers and subordinates that did and did not take the course. The evaluation was done before the training and two years after. The results were then computer tabulated to create statistical averages of how the training works through the perceptions of the employees, their managers and subordinates. Reports were then created from that data. Advantages of this evaluation method includes involvement of all members of an organization, use of a control (the untrained sample) to ensure accuracy of results and high return rate because of the group approach to completing the questionnaire. Disadvantages include a management commitment and travel expenses if the respondents are not in a common geographical area [Salinger & Roberts, 1984].

Office of Personnel Management: The OPM devised a generic training evaluation program called the Participant Action Plan Approach (PAPA). The evaluation system asked each participant to identify certain behaviors or actions related to the course that he/she would like to change. Here the evaluator would either interview the trainee (by phone or in person) or use questionnaires (six months after the

training) to determine whether the actions were carried out. The evaluation also checked for results of tried action plans and why untried action plans were not attempted. The questionnaire also asked how the course could be changed to better teach the trainee. Problems and obstacles that deterred the trainee from carrying out the plan were also identified. The results were computer tabulated. Results were used to create reports and eventually modify the course. The program provides data primarily concerning behavioral changes brought about by the course. Advantages of this program include ability to be used for many courses and it can accurately describe what actions came from the training. Disadvantages include the need of self-reports and skilled interviewers, large amount of time needed and it does not involve supervisors and subordinates in the evaluation process [Salinger & Roberts, 1984].



1993 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA

*** REPORT NOT AVAILABLE AT PUBLICATION TIME ***

PREPARED BY: Dr. Sherif A. Sherif
ACADEMIC RANK: Associate Professor
UNIVERSITY AND DEPARTMENT: University of Florida
Department of Mechanical Engineering
NASA/KSC
DIVISION: -----
BRANCH: -----
NASA COLLEAGUE: -----
DATE: -----
CONTRACT NUMBER: University of Central Florida
NASA-NGT-60002 Supplement: 11

PRECEDING PAGE BLANK NOT FILMED

—

—

—

1994018774

N94-23247

1993 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA

417-63
197203
p. 20

EXPANDING THE KATE TOOLBOX

PREPARED BY:	Dr. Stan J. Thomas
ACADEMIC RANK:	Associate Professor
UNIVERSITY AND DEPARTMENT:	Wake Forest University Department of Mathematics and Computer Science
NASA/KSC	
DIVISION:	Engineering Development
BRANCH:	Artificial Intelligence
NASA COLLEAGUE:	Peter Engrand
DATE:	August 6, 1993
CONTRACT NUMBER:	University of Central Florida NASA-NGT-60002 Supplement: 11

Acknowledgements

I would like to thank Peter Engrand of NASA, Charlie Goodrich, Chuck Pepe and all my INET colleagues for their cooperation and technical support this Summer. Thanks and best wishes also go to Li Yang as she continues this work. I would also like to extend both compliments and thanks to Dr. Hosler and Ms. Stiles for their professional management of the 1993 summer faculty program.

Abstract

KATE is a model-based software system developed in the Artificial Intelligence Laboratory at the Kennedy Space Center for monitoring, fault detection, and control of launch vehicles and ground support systems. In order to bring KATE to the level of performance, functionality, and integratability needed for firing room applications, efforts are under way to implement KATE in the C++ programming language using an X-windows interface.

This report describes two programs which have been designed and added to the collection of tools which comprise the KATE toolbox. The first tool, called the schematic viewer, gives the KATE user the capability to view digitized schematic drawings in the KATE environment. The second tool, called the model editor, gives the KATE model builder a tool for creating and editing knowledge base files. The body of this report discusses design and implementation issues having to do with these two tools. It will be useful to anyone maintaining or extending either the schematic viewer or the model editor.

Summary

The Knowledge-based Autonomous Test Engineer (KATE) system is a model-based software system which has been developed in the Artificial Intelligence Laboratory at the Kennedy Space Center over the last decade. It is designed for monitoring, fault detection, and control of launch vehicles and ground support systems. In order to bring KATE to the level of performance, functionality, and integratability needed for firing room applications, efforts are currently under way to implement KATE in the C++ programming language using an X-windows interface. The version of KATE currently under development is called KATE-C; however, we will omit this distinction and refer to KATE generically.

This report commences with a brief introduction to the fundamental principles behind the operation of KATE. Emphasis is placed on the structure and importance of KATE's knowledge-base. We then describe two programs which have been designed and added to the collection of tools comprising what is called the KATE toolbox. The first tool, called the schematic viewer, gives the KATE user the capability to view digitized schematic drawings in the KATE environment. This tool was designed to illustrate the potential for integrating real-life schematics into the operation of KATE. The second tool, called the model editor, gives the KATE model builder a tool for creating and editing knowledge base files. Without such a tool the KATE model builder has to create and edit knowledge bases outside of the KATE environment using a traditional text editor. The body of this report discusses design and implementation issues having to do with these two software tools which have been designed and prototyped this summer. It will be useful to anyone maintaining or extending either the schematic viewer or the model editor.

TABLE OF CONTENTS

<u>Section</u>	<u>Title</u>
I	AN INTRODUCTION TO MODEL-BASED REASONING
1.1	Basic Principles
1.2	Knowledge-bases
1.2.1	The KATE Knowledge-base
1.2.1.1	High level system knowledge
1.2.1.2	Middle level system knowledge
1.2.1.3	Low level system knowledge
II	THE SOFTWARE ENVIRONMENT
2.1	Programming Language and Graphics Interface
2.2	Frame Utilities
III	THE SCHEMATIC VIEWER FRAME UTILITY
3.1	Requirements
3.2	Achievements
3.3	Enhancements
IV	THE MODEL EDITOR
4.1	Motivation
4.2	Design
4.3	Requirements
4.3.1	The Flatfile Editor
4.3.2	The Middle Level Editor
4.4	Implementation of the Flatfile Editor
4.4.1	Overview
4.4.2	Operation
4.4.3	Status
V	REVIEW AND RECOMMENDATIONS
5.1	Concerns
5.2	Other Remarks

REFERENCES

LIST OF ILLUSTRATIONS

<u>Figure</u>	<u>Title</u>
1-1	Overview of Model-based Reasoning
3-1	The Schematic Viewer within the KATE User Interface
4-1	Preliminary Design for a Model Editing Environment
4-2	Revised Design for the Model Editor
4-3	The ECLSS Flatfile in the Flatfile Editor

AN INTRODUCTION TO MODEL-BASED REASONING

1.1 BASIC PRINCIPLES

The premise of model-based reasoning is extremely simple. A computer simulation model of a physical system is constructed from a knowledge-base representing the components of the system and their interconnections. The physical system, which must have numerous sensors, is put into operation. As the physical system operates, sensor readings are compared to their predicted values from the simulation model. As long as there are no significant discrepancies between predicted values and actual sensor readings, nothing is done. When a significant discrepancy occurs, the model-based reasoning system carries out whatever actions are necessary to alert a human that a problem has occurred. This aspect of model-based reasoning is simply known as monitoring.

If model-based reasoning systems were only capable of monitoring, they would be of limited utility. Fortunately, model-based reasoning systems such as KSC's Knowledge-based Autonomous Test Engineer (KATE) have other powerful capabilities. Among the most interesting is failure diagnosis. Once a significant discrepancy has been identified in the monitoring stage, KATE utilizes its internal representation of the physical system in an effort to identify failures which could have led to the conflict between predicted values and actual sensor values. A significant difference between this reasoning process and traditional process control techniques is KATE's ability to include sensors themselves in the diagnostic process. A high-level overview of the monitoring process and its relationship to diagnosis is illustrated in Figure 1-1.

In addition to their ability to monitor and diagnose complex systems, model-based reasoning systems such as KATE have the potential for several other very useful functions. If the computer has the ability to issue commands to the physical system, it should be possible to describe a desired state of the physical system and have the reasoning system determine what commands can be issued to achieve that state. If the physical system has redundant pathways and components, as is frequently the case in NASA systems, the model-based system can often determine how to continue operation of the physical system after some component or components have failed. It is also possible to have such a model-based reasoning system construct an explanation of the steps taken to identify a failed component or to achieve a specific objective.

In addition to their operational use, model-based reasoning systems have great potential as training tools. An instructor or student can create failure scenarios in the simulation environment to test the student's ability to respond to almost any failure in the actual hardware. Another potential use for model-based reasoning systems is to

determine the adequacy of the sensors in a complex system before it is built. For a more in-depth introduction to model-based reasoning see [1].

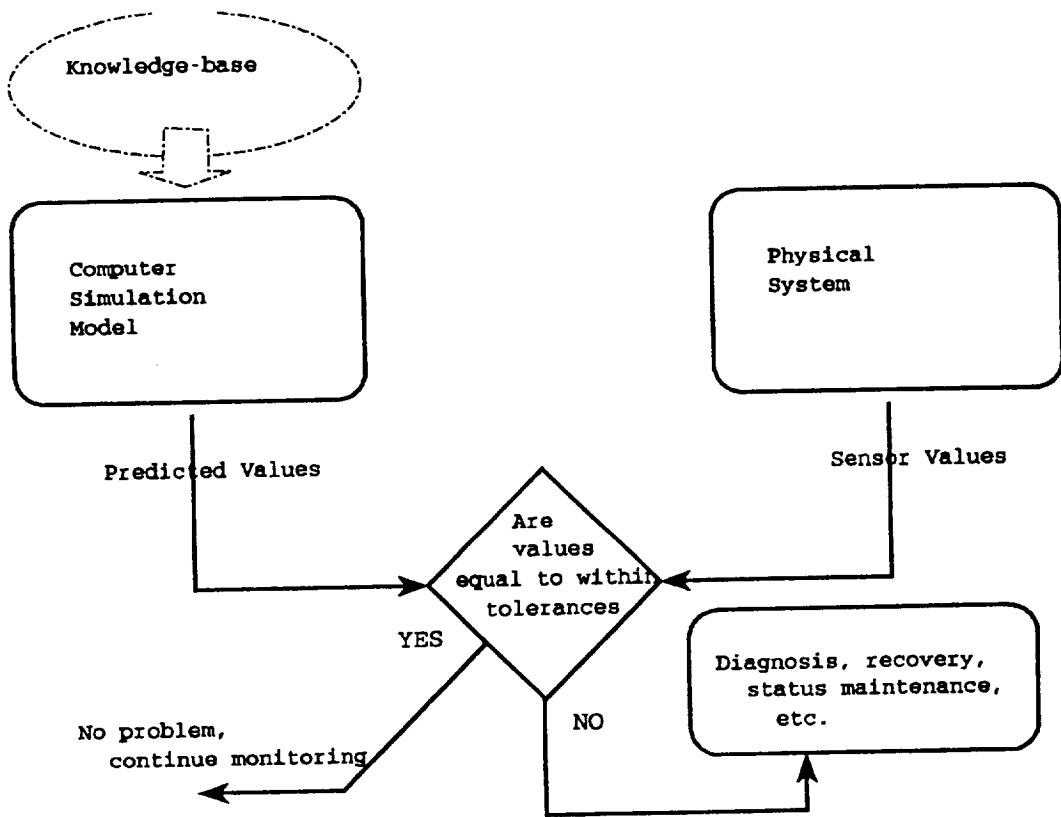


Figure 1-1. Overview of Model-based Reasoning

1.2 KNOWLEDGE-BASES

In order for a model-based reasoning system to function, it must have information about the structure and operation of the physical system to be modeled. We call such information about the real world the system's knowledge-base. As characterized in [2] for the KATE system:

The KATE knowledge base contains vital information about the physical system that KATE is controlling or monitoring. This information is the raw material used by KATE to construct a simulation model that mimics the system's structure and function. Objects in the model have a one-to-one correspondence with parameters, commands, sensors, and other components in the physical system. The knowledge-base is referenced by KATE in the

same way that schematic diagrams and operating specifications are used by system engineers.

1.2.1 THE KATE KNOWLEDGE-BASE. In order to lay the groundwork for topics discussed later in this report, we elaborate upon the organization of KATE's knowledge-base—a three-level hierarchical structure. Such hierarchical structures are typical of the organization of knowledge-bases used for model-based reasoning systems.

1.2.1.1 High level system knowledge. The so-called "top-level" of KATE's knowledge base represents information about very broad classes of system components. For the systems with which KATE is currently used, these classes are commands, measurements, components, pseudo objects, display function designators and so called synchronization objects. For operational efficiency, generic knowledge about the structure and function of these high level classes of objects is hard-coded into the C++ implementation of KATE. This means that changes to KATE's top-level knowledge of system component classes requires possible extensive modifications to the source code. Fortunately, such changes occur infrequently.

1.2.1.2 Middle level system knowledge. The so-called "mid-level" of KATE's knowledge-base represents information about specific types of system components. For example, this class contains knowledge about components such as pumps, relays, valves, and tanks in addition to pseudo objects such as pressures and admittances. Each middle level is an example of some top-level class described above and inherits certain properties from the top-level class. Again, for efficiency reasons, the mid-level of KATE's knowledge-base is represented in C++ header and source code files which are compiled into the corpus of KATE at compile time. However, modifications to the content of this level have no effect on the body of the KATE system, only the classes of components available for subsequent modeling. This level of the knowledge-base has a regular, predictable, organization and syntax which makes it easy to add to.

1.2.1.3 Low level system knowledge. The lowest level of KATE's knowledge-base is stored in what are referred to as "flatfiles". This is the information about the actual physical components in a system being modeled. Each object at this level is an instance of some mid-level class and inherits properties from that abstract class, which inherits knowledge from its parent class.

The flatfiles representing low-level knowledge are textfiles (ASCII files) with a well-defined keyword-based syntax. They are read by KATE at run-time in order to construct an internal representation of the physical system to be modeled.

II

THE SOFTWARE ENVIRONMENT

2.1 PROGRAMMING LANGUAGE AND GRAPHICS INTERFACE

Several implementations of the general principles underlying KATE have been carried out over the past decade under various titles, for example, the LOX [3] and ECS[2] systems. These systems were originally developed in LISP—the traditional language for rapid prototyping of AI systems—as a proof of concept.

In order to bring KATE to the level of performance and functionality needed for firing room applications, current efforts focus on simultaneously porting KATE to the C++ programming language while advancing its capabilities. C++ is an object-oriented language developed at Bell Laboratories. It is a derivative of the C programming language with features which encourage the writing of modular, reusable code. The work described in this report required the investigator to achieve proficiency in the C++ programming language.

For a software system as complex as KATE to be accepted and used by firing room personnel it must have an excellent user interface. At the same time, the KATE environment must be accessible from several different hardware platforms. In order to achieve both of these goals, the user interface for all KATE modules currently under development must adhere to the Motif/X-Windows graphics user interface standards. As with C++, it was necessary for the investigator to develop proficiency with these tools in order to carry out the work described herein.

2.2 FRAME UTILITIES

In an effort to standardize the user interface of KATE, user interface components are implemented as so-called "frame utilities". A frame utility interacts with the main body of KATE in a specific way and inherits graphic and functional properties from the KATE user interface environment. Working with frame utilities provides conveniences for the implementor but, at the same time, places constraints on the appearance and operation of the code being developed.

III

THE SCHEMATIC VIEWER FRAME UTILITY

3.1 REQUIREMENTS

The first KATE utility developed this summer is known as the Schematic Viewer Frame Utility (SVFU). The goal was to add to the set of tools available in the KATE environment a way for the KATE user to view schematic drawings. Specific requirements for this program were:

- o The SVFU must operate as a frame utility.
- o The SVFU should allow a KATE user to view digitized schematic drawings within the KATE environment.
- o The SVFU should allow a KATE user to view system overview drawings created using a "painting" program.
- o The SVFU should be useful for demonstrating the potential for accessing digitized information within the KATE environment.
- o The SVFU should be designed and implemented in such a way that KATE model and hardware values can be superimposed on the drawings while KATE is executing.

3.2 ACHIEVEMENTS

A frame utility was developed which met the requirements listed above. Figure 2-1 shows the appearance of the SVFU within the KATE user interface. In this example a schematic drawing was digitized and loaded into the SVFU. Any image file in either the X-windows bitmap (.xbm) or X-windows pixmap (.xpm) format can be loaded and viewed. Using the Portable Pixmap library from MIT, almost every graphics and image file format can be converted to a bitmap or pixmap which can then be viewed in KATE. The SVFU allows the user to load multiple image files and to then switch among them quickly and easily. The utility has been demonstrated to potential KATE users and has been well received. There has been discussion of the possibility of having digitized schematic drawings stored on compact disc for loading into the SVFU.

3.3 ENHANCEMENTS

Work has begun on superimposing KATE values on top of drawings in the SVFU. This should be a very useful tool for console operators using KATE.

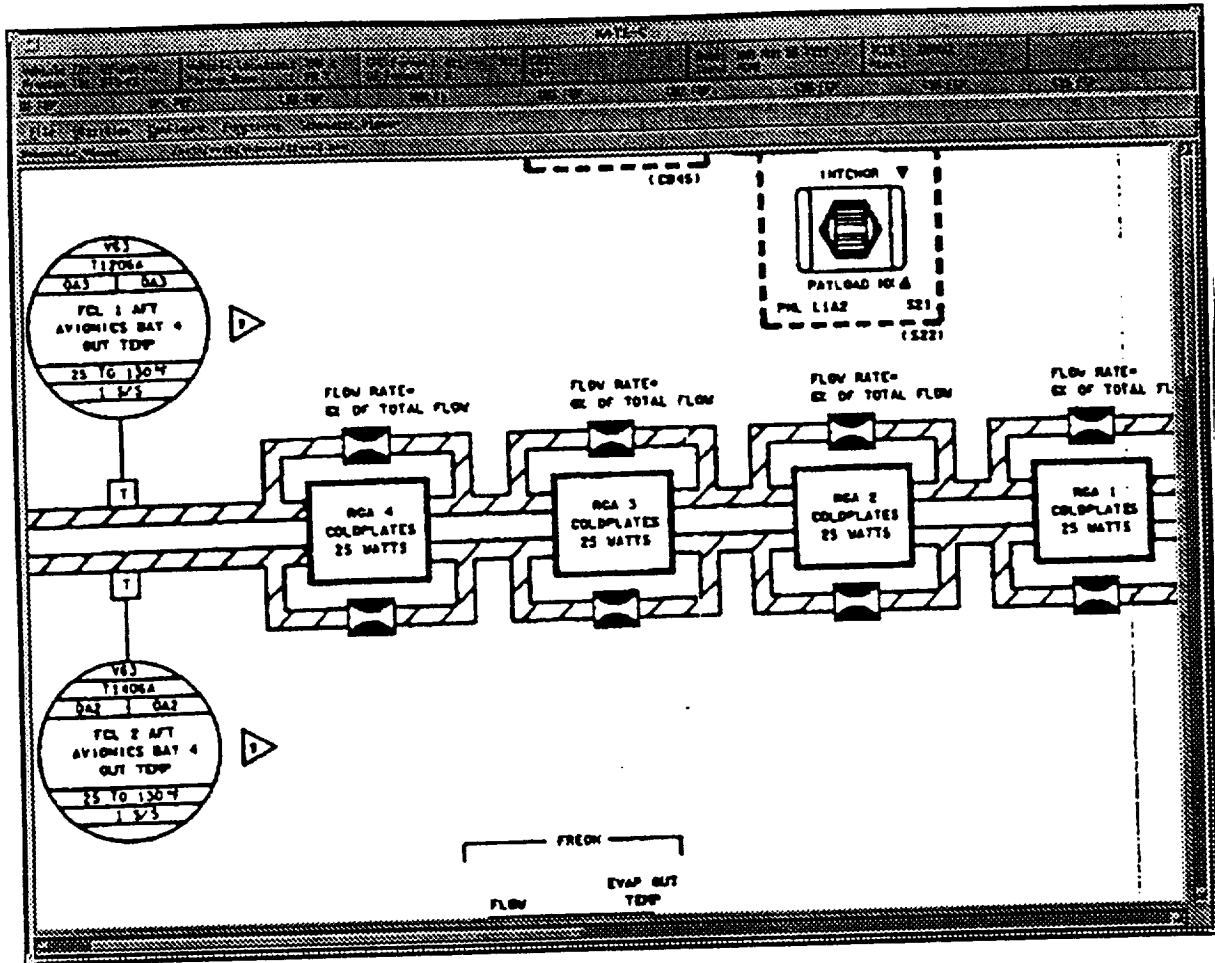


Figure 3-1. The Schematic Viewer within the Kate User Interface

Loading large files into the SVFU is time consuming. The files loaded into the SVFU need to be in a compressed format. This would greatly reduce the storage space required for storing large collections of digitized schematics and would improve the time to load a schematic into the system. This modification will require the SVFU to read image files stored in at least one widely used compression format. The ability to read various image file formats, compressed and uncompressed, would be ideal.

The initial version of the SVFU does not support the quick-loading feature expected of a frame utility. Several menu buttons for such things as "Help" options are not functional. The program has not been thoroughly checked for potential memory leaks. These are all issues which can be resolved quickly.

THE MODEL EDITOR

4.1 MOTIVATION

Constructing a KATE knowledge-base is a critical and time-consuming task. Starting with the pre-defined top-level KATE classes and a pre-defined library of mid-level components, the model builder proceeds along the following lines:

- o Gather together schematics and engineering documents for the physical system to be modeled.
- o Study the target system to gain an understanding of its principal components and their interactions.
- o Determine whether or not the existing middle level component classes are adequate for the system to be modeled. If not, add new component classes. This requires at least some C++ code to be written.
- o Construct a flatfile for the physical components, commands and measurements in the physical system.
- o Specify the interconnections among the components in the flatfile.
- o Add to the pseudo objects to represent logical functions of groups of components in the system.

There are currently no tools to assist the KATE model builder. All work is done using a text editor and there is no way for the model builder to view the model under construction except as a collection of text files. This investigator undertook to design and implement sophisticated, graphically oriented tools to assist in the process of constructing KATE knowledge-bases.

4.2 DESIGN

With the benefit of preliminary design documents and discussions with INET engineers an initial design for a knowledge-base editing environment was developed. This design is illustrated in Figure 4-1. The principal feature is that of two distinct editing environments, one for the flatfile level of a model and another for the middle level components. Henceforth, we shall refer to those portions of an editing environment which have to do with editing the flatfile level of a model as the flatfile editor (FFE) and to those portions which pertain to editing the middle level components as the middle level editor (MLE). In Figure 4-1 the FFE obtains a description of the middle level of a knowledge base by reading a file containing enough information about each middle level component to allow a flatfile instance of a component class to be edited intelligently. The MLE is responsible for editing and updating not only this middle level description file but also the mid-level C++ header and code files.

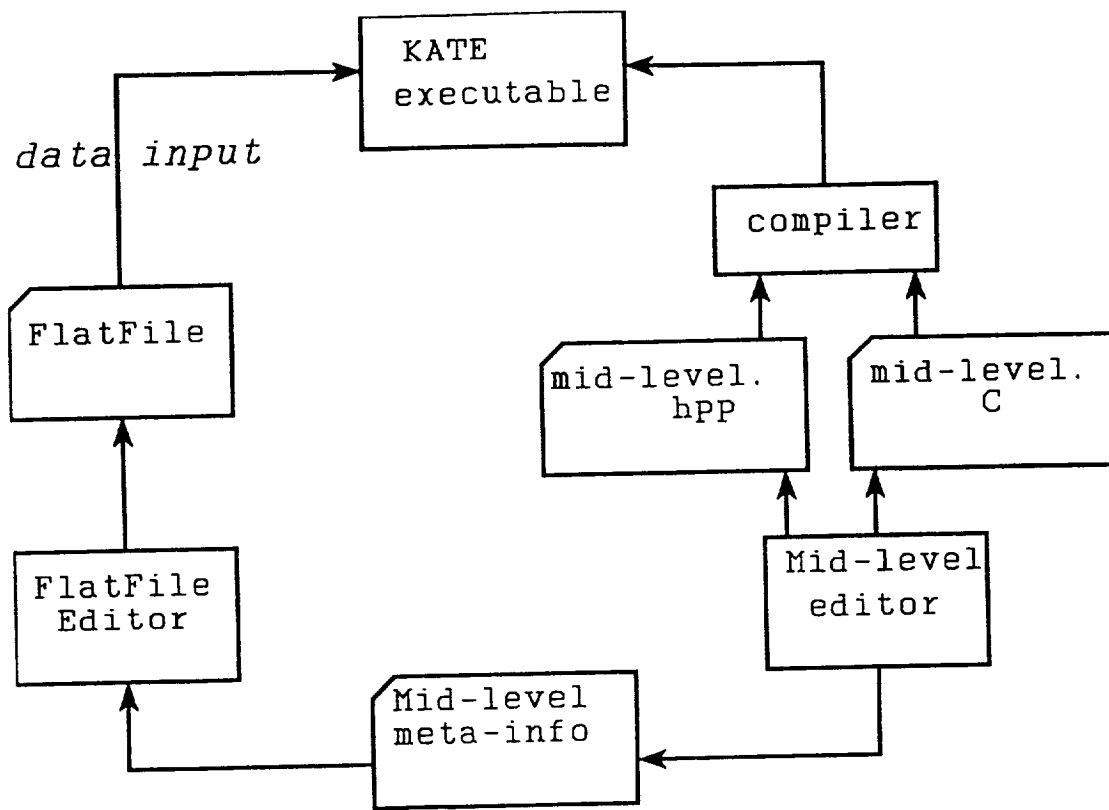


Figure 4-1. Preliminary Design for a Model Editing Environment

Upon further study it was decided that it might be possible to eliminate the additional file of meta-level descriptions of the middle level components by gleaning the required information directly from the mid-level header (.hpp) files. This revised design, illustrated in Figure 4-2, was pursued for the remainder of the project.

4.3 REQUIREMENTS

4.3.1 THE FLATFILE EDITOR. With the assistance of existing design documents and with an overall design plan in hand, specific requirements for the FFE portion of the model editor were developed. They are:

- o The flatfile editor should be a frame utility. In the short term this requirement is not critical. However, there are two compelling reasons for having the flatfile editor tightly integrated into the KATE user interface. First, the simulation capability of KATE could be used to assist in the verification of a model while still under development. This could greatly enhance the ability of the model builder to create a thorough, accurate model

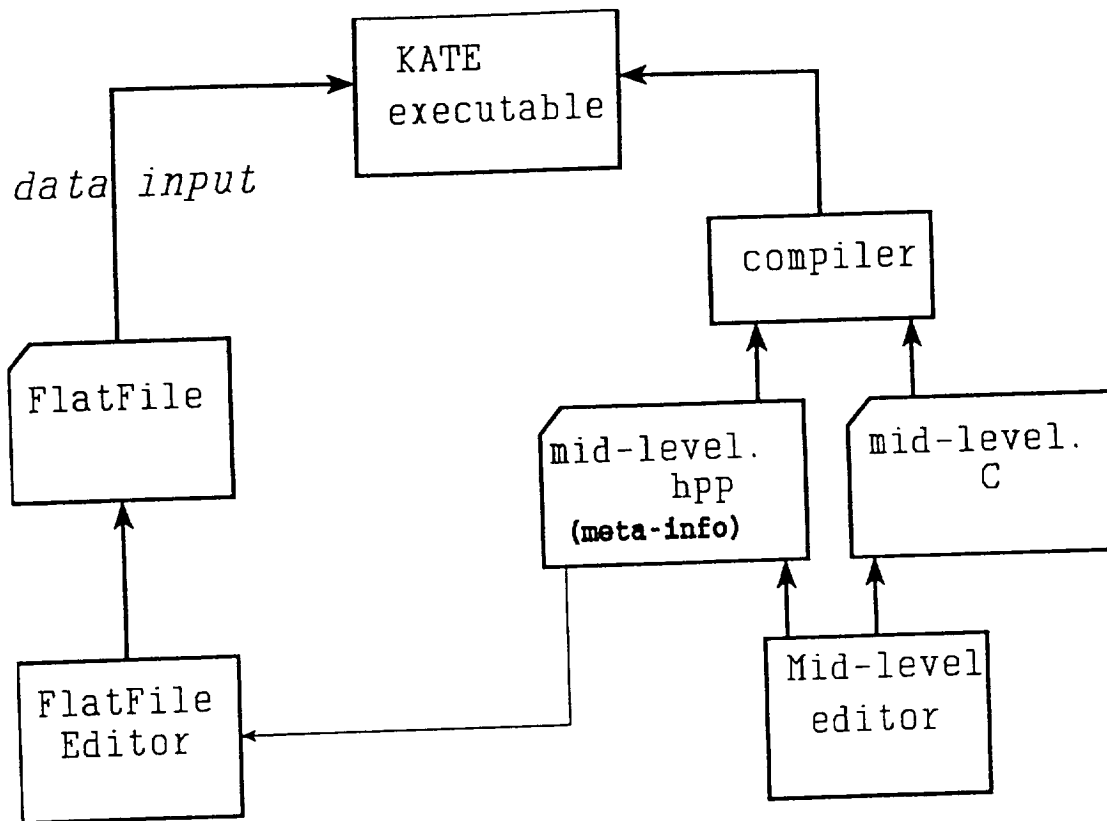


Figure 4-2. Revised Design for the Model Editor

- without repeatedly having to reload and test the model. On the other hand, a tightly integrated environment could be used to make small corrections to a model already loaded into KATE.
- o The flatfile editor should react transparently and correctly to changes in the middle level of the knowledge base. Practically speaking, this means that the connection shown in Figure 4-2 between the middle level header file (mid-level.hpp) and the flatfile editor should be the only connection needed between knowledge about the middle level of a knowledge base and the flatfile editor.
 - o The flatfile editor should be easy to modify when the top level of the knowledge base changes. Although any good software should be easy to modify we are making a slightly more specific requirement here.

4.3.2 THE MIDDLE LEVEL EDITOR. The only specific requirement for the middle level editor is that it should edit both the C++ header and C++ source code files composing the middle level of a knowledge base. Given the software organization of KATE in C++ it is not necessary for this editor to be a tightly integrated part of the

KATE environment. Any change to the middle level of the knowledge base requires the entire KATE system to be recompiled and linked. At this time no further work has been done on this portion of the editing environment. The remainder of this report deals exclusively with the flatfile editor.

4.4 IMPLEMENTATION OF THE FLATFILE EDITOR

A significant portion of the flatfile editor has been implemented. The following discussion will describe the major components of that program at a level which would be useful to someone endeavoring to continue development of this software.

4.4.1 OVERVIEW. At the highest level, the FFE works as a frame utility. The principal C++ class used is called the `FFFrameUtilityClass`. Instances of this class manage three principal subdivisions of the screen along with a list of structures related to the individual files loaded into the editor at any point in time.

The FFE's portion of the screen is subdivided into a menu bar, a list area and a drawing area. The menu bar is used to select editing commands and options for the editor. The list area is used to list the names of the objects in the flatfile currently being edited. Several options have been proposed for restricting and organizing the content of this list window. The drawing area can currently only be used to display a tree-like representation of the flatfile being edited. This drawing area conceivably could be used to display an icon-based representation of the flatfile being edited.

The `FFFrameUtilityClass` also maintains a list called the `FlatFileList`. Each item on this list is itself an instance of a C++ class containing information specific to a single flatfile. Along with obvious entries such as the name of the flatfile, each instance contains several entries including a list of the objects in the flatfile and a flag indicating whether or not any editing changes have been made since the file was saved to disk.

Two other important classes of objects are the `TopLevelClass` and the `MidLevelClass` objects. The `TopLevelClass` represents useful information about the editing of top-level KATE classes. The `MidLevelClass` represents similar information for each mid-level component class. For example, the `MidLevelClass` instance for a Pump would contain, among many other items, the names of the input values expected for a Pump. The information for the `MidLevelClass` is gathered almost exclusively at run time by scanning through a mid-level header file. Interestingly enough, the C code to carry out this scanning is generated by the UNIX tool *lex*. If the syntax of the middle level KATE classes changes, the code which scans the mid-level header file can be modified by making simple changes to the *lex* file describing the structure of header files.

4.4.2 OPERATION. When the FFE is loaded, a pane is obtained within the KATE runtime user interface. At this point the user has two choices—load an existing flatfile

for editing or create a new flatfile from scratch. If the user chooses to load an existing file, the file is read into a KATE knowledge-base structure and this structure is then used to create a list of the flatfile objects for editing. A list of the flatfile contents is shown in the list area and a tree diagram is automatically generated. The user can subsequently either select items for editing or add additional items to the flatfile under consideration. Whenever the user clicks on an item in either the list window or the tree window, an edit template is generated containing known information about the component. The user can then modify the item by typing in the template. Similarly, if the user wants to add items to a flatfile, he obtains a list of top-level classes and from a class obtains a list of all the known mid-level components for that class. Selecting one of these mid-level classes results in a blank template for the user to edit. In both editing situations, the editing templates popped up are constructed dynamically to correspond to the specific class of object being edited. The information to do this is derived from the TopLevelClass and MidLevelClass mentioned previously.

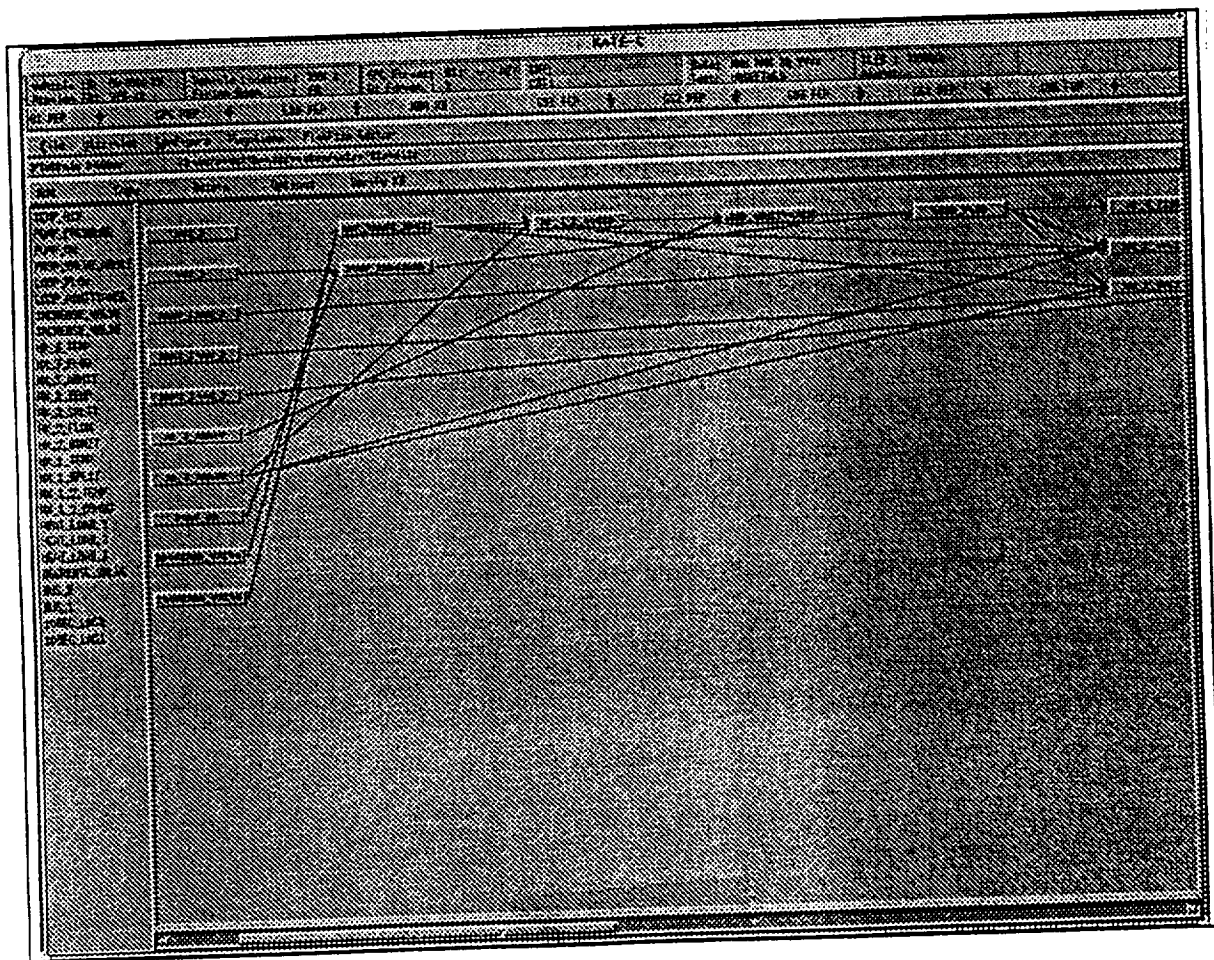


Figure 4-3. The ECLSS Flatfile in the Flatfile Editor

At any time during the editing process the user can elect to save the edited file, switch to editing a different file, or use other KATE utilities such as the Schematic Viewer Frame Utility. Figure 4-3 illustrates an editing session. The names of the objects in the flatfile are listed on the left and a tree drawing is on the right. The menu bar can be seen just above these areas.

4.4.3 STATUS. Much of the development effort this summer has been directed toward constructing a well-conceived foundation upon which the flatfile editor can be implemented. The major functions of the editor have been completed but many details remain. The automatic tree drawing capability of the FFE is in its infancy. There are probably a number of special cases in which the popup editing templates do not have exactly the correct editing fields displayed. It was not possible to incorporate the newly defined Synchronization Objects added to KATE this Summer. Also, the class of so-called TimeDependentObjects has not been dealt with. The Verify option on the menu bar is not implemented.

REVIEW AND RECOMMENDATIONS

5.1 CONCERNS

For the most part the design illustrated in Figure 4-2 has proven to be a satisfactory approach to the implementation of the FFE. We believe this organization will prove the easiest and most flexible editing organization to maintain as KATE evolves. There are however some significant problems which should be addressed if the FFE is to have long-term viability.

First of all, the scanning of the middle level header file (`mid-level.hpp`) is done in a very *ad hoc* manner. The structure of several header files was examined and certain rules of thumb about their structure were adequate to implement a scanner using *lex*. However there is no reason that a model builder constructing a middle level header file would necessarily stick with the same format. Any significant change to the structure of the middle level header files will necessitate updates to the *lex* grammar file. For long-term viability of the FFE an actual grammar should be constructed for the header files and a genuine parser should be developed. The middle level knowledge base is also being broken up into multiple files. This will require some immediate minor restructuring of the FFE to accomodate this change.

The second major concern is that the current prototype FFE does not do a good job with knowledge base constructs outside of its range of knowledge. For example, if a flatfile containing Synchronization objects is read into the FFE and then written back to disk, those Synchronization objects will not be written back to disk because as far as the FFE is concerned, they never existed in its realm of knowledge. A similar statement applies to "comment" fields within a flatfile which is read, processed, and written back to disk. Currently we do not envision a simple solution to this problem since the FFE actually works solely with an internal representation of the flatfile and does not retain any other type of "image" of the file being edited.

5.2 OTHER REMARKS

A significant amount of work has been completed toward the design and implementation of a knowledge-base editing environment. The FFE integrates very well into the KATE user interface environment and will, when complete, be of significant benefit to KATE model builders. Except for the concerns mentioned above, we expect development to continue along the lines in which it is currently headed. Much of the remaining work should be in the nature of flushing out and completing modules already partially developed. With the exception of the icon drawing components, the FFE infrastructure is complete.

REFERENCES

- [1] Steven L. Fulton and Charles O. Pepe, "An Introduction to Model-Based Reasoning", *AI Expert*, January 1990, pp. 48-55.
- [2] Charles O. Pepe, et. al., *KATE - A Project Overview and Software Description*, Boeing Aerospace Report, Boeing Aerospace Operations, Mail Stop FA-78, Kennedy Space Center, Florida.
- [3] Delaune, C.I., Scarl, E.A., and Jamieson, J.R., "A Monitor and Diagnosis Program for the Shuttle Liquid Oxygen Loading Operation", *Proceedings of the First Annual Workshop on Robotics and Expert Systems*, Johnson Space Center, Houston, TX, June 1985.

1994018775
N94-28848

1993 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA

18-47
197204
P-33

HURRICANE RISK ASSESSMENT: ROLLBACK OR RIDE OUT

PREPARED BY:	Mr. Richard A. Wohlman
ACADEMIC RANK:	Instructor
UNIVERSITY AND DEPARTMENT:	Western Carolina University Department of Mathematics and Computer Science
NASA/KSC	
DIVISION:	Advanced Projects Office
BRANCH:	Atmospheric Sciences
NASA COLLEAGUE:	Jim Nicholson
DATE:	August 8, 1993
CONTRACT NUMBER:	University of Central Florida NASA-NGT-60002 Supplement: 11

ACKNOWLEDGEMENTS

Just a moment to thank some folks-

To the UCF people, Ray and Kari, pretty smooth operation you two.

To my NASA colleague Jim Nicholson. Even after the move to the other end of the world, I managed to get something done!

To Rich Hall for his knowledge and assistance in Turbo Pascal.

To all of the other Faculty Fellows for good fellowship and encouragement and toleration!

To my fellow paraplane pilots, Martha, Greg and Keith, when are we gonna do this again?

And to my special friends and troublemakers Keith (yeah, I know he shows up twice) and Harry a truly sincere thanks for the boost to not only my ego but to fanning the fires of intellectual pursuits.

This has been two of the most exciting and enjoyable summers of my life, and I congratulate all who have had the opportunity to participate and contribute. My best wishes go to each and every one of you. Thank you again.

RAW
8/8/93

ABSTRACT

Winds in excess of 74.5 knots could cause severe damage to a space shuttle on the launch pad. Current plans exist for rollback to the Vehicle Assembly Building, but require 48 hour leadtime to implement. Decisions based upon cost/loss are evaluated to ascertain whether predetermined forecast probabilities for rollback/rideout decisions can be made far in advance of hurricane season for use in decision making.

SUMMARY

KSC lies in the hurricane belt, and there exists the possibility that a hurricane with winds of at least 74.5 knots might strike the Center. Should a Space Shuttle be on the launch pad during winds of this nature, severe damage up to and including the loss of the vehicle could occur. Therefore, in the event that winds associated with a hurricane are forecast to be in excess of 74.5 knots, procedures are in place to roll the mobile launch platform with the shuttle on top back to the relative safety of the Vehicle Assembly Building.

This plan requires 48 notification in the advance of the onset of the winds, with the last eight of those hours being devoted to the actual movement of the shuttle from the pad to the VAB.

Managers make the rollback/rideout decision based upon National Hurricane Center forecasts, which give numerical percentages of strike. At present, there is no predetermined probability value which triggers the rollback decision, and at present the decision is extremely subjective.

With a Bayesian analysis of sequential events and their modification to current probabilities, along with a cost versus loss study, an objective value from which to base a decision might be reached. In addition, examination of past information concerning similar hurricane paths might be useful in fine tuning the probability forecasts

TABLE OF CONTENTS

SECTION I INTRODUCTION	1
1.1 BACKGROUND AND PROBLEM IDENTIFICATION	2
1.2 PREVIOUS STUDIES	2
1.3 NEW TOOLS FOR THE DECISION MAKER	5
SECTION II DECISION MAKING	7
2.1 SIMPLE COST VERSUS LOSS	7
2.2 SEQUENTIAL DECISIONS	8
SECTION III COMPUTER TRACK ANALYSIS	12
3.1 PROGRAM DEVELOPMENT AND REQUIREMENTS	12
3.2 DATA AND SUBROUTINES	13
3.3 PROGRAM PICKTRAK.PAS	14
3.4 PROGRAM PLOTYR.PAS	14
IV. CONCLUDING REMARKS	15
APPENDICES	16

SECTION I INTRODUCTION

1.1 BACKGROUND AND PROBLEM IDENTIFICATION

The Kennedy Space Center, located on the east coast of the Florida peninsula has been subjected to 64 tropical storms in years 1886 through 1992. Of these, 26 were of hurricane strength. Because winds in excess of 74 knots would cause extreme damage to a shuttle on the launch pad, protective measures in the form of rolling the shuttle back to the Vehicle Assembly Building are planned in the event of such winds. That rollback decision has to be made by shuttle managers in sufficient time to allow the actual rollback, and the current timeline (appendix 1) requires that decision be made 48 hours prior to the onset of damaging winds.

The primary responsibility of hurricane forecasting lies with the National Hurricane Center (NHC) who distributes forecasts via alerts, watches and warnings to the various meteorological stations around the country. Weather support to the Kennedy Space Center is provided directly by personnel at the Cape Canaveral Air Force Station Forecast Facility. These meteorologists provide the manager with not only forecasts, but rationale, track error analysis, and model confidence.

While hurricane forecasting has become quite sophisticated in recent years, there are still sufficient forecast positioning errors to preclude a cut and dried decision to rollback or to ride out. For example, a recent National Hurricane Center study shows RMS position errors to be about 170 nautical miles at the 24 hour point. As the radius of maximum winds in a 95 knot hurricane is less than 25 nautical miles, a storm of that nature moving inland at 70 miles to the north or south of the center would produce winds well below the maximum allowable. Thus, even though the forecast might place the storm directly atop KSC in 24 hours, managers cannot use that forecast as the sole source for decision making.

Because rolling the shuttle back to the Vehicle Assembly Building is such a costly event, both in terms of launch delays, shuttle processing schedules, and redundant operations, the rollback option is taken only after much deliberation. And, because hurricane strikes at the cape are relatively rare occurrences, managers have not been able to create any

kind of track record of past decisions and the criterion upon which those decisions were based. Managers rely on advice provided by the weather support personnel which in turn is largely based on Hurricane Center forecasts. One problem arises with the probability forecast itself which is given as a probability of landfall at a certain area. Does the manager protect if the probability is 5%? What about 10%? At present, there is no black and white set of rules for making the rollback decision based strictly upon the strike probability.

The situation is further muddled by the aforementioned forecast position errors. Obviously, the farther into the future you forecast, the larger the position errors. This necessarily produces low strike probabilities for specific locations due to the uncertainties involved.

Simply put, the manager needs some method to yield a set of guidelines for use in making the rollback/rideout decision at the 48 hour point.

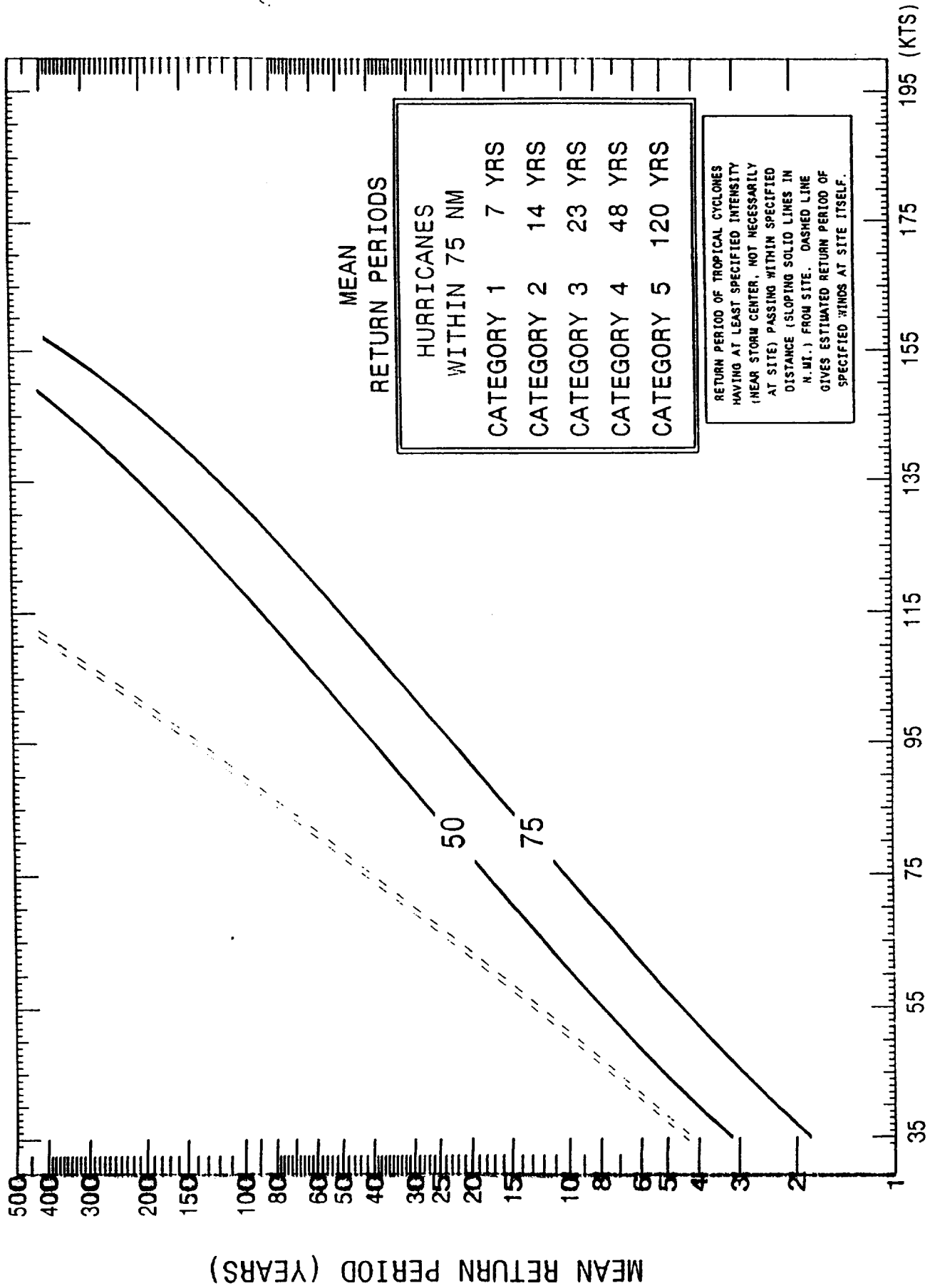
1.2 PREVIOUS STUDIES

Investigation into the work done in the area was conducted during the summer of 1992 here at the Kennedy Space Center. A detailed breakdown of several approaches is discussed in the report of that work. The bulk of the data was found in the NOAA Technical Memorandum NWS NHC 38 by Charles J. Neumann titled 'The National Hurricane Center Risk Analysis Program (HURISK)'.

HURISK is a computer program developed to find statistical data concerning, strangely enough, hurricane risk assessment. When conducted for a specific location, it provides data about the frequency, strength and movement of tropical storms through a circle 75 nautical miles in radius around the selected site. Of particular interest are the data computed for the mean return periods (chart 1) of tropical storms with winds of a specified amount, and the probability of at least x storms passing within 75 nautical miles of KSC over n consecutive years (chart 2). This second chart is important as it also yields the probability of going for n consecutive years without having a storm of a specified category. See the source document for detailed explanation.

In addition, computerized tropical storm track data were obtained which contained information similar to that found in the NOAA Historical Climatology Series 6-2, Tropical Cyclones of the North Atlantic Ocean, 1871-1986'. These data were examined in a rather cursory

SITE: CAPE KENNEDY, FL



PROBABILITY OF AT LEAST X TROPICAL CYCLONES (≥ 64 KNOTS) PASSING WITHIN 75 N.M.I. OF SPECIFIED SITE OVER N CONSECUTIVE YEARS

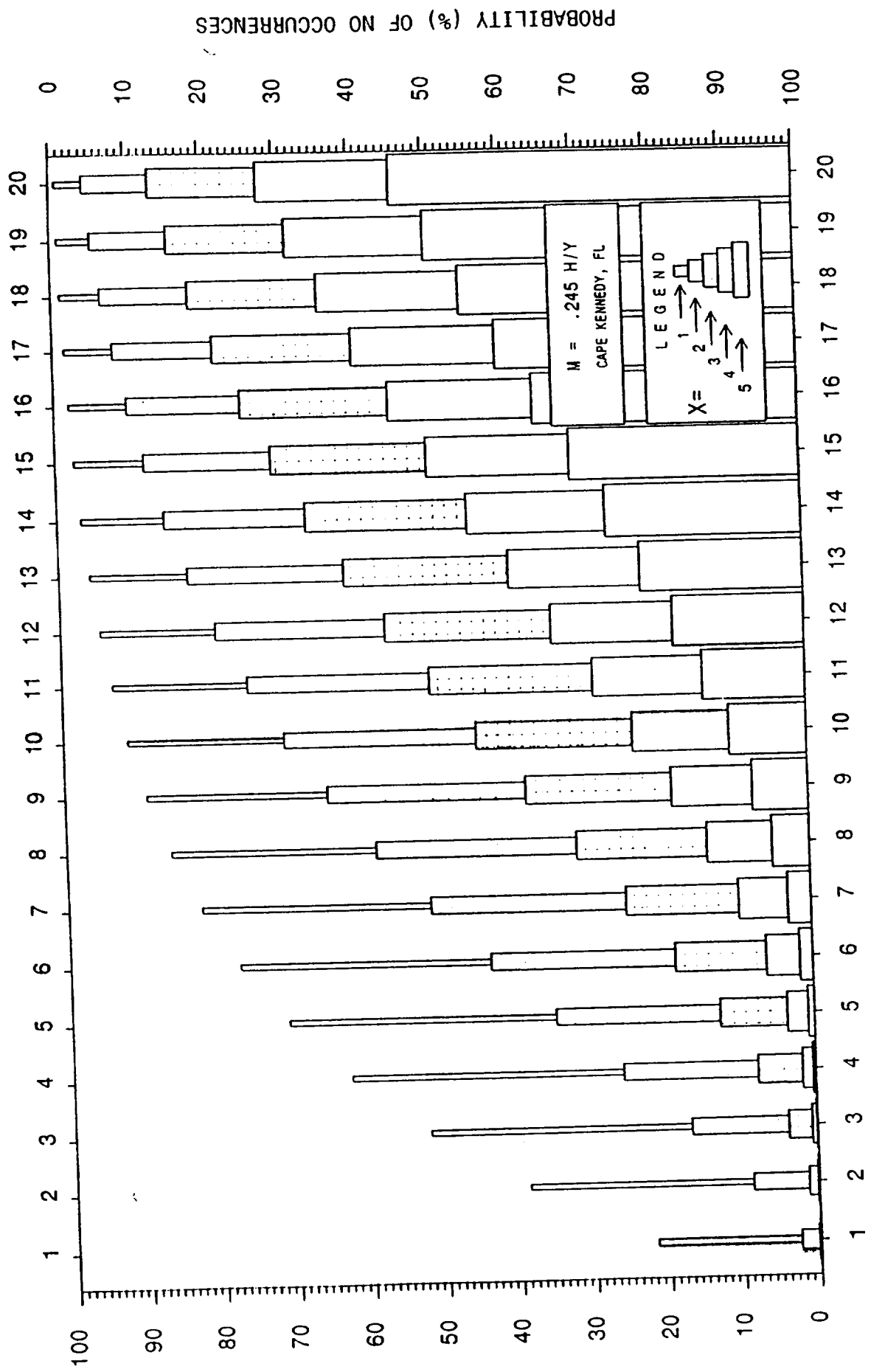


CHART 2

manner as rudimentary programs were produced to select specific storm data. It was concluded that these data should be more closely examined in a graphical format, and that by selecting storms similar in nature to the ongoing storm, it might be possible to detect trends from statistical fields generated from those similar storms.

1.3 NEW TOOLS FOR THE DECISION MAKER

One approach to giving the decision maker a go/no-go kind of decision is by using a cost versus loss analysis. Typically, this kind of analysis is used by managers in determining whether to initiate some kind of protection scheme strictly from a monetary point of view. This method determines the most cost effective course of action based solely on the dollars involved, and ignores the cost of public opinion, loss of capability, and the downstream effects of the loss itself. Additionally, by applying a Bayesian analysis to the probability of having a damaging hurricane throughout the season, new thresholds can be developed which yield probabilities for using the cost versus loss decision making process.

In the event that protection methods are initiated, costs can be considered to be delays in the shuttle flow process. In fact, these delays are not simply the time it takes to enact some protection to the shuttle, but must include the time that it takes to bring the shuttle to the same place in the processing timeline that it was in prior to initiating that protection. The very nature of the differences in payload requirements themselves makes it difficult to obtain dollar values for the protection operations. For example, a spacelab mission which is loaded horizontally in the Orbiter Processing Facility would not need to be removed from the orbiter while still on the pad. However, a Tracking and Data Relay Satellite, which is inserted vertically into the payload bay once the shuttle moves to the launch pad might have to be removed from the orbiter prior to rolling the orbiter back to the VAB. For the purposes of this study, cost will be assumed to \$2 million per day¹.

In addition to the statistical study, a graphical package was under development to assist meteorologists in reviewing historical tropical storm data. The computer programs are designed to permit individual selection of storms based upon specified criterion, to allow the

¹ Because the costs are so variable, perhaps a better currency for decision making is days instead of dollars.

forecaster to create statistical data on storms which show similarities to the current storm. Several interesting problems needed to be overcome, and work continues in this area. Once complete, these should provide the forecaster the capability to modify the forecast based upon movement, and a statistical study of forecast and observed motion vectors of the current storm. At any rate, graphical representation of historical data is accomplished, and should help in storm track visualization.

SECTION II DECISION MAKING

2.1 SIMPLE COST VERSUS LOSS

The cost versus loss approach has been applied to decision making with respect to meteorological events for some time, as statistical methods for prediction commonly produce probabilities for the occurrence of a specific event. As is shown in the diagram below, a decision is made at a specific point in time to protect assets, or not to protect assets. The forecast event then either occurs or does not occur. Costs are simply the value in the currency chosen. These costs are usually given in dollar values, but can be man hours, days of delays, or some other currency. Losses are usually expressed in dollars, and represent the replacement cost of the item.

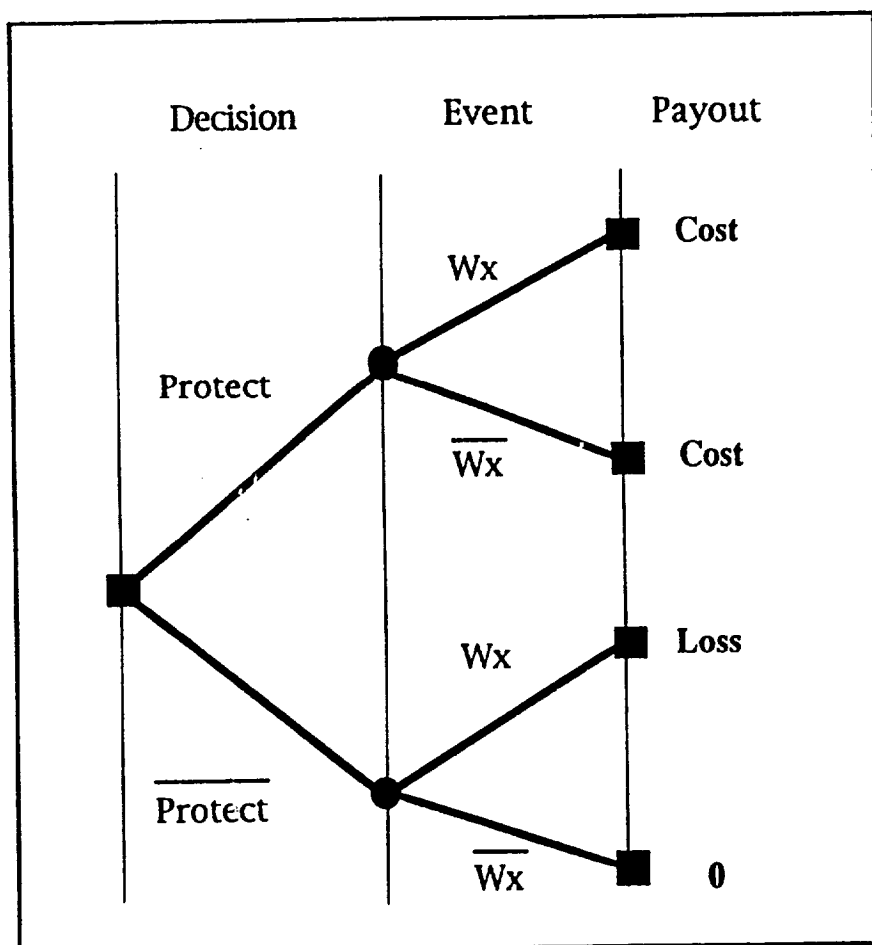


Diagram 1. Simplified Cost versus Loss decision making tree

The decision is made as follows:

$$\text{If } p < \frac{C}{L} \text{ then do not protect}$$

where p is the forecast probability of the meteorological event, C is the cost of protection, and L represents the loss value. Should the p be greater than the cost/loss ratio, then one would protect.

In the case of protecting the shuttle, assume that C/L is equal to $1/500$, a somewhat arbitrary figure². Studies³ have shown that hurricanes have passed within 75 nautical miles on 26 occasions since 1886. If one considers hurricane season to last from 1 June through November (183 days), and hurricanes to affect the area for only one day each, the probability of being hit on any one day in hurricane season is found to be 26 in 19398. In a probability to cost/loss comparison:

$$0.0013403 < 0.002$$

Thus, one would not protect on a day to day basis since the probability of the event is less the cost/loss ratio⁴. Had the probability been greater than the C/L , it would be unreasonable to expect that rollback should occur on each day of the season, there would be no launches at all!

2.2 SEQUENTIAL DECISIONS

The record of tropical storm activity from 1886 through 1991 shows that of 970 storms, 614 reached hurricane strength. This averages out to about 9 storms per year, with nearly 6 reaching hurricane strength. Rarely are hurricanes dependent upon one another, and their existence can be thought of as independent sequential events. The decision making tree for more than one event becomes complicated after very

² Although not unreasonable this number is optimistic at best. In the currency of days: one day of lost processing time to 500 days to replace a shuttle.

³ NOAA TM NWS NHC 38, reprinted Aug 91.

⁴ In reality, it is unreasonable to expect that a hurricane with winds in excess of 74.5 knots would spontaneously appear close enough to do damage over the course of 24 hours, so this daily probability is really only useful in determining climatological probabilities, and works out to about a one in four chance of having a hurricane in any given year.

few iterations. Thus, the example of such a tree is shown below with associated payouts for only two events.

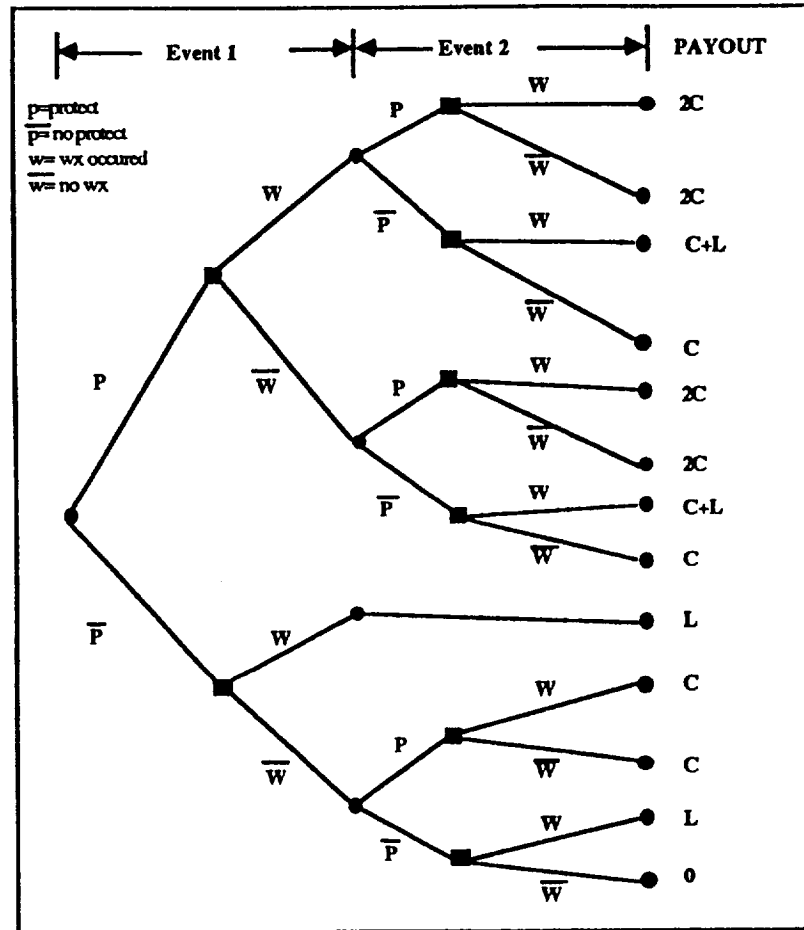


Diagram 2. Decision tree for 2 events

For each additional event, another four pronged branch is added to the tree at all locations where loss does not occur. For n events, the number of final nodes would be:

$$\text{number of nodes} = 3^n + \sum_{x=1}^n x^2$$

where the summation gives the number of those nodes which end in losses. Only one node yields no payouts, occurs when you never protect and never have the occurrence.

This decision tree is based upon having but one orbiter on a launch pad at a time, and assumes that after a loss shuttle operations would cease. In reality this is probably the case, but it could be argued that since the

loss of the orbiter is known to be due to a management decision as opposed to an engineering flaw, operations might continue shortly thereafter. Were this scenario to take place, the number of nodes would increase to:

$$\text{number of nodes} = 4^*$$

We shall use the first example to generate the decision probabilities. This is done by using 'backward induction' from the endpoint to the initial condition. In our example, consider the case where the $p > (C/L)$. On day 2, the decision maker should protect. Thus, if the decision maker decides to protect on day 1, the cost for the two days is $2C$. If the decision maker does not protect on day 1, he incurs a risk of pL of the loss and saves $-(1-p)C$ by not protecting. Thus, his expected payoff (or savings) will be:

$$-pL - (1-p)C$$

Remember, L and C are payouts, and are negative numbers which makes the overall number positive⁵. Thus, the decision maker should protect on day 1 if:

$$2C > -pL - (1-p)C$$

Which reduces to:

$$p > C / (L - C)$$

The manager should protect on the first day based on its new relationship. In tabular form the threshold values for each of the possible situations are as follows.

Case	Day 1	Day 2	Payoff	Optimal when
1	p	p	$-2C$	$C/(L-C) < p \leq 1$
2	p	not p	$-C-pL$	never ⁶
3	not p	p	$-pL-(1-p)C$	$C/L < p < C/(L-C)$
4	not p	not p	$-[1-(1-p)^2]L$	$p \leq C/L$

⁵ Unless the probability of the event is .5 in which case one should always protect.

⁶ Since the payoff for case 3 is less than or equal the payoff of case 2 for all cases, case 2 is never optimal.

If we assume values bases for cost and loss based upon dollar values of \$3 billion for loss and as mentioned above and \$2 million for cost of rollback per day with five lost days,⁷ for a cost of \$10 million. If the probability of strike is .01 the table looks like:

Case	Day 1	Day 2	Payoff	Optimal when
1	.01	.01	20	$0.0033444 < p \leq 1$
2	.01	.99	40	never ⁸
3	.99	.01	39.9	$0.003333 < p < 0.0033444$
4	.99	.99	59.7	$p \leq 0.0033333$

It becomes obvious that because of the large disparity between the cost of protection versus loss, protection should be implemented whenever forecasts indicate strike probability is greater than some number less than one. As the cost of protection increases, however, threshold forecast probabilities also increase, albeit slowly. However, using the above tables in a spreadsheet form, one could easily tailor the optimization given the revised cost and loss values based upon scheduled launches and payloads, and even the two shuttle situation.

⁷ Two days for rollback, one day during storm, and two days to get back to initial conditions. Best case scenario.

⁸ Since the payoff for case 3 is less than or equal the payoff of case 2 for all cases, case 2 is never optimal.

SECTION III COMPUTER TRACK ANALYSIS

Another study was initiated using historical data of hurricane and tropical storm tracks. These data were acquired last summer, and comprise track data for storms from 1886 through 1991. Storm data is in ASCII text format, and contains six hourly information containing position, wind, and sea level pressure (Appendix 1).

One goal set last year was to create the ability to examine these data in a graphical form. It was felt that by selecting specific storms similar in nature to the current storm, a probability study on forward motion could be undertaken which could then be used along with the NHC warning⁹ to afford local forecasters another tool for determination of landfall. In addition, the computer access to track data would allow for numerous additional courses of study. The source code for these programs are contained in the appendices and program documentation is integral to the source code.

3.1 PROGRAM DEVELOPMENT AND REQUIREMENTS

Programming was written in Borland's Turbo Pascal on an MS-DOS laptop computer. The goal was to create executable programs that would be easily transportable to that type of machine, and to the extent of the programs written, this goal has been successfully met. Machine requirements are minimal, an MS-DOS machine with one megabyte RAM, and VGA display. As Borland also has Pascal available for the Apple Macintosh machine, the source code can be transferred to a Mac, then compiled and run. Since data files are also transferable (being data files) no significant revision of source code is anticipated.

One note about the programming environment. It was chosen because it satisfied several criterion. First, it was easy to edit code in the Borland environment. Conversations with computer programmers at the center pointed out that fact. Secondly, compilation linking and running are accomplished very easily with syntax errors found and easily corrected, as the environment places the cursor at the error in the edit mode for immediate correction. Finally, the graphical capabilities were greatly enhanced by the extensive library of commands and features available

⁹ NHC has recently incorporated a method whereby historical forecasts of similar storms are used in obtaining forecast guidance, and help in assessing the confidence level in the current forecast.

in the environment. I found it extremely easy to write code to accomplish my task in reasonable time. Having no Pascal experience (but speaking several other languages) the rather steep learning curve was possible only due to the user friendliness of the language.

3.2 DATA AND SUBROUTINES

Each program uses one or both of two data files. One, distributed through shareware, is a database which originally was produced by the Central Intelligence Agency. The documentation accompanying that program, PLOTMWDB, contains an interesting history of the data. It also provides a scheme for unpacking and reading data. Generally, data points are described by type (coastline, river, island, etc.) and coordinates given in minute form. Conversion of these data to pixel address points for screen display was a particularly satisfying success. The procedure named 'MAKEMAP' accomplishes this.

The storm track file, as mentioned, is a text file. Its conversion to integer and real number data proved to be another interesting challenge which was overcome primarily with the use of three procedures which convert two, three and four sequential text characters into integer format for use in the storm plotting routines. That conversion to pixel points makes use of several procedures, one which converts from decimal degrees to minutes, determines if the point is visible in the defined visible screen window. If so, it is plotted, if not, it is discarded. Each visible storm fix is plotted along with adjoining line segments with different colored circles at the fix points representing storm strengths, namely, depression, storm and hurricane. Minor code modifications would permit additional differentiation, for example, based upon category level. The storm data acquisition and plotting is accomplished with a procedure named 'GETSTORM', and appears in several programs as does the map making program.

Another interesting problem was in determining the best way to erase a storm track in the storm selection program. Using the 'exclusive or' feature in the graphing package, writing the screen with the track twice effectively returns it to the original color, in effect erasing it. This required reading a storm data file, and saving it to a temporary location so that it could be reread and plotted. Then, calling a write procedure to plot the storm track, and then calling the write procedure again to rewrite it which in effect erased it. This method was required by the sequential read nature of the original storm track data file and the inability to back step easily.

The following is a brief summary of the two programs which evolved over the course of the summer. The first selects data, the second plots selected years.

3.3 PROGRAM PICKTRAK.PAS

As the name implies, this program was created to allow the user to select the storms of his choice. In its present form, it requests a storm data file name (source file) and then asks for the output file name. After generating a map of the area it sequentially plots the storm track, gives general information concerning that storm, and asks if the user wants to include it in the output file. Any keystroke but 'Y' (not case sensitive) bypasses the storm. This program has the potential of being modified to accomplish many different selection schemes. Yearly data, storms of specified strength, storms passing within x miles of a location, probability distribution functions based upon historical data may all be fairly easily woven into the code.

3.4 PROGRAM PLOTYR.PAS

Designed to allow the user to examine storm data for a given year. On execution, the program requests a file name which contains the storm data which the forecaster is interested in displaying. It then asks for a forward storm speed. This velocity is used to generate event circles which are centered about KSC, and represent decision points for (from center outward) 8,12,24,36,48 and 72 hours. These correspond to the rollback timeline event critical points (chart 3). The program then shifts to graphics mode, generates the map of the region, and plots the storm tracks for the selected data file. This program has the potential to allow selection of specific storm criterion with minor revisions in the source code.

IV. CONCLUDING REMARKS

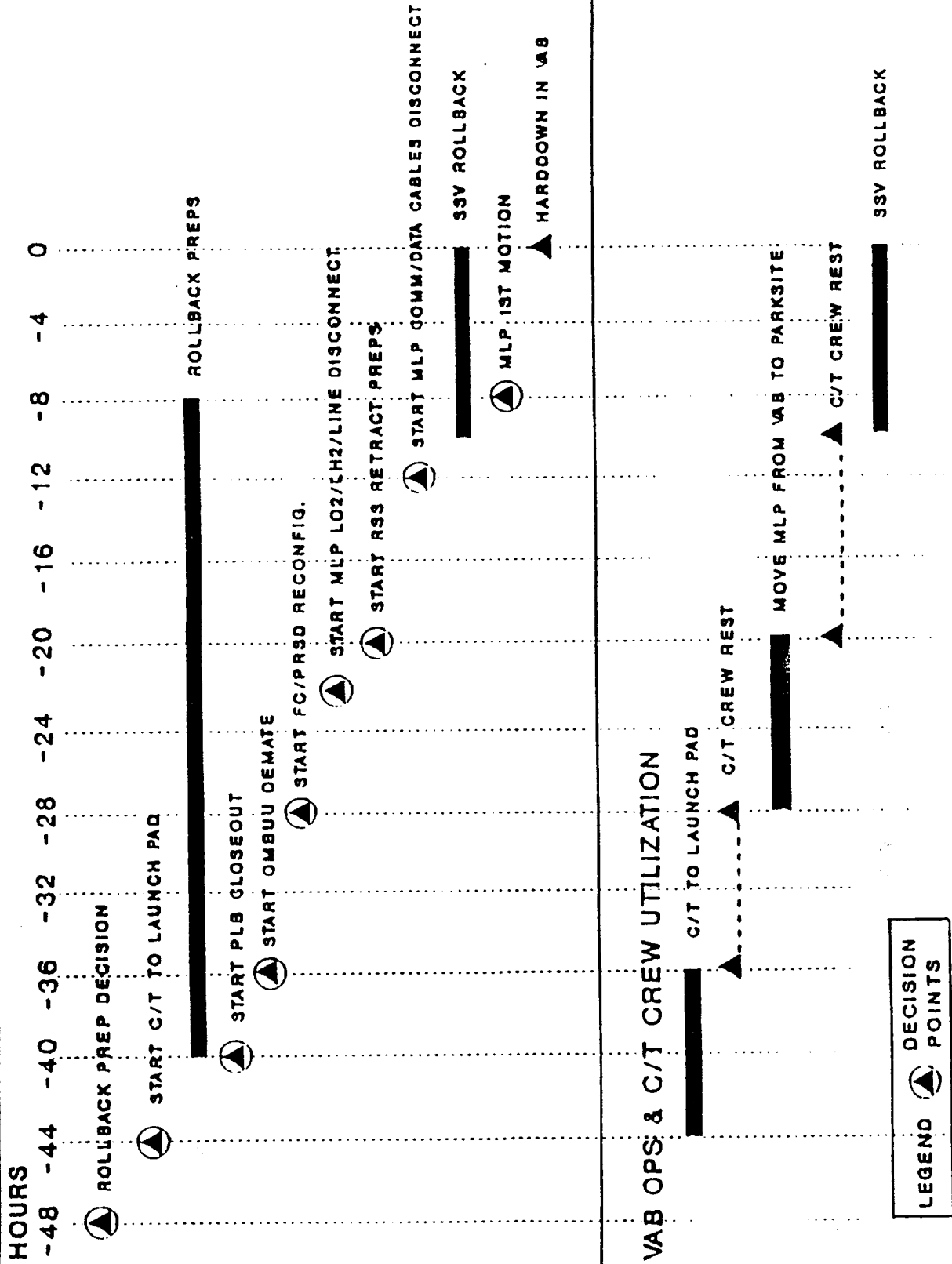
Determination of a fixed probability value for making a rollback decision varies with the number of storms that are forecast to occur at the center, and is very dependent upon the initial values one places on the cost of operations and the loss of the vehicle. Cost is difficult at best to quantify, because of the variability of situations that might exist; one shuttle vs. two, recently moved shuttle, processing accomplished, and even availability of space in the VAB add varying degrees of complexity to the problem. Loss, typically thought of as strictly a dollar value, is again fogged by the value of the space shuttle in the eyes of the nation, and its value in terms of a national resource. Replacement cost might be prohibitive in this period of economic belt tightening, and the adverse effect of such a loss on NASA's reputation and effectiveness would be disastrous. For this reason, the author feels that rollback decisions should be made at a low probability of strike, on the order of 1 or 2 percent.

The time to make those decisions can be fairly easily determined from the NHC warnings. Forward speed times 48 gives the decision circle to use.

The computer packages developed through the course of the study can be used by forecasters to examine more closely the historical data. They also provide a strong departure point for additional studies that might improve hurricane forecasting. Some of those would be the development of a statistical analysis of selected historical storm tracks. For example, one of the tasks not accomplished this summer is a 'forward scatter' capability based on selected storms. It was felt that if a significant number of storms from the historical record were analyzed, probabilities based upon behavior might become evident. For example, if a storm 300 Nautical miles due east were to begin tracking from a heading of 270° to 290° , what percent of similar storms then turned back westward? In other words, what are the chances that the current storm will 'get us'? This methodology could be verified by using real storms which could be shifted so that their track passed within a specified number of miles of the center. The value of the success of this method would allow managers to live with a higher probability of strike.

APPENDICES

SSV ROLLBACK MILESTONE SCHEDULE




```
PROGRAM PICKTRAK (INPUT,OUTPUT);
```

```
USES Graph,DOS,CRT;
```

```
(.....INITIALIZATIONS.....)
```

```
TYPE POINT = RECORD
```

```
CODE : INTEGER;  
LAT : INTEGER;  
LON : INTEGER;  
END;
```

```
VAR
```

```
NEWPOINT : POINT;  
INFILSPT : FILE OF POINT;  
INFILPT_NAME : STRING[14];  
INFIL : FILE OF CHAR;  
IN_FILE : STRING[14];  
OUTFILE : FILE OF CHAR;  
OUT_FILE : STRING[14];
```

```
HDRFILE : TEXT;  
DATA : ARRAY[1..30,1..81] OF CHAR;  
HEADLINE : STRING[80];  
DUMMY : STRING[80];  
TEMP : STRING[53];  
C,D : CHAR;
```

```
VEL : REAL;  
MaxX : INTEGER;  
MaxY : INTEGER;  
GDVR : INTEGER;  
GRMODE : INTEGER;  
PC : INTEGER;  
i,j,k,m,n,r : INTEGER;  
T,U,V,X,JJ : INTEGER;  
minlon : INTEGER;  
OLDLAT,OLDLON : INTEGER;  
MINLAT,MINLON : INTEGER;  
HERELAT,HERELON : INTEGER;
```

```
MaxColor : WORD;  
MAP_FILE : FILE OF POINT; (symbolic name for input map data file)  
FILE_NAME : STRING[16]; (stores map file name input by user)
```

```
( MIN_LAT : REAL; (stores north, south,)  
( MAX_LAT : REAL; (east, and west limits of area)  
( MIN_LON : REAL; (to be displayed - )  
( MAX_LON : REAL; (input by user)
```

```
( MINMAX.MINXMIN: REAL; (max and min latitude and longitude)  
( MINMAX.MINYMIN: REAL; (values used in window generating routine)
```

```
MO,DY,YR : INTEGER;  
NMR,LISE : INTEGER;  
NAME : STRING[10];  
LT : ARRAY[1..100] OF INTEGER;  
LN : ARRAY[1..100] OF INTEGER;  
WN : ARRAY[1..100] OF INTEGER;  
SP : ARRAY[1..100] OF INTEGER;
```

```
VISIBLE,OLDVIS :BOOLEAN;
```

```
(..... INITIALIZE .....
```

```
PROCEDURE initialize;  
begin
```

```
(initialize the graphics)  
DetectGraph(GDVR,GRMODE);  
InitGraph(GDVR,GRMODE, '');
```

```
(gets max x and y values for screen)
```

```
MaxX := GetMaxX;  
MaxY := GetMaxY;  
MaxColor := GetMaxColor;  
SetViewPort(0,0,MaxX,MaxY,ClipOn);  
SetLineStyle(SolidIn,0,NormWidth);  
SetColor(MAXCOLOR);
```

```
end;
```

```
(..... ERASEBLOCK .....
```

```
PROCEDURE ERASEBLOCK(LEFT,TOP,RIGHT,BOTTOM:INTEGER);
```

```
VAR X,Y,Z : INTEGER;  
BEGIN
```

```
Z:=GETBKCOLOR;  
FOR X:= LEFT TO RIGHT DO  
FOR Y:=TOP TO BOTTOM DO  
PUTPIXEL(X,Y,Z);
```

```
END;
```

```
(..... BORDER .....
```

```
PROCEDURE border; (border of screen output)
```

```
begin  
SetColor(15);  
SetLineStyle(SolidIn,0,ThickWidth);  
Rectangle(0,0,maxx,maxy-120);  
SetColor(RED);  
SetLineStyle(SolidIn,0,NormWidth);
```

```
(draw thick border in white)
```

```
(around thin border in grey)
```

```

REPEAT
  READ(INFILE,C);
UNTIL C=CHR(10);

```

END;

(..... MAKE TWO)

```

PROCEDURE MAKE TWO(T,U:INTEGER);

```

VAR P,Q : INTEGER;

BEGIN

```

IF HEADLINE(T)=' ', THEN P:=0
ELSE P:=(INTEGER(HEADLINE(T))-48);
IF HEADLINE(U)=' ', THEN Q:=0
ELSE Q:=(INTEGER(HEADLINE(U))-48);
X:=P*10+Q;

```

END;

(..... MAKE THREE)

```

PROCEDURE MAKE THREE(T,U,V:INTEGER);

```

VAR P,Q,R : INTEGER;

BEGIN

```

IF HEADLINE(T)=' ', THEN P:=0
ELSE P:=(INTEGER(HEADLINE(T))-48);
IF HEADLINE(U)=' ', THEN Q:=0
ELSE Q:=(INTEGER(HEADLINE(U))-48);
IF HEADLINE(V)=' ', THEN R:=0
ELSE R:=(INTEGER(HEADLINE(V))-48);
X:=P*100+Q*10+R;

```

END;

(..... MAKE FOUR)

```

PROCEDURE MAKE FOUR(T,U,V,W:INTEGER);

```

VAR P,Q,R,S : INTEGER;

BEGIN

```

IF HEADLINE(T)=' ', THEN P:=0
ELSE P:=(INTEGER(HEADLINE(T))-48);
IF HEADLINE(U)=' ', THEN Q:=0
ELSE Q:=(INTEGER(HEADLINE(U))-48);
IF HEADLINE(V)=' ', THEN R:=0
ELSE R:=(INTEGER(HEADLINE(V))-48);
IF HEADLINE(W)=' ', THEN S:=0
ELSE S:=(INTEGER(HEADLINE(W))-48);
X:=P*1000+Q*100+R*10+S;

```

END;

```

Rectangl(3,3,maxx-3,maxy-123);
SetColor(white)

```

end;

(..... LATLONCON)

```

Procedure latloncon (y,x:longint);
  (this function converts from map minute coordinates to
  window coordinates for plotting lat/lon on the screen
  and determines if it is in the visible window)

```

begin

```

IF Y<300 THEN VISIBLE:=FALSE (minlat = equator)
ELSE IF Y>2100 THEN VISIBLE:=FALSE (maxlat = 40N)
ELSE IF X<-6600 THEN VISIBLE:=FALSE (maxlon = 110W)
ELSE IF X>-1199 THEN VISIBLE:=FALSE (minlon = 20W)
ELSE VISIBLE:=TRUE;

```

```

winlon:=(x+5500) div 5;
winlat:=(y-2100) div -5);

```

end;

(..... GRIDLINES)

```

Procedure gridlines;

```

BEGIN

```

SetColor(black); (linecolor is dark gray)
SetLineStyle(1,0,1); (line style is dotted)
m:= -5000; (dummy lon for latloncon routine)
for i := 1 to 8 do
  begin
    n:=i*300; (figures winlat values from
    latloncon(n,m); (5N to 40N)
    if visible then
      begin
        MoveTo(0,Winlat); (draws the line)
        LineTo(639,Winlat);
      end
    end
  end

```

end;

```

n:=650; (dummy lat for latloncon routine)
for i := 1 to 20 do
  begin
    m:= -900*(-i*300); (figures winlon values from
    latloncon(n,m); (20W to 115W)
    MoveTo(Winlon,0); (draws line)
    LineTo(winlon,479);
  end;

```

end;

(..... ENDOFFLINE)

```

PROCEDURE ENDOFFLINE;

```

BEGIN

```

  READ(INFILE,C);
  IF C<>CHR(10) THEN
    BEGIN

```

```

(..... WRITEHEADER .....}

PROCEDURE WRITEHEADER;
BEGIN
  FOR I:=1 TO 52 DO
    WRITE(HDRFILE,TEMP[I]);
  WRITELN(HDRFILE,TEMP[53]);
END;
(..... WRITDATA .....}

PROCEDURE WRITDATA(X:INTEGER);
VAR
  I,J : INTEGER;
BEGIN
  FOR I:=1 TO X DO
    BEGIN
      FOR J:=1 TO 79 DO WRITE(HDRFILE,DATA[I,J]);
      WRITELN(HDRFILE,DATA[I,80]);
    END;
  WRITELN(HDRFILE,.....);
END;
(..... PLOTPOINT .....}

PROCEDURE PLOTPOINT(X,Y:INTEGER);
BEGIN
  SetLineStyle(SolidIn,0,ThickWidth);
  LATLONGCON(X,Y);
  IF VISIBLE THEN
    BEGIN
      IF OLDVIS THEN
        MOVETO(OLDLON,OLDLAT);
        LINETO(WINLON,WINLAT);
      END;
      OLDLAT:=WINLAT;
      OLDOLON:=WINLON;
      OLDVIS:=TRUE;
    END
  ELSE
    BEGIN
      OLDVIS:=FALSE;
    END;
END;
(..... PLOTKSC .....}

PROCEDURE PLOTKSC;

```

```

VAR
  NM : INTEGER;
BEGIN
  SETFILLSTYLE(SOLIDFILL,RED);
  HERELAT:=ROUND(28.3833*60);
  HERELON:=ROUND(80.5833*-60);
  LATLONGCON(HERELAT,HERELON);
  FILLELLIPSE(WINLON,WINLAT,2,2); (PLOTS KSC AS A RED DOT)
  HERELON:=WINLON;
  HERELAT:=WINLAT;
  NM:=12; (NM=1 DEG = 60NM)
  SETCOLOR(WHITE);
  CIRCLE(HERELON,HERELAT,ROUND(NM*VEL*9/50)); (PLOTS EVENT CIRCLES)
  CIRCLE(HERELON,HERELAT,ROUND(NM*VEL*12/60)); (DISTANCES: 100,300,AND 500)
  CIRCLE(HERELON,HERELAT,ROUND(NM*VEL*24/60));
  CIRCLE(HERELON,HERELAT,ROUND(NM*VEL*36/60));
  CIRCLE(HERELON,HERELAT,ROUND(NM*VEL*48/60));
  CIRCLE(HERELON,HERELAT,ROUND(NM*VEL*72/60));
END;
(..... PLOTTRACK .....}

PROCEDURE PLOTTRACK(X:INTEGER);
BEGIN
  VISIBLE:=FALSE;
  OLDVIS:=FALSE;
  X:=X*4;
  FOR I:=1 TO X DO
    BEGIN
      PLOTPOINT(LT[I]*6,LN[I]*-6);
    END;
  END;
(..... CLRTRK .....}

PROCEDURE CLRTRK;
VAR
  I : INTEGER;
BEGIN
  FOR I:=1 TO 100 DO
    BEGIN
      LT[I]:=0;
      LN[I]:=0;
    END;
  END;
(..... GETSTORM .....}

PROCEDURE GETSTORM;
BEGIN

```



```

J:=(J-1)*17;
MAKETHREE(13+J,14+J,15+J);
L1+((JJ-1)*4)]=X;
MAKETHREE(17+J,18+J,19+J);
LW1+((JJ-1)*4)]=X;
MAKETHREE(21+J,22+J,23+J);
WN1+((JJ-1)*4)]=X;
IF WN1<35 THEN PC:=3
ELSE IF WN1<65 THEN PC:=14
ELSE PC:=60;
MAKEFOUR(25+J,26+J,27+J,28+J);
SP1+((JJ-1)*4)]=X;

END;
FOR I:=1 TO 10 DO
  DATA(JJ+1,I)]='';
END;
ENDOFFLINE;
PLOTTRACK(K);
OUTTEXTY(0,410,'DO YOU WANT THIS TRACK? (Y or N) ');
D:=READKEY;
OUTTEXTY(264,410,D);
IF (D=CHR(89)) OR (D=CHR(121)) THEN
  BEGIN
    WRITEHEADER;
    WRITEDATA(K);
  END;
PLOTTRACK(K);
CLRTRK;
ERASEBLOCK(0,380,300,450);

CLOSE(INFILE);
CLOSE(HDRFILE);
END;
/..... MAKEMAP .....

```

```

ASSIGN(INFILE,IN_FILE);
ASSIGN(HDRFILE,OUT_FILE);
RESET(INFILE);
REWRITE(HDRFILE);

WHILE NOT EOF(INFILE) DO
  BEGIN
    OLDVIS:=FALSE;
    VISIBLE:=FALSE;
    FOR I:=1 TO 52 DO
      BEGIN
        READ(INFILE,HEADLINE[I]);
        TEMP[I]=HEADLINE[I];
      END;
    READ(INFILE,C);
    HEADLINE[53]=C;
    TEMP[53]=C;
    ENDOFFLINE;
    MAKETWO(20,21);
    K:=X;
    MAKEFOUR(31,32,33,34);
    NMBR:=X;
    MAKETWO(7,8);
    MO:=X;
    MAKETWO(10,11);
    DY:=X;
    MAKEFOUR(13,14,15,16);
    YR:=X;
    OUTTEXTY(0,380,'STORM START DATE: ');
    FOR I:=7 TO 16 DO
      OUTTEXTY(90+8*I,380,HEADLINE[I]);
    OUTTEXTY(0,390,'STORM NUMBER: ');
    FOR I:=1 TO 4 DO
      OUTTEXTY(90+8*I,390,HEADLINE[I+30]);
    OUTTEXTY(0,400,'THIS STORM IS');

    FOR I:=1 TO 10 DO
      BEGIN
        NAME[I]=HEADLINE[35+I];
        OUTTEXTY(110+8*I,480,NAME[I]);
      END;
    FOR JJ:=1 TO K DO
      BEGIN
        FOR I:=1 TO 81 DO
          BEGIN
            READ(INFILE,C);
            HEADLINE[I]=C;
            DATA(JJ,I)=C;
          END;
          READ(INFILE,C);
          MAKETWO(7,8);
          MO:=X;
          MAKETWO(10,11);
          DY:=X;
          FOR I:=1 TO 4 DO
            BEGIN

```

```

          ASSIGN(INFILEPT,'NEWPOINT.DAT');
          RESET(INFILEPT);
          WHILE NOT EOF(INFILEPT) DO
            BEGIN
              READ(INFILEPT,NEWPOINT);
              WITH NEWPOINT DO
                IF CODE > 5 THEN
                  BEGIN
                    LATLONCON(LAT,LON);
                    IF VISIBLE THEN
                      BEGIN
                        MOVETO(winlon,winlat);
                        oldlat:=winlat;
                        oldlon:=winlon;
                        oldvis:=visible;
                      END

```

```

oldlon:=0;
oldvis:=false;

MAKEMAP;
GRIDLINES;
BORDER;
PLOTKSC;
SETWRITEMODE(XORPUT);
GETSTORM;
OUTTEXTXY(1,400,'THAT COMPLETES THE DATA. PRESS RETURN TO EXIT');
READLN(DUMMY);
CloseGraph;

END. (MAIN PROGRAM)

```

```

ELSE CLDVIS:=FALSE
END
ELSE
  BEGIN
    LATLONGCON(LAT, LON);
    IF NOT VISIBLE THEN
      OLDVIS :=FALSE
    ELSE IF NOT OLDVIS THEN
      BEGIN
        MOVETO(WINLON, WINLAT);
        OLDLAT:=WINLAT;
        OLDLON:=WINLON;
        OLDVIS:=TRUE;
      END
    ELSE
      BEGIN
        LINETO(WINLON, WINLAT);
        oldlat:=winlat;
        oldlon:=winlon;
        oldvis:=TRUE;
      END
    END
  END
END;
OLDVIS:=FALSE;
SETFILLSTYLE(SOLIDFILL,CYAN);
FLOODFILL(600,100,WHITE);
FLOODFILL(10,300,WHITE);
CLOSE(INFILEPT);
LATLONGCON(610,-4300);
FLOODFILL(WINLON,WINLAT,WHITE);
SETFILLSTYLE(SOLIDFILL,CYAN);
LATLONGCON(1620,-4850);
FLOODFILL(WINLON,WINLAT,WHITE);
LATLONGCON(700,-5130);
FLOODFILL(WINLON,WINLAT,WHITE);

```

```

END;
(..... MAIN PROGRAM ..... )
BEGIN
  CLRSCLR;
  WRITELN('WHAT IS THE NAME OF YOUR STORM DATA FILE?');
  READLN(IN_FILE);
  WRITELN('WHERE WOULD YOU LIKE TO SAVE YOUR SELECTED STORMS?');
  READLN(OUT_FILE);
  (WRITELN('HOW FAST IS THE CURRENT STORM MOVING?');
  READLN(VEL);
  IN_FILE:='TEST.DAT';)
  INITIALIZE;
  BORDER;
  SETFILLSTYLE(SOLIDFILL, GREEN);
  FLOODFILL(100,100,WHITE);
  oldlat:=0;
  (fills screen with GREEN)

```

```

: ARRAY[1..100] OF INTEGER;
: ARRAY[1..100] OF INTEGER;
: ARRAY[1..100] OF INTEGER;
: ARRAY[1..100] OF INTEGER;
      (used in data read routine)

HEADLINE : STRING(80);
NEWLINE  : STRING(80);
      (flags to point points)

VISIBLE,OLDVIS : 800LEAN;

(..... INITIALIZE ..... )

PROCEDURE Initialize;
begin
  (initialize the graphics)
  DetectGraph(GDVR,GRMODE);
  InitGraph(GDVR,GRMODE,' ');

  (gets max x and y values for screen)
  MaxX := GetMaxX;
  MaxY := GetMaxY;
  MaxColor := GetMaxColor;
  SetViewPort(0,0,MaxX,MaxY,ClipOn);
  SetLineStyle(SolidIn,0,NormWidth);
  SetColor(MAXCOLOR);

end;

(..... BORDER ..... )
PROCEDURE border;
begin
  SetColor(15);
  SetLineStyle(SolidIn,0,ThickWidth);
  Rectangle(0,0,MaxX,MaxY-120);
  SetColor(RED);
  SetLineStyle(SolidIn,0,NormWidth);
  Rectangle(3,3,MaxX-3,MaxY-123);
  SetColor(white);
end;

(..... LATLONCON ..... )
Procedure latloncon (y,x:longint);
(This function converts from map minute coordinates to
 window coordinates for plotting lat/lon on the screen
 and determines if it is in the visible window)
begin
  IF Y<300 THEN VISIBLE:= FALSE (minlat = equator)
  ELSE IF Y>2400 THEN VISIBLE:= FALSE (maxlat = 40N)
  ELSE IF X<6600 THEN VISIBLE:= FALSE (maxlon = 110W)
end;

```

```

PROGRAM PLOTYR (INPUT,OUTPUT);
(This program will plot the storm tracks from data stored in a file
 which the user may select. The program will prompt you for a year)
USES Graph,DOS,CRT,PRINTER;
(.....INITIALIZATIONS.....)

```

```

CONST
ESC = #18;

TYPE
POINT = RECORD
CODE : INTEGER;
LAT : INTEGER;
LON : INTEGER;
END;

CONST
OUT=10;

VAR
NEWPOINT : POINT;
INFLEPT : FILE OF POINT;
INFLEPT_NAME : STRING(14);
INFLE : FILE OF CHAR;
IN_FILE : STRING(14);

DUMMY : STRING(80); (dummy variable for keyboard input purposes)

C : CHAR;
VEL : REAL;
MaxX : INTEGER;
MaxY : INTEGER;
GDVR : INTEGER;
GRMODE : INTEGER;
PC : INTEGER;
i,j,k,a,n,c : INTEGER;
T,U,V,X,JJ : INTEGER;
OLDLAT,OLDLON : INTEGER;
WINLAT,WINLON : INTEGER;
HERELAT,HERELOM : INTEGER;

MaxColor : WORD;

MAP_FILE : FILE OF POINT; (symbolic name for input map data file)
FILE_NAME : STRING(16); (stores map file name input by user)
WINMAXX,WINMINX : REAL; (max and min latitude and longitude)
WINMAXY,WINMINY : REAL; (values used in window generating routine)

NO.DY,YR : INTEGER;
NMOR,LIFE : INTEGER;
NAME : STRING(10);
(variables for storm)
(data, selfevident)

```

```

(used for reading in data from storm file)
(Max x screen size for graphics)
(Max y screen size for graphics)
(graphics driver)
(for coloring position dots)
(used in the position routine)
(lat and lon for position)
(lat/lon for KSC)
(max color value for graphics)

```

```

ELSE P:=(INTEGER(HEADLINE(T))-48);
IF HEADLINE(U)=' ', THEN Q:=0;
ELSE Q:=(INTEGER(HEADLINE(U))-48);
X:=P*10+Q;

END;
(..... MAKETHREE .....
)

PROCEDURE MAKETHREE(T,U,V:INTEGER);
VAR P,Q,R : INTEGER;
BEGIN
IF HEADLINE(T)=' ', THEN P:=0;
ELSE P:=(INTEGER(HEADLINE(T))-48);
IF HEADLINE(U)=' ', THEN Q:=0;
ELSE Q:=(INTEGER(HEADLINE(U))-48);
IF HEADLINE(V)=' ', THEN R:=0;
ELSE R:=(INTEGER(HEADLINE(V))-48);
X:=P*100+Q*10+R;
END;
(..... MAKEFOUR .....
)

```

```

PROCEDURE MAKEFOUR(T,U,V,W:INTEGER);
VAR P,Q,R,S : INTEGER;
BEGIN
IF HEADLINE(T)=' ', THEN P:=0;
ELSE P:=(INTEGER(HEADLINE(T))-48);
IF HEADLINE(U)=' ', THEN Q:=0;
ELSE Q:=(INTEGER(HEADLINE(U))-48);
IF HEADLINE(V)=' ', THEN R:=0;
ELSE R:=(INTEGER(HEADLINE(V))-48);
IF HEADLINE(W)=' ', THEN S:=0;
ELSE S:=(INTEGER(HEADLINE(W))-48);
X:=P*1000+Q*100+R*10+S;
END;
(..... PLOTPOINT .....
)
PROCEDURE PLOTPOINT(X,Y:INTEGER);
BEGIN
SETFILLSTYLE(SOLIDFILL,PC);
LATLONGON(X,Y);
IF VISIBL THEN
BEGIN
IF OLDVIS THEN
BEGIN
MOVETO(OLDOLON,OLDLAT);
LINETO(WINLON,WINLAT);
END;
END;

```

```

ELSE IF X>=1199 THEN VISIBL:= FALSE (winlon = 20M)
ELSE VISIBL:=TRUE;
winlon:= (x+6000) div 5;
winlat:= ((y-2100) div -5);

end;
(..... GRIDLINES .....
)
Procedure gridlines;
BEGIN
SetColor(Black);
SetLineStyle(1,0,1);
n := -5000;
for i := 1 to 3 do
begin
n:=i*200;
latloncon(n,m);
if visible then
begin
MoveTo(0,Winlat);
LineTo(639,Winlat);
end
end;
n:=650;
for i := 1 to 20 do
begin
m := -900+(-i*200);
latloncon(n,m);
MoveTo(Winlon,0);
LineTo(winlon,479);
end;
end;
(..... ENDOFLINE .....
)
PROCEDURE ENDOFLINE;
BEGIN
READ(INFILE,C);
IF C<>CHR(10) THEN
BEGIN
REPEAT
READ(INFILE,C);
UNTIL C=CHR(10);
END;
(..... MAKETWO .....
)
PROCEDURE MAKETWO(T,U:INTEGER);
VAR P,Q : INTEGER;
BEGIN
IF HEADLINE(T)=' ', THEN P:=0

```

```

PROCEDURE GETSTORM;
  (gets storm data)
  (selects and resets storm files)
BEGIN
  ASSIGN(INFILE,IN_FILE);
  RESET(INFILE);
  WHILE NOT EOF(INFILE) DO
  BEGIN
    OLDVIS:=FALSE;
    VISIBLE:=FALSE;
    FOR I:=1 TO 52 DO
      BEGIN
        READ(INFILE,HEADLINE[I]);
        END;
        READ(INFILE,C);
        HEADLINE[53]:=C;
        ENDOFLINE;
        MAKE TWO(20,21);
        X:=X;
        MAKEFOUR(31,32,33,34);
        NBR:=X;
        MAKE TWO(7,9);
        MO:=X;
        MAKE TWO(10,11);
        DY:=X;
        MAKEFOUR(13,14,15,16);
        YR:=X;
        FOR I:=1 TO 10 DO
          BEGIN
            NAME[I]:=HEADLINE[35+I]
            END;
          FOR JJ:=1 TO K DO
            BEGIN
              FOR I:=1 TO 81 DO
                BEGIN
                  READ(INFILE,C);
                  HEADLINE[I]:=C;
                  END;
                READ(INFILE,C);
                MAKE TWO(7,8);
                MO:=X;
                MAKE TWO(10,11);
                DY:=X;
                FOR I:=1 TO 4 DO
                  BEGIN
                    J:=(I-1)*17;
                    MAKE THREE(13+J,14+J,15+J);
                    LN(I+(JJ-1*4)):=X;
                    MAKE THREE(17+J,18+J,19+J);
                    LN(I+(JJ-1*4)):=X;
                    MAKE THREE(21+J,22+J,23+J);
                    WN(I+(JJ-1*4)):=X;
                    IF WN[I]<35 THEN PC:=3
                      ELSE IF WN[I]<65 THEN PC:=14
                  END;
                END;
              END;
            END;
          END;
        END;
      END;
    END;
  END;

```

```

FILELIPSE(WINLON,WINLAT,2,2);
OLDLAT:=WINLAT;
OLDLON:=WINLON;
OLDVIS:=TRUE;
END
ELSE
BEGIN
  OLDVIS:=FALSE;
  END;
END;
END;
(..... PLOTASC .....
PROCEDURE PLOTASC;
VAR NM : INTEGER;
BEGIN
  SETFILLSTYLE(SOLIDFILL,5);
  HERELAT:=ROUND(29.3833*50);
  HERELON:=ROUND(90.5833*60);
  LATLONCON(HERELAT,HERELON);
  FILELIPSE(WINLON,WINLAT,2,2);
  HERELON:=WINLON;
  HERELAT:=WINLAT;
  NM:=12;
  SETCOLOR(WHITE);
  CIRCLE(HERELON,HERELAT,ROUND(NM*VEL*8/60));
  CIRCLE(HERELON,HERELAT,ROUND(NM*VEL*12/60));
  CIRCLE(HERELON,HERELAT,ROUND(NM*VEL*24/60));
  CIRCLE(HERELON,HERELAT,ROUND(NM*VEL*36/60));
  CIRCLE(HERELON,HERELAT,ROUND(NM*VEL*48/60));
  CIRCLE(HERELON,HERELAT,ROUND(NM*VEL*72/60));
  (NM=1 DEG = 60NM)
  (PLOTS EVENT CIRCLES)
  END;
  (..... ERASETRACK .....
PROCEDURE ERASETRACK;
  (after track is plotted, replotted to)
  (erase it)
BEGIN
  K:=K*4;
  FOR I:=1 TO K DO
    BEGIN
      PLOTPOINT(LT[I]*6,LN[I]*-6);
      END;
    READLN(DUMMY);
    CLOSEGRAPH;
    FOR I:=1 TO 20 DO
      BEGIN
        WRITELN('LAT=' ,LT[I], ' LON=' ,LN[I]);
        END;
      END;
    END;
  (..... GETSTORM .....

```



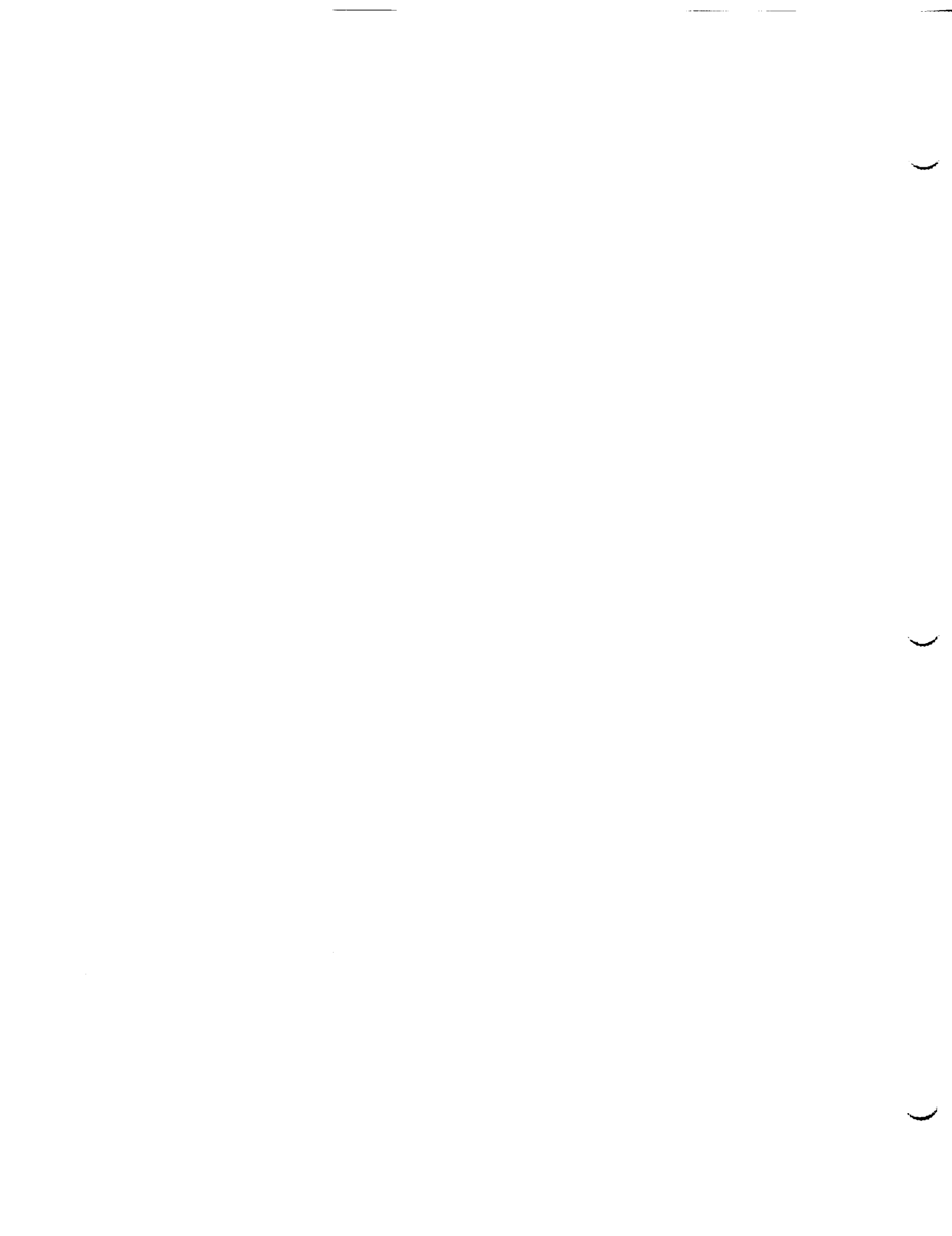
```

UNTIL J>= MAXX-1;
WRITE(LST, #SOC);
..... MAIN PROGRAM .....
END;
(.....)

BEGIN
  CLRSCL;
  WRITELN('WHAT IS THE NAME OF YOUR STORM DATA FILE?');
  READLN(IN_FILE);
  WRITELN('HOW FAST IS THE CURRENT STORM MOVING?');
  READLN(VEL);
  INITIALIZE;
  BORDER;
  SETFILLSTYLE(SOLIDFILL, GREEN);
  FLOODFILL(100, 100, WHITE);
  oldat:=0;
  oldlon:=0;
  oldvis:=false;

  MAKEMAP;
  GRIDLINES;
  BORDER;
  PLOTKSC;
  SETWRITEMODE(XORPUT);
  GETSTORM;
  MoveTo(0, 400);
  OUTTEXT('STORM DATA USING DATA FILE ');
  FOR I:=1 TO 11 DO OutText (IN_FILE[I]);
  READLN(DUMMY);
  (PRINTGRAPH;
  OUTTEXTXY(0, 420, 'Press return to exit. ');
  CloseGraph;
  END. (MAIN PROGRAM)

```



1994018776

N94-28249

1993 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA

49-15

197205

P 30

FLAME TRENCH ANALYSIS OF NLS VEHICLES

PREPARED BY:	Dr. Nuri Zeytinoglu
ACADEMIC RANK:	Associate Professor
UNIVERSITY AND DEPARTMENT:	Purdue University Engineering and Technology Department
NASA/KSC	
DIVISION:	Mechanical Engineering
BRANCH:	Special Projects
NASA COLLEAGUE:	Eric Thaxton
DATE:	August 13, 1993
CONTRACT NUMBER:	University of Central Florida NASA-NGT-60002 Supplement: 11

ACKNOWLEDGEMENTS

I would like to express my deep appreciation for being selected as a 1993 NASA/ASEE Summer Faculty Fellow, to Dr. Ramon Hosler of University of Central Florida and to Dr. Gary Lin of Kennedy Space Center. Special thanks to my NASA colleague Mr. Eric Thaxton who provided great amount of help and information about the problem. Thanks to Mrs. Kari Stiles for doing a wonderful job in all the administrative related matters. I also wish to thank Mr. Rao Caimi for the trip to Pad 39B and his help on computer related problems. I am also beholden to the rest of the Special Projects Division members for lending a helping hand whenever I needed one.

ABSTRACT

The present study takes the initial steps of establishing a better flame trench design criteria for future National Launch System vehicles. A three-dimensional finite element computer model for predicting the transient thermal and structural behavior of the flame trench walls was developed using both I-DEAS and MSC/NASTRAN software packages. The results of JANNAF Standardized Plume Flowfield calculations of sea-level exhaust plumes of SSME, STME and ASRM were analyzed for different axial distances. The results of sample calculations, using the developed finite element model, are included. The further suggestions are also reported for enhancing the overall analysis of the flame trench model.

SUMMARY

Pads 39A and B are being studied to determine if a new family of vehicles can be launched without damaging the existing structure. The new vehicles are designated as NLS (National Launch System). NLS is a family of vehicles which are built around a common core using different external booster configurations. The engines for the core and boosters have not been determined yet. Options include using old shuttle main engines (SSME), modified Saturn V engines (F-1A), Russian RD-170 engines, and solid fuel boosters. The flame trench capacity will be an important factor in the decision of these available options. Recently, the flame trench has been analyzed using simplifying assumptions for one of the NLS configurations at the Kennedy Space Center. Due to the importance of this problem an accurate, reliable solution is required.

The purpose of this summer research study is to investigate the capacity of the existing flame trench and to establish a better flame trench design criteria for future NLS launch vehicles. The solution of the problem requires the following steps:

1. A three-dimensional finite-element model of the flame trench walls.
2. The exit flow conditions of the rocket engine nozzles.
3. The complete analysis of the sea-level plume characteristics.
4. The plume impingement code to predict the heating and pressure rates on the flame trench.

The above steps are carried out as far as possible in the limited time available. As a first step three-dimensional finite element computer model is developed for predicting the transient thermal and structural behavior of the flame trench using I-DEAS and MSC/NASTRAN software packages. Secondly, the JANNAF Standardized Plume flowfield (SPF) is used to obtain the gas dynamic structure of sea-level exhaust plumes of different rocket exhaust flow fields. The analysis is also needed to gain a better understanding of heating rates and pressure rates on the flame trench walls. Since the steps described above are not completed fully at the present time, the results presented in this report should be looked at in a qualitative manner. However, these results will be used later to improve the required solution. The further suggestions are also reported for improving the overall analysis of the flame trench. It was concluded that the three-dimensional finite-element model of the flame trench can be tested for any given NLS configuration with slight modifications explained in section V.

TABLE OF CONTENTS

- I. INTRODUCTION
- II. DESCRIPTIONS OF THE COMPUTER CODES REQUIRED FOR THE STUDY
 - 2.1 RAMP CODE
 - 2.2 JANNAF CODE
 - 2.3 SFPLIMP CODE
 - 2.4 I-DEAS FEM PACKAGE
 - 2.5 MSC/NASTRAN CODE
- III. METHODOLOGY
 - 3.1 FINITE ELEMENT MODEL
 - 3.2 HEATING LOADS AND PRESSURE LOADS
- IV. RESULTS
- V. CONCLUSIONS AND RECOMMENDATIONS
- APPENDIX
- REFERENCES
- TABLES
- FIGURES

LIST OF ILLUSTRATIONS

Figure	Title
1	General Overview of PAD 39B
2	A Schematic Representation of the Plume Flowfield Regions
3	Computational Scheme
4	3-D Model of the Flame Trench
5	Restraint Set Number 1 for Static Run
6	Displacement Contour Lines of Constant Pressure Loading (P=2 atm)
7	Deformed Shape of the Model with Constant Pressure Loading (P=2 atm)
8	Stress Contour Levels of Constant Pressure Loading (P=2 atm)
9	Steady-State Temperature Contours of the FEA Model
10	Displacement Contour Lines of Thermal Loads
11	Stress Contour Levels of Thermal Loads

LIST OF TABLES

Table	Title
1	The Face Heat Influx Distribution on the Brick Elements using I-DEAS software
2	Results of the Transient Analysis with MSC/NASTRAN Program
3	Sea Level ASRM, SSME and STME Exhaust Plumes Axial Mach Number Distributions
4	Sea Level ASRM, SSME and STME Exhaust Plumes Axial Temperature Distributions
5	Sea Level ASRM, SSME and STME Exhaust Plumes Axial Pitot Pressure Distributions

LIST OF ABBREVIATIONS AND ACRONYMS

ASRM	- Advanced Solid Rocket Motors
FEM	- Finite Element Modeling
I-DEAS	- Integrated Design Engineering and Analysis System
JANNAF	- Joint Army, Navy, NASA, Air Force
LC	- Launch Complex
MLP	- Mobile Launch Platform
NLS	- National Launch System
RAMP	- Reacting Multiphase
SFPLIMP	- Source Flow Plume Impingement Program
SPF	- Standardized Plume Flowfield
SRM	- Solid Rocket Motors
SSME	- Space Shuttle Main Engine
SSV	- Space Shuttle Vehicle
STME	- Space Transportation Main Engines

I. INTRODUCTION

The existing flame trench at Launch Pad 39A was built in 1967 for the Apollo space program. An identical launch pad was built in 1969 and named Pad 39B. Both pads were converted to be used with the space shuttle in 1981. The flame trench has not been modified since its construction. Pads 39A and B are being studied to determine if a new family of vehicles can be launched without damaging the existing structure. The new vehicles are designated as NLS (National Launch System). NLS is a family of vehicles which are built around a common core using different external booster configurations. The engines for the core and boosters have not been determined. Options include using old shuttle main engines (SSME), modified Saturn V engines (F-1A), Russian RD-170 engines, and solid fuel boosters. The flame trench capacity will be an important factor in the decision process.

The flame trench is built of reinforced concrete lined with a refractory coating and alumina fire brick. It is 58 feet wide and 42 feet deep. There is a parabolic flame deflector directly under the vehicle to divert the flow to the north and south exists from the flame trench.

The flame trench is limited in temperature and pressure capacity. The effects of an exhaust plume in the trench have not been studied. Excessive amounts of pressures could cause the trench walls to crack and excessive amounts of temperatures could damage the refractory coating.

KSC-STD-Z-0012B, Standard for Flame Deflector Design, is the only design criteria standard currently available to the flame trench. This standard establishes minimum design requirements for the height, width and distance away from the engines that the flame deflector must comply with. Currently, the deflector width has been empirically established as a minimum of 1.6 times the sum of all the vehicle engine nozzle diameters impinging on a common surface. The width and depth of the flame trench must then be sized to accommodate the flame deflector.

The purpose of this study is to investigate the capacity of the existing flame trench for different exhaust configurations and to establish a better flame trench design criteria for future NLS launch vehicles. The study is carried out based on the drawing number 79K10338 and Figure 1 shows the general overview of the PAD39B.

Main thrust of the summer research study can be summarized as:

(1) Survey literature of available computer codes for the problem solution including the nozzle design, plume flow, plume impingement and finite-element programs.

(2) Developing a three-dimensional finite-element model using I-DEAS and MSC/NASTRAN softwares for the flame trench walls.

(3) Using Joint Army, Navy, NASA, Air Force (JANNAF) Standardized Plume Flowfield code (SPF) for predicting the gas dynamic structure of sea-level rocket exhaust plumes.

(4) Better prediction of the pressure and heating rates on the flame trench walls using Source Flow Plume Impingement Program (SFPLIMP).

II. DESCRIPTIONS OF THE COMPUTER CODES REQUIRED FOR THE STUDY

The computer codes required to solve the problem are briefly described below:

2.1 RAMP CODE

The Reacting multiphase (RAMP) code was developed by Lockheed-Huntsville under government funding and used to solve a wide variety of problems associated with real gas, supersonic, compressible flow (Ref. 1). The code is capable of performing rocket nozzle flow calculations and has a direct interface with the JANNAF CODE. The code is fully coupled in that it considers the exchange of momentum and energy between the gas particle plumes. The results of the nozzle solution via RAMP code must be presented in the form of radial distributions of properties at the exit plane as being the initial conditions for the JANNAF code.

2.2 JANNAF CODE

There has been remarkable progress in developing mathematical simulations of the complicated structure and behavior of the plume phenomena in the past decade. A number of review papers and computer programs are available to give the researcher a general overview of the state-of-the-art in rocket exhaust plume analysis (Refs. 2 through 4).

Joint Army, Navy, NASA, Air Force (JANNAF) Standardized Plume Flowfield (SPF) is a computer code which is accepted as government and industry standard for predicting the gas dynamic structure of two-phase low altitude rocket exhaust plumes. Figure 2 illustrates a schematic of the plume flowfield regions. The computational methodology employed in the latest version (SPF-II) of the JANNAF code is described in Ref. 5.

The code mainly contains three principal components: Processor, Shock Capturing and Turbulent Mixing component. Processor component reads the user input data (run parameters, chemical systems, initial conditions and external flow conditions) and data bank (JANNAF thermodynamic and chemical kinetic data for the species). It also creates the input files required for the shock-capturing and turbulent mixing components. Additionally, the Processor uses laminar viscosity and thermal conductivity data embedded in a subroutine.

The Shock-Capturing Component of the code provides an inviscid, frozen chemistry solution of the plume nearfield region. It uses shock-capturing to account for the detailed shock structure, and it fully accounts for gas/particle interactions. The Shock-Capturing Component is utilized for a distance of approximately two inviscid cells which is coded as three times the distance to the Mach disc.

The Turbulent Mixing Component of SPF provides a viscous, turbulent mixing solution of the plume shear layer with finite-rate chemistry. Gas/particle interactions are fully accounted for, with the local fluid viscosity and thermal conductivity. Code offers different forms of turbulence modeling through this component.

2.3 SFPLIMP CODE

The Source Flow Plume Impingement Program (SFPLIMP) is a program designed to compute the forces, moments and heating rates caused by plume impingement on bodies immersed into plume flow fields. The code is capable of modeling on a wide range of configurations. In addition to the plume and processing data, user must provide a three-dimensional representation of a target geometry. The code is capable of modeling several subshapes including flat plates (right triangles, rectangles and circles) and three-dimensional shapes (circular cones, circular cylinders, spheres and polynomials of revolution).

2.4 I-DEAS FEM PACKAGE

I-DEAS Finite Element Modeling (FEM) allows the user to build a finite element model, including physical and material properties, loads, and boundary conditions. This package has excellent pre-processing and post-processing features. Solving the model generally runs in I-DEAS Model Solution. The model solution can solve problems in linear statics and in steady-state heat transfer. For other types of analysis such as transient heat transfer, I-DEAS Model Solution can write the entire FEA model to outside solver package.

2.5 MSC/NASTRAN CODE

MSC/NASTRAN is a large scale general purpose digital computer program which solves a wide variety of engineering problems by the finite element method. This code can analyze linear, non-linear and transient heat transfer with constant or temperature-dependent convective and radiative boundary conditions. The code can use isotropic and anisotropic temperature dependent thermal conductivity material properties. It also provides a user-selected difference parameter for stability in transient solution algorithm.

MSC/NASTRAN operates in a batch mode. Input to the batch process is usually in the form of a card deck or card image file. The data deck is constructed preparing the cards for executive control deck. The purpose of this deck is to identify the job and the type of solution to be performed. It also declares the general conditions under which the job is to be executed, such as maximum time allowed, type of system diagnostics desired etc. The user next sets the data cards for the case control deck. This deck defines the load case, selects the data from the Bulk Data Deck and makes the output requests for plotting. The Bulk Data Deck is the final step of organizing the card decks. This deck contains the majority of the input data for the file. It includes all the data necessary to describe the thermal model and its loading conditions.

III. METHODOLOGY

Figure 3 shows the computational scheme required for the problem solution. The JANNAF computer code requires the exit conditions of the engine nozzle to start the calculations for sea-level plume definition of the proposed exhaust configuration. The results from the JANNAF code must be run to predict the loading conditions of the finite element model. The I-DEAS and MSC/NASTRAN codes must be then used to determine the capacity of the flame trench for different NLS configurations.

3.1 FINITE ELEMENT MODEL

The flame trench facing to the north side of Pad 39B is modeled by a three-dimensional finite element model starting at the base of the exhaust deflector, running approximately 50 feet toward the north along the centerline of the plume footprint. Figure 4 shows the three-dimensional model used in the study. The model is built of from 280 nodes and 162 solid brick type of elements. The model uses physical and thermal material properties of concrete and brick. The concrete stiffeners shown in drawing 79K10338 were modeled with the restraint set as shown in Figure 5.

3.2. HEATING LOADS AND PRESSURE LOADS

The impingement of the Space Shuttle Main Engines (SSME) and Solid Rocket Motors (SRM) or Advanced Solid Rocket Motors (ASRM) gaseous exhaust plumes on Launch Complex 39 (LC-39) results in pressure loads and convective heating to numerous components. Reference 6 presents the results of a study to determine the thermal and pressure environment imposed on elements of LC-39 by impingement of the Space Shuttle Vehicle (SSV) exhaust plumes at the sea level for a maximum north drift trajectory. The plume properties given in this reference are used to establish pressures and heating loads to objects immersed in the ASRM or SSME exhaust plume. Reference 6 do not take into account for radiation from plume exhaust turned by the Mobile Launcher Platform (MLP) and recirculation. Prediction of this water-cooled launch pad radiation is not readily computed due to the extreme changes of rocket plume flow field properties and recirculated flow field environments during launch.

Both SRM and SSME flame trenches are subjected to heating due to vehicle exhaust gases that have flowed down the trench after having impinged on the flame deflector. A rigorous calculation of the thermal and pressure environments of objects subjected to indirect impingement requires a three-dimensional analysis of the flow field which would be far beyond this study due to time limitation. Therefore, a simplified approach is used for the heating and pressure loads required for the study. The heating rates were calculated for each point on the flame trench as described in Reference 6, assuming that the heating rate at the point where the plume initially strikes the flame deflector varies with the distance down the flame deflector and flame trench, taken to the 0.2 power.

The heat loading conditions, used in the present study, include the face heat influx distribution to fifty-four of the elements as presented in Table 1 as explained in the previous paragraph and face convection applied to the back surface of the model. The heat transfer coefficient for face convection was calculated based on the recommendation of Churchill and Chu (Ref. 7).

The results of the gaseous and particle plume impingement pressure loads from the plumes are also included in Reference 6. It is assumed that the gaseous flow and particles travel parallel to the trench walls. The estimates of pressure loads to the flame trench are presented as a result of indirect impingement. The face pressure level of 2 atmospheres is used as recommended in Reference 6.

IV. RESULTS

Since the I-DEAS software package can not provide transient temperatures and better predictions of the time dependent loads are required, the calculations presented in this section should be looked at qualitatively.

Figure 6 depicts the displacement contour lines for constant applied face pressure loading of 2 atmospheres for the static run only. Contour level 6 indicates the maximum displacement of .002 inches. The maximum displacement contour levels occur close to the midpoint of the two concrete stiffeners at the center of the model. Following program execution the deformed shape of the structure is shown in Figure 7. It is notable that the structure deforms in a manner which agrees with its expected deformation. Figure 8 shows the stress contour levels for the same loading condition.

Figure 9 illustrates the steady-state temperature contours of the resulting model based on the applied face heat flux and face convective heat transfer. Maximum temperatures of 2600 degrees Fahrenheit were observed at alumina brick close to the deflector as expected. Appendix provides the necessary MSC/NASTRAN data cards for transient heat transfer analysis. Table 2 presents the results of the transient analysis with MSC/NASTRAN finite element program during the first three seconds of launch.

Both I-DEAS and MSC/NASTRAN are capable of computing thermal stress analysis due to the thermal loading conditions. In this process, the first run (thermal analysis) will calculate the temperature distribution in the model due to the given thermal loads and boundary conditions. The second run (structural analysis) uses these calculated temperatures as temperature loads to a structural analysis to calculate the displacements and stresses caused by the temperatures. Figure 10 and 11 illustrate the displacement and stress contour lines, respectively.

The sea level plume definitions were calculated with the JANNAF code using an exit plane start line generated with the RAMP code. Tables 3 through 5 include the plume properties (Mach number, temperature and pressure) of ASRM, SSME and STME sea-level plumes for different axial distances at a radial distance of 2 feet.

V. CONCLUSIONS AND RECOMMENDATIONS

A computational scheme has been proposed and partially developed for the flame trench capacity of different NLS exhaust configurations. Three-dimensional finite element model presented in this study can be used effectively to analyze thermal and structural acceptability of the flame trench for proposed loading conditions. Further improvements on the FEA model may include: addition of the reinforced concrete material conditions to the model, addition of the temperature dependence to the properties and the convergence check of the results with a better mesh refinement of the FEA model.

Suggestions for completion of the problem are listed below:

- Use of reacting multiphase nozzle code for Saturn V and Russian RD-170 engines.
- Run of the JANNAF Plume code for Saturn V and Russian RD 170 engine exit nozzle conditions.
- Calculations of the impingement pressure loads, convective heating loads and radiation heating loads on the flame trench walls using SPFLIMP code.
- Setting temperature and stress limits for the flame trench model.

APPENDIX

```

ID DIRECT TRANSIENT,                                P3DMOD1
$
$   MAXIMUM CPU TIME ALLOWED FOR THE JOB
$
TIME 10 $ UNITS ARE IN MINUTES
$
$   THE THERMAL ANALYZER PORTION OF MSC/NASTRAN IS TO BE USED
$
APP HEAT
$
$   THE TRANSIENT ALGORITHM IS TO BE USED
$
SOL 89
CEND $ END OF EXECUTIVE CONTROL DECK
$
$*****
$   END OF EXECUTIVE CONTROL ... START CASE CONTROL
$*****
$
TITLE=TRANSIENT PROBLEM      ---   UNITS=BG
$   REQUEST SORTED AND UNSORTED CARD ECHO=S TO SEE COMMENTS
$   IF THIS CARD IS OMITTED ONLY THE SORTED BULK DATA WILL APPEAR
$   COMMENT CARDS WOULD THEN NOT BE PRINTED
ECHO=BOTH
$
$   SPECIFY THE SUPERELEMENTS TO BE RUN. THIS CARD IS REQUIRED
$   AND SPECIFIES ALL FOR THIS MODEL THERE IS BUT A SINGLE
$   SUPERELEMENT WHICH ENCOMPASSES THE ENTIRE MODEL
SEALL=ALL
DLOAD=300
TEMP(ESTI)=400
IC=400
TSTEP=500
$   SELECT THE OUTPUT DESIRED
$   OUTPUT
THERMAL=ALL
SUPER=ALL
FLUX=ALL
SUBTITLE=CASE SET 2, LOAD SET 2
$   END CASE CONTROL DECK
BEGIN BULK
***** BULK DATA RECEIVED FROM THE I-DEAS *****
*****          1085 LINES OF ENTRY          *****
TLOAD2,300,2,,0.,1.+6,0.,0.,+TL1
+TL1,0.,0.
TEMPD,400,70.
TSTEP,500,10,.2,1
PARAM AUTOSPC YES
PARAMPOST -2
ENDDATA

```

REFERENCES

1. Smith, S.D., "Improved Rocket Engine Nozzle/Plume Code" Pages 183-196.
2. Dash, S.M., "Recent Developments in the Modeling of High Speed Jets, Plumes and Wakes, "AIAA Paper 85-1616, presented at AIAA 18th Fluid Dynamics Plasma Dynamics and Laser Conference, July 1985.
3. Wolf, D.E., and Dash, S.M., "Interactive Phenomena in a Supersonic Jet Mixing Plumes, Part I: Phenomenology and Numerical Modeling Technique," AIAA Journal, Vol.22, No. 7, July 1984, pp. 905-913.
4. Dash, S.M. and Wolf, D.E., "Advances in Two-Phase Flow Modeling for the JANNAF Standard Plume Flowfield Model (SPF)", JANNAF 13th Plume Technology Meeting, CPIA Pub. 357, Vol. II, April 1982, pp. 41-72.
5. Dash, S.M., Pergament, H.S., Wolf, D.E., Sinha, N. and Taylor, M.W., "The JANNAF Standardized Plume Flowfield Code Version II, Volume I," U.S. Army Missile Command, TR CR-RD-SS-90-4, July 1990.
6. Environment and Test Specification Levels Ground Support Equipment for Space Shuttle System Launch Complex 39, GP-1059 Volume IV, Engineering and Development Directorate, April 1992.
7. Fundamentals of Heat and Mass Transfer, Incropera, F.P. and Dewitt, D.P., Second Edition, Wiley Publ., 1985.

TABLE 1

The Face Heat Influx Distribution on the
Brick Elements using I-DEAS software

HEATING RATE (Ref. 6)
BTU/(FEET**2 SEC)

DISTANCE FROM THE DEFLECTOR BASE (FEET)	DISTANCE FROM THE FLOOR (FEET)					
	3.5	10.5	12.5	24.5	31.5	38.5
2	240.16	282.44	372.68	372.68	282.44	240.56
6	192.81	226.76	299.20	299.20	226.76	192.81
10	174.04	204.69	270.09	270.09	204.69	174.04
15	160.55	188.83	249.16	249.16	188.83	160.55
21	150.12	176.57	232.97	232.97	176.57	150.12
27	142.70	167.83	221.45	221.45	167.83	142.70
33	137.08	161.22	212.73	212.73	161.22	137.08
40.5	131.91	155.14	204.71	204.71	155.14	131.91
47.5	128.86	151.54	199.96	199.96	151.54	128.86

TABLE 2
Results of the Transient Analysis with
MSC/NASTRAN Program

TIME (SEC)	TEMPERATURE (DEGREES IN FAHRENHEIT)				
	DISTANCE FROM THE DEFLECTOR BASE				
	X=0'	X=12'	X=24'	X=36'	X=50'
0.0	100	100	100	100	100
0.2	227	185	174	168	165
0.4	505	371	337	318	309
0.6	731	523	471	441	426
0.8	893	634	568	530	511
1.0	1037	734	655	610	588
1.2	1160	819	730	679	654
1.4	1272	898	800	743	713
1.6	1376	972	864	802	772
1.8	1474	1043	926	859	826
2.0	1568	1111	985	914	878
2.2	1659	1178	1043	967	929
2.4	1747	1243	1100	1019	979
2.6	1834	1308	1157	1071	1029
2.8	1920	1373	1212	1122	1078
3.0	2003	1436	1268	1173	1126

TABLE 3

Sea Level ASRM, SSME and STME Exhaust Plumes
Axial Mach Number Distributions

Axial Distance (FEET)	MACH NUMBER		
	ASRM	SSME	STME
50	2.363	1.466	2.690
100	2.363	1.145	1.441
150	2.358	0.615	0.812
200	2.096	0.387	0.494
250	1.758	0.278	0.344
300	1.467	0.216	0.262
350	1.224	0.177	0.211
400	1.026	0.150	0.177
450	0.869	0.131	0.153
500	0.743	0.116	0.134
550	0.642	0.105	0.120
600	0.561	0.096	0.109

TABLE 4

Sea Level ASRM, SSME and STME Exhaust Plumes
Axial Temperature Distributions

Axial Distance (FEET)	TEMPERATURE (DEGREES RANKINE)		
	ASRM	SSME	STME
50	3879	4918	4060
100	3879	4564	4849
150	3898	3198	4194
200	4233	2134	2890
250	4607	1597	2080
300	4848	1308	1635
350	4973	1134	1371
400	4988	1021	1202
450	4904	943	1086
500	4740	885	1003
550	4496	842	941
600	4175	808	893

TABLE 5

Sea Level ASRM, SSME and STME Exhaust Plumes
Axial Pitot Pressure Distributions

Axial Distance (FEET)	PITOT PRESSURE (PSIA)		
	ASRM	SSME	STME
50	105.8	51.9	97.6
100	105.8	31.4	43.2
150	105.3	18.6	21.8
200	84.5	16.2	17.2
250	61.6	15.5	15.9
300	45.4	15.2	15.4
350	34.3	15.0	15.2
400	27.3	14.9	15.0
450	23.1	14.9	14.9
500	20.5	14.8	14.9
550	18.9	14.8	14.9
600	17.9	14.8	14.8

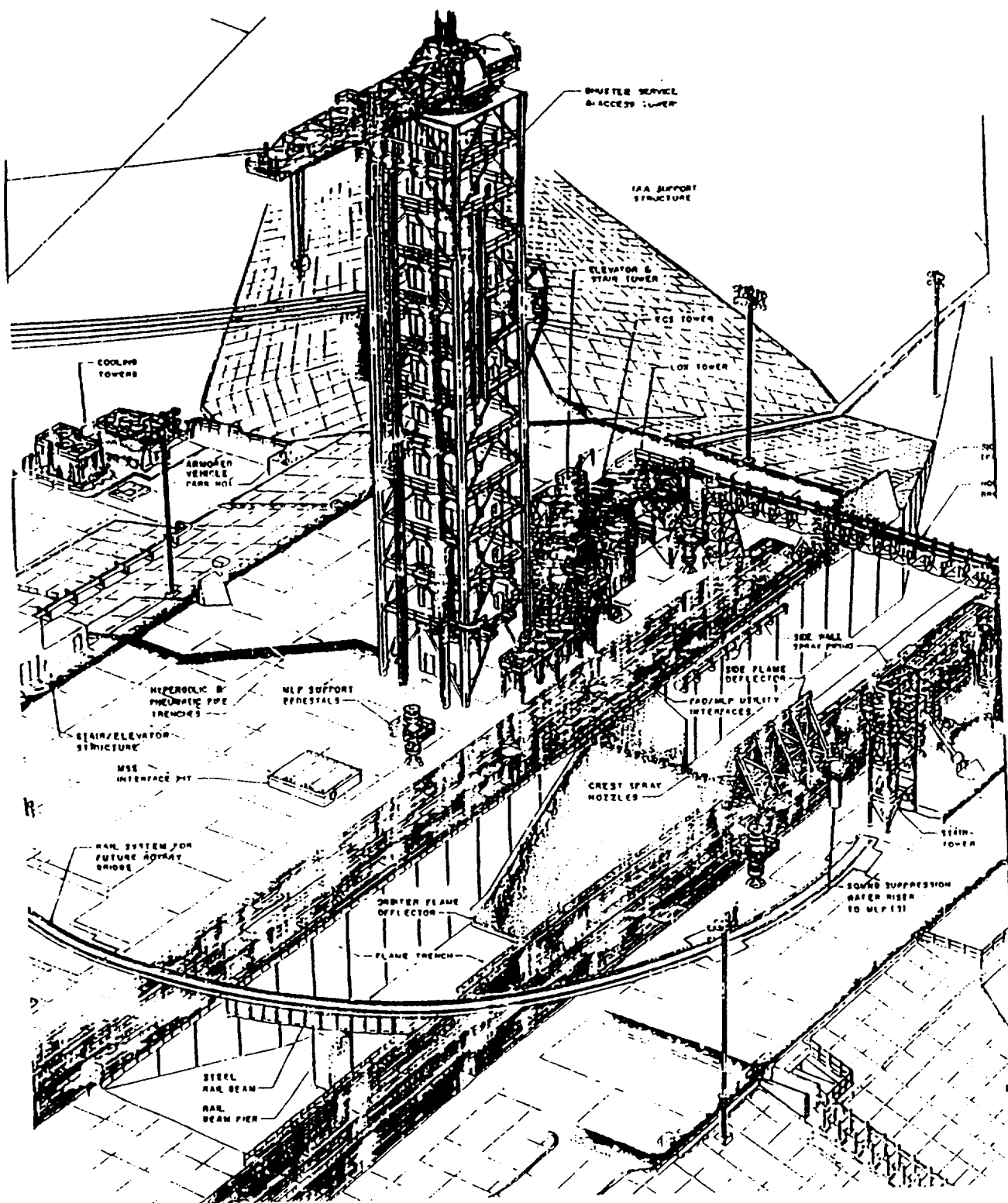


Figure 1 General Overview of PAD 39B

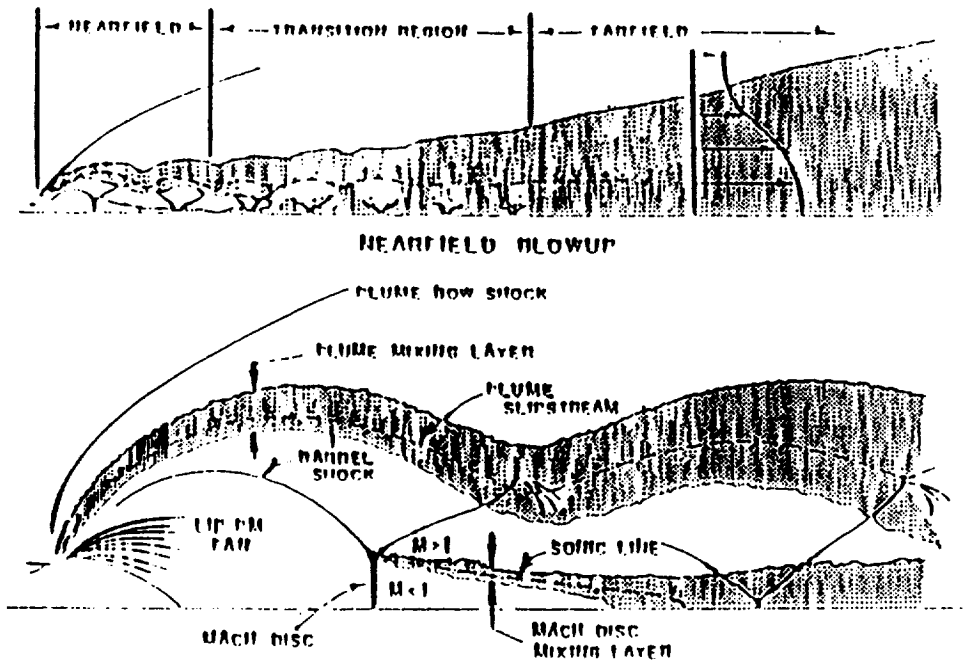


Figure 2 A Schematic Representation of the Plume Flowfield Regions

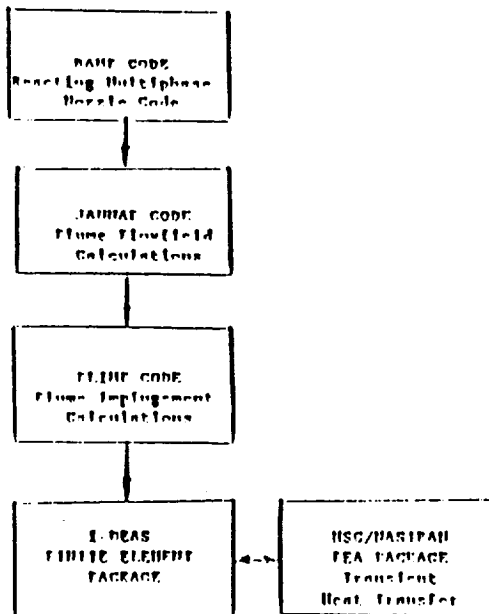


Figure 3 Computational Scheme

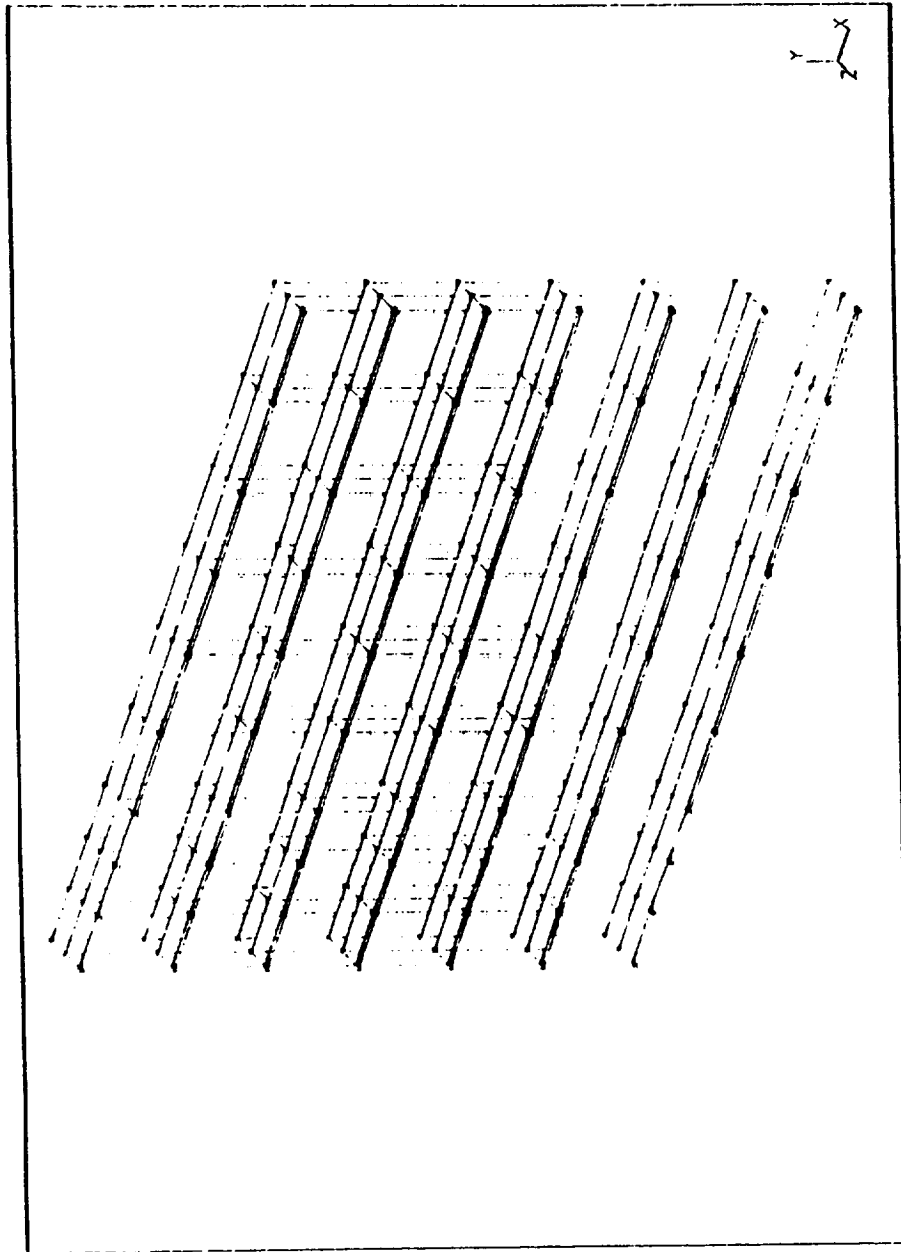


Figure 4 Three-Dimensional Model of the Flame Trench

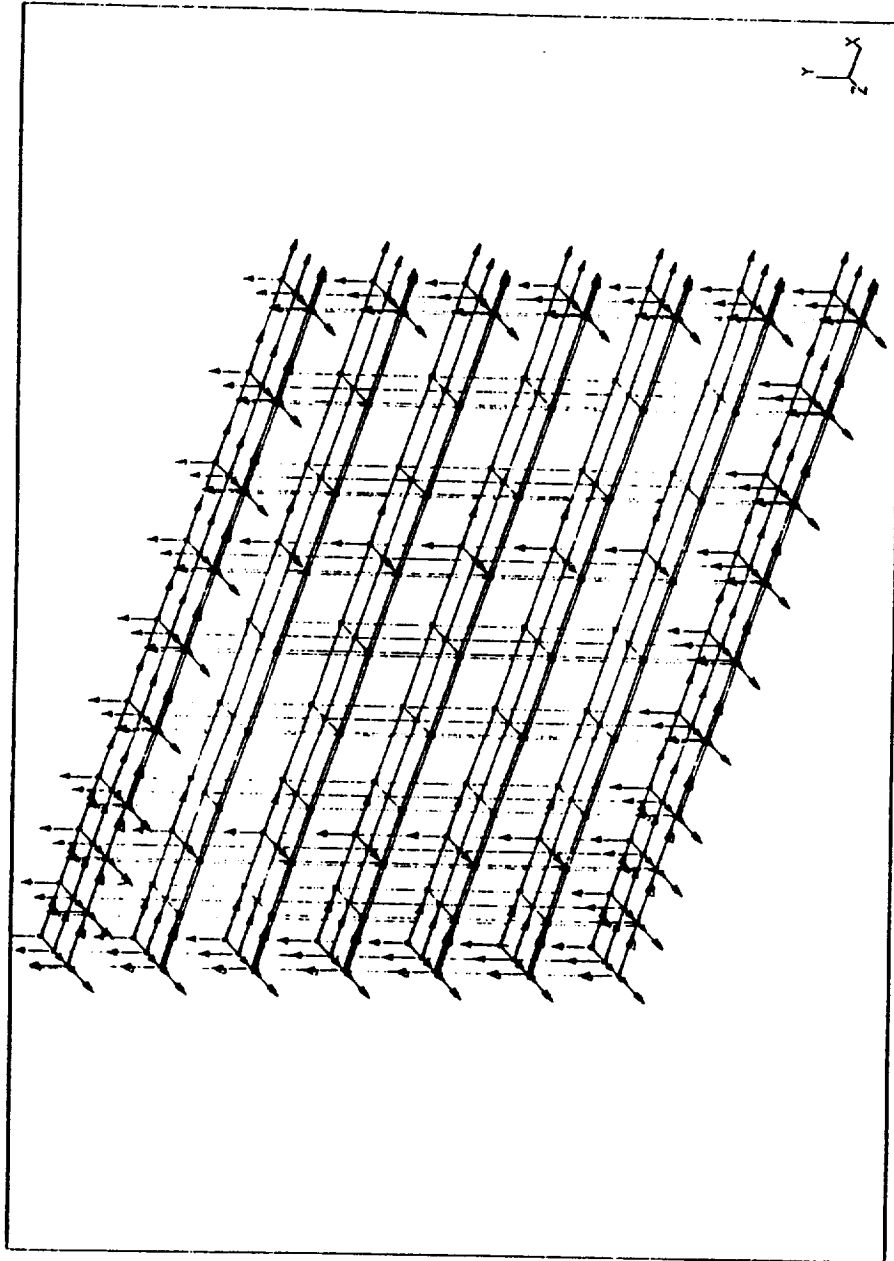


Figure 5 Restraint Set Number 1 for Static Run

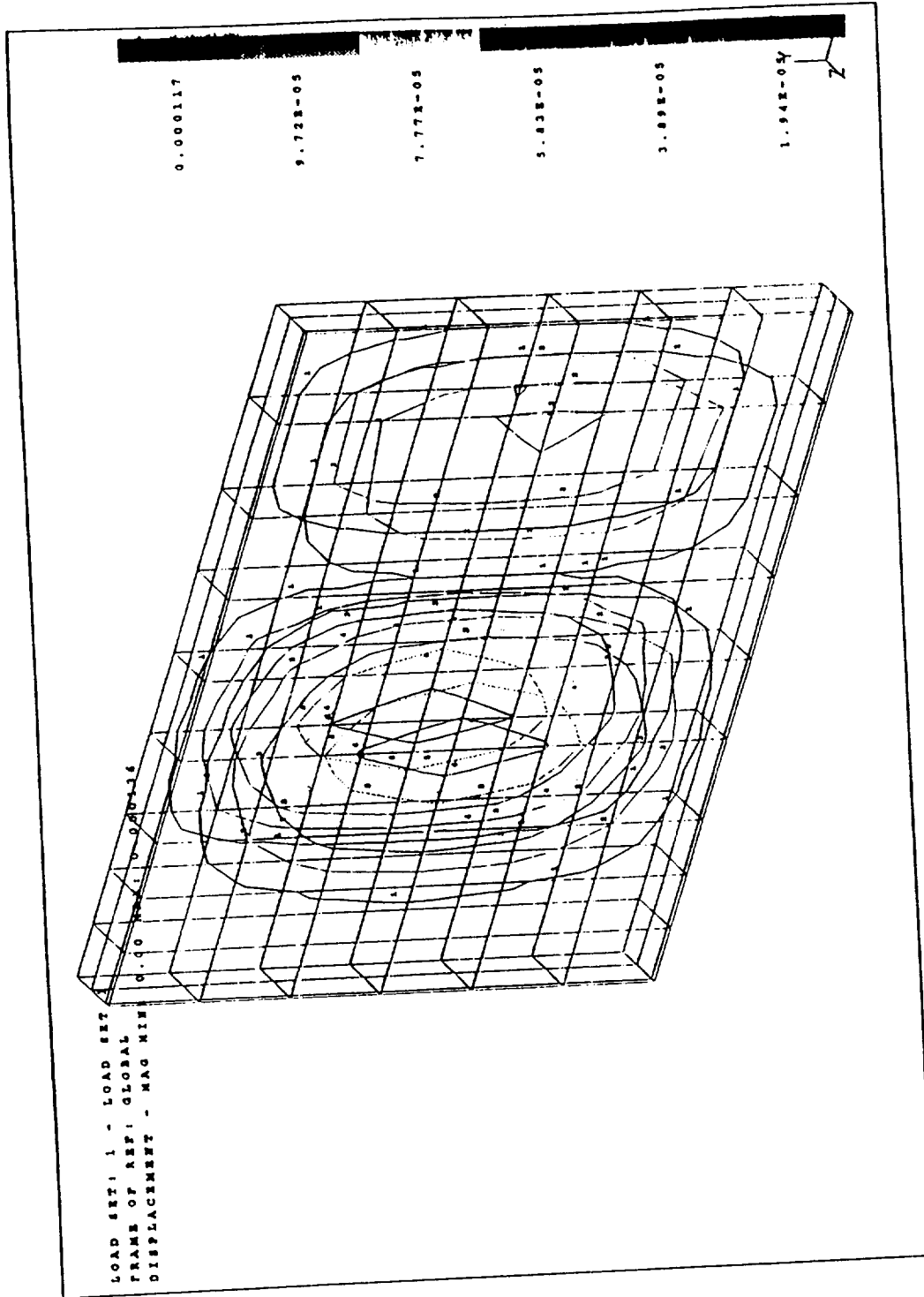


Figure 6 Displacement Contour Lines of Constant Pressure Loading

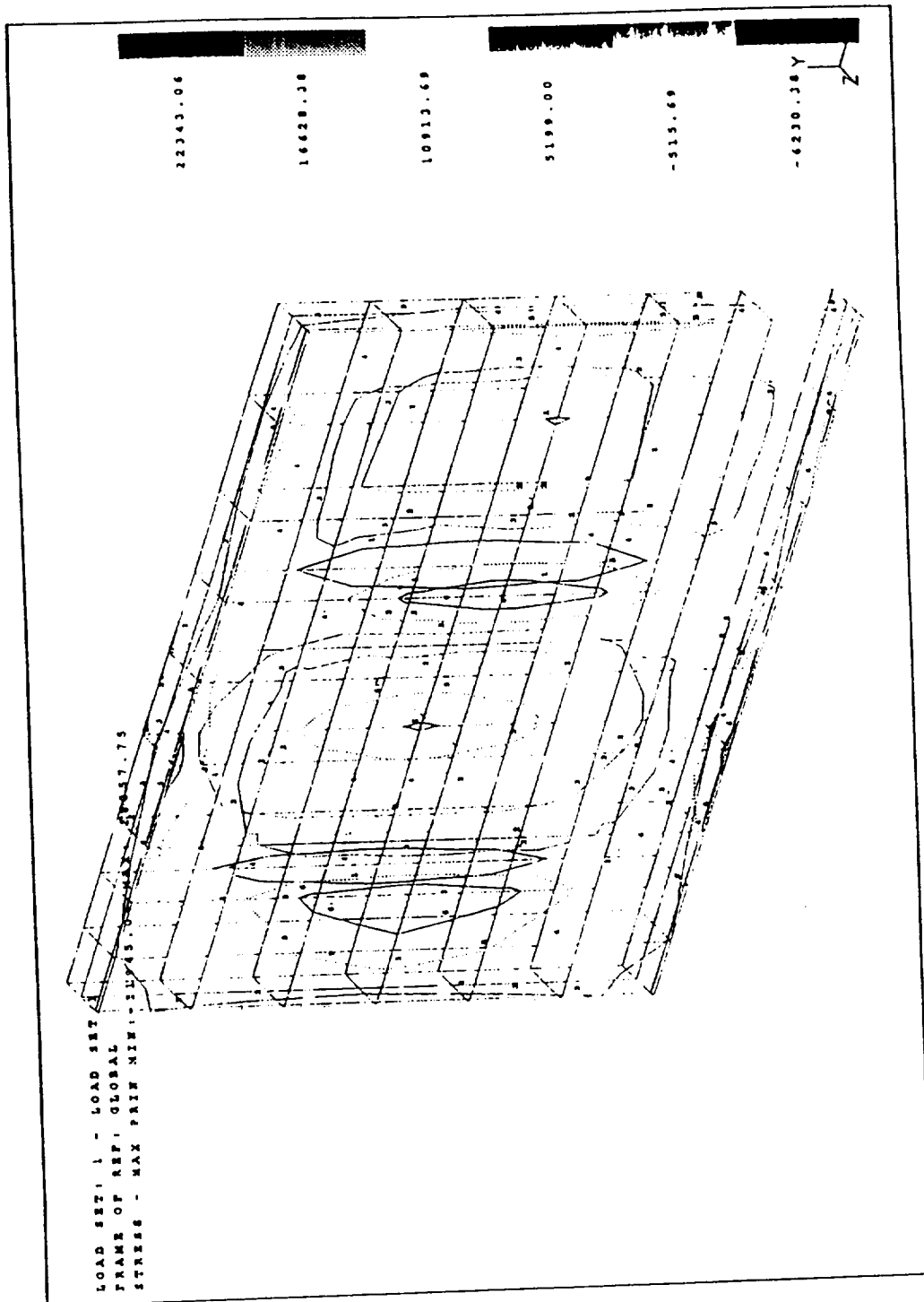


Figure 8 Stress Contour Levels of Constant Pressure Loading

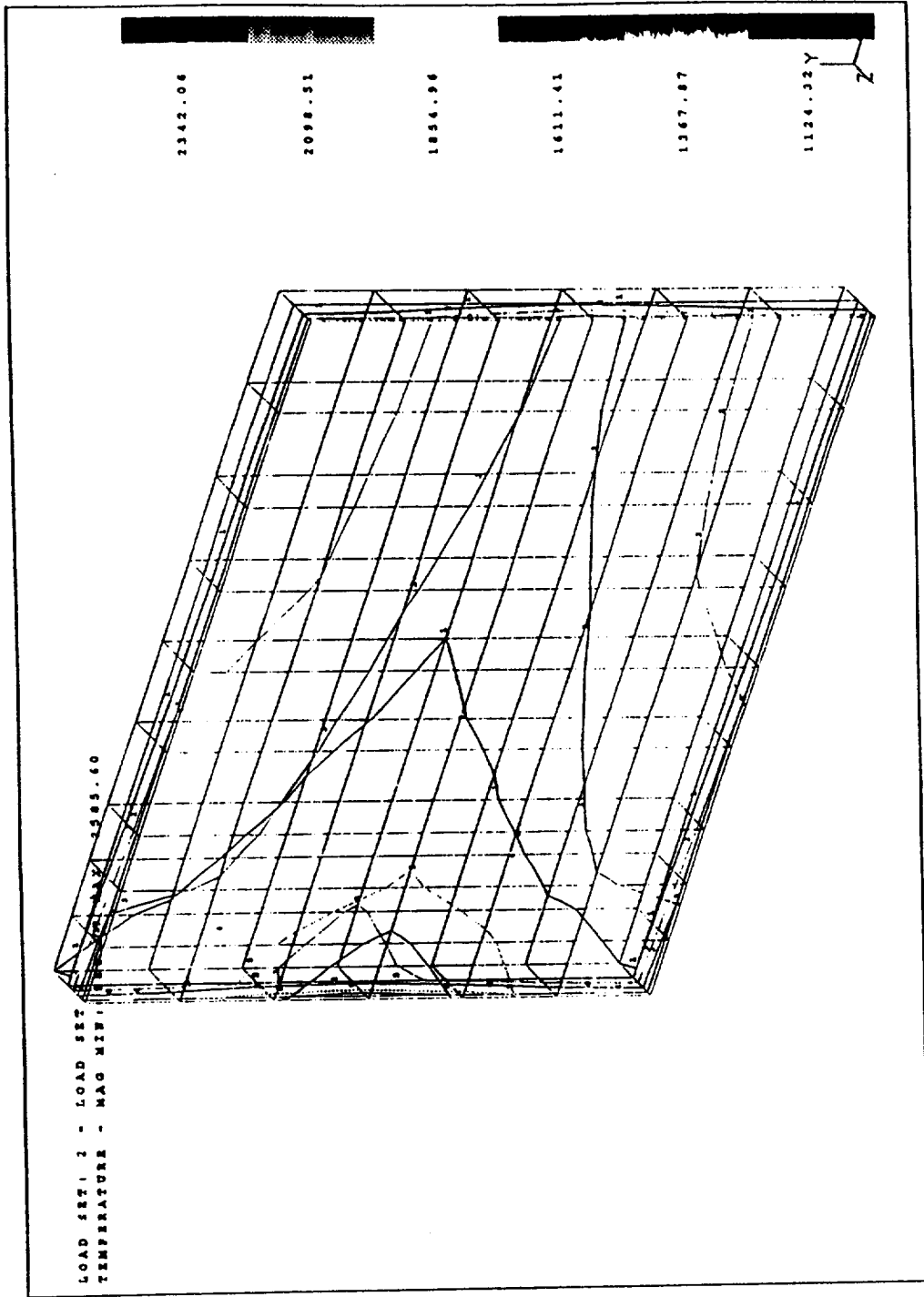


Figure 9 Steady-State Temperature Contours of the FEA Model

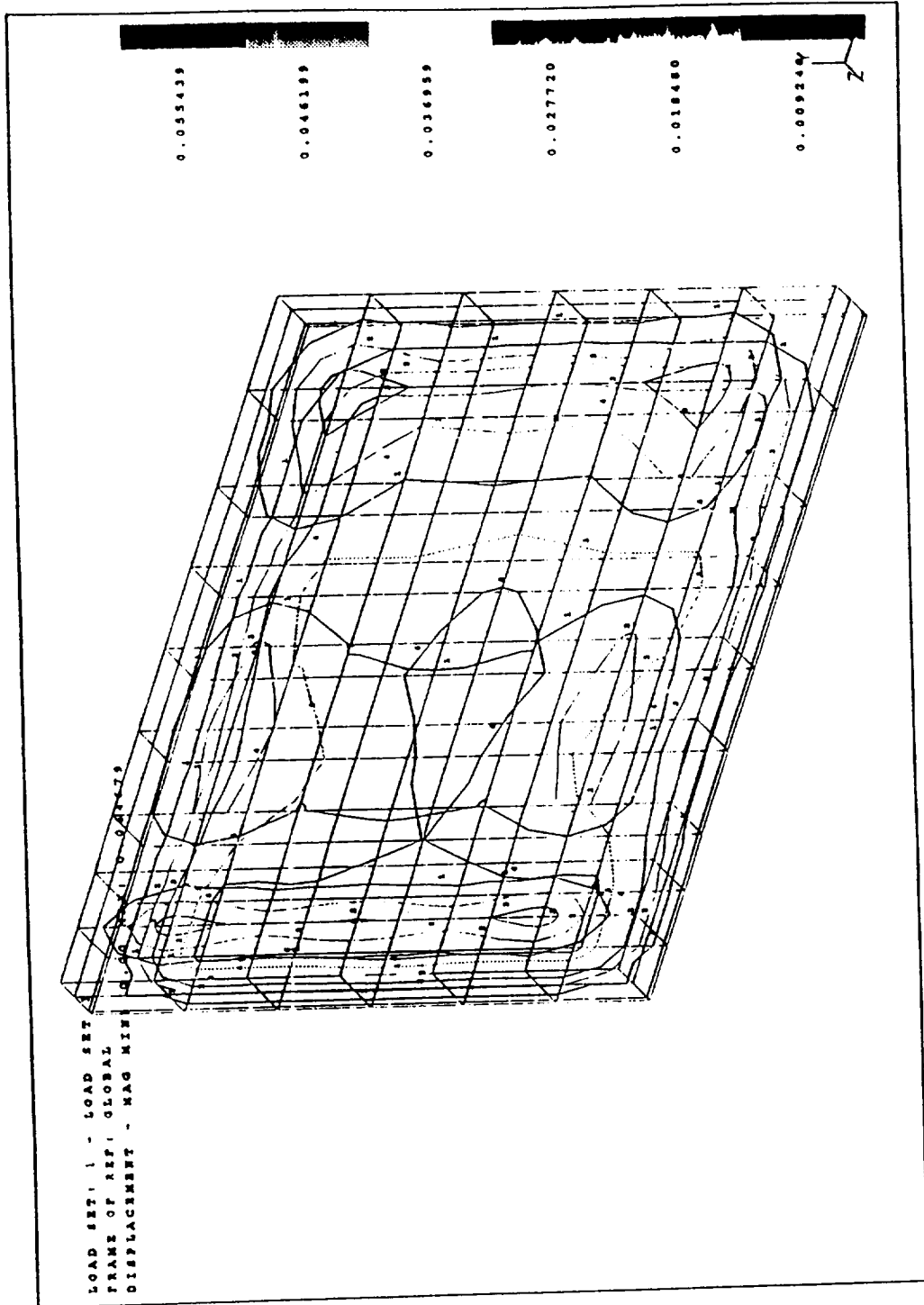


Figure 10 Displacement Contour Lines of Thermal Loads

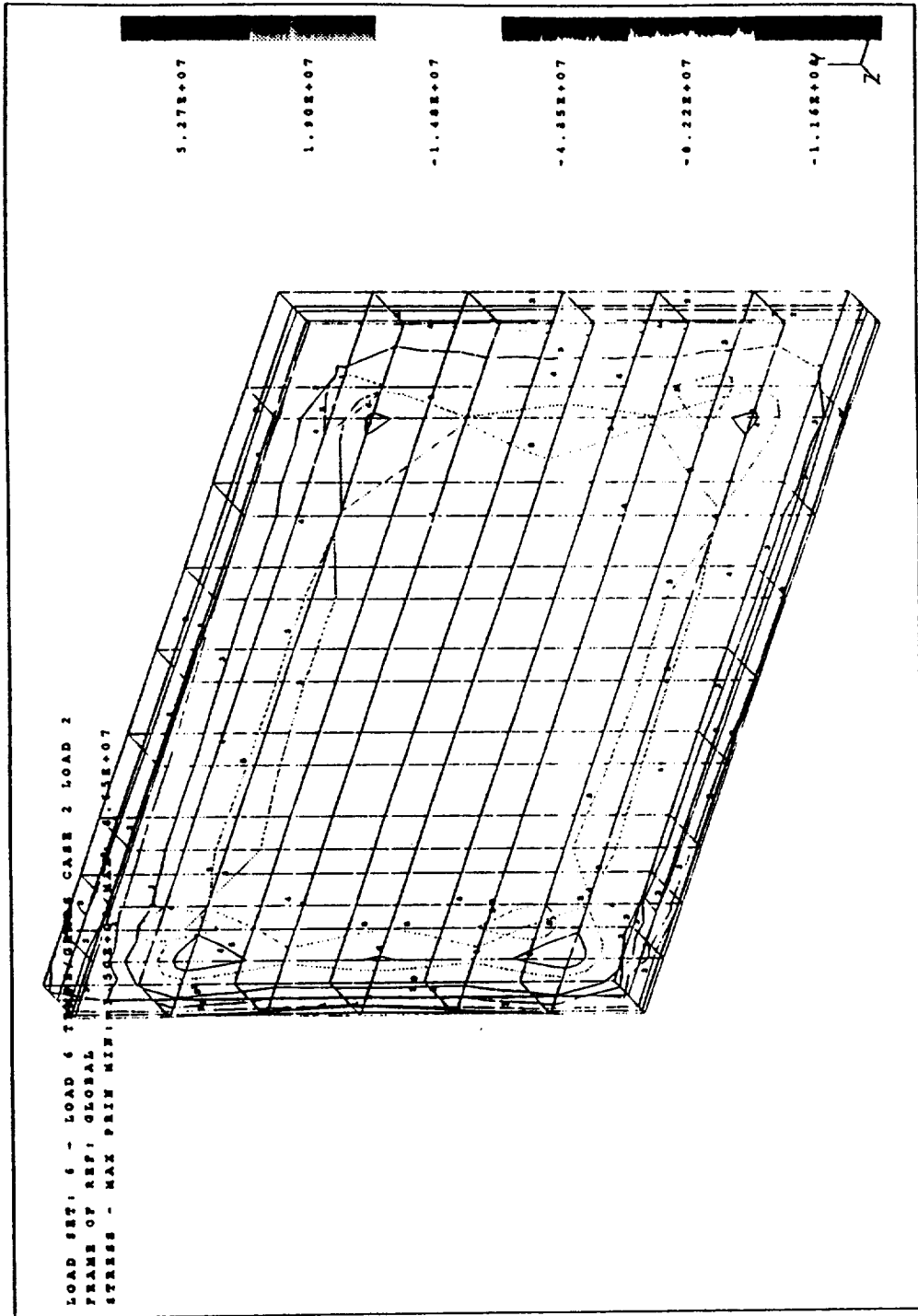


Figure 11 Stress Contour Levels of Thermal Loads

1994018777
N94-22250

1993 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA

520-60

197206

p. 29

SWITCH CONFIGURATION FOR MIGRATION TO OPTICAL FIBER NETWORK

PREPARED BY:	Dr. George W. Zobrist
ACADEMIC RANK:	Professor
UNIVERSITY AND DEPARTMENT:	University of Missouri - Rolla Department of Computer Science
NASA/KSC	
DIVISION:	Data Systems Division
BRANCH:	CAD/CAE Systems
NASA COLLEAGUE:	Fernando Rico-Cusi
DATE:	August 13, 1993
CONTRACT NUMBER:	University of Central Florida NASA-NGT-60002 Supplement: 11

ACKNOWLEDGMENTS

This is to acknowledge the support of Mr. F. Rico-Cusi and Mr. Hank Perkins of DL-DSD-22 and Mr. Shawn Riley and Mr. Mark Sorger of EG&G. They were very helpful in the initial guidance of this research effort and in obtaining the necessary resources. Also, Mr. Bryan Boatright and Mr. Jerry Barnes of DL-ESS-13 for providing insight into the overall KSC network and future plans. Ms. Joy Maldonade, Mr. Patrick O'Rouke and Mr. Ed Bertot of DL-DSS-22 for additional support.

Additionally, I wish to acknowledge the administrative support of Dr. Ray Hosler and Ms. Kari Stiles, University of Central Florida, and Mr. Tom Barron, Training Section of NASA/KSC.

ABSTRACT

The intent of this study is to investigate the migration of an Ethernet LAN segment to fiber optics. At the present time it is proposed to support an Fiber Distributed Data Interface (FDDI) backbone and to upgrade the VAX cluster to fiber optic interface. Possibly some workstations will have an FDDI interface. The remaining stations on the Ethernet LAN will be segmented.

The rationale for migrating from the present Ethernet configuration to a fiber optic backbone is due to the increase in the number of workstations and the movement of applications to a windowing environment, extensive document transfers, and compute intensive applications.

SUMMARY

The Computer Aided Design/Computer Aided Engineering (CAD/CAE) graphics network at Kennedy Space Center is composed of several Local Area Networks (LAN). These LAN's are interconnected through either bridges or routers. At present this LAN is an Ethernet network. The design/engineering workstations are various Intergraph and Digital Equipment Corporation products, mainly. The host is a VAX cluster and there are several Intergraph servers, for plotting/printing/disk.

The rationale for migrating from the present Ethernet configuration to a fiber optic backbone is due to the increase in the number of workstations and the movement of the applications to a windowing environment, extensive document transfers, and compute intensive applications.

In a NASA/KSC report presented in 1988 the Ethernet utilization was under 5% and there were only fourteen (14) Intergraph workstations on the Headquarters LAN. At present utilization of 60-70 % has been observed in short bursts and 30-40 % averaged over longer time periods. There are presently 58 workstations on the NASA/KSC HQ LAN.

The proposed solution is to use intelligent switches. There are various configurations for high speed intelligent switching bridges. They are used by to bridge multiple LAN's, either FDDI, Ethernet, Token Ring or others. The intelligent switch offers many advantages over shared channel LAN's. The advantages include an increase in the bandwidth, latency (propagation delay) reduction, an increase in connectivity, and better traffic management.

This configuration should also increase throughput due to the Ethernet LAN segmentation and the installation of FDDI controllers for the VAX cluster, various Intergraph servers, and several VAX workstations which have a high workload. One also has the option to privatize Ethernet workstations, if the load demands. It should also be noted that other developers have reported that until all workstations are upgraded to FDDI a sizable increase in throughput is usually not recognized, this is due not only to the 10 Mb/s output of the Ethernet controller, but applications are not taking advantage of the higher bandwidth available from FDDI.

TABLE OF CONTENTS

<u>Section</u>	<u>Title</u>
1.	INTRODUCTION
2.	NASA/KSC CAD/CAE GRAPHICS NETWORK CONFIGURATION
3.	RATIONALE FOR MIGRATION TO FDDI
4.	ETHERNET TECHNOLOGY
5.	FDDI TECHNOLOGY
6.	FDDI COMPONENTS
7.	FDDI TOPOLOGIES
8.	FDDI INTERCONNECTION DEVICES
9.	INTELLIGENT SWITCH
10.	WIRING SYSTEM TOPOLOGY
11.	CONFIGURATION FOR MIGRATION TO FDDI
12.	SUMMARY
13.	REFERENCES

LIST OF FIGURES

<u>Figure</u>	<u>Title</u>
1.	VAXcluster - Computer Room
2.	NASA/KSC CAD/CAE LAN - HQ's Building
3.	FDDI Concentrator
4.	Dual Ring Topology
5.	Tree of Concentrators
6.	Dual Homing
7.	SYNERNETICS Intelligent Switch
8.	Wiring Topology
9.	SYNERNETICS Ethernet Switching Module
10.	SYNERNETICS Generic FDDI/Ethernet LAN

1. INTRODUCTION

The Computer Aided Design/Computer Aided Engineering (CAD/CAE) graphics network at Kennedy Space Center is composed of several Local Area Networks (LAN). These LAN's are interconnected through either bridges or routers. At present this LAN is an Ethernet network. The design/engineering workstations are various Intergraph and Digital Equipment Corporation products, mainly. The host is a VAX cluster and there are several Intergraph servers, for plotting/printing/disk.

The workstations use the VAX cluster for their work environment. There are various protocols on the LAN, mainly Transport Control Protocol/ Internet Protocol (TCP/IP) and DecNet, with some XNS (Intergraph protocol).

The intent of this study is to investigate the migration of the Headquarters Building portion of the LAN to fiber optics. At the present time it is proposed to support an Fiber Distributed Data Interface (FDDI) backbone and to upgrade the VAX cluster and Intergraph servers to a fiber optic interface. Possibly some workstations will have an FDDI interface.

In the sections that follow, the following items will be discussed. A review of the NASA/KSC CAD/CAE graphics network configuration, rationale for migration, Ethernet and FDDI principles and nomenclature, intelligent switch concepts, structured cabling and resultant configuration, and presentation of the proposed LAN re-configuration.

2. NASA/KSC CAD/CAE GRAPHICS NETWORK CONFIGURATION

The NASA/KSC CAD/CAE graphics network (1) configuration is composed of a VAX cluster utilizing a Star Coupler tying together a VAX 11/780, VAX 6000-610, and a VAX 6000-510. The VAX 780 and VAX 6510 are to be replaced with an ALPHA 7610 AXP and an ALPHA 4000 AXP (fig. 1). It has not been decided whether the host machines that are presently on-line will be removed or kept in the cluster. The VAX cluster is interfaced to the workstation environment through an Ethernet LAN, and by Bridges/Routers to workstations that are not situated at the Headquarters building (fig. 2).

The NASA/KSC CAD/CAE LAN presently provides connectivity for the CAD/CAE workstations, which are Intergraph, DEC, and PC's. The network communicates between HQ's, O&C, EDL,

CIF, and the Merritt Island Courthouse (MICH) on Broadband Communication Distribution System (BCDS) Channel FM1. There is also a gateway to NSI-DECnet network.

There are several DEC workstations in the Mechanical Engineering area and Boeing has a DEC workstation. These are VAXstation 4060's and 3176's.

The Headquarters CAD/CAE LAN is a single segment Ethernet network and their is presently an FDDI fiber optic ring for the Kennedy Metropolitan Area Network (KMAN). KMAN is to provide connectivity to other sites (in the future) and presently to off-KSC sites.

3. RATIONALE FOR MIGRATION TO FDDI

The rationale for migrating from the present Ethernet configuration to a fiber optic backbone is due to the increase in the number of workstations and the movement of the applications to a windowing environment, extensive document transfers, and compute intensive applications.

In a NASA/KSC report (2) presented in 1988 the Ethernet utilization was under 5% and there were only fourteen (14) Intergraph workstations on the Headquarters LAN. At present utilization of 60-70 % has been observed in short bursts and 30-40 % averaged over longer time periods. There are presently 58 workstations on the NASA/KSC HQ LAN.

This is then the rationale for obtaining an increase in bandwidth to relieve present congestion and provide the capabilities for future growth. It should be noted that in network communications terminology bandwidth is the amount of data that can be transmitted over a channel in bits/second. This is a different definition than used in electrical engineering terminology.

There are several alternatives for providing greater bandwidth for the CAD/CAE LAN. One is through segmentation, this is a reconfiguration of the LAN network into segments whereby one tries to keep traffic local to the segment and only obtain access to other segments if needed. This results in usage of Bridge/Routers to connect the various segments. Propagation delay will be increased every time a Bridge/Router is introduced into the network. Propagation delay is the amount of time between the time the message is sent from the source to being received by the intended destination. In the LAN being investigated it is presumed that most traffic is between the workstations and the VAXcluster, thereby segmenting would not alleviate the

problem to a great degree, since the channel would be utilized between the workstation and the VAXcluster. It can be concluded that poor performance of the LAN under consideration is due to inadequate channel bandwidth and hence a network utilizing an FDDI backbone is warranted.

Segments can be composed of the entire LAN on one segment to having each workstation/host on a segment with appropriate bridging/routing. The entire LAN on one segment is suitable if the traffic does not warrant segmentation, i.e., keeping traffic local. The other extreme is probably useful if just certain workstations/hosts are provided with a private segment, i.e., those components which produce/receive the most traffic, or have highly localized traffic within a workgroup.

A "rule of thumb" to determine whether segmentation will increase LAN capacity is the 80-20 rule, i.e., if local traffic is 80 % of the traffic generated and remote traffic is 20 % then the LAN capacity will increase by segmentation of that workgroup. Of course, segmentation can sometimes increase congestion, if the remote traffic must take several "hops" across different segments to reach the destination. An increase of bandwidth could occur on other segments if there is considerable remote traffic from different sources traversing a common segment to get to their destination.

Another approach, i.e., as compared to segmentation, is the concept of Intelligent Switching. Intelligent switches are able to accommodate Ethernet and FDDI modes and able to switch, between segmented networks either internally or externally, at a very rapid rate. This not only reduces or propagation delays, but allows one to migrate to FDDI rather than configuring for fiber optics entirely.

They also provide concurrent communications between workgroups and can match different bandwidth LAN's through the switch interface. In general one can achieve a high-through-put, low-propagation delay (latency), and transparent communication between end-stations.

In the case of the NASA/CAE LAN this is a reasonable solution for several reasons. One, most of the workstations are not upgradeable to FDDI controllers and the cost would also be prohibitive. Two, the system is not yet saturated but if the workload increases in the future it will be needed. Thirdly, there is a movement to FDDI configurations at KSC and planning to migrate should be initiated.

4. ETHERNET TECHNOLOGY

Ethernet (IEEE 802.3 Carrier Sense Multiple Access with Collision Detection - CSMA/CD) (3) provides the services of the lower two layers in the International Standards Organization (ISO) Open Systems Interconnection (OSI) model for network protocols. There are seven layers in this model.

The layers and a brief description of their functions follow. The lowest layer is the Physical layer which is concerned with transmitting the bits over the transmitting medium, the next layer is the Data Link layer which is concerned with preparing the line for transmission and framing the packets so that there is a delineation of the packet boundaries, addressing, and error detection. This is the layer, along with the Physical layer, for which Ethernet is used. The next layer is the Network layer, this layer determines how packets are routed through the sub-networks. Above this layer is the Transport layer, which mainly fragments the packet into smaller units, if needed, and insures that these fragments will be correctly put back together. The next layer is the Session layer, which is basically the user's interface to the network. The other two layers are the Presentation and Application layers. They are used for tasks, such as data compression and data distribution, respectively.

The Physical layer characteristics for Ethernet are:

- Data rate: 10 Million bits/second
- Maximum station separation: 2.8 Kilometers
- Maximum number of stations: 1024
- Medium: Twisted pair, Coaxial cable, Optical fiber
- Logical topology: Bus
- Physical topology: Star, Bus, Hierarchical Star
- Maximum frame size: 1518 bytes
- Frames on LAN: Single

The Data Link characteristics are:

- Link control procedure: Fully distributed peer protocol, with statistical contention resolution
- Message protocol: Variable frame size, "best effort delivery"

Ethernet is a carrier sense protocol, i.e., all stations monitor the cable during their transmission, terminating transmission immediately if a collision is detected. When an Ethernet station wishes to transmit a packet a carrier

sense is performed forcing the station to defer if any transmission is in progress. If there is no station sensed to be transmitting then the sender can transmit after an appropriate delay. It is possible that two, or more, stations will sense the channel idle at the same time and begin transmitting. This has the possibility of producing a collision. The station will continue monitoring and sense this collision. When a collision is detected the station will stop transmitting and will reschedule a re-transmission at a later time. Re-transmission time is random and is selected using a binary exponential backoff algorithm.

5. FDDI TECHNOLOGY

FDDI is a token passing technology that uses a timed token protocol (4). There can be multiple frames on the network which is configured as a logical dual ring, or a dual ring of trees. The media standard is presently optical fiber, although transmission of the packet over copper is also being considered and should be in the standard, in the future. The designation for the later is Copper Distributed Data Interface (CDDI). The bandwidth is 100 Mb/s. Of course the transmission distance for a predetermined db loss is greater with a fiber optic cable, as compared to a copper cable. There is also concern with cross-talk and radiation with the copper media. These concerns are being addressed, mainly through twisted pair and shielding.

As in Ethernet FDDI provides services at the lower two levels of the OSI model.

A brief description of the communication on the ring is described below.

There are dual fibers in the ring network. One fiber is used as the primary transmission path and the other is the secondary transmission path. The dual rings are independent until a fault appears either in the cable or a component. At that time the rings are joined together to isolate the fault and continue communications. This is called "wrapping". If two faults occur then systems can become isolated.

Stations (components) are connected in series to form a ring. Data is transmitted serially from a station to the next station downstream. Each station performs the repeater function to regenerate the data frame and re-transmit to the next station, while possibly inserting data frames of it's own.

The right to transmit data frames is controlled by a token. The token is passed from station to station, and when a station has the token it can add data to the transmission frame as long as it is allowed to hold the token, this is determined by a system parameter called the token holding time. When a frame returns to the sending station it is stripped from the ring.

In the FDDI standard one can have multiple frames on the LAN, as compared to the IEEE 802.5 communications ring. In IEEE 802.5 once a station has captured the token and inserted data onto the ring, it has the channel captured until the receiving station removes the data and re-sends an appropriate token.

The fiber optic cable can be broadly classified as either Multimode (MMF) or Singlemode (SMF) fiber. Multimode optic fiber means that there will be multiple modes transmitted on the cable. The core sizes are typically, 50, 62.5, or 100 microns. This creates modal dispersion which affects both the distance and the achievable bandwidth. MMF can use inexpensive light sources, as compared to SMF. The main result from using MMF cable is that maximum station separation is considerably less than with SMF cable, but attendant costs for transmission components are cheaper.

SMF has only a single mode of transmission. The core size varies from 8-10 microns, laser's must be used for transmission and there must be precision alignment. Station separation can reach 20 Kilometers, or more with SMF. Normally, SMF cables are used for intra-building connections.

The characteristics of FDDI are:

- Data rate: 100 Million bits/second
- Maximum station separation: 2 Kilometers for MMF or
greater than 20 Kilometers for SMF
- Maximum number of stations: 500
- Medium: Optical fiber and copper shielded twisted pair
- Logical topology: Dual Ring or Dual Ring of Trees
- Physical topology: Star, Ring, Hierarchical Star
- Maximum frame size: 4500 bytes
- Frames on LAN: Multiple
- Access Method: Timed-token passing

6. FDDI COMPONENTS

The FDDI concentrator (3,5) has a basic role in the various FDDI topologies, which will be discussed later. The concentrator is a component that provides a connection point for various end stations in the network and other FDDI devices (fig. 3).

There are four port types used in FDDI components:

- o Port A connects to the incoming primary ring and the outgoing secondary ring of the FDDI dual ring.
- o Port B connects to the outgoing primary ring and the incoming secondary ring of the FDDI dual ring.
- o Port M connects a concentrator to a workstation/host or another concentrator. This port is only implemented in a concentrator.
- o Port S connects a singly attached component to a concentrator.

One can have either dual attached concentrators (DAC) or single attached concentrators (SAC). The DAC is used in the dual ring, while the SAC is connected to the upper level ring through a DAC. A DAC will connect to the dual ring via its A and B ports, while the SAC accesses the dual ring via its S port which is connected to an upper level M port.

The concentrator is a Physical level repeater, in that it allows the connection of hosts/servers/workstations and/or other concentrators to the FDDI dual ring.

There are two types of connections to the FDDI dual ring. One is the single attached station (SAS) and the other is the dual attached station (DAS). The SAS has one S type port and connects to the dual FDDI ring via a M type port of a concentrator. The concentrator that the SAS is connected to has a design feature that allows a defective SAS to be isolated if it faults, through "wrapping".

Dual attached stations connect to both the primary and secondary rings of the FDDI dual ring. A DAS has two ports an A port and a B port. The A port connects to another station's B port, while the B port connects to another station's A port. A DAS doesn't require a concentrator to connect to the FDDI ring like the SAS. A disadvantage of using DAS stations is that when two, or more, DAS's fail in the ring then the stations become isolated. Also, if

several DAS stations are "turned off", without having optical bypass switches, the FDDI network will become isolated. A DAS can be connected to a concentrator by using the B port for the connection.

FDDI controllers connect end user stations to the FDDI network. The controllers can be either SAS or DAS. The controller has an interface to a particular workstation/host or server system bus. It should be noted that not all components have FDDI controllers available, especially if their architecture is of an earlier generation. They can also be quite expensive in comparison to the value of the component to be upgraded.

7. FDDI TOPOLOGIES

The FDDI standard permits a diverse set of topologies (5). In this report several will be discussed. The ones to be presented are the following:

o Standalone FDDI Concentrator Topology

This topology consists of a single attached concentrator from which components can be attached. These stations can be either SAS or DAS. This topology can be used to attach endstations together that form a workgroup. This is a useful topology if the building is already prewired for fiber and one does not wish to install a dual ring FDDI backbone (fig. 3).

o Dual FDDI Ring Topology

This topology consists of dual attached stations being connected directly into the dual FDDI ring. This topology can be useful if there are a small number of stations. The disadvantage is that it does not lend itself well to additions/moves/changes. If a station is disconnected then one has caused a fault in the ring, if this is a single fault "wrap around" will occur. If there is more than one fault then stations will be isolated. This topology should not be used for DAS connections, unless there is little risk of the users disturbing the connection. The utilization of dual attached concentrators in a dual ring topology is an accepted topology, since they are presumed to be in service at all times and this also will provide for expansion (fig. 4). This will be discussed later.

o Tree of Concentrators

In this topology concentrators are wired in a hierarchical star topology. One concentrator is designated as the root of the tree, and other components are connected to it. These components can be SAS, DAS, or other concentrators. In this topology the LAN is not disrupted by disconnecting stations, and/or their movement to other locations. Additional stations can be added by attaching another level of concentrators. This topology is best suited for a single building FDDI LAN. This topology tends to be more manageable, since concentrators tend to isolate faults from the rest of the system (fig.5).

o Dual Ring of Trees

This topology is a merging of the dual ring and the tree of concentrator configurations. Basically, one connects distinct tree of concentrator topologies in a dual ring. This dual ring functions as the FDDI backbone. This is a suitable topology when buildings will be connected via a FDDI backbone. It has the advantages and disadvantages of both merged types. It is the most flexible topology since it has both fault tolerance and increases the availability of the FDDI dual ring backbone (fig. 4 and 5).

o Dual Homing

The FDDI standard allows for redundant paths in tree topologies. This redundancy is called "dual homing". Basically, this topology is constructed as follows. The dual ring has at least two concentrators in it. An M port from one concentrator is connected to the A port of a DAS (which is in the next hierarchical level) and an M port of the other concentrator is connected to the B port of the DAS. The B port connection is the primary connection to the DAS, while the A port connection of the DAS is the backup connection. If the primary connection fails then the backup connection will be activated. This concept can also be applied to two M port connections from the same concentrator. The concept can be further diversified by putting the concentrators that are being used in the dual homing concept on different FDDI rings. This concept can be utilized for single attached stations by installing two single attachment interface cards. This configuration is not supported in the FDDI standard (fig. 6).

8. FDDI INTERCONNECTION DEVICES

Bridges and Routers are interconnection devices for FDDI networks (5). Bridges act as Data Link relay between networks. In our case the bridge would have to connect FDDI to Ethernet LAN's. Bridges are protocol-independent and basically store and forward devices.

Basic FDDI to Ethernet bridge functions include:

- o Source address tracking: This will allow the bridge to determine when to forward a packet and when to keep local.
- o Frame forwarding and filtering: Frame forwarding is achieved through table lookup, while filtering is the process whereby frames are prevented from crossing the bridge.
- o Spanning tree: The logical topology of an extended LAN must be loop free, i.e., there must be a single path between all attached stations. To prevent logical loops bridges must form a logical configuration called a spanning tree.
- o Translation: The bridge must modify the fields of the forwarded frame to make it compliant with the format of the network to which it is being sent.
- o Fragmentation: Frames may have to be fragmented when moving across the bridge boundary. FDDI frames can be up to 4500 bytes, while Ethernet frames are restricted to a maximum of 1518 bytes. Therefore, the frame from FDDI to Ethernet may have to be fragmented.
- o Bridge management: Ability to monitor conditions and implement filters on extended networks.

9. INTELLIGENT SWITCH

There are various configurations for high speed intelligent switches (6). They are used to interconnect multiple LAN's, either FDDI, Ethernet, Token Ring or others. The intelligent switch offers many advantages over shared channel LAN's. The advantages include an increase in the bandwidth, latency (propagation delay) reduction, an increase in connectivity, and better traffic management.

Depending upon the vendor the switch may/may not interconnect various communication standards internally. Some of the configurations are:

- o Ethernet to Ethernet switching
- o FDDI to FDDI switching
- o FDDI to Ethernet to Token Ring switching externally
- o FDDI to Ethernet switching internally

There are switches which allow only Ethernet to Ethernet or Token Ring to be internetworked by bridging.

The FDDI to FDDI switching configuration is basically a FDDI concentrator. One can typically purchase FDDI line cards with two (2) or more ports. These ports would support SAS or DAS devices, or presumably SAC or DAC concentrators. Through the purchase of appropriate bridges FDDI and Ethernet segments can be interconnected. These switches can set up concurrent connections to obtain an aggregate throughput much higher than a single segment could obtain. These switches achieve low latency by not utilizing the store and forward concept, but to use cut-through forwarding. This technique forwards a packet as soon as the destination address is determined from the header.

Another switching configuration has backplanes for various communication LAN's, such as FDDI, Ethernet, and Token Ring, i.e., internal to the switch is a collapsed network backbone appropriate to the technology to be utilized. One could connect SAS/DAS stations to the internal FDDI dual ring, Ethernet workstations to the internal Ethernet LAN via an Ethernet card, and also suitable connections for Token Ring. Then externally through appropriate bridging one can interconnect the various communication configurations. This offers more flexibility than an FDDI switch, but the switching speed is determined by the manufacturer of the bridge.

Both of the above solutions offer the advantage of non-proprietary solutions for obtaining interconnectivity. Therefore, if bridge technology is upgraded one could purchase a bridge with that technology and integrate into the network. The disadvantage is that numerous components will have to be purchased to integrate the network into the switch environment and overall propagation delay will probably be larger, than if the bridging/interconnection is integrated into the switch.

Another type of switch can be called the intelligent switch, in that the internal configuration is such that FDDI can be integrated with Ethernet communications (fig. 7). The concept is to have a collapsed FDDI backbone internal to the switch and be able to bridge from external FDDI or Ethernet stations through the FDDI backbone. There is also the possibility of switching at the module level without going through the FDDI backbone for the Ethernet module. The FDDI module must go through the FDDI backbone internal to the switch.

Each Ethernet module contains ports which can have either Ethernet LAN segments connected or a private Ethernet channel, i.e., an end-station. Ethernet segments attached to a unique module are switched by an internal bridging function to the appropriate output port. Ethernet segment connections for ports on separate modules must go through the FDDI internal backbone to arrive at the destination address. The same is true for FDDI SAS/DAS connections.

This allows very sophisticated interconnections between dissimilar LAN segments and also allows gradual migration to FDDI devices as bandwidth needs increase. The communication between Ethernet and FDDI is transparent. Due to the usage of the FDDI internal backbone (backplane) there is a maximum of two low latency "hops" between any two stations.

Normally, a switch will have filtering capability based on; source address, destination address, protocol type, or some combination of these attributes. This can be usually done on a per port basis, or workgroup. Some routing functions can be obtained through this capability.

10. WIRING SYSTEM TOPOLOGY

The wiring topology (fig. 8) is divided into several subsystems (5):

- o Campus backbone system
- o Building backbone subsystem
- o Horizontal subsystem
- o Work area group subsystem

The campus backbone system links clusters of buildings together within a site. The building backbone subsystem provides the link between the campus backbone and the horizontal areas. The horizontal subsystem provides the connection from the building backbone to the work area group. The work area group subsystem connects the user end

stations to the LAN system. Typically, there is also an administrative subsystem for managing the other subsystems and their various components.

One could propose that the FDDI KMAN which is currently in operation would be the campus backbone and the configuration which is proposed in the next section would comprise the building/horizontal/work area subsystems. This is not completely true, since there are numerous other LAN's for communication, specifically the BCDS. In fact, the BCDS is proposed as a backup network for the proposed configuration.

11. CONFIGURATION FOR MIGRATION TO FDDI

The present Ethernet LAN in the Headquarters building is a single segment LAN with bridge/router connections to other CAD/CAE LAN's and other parts of the KSC network. To provide capabilities for migration to FDDI when resources permit and loading necessitates, the intelligent switching configuration is proposed.

This configuration consists of a Building switch and several Workgroup switches for the third floor of the building, and a proposed switch for the first floor, where an additional LAN is present. This configuration will allow migration to FDDI when workstations are upgraded to FDDI. It will also allow the Ethernet LAN to be segmented, which should provide greater access for each segment to the VAX cluster. Components of the VAX cluster and the various servers have FDDI controllers available and hence will be integrated into the building switch through a concentrator module. The connection to the Metropolitan Area Network will be provided by a Router.

The VAX cluster has the capability of utilizing both Ethernet and FDDI controllers, where one would be designated as the primary and the other the secondary. By utilizing this option the existing network would remain in usage. This would provide a backup if the building switching hub is out of service, or if there is overloading both entry points to the network could possibly be utilized.

The proposed intelligent switch is from the Synernetics Corporation and has four modules available (7):

- o System Processor Module (SPM)
- o FDDI Enterprise Access Module (FEAM)
- o FDDI Concentrator Module (FCM)
- o Ethernet Switching Module (ESM)

The SPM module is dedicated to the management of the system and it continually monitors the system and is used to configure the system. This module is required.

The FEAM provides A/B ports for connecting the switch to an FDDI backbone.

The FCM is an FDDI concentrator and allows one to connect end stations and other intelligent switches to the FDDI backbone. The FCM has either four (4) or six (6) M ports

The ESM has eight (8) ports for Ethernet connections which can be switched and a fully translational Ethernet to FDDI bridge which can forward messages to other ESM modules or to FDDI stations via the FDDI collapsed backbone internal to the intelligent switch. Messages can be switched between ports on an ESM without going through the FDDI backbone (fig. 9).

There are two switch sizes, one has four (4) slots and the other has twelve (12) slots available. The modules can be utilized in both switches.

To achieve reliability in the switch environment, the horizontal switches are dual homed off separate concentrator modules in the building switch. This will allow alternate access if one of the concentrator modules becomes inoperative. The building switch also has provision for dual power supplies. One could utilize another horizontal switch dual homed off of the building switch to integrate the VAX cluster and various servers into the network, but this would provide another failure point and would reduce the reliability. Another possible solution to provide FDDI redundancy for the VAX cluster would be to provide two SAS FDDI controllers and dual home into these, if the VAX's can operate in this manner. At present there is no DAS FDDI controllers available for the VAX cluster, there are DAS FDDI controllers available for the Intergraph servers that are in the present Ethernet network and for some of the Intergraph workstations.

Spare modules and a spare switch are to be purchased to ensure that minimal down time would result from a failure. Since this network is not mission critical these precautions should be sufficient at this time.

This configuration, which is a dual ring of trees topology, should provide a reliable network for the CAD/CAE design/development environment and also provide a path for a

complete migration to a FDDI network environment as resources become available and/or the network becomes saturated (fig. 10).

13. SUMMARY

The proposed intelligent switch configuration provides a path for migration from Ethernet to FDDI as resources become available and loading necessitates greater bandwidth. The intelligent switch configuration also provides a gateway to Asynchronous Mode Transmission (ATM) when it becomes available in the future. The proposed solution also is redundant due to dual homing of the horizontal switches, purchase of spares, and dual power supplies on the building switch.

This configuration should also increase throughput due to the Ethernet LAN segmentation and the installation of FDDI controllers for the VAX cluster, various Intergraph servers, and several VAX workstations which have a high workload. One also has the option to privatize Ethernet workstations if the load demands. It should also be noted that other developers have reported that until all workstations are upgraded to FDDI a sizable increase in throughput is usually not recognized, this is due not only to the 10 Mb/s output of the Ethernet controller, but applications are not taking advantage of the higher bandwidth available from FDDI.

Several items should be addressed in future work:

- o A simulation of the proposed system should be done to determine throughputs/latency/utilization and a comparison to various alternative systems, so that the configuration can be "fine-tuned".
- o A reliability study of various configurations so that one can intelligently predict the redundancy that is required for a certain level of reliability.
- o There should be a network image constructed, which will be updated automatically, if possible, as additions / deletions / movement / upgrades of components are done.

14. REFERENCES

1. Zobrist, G. W., "Performance Evaluation of NASA/KSC CAD/CAE Graphics Local Area Network", NASA/KSC Research Report, NASA Grant NGT-60002, October 1988.
2. Kennedy Space Center Network Handbook, KSC-DL-3572, March 17, 1993.
3. Stallings, W., Handbook of Computer Communications Standards: Volume 2 - Local Network Standards, Howard W. Sams Company, Indianapolis, IN, 1987.
4. Ross, F. E., "FDDI - a Tutorial", IEEE Communications Magazine, Vol. 24, No. 5, May 1986, pp. 10-17.
5. Mirchandani, S. and Khanna, R., FDDI - Technology and Applications, John Wiley and Sons, New York, NY, 1993.
6. Herman, J. and Serjak, C., "ATM Switches and Hubs Lead the Way to a New Era of Switched Internetworks", Data Communications, March 1993, pp. 69-84.
7. Switching Hubs: An Integral Part of ATM Implementation, Synernetics Corporation, 85 Rangeway Road, North Billerica, MA 01862.

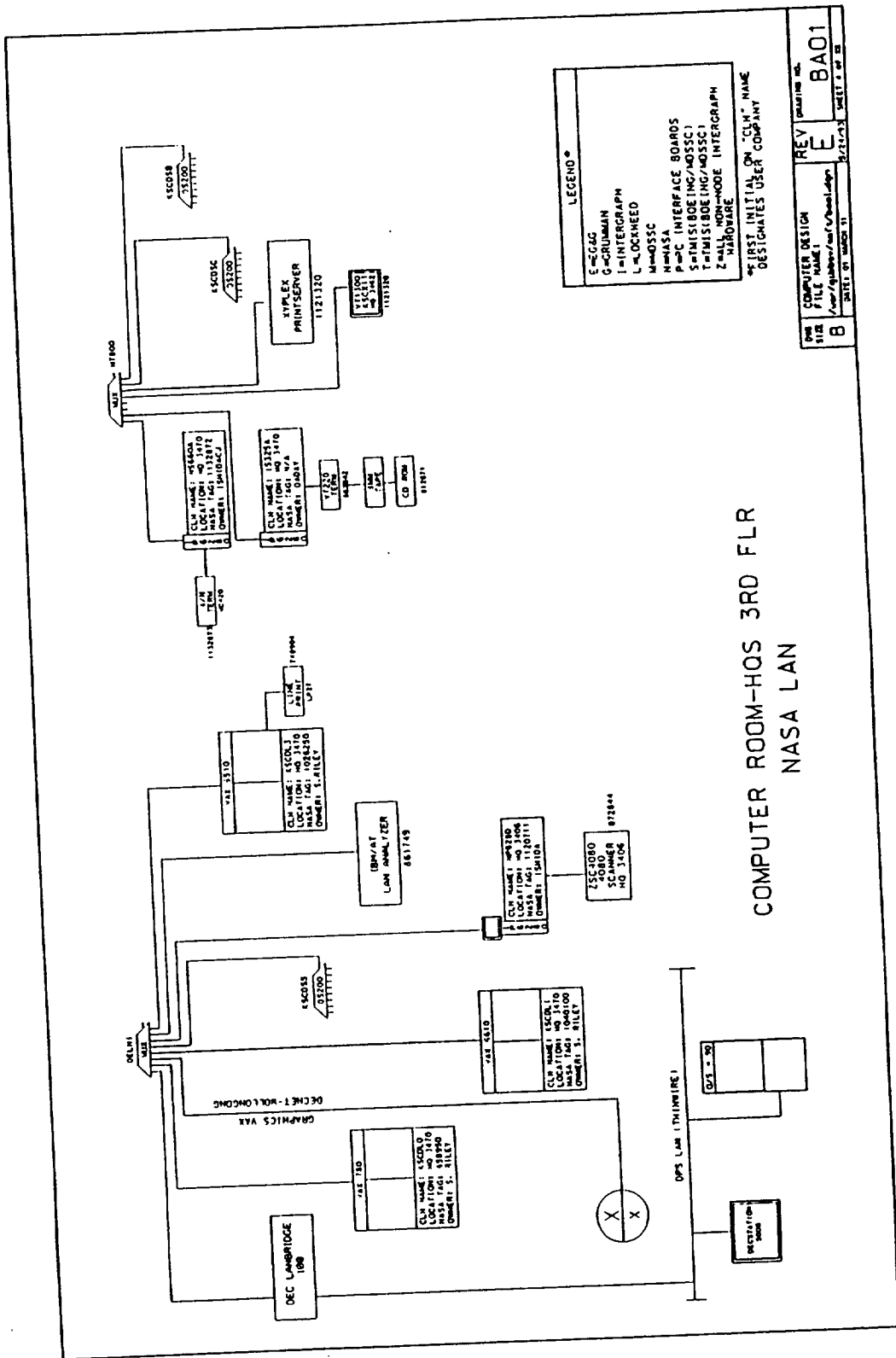
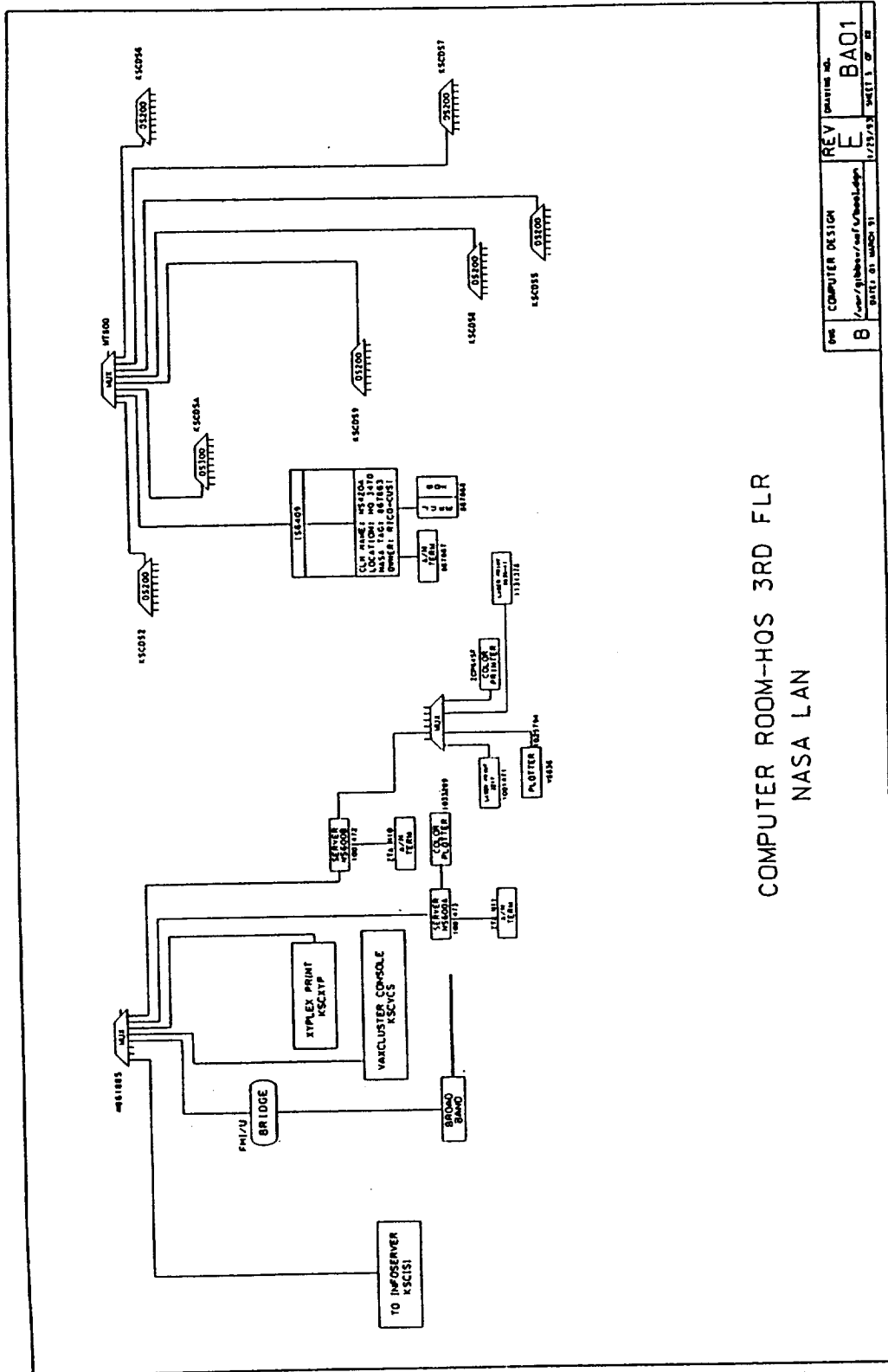


Figure 1. VAXcluster - Computer Room



COMPUTER ROOM-HQS 3RD FLR
NASA LAN

REV	COMPUTER DESIGN	DATE: 01 MAR 83
B	user/rybar/caf/whed/ldp	
E		
BA01		

Figure 1. VAXcluster - Computer Room (Continued)

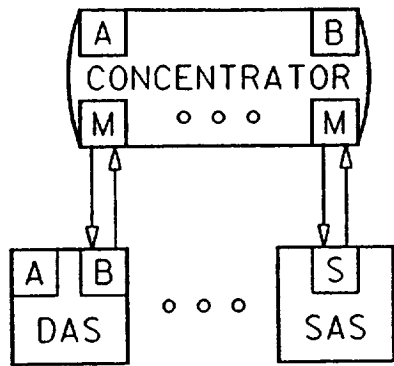


Figure 3. FDDI Concentrator

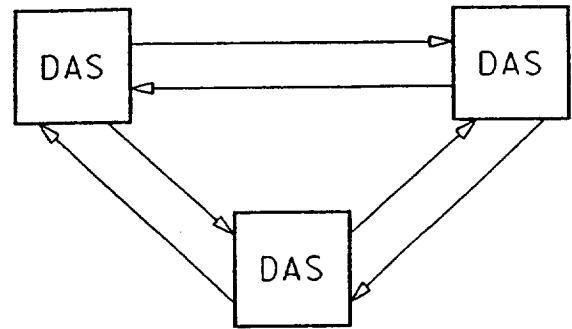


Figure 4. Dual Ring Topology

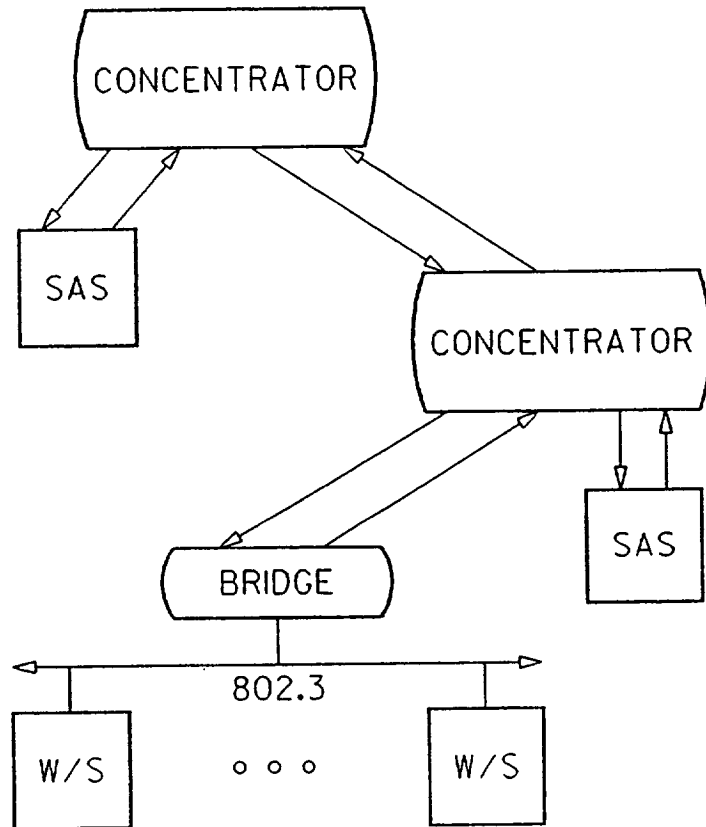


Figure 5. Tree of Concentrators

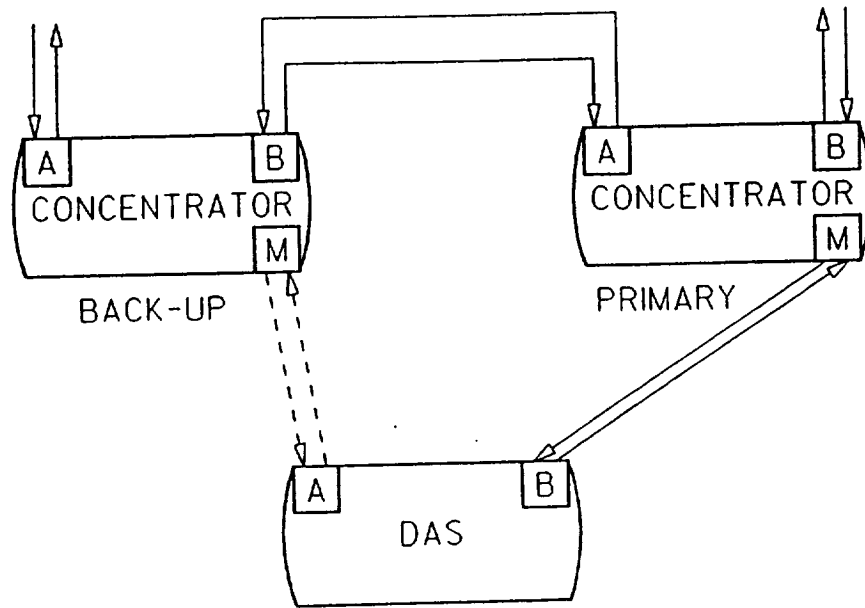


Figure 6. Dual Homing

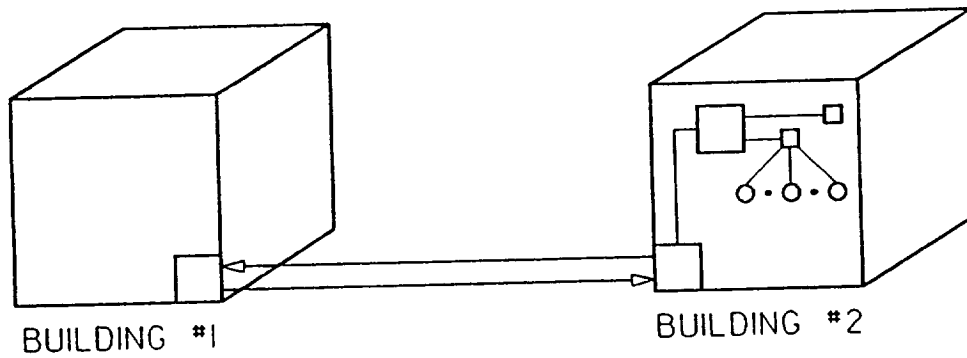


Figure 8. Wiring Topology

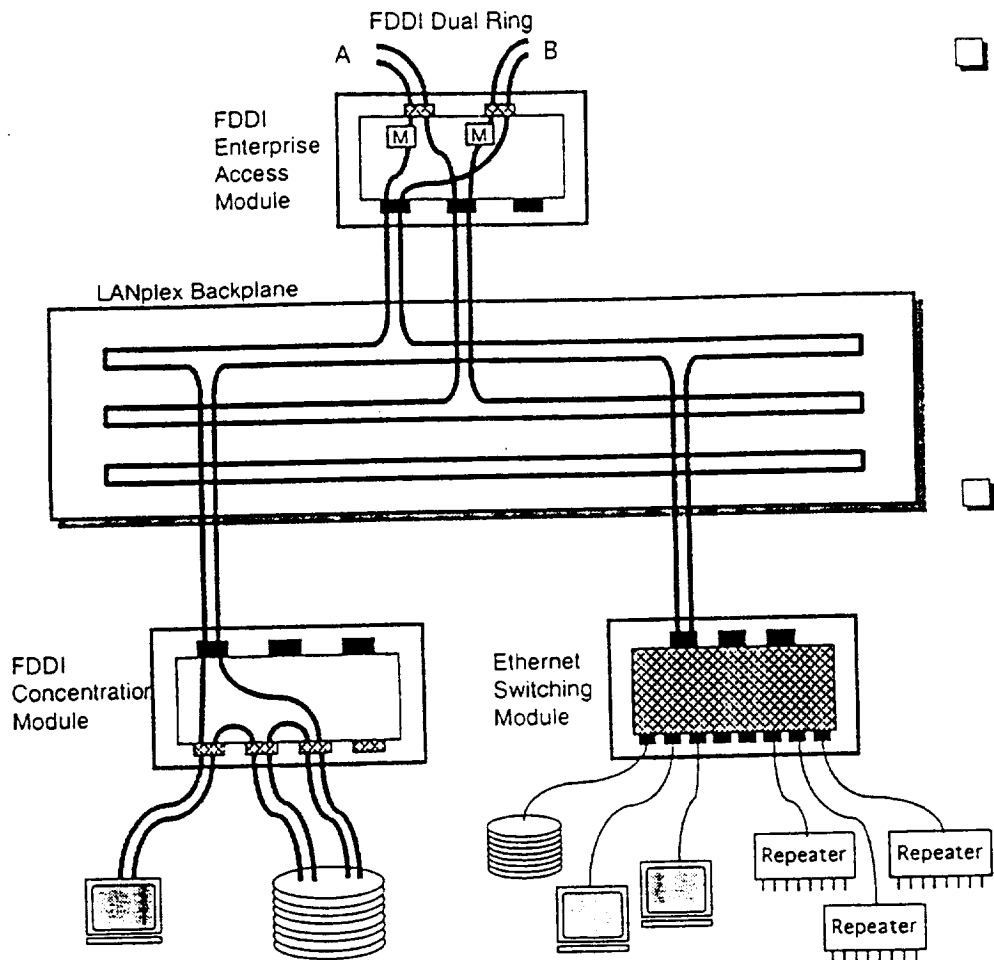


Figure 7. SYNERNETICS Intelligent Switch

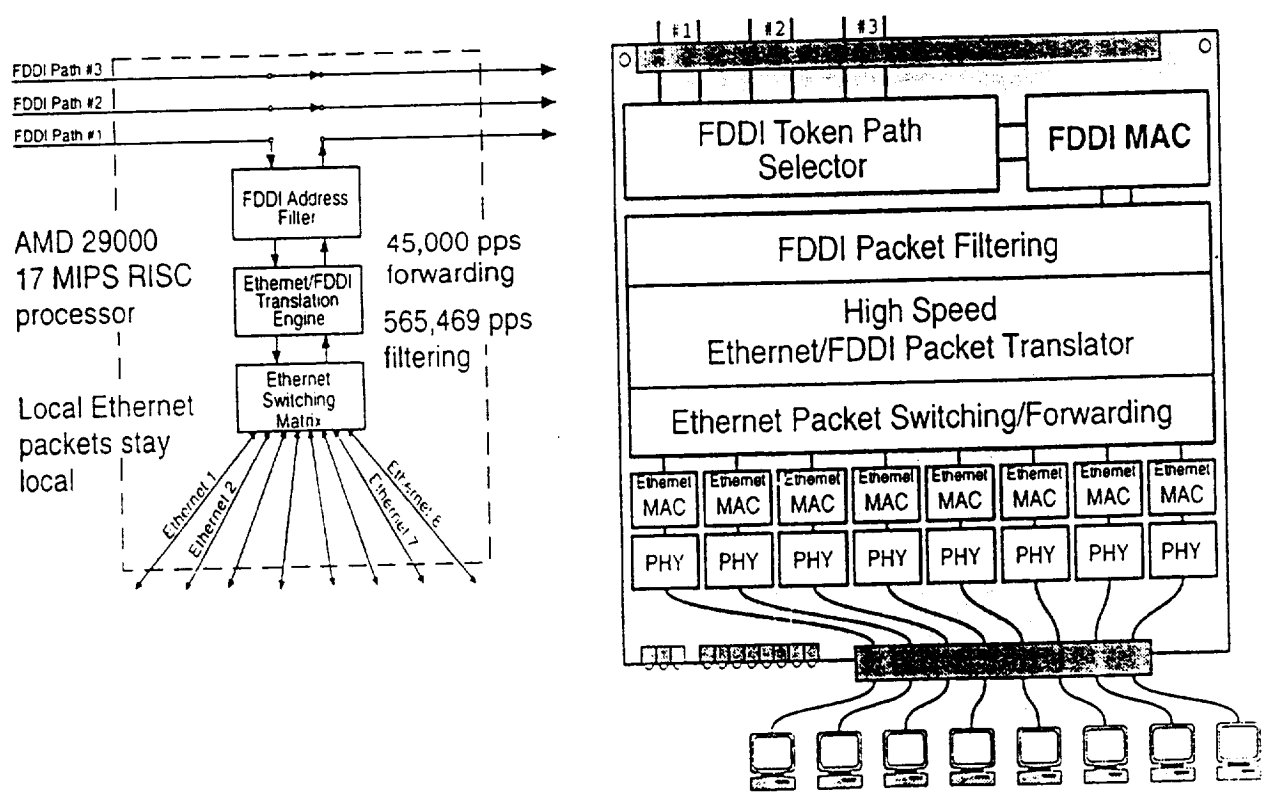


Figure 9. SYNERNETICS Ethernet Switching Module

NASA CAD/CAE FIBER OPTIC LAN

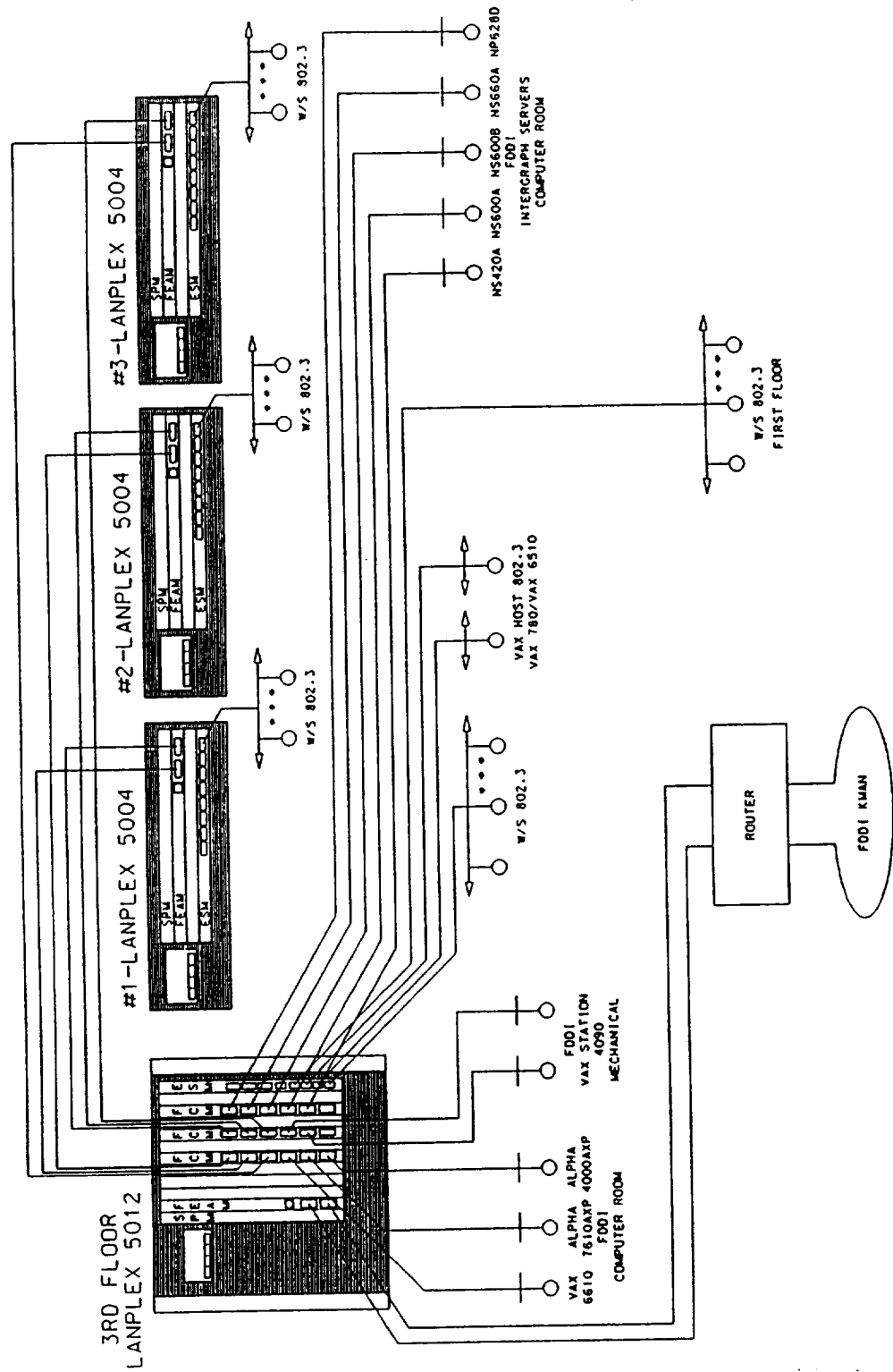


Figure 10. SYNERNETICS Generic FDDI/Ethernet LAN

