UNITED STATES AIR FORCE

SUMMER RESEARCH PROGRAM -- 1996

GRADUATE STUDENT RESEARCH PROGRAM FINAL REPORTS

VOLUME 6

ARNOLD ENGINEERING DEVELOPMENT CENTER
UNITED STATES AIR FORCE ACADEMY
WILFORD HALL MEDCIAL CENTER
AIR LOGISTIC CENTERS

RESEARCH & DEVELOPMENT LABORATORIES
5800 Uplander Way
Culver City, CA 90230-6608

Program Director, RDL Gary Moore	Program Manager, AFOSR Major Linda Steel-Goodwin
Program Manager, RDL Scott Licoscos	Program Administrator, RDL Johnetta Thompson

Program Administrator, RDL
Rebecca Kelly

Submitted to:

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
Bolling Air Force Base
Washington, D.C.
December 1996

AQM 01-06-1272

20010321 053

REPORT DOCUMENTATION PAGE

AFRL-SR-BL-TR-00-

0735

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions,
the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including s
Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, P

nd reviewing
Information

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
December, 1996

3. REP

4. TITLE AND SUBTITLE
1996 Summer Research Program (SRP), Summer Faculty Research Program (SFRP), Final Reports, Volume 6, Arnold Eng. Development Center, US Air Force Academy, Wilford Hall Medical Center, and Air Logistic Centers

5. FUNDING NUMBERS
F49620-93-C-0063

6. AUTHOR(S)
Gary Moore

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Research & Development Laboratories (RDL)
5800 Uplander Way
Culver City, CA 90230-6608

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Office of Scientific Research (AFOSR)
801 N. Randolph St.
Arlington, VA 22203-1977

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
Approved for Public Release

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The United States Air Force Summer Research Program (USAF-SRP) is designed to introduce university, college, and technical institute faculty members, graduate students, and high school students to Air Force research. This is accomplished by the faculty members (Summer Faculty Research Program, (SFRP)), graduate students (Graduate Student Research Program (GSRP)), and high school students (High School Apprenticeship Program (HSAP)) being selected on a nationally advertised competitive basis during the summer intersession period to perform research at Air Force Research Laboratory (AFRL) Technical Directorates, Air Force Air Logistics Centers (ALC), and other AF Laboratories. This volume consists of a program overview, program management statistics, and the final technical reports from the SFRP participants at the Arnold Engineering Development Center, US Air Force Academy, Wilford Hall Medical Center, and Air Logistic Centers.

14. SUBJECT TERMS
Air Force Research, Air Force, Engineering, Laboratories, Reports, Summer, Universities, Faculty, Graduate Student, High School Student

15. NUMBER OF PAGES

16. PRICE CODE

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

PREFACE

Reports in this volume are numbered consecutively beginning with number 1. Each report is paginated with the report number followed by consecutive page numbers, e.g., 1-1, 1-2, 1-3; 2-1, 2-2, 2-3.

This document is one of a set of 16 volumes describing the 1996 AFOSR Summer Research Program. The following volumes comprise the set:

VOLUME	TITLE
1	Program Management Report
	Summer Faculty Research Program (SFRP) Reports
2A & 2B	Armstrong Laboratory
3A & 3B	Phillips Laboratory
4	Rome Laboratory
5A, 5B, & 5C	Wright Laboratory
6	Arnold Engineering Development Center, United States Air Force Academy, Wilford Hall Medical Center and Air Logistics Centers
	Graduate Student Research Program (GSRP) Reports
7A & 7B	Armstrong Laboratory
8	Phillips Laboratory
9	Rome Laboratory
10A & 10B	Wright Laboratory
11	Arnold Engineering Development Center, United States Air Force Academy, Wilford Hall Medical Center, and Wright Patterson Medical Center
	High School Apprenticeship Program (HSAP) Reports
12A & 12B	Armstrong Laboratory
13	Phillips Laboratory
14	Rome Laboratory
15A&15B	Wright Laboratory
16	Arnold Engineering Development Center

APPENDICIES:

GSRP FINAL REPORTS

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

INTRODUCTION

The Summer Research Program (SRP), sponsored by the Air Force Office of Scientific Research (AFOSR), offers paid opportunities for university faculty, graduate students, and high school students to conduct research in U.S. Air Force research laboratories nationwide during the summer.

Introduced by AFOSR in 1978, this innovative program is based on the concept of teaming academic researchers with Air Force scientists in the same disciplines using laboratory facilities and equipment not often available at associates' institutions.

The Summer Faculty Research Program (SFRP) is open annually to approximately 150 faculty members with at least two years of teaching and/or research experience in accredited U.S. colleges, universities, or technical institutions. SFRP associates must be either U.S. citizens or permanent residents.

The Graduate Student Research Program (GSRP) is open annually to approximately 100 graduate students holding a bachelor's or a master's degree; GSRP associates must be U.S. citizens enrolled full time at an accredited institution.

The High School Apprentice Program (HSAP) annually selects about 125 high school students located within a twenty mile commuting distance of participating Air Force laboratories.

AFOSR also offers its research associates an opportunity, under the Summer Research Extension Program (SREP), to continue their AFOSR-sponsored research at their home institutions through the award of research grants. In 1994 the maximum amount of each grant was increased from $20,000 to $25,000, and the number of AFOSR-sponsored grants decreased from 75 to 60. A separate annual report is compiled on the SREP.

The numbers of projected summer research participants in each of the three categories and SREP "grants" are usually increased through direct sponsorship by participating laboratories.

AFOSR's SRP has well served its objectives of building critical links between Air Force research laboratories and the academic community, opening avenues of communications and forging new research relationships between Air Force and academic technical experts in areas of national interest, and strengthening the nation's efforts to sustain careers in science and engineering. The success of the SRP can be gauged from its growth from inception (see Table 1) and from the favorable responses the 1996 participants expressed in end-of-tour SRP evaluations (Appendix B).

AFOSR contracts for administration of the SRP by civilian contractors. The contract was first awarded to Research & Development Laboratories (RDL) in September 1990. After completion of the 1990 contract, RDL (in 1993) won the recompetition for the basic year and four 1-year options.

2. PARTICIPATION IN THE SUMMER RESEARCH PROGRAM

The SRP began with faculty associates in 1979; graduate students were added in 1982 and high school students in 1986. The following table shows the number of associates in the program each year.

YEAR	SRP Participation, by Year			TOTAL
	SFRP	GSRP	HSAP	
1979	70			70
1980	87			87
1981	87			87
1982	91	17		108
1983	101	53		154
1984	152	84		236
1985	154	92		246
1986	158	100	42	300
1987	159	101	73	333
1988	153	107	101	361
1989	168	102	103	373
1990	165	121	132	418
1991	170	142	132	444
1992	185	121	159	464
1993	187	117	136	440
1994	192	117	133	442
1995	190	115	137	442
1996	188	109	138	435

Beginning in 1993, due to budget cuts, some of the laboratories weren't able to afford to fund as many associates as in previous years. Since then, the number of funded positions has remained fairly constant at a slightly lower level.

3. RECRUITING AND SELECTION

The SRP is conducted on a nationally advertised and competitive-selection basis. The advertising for faculty and graduate students consisted primarily of the mailing of 8,000 52-page SRP brochures to chairpersons of departments relevant to AFOSR research and to administrators of grants in accredited universities, colleges, and technical institutions. Historically Black Colleges and Universities (HBCUs) and Minority Institutions (MIs) were included. Brochures also went to all participating USAF laboratories, the previous year's participants, and numerous individual requesters (over 1000 annually).

RDL placed advertisements in the following publications: *Black Issues in Higher Education, Winds of Change,* and *IEEE Spectrum.* Because no participants list either *Physics Today* or *Chemical & Engineering News* as being their source of learning about the program for the past several years, advertisements in these magazines were dropped, and the funds were used to cover increases in brochure printing costs.

High school applicants can participate only in laboratories located no more than 20 miles from their residence. Tailored brochures on the HSAP were sent to the head counselors of 180 high schools in the vicinity of participating laboratories, with instructions for publicizing the program in their schools. High school students selected to serve at Wright Laboratory's Armament Directorate (Eglin Air Force Base, Florida) serve eleven weeks as opposed to the eight weeks normally worked by high school students at all other participating laboratories.

Each SFRP or GSRP applicant is given a first, second, and third choice of laboratory. High school students who have more than one laboratory or directorate near their homes are also given first, second, and third choices.

Laboratories make their selections and prioritize their nominees. AFOSR then determines the number to be funded at each laboratory and approves laboratories' selections.

Subsequently, laboratories use their own funds to sponsor additional candidates. Some selectees do not accept the appointment, so alternate candidates are chosen. This multi-step selection procedure results in some candidates being notified of their acceptance after scheduled deadlines. The total applicants and participants for 1996 are shown in this table.

1996 Applicants and Participants			
PARTICIPANT CATEGORY	TOTAL APPLICANTS	SELECTEES	DECLINING SELECTEES
SFRP	572	188	39
(HBCU/MI)	(119)	(27)	(5)
GSRP	235	109	7
(HBCU/MI)	(18)	(7)	(1)
HSAP	474	138	8
TOTAL	1281	435	54

4. SITE VISITS

During June and July of 1996, representatives of both AFOSR/NI and RDL visited each participating laboratory to provide briefings, answer questions, and resolve problems for both laboratory personnel and participants. The objective was to ensure that the SRP would be as constructive as possible for all participants. Both SRP participants and RDL representatives found these visits beneficial. At many of the laboratories, this was the only opportunity for all participants to meet at one time to share their experiences and exchange ideas.

5. HISTORICALLY BLACK COLLEGES AND UNIVERSITIES AND MINORITY INSTITUTIONS (HBCU/MIs)

Before 1993, an RDL program representative visited from seven to ten different HBCU/Mis annually to promote interest in the SRP among the faculty and graduate students. These efforts were marginally effective, yielding a doubling of HBCI/MI applicants. In an effort to achieve AFOSR's goal of 10% of all applicants and selectees being HBCU/MI qualified, the RDL team decided to try other avenues of approach to increase the number of qualified applicants. Through the combined efforts of the AFOSR Program Office at Bolling AFB and RDL, two very active minority groups were found, HACU (Hispanic American Colleges and Universities) and AISES (American Indian Science and Engineering Society). RDL is in communication with representatives of each of these organizations on a monthly basis to keep up with the their activities and special events. Both organizations have widely-distributed magazines/quarterlies in which RDL placed ads.

Since 1994 the number of both SFRP and GSRP HBCU/MI applicants and participants has increased ten-fold, from about two dozen SFRP applicants and a half dozen selectees to over 100 applicants and two dozen selectees, and a half-dozen GSRP applicants and two or three selectees to 18 applicants and 7 or 8 selectees. Since 1993, the SFRP had a two-fold applicant increase and a two-fold selectee increase. Since 1993, the GSRP had a three-fold applicant increase and a three to four-fold increase in selectees.

In addition to RDL's special recruiting efforts, AFOSR attempts each year to obtain additional funding or use leftover funding from cancellations the past year to fund HBCU/MI associates. This year, 5 HBCU/MI SFRPs declined after they were selected (and there was no one qualified to replace them with). The following table records HBCU/MI participation in this program.

SRP HBCU/MI Participation, By Year

YEAR	SFRP		GSRP	
	Applicants	Participants	Applicants	Participants
1985	76	23	15	11
1986	70	18	20	10
1987	82	32	32	10
1988	53	17	23	14
1989	39	15	13	4
1990	43	14	17	3
1991	42	13	8	5
1992	70	13	9	5
1993	60	13	6	2
1994	90	16	11	6
1995	90	21	20	8
1996	119	27	18	7

6. SRP FUNDING SOURCES

Funding sources for the 1996 SRP were the AFOSR-provided slots for the basic contract and laboratory funds. Funding sources by category for the 1996 SRP selected participants are shown here.

1996 SRP FUNDING CATEGORY	SFRP	GSRP	HSAP
AFOSR Basic Allocation Funds	141	85	123
USAF Laboratory Funds	37	19	15
HBCU/MI By AFOSR (Using Procured Addn'l Funds)	10	5	0
TOTAL	188	109	138

SFRP - 150 were selected, but nine canceled too late to be replaced.
GSRP - 90 were selected, but five canceled too late to be replaced (10 allocations for the ALCs were withheld by AFOSR.)
HSAP - 125 were selected, but two canceled too late to be replaced.

7. COMPENSATION FOR PARTICIPANTS

Compensation for SRP participants, per five-day work week, is shown in this table.

1996 SRP Associate Compensation

PARTICIPANT CATEGORY	1991	1992	1993	1994	1995	1996
Faculty Members	$690	$718	$740	$740	$740	$770
Graduate Student (Master's Degree)	$425	$442	$455	$455	$455	$470
Graduate Student (Bachelor's Degree)	$365	$380	$391	$391	$391	$400
High School Student (First Year)	$200	$200	$200	$200	$200	$200
High School Student (Subsequent Years)	$240	$240	$240	$240	$240	$240

The program also offered associates whose homes were more than 50 miles from the laboratory an expense allowance (seven days per week) of $50/day for faculty and $40/day for graduate students. Transportation to the laboratory at the beginning of their tour and back to their home destinations at the end was also reimbursed for these participants. Of the combined SFRP and

GSRP associates, 65 % (194 out of 297) claimed travel reimbursements at an average round-trip cost of $780.

Faculty members were encouraged to visit their laboratories before their summer tour began. All costs of these orientation visits were reimbursed. Forty-five percent (85 out of 188) of faculty associates took orientation trips at an average cost of $444. By contrast, in 1993, 58 % of SFRP associates took orientation visits at an average cost of $685; that was the highest percentage of associates opting to take an orientation trip since RDL has administered the SRP, and the highest average cost of an orientation trip. These 1993 numbers are included to show the fluctuation which can occur in these numbers for planning purposes.

Program participants submitted biweekly vouchers countersigned by their laboratory research focal point, and RDL issued paychecks so as to arrive in associates' hands two weeks later.

In 1996, RDL implemented direct deposit as a payment option for SFRP and GSRP associates. There were some growing pains. Of the 128 associates who opted for direct deposit, 17 did not check to ensure that their financial institutions could support direct deposit (and they couldn't), and eight associates never did provide RDL with their banks' ABA number (direct deposit bank routing number), so only 103 associates actually participated in the direct deposit program. The remaining associates received their stipend and expense payments via checks sent in the US mail.

HSAP program participants were considered actual RDL employees, and their respective state and federal income tax and Social Security were withheld from their paychecks. By the nature of their independent research, SFRP and GSRP program participants were considered to be consultants or independent contractors. As such, SFRP and GSRP associates were responsible for their own income taxes, Social Security, and insurance.

8. CONTENTS OF THE 1996 REPORT

The complete set of reports for the 1996 SRP includes this program management report (Volume 1) augmented by fifteen volumes of final research reports by the 1996 associates, as indicated below:

1996 SRP Final Report Volume Assignments

LABORATORY	SFRP	GSRP	HSAP
Armstrong	2	7	12
Phillips	3	8	13
Rome	4	9	14
Wright	5A, 5B	10	15
AEDC, ALCs, WHMC	6	11	16

APPENDIX A – PROGRAM STATISTICAL SUMMARY

A. Colleges/Universities Represented

Selected SFRP associates represented 169 different colleges, universities, and institutions, GSRP associates represented 95 different colleges, universities, and institutions.

B. States Represented

SFRP -Applicants came from 47 states plus Washington D.C. and Puerto Rico. Selectees represent 44 states plus Puerto Rico.

GSRP - Applicants came from 44 states and Puerto Rico. Selectees represent 32 states.

HSAP - Applicants came from thirteen states. Selectees represent nine states.

Total Number of Participants	
SFRP	188
GSRP	109
HSAP	138
TOTAL	435

Degrees Represented			
	SFRP	GSRP	TOTAL
Doctoral	184	1	185
Master's	4	48	52
Bachelor's	0	60	60
TOTAL	188	109	297

SFRP Academic Titles	
Assistant Professor	79
Associate Professor	59
Professor	42
Instructor	3
Chairman	0
Visiting Professor	1
Visiting Assoc. Prof.	0
Research Associate	4
TOTAL	188

Source of Learning About the SRP		
Category	Applicants	Selectees
Applied/participated in prior years	28%	34%
Colleague familiar with SRP	19%	16%
Brochure mailed to institution	23%	17%
Contact with Air Force laboratory	17%	23%
IEEE Spectrum	2%	1%
BIIHE	1%	1%
Other source	10%	8%
TOTAL	100%	100%

APPENDIX B – SRP EVALUATION RESPONSES

1. OVERVIEW

Evaluations were completed and returned to RDL by four groups at the completion of the SRP. The number of respondents in each group is shown below.

Table B-1. Total SRP Evaluations Received

Evaluation Group	Responses
SFRP & GSRPs	275
HSAPs	113
USAF Laboratory Focal Points	84
USAF Laboratory HSAP Mentors	6

All groups indicate unanimous enthusiasm for the SRP experience.

The summarized recommendations for program improvement from both associates and laboratory personnel are listed below:

A. Better preparation on the labs' part prior to associates' arrival (i.e., office space, computer assets, clearly defined scope of work).

B. Faculty Associates suggest higher stipends for SFRP associates.

C. Both HSAP Air Force laboratory mentors and associates would like the summer tour extended from the current 8 weeks to either 10 or 11 weeks; the groups state it takes 4-6 weeks just to get high school students up-to-speed on what's going on at laboratory. (Note: this same argument was used to raise the faculty and graduate student participation time a few years ago.)

2. 1996 USAF LABORATORY FOCAL POINT (LFP) EVALUATION RESPONSES

The summarized results listed below are from the 84 LFP evaluations received.

1. LFP evaluations received and associate preferences:

Table B-2. Air Force LFP Evaluation Responses (By Type)

		How Many Associates Would You Prefer To Get ?								(% Response)			
		SFRP				GSRP (w/Univ Professor)				GSRP (w/o Univ Professor)			
Lab	**Evals Recv'd**	**0**	**1**	**2**	**3+**	**0**	**1**	**2**	**3+**	**0**	**1**	**2**	**3+**
AEDC	0	-	-	-	-	-	-	-	-	-	-	-	-
WHMC	0	-	-	-	-	-	-	-	-	-	-	-	-
AL	7	28	28	28	14	54	14	28	0	86	0	14	0
FJSRL	1	0	100	0	0	100	0	0	0	0	100	0	0
PL	25	40	40	16	4	88	12	0	0	84	12	4	0
RL	5	60	40	0	0	80	10	0	0	100	0	0	0
WL	46	30	43	20	6	78	17	4	0	93	4	2	0
Total	**84**	**32%**	**50%**	**13%**	**5%**	**80%**	**11%**	**6%**	**0%**	**73%**	**23%**	**4%**	**0%**

LFP Evaluation Summary. The summarized responses, by laboratory, are listed on the following page. LFPs were asked to rate the following questions on a scale from 1 (below average) to 5 (above average).

2. LFPs involved in SRP associate application evaluation process:
    a. Time available for evaluation of applications:
    b. Adequacy of applications for selection process:
3. Value of orientation trips:
4. Length of research tour:
5 a. Benefits of associate's work to laboratory:
    b. Benefits of associate's work to Air Force:
6. a. Enhancement of research qualifications for LFP and staff:
    b. Enhancement of research qualifications for SFRP associate:
    c. Enhancement of research qualifications for GSRP associate:
7. a. Enhancement of knowledge for LFP and staff:
    b. Enhancement of knowledge for SFRP associate:
    c. Enhancement of knowledge for GSRP associate:
8. Value of Air Force and university links:
9. Potential for future collaboration:
10. a. Your working relationship with SFRP:
    b. Your working relationship with GSRP:
11. Expenditure of your time worthwhile:

(Continued on next page)

12. Quality of program literature for associate:
13. a. Quality of RDL's communications with you:
    b. Quality of RDL's communications with associates:
14. Overall assessment of SRP:

Table B-3. Laboratory Focal Point Reponses to above questions

	AEDC	*AL*	*FJSRL*	*PL*	*RL*	*WHMC*	*WL*
# Evals Recv'd	0	7	1	14	5	0	46
Question #							
2	-	86 %	0 %	88 %	80 %	-	85 %
2a	-	4.3	n/a	3.8	4.0	-	3.6
2b	-	4.0	n/a	3.9	4.5	-	4.1
3	-	4.5	n/a	4.3	4.3	-	3.7
4	-	4.1	4.0	4.1	4.2	-	3.9
5a	-	4.3	5.0	4.3	4.6	-	4.4
5b	-	4.5	n/a	4.2	4.6	-	4.3
6a	-	4.5	5.0	4.0	4.4	-	4.3
6b	-	4.3	n/a	4.1	5.0	-	4.4
6c	-	3.7	5.0	3.5	5.0	-	4.3
7a	-	4.7	5.0	4.0	4.4	-	4.3
7b	-	4.3	n/a	4.2	5.0	-	4.4
7c	-	4.0	5.0	3.9	5.0	-	4.3
8	-	4.6	4.0	4.5	4.6	-	4.3
9	-	4.9	5.0	4.4	4.8	-	4.2
10a	-	5.0	n/a	4.6	4.6	-	4.6
10b	-	4.7	5.0	3.9	5.0	-	4.4
11	-	4.6	5.0	4.4	4.8	-	4.4
12	-	4.0	4.0	4.0	4.2	-	3.8
13a	-	3.2	4.0	3.5	3.8	-	3.4
13b	-	3.4	4.0	3.6	4.5	-	3.6
14	-	4.4	5.0	4.4	4.8	-	4.4

3. 1996 SFRP & GSRP EVALUATION RESPONSES

The summarized results listed below are from the 257 SFRP/GSRP evaluations received.

Associates were asked to rate the following questions on a scale from 1 (below average) to 5 (above average) - by Air Force base results and over-all results of the 1996 evaluations are listed after the questions.

1. The match between the laboratories research and your field:
2. Your working relationship with your LFP:
3. Enhancement of your academic qualifications:
4. Enhancement of your research qualifications:
5. Lab readiness for you: LFP, task, plan:
6. Lab readiness for you: equipment, supplies, facilities:
7. Lab resources:
8. Lab research and administrative support:
9. Adequacy of brochure and associate handbook:
10. RDL communications with you:
11. Overall payment procedures:
12. Overall assessment of the SRP:
13.
    a. Would you apply again?
    b. Will you continue this or related research?
14. Was length of your tour satisfactory?
15. Percentage of associates who experienced difficulties in finding housing:
16. Where did you stay during your SRP tour?
    a. At Home:
    b. With Friend:
    c. On Local Economy:
    d. Base Quarters:
17. Value of orientation visit:
    a. Essential:
    b. Convenient:
    c. Not Worth Cost:
    d. Not Used:

SFRP and GSRP associate's responses are listed in tabular format on the following page.

Table B-4. 1996 SFRP & GSRP Associate Responses to SRP Evaluation

	Arnold	Brooks	Edwards	Eglin	Griffis	Hanscom	Kelly	Kirtland	Lackland	Robins	Tyndall	WPAFB	average
# res	6	48	6	14	31	19	3	32	1	2	10	85	257
1	4.8	4.4	4.6	4.7	4.4	4.9	4.6	4.6	5.0	5.0	4.0	4.7	4.6
2	5.0	4.6	4.1	4.9	4.7	4.7	5.0	4.7	5.0	5.0	4.6	4.8	4.7
3	4.5	4.4	4.0	4.6	4.3	4.2	4.3	4.4	5.0	5.0	4.5	4.3	4.4
4	4.3	4.5	3.8	4.6	4.4	4.4	4.3	4.6	5.0	4.0	4.4	4.5	4.5
5	4.5	4.3	3.3	4.8	4.4	4.5	4.3	4.2	5.0	5.0	3.9	4.4	4.4
6	4.3	4.3	3.7	4.7	4.4	4.5	4.0	3.8	5.0	5.0	3.8	4.2	4.2
7	4.5	4.4	4.2	4.8	4.5	4.3	4.3	4.1	5.0	5.0	4.3	4.3	4.4
8	4.5	4.6	3.0	4.9	4.4	4.3	4.3	4.5	5.0	5.0	4.7	4.5	4.5
9	4.7	4.5	4.7	4.5	4.3	4.5	4.7	4.3	5.0	5.0	4.1	4.5	4.5
10	4.2	4.4	4.7	4.4	4.1	4.1	4.0	4.2	5.0	4.5	3.6	4.4	4.3
11	3.8	4.1	4.5	4.0	3.9	4.1	4.0	4.0	3.0	4.0	3.7	4.0	4.0
12	5.7	4.7	4.3	4.9	4.5	4.9	4.7	4.6	5.0	4.5	4.6	4.5	4.6
Numbers below are percentages													
13a	83	90	83	93	87	75	100	81	100	100	100	86	87
13b	100	89	83	100	94	98	100	94	100	100	100	94	93
14	83	96	100	90	87	80	100	92	100	100	70	84	88
15	17	6	0	33	20	76	33	25	0	100	20	8	39
16a	-	26	17	9	38	23	33	4	-	-	-	30	
16b	100	33	-	40	-	8	-	-	-	-	36	2	
16c	-	41	83	40	62	69	67	96	100	100	64	68	
16d	-	-	-	-	-	-	-	-	-	-	-	0	
17a	-	33	100	17	50	14	67	39	-	50	40	31	35
17b	-	21	-	17	10	14	-	24	-	50	20	16	16
17c	-	-	-	-	10	7	-	-	-	-	-	2	3
17d	100	46	-	66	30	69	33	37	100	-	40	51	46

4. 1996 USAF LABORATORY HSAP MENTOR EVALUATION RESPONSES

Not enough evaluations received (5 total) from Mentors to do useful summary.

5. 1996 HSAP EVALUATION RESPONSES

The summarized results listed below are from the 113 HSAP evaluations received.

HSAP apprentices were asked to rate the following questions on a scale from
1 (below average) to 5 (above average)

1. Your influence on selection of topic/type of work.
2. Working relationship with mentor, other lab scientists.
3. Enhancement of your academic qualifications.
4. Technically challenging work.
5. Lab readiness for you: mentor, task, work plan, equipment.
6. Influence on your career.
7. Increased interest in math/science.
8. Lab research & administrative support.
9. Adequacy of RDL's Apprentice Handbook and administrative materials.
10. Responsiveness of RDL communications.
11. Overall payment procedures.
12. Overall assessment of SRP value to you.
13. Would you apply again next year? Yes (92 %)
14. Will you pursue future studies related to this research? Yes (68 %)
15. Was Tour length satisfactory? Yes (82 %)

	Arnold	Brooks	Edwards	Eglin	Griffiss	Hanscom	Kirtland	Tyndall	WPAFB	Totals
# resp	5	19	7	15	13	2	7	5	40	113
1	2.8	3.3	3.4	3.5	3.4	4.0	3.2	3.6	3.6	3.4
2	4.4	4.6	4.5	4.8	4.6	4.0	4.4	4.0	4.6	4.6
3	4.0	4.2	4.1	4.3	4.5	5.0	4.3	4.6	4.4	4.4
4	3.6	3.9	4.0	4.5	4.2	5.0	4.6	3.8	4.3	4.2
5	4.4	4.1	3.7	4.5	4.1	3.0	3.9	3.6	3.9	4.0
6	3.2	3.6	3.6	4.1	3.8	5.0	3.3	3.8	3.6	3.7
7	2.8	4.1	4.0	3.9	3.9	5.0	3.6	4.0	4.0	3.9
8	3.8	4.1	4.0	4.3	4.0	4.0	4.3	3.8	4.3	4.2
9	4.4	3.6	4.1	4.1	3.5	4.0	3.9	4.0	3.7	3.8
10	4.0	3.8	4.1	3.7	4.1	4.0	3.9	2.4	3.8	3.8
11	4.2	4.2	3.7	3.9	3.8	3.0	3.7	2.6	3.7	3.8
12	4.0	4.5	4.9	4.6	4.6	5.0	4.6	4.2	4.3	4.5
Numbers below are percentages										
13	60%	95%	100%	100%	85%	100%	100%	100%	90%	92%
14	20%	80%	71%	80%	54%	100%	71%	80%	65%	68%
15	100%	70%	71%	100%	100%	50%	86%	60%	80%	82%

TURBULENCE STATISTICS AND ENERGY BUDGET OF A TURBULENT SHEAR LAYER

Saad A. Ahmed
Associate Professor
Mechanical Engineering Department
KFUPM, Dhahran 31261, KSA

Final report for
Summer Faculty Research Program
Arnold Engineering and Development Center

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, DC

and

Arnold Engineering and Development Center

September, 1996

TURBULENCE STATISTICS AND ENERGY BUDGET OF A TURBULENT SHEAR LAYER

Saad A. Ahmed
Associate Professor
Mechanical Engineering Department
KFUPM, Dhahran 31261, KSA

ABSTRACT

A two-component laser Doppler velocimeter "LDV" was employed to measure the flowfield of a generic tactical missile configuration in simulated powered flight. Mean velocities, normal and shear stresses, triple products, and higher order statistics measured by the LDV system working, in the off-axis forward scatter mode were analyzed. A balance of the turbulence energy equation has been performed in order to get a detailed insight into the turbulent shear layer behavior. The turbulent kinetic energy terms: production, diffusion, and convection terms were computed directly from the experimental data using central difference, while the viscous dissipation term was obtained from the balance of the kinetic energy equation. The analysis of the data was successful in identifying the various areas of interest in the flowfield where different turbulent transport mechanisms dominate. In addition, the data from this study will be available for upgrading advanced numerical codes.

TURBULENCE STATISTICS AND ENERGY BUDGET OF A TURBULENT SHEAR LAYER

Saad A. Ahmed

1. INTRODUCTION

The flowfield in the vicinity of the base of a missile has been a subject of interest in the past few years. The complexity and significance of the flowfield in this region make missile base-flow predictions difficult and essential. The adverse effects of base drag on tactical missile cruise performance is well known. Plume induced separation of missile afterbody flow would cause severe degradation of the stability characteristics of a vehicle. Also, the effectiveness of control surfaces could be reduced and catastrophic failure might result. Although substantial literature exists (see Dutton et al.[4], 1995, for more details), it is neither complete nor sufficient to match the progress that has been made in computational fluid dynamics capabilities and computers in recent years. This lack of detailed experimental data has been a major restriction to the development and validation of numerical prediction codes that model the flowfield of missile base flow. A high quality and detailed experimental data utilizing non-intrusive diagnostic methods are needed for development and validation of both new and improved / or modified current computational methods (see Kennedy et al.[7], 1992).

2. EXPERIMENTAL SET-UP

The experiments were conducted in AEDC Aerodynamic Wind Tunnel (1T) which is a continuous flow, and an open circuit wind tunnel. It can operate at a constant stagnation pressure of 1.35 atm over a Mach number range from 0.2 to 1.5. The stagnation temperature could be varied from 300 K to 323 K above ambient temperature. The side walls were fitted with large flat and polished optical quality glass windows. The floor and ceiling portions of the test sections were perforated in order to reduce wall interference at the test section. Compressed air for the model jet plume was obtained from the VKF which could provide dry air, continuously at 200 atm. The air was supplied to the base nozzle through passages machined in the mounting strut. The tunnel was operated at a free stream Mach No. of 1.4 with a stagnation pressure of 137.5 kPa and a stagnation temperature of 355.6 K, or 640 °R. The propulsive jet was operated at a pressure of 2.41×10^6 Pa (or 350 psia, NPR = 330, total nozzle exit pressure divided by static free stream pressure) and a Mach Number of 2.7. The high operating pressure resulted in highly under expanded plume and a region of plume induced flow separation could develop around the base.

The model was a 57 cm long, 1.27 cm diameter (L/D = 45), circular ogive cylinder mounted on struts from the side walls (see Fig. 1). The details of the model and the tunnel are reported earlier by Walker and Helstley[10], 1993.

2.1 LDV System

The laser Doppler velocimeter (LDV) system used was a two-component/ two-color (green and blue) operating in the off-axis forward scatter mode. The beams were frequency shifted using Bragg cells to determine the direction of the flow easily. The laser source was a 5-Watt argon ion laser, water cooled type. Fringes were oriented at ± 45° relative to the tunnel axis to minimize fringe bias effects and to improve the resolution of the system (see Ahmed et al., 1992). The sampling volume was approximately 0.3 mm in diameter and 2.0 mm long. Probe volume location was expressed in x, y, and z coordinates relative to an origin at the model nozzle exit center. Seed particles were injected both into the free stream flow and into the high pressure air jet utilizing liquid atomizers or fluidized bed. The details of the system and its operational procedure were reported by Walker and Helstley[10], 1993.

3 METHODOLOGY

A two-component LDV system was used to measure accurately $\overline{U}$, $\overline{V}$, $\overline{u'^2}$, $\overline{v'^2}$, $\overline{u'v'}$, $\overline{u'^3}$, $\overline{u'^2v'}$, $\overline{u'v'^2}$, and $\overline{v'^3}$ in the flowfield downstream of the base of a generic tactical missile configuration in simulated powered flight. The data are unique in their details, high quality and accuracy, no other similar data are available. All the measured values of velocity components, turbulence statistics, and turbulent kinetic energy "K" budgets are examined but not reported here due to its classification.

In principle, the missile base flow has free turbulent shear layers in addition to different types of shocks; such as barrel shocks, oblique shocks and Mach disks. Its characteristics will be dominated by the characteristics of these shear layers. To avoid the problem of classification, the analysis and the presentation will be focused only on turbulent shear layers rather than missile base flow (see Fig. 2). This shear layer could be generated by simple devices, such as a swirler or a simple jet. The analysis and the data presented here are for a shear layer generated by a free vortex swirler whose data is available in the open literature and reported earlier by Ahmed and Nejad[1], 1992 (the measurements were taken by the author at Wright Laboratory "WL" while the author was employed by WL). The analysis of the swirler data (i.e., WL data) will also be applicable to missile base flow and many of the features noticed in the swirling shear layer will be applicable to missile base flow shear layers (i.e., AEDC data). Swirler data are also detailed data and the objectives of WL data is similar to AEDC data (i.e., detailed experimental data are needed to help in the understanding of complex flows and to check the validity of advanced numerical codes).

3.1 WL Experimental Set-up

3.1.1 Combustor Model. A modular can-type research combustor was designed and constructed (see Ahmed and Nejad[1], 1992). Special care was taken to ensure that the fabricated model satisfied the axisymmetric nature of the flowfield. In addition, a superb provision was provided on the model to facilitate the use of a two-component fiber-optic laser Doppler velocimeter (LDV). The model consisted of three main sections (see Figure 3): inlet section, combustion chamber, and window assemblies.

The inlet section is made of a settling chamber, a plexiglas inlet pipe, and a cylindrical teflon swirler housing. This swirler housing has the ability of being positioned relative to the measurement station in the combustion chamber. This is made possible by placing the entire inlet assembly on a traversing mechanism controlled by a stepping motor. An average inlet velocity of 16.0 ± 0.4 *m/s* (large enough to ensure turbulent flow in the combustor, at a Reynolds number - based on the outer diameter of the swirler- of 1.5×10^5) was maintained throughout all experiments.

The combustion chamber is made up of a plexiglas tube which terminates into an exhauster of a larger diameter. The measurement station was designed to accommodate two different window assemblies to enable the measurement of the three component velocity vectors with a two component LDV system, see Figure 3.

The window assembly is made up of a flat optical quality quartz window which is installed in a frame made from plexiglas. This frame, fixed to two brackets, is attached to the combustor section as illustrated in Figure 3a.

The swirler is a free vortex swirler and is similar to the one used by Buckley et al.[2], 1983. The current facility employed a combustor of 150 mm ID. Swirler dimensions are 19 mm ID (central hub) and 101.6 mm OD. The swirler has 12 curved vanes designed to give a mean airflow turning angle of Θ^{O} (see Buckley et al.[2], 1983).

3.1.2 Laser Velocimeter. The velocity measurements for this study were done using a TSI Inc. 9100-7 four beam, two color, back scatter LDV system. The LDV system had two TSI 9180-3A frequency shifters to provide directional sensitivity. The optical head was mounted on a three axis traversing table with a resolution of 0.0025 *mm*. The features of the optical head include:

1. Setting the fringe inclinations at 45.7° and 135.17° to the combustor centerline.
2. The approximate measurement volume dimensions were 390 μm in length and 60 μm diameter based on the $1/e^2$ intensity points.

The chemical seeder developed by Craig et al.[3], 1984 was used for the present study. This chemical seeder produced micron size Titanium dioxide particles. These particles were injected into the upstream settling chamber in order to ensure that the flow was uniformly seeded.

3.2 Results

The present study is a part of an ongoing effort aimed at better understanding of the flow features in a confined isothermal swirling flowfield. To ensure a complete mapping of the flowfield and to confirm symmetry of the flow, detailed measurements were made in the horizontal and vertical planes, see Figures 4 to 10. All measurements are made non-dimensional (e.g., the inlet upstream centerline mean velocity " $\overline{U_o}$ ", and the step height "H" are used for this purpose). Since the flow is nearly axisymmetric, the results presented here are for the lateral half section of the combustor.

3.3 Discussions

3.3.1 Stream Function and Turbulence Intensity Contours.

Figure 4 shows the contour plots of normalized stream function ψ . Zero values of ψ depicts the corner recirculation zone "CRZ" boundaries while $\psi = 1$ is the boundary of another recirculation region, known as the central toroidal recirculation zone "CTRZ". As could be easily shown from the contours of the normalized stream function, the flow reattached earlier in the free vortex "FV" case than a simple dump "SD" case. Specifically, the points of reattachments for the FV case is $x = 4.3H$ compared to approximately about 8H for the no swirl case. The swirler is responsible for shortening the CRZ. Furthermore, the introduced swirler caused the flow to recover (i.e., uniform flow) shortly after the reattachment. In summary, the free vortex flow has two shear layers (inner layer around the CTRZ and outer layer around the CRZ).

The contour plots of turbulent velocity profiles are shown in Figures 5a through 5c. For all three components of turbulent velocities, two peaks characterized the swirling flow around the inner and outer shear layers (i.e., the larger is seen in the boundaries of CTRZ, and the other occurred in the shear layer of CRZ). Turbulence activities are reduced downstream of the reattachment point. Also, downstream of the reattachment point, the peak value of the axial turbulent intensity is observed to move towards the combustor centerline as it decays in strength and grows in size, indicating a progressive development of the outer shear layer. Turbulence activity is seen to be more concentrated in the central shear layer and the centers of the maximum contour values, for the two regions where the peak values occurred, are observed to be at $x / H = 3.0$.

The features of the radial and tangential turbulent velocity contours are similar to those described above for the axial turbulent velocity. However, the contours of the axial turbulent velocity indicate a limited flow recovery in the axial direction downstream of the reattachment point while flow did not recover completely in the radial and tangential directions. This is obvious from the radial and tangential turbulent velocity contours.

3.3.2 Triple Velocity Product Correlations. The spatial derivatives of the triple velocity correlations contribute to the diffusion terms of the transport equation of turbulent kinetic energy "K". Some of these velocity products are shown in Figures 6a through 6g, others were not measured because the LDV system was not a 3D system. Kasagi and Matsunaga[6] (1995) reported that the accurate prediction of the triple velocity products is one of the striking issues in the second-order closure modeling. For virtually all the triple products shown, the swirling flow demonstrate higher magnitudes when compared to the no-swirl flow case. The triple products, as could be easily discerned from the referred Figures, have certain common characteristics. For example, the triple products are antisymmetric about the centerline of the shear layer. They peak to different values on both sides of the shear layer and their values diminishes at the edges of the shear layer. These peaks are more prominent before the reattachment point (near field) while they decrease considerably downstream of the reattachment point (far field). A similar behavior of the triple velocity product terms was observed by Kasagi and Matsunaga[6] (1995) for the backward step case.

3.3.3 Turbulent Kinetic Energy (TKE) Equation. The present experimental data were analyzed for the turbulent kinetic energy terms (i.e., production, diffusion, convection, and viscous dissipation). Since axisymmetry was assumed for the flow (this was seen as a valid assumption from the experimental results), a two-dimensional cylindrical coordinate system was used. The domain of interest was from the swirler inlet to the last measurement plane, that is, $x/H = 0.0 - 18.0$ and a grid of 25 x 12 (radial x axial) was used. The turbulence energy equation in cylindrical coordinates (for steady, axisymmetric, and constant density flow), employed in the determination of the turbulent kinetic energy terms in this analysis, could be written in cylindrical coordinates x, r, θ as (see Kasagi and Matsunaga[6], 1995):

$$
\begin{aligned}
\overline{U}\frac{\partial k}{\partial x}+\overline{V}\frac{\partial k}{\partial r} = -\Bigg[& \overline{v'^2}\frac{\partial \overline{V}}{\partial r}+\overline{u'v'}\frac{\partial \overline{U}}{\partial r}+\overline{u'v'}\frac{\partial \overline{V}}{\partial x}+\overline{u'^2}\frac{\partial \overline{U}}{\partial x}+\overline{w'^2}\frac{\overline{V}}{r}+\overline{u'w'}\frac{\partial \overline{W}}{\partial x}+\overline{v'w'}\frac{\partial \overline{W}}{\partial r}+\overline{v'w'}\frac{\overline{W}}{r}\Bigg]- \\
& \frac{\mu}{\overline{\rho}}\left[\overline{\left(\frac{\partial v'}{\partial r}\right)^2}+\overline{\left(\frac{\partial w'}{\partial r}\right)^2}+\overline{\left(\frac{\partial u'}{\partial r}\right)^2}+\overline{\left(\frac{\partial v'}{\partial x}\right)^2}+\overline{\left(\frac{\partial w'}{\partial x}\right)^2}+\overline{\left(\frac{\partial u'}{\partial x}\right)^2}\right]+ \\
& \frac{1}{\overline{\rho}}\frac{\partial}{\partial x}\left[\mu\frac{\partial k}{\partial r}-\overline{\rho}\frac{\overline{u'v'^2}}{2}-\overline{\rho}\frac{\overline{u'w'^2}}{2}-\overline{\rho}\frac{\overline{u'^3}}{2}-\overline{u'p'}\right]+ \\
& \frac{1}{\overline{\rho}r}\frac{\partial}{\partial r}\left[\mu r\frac{\partial k}{\partial r}-\overline{\rho}r\frac{\overline{v'^3}}{2}-\overline{\rho}r\frac{\overline{v'w'^2}}{2}-\overline{\rho}r\frac{\overline{u'^2v'}}{2}-r\overline{v'p'}\right]+\left[\frac{\mu}{\overline{\rho}}\frac{\overline{v'^2}}{r^2}+\frac{\mu}{\overline{\rho}}\frac{\overline{w'^2}}{r^2}\right]
\end{aligned}
\tag{1}
$$

This form of the equation is essentially the same as the one given by Gould et al.[5], 1990. The TKE equation is composed of the following terms:

1. axial and radial convection of turbulent kinetic energy:

$$\overline{U}\frac{\partial k}{\partial x}+\overline{V}\frac{\partial k}{\partial r} \tag{2}$$

2. production of turbulent kinetic energy:

$$\left[\overline{v'^2}\frac{\partial \overline{V}}{\partial r}+\overline{u'v'}\frac{\partial \overline{U}}{\partial r}+\overline{u'v'}\frac{\partial \overline{V}}{\partial x}+\overline{u'^2}\frac{\partial \overline{U}}{\partial x}+\overline{w'^2}\frac{\overline{V}}{r}+\overline{u'w'}\frac{\partial \overline{W}}{\partial x}+\overline{v'w'}\frac{\partial \overline{W}}{\partial r}+\overline{v'w'}\frac{\overline{W}}{r}\right] \tag{3}$$

3. viscous dissipation of turbulent kinetic energy:

$$\frac{\mu}{\bar\rho}\left[\overline{\left(\frac{\partial v'}{\partial r}\right)^2}+\overline{\left(\frac{\partial w'}{\partial r}\right)^2}+\overline{\left(\frac{\partial u'}{\partial r}\right)^2}+\overline{\left(\frac{\partial v'}{\partial x}\right)^2}+\overline{\left(\frac{\partial w'}{\partial x}\right)^2}+\overline{\left(\frac{\partial u'}{\partial x}\right)^2}\right] \tag{4}$$

4. turbulent diffusion of kinetic energy:

$$\frac{1}{\bar\rho}\frac{\partial}{\partial x}\left[-\bar\rho\frac{\overline{u'v'^2}}{2}-\bar\rho\frac{\overline{u'w'^2}}{2}-\bar\rho\frac{\overline{u'^3}}{2}\right]+\frac{1}{\bar\rho r}\frac{\partial}{\partial r}\left[-\bar\rho r\frac{\overline{v'^3}}{2}-\bar\rho r\frac{\overline{v'w'^2}}{2}-\bar\rho r\frac{\overline{u'^2v'}}{2}\right] \tag{5}$$

5. viscous diffusion of turbulent kinetic energy:

$$\frac{1}{\bar\rho}\frac{\partial}{\partial x}\left[\mu\frac{\partial k}{\partial r}\right]+\frac{1}{\bar\rho r}\frac{\partial}{\partial r}\left[\mu r\frac{\partial k}{\partial r}\right]+\left[\frac{\mu}{\bar\rho}\frac{\overline{v'^2}}{r^2}+\frac{\mu}{\bar\rho}\frac{\overline{w'^2}}{r^2}\right] \tag{6}$$

6. pressure diffusion of turbulent kinetic energy:

$$\frac{1}{\bar\rho}\frac{\partial}{\partial x}\left[-\overline{u'p'}\right]+\frac{1}{\bar\rho r}\frac{\partial}{\partial r}\left[-r\overline{v'p'}\right] \tag{7}$$

All the terms of the turbulent kinetic energy were calculated directly from the experimental data using central difference, except the pressure diffusion terms and the viscous dissipation terms. The pressure diffusion terms could be neglected without incurring any significant error as their magnitudes were considerably small when compared with other terms (see Kasagi and Matsunaga[6], 1993). Also, some terms of the equation were not measured and assumptions were made to account for these terms (i.e., $\overline{v'w'^2} = \overline{v'^3}$, $\overline{v'w'} = \overline{v'^2}$). The total viscous dissipation term could be obtained by algebraic balance of the turbulent kinetic energy equation.

3.3.4 Turbulent Kinetic Energy (TKE) Terms. Figures 7 through 10 show the profiles of the TKE terms downstream of the step. Figure 7, shows profiles of the total non-dimensional production term "PT". The plots show that the FV profiles have two peaks. Generally, the values of the peaks decrease downstream of the step, dying out completely downstream of $x = 6H$. The major peak is closer to the center of the combustor while the minor one occurs closer to the step plane. This indicates that the swirler enhances the production of TKE and, in addition, enables thorough mixing of the incoming flow. The production term makes positive contributions to the TKE and its peak (i.e., maximum production rate) decreases with increasing streamwise distance.

The production term $-\overline{u'v'}\frac{\partial\overline{U}}{\partial r}$ was found to dominate the total production term in a separate comparison not shown here. This is also in agreement with the previously reported data of Kasagi and Matsunaga [6] (1995), Ninomiya and Kasagi[8] (1993), and Sami[9] (1967), and a host of others. On the other hand, the $-\overline{u'v'}\frac{\partial\overline{V}}{\partial x}$ production term comparatively contributed insignificantly to the total production of TKE while the $-\overline{u'^2}\frac{\partial\overline{U}}{\partial x}$ and $-\overline{v'^2}\frac{\partial\overline{V}}{\partial r}$ terms did make opposite contributions, of the same order of magnitude, which nullify the effect of their sum on the total production of TKE.

The total non-dimensional convection term "CT" profiles are shown in Figure 8. In a similar pattern with PT, the CT is characterized with a relatively small peak observed around the step plane and this peak diminishes downstream of the reattachment point. Hence, energy is convected, and as such carried, from around the center of the shear layer toward the shear layer edges. The introduction of the swirler created another peak, located between the hub and the step, up to an axial distance of $x = 2H$. This additional peak merges with the other peak and form a larger size peak which covers a big portion of the combustor. This is an indication of increase in energy transportation and hence distribution. In addition, turbulent energy convection before the reattachment point is opposite in sign (i.e., loss) to that after the reattachment point, from $x/H = 5$ onward. However, the magnitude of the peak after flow reattachment is less compared to its size before reattachment.

The turbulent total non-dimensional diffusion term "DT", whose profiles are shown in Figure 9, redistributes the energy in the vicinity of the shear layer- from the centers to the outer parts of the shear layer. These activities are observed in the vicinity of CRZ boundaries and energy is also seen to be redistributed close to the CTRZ boundaries. This occurrence is primarily due to the swirl introduction and this shows the importance of stirring the flow at the inlet of the combustor. The intensity of energy diffusion, at the step plane and at the boundary of the CTRZ, decreases with increasing streamwise distance.

The total non-dimensional viscous dissipation term "VDT", obtained by balancing the TKE transport equation, is seen to make a negative contribution to the TKE (i.e., loss) and portrays a similar but opposite profile style as the production term, although slightly smaller in magnitude, see Figure 10. The maximum rate of energy dissipation was observed in the vicinity of the CTRZ and it vanishes downstream of the reattachment point.

4. CONCLUSIONS

A two component LDV system was utilized to measure the flowfield turbulence characteristics of a free vortex swirler. As a result, a database of the turbulence statistics was established over a fairly wide region of 0.38H to 24H downstream the step including the regions of major interest (i.e., corner recirculation and central toroidal recirculation zones). The Reynolds normal and shear stresses peak in the shear layers upstream of the reattachment points. The budgets of turbulent kinetic energy K are calculated from the distributions of velocity correlations using central differences. The mechanism of the turbulence energy exchange is investigated during the course of this study.

In summary, for all activities of energy production, convection, diffusion and dissipation, the swirled flow demonstrated higher values compared to the no swirl case. In particular, the swirled flow was shown to enhance the production of TKE as larger values of the production terms were recorded close to the centerline of the combustor due to the existence of CTRZ. The FV was shown to portray an additional activity close to the centerline of the combustor. These additional activities serve very important purposes. For instance, a central toroidal recirculation zone, generated close to the center of the combustor, enhances combustor performance by serving as another flame anchoring mechanism. The swirler also reduced the flow reattachment point considerably. As a result, ensuring that flow recovers shortly after the reattachment point. Furthermore, energy was more distributed on either side of the shear layer in the region close to the centerline of the combustor and flow separation areas. Thus, it is now abundantly clear that a swirler in the combustor will have important effects on the flow characteristics. Also, these detailed information should be of value for further development of second order closure models.

5. ACKNOWLEDGMENTS

The author would like to acknowledge and thank Mr. F. Helstley for use of his data and for all the help he provided. Thanks are also due to Mr. W. D. Williams for his support in the laboratory. This investigation was performed at Arnold Engineering and Development Center under the Summer Faculty Research Program supported by AFOSR. The results reported here were for an experiment performed at the Wright Laboratory and the data are available in the open literature.

REFERENCES

1 Ahmed, S.A., and Nejad, A.S. "Velocity Measurements in a Research Combustor. Part I. Isothermal Swirling Flow," Experimental Thermal and Fluid Science Journal, Vol. 5, pp. 162-174, 1992.

2 Buckley, P.L., Craig, R.R., Davis, D.L. and Schwartzkopf, K.G., "The Design and Combustion Performance of Practical Swirlers for Integral Rocket / Ramjets," AIAA Journal, Vol. 21, No. 5, pp. 743-740, 1983.

3 Craig, R.R., Nejad, A.S., Hahn, E.Y., and Schwartzkopf, K.G., "A General Approach for Obtaining Unbiased LDV Data in Highly Turbulent Non-reacting and Reacting Flows," AIAA Paper No. 84-0366, 1984.

4 Dutton, J.C., Herrin, J.L., Molezzi, M.J., Mathur, T. and Smith, K.M. “Recent Progress on High-Speed Separated Base Flow,” AIAA 95-0472, 1995.

5 Gould, R.D., Stevenson, W.H. and Thompson, H.D., "Investigation of Turbulent Transport in an Axisymmetric Sudden Expansion," AIAA Journal, Vol. 28, pp. 276-283, 1990.

6 Kasagi, N. and Matsunaga, A. "Three-dimensional particle tracking velocimetry measurement of turbulence statistics and energy budget in a backward-facing step flow," Int. J. heat and Fluid Flow, Vol. 16, pp. 477-485, 1995.

7 Kennedy, K. D., Walker, B. J., Mikkelsen, C. D., and Heltsley, F. L., "Measurements and Predictions for Velocity and Turbulent Kinetic Energy in Air/Air Plumes," AIAA 92-2782, 1992.

8 Ninomiya, N and Nobuhide K., "Measurement of the Reynolds Stress Budgets in an Axisymmetric Free Jet with the Aid of Three Dimensional Particle Tracking Velocimetry," Ninth Symposium on Turbulent Shear Flows, Kyoto, Japan. August 16-18, 1993.

9 Sami S., "Jet Diffusion in the Region of Flow Establishment," Journal of Fluid Mech., 27, pp. 231-252, 1967.

10 Walker, B. J., and Heltsley, F. L., "Nonintrusive Diagnostics of Supersonic Missile Afterbody Flows." 15th ICIASF 93, St. Louis, France, 1993.

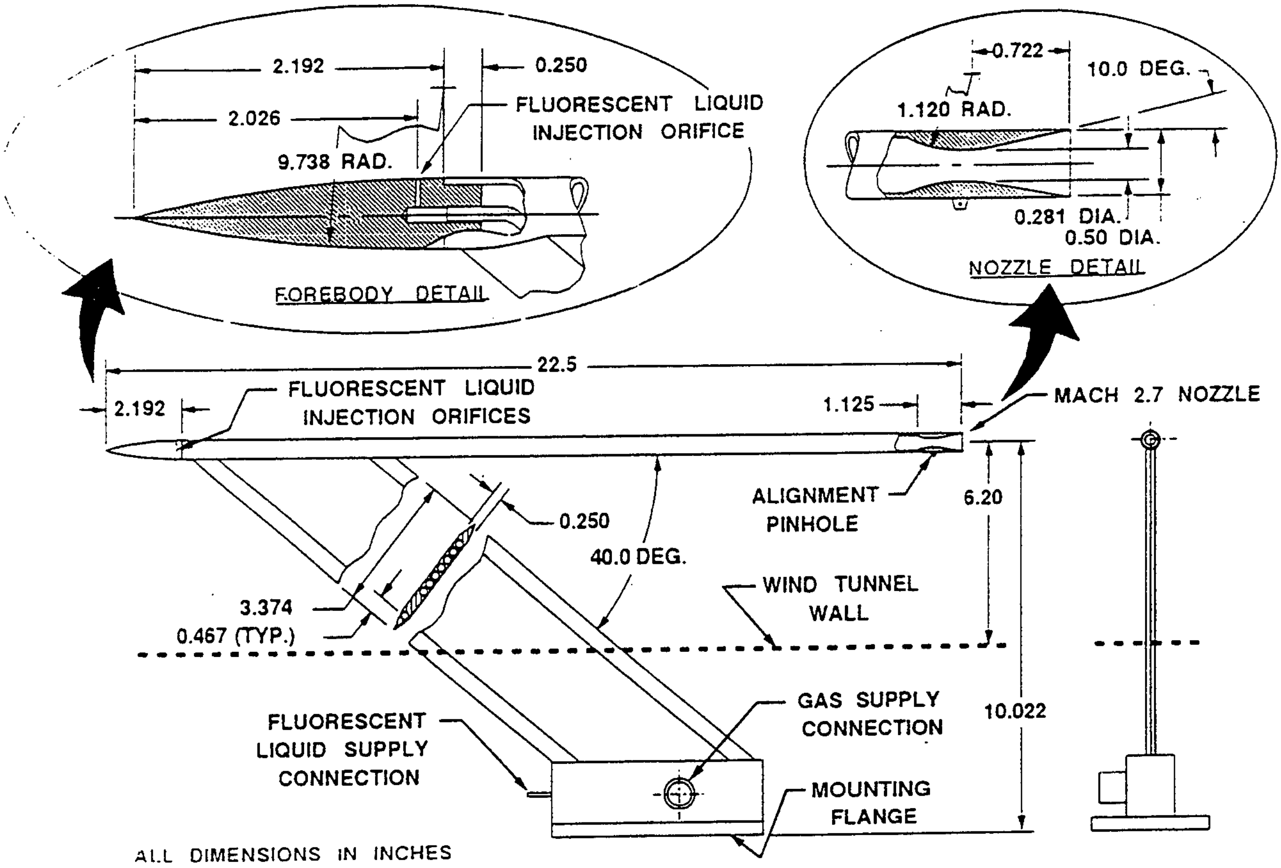


Fig. 1 Schematic of Small Generic Zero-Base Missile Model

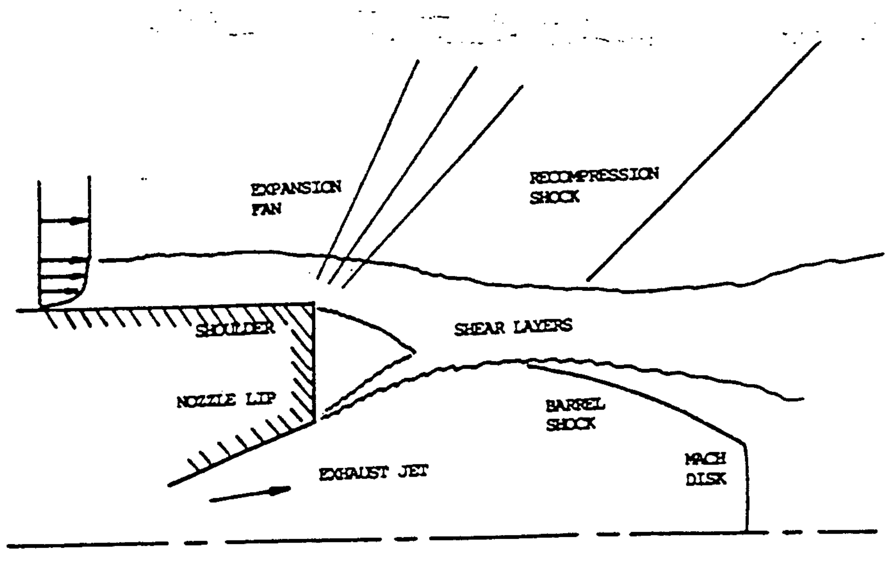
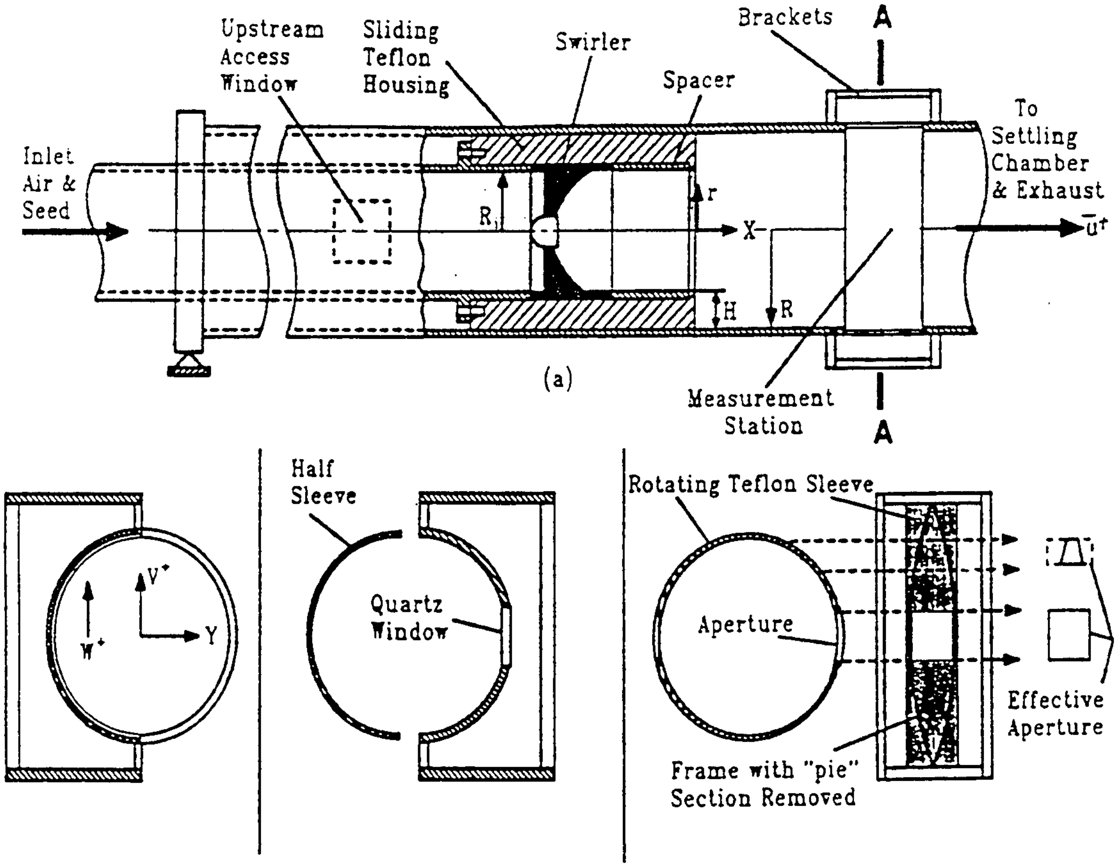


Fig. 2 Missile Base Flow Characteristics



(b) Section A - A

(c) Assembly for Measurements in the U-V Plane

(d) Assembly for Measurements in the U-W Plane

Fig. 3 Schematic of the Dump Combustor Model

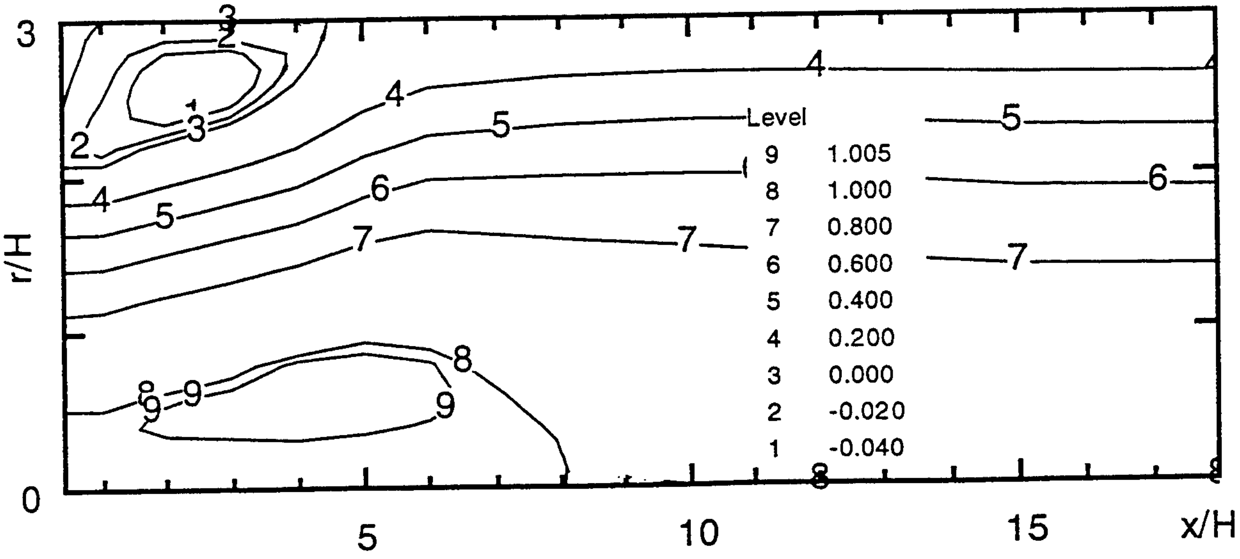


Fig. 4 Contours of Non-dimensional Stream Function

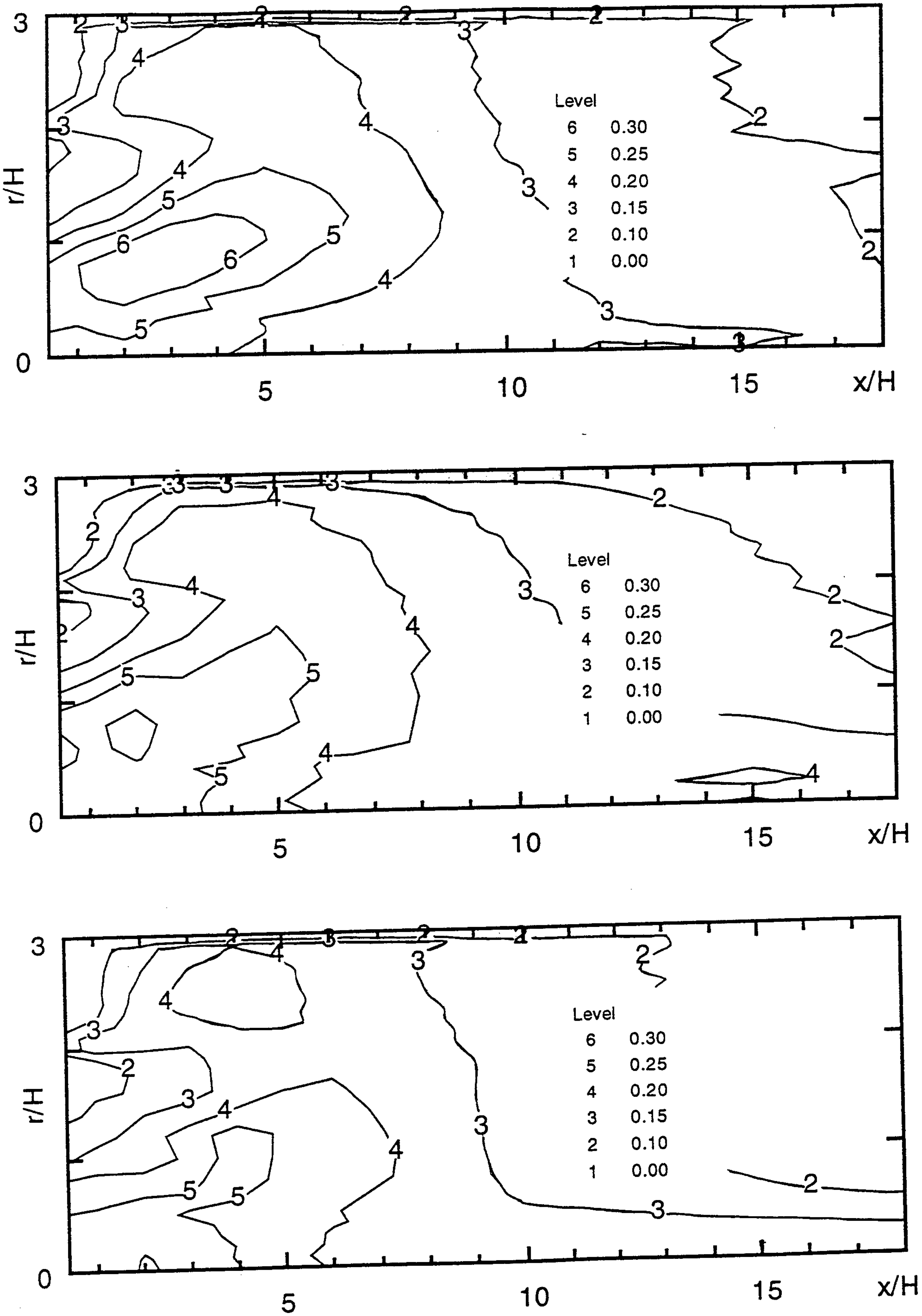
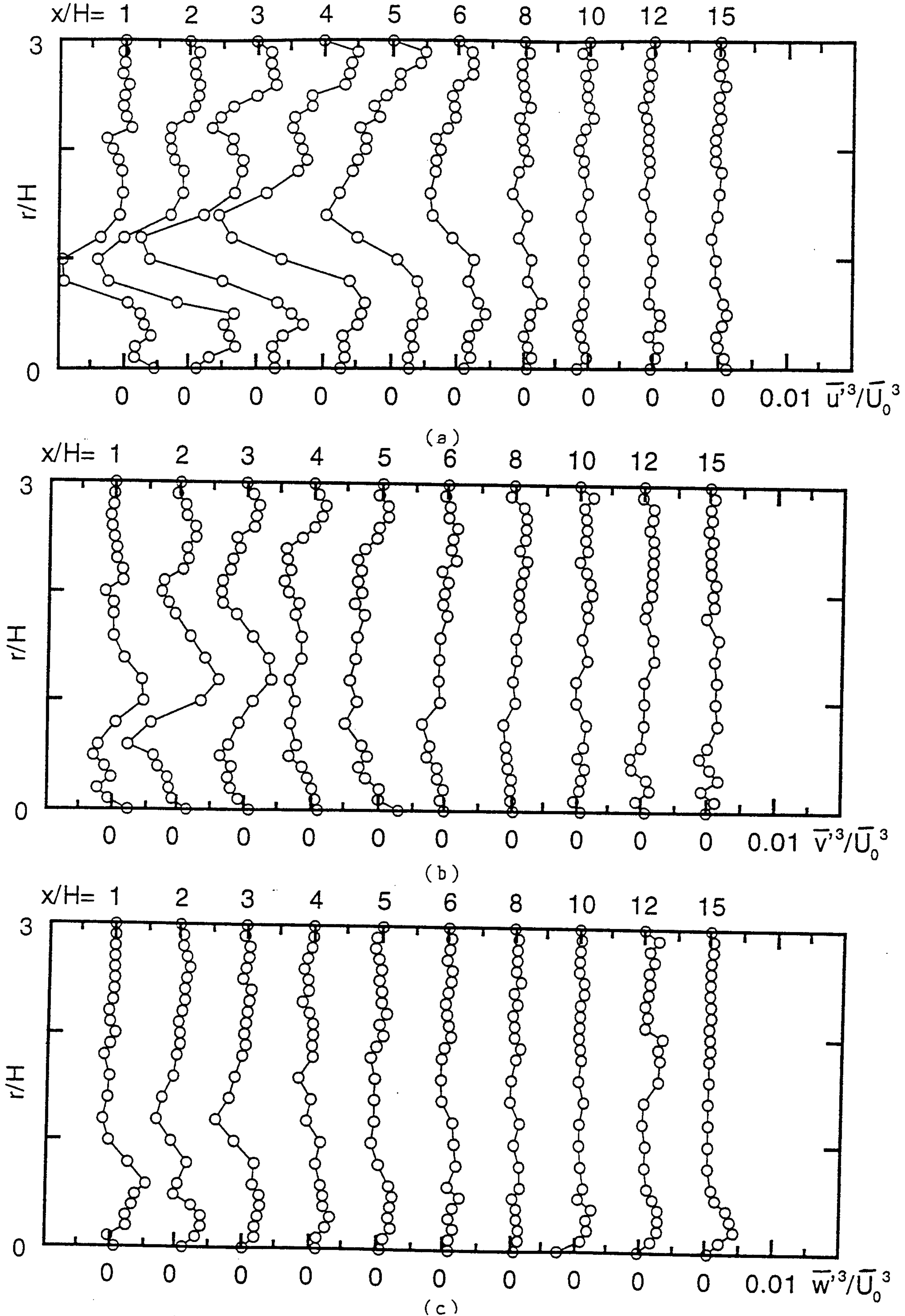


Fig. 5 Contours of Axial, Radial, and Tangential Turbulence Intensity; Respectively

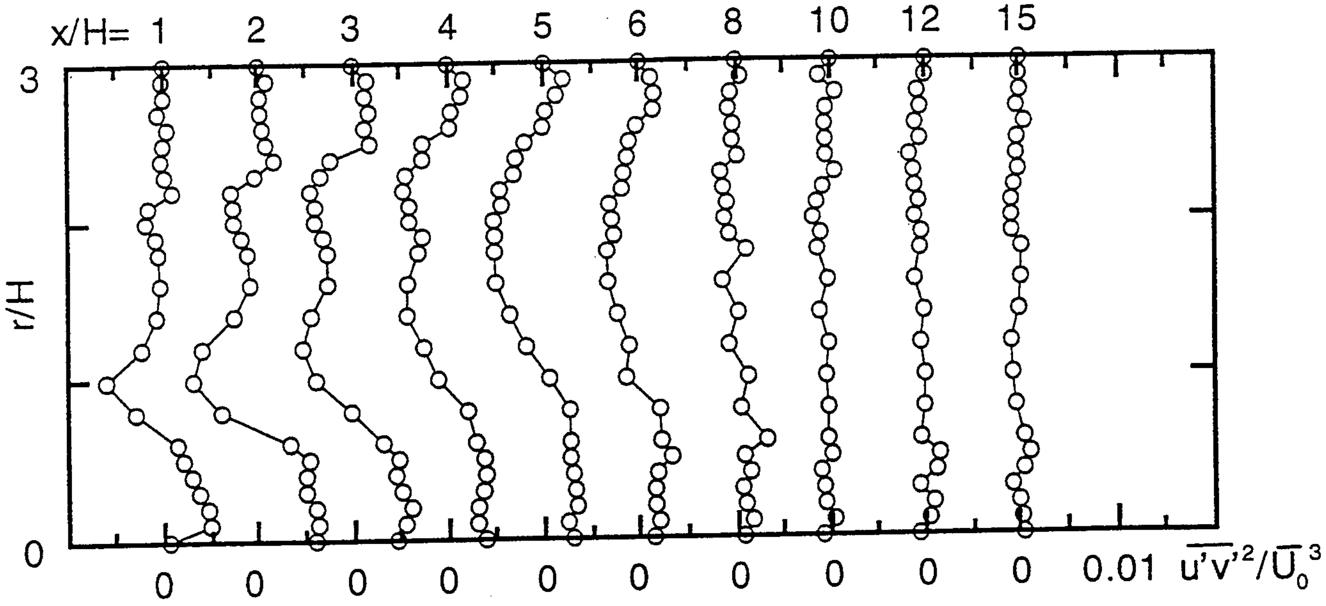

x/H= 1 2 3 4 5 6 8 10 12 15
3
r/H
0
0 0 0 0 0 0 0 0 0 0 0.01
$\overline{u'^3}/\overline{U}_0^3$
$\overline{v'^3}/\overline{U}_0^3$
$\overline{w'^3}/\overline{U}_0^3$

(a)

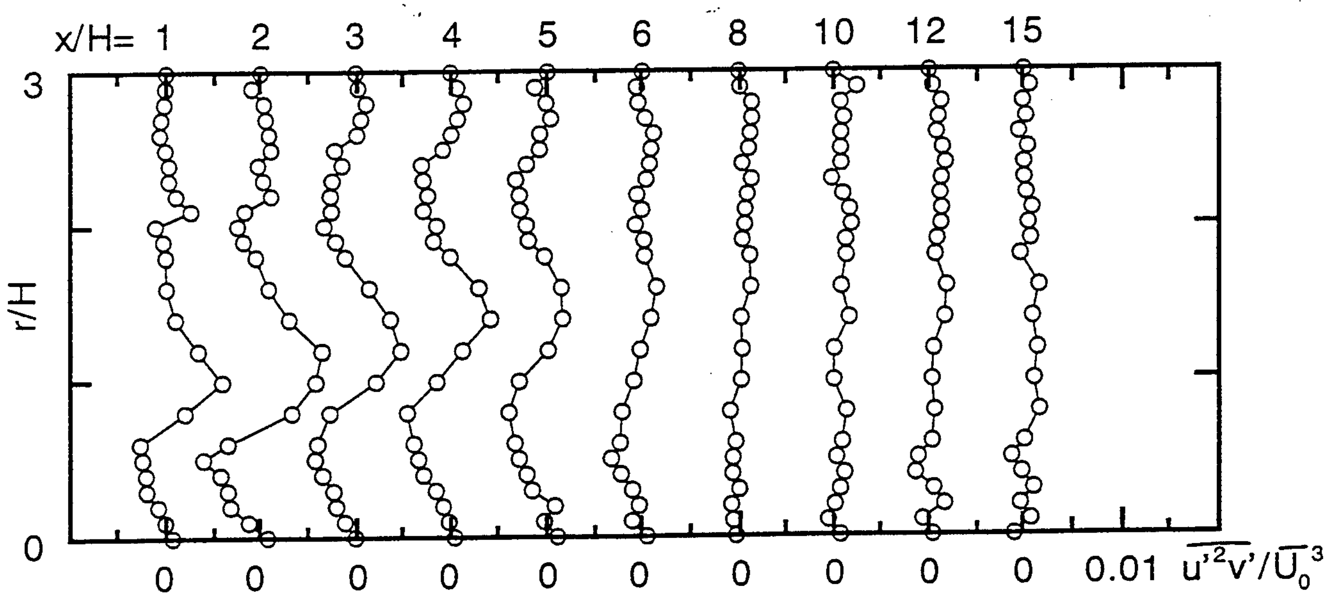
(b)

(c)

Fig. 6

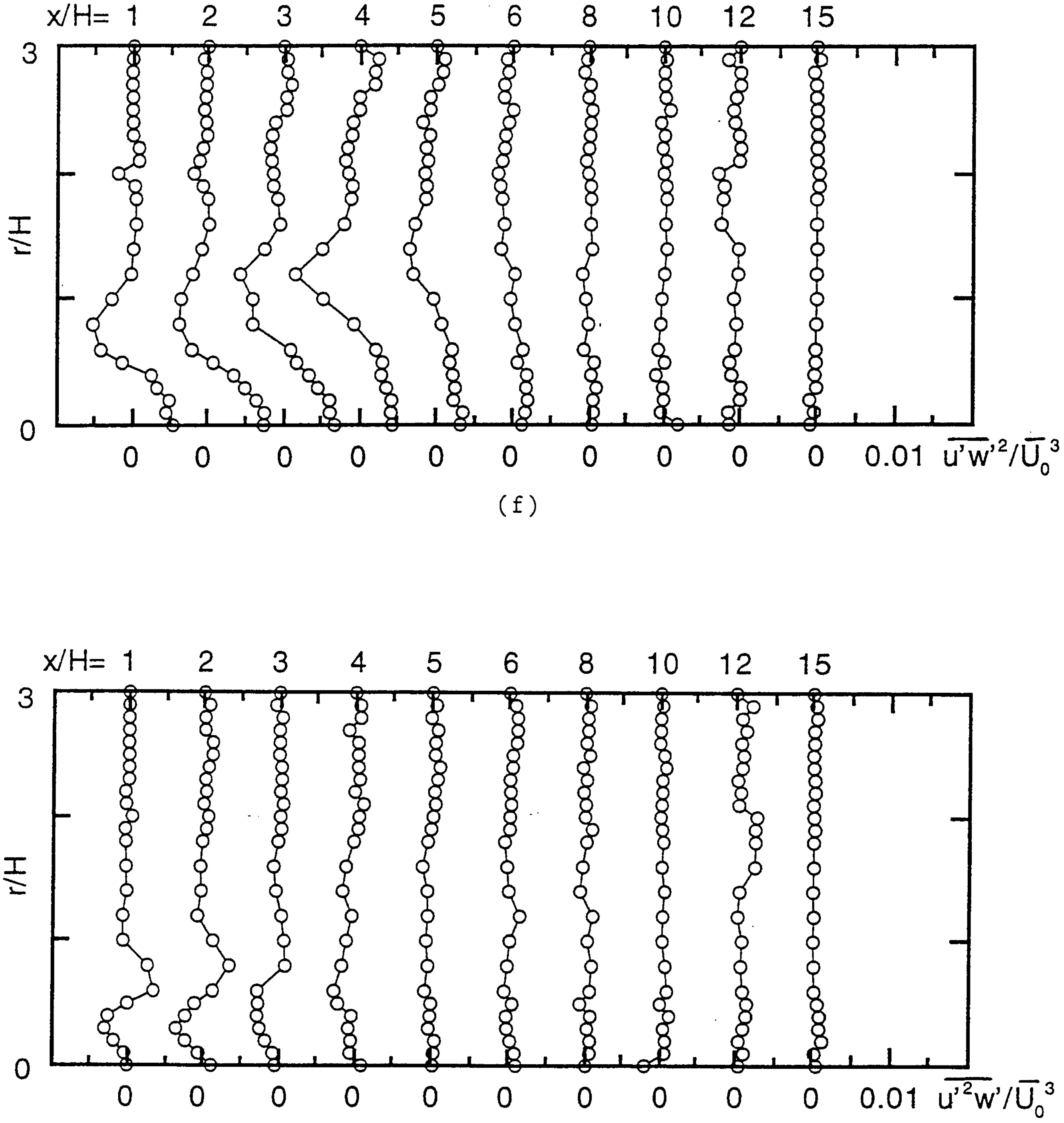

x/H= 1
2
3
4
5
6
8
10
12
15
3
r/H
0
0
0
0
0
0
0
0
0
0
0
0.01
u'v'²/Ū₀³

(d)


x/H= 1
2
3
4
5
6
8
10
12
15
3
r/H
0
0
0
0
0
0
0
0
0
0
0
0.01
u'²v'/Ū₀³

(e)

Fig. 6



(f)

(g)

Fig. 6 Evolution of Nondimensional Triple Velocity Profiles

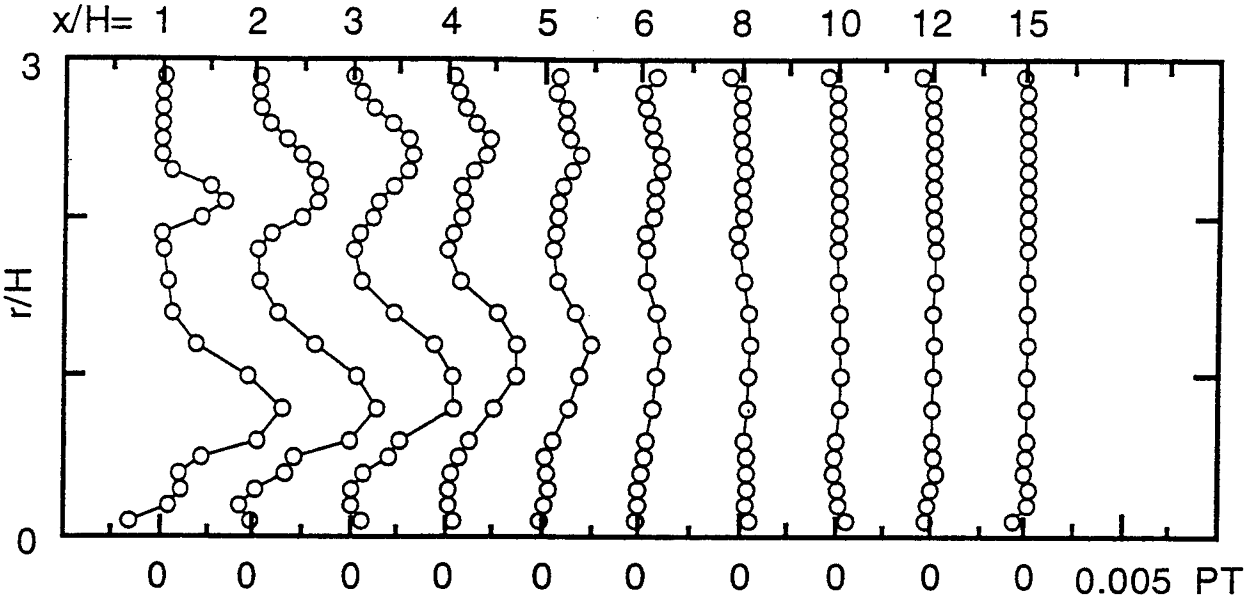


Fig. 7 Evolution of Normalized Production of Turbulent Kinetic Energy

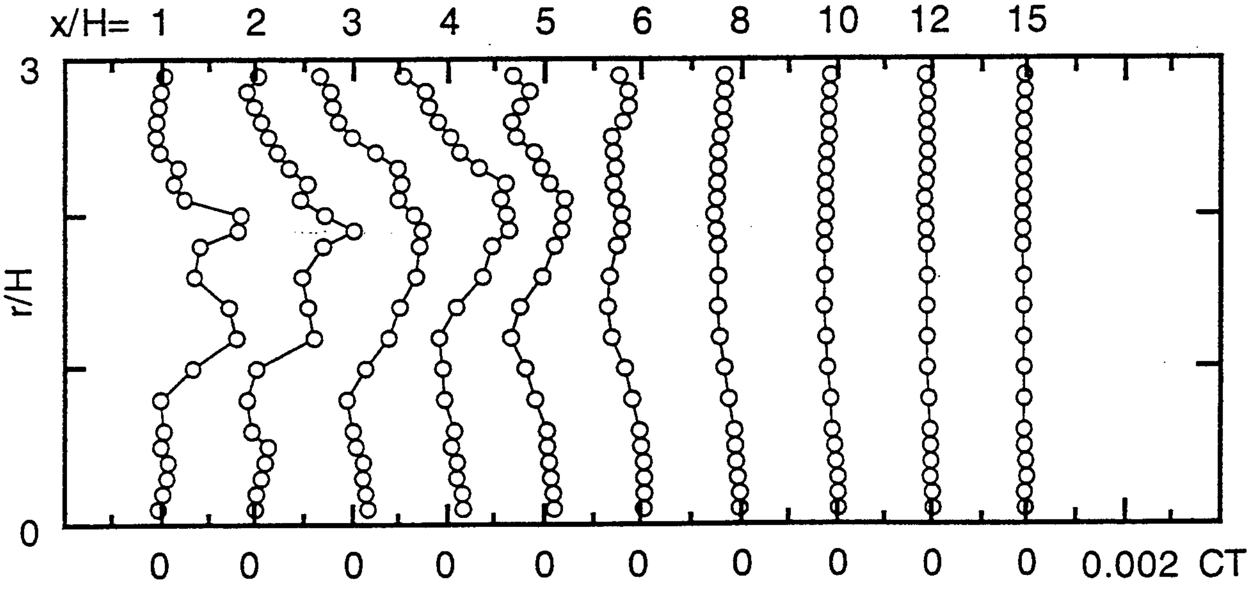


Fig. 8 Evolution of Normalized Convection of Turbulent Kinetic Energy

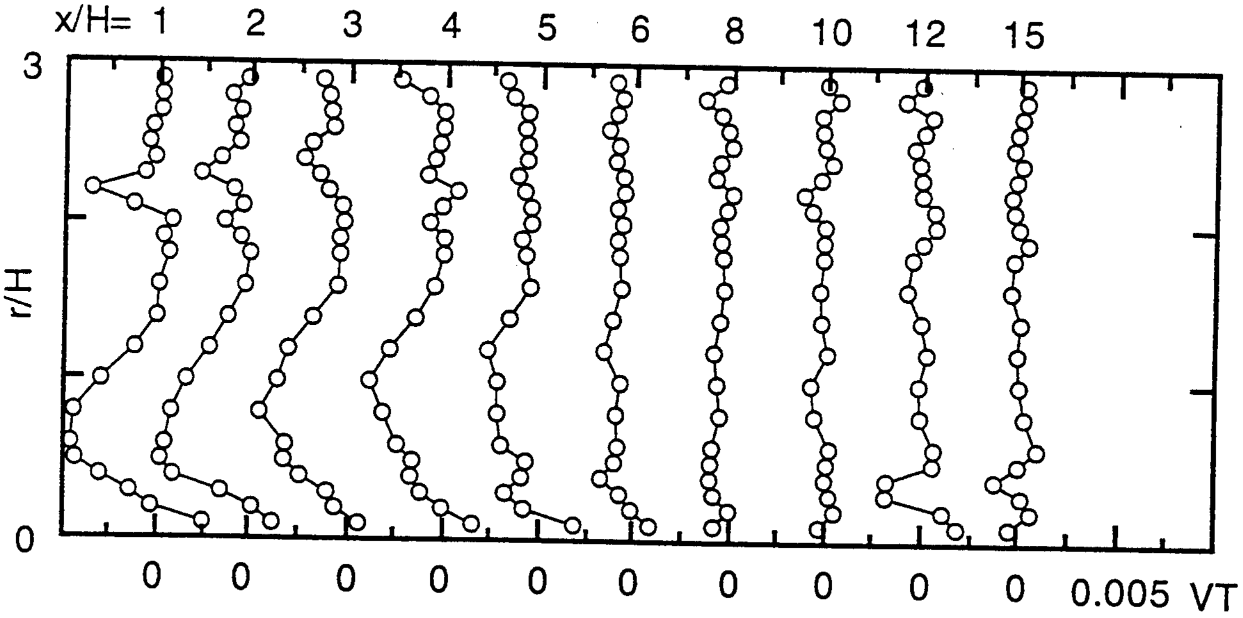


Fig. 9 Evolution of Normalized Diffusion of Turbulent Kinetic Energy

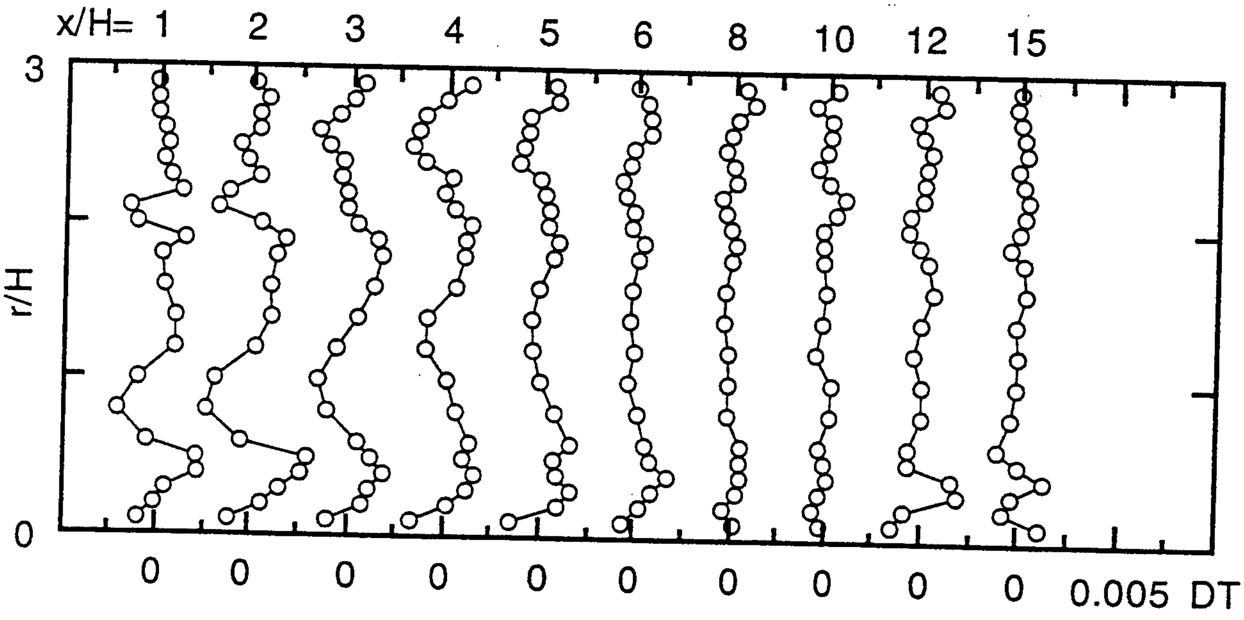


Fig. 10 Evolution of Normalized Viscous Dissipation of Turbulent Kinetic Energy

Turbine Engine Blade Vibration Analysis System

Csaba Biegl
Assistant Professor
Department of Electrical Engineering

Vanderbilt University
Nashville TN 37235

Final Report for:
Summer Research Program
Arnold Engineering Development Center

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington DC

and

Arnold Engineering Development Center

September 1996

TURBINE ENGINE BLADE VIBRATION ANALYSIS SYSTEM

Csaba Biegl
Research Assistant Professor
Vanderbilt University
Department of Electrical Engineering

Abstract

This report describes a system which is capable of reconstructing and displaying the actual movement of a vibrating compressor blade in a turbine engine. The system combines the outputs of real-time strain gage vibration measurements with a finite element model of the blade to accomplish this task. Various error analysis tools help the user to verify the results. The system is integrated with a graphical user interface for ease of operation and fast, on-line display of the results. The report discusses the design specifications and implementation of the system and concludes with initial application results.

TURBINE ENGINE BLADE VIBRATION ANALYSIS SYSTEM

Csaba Biegl

Introduction

The accurate measurement of the blade vibrations in operating turbine engines is an important design tool for controlling harmful component stresses. Unfortunately, measuring the blade vibrations in operating engines is a very complicated task. Typical measurements techniques, which include strain gages and various interference free methods (laser interferometry, etc..) only give displacements at a few points on the blade surface. It is not guaranteed that the highest stress level will occur at the instrumented locations.

A possible solution is to use measured vibration data together with the finite element model of the blade. The finite element model will provide the shapes and resonant frequencies of the various vibration modes of the blade. Knowing these modes, the measured signals can be used to determine the contribution of each mode to the overall blade vibration in the operating engine. As the finite element analysis produces detailed mode shape and stress information for every mode, it is also possible to reconstruct the displacements and stresses over the whole blade surface.

This report describes the theory, design specification and prototype development of a system which intends to solve this problem.

The CADDMAS system

One of the possible sources for measured blade vibration data is AEDC's CADDMAS (Computer Aided Dynamic Data Monitoring and Analysis System) system [1]. CADDMAS is a high speed parallel data acquisition, processing and storage system. It can process a large number of incoming analog signals with a bandwidth of up to 50 kHz. Various near real-time plots can be generated from the sampled data. These include time, frequency spectrum, spectrum envelope and Campbell plots. Data can also be propagated over the network and/or saved in various formats.

Finite element modeling

Finite element modeling is a frequently used design technique in mechanical engineering. The analyzed object's shape is digitized using a suitable grid generation technique. Adjacent grid points are then used to form small solid objects, called elements, from which the whole object can be reproduced. After specifying the properties of the material filling the volume of the elements, various finite element analysis programs can be used to determine the mechanical properties of the whole object. Typical analysis goals and results include stress patterns, vibration properties, etc.. From the point of view of the current effort, the vibration analysis results of the finite element modeling process are the most important. These results are typically specified in the form of mode shapes and mode frequencies. Mode shapes are surface wave forms at which the analyzed object is capable of stable oscillation. The mode frequencies are the natural frequencies at which these oscillations can occur. Most objects are capable of stable oscillation in various (half, one and a half, etc.. wave bending and rotation) modes. Typical analysis tools will compute and output a number of these modes simultaneously.

Several finite element modeling and analysis tools can be used to model turbine engine blades. The most commonly used ones are ANSYS and ALGOR, both of which are general purpose commercial packages, and BLADE-GT which is a specialized tool designed for blade analysis only.

Blade vibration analysis system -- operation principles

The goal of the blade mode visualization system is to reconstruct the displacements and stresses occurring on a vibrating blade inside an operating engine. To accomplish this goal, the system uses the finite element model of the blade and strain gage measurements of vibration signals acquired on-line during an actual engine test. The data samples are processed in the following stages:

1. *Preprocessing:* this involves performing a Fast Fourier Transform (FFT) on the incoming time domain data (this step may be unnecessary if the data source is CADDMAS, as CADDMAS can also provide frequency domain data), filtering and downsampling. Downsampling is necessary because it would be impossible to rerun the analysis for every time sample. These operations have to be performed for every strain gage signal.

2. *Frequency matching:* this step assigns the energy in the FFT peaks to the various blade modes. Ideally, all peaks in the FFT data should correspond to vibration mode frequencies. To account for small variations, a frequency tolerance band is set up around the nominal mode frequencies. Even allowing for these tolerances, it is expected that some FFT peaks will not be matched. These will be stored in so called *FFT reject* data samples. These FFT reject samples can serve as a good indicator of how well the finite element model matches the actual blade.
3. *Mode amplitude calculation:* once all strain gages' contribution to the frequency band of all modes is known, it is possible to determine the amplitude of each mode in the blade's overall vibration. This step uses precomputed sensitivity data for every mode at every gage's location. This sensitivity data is calculated from the mode shape, it is the signal amplitude a given gage at a given location and orientation should produce assuming the mode is present with unit amplitude. Mode amplitudes derived from several gages' measured data and sensitivity will typically disagree somewhat. The mode amplitude's best estimate is obtained using a least square fit using all available strain gage data. The resulting least square error term is also a good indicator of model fidelity.
4. *Blade shape and stress calculations:* The shape of the vibrating blade can be reconstructed by adding the scaled (with the mode amplitude) mode shape displacements to the coordinates of the finite element model nodes. Surface stress levels can be calculated by determining the distance change between adjacent surface grid points and scaling it with material property constants.

Blade vibration analysis system -- design criteria

The analysis system has to reliably and accurately reconstruct the vibrator behavior of a turbine engine blade during an engine test. If this is impossible, it should give clear indications of this fact. These requirements resulted in the following design criteria:

1. *Use finite element model analysis results from different packages:* Currently ADC uses three different packages for finite element modeling analysis of turbine blades: ALGERIA, ANISES, BLADE-GT. Users should be able to import data from any of these packages into the blade mode visualization

system. Not all the data which is used during the analysis is necessary; only the geometry (grid), mode shape and mode frequency data needs to be imported.

2. *Display various aspects of the measured and reconstructed vibration data:* To accurately judge the quality of the reconstructed blade vibrations, it is necessary to provide additional graphical information besides the image of the blade itself. Such data displays include:
    - The animated oscillating blade: This is a three dimensional image which shows a slowed down sequence of the actual movement of the blade.
    - Stress level displays: This is a contour plot of the stress levels experienced by the blade. Both peak stresses and animated stress level displays should be available.
    - Displacement displays: This is a two dimensional projection of the blade superimposed with the surface displacement component which is parallel with the axis of the projection. Such displays are especially important because these allow validation of the system with laser interferometry data.
    - Time and frequency domain plots of the strain gage signals.
    - Frequency domain reject strain gage signals.
    - Mode amplitudes, frequencies and errors: These are bar plots of calculated mode parameters.
    - Reconstructed strain gage time domain data: This data is generated by summing sine waves whose amplitude and phase is determined from the calculated mode parameters. If the model is accurate (i.e. there are no significant FFT reject peaks and the mode amplitude least square error term is small) then these time samples should match the original time data from the strain gages closely.
3. *Data source:* The system should operate either on-line, with CADDMAS as the data source, or off-line using saved strain gage time samples. During on-line operation CADDMAS sends FFT peak data (i.e. only the first few largest FFT peaks) at every sampling period and, occasionally, a full time domain sample.
4. *Near real-time operation:* It is not necessary to recalculate mode parameters (amplitude, actual frequency) after every sampled time block. It is sufficient to update these parameters in intervals of a

few seconds. However, all input data should be processed and largest peaks preserved, so that during the next update period they can be accounted for. This way short term vibration (and stress) peaks will not go unnoticed.

Blade vibration analysis system -- system design

Based on the above requirements an initial version of the blade vibration analysis system was implemented. The system was built using components from the universal simulation graphical user interface which was partly developed during last year's AFFAIRS Summer Research Program [2]. This user interface library provides many of the services needed by the blade visualization application, including the plot engine, plot window management, hard copy, data base services and configuration management. Some enhancements were necessary to adapt the three dimensional plot capabilities of the user interface library to the requirement of displaying solid objects instead of three dimensional function surfaces for which it was originally designed. To complete the blade analysis system the following extra components were implemented and integrated with the user interface (see the functional design in Figure 1):

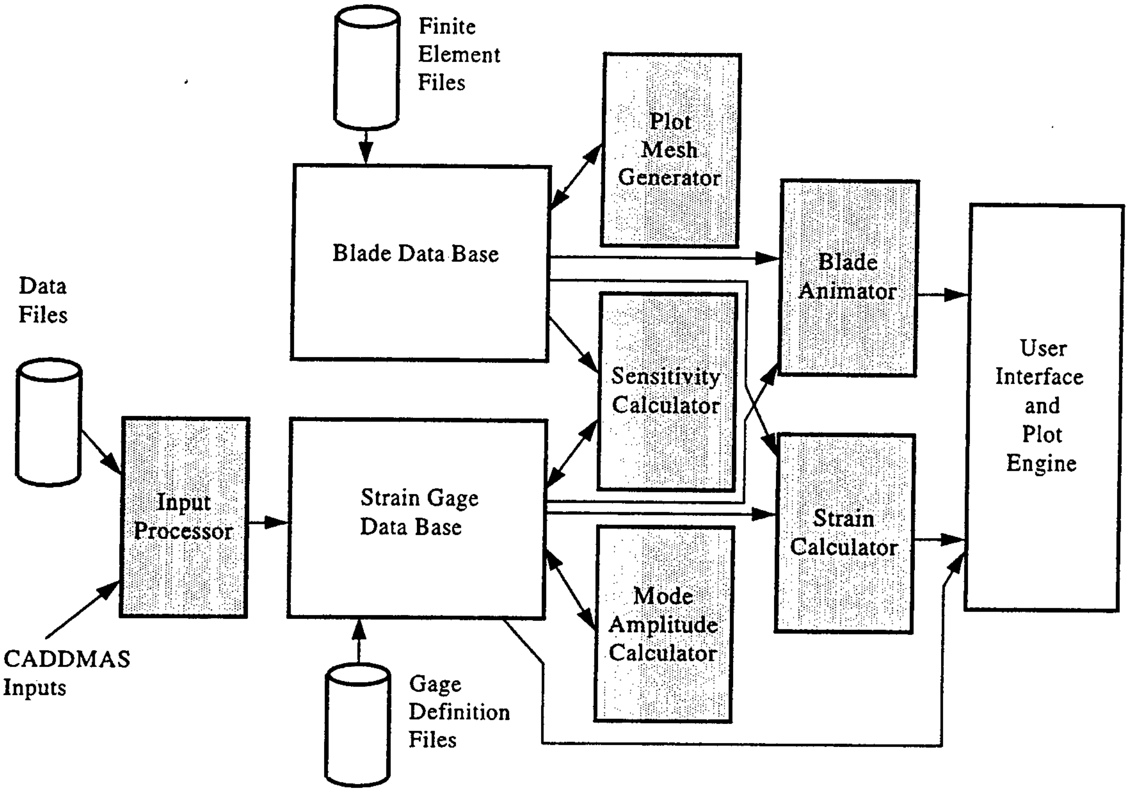


Figure 1. Design of the Blade Analysys System

1. *Blade data base:* This data base contains the finite element model information necessary to run the system. This includes the nodes of the model, the solid elements formed from the nodes and the vibration mode shapes and frequencies. This data base is set up by reading files generated by either of the supported finite element packages: ALGERIA ANISES or BLADE-GT. BLADE-GT. support is not complete due to a peculiarity in BLADE-GT.: Although it models the blade as a three dimensional solid object, the mode shapes it calculates are output only in two dimensions for the center plane of the blade. To use these modes in the blade mode visualization system, the mode shape displacements have to be projected back onto the three dimensional blade surface. Currently this is not done as it is expected that the first applications of the system will use either ANSYS or ALGOR models.

   The blade data base also contains additional information which is computed after loading the finite element model files. This includes element faces, face normals and midpoints, and face adjacencies for external (visible) faces. This information is used to speed up the displaying of the blade.

Any additional data which is typically found in the element model outputs (like material properties, etc..) is ignored by the system.

2. *Strain gage data base:* Besides the finite element model and the strain gage measurement data, the system also needs to know the location and orientation of the strain gages to properly reconstruct the blade vibration modes. The location of the gages is specified in a file supplied by the user. This file is read into the strain gage data base when the system is started up. Currently two types of strain gages are supported: single gage and a three gage 0-45-90 degree rosette. The strain gage data base also contains the current measured data record for each gage and computed gage related parameters, like the gages' sensitivity for different vibration modes.
3. *Plot mesh generator:* Various finite element packages use different conventions for specifying the coordinates of the model's nodes and elements. To display the three dimensional surface of the blade efficiently, the plot package of the graphical user interface expects only the surface points of the model specified in a well defined mesh order. After the blade has been loaded in the data base, the plot mesh generator computes this ordering, along with other useful information. It performs the following operations:
    - Identify element faces: different finite element modeling tools output the corner nodes of elements (e.g. bricks, etc..) in different order. For this reason the plot mesh generator does not assume any well defined order, but determines the element faces itself.
    - Classify all element faces as either external (visible) or internal (joined with an other element's face).
    - For all external faces compute face midpoint (this is the average of the face coordinates) and the face normal vector. This information is used by the strain gage sensitivity calculations.
    - Join neighboring visible faces into an ordered mesh suitable for efficient rendering. Optionally, the mesh can be generated from only a subset of the model's surface nodes. This is useful if the model resolution is so high that it would make a wire frame display of the object hard to interpret due to too many lines. This option also speeds up the display of the blade.

To properly process finite element models which represent objects with curved surfaces, the plot mesh generator allows element faces to be slightly non-planar.

4. *Sensitivity calculator:* Once both the strain gage location file and the finite element model files are loaded it is possible to calculate the various gages' sensitivity to the different modes of the blade. This calculation is done in the following steps:

   - Calculate gage end point coordinates from the gages' location (typically the center is specified), orientation and length. Rosette component gages are treated as individual gages.
   - For both gage end points find the element face (must be external) on which the end point is located. On curved surfaces it is possible that the gage specification is not fully accurate. In such cases the end point is mapped onto the nearest face using perpendicular projection.
   - For both end points calculate interpolation coefficients from the corners of the face to the end point.. A weighted sum of the face's corner point coordinates with these coefficients should reproduce the gage's end point coordinates.
   - Calculate the gage's sensitivity to each blade mode. This is done by applying the mode's displacements (with unit amplitude) to the end point face corner points. Using the interpolation factors calculated above it possible to determine the new coordinates of both gage end points. The gage's sensitivity to the given mode is defined as the end point distance change between rest and when the mode is active with unit amplitude.

   Once all gage sensitivities for all modes are calculated, they are stored in the strain gage data base.

5. *Input processor:* This component preprocesses the incoming strain gage signals and places the data into the strain gage data base. The input processor accepts either on-line data from CADDMAS (via a network interface) or can read measurement data stored in a file. Either way, the data is Fourier transformed and converted to absolute value and phase angle format. The time and spectrum magnitude and phase records are then stored in the strain gage data base. If the system operates off-line and the input data file contains multiple time sample records then the input processor allows the user to step forward and backward through the records in the input. Any time a different record is selected, the above calculations are repeated. During on-line operation CADDMAS sends two types of data records

to the system: full time sample records and frequency domain peaks. The input processor uses this data the following way:

- When a full time sample record is received, it is Fourier transformed and the spectrum magnitude and phase is calculated.
- Any additional frequency domain peaks are used to update the calculated spectrum magnitude record: if the peak sent by CADDMAS is bigger than the one already in the magnitude record at the proper frequency location then the new peak replaces the old one.
- When a new set of time domain samples arrives, the previous set is copied into the strain gage database, vibration mode parameters are recalculated, and the collection process outlined above starts again.

This on-line processing scheme makes it possible to run the rest of the system (i.e. mode amplitude calculations, etc..) at reasonable update rates. In this application CADDMAS is programmed to send complete time records once in a few seconds, but frequency domain peak updates may arrive much more frequently. On the other hand, no short transient vibration peak is lost, its processing is simply postponed until the next full update.

6. *Mode amplitude calculator:* The mode amplitude calculator is executed every time the input processor places a new set of preprocessed samples into the strain gage data base. It performs the following steps:

   - Determine each mode's measured amplitude in every gage's input. This is done using the spectrum magnitude records calculated from the gage inputs. Every mode is assumed to have a nominal frequency (this is a finite element analysis output) and a frequency tolerance band around it (which is controllable from the graphical user interface). The sum of all spectrum magnitude record points which fall into this band is used as the mode's measured amplitude at the given gage. The mode's measured frequency is the frequency value at which the largest peak was found.
   - Calculate mode amplitudes. This is done by selecting a mode amplitude value which minimizes the square error term:

$$ERR_i = \Sigma(m_{ij} - a_i\, s_{ij})^2$$

Where

i: mode index

j: strain gage index

m_{ij}: measured amplitude in the i-th mode's frequnecy band at the j-th gage

s_{ij}: j-th gage's sensitivity to i-th mode

a_i: i-th mode's amplitude

ERR_i: i-th mode's least square error term

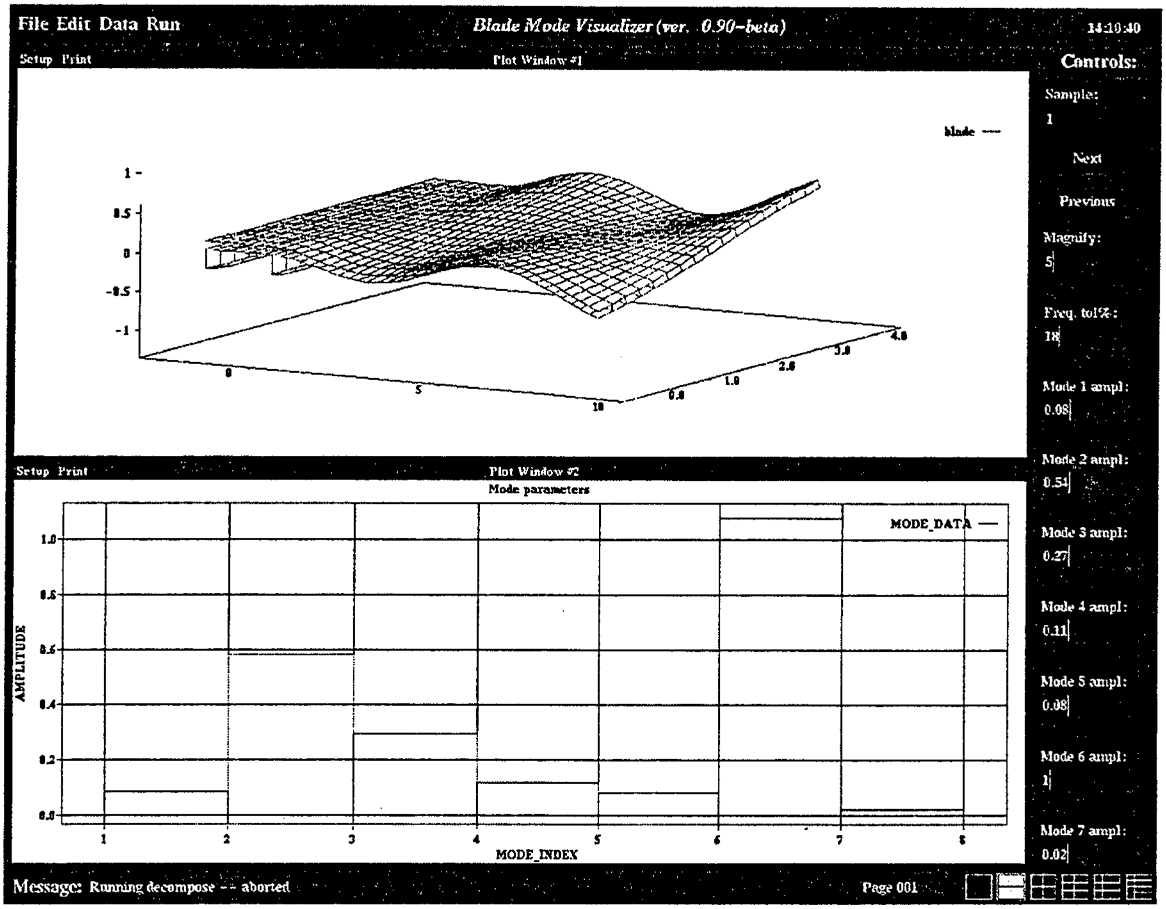
- Calculate and store data records which, when plotted, can be used to judge the quality of the mode amplitude calculations and thus can serve as an indicator for the finite element model validity. These data records include:
    - Spectrum magnitude reject (FFT reject) records: These are magnitude samples with zeros in the gages' frequency tolerance band, the original data elsewhere. Large peaks in magnitude reject plots indicate that the nominal mode frequencies differ from the measured signal spectrum peaks. Slight discrepancies can possibly be corrected with a wider frequency tolerance band, but large peaks probably signal the need to repeat the finite element analysis with different parameters.
    - Mode parameter records: These can be used to set up bar plots of mode amplitudes and frequencies.
    - Mode amplitude error terms: These are the residual least square error terms after the optimal mode amplitude has been chosen by minimizing the equation above. These are available for plotting either on a per gage or a per mode basis, or all of them can be displayed on a three dimensional surface plot. Large error terms mean that the ratio of the measured signals differs significantly from the ratio of calculated sensitivities. This may again mean invalid finite element analysis results, but can also be the sign of inaccurate gage positioning or a failing gage.
    - Reconstructed time data: These are time sample records generated by summing sine waves for each mode. Their frequency is the mode's measured frequency, the phase

is taken directly from the corresponding gage's spectrum phase records. The sine wave's amplitude is determined using the calculated mode amplitude and the gage's sensitivity to the given mode. These reconstructed time samples can be plotted against the original time domain measured data. If there were no significant frequency reject errors (i.e. most of the spectrums' energy contributed to the mode amplitude calculations) and the mode error terms were small (i.e. the measured signal ratios were in good agreement with the sensitivity ratios.) then the reconstructed time data should match the original time samples quite well.

7. *Blade animator:* This component generates the three dimensional display of the blade onto which a weighted sum of the mode shape displacements is superimposed. This display uses the calculated mode amplitudes which are applied at frequencies whose ratio matches the ratio of true mode frequencies. Thus this display is a realistic, but slowed down, image of the vibrating blade. The blade animator is also capable of running in a stand-alone mode when the system does not process any strain gage data, but the user is able to control the mode amplitudes directly from the graphical user interface.
8. *Strain calculator:* Material stresses are calculated along the surface grid lines of the finite element model. Stresses are assumed to be proportional to the length change in these lines. The stress map generated this way is then scaled using the strain gage input data (which is assumed to be the stress value in real engineering units at the gage location) to physical dimensions. These stress maps can be plotted superimposed onto two dimensional projections of the blade from various directions. The same plotting technique can also be used for surface displacements. Both of these plot categories are available either as animated plots (at the same rate and time scale as the animated blade plot) or peak displays.

Implementation of the system

The first version of the blade vibration analysis system has been implemented during the Summer Program. It is available on Intel PC-s which run a 32 bit version of Windows (95 or NT) or UNIX workstations running the X11 system. The system has been tested with finite element models of various objects. The most complete testing was done using the data from a previous AEDC vibration test which used a flat blade. Both a finite element model and actual multi-channel strain gage measurement data were available from this test, which made it possible to run all functions of the blade vibration analysis system. (Other available AEDC finite element models could only be displayed off-line due to the lack of actual strain gage data.) Captured screen images and plot hard copies from this test can be seen in figures 2 through 4.



Sample: 1
Magnify: 5
Freq. tol%: 18
Mode 1 ampl: 0.08
Mode 2 ampl: 0.54
Mode 3 ampl: 0.27
Mode 4 ampl: 0.11
Mode 5 ampl: 0.08
Mode 6 ampl: 1
Mode 7 ampl: 0.02

Message: Running decompose -- aborted

Figure 2. Captured user interface screen with blade image and mode amplitudes

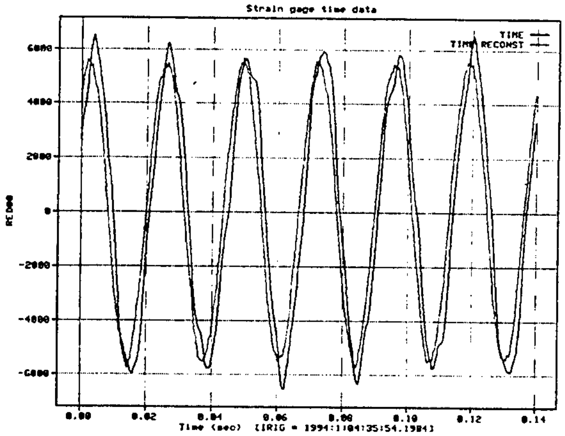


Figure 3. Example reconstructed time data

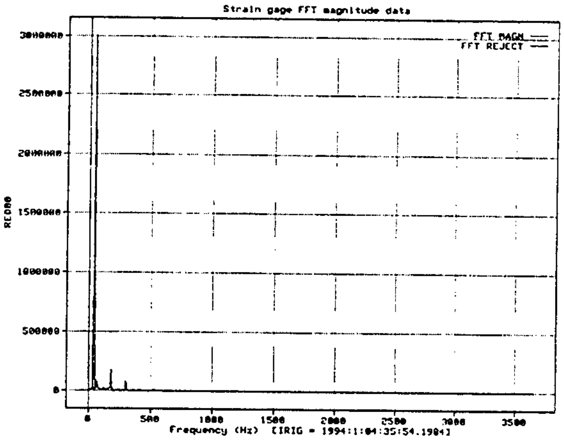


Figure 4. Example FFT magnitude data and FFT reject plot

Conclusions, future work

It is anticipated that the system will be used in an actual blade test in the near future. This will not be a complete engine test, but a laser interferometry test of a single blade with forced vibration. Comparing the laser interferometry results to the surface displacements generated by the vibration analysis system will give an excellent opportunity to validate the whole process. According to current plans, an application of the system for a real engine test is not far away either.

Actual use of the system will certainly reveal areas which require refinements. As far as major development efforts are concerned, the next step is to better integrate the blade vibration analysis system with a finite element modeling solver code. This is necessary because changes in operating turbine engines may necessitate a rerun of the finite element analysis. For example a change in engine speed will change the centrifugal forces on the blade. This can result in different mode shapes and frequencies. Finite element modeling tools can account for this by allowing the user to specify preload forces on the elements. Changing aerodynamic forces, component temperatures, etc.. all can have similar effects.

Currently the finite element modeling is done off line and the results are imported into the system. By incorporating a finite element model solver into the system it would be much easier to rerun the analysis whenever the model and the measured data become inconsistent. Additionally, it would also be possible to develop automatic adaptive mechanisms which detect when the model and the measured start to diverge and automatically trigger an analysis rerun.

A similar adaptation method may be necessary to deal with strain gage failures. The interior of operating turbine engines is a very hostile environment for strain gages. It is not uncommon to lose almost all gages by the end of an engine test period. As the system uses all available strain gage data to determine the mode amplitudes, a gage failure may introduce significant error terms. Such errors should not be classified as events which require a rerun of the finite element analysis, but instead the system should continue operation without using data from the failed gage.

Finally, if proven successful, the system should be integrated into a larger computerized engine test environment which provides comprehensive data with different types of parameters and analysis tools in a networked environment for a variety of users.

References

1. Bapty T., Abbott, B: "Parallel Signal Processing Architecture for Turbine Engine Testing", *Final Report for 1991 AFOSR Summer Graduate Student Research Program*

2. Biegl, C.: "Towards The Real-Time Execution Of Turbine Engine Simulation Programs", *Final Report for 1995 AFOSR Summer Faculty Research Program*

Laser Vapor Screen Flow Visualization Technique

Frank G. Collins
Professor
Department of Mechanical and Aerospace Engineering

The University of Tennessee Space Institute
Tullahoma, TN 37388-8898

Final Report for:
Summer Research Program
Arnold Engineering Development Center

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington DC

and

Arnold Engineering Development Center

September 1996

LASER VAPOR SCREEN FLOW VISUALIZATION TECHNIQUE

Frank G. Collins
Professor
Department of Mechanical and Aerospace Engineering
and Engineering Mechanics
The University of Tennessee Space Institute

Abstract

The vapor screen flow visualization method is fully described. Naturally occurring vaporization patterns are used to illustrate flow features that could be made visible using the method in the wind tunnel. Natural vaporization usually occurs in vortices on vehicles that are at large angles-of-attack. Previous wind tunnel use of the vapor screen method is examined with emphasis on the improvements that have been made over time, leading to the present use of a laser light sheet to illuminate the water droplets. The system used at NASA Langley is described in some detail, including the optics. Equations needed for analyzing the humidity conditions in a wind tunnel are given. The humidity conditions of previous tests are analyzed in some detail. It is shown that the method depends on the total amount of water vapor in the tunnel circuit rather than the specific humidity. Relative humidity approaching 1.0 in the stilling chamber is required to apply the technique to subsonic flow fields. Mie scattering patterns for various droplet sizes are computed to illustrate the benefits of viewing the laser light sheet from 50° rather than 90°.

LASER VAPOR SCREEN FLOW VISUALIZATION TECHNIQUE

Frank G. Collins

Introduction

Vapor screen is a means for making certain aspects of a flow field within a plane in a wind tunnel visible. It is particularly useful for exploring the properties of flows that have embedded vortices, such as flows over aircraft with strakes or other sharp-edged geometries that generate strong vortices or missile configurations at high angles of attack.

The vapor screen method operates by introducing an intense sheet of light into a wind tunnel, which is then scattered by water droplets (or sometimes ice crystals) that have been generated in the flow field. Scattered light is recorded at an appropriate angle to the light sheet by a video or still camera. The light sheet is generally introduced perpendicular to the free stream flow direction. Regions of concentrated vorticity are observed as white patches in subsonic flow or dark areas within a white sheet in supersonic flow.

Vapor screen requires that the local relative humidity exceeds 1.0 (in this report relative humidity is given as a decimal) and that the air becomes supersaturated. The supersaturated water vapor will then nucleate and condense to form small water droplets. The accompanying latent heat release brings the flow to a new equilibrium condition at a higher stagnation temperature, which is approximately saturated. The scattered light intensity distribution can be described by Mie scattering theory assuming spherical water droplets. If the number and size of the droplets are large enough, the light scattered by the water droplets will be sufficient to be recorded by the camera and a vapor screen image will be recorded. The success of the vapor screen method depends on both generating great numbers of large water droplets at the appropriate point in the flow field and on introducing an intense sheet light sheet into the wind tunnel.

In subsonic flow the free stream relative humidity must be increased sufficiently to allow condensation to occur within local regions of the flow over a model, such as in the core of vortices. These regions appear as white regions within the light sheet. Local condensation occurs because the local temperature in the vortex is decreased; the relative humidity is much more sensitive to changes in the local static temperature than to changes in pressure (Ref. 1). In supersonic flow the entire flow field saturates within the nozzle, allowing droplets to be formed in the free stream upstream of the model. The droplets are then partially removed by the intense circular motion within the core of the vortices, which appear as dark regions, devoid of water droplets, contained within a white sheet. The vapor screen method has recently found great value in studying vortical flows for free-stream Mach numbers greater than 0.6 (Ref. 2 to 5); it has also been used with some success at lower Mach numbers (Ref. 6).

Campbell, et al. (Ref. 1) collected around 300 photographs of aircraft which displayed natural water condensation while the aircraft were in flight. These photographs were collected to obtain a more complete understanding of flow features that can be observed due to condensation. Naturally-occurring patterns could be classified as spanwise "gull" patterns from swept wing-bodies, expansion and shock patterns for transonic flow, and leading-edge separation-induced vortex flow from slender wings and bodies. Generally the vehicles were at high angles-of-attack. The same types of flow features should be observable by the vapor screen method in the wind tunnel. Unfortunately, no atmospheric temperature, pressure, or humidity data accompanied the photographs.

The authors used existing CFD solutions for similar aircraft configurations to obtain flow field information and then added a humidity computation to the solution assuming that the atmosphere had a temperature of 280 K and a relative humidity of 0.85. This relative humidity value is somewhat higher than the average relative humidity given by the Standard Atmosphere (Ref. 7). They concluded that homogeneous nucleation of water vapor did not occur before the relative humidity reached 4.0 locally, and that otherwise heterogeneous nucleation occurred. Once condensation occurred, then it was thought that the relative humidity remained in the range 1 to 4 as a result of the latent heat release.

Some of their results are of present interest. For flow over a transonic airfoil, they predicted that the relative humidity reached 20.0 just before the shock wave on the upper airfoil surface. The relative humidity decreased through the shock due to the rapid increase in temperature, even though the pressure also increased. Since the temperature was higher in the boundary layer, the relative humidity was very low there. Supercritical conditions, with the relative humidity greater than 1.0, were predicted to extend far upward from the wing surface in the region of transonic flow, explaining the observation of large patches of condensation in some regions of transonic flow conditions.

For M = 0.3 flow over a delta wing at 25° angle-of-attack, the relative humidity increased to 1.25 in some portions of the vortices but it was predicted to be less than 1.0 within the vortex core, which appeared clear in the photograph. This may offer an additional explanation to why the core of vortices in supersonic flow are clear (or appear as dark spots in vapor screen images) other than due to centrifugal effects on the core droplets. The computer program predicted that within the vortex core the temperature increased and the pressure decreased.

Condensation was found to occasionally occur in the streamwise vorticity that is shed from the leading edge of a wing at high angle-of-attack. These vortices are caused by a three-dimensional cross-flow instability. The vortices were observed to disappear at mid-chord, not because they no longer existed, but because the local temperature increased in the compression above the wing causing the droplets to evaporate. Condensation is also observed to disappear at the point where a vortex bursts, probably due to the fact that there is an adverse pressure gradient and deceleration of the flow which causes a rise in temperature and pressure great enough to evaporate the droplets. The same occurs for flows that pass through shock waves. This suggest that caution should be exercised when comparing flows visualized by condensation with those obtained by seeding with dyes or smoke, because whereas the seeding particles cannot disappear, the water vapor can both condense and evaporate.

Previous Applications of Vapor Screen Flow Visualization

The vapor screen flow visualization method was developed by Allen and Perkins (Ref. 8) in 1950 to examine the vortex system created by an ogive-cylinder at high angles-of-attack in supersonic flow ($M = 2$) in the NASA Ames 1- by 3-foot supersonic wind tunnel. They added water vapor to the tunnel, in a manner not described, which condensed in the nozzle to form a fine fog in the test section. The fog was illuminated by a narrow beam of white light from a mercury vapor lamp. The light beam entered the tunnel from one side and the scattered light was photographed from the other, at approximately a 45 degree angle to the light beam. The vortices were observed as dark patches in a white sheet. It was speculated that the droplets were removed from the core of the vortex by the circular motion within the vortices, thus leaving a dark patch in the light sheet. Shock waves were also observed as a change in light intensity.

The vapor screen method was thoroughly investigated later by McGregor (Ref. 9) in the 3-Foot R. A. E. Bedford tunnel. While most of his investigation was performed on the flow over delta wing in the Mach number range of 1.3 to 2.0, some examination was made using the same model at Mach numbers of 0.8-0.85. The tunnel air was very thoroughly dried (dew point was reduced to -40 °C ; $\omega = 0.000078$, where ω is the specific humidity) and then measured amounts of water vapor were added to the circuit. For supersonic flow the air/vapor mixture cooled in the nozzle to form a fog in the test section. The fog was illuminated by a beam of white light from a 1000 watt mercury-vapor lamp. At the center of the test section the light beam measured 0.15" by 9". The degree of condensation was found to depend upon the amount of water in the tunnel circuit and not on the value of the specific humidity. Rather the frost point was found to give a unique relation between the amount of water in the tunnel circuit and the tunnel volume. This observation is explained later. Scattered light could be initially observed at a supersaturation ratio of just greater than 15, but useful photographs could not be obtained until additional water had been added to the circuit. The light scattering was recorded from downstream, at 90° to the light sheet, by film using exposures as long as 80 seconds. The best results were obtained with what appeared to be quite thin fog as viewed from the side of the tunnel.

McGregor used measurements of the diffraction of the white light by the water droplets to estimate that the mean droplet radius was 0.74 μm at $M = 1.41$ ($\omega = 0.0092$, $\phi_o = 0.064$, where ϕ_o is the relative humidity in the stilling chamber). The fact that he was able to clearly measure the angle of the first maxima of the different colors indicates that the droplets must have been of very uniform size. The droplet density was $6 \times 10^{11}/m^3$. The droplets were found to increase in size with increasing humidity. No droplet size measurements were made for a subsonic flow.

Considerably more water vapor was required to generate a useful vapor screen image at subsonic Mach numbers. At $M = 0.8$, the specific humidity had to be increased to 0.0588 with a stilling chamber relative humidity was 0.78, but a specific humidity of only 0.0394 ($\phi_o = 0.53$) was required at $M = 0.85$. The droplet size was much larger in subsonic flow than in supersonic flow but the size was not reported.

Before continuing with this review of vapor screen use, it is of interest to know the size of the water droplets. Durbin (Ref. 10) used light scattering to measure the size of water droplets in the NASA Langley 11-inch hypersonic wind tunnel test section; the droplets had formed in the nozzle. He inferred the particle size from measured scattered light intensity at two angles and from measured light attenuation. At $M = 6.86$ the droplets had a radius of 50 nm with a concentration of $10^{16}/m^3$. Stein and Wegener (Ref. 11) found a mean droplet radius of 3.5 nm for condensing flow through a $M = 2$ nozzle, with $\omega = 0.0032$. They measured the mean droplet diameter and concentration as a function of the specific humidity. The mean radius was found to increase with increasing specific humidity from 2 to 6 nm while the number density decreased until $\omega = 0.003$ and then it remained relatively constant at $4 \times 10^{18}/m^3$. They could not tell whether the droplets were water vapor or condensed air constituents.

Size distributions have been measured in clouds (see Ref. 12). While the size varies with the cloud, many have mean radii less than 10 μm. The radius cannot be greater than about 40 μm (after that size the droplets grow by coalescence). In a wind tunnel their size in the test section depends upon the specific humidity and the time since nucleation; thus, there is a scale effect when applied to small wind tunnel models.

Keener (Ref 13) performed a comprehensive examination of flow separation patterns on various symmetrically-shaped forebodies using several flow visualizations techniques, including vapor screen. The tests were conducted in the NASA Ames 6- by 6-Foot Transonic/Supersonic Wind Tunnel over the Mach number range 0.1 to 2 with angle-of-attack to 86°. A light sheet was formed by focusing the output from six 1000-watt mercury-vapor lamps on each side of the tunnel and introducing the light from both sides of the tunnel simultaneously. Still photographs were taken of the vapor screen images. Photographs were published from flows in the Mach number range of 0.6 to 2.0, for angle-of-attack equal to and greater than 30°. The method of controlling the humidity in the tunnel was not described. It was thought that for subsonic a free-stream Mach number the water condensed within the vortices, which appeared white, and for a supersonic Mach number the water condensed in the nozzle, forming ice crystals which were centrifuged out of the vortex centers leaving a dark spot. Transonic conditions gave a combination of light and dark areas, depending on the local temperature and pressure (condensation or evaporation).

Erickson (Ref. 2, 4, 14, 15) performed a series of tests in the 7- by 10-Foot Transonic Tunnel at David Taylor Research Center. These studies were performed on various generic delta wing configurations and on the F/A-18 fighter. The tests were conducted over a wide subsonic Mach number range and to large angles-of-attack. Some models included leading-edge flap deflections and forebody strakes, which generated additional vortices. Originally the tunnel humidity was controlled by putting water vapor in the tunnel before the test (Ref, 2) but later it was injected into the settling chamber with a spray nozzle (Ref. 4). The light sheet was generated by an 18-watt argon-ion laser, located outside the tunnel (the tunnel was operated sub-atmospheric which was not compatible with the laser). The laser light was steered into the tunnel with mirrors. The models were painted black to increase contrast (in Ref. 4 the tunnel interior was also painted).

The laser light was admitted from various directions for the different tests. Initially (Ref. 2) it was admitted from one side of the tunnel but the model caused shadows on the far side and so in later tests it was admitted from above the model, through the tunnel ceiling during the tests (Ref. 4). In all cases the plane of the laser light was somewhat adjustable in the axial direction and with respect to the plane of the model, to keep it approximately perpendicular to the model plane. The scattered light was generally photographed or recorded on video tape from an angle on the order of 45° from the forward scattering direction as well as by a sting-mounted camera at 90° to the light plane. Condensation in vortices could frequently be observed with only the wind tunnel test section interior lights (Ref. 4). Photographs were published of vapor screen images for Mach numbers as low as 0.4.

Erickson performed similar tests in the NASA Ames 6- by 6-Foot Transonic/Supersonic Wind Tunnel (Ref. 3). Water was injected into the tunnel from a single orifice on the tunnel ceiling. The laser light sheet was introduced through the side of the tunnel and the vapor screen image was recorded from downstream.

Most recently Erickson has added an improved version of the laser vapor screen (LVS) flow visualization technique to the 8-Foot Transonic Pressure Tunnel (TPT) and the 7- by 10-Foot High-Speed Tunnel (HST) at NASA Langley (Ref. 6). While most of the details of the installation apply to the TPT, some information was given for the HST. The TPT is a continuous flow. closed-return tunnel, stagnation pressure 0.25 to 2.0 atmospheres, and Mach number 0.2 to 1.2. Water was added from six atomizing nozzles located in the ceiling of the diffuser; deionized water was used. The full-cone spray patterns encompassed most of the diffuser cross section under tunnel-off conditions. The diffuser section was chosen for the water spray nozzles because of the relatively high cross flow velocity in the diffuser that would assist in the atomization of the water droplets. It could not be determined if complete vaporization occurred in the diffuser. The LVS method was used over the Mach number range of 0.6 to 1.2. It is estimated that a test section relative humidity of 0.98 was required in the test section at $M = 0.6$ for successful application of the vapor screen. The HST is a continuous-flow, subsonic/transonic ($M = 0.06$ to 0.94) tunnel that operates at ambient pressure and temperature and continuously exchanges air with the atmosphere. The humidity was measured with a hygrometer in the settling chamber. A measured relative humidity in the settling chamber as low as 0.30 yielded satisfactory vapor screen images and frequently is was unnecessary to add additional water to the flow. LVS images were obtained on vortex-dominated models at high angles-of-attack for $M= 0.2$ to 0.85. The optical details of the LVS installations in these tunnels will be described in a later section.

The modern vapor screen application, such as described in Ref. 6, uses a laser to generate the light sheet. It is thus referred to as the laser vapor screen (LVS) method. A typical modern system would consist of the following elements:

- a multi-mode high-power laser, using all spectral lines to maximize the output power;
- a fiber optic cable (single fiber, step-index, multimode);
- laser sheet optics, including light sheet steering controls;
- video equipment for recording the scattered light;
- water injection system to increase (and control) the specific humidity in the wind tunnel;
- means for measuring the relative humidity in the free stream of the tunnel test section.

The laser sheet optics must have the capability to control the focus (thickness) of the laser sheet and the spread (width) of the laser sheet. Since a laser and fiber optics is an extended luminous source, as opposed to a point source, the laser light can only be imaged to a waist.

The NASA Langley laser vapor screen system (Ref. 6) uses a fiber optic cable to introduce the laser light sheet into the wind tunnel at the desired location. The system consists of the following elements:

- argon-ion laser with up to 15 watts of power, using all lines, operating in multimode fashion to maximize the output power; optics designed for beam diameters from 800 to 1800 μm;
- fiber optic, single, 200 μm core diameter, step-index, multimode, armored and contained in copper tubing for strain relief, numerical aperture = 0.22, usable to 20 watts of input power;
- laser-to-fiber coupler (SMA connector plus a focusing lens with a reflective coating which is connected to vertical, horizontal, and angular adjustment blocks);
- laser sheet projection optics, which consists of a SMA connector to the fiber optic, a translating achromatic double lens, 3.15 inch focal length, corrected for coma and to cancel chromatic aberration, and two identical plano-convex cylindrical lenses, one translating and one fixed (generates the light sheet and controls the sheet divergence (spread angle) from 4° to 22° and focus (sheet thickness);
- plane mirror on the top of the tunnel for axial positioning of the light sheet.

The laser light sheet is introduced through the ceiling of each tunnel and the scattered light is recorded with a color video camera located downstream of the model on the sting. The fiber optic output is an extended luminous source (not a point source) and can only be imaged to a waist.

Translation of the achromatic lens relative to the cylindrical lenses controls the light sheet thickness or focus (minimum thickness of 0.06 inches at the model surface).

Wind Tunnel Humidity Analysis

Some preliminary tests of the laser vapor screen method were performed in the 16T wind tunnel at AEDC. The equations used to determine the humidity conditions in that tunnel will be given in this section.

16T is a closed-circuit, continuous-flow subsonic/transonic wind tunnel that can be operated with the stagnation pressure above or below the atmospheric pressure. The humidity is controlled by passing the tunnel make-up air through a dryer. The temperature is controlled by a heat exchanger (cooler) located upstream of the turning vanes that are upstream of the stilling chamber. The water flow through the cooler is carefully controlled so that the air temperature distribution, measured by thermocouples placed downstream of the heat exchanger, is uniform. A uniform test section temperature is a requirement for the successful application of the vapor screen method.

The AEDC 16T tunnel conditions are given by making measurements in the stilling chamber. The measured quantities include the stilling chamber pressure, p_o, the stilling chamber temperature, $T_0 = T_{vp}$, and the stilling chamber dew point temperature, T_{dp}. The test section Mach number, M, is computed. It was necessary to use these measurements to determine the state of the air/water vapor mixture in the stilling chamber and test section. The computations proceed as follows (the general equations for humidity are derived in Ref. 21):

1. The partial pressure of the water vapor at the dew point temperature in the stilling chamber is computed from the equation (the equations given in Ref. 16 were used rather than fitting the data given in Ref. 17)

   $$T_{dp} = -17.958 + 40.582P + 3.7421P^2 + .3938P^3$$

   if $T_{dp} \leq 32$ °F and the equation

   $$T_{dp} = -21.001 + 38.964P + 8.122P^2$$

   if $T_{dp} > 32$ °F, where $P = \log_{10} p_{dp}$ These nonlinear equations are solved for p_{dp} using a Newton-Raphson procedure. In these and the following equations, unless explicitly stated. the temperatures are given in degrees F and the pressures in psf. The following equation, obtained by fitting a cubic equation to the vapor pressure data for ice in Ref. 17, can be used for temperatures less than 32 °F (within the range -55 °F to 32 °F). Comparison of the use of this equation with the first equation is excellent.

$$T_{vp_i} = -18.009 + 40.5606P + 3.8148P^2 + .38861P^3$$

where $P = \log_{10} p_{vp}$. The equations given above are approximations to the very nonlinear variation of water vapor pressure with temperature.

2. Compute the partial pressure of the water vapor at the stilling chamber temperature using the same equation, except with $T_{vp} = T_0$ and $P = \log_{10} p_{vp}$.
3. The relative humidity in the stilling chamber is computed using the pressures computed above from the equation

$$\phi = p_{dp} / p_{vp}.$$

4. The specific humidity in the tunnel is computed from the equation

$$\omega = 0.622 p_{dp} / \left(p - p_{dp}\right).$$

5. The test section pressure and temperature are computed from the ideal gas compressible flow equations

$$p_{ts} = p\left(1 + 0.2M^2\right)^{-3.5}$$

and

$$T_{ts} = -459.67 + T_{vp}\left(1 + 0.2M^2\right)^{-1}.$$

6. The partial pressure of the water vapor in the test section at temperature T_{ts}, p_{vts}, is computed from the equation

$$p_{vp_{ts}} = \omega p_{ts} / (0.622 + \omega)$$

where the specific humidity, ω, has the value computed from step 4.

7. The test section dew point pressure, p_{dpts}, corresponding to the test section temperature, T_{ts}, is computed from the same equations used in step 1.
8. Finally, the relative humidity in the test section is given by the equation

$$\phi_{ts} = p_{vp_{ts}} / p_{dp_{ts}}.$$

A number of assumptions have been made in writing these equations. It has been assumed that moist air is a mixture of calorically and thermally perfect gases (see Ref. 18 to 21). The equilibrium between the liquid and its vapor is not influenced by the presence of the air. The partial pressure of the water vapor is equal to the saturation pressure corresponding to the temperature of the mixture (assuming a planar interface between the two phases). The dew point is the temperature at which the water vapor condenses on a plane when it is cooled at constant pressure. The partial pressure of the vapor, and thus its mole fraction, remains constant until condensation begins. $\gamma = 1.4$ is assumed for the air/vapor mixture (see Ref. 22).

An approximation to the test section relative humidity for given stagnation chamber relative humidity was given in Ref. 1. Using the ideal gas relations for p_{ts}/p_o, T_{ts}/T_o for given Mach number, the expression is

$$\phi_{ts} = \phi_o \left(\frac{p_{ts}}{p_o} \right) 10^{-\left\{ \frac{2263}{T_o(K)} \left(\frac{T_{ts}/T_o - 1}{T_{ts}/T_o} \right) \right\}}.$$

This equation is not accurate for very large values of relative humidity but it does illustrate the fact that the relative humidity is directly proportional to pressure but logrithmically proportional to temperature. Thus the relative humidity is more sensitive to changes in temperature than to changes in pressure.

Analysis of Previous Vapor Screen Applications

In the wind tunnel there is a functional relationship between the specific humidity and the pressure, temperature, and relative humidity in the stilling chamber. That relationship can be obtained as follows (in the following no subscript will be placed on the quantities but they can be understood to apply to the stilling chamber conditions). Both the air and the water vapor are assumed to be perfect gases. Then the pressure is the sum of the partial pressures of air (subscript a) and water vapor (subscript v), at the temperature T. That is, for a mixture volume V,

$$p = p_a + p_v = \frac{m_a R_a T}{V} + \frac{m_v R_v T}{V}.$$

In this equation R_a and R_v are the specific gas constants for air and water vapor, respectively, and m_a and m_v are the masses of air and water vapor in the volume V.

Using the definitions of the specific and relative humidity in the forms (Ref. 21)

$$\omega = m_v / m_a; \quad \phi = p_v(T)/p_s(T),$$

where $p_v(T)$ is the vapor pressure of the water vapor at the temperature T and $p_s(T)$ is the saturation pressure at the same temperature, then the following relation can be derived

$$\omega = \frac{p_v(T)}{p - p_v(T)}\left(\frac{R_a}{R_v}\right) = \frac{\phi p_s(T)}{p - \phi p_s(T)}\left(\frac{R_a}{R_v}\right)$$

with slope

$$\frac{d\omega}{d\phi} = \frac{\omega p}{\phi(p - \phi p_s(T))}.$$

These equations apply at any point in the wind tunnel. The specific humidity is constant until condensation occurs.

Some aspects of this relationship can best be displayed through graphs. Figure 1 shows the relationship between specific humidity and relative humidity in the stilling chamber for given stilling chamber temperature and two values of the stilling chamber pressure. The slope, given by the equation above, is essentially constant for a wide range of relative humidity for a given stilling chamber pressure. Thus, for a given pressure and temperature there is a nearly linear relationship between specific and relative humidity. The specific humidity required for a given relative humidity increases with decreasing stagnation pressure. This is illustrated for fixed stilling chamber conditions in Figure 2.

Application of these relations to the wind tunnel can best be illustrated by plotting relative humidity versus Mach number for given stilling chamber pressure, temperature, and relative humidity (Figure 3). In this figure ϕ is the stilling chamber relative humidity. As the flow accelerates through the tunnel nozzle, the pressure and temperature drop, causing the test section relative humidity to increase. Since the local relative humidity near the model must be greater than 1.0, it can be seen that minimum levels of specific humidity (or equivalently, stilling chamber relative humidity) are required before any condensation could occur near a model. The curves in Figure 3 were created for $p_o = 1$ atm., $T_o = 100\ ^oF$. As the stagnation pressure is lowered, the specific humidity required to produce the same relative humidity increases (Figure 2). However, the mass of water vapor in the tunnel circuit remains essentially constant if the stilling chamber relative humidity is held constant.

McGregor (Ref. 9) was the first to observed that the degree of condensation in the wind tunnel depends only on the amount of water in the tunnel circuit. This can be explained as follows. Assume conditions of fixed stilling chamber temperature and relative humidity and consider a change in the stilling chamber pressure. From Figure 2 it is seen that the specific humidity decreases with increasing stagnation pressure for fixed relative humidity in the stagnation chamber. The mass of air in the tunnel is related to the pressure and temperature by the perfect gas law. Assuming, for the sake of argument, that the entire tunnel is at the stilling chamber conditions, the mass of air in the tunnel changes with change in stilling chamber pressure as

$$m_{a_1}/m_{a_2} = p_{o_1}/p_{o_2}\,.$$

Since $m_a = m_v/\omega$, the change in the mass of air at constant stilling chamber temperature is given by

$$\frac{m_{a_1}}{m_{a_2}} = \frac{m_{v_1}}{m_{v_2}}\frac{\omega_2}{\omega_1} = \frac{p_{o_1}}{p_{o_2}}.$$

From Figure 2, ω_1/ω_2 is roughly equal to $(p_{o1}/p_{o2})^{-1}$ and thus

$$\frac{m_{v_1}}{m_{v_2}} = \frac{p_{o_1}}{p_{o_2}}\frac{\omega_1}{\omega_2} \approx 1.$$

Thus the mass of vapor in the circuit is roughly constant, proving the observation that the degree of condensation depends only on the amount of water in the tunnel and not on the specific humidity.

A number of additional observations of previous investigators are worthy of discussion at this point, especially those relating to the required levels of stilling chamber relative humidity. McGregor (Ref. 9) plotted the amount of water required to produce 15-fold supersaturation in his tunnel. A further increase in water was required before the first condensation could be detected by the eye and even more water was required to obtain suitable photographs. This indicates that the water droplets must grow to a suitable size before the vapor screen visualization is useful. In a subsonic tunnel, this means that the droplets must progress downstream for a finite period of time before they can be usefully observed. Williams, et al. (Ref. 23) observed this effect in a subsonic tunnel with Rayleigh scattering from water droplets. Their tunnel was operated at fixed stilling chamber conditions and the test section Mach number was varied by varying the downstream pressure. During one test the relative humidity reached 1.0 at test section conditions corresponding to M= 0.3 and the first scattering signal was observed at M=0.35, where the relative humidity was 1.126. However, the droplets could not be observed visually until the flow reached conditions such that $M=0.65$, where the relative humidity was 2.70, and the local flow temperature was 26.8 °F below the local dew point temperature. The vapor was observed from a generally forward direction using only tunnel building lights. It was estimated that the flow time from the point in the nozzle where the flow saturated to the nozzle exit where the droplets were seen was 2 ms. This time is characteristic of the temperature gradient existing in their wind tunnel and should not be extrapolated to tunnels or flow fields with other temperature gradients.

The work of McGregor (Ref. 9) also allows the computation of the relative humidity in the stilling chamber and in the test section under optimum viewing conditions. At a tunnel test section Mach number of 0.8, the stilling chamber and test section relative humidities were 0.78 and 4.62, respectively. The stilling chamber temperature was 40 °F below the dew point temperature. At $M=0.85$, the relative humidities were 0.54 and 4.01, respectively and the delta temperature was 32.6 °F. McGregor observed the vapor screen with a light sheet from a mercury vapor lamp. A laser light sheet would have allowed the vapor sheet to have been observed with a lesser amount of water in the tunnel circuit.

The relative humidity in the Langley tunnel under good vapor screen conditions can only be estimated from the work of Erickson and Inenaga (Ref. 6). At a tunnel Mach number of 0.6, assuming that the stagnation pressure is one atmosphere and that the stagnation temperature is 80 °F, the test section relative humidity would be 0.35 and the test section relative humidity would be 0.98. Assuming the same stagnation conditions (they were not stated in the reference), then at a test section Mach number of 1.2, the stilling chamber relative humidity would be 0.07 and the test section relative humidity would be 8.07.

Mie Scattering Patterns

It will be assumed that the water droplets in the wind tunnel flow fields are small enough to be considered to be spherical. Assuming that this is true, Mie scattering patterns from individual scattering spheres were computed using the program MIE01.FOR given in Bohren and Huffman (Ref. 24), as modified by Dr. Robert A. Reed, Sverdrup Technology, Inc., AEDC Group. The code computes k^2S_{11}, where k is the wavenumber and S_{11} is the component of the amplitude scattering matrix (part of the Mueller matrix for the scattering of a single particle) that gives the intensity ratio.

For an incident wave of intensity I_o watts/cm^2, the scattered intensity (watts/sr) is k^2S_{11}. A laser light wavelength of 0.515 μm was in the computations used along with the index of refraction for water of $n = 1.335 + 0\ i$ (Ref. 25 & 26). In this case $k^2 = 148.85$. Scattering intensity distributions were computed for droplet radii from 0.1 μm to 12.5 μm. Some of the results are compared in Figure 4. It can be seen that observing the scattered light at 50° from the forward direction would increase the observed scattered intensity by a factor greater than 10 for all droplet sizes, making it much easier to record the vapor screen images. That was the angle (approximately) that had been used by many previous investigators.

Conclusions

To use the vapor screen flow visualization method for subsonic flow, the relative humidity must be increased to almost 100% in the test section. The relative humidity and the success of the vapor screen method depend on the total amount of water in the tunnel circuit rather than the value of the specific humidity. At low stilling chamber pressures this requires a large value for the specific humidity.

A very concentrated and intense sheet of laser light is mandatory for vapor screen flow visualization. NASA Langley used up to a 15 watt argon ion laser with more than 70% of the energy getting into the sheet in the tunnel. The sheet had a thickness of 0.06 in. at the model and a width spread angle of 4° to 22° (in 7- by 10-Foot or 8-Foot diameter tunnels).

Mie scattering suggests that it would be easier to observe the scattered light from an angle closer to the forward scattering direction than 90°, such as 50°. The scattering intensity is increased by an order of magnitude at this lower angle for droplets of all sizes.

The moisture conditions in a wind tunnel that will allow condensation to occur within model flow features of interest need to be determined. Experience with vaporization in atmospheric flows is useful for predicting those flow features that can be easily observed with the laser vapor screen flow visualization method by controlling the global tunnel humidity level. Local increases in relative humidity with the use of streamtube water streams may be successful for making other features visible by allowing local regions to saturate without fogging the entire tunnel.

References

1. J. F. Campbell, J. R. Chambers, and C. L. Rumsey, "Observation of Airplane Flow Fields by Natural Condensation Effects," AIAA 88-1091, January, 1988.
2. G. E. Erickson and L. W. Rogers, "Effects of Forebody Strakes and Mach Number on Overall Aerodynamic Characteristics of Configuration with 55°Cropped Delta Wing," NASA TP-3253, November 1992.
3. G. E. Erickson, "Wind Tunnel Investigation of the Interaction and Breakdown Characteristics of Slender-Wing Vortices at Subsonic, Transonic, and supersonic Speeds," NASA TP-3114, November 1991.
4. G. E. Erickson, "Wind Tunnel Investigation of Vortex Flows on F/A-18 at Subsonic Through Transonic Speeds," NASA TP-3111, December 1991.
5. G. E. Erickson and D. G. Murri, "Forebody Strakes for High-Angle-of-Attack Vortex Flow Control: Mach Number and Strake Performance Effects," *High-Angle-of-Attack Technology*, Vol. I, J. R. Chambers, W. P. Gilbert, and L. T. Nguyen, Editors, NASA CP-3149, Part 1, 1992, pp. 381-480.
6. G. E. Erickson and A. S. Inenaga, "Fiber-Optic-Based Laser Vapor Screen Flow Visualization System for Aerodynamic Research in Larger Scale Subsonic and transonic Wind Tunnels," NASA TM 4514, January 1994.
7. Anonymous, U. S. Standard Atmosphere, 1976, NOAA -S/T 76-1562, October 1976.
8. H. J. Allen and E. W. Perkins, "A Study of Effects of Viscosity on Flow Over Slender Inclined Bodies of Revolution," NACA Report 1048, 1951.
9. I. McGregor, "The Vapor-Screen Method of Flow Visualization," J. Fluid Mech. Vol. 11, Part 4, December, 1961, pp. 481-511.
10. E. J. Durbin, "Optical Methods Involving Light Scattering for Measuring Size and Concentration of Condensation Particles in Supercooled Hypersonic Flow," NACA TN 2441, August, 1951.
11. G. D. Stein and P. P. Wegener, "Experiments on the Number of Particles Formed by Homogeneous Nucleation in the Vapor Phase," J. Chem. Phys., Vol. 46, pp. 3685-3686, 1967.
12. R. Williams and P. J. Wojtowicz, "A Simple Model for Droplet Size Distribution in Atmospheric Clouds," J. Appl. Meteorol., Vol. 21, 1982, pp. 1042-1044.

13. E. R. Keener, "Flow-Separation Patterns on Symmetric Forebodies," NASA TM 86016, January 1986.

14. G. E. Erickson, L. W. Rogers, J. A. Schreiner, and D. G. Lee, "Subsonic and Transonic Vortex Aerodynamics of a Generic Forebody Strake-Cropped Delta Wing Fighter," AIAA-88-2596, June 1988.

15. G. E. Erickson, L. W. Rogers, J. A. Schreiner, And D, G. Lee, "Further Studies of the Subsonic and Transonic Vortex Flow Aerodynamics of a Close-Coupled Forebody-Slender Wing Fighter," AIAA-88-4369, August 1988.

16. Anonymous, "Tunnel 16T Test Conditions Standard Program," Program number PDP 10001, Arnold Engineering Development Center.

17. J. Hilsenrath, et al., "Tables of Thermodynamic Properties of Gases," NBS Circular 564, November 1, 1955.

18. J. A. Goff and S. Gratch, "Thermodynamic Properties of Moist Air," Heating, Piping & Air Conditioning, June, 1945, pp. 334-348.

19. J. A. Goff and S. Gratch, "Low-Pressure Properties of Water from -160 to 212 F," Heating, Piping & Air Conditioning, February, 1946, pp. 125-136.

20. J. A. Goff, "Standardization of Thermodynamic Properties of Moist Air," Heating, Piping & Air Conditioning, November, 1949, pp. 118-128.

21. G. V. Van Wylen and R. E. Sonntog, <u>Fundamentals of Classical Thermodynamics</u>, 3rd Edition, Wiley, New York, 1986.

22. P. P. Wegener and L. M. Mach, "Condensation in <u>Supersonic</u> and Hypersonic Wind Tunnels," <u>Advances in Applied Mechanics</u>, Vol. V, 1958, pp. 307-445.

23. W. D. Williams, D. W. Sinclair, and L. L. Price, "Laser-Raman/Rayleigh Flow Diagnostic Techniques Applied to Subsonic Flow," AEDC-TR-80-20, October 1980.

24. C. F. Bohren and D. R. Huffman, <u>Absorption and Scattering of Light by Small Particles</u>, Wiley-Interscience, New York, 1983.

25. G. M. Hale and M. R. Querry, "Optical Constants of Water in the 200-nm to 200-μm wavelength Region," Appl. Optics, Vol. 12, No. 3, 1973, pp. 555-563.

26. S. G. Warren, "Optical Constants of Ice from the Ultraviolet to the Microwave," Appl. Optics, Vol. 23, No. 2, 1984, pp. 1206-1223.

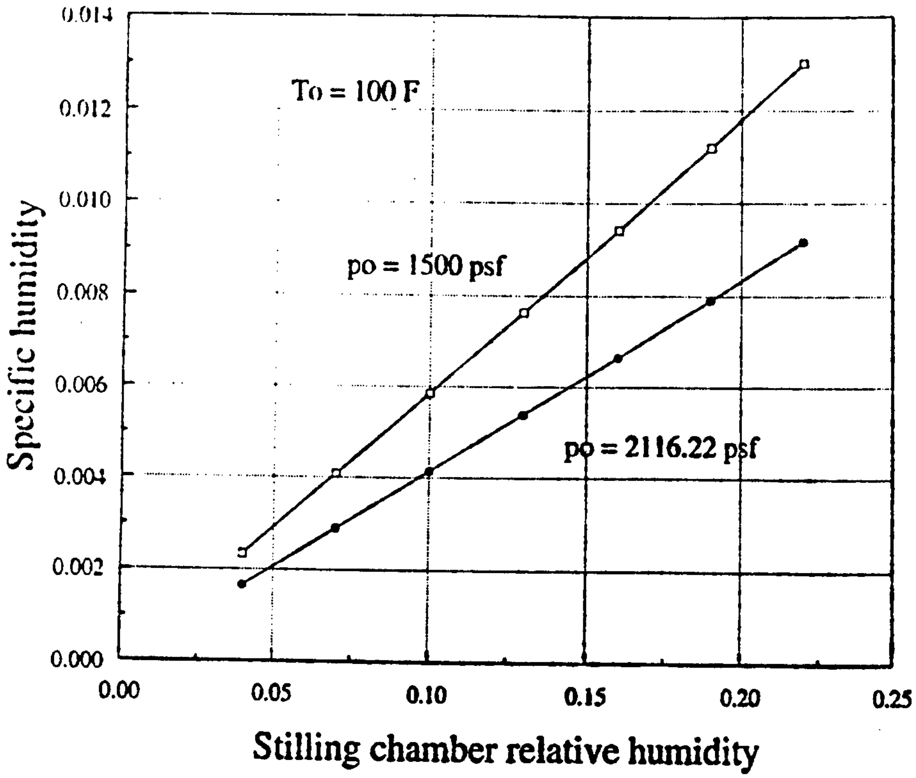


Figure 1. Specific humidity v. stilling chamber relative humidity.

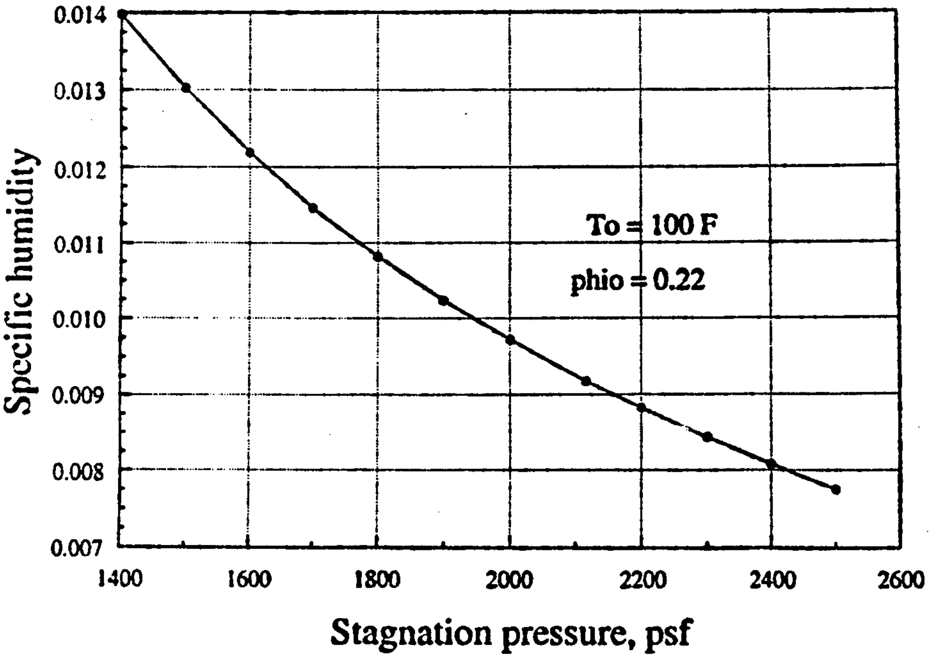


Figure 2 Specific humidity vs. stagnation pressure.

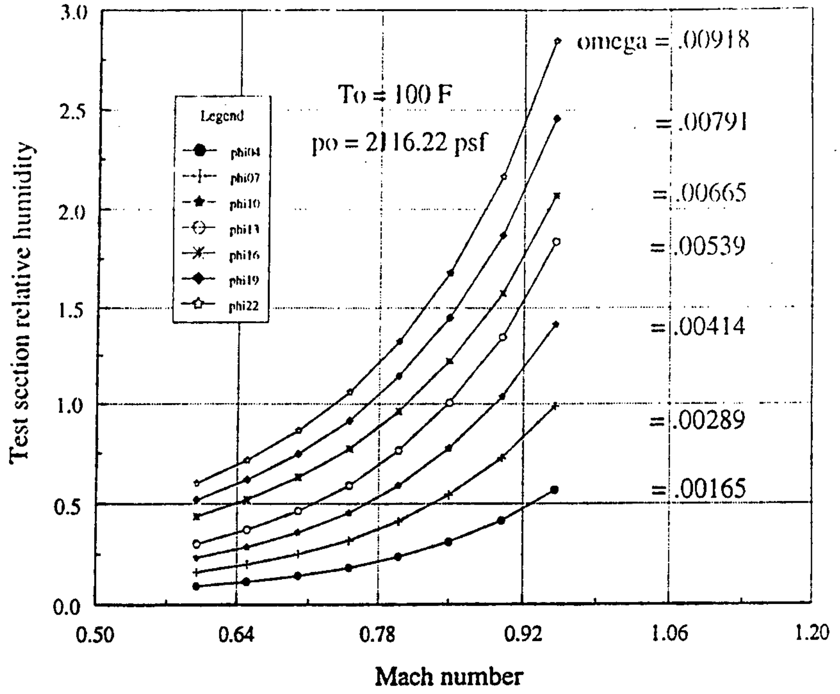


Figure 3. Test section relative humidity for constant stagnation relative humidity

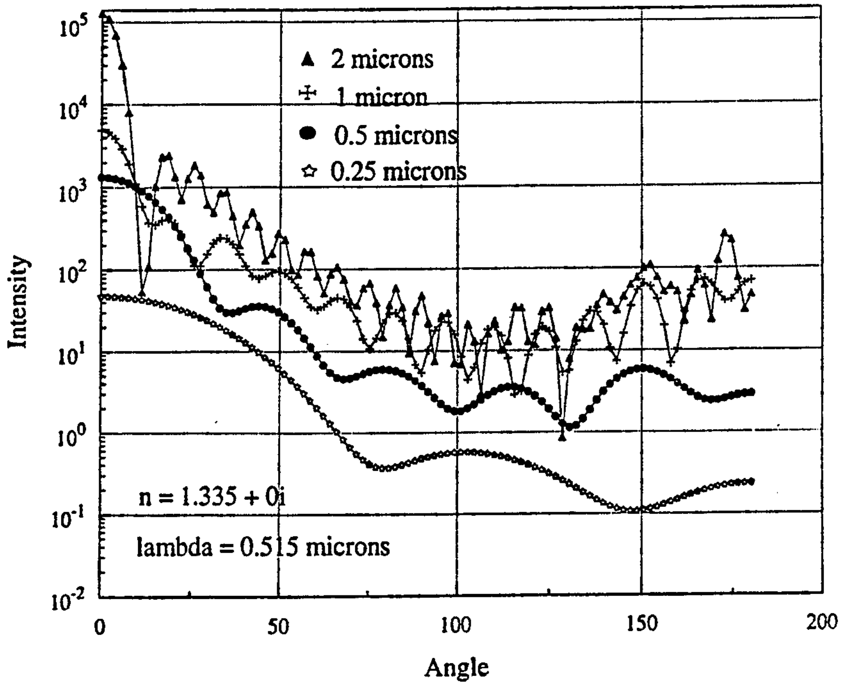


Figure 4. Mie scattering from spherical water droplets.

Randolph Peterson's report was unavailable at the time of publication.

A PROCESS FOR SETTING UP
COMPUTATION OF SWIRLING FLOWS
IN THE AEDC H-3 ARC HEATER

Robert L. Roach
Assistant Professor
Mechanical and Aerospace Engineering

University of Tennessee Space Institute
B.H. Goethert Parkway
Tullahoma, TN 37388-8897

Final Report for:
Summer Faculty Research Program
Arnold Engineering & Development Center

Sponsored by:
Air Force Office of Scientific Research
Boling Air Force Base, DC
and
Arnold Engineering & Development Center

September 1996

A PROCESS FOR SETTING UP COMPUTATION OF SWIRLING FLOWS IN THE AEDC H-3 ARC HEATER

Robert L. Roach
Assistant Professor
Mechanical and Aerospace Engineering

Abstract

The work reported here is concerned with setting up a computational method for determining the flowfield in arc heater type wind tunnels similar to the AEDC H-3 Arc Heater. The aim is to examine the relationship between the swirl flow, induced by injection ports in the intersegment slots and the tendency of the arc to impinge upon the walls. To do this, it is necessary to compute the flowfield in the arc heater geometry. Since the method uses the full Navier-Stokes equations, grid generation and initial condition generation programs were devised for future parametric studies to be performed. A checkout case was run for one set of stagnation conditions to ensure that the process can now be used for a series of parametric evaluations.

A PROCESS FOR SETTING UP COMPUTATION OF SWIRLING FLOWS IN THE AEDC H-3 ARC HEATER

Robert L. Roach

Introduction

A large, state-of-the-art segmented electric arc heater, designated H-3, has been developed at the USAF AEDC. This hypersonic wind tunnel facility is capable of very high stagnation pressures and has operated at enthalpies as high as 3,000 BTU/lbm at total pressures as high as 107 atm. The H-3 heater is a 50% geometric upgrade from the existing H-1 heater and is designed to operate at 80 MW. It is also designed to operate with pressures as high as 150 atm while the pressure vessel is capable of 200 atm. A schematic of the H-3 heater is shown in Fig. 1.

The H-3 heater consists of a large number of water-cooled copper segments which use mica spacers between the segments providing electrical isolation of each segment. The mica spacers have a larger inner diameter than the copper segments producing a slot between the segments. The spacers also have channels machined into them to produce air injection ports in the slots.

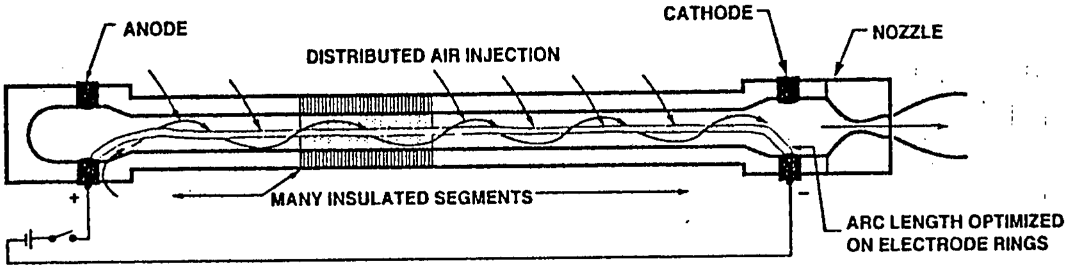


Fig. 1. Schematic of the H-3 Arc Heater

The main chamber is a straight bore three inches in diameter and approximately 88 inches long with seven transition segments and an electrode 4.5 inches in diameter and 4.5 inches long at each end of the bore. The total number of segments is 216.

While electric arc heaters have been proven to be the only reliable method to heat air to high temperatures for a duration of several minutes, the unprecedented size and capabilities of the H-3 heater bring with it some new challenges. In this case, there is evidence that the electric arc, supposedly confined to the heater centerline by a pressure gradient induced by swirling the flow, in fact does not stay confined and occasionally strikes the side walls causing some damage. Since the arc confinement is done by fluid mechanical means, the arc impingement problem is thought to be a defect in the flow field. Hence it was undertaken to study the flow field in the H-3 heater. In particular the aim is to examine the relationship between the swirl flow, induced by injection ports in the intersegment slots and the tendency of the arc to impinge upon the walls. It is thought that regions of recirculation in the vicinity of the injection ports may be inducing such impingements.

To determine the nature of the flow in the vicinity of the injection ports and slot region, it is necessary to compute the flowfield in the main chamber and between two segments including the slot region and injection ports. The flow is compressible, consists of swirling flow with nonzero vorticity, and viscpus interactions with the solid surfaces. Such a computation requires the solution of the full Navier-Stokes equations.

Now the flow in the chamber includes the electric arc, however, since the main interest is flow anomalies in the region near the slot openings well away from the centerline, it was first desired to set up a computational procedure for the cold flow problem. That is, flow in the H-3 heater without the electric arc. This represents a considerable simplification and allows for an analysis with only the compressible Navier-Stokes equations without the MHD equations being added.

Thus the goal of this project was to set up a procedure to generate grids for segments in the H-3 electric arc heater, formulate initial conditions based on input from the PC-based SWIRLARC code, and to set up a CFD code for a parametric study. In the remainder of this report, the Navier-Stokes solver, grid generation programs and initial condition programs are described. The flowchart in Chart 1 shows the overall process. Each of the steps is described in more detail below.

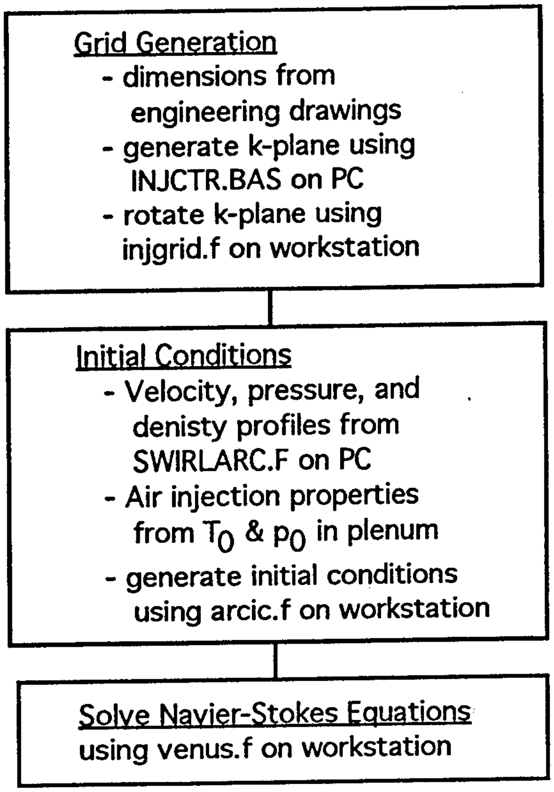


Chart 1. Process of find the cold flow solution in the AEDC H-3 Electric Arc Heater.

The Solver

A Navier-Stokes solver, VENUS [1] was used for the flow computations. VENUS (Vectorized Euler-Navier-Stokes Unsteady Solver) is an Approximate Factorization (AF) algorithm which uses 4th-order compact differencing for the evaluation of all explicit derivatives and a 2nd-order evaluation of the implicit derivatives. When the steady state is reached, the implicit side is zero leaving the solution spatially 4th-order accurate. For unsteady problems, 2nd-order time accuracy is used and inner iterations can be used to drive the implicit side to zero, keeping 4th order spatial accuracy in the unsteady solution.

Grid - injctr.bas, injgrid.f

The grid was generated using a specially-written program which runs on a PC and which creates the node points for a streamwise cross-section (k-plane). The final grid is created by rotating this plane around the centerline. The procedure for creating a grid for the region between two segments of a typical arc-heater is given below. The computational domain spans, in the axial direction, over parts of two copper segments so that the slot between the segments is roughly in the center. Thus, when using the engineering drawings of a single segment, the dimensions from both edges are combined to give the required dimensions for the grid generation program. These dimensions are shown in Fig. 2.

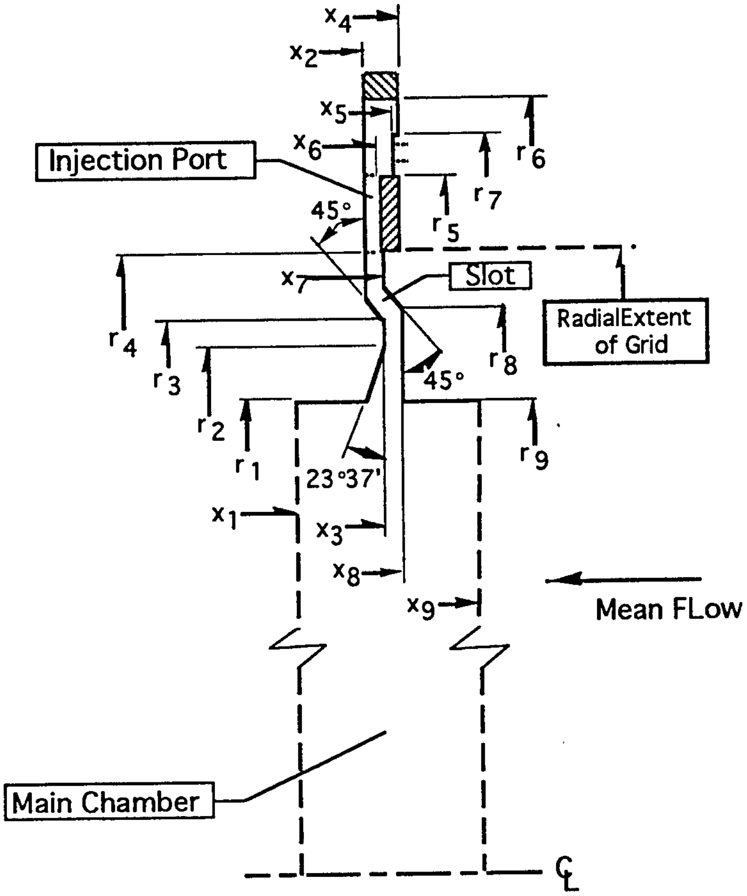


Fig. 2 Flow Geometry on a typical azimuthal plane.

Step 1 - Dimensions from engineering drawings are used to find the required radii and axial positions of the surfaces shown in Fig. 2.

Step 2 - The total number of grid points in each of the coordinate directions must be chosen along with the three index locations of the main surfaces. These main surfaces are the upstream and downstream faces of the injection slot and the inner radius of the chamber. Since the injection slot faces are perpendicular to the axial direction, they are located at 'i' stations and are given the names 'IS1' and 'IS2'. The Chamber inner radius is perpendicular to the j-direction and hence has the index designation 'JS1'. Other index locations include the i- and k-index locations of the injection slots on the outer radius of the grid. For the present geometry, these were designated as 'i_{s1}', 'i_{s2}, 'k_{s1}', 'k_{s2}', 'k_{s3}', and 'k_{s4}'. With these chosen the PC-based program 'injctr.bas' is run to create the grid cross section corresponding to a k-plane. All k-planes are identical and the 3D grid generation program merely takes the cross section data file and creates the full grid. The output file from the BASIC program is uploaded to a workstation where the 3D grid generation program, initial condition program, and flow code program is located.

Step 3 - Using the k-plane file uploaded from the PC-based, BASIC program, the program 'injgrid.f' is run to create the full 3D grid. The program creates different circumferential stations by rotating the uploaded k-plane about the centerline axis. The number of circumferential stations depends on the accuracy required but is usually a number similar to the number of stations in the other two directions. Here, the number of stations chosen was to try to match the openings of the injection ports. Since these spanned a certain number of degress, the number of circumferential stations was chosen to fit an even number of stations over that span of arc length. Views of the computational grid for a segment of the H-3 Arc Heater are shown in Fig. 3-4.

Fig. 3 shows the overall computational domain consisting of one quarter of a segment region. The main chamber and slot between the segments are shown. Some of the important indices are also indicated along with the orientation of the Cartesian and cylindrical coordinates. The index jb1 indicates the inner radius of a copper segment and represents the radial location of the inner chamber radius.

Detail of the air injection ports is shown in Fig. 4. There are two ports in the computational domain, as indicated by the k-indices with 's' in the subscript. A closeup of one of the ports shows the i-indices of the slots. Also shown are the i-indices of the edges of the slot.

Initial Conditions - arcjetic.f

Initial Conditions in the Main Chamber

The initial conditions were created in the FORTRAN program 'arcjetic.f' using input from the SWIRLARC [2] PC-based computer program. SWIRLARC computes an axisymmetric flow in an arc heater. Radial profiles of the dependent variables are provided at various axial stations as part of the output of the computation. These profiles are thus available to be used as initial conditions for a genuinely 3D computation. The one change necessary, to use the SWIRLARC profiles, is to zero out the radial velocity component at the inflow. To mimic the gas injection from the side walls, SWIRLARC assumes the injected gas velocity to be uniform along the length of

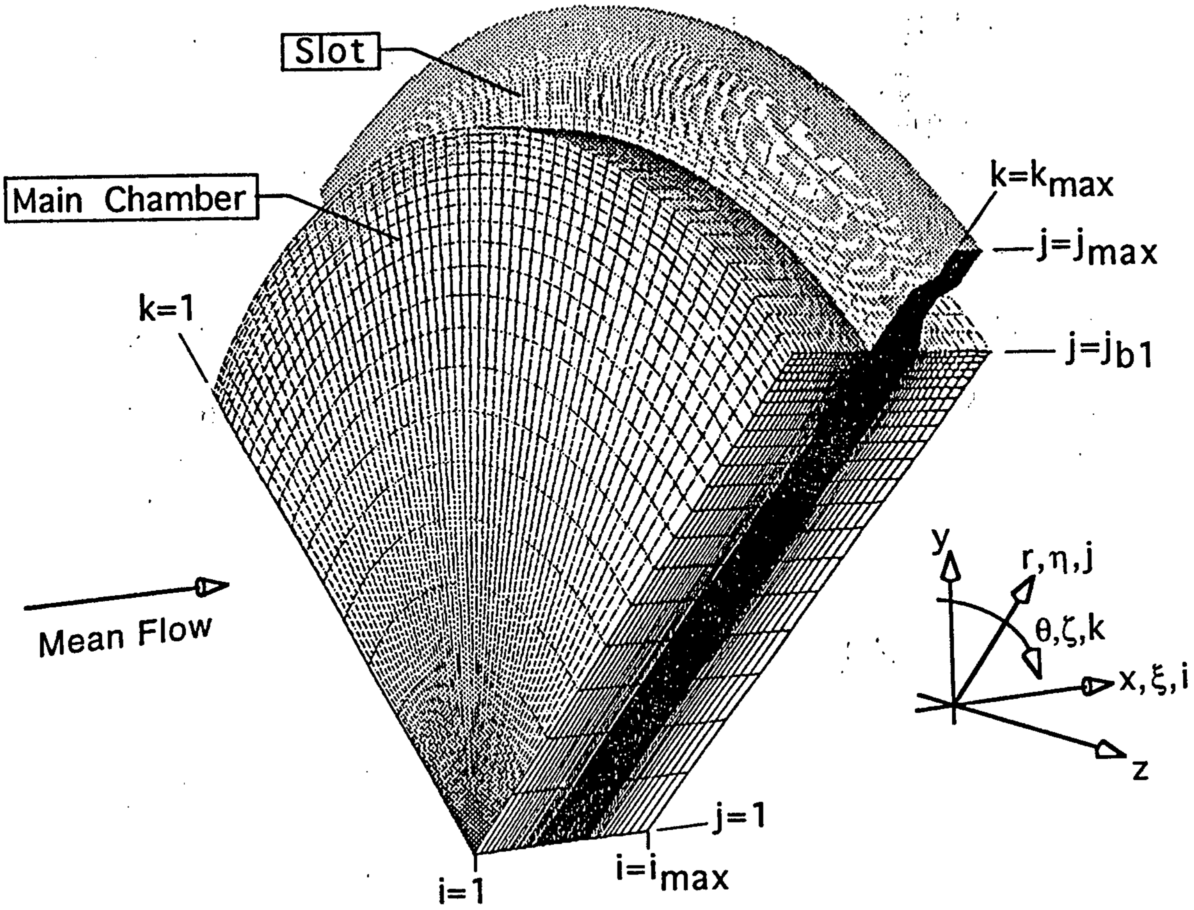


Fig. 3 Overall computational domain and grid for flow in one-quarter of region between segments of the AEDC H-3 Arc Heater.

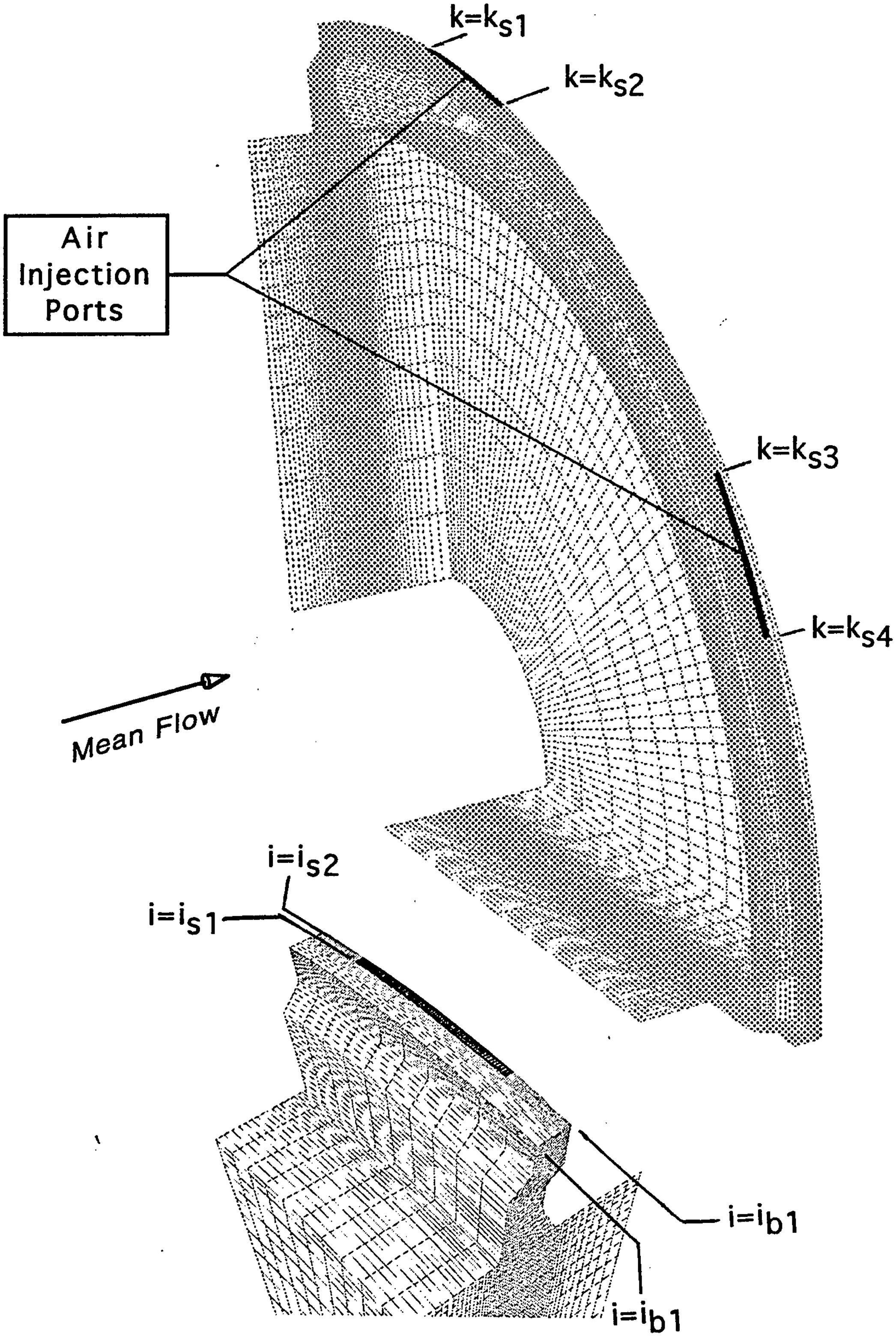


Fig. 4 Detail of grid near injection ports of the computation domain.

the walls and not at discrete locations corresponding to the slots. As a consequence, the normal velocity on the real solid surfaces was not assumed to be zero in the SWIRLARC computation. In order to properly compute the effects of the slots, this is changed to match the real conditions in the present computation. In this case, the radial velocity is quite small and is safely set to zero for the inflow.

The solution at the x = 2.0 meters station was used for the inflow profiles. In addition, the profiles at the station x = 2.41 meters were also available for the downstream initial conditions. Again, the radial velocity was set to zero to account for the no slip walls. Since the SWIRLARC program uses a different grid for its computations, these profiles must be interpolated. This is done using parametric cubic splines.

Once the inflow profiles have been set, the same profiles are assumed in the remainder of the chamber. Hence the profiles of the dependent variables at every station downstream is set equal to those at the inflow.

Initial Conditions in the Slot

Injection ports exist along the outer radius of the slot. These ports are rectangular and their locations are specified by the grid indices (IS1,KS1)-(IS2,KS2) and so forth for as many ports as are in the domain. In the results reported, there were two ports.

Initial conditions in the slot depend on the flow from the injection ports. For the present, these ports were assumed to be aligned tangentially to the slot outer radius. Gas flow through the injection ports was assumed to be an isentropic expansion from the difference between the stagnation pressure of the injected gas and the stagnation pressure from the main chamber. As the flow is subsonic, the mass flow rate must also be specified. Finally, parabolic profiles are assumed for the tangential velocity in two directions across the slot. With these assumptions a maximum velocity is found an placed into the program. Computation of maximum velocity should be automated since it can be found, but this has not yet been done.

Initial conditions from the injection ports are found by resolving the tangentially injected gas into the y- and z-velocity components. These are multiplied by an average gas density. The axial component of the velocity (x-direction here) is assumed to be zero. The remaining variables in the injection port faces are set equal to those along the remainder of the slot outer radius. The solid walls in the slot are set with no slip conditions (all velocities zero) and density and energy set equal to their stagnation values.

Conditions on the two circumferential planes in the slot are also set equal to the stagnation conditions with zero velocities. This condition, unfortunately may be responsible for the slow convergence of the flow in the slot region. The values of the velocities at convergence will be quite far from zero. Since the initial conditions of the interior depend on the values of the slot boundaries, any conditions which are far from the final conditions will slow the convergence. Hence, it is recommended that these conditions be changed to something more resembling the final solution. Perhaps some fraction of the injection velocity distributed over the region would be reasonable.

The final boundary of the slot region is the opening to the main chamber along the slot inner radius which abuts the main chamber. The conditions on this boundary have already been given by the initial conditions in the main chamber.

With all boundaries in the slot specified, the conditions in the volume are filled in using a three-dimensional transfinite Interpolation [Gordon and Hall]. This technique produces smooth values of the dependent variables in the volume using information from all the boundaries.

Boundary Conditions

Inflow Boundaries

Inflow boundaries are located at the entrance to the main chamber and the injection ports in the circumferential slot. In both cases the inflow is subsonic meaning that there is one characteristic leaving from the interior. This requires the adjustment of one of the flow variables. In this case, the total energy was allowed to relax using the boundary condition formulation of Rudy and Strikwerda [Rudy & Strikwerda] which is based a type of Riemann invariant argument. In their formulation the sound speed on the inflow boundary is computed using the values of the dependent variables at the old time level. A new inflow pressure is then computed based on the difference between the normal velocity between the inflow and the next station:

$$P_1 = P_2 - \rho_1 a_1 (u_2 - u_1)$$

With this new pressure, a new inflow total energy is computed.

Outflow Boundary

The only outflow boundary is in the main chamber. It lies geometrically close to the inflow boundary also in the main chamber. Since the outflow boundary is subsonic, one characteristic crosses the boundary from outside the domain and hence one quantity should either be held fixed or some invariant quantity should be held. Much has been written in the literature on these boundaries given such considerations. Usually it is desired to place the outflow boundary far from the main parts of the flow so that any errors propagating in from the quantities that are held fixed don't disrupt the solution. For internal flows, this is very difficult to do without putting a sonic throat downstream. For these reasons, and for a more stable computation from the initial conditions, a simple first-order extrapolation of all variables from the station upstream of the outflow boundary was used. This has the advantage of being a rather dissipative condition which then damps any disturbances from the outflow boundary. It has the disadvantage of not allowing the imposition of a lower downstream pressure. As the solution began to settle, the outflow boundary was changed to a more proper imposition of the downstream pressure. In this case, the pressure at the outflow was computed using the non-reflecting condition of Rudy & Strikwerda [Rudy & Strikwerda] which depends on the unsteady conditions happening at the outflow and the downstream imposed pressure. Thus the outflow pressure was computed using:

$$p_{imax}^{n+1} = \frac{1}{1+\alpha\Delta t}\left[p_{imax}^{n} + \alpha\Delta t p_{out} + \rho_{imax}^{n} a_{imax}^{n}\left(u_{imax}^{n+1} - u_{imax}^{n}\right)\right]$$

where a_{imax}^{n} is the outflow boundary sound speed at the nth (old) time level, and α is a relaxation parameter found to be optimal in the range of [.3,.5]. Note that the axial velocity must be found before this pressure can be computed. The axial velocity and the other velocity components, along with the density are all extrapolated. With the pressure found from the above expression, the total energy is then computed from the other variables.

Solid Wall Boundaries

No-slip, adiabatic solid wall boundary conditions were used. Thus the velocities were held fixed at zero. The zero-gradient density and pressure were computed to second order by expressions like:

$$\rho_1 = \frac{(\rho_2 \Delta x_3^2 - \rho_3 \Delta x_2^2)}{\Delta x_3^2 - \Delta x_2^2}$$

The total energy is then computed from the new pressure.

Centerline Boundary

The centerline boundary is a geometric axis of symmetry so that only the tangential velocity component is nonzero. Thus the two normal components were held fixed at zero. The tangential velocity and the thermodynamic variables are all found by an averaging procedure. In this case, the value at the centerline for each of the k-planes is found using the same zero gradient expression as is used at solid walls. Since all these values are coincident for any given axial station, these values are averaged. The average value is then given to all of the coincident points. This process is done for the axial velocity, the density, and the pressure. The total energy is then computed from these variables.

Periodic Boundaries

The periodic boundaries are the k = 1 and k = kmax planes. Since advantage has been taken of the symmetry of the problem, these planes are involved in the imposition of periodic conditions on the flow. These two planes are located at one station beyond the quarter-plane region as shown in Fig. 5. Since the conditions at planes k = 2 and k = kmax-1 are supposed to be the same, then it makes sense that the conditions for plane k = 1 should match those of the plane just before k = kmax-1. Similarly, the conditions on the plane k =kmax should match those of plane k = 3. Thus, after the interior points are computed, the thermodynamic variables and axial velocities on the k = 1 and k = kmax planes are set equal to the values on their respective matching planes. The velocity components perpendicular to the axial direction, though, must be rotated in each case by 90 degrees due to the different orientation of the planes.

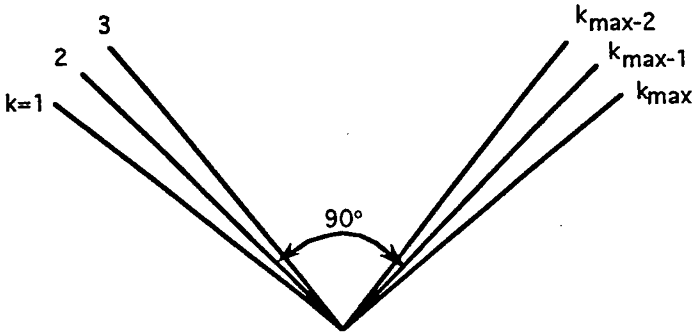


Fig. 5. Orientation of the planes in the k-direction showing the periodic planes, k =2 and k= k_{max-1}.

Results

The entire process of setting up the grid, initial conditions and solving the flow was carried for a typical set of conditions for the H-3 arc heater. A stagnation chamber pressure of 100 Atm was used with a stagnation pressure of 125 Atm for the copper segment plenum. Similar conditions were used to run the SWIRLARC program to generate initial profiles for a segment at x = 2.0 meters and x = 2.41 meters downstream from the anode. The first of these was used to generate profiles for the inflow and the second was used to provide a back pressure for the downstream boundary condition.

For the geometry and initial conditions described, the VENUS code was run to convergence which was defined to be a reduction in the maximum change of the density by five orders of magnitude from the initial maximum change. The results for the grid and given conditions are shown in Fig. 6-8.

Fig. 6 shows the velocity vectors on one i-station near the outer radius of the slot and includes some of the radii in the main chamber. The high velocity from the injection port is evident at the top and center of the figure. This high velocity diminishes rapidly as the flow spreads out into the upper portion of the slot and as it fills in the velocity profiles further around the circumference of the slot region. The effects of the injection port on the gas flow into the main chamber is evident from the different orientation of the vectors near the slot opening at different circumferential locations. Also evident is a seeming divergence of vectors in the upper part of the chamber near the slot opening. This is further evidence of a swirling flow in the slot opening near the main chamber. Further evidence of this is seen in Fig. 7.

Fig. 7 shows the velocity vectors at the central k-station and is typical of the velocity vectors at any k-station due to the near axisymmetry. It can be seen that there is a recirculation region inside the slot, near the opening to the main chamber. In addition there are some rather large velocity magnitudes near the downstream corner. These large magnitudes are thought to be larger than actual magnitudes and are likely a result of the low grid resolution. Since the purpose of this work was merely to set up a procedure to perform a parametric study, this solution should only be used as a guide for future use of the procedure. In this case, it is clear that closer grid clustering in the radial direction near the slot opening and along the main chamber walls should be used than was used in this example.

The swirling flow in the slot is most easily seen in the particle traces shown in Fig. 8. Particles were released by the tracing program near the recirculation region and were captured by the vortical flow, producing the paths shown. Rotation of the flow is clearly seen as the particle paths twist around each other for the entire k-direction.

Conclusions

A process has been set up for solving the Navier-Stokes equations for the cold flow in the vicinity of one segment of the AEDC H-3 Electric Arc Heater facility. The process includes creating the grid and initial conditions. The grid generation uses the original engineering drawings to get relevant dimensions of the arc heater segments and a two step process to create the three-dimensional grid. This part could be made more automated to allow the user to more easily cluster points near walls or other interesting regions. In addition, the PC-based INJCTR.BAS code originally determined the dimensions in the reverse direction to the flow and was corrected by a simple subtraction from the maximum axial coordinate. However, this should be corrected in a more automated environment. In spite of these limitations, the PC-based code does allow for much flexibility in creating the grids.

The initial conditions are generated from profiles in the main chamber provided by the SWIRLARC.FOR program. One problem with these profiles is that they are based on a porous wall to approximate the flow entering the main chamber through the slot region. As a consequence, the nonzero velocity components at the wall must be zeroed out as well as the entire radial velocity component. As a consequence, the inflow conditions will not immediately satisfy the continuity equation. This is later corrected as the flow develops since the upstream density can be adjusted instead of the total energy, but the effect is very small. Other possible problems with the SWIRLARC program include an ambiguity in the mass flux through the main chamber as opposed to the mass flux through the injection ports. In a more automated environment, this ambiguity should be removed.

Finally there is the solution itself. The VENUS code used here is a reliable 4th-order, flow solver and its solution, for the conditions given here, indicates some inadequacies in the present concept of the process which can be easily overcome. These include the fact that the flow leaving the main chamber had a part which was re-entrant meaning that the flow was reversed at the exit plane. Since the flows in nearby segments should be similar, this would imply an out flow on the inflow plane, which was not provided. Hence, either the real flow has many such reverse flow regions, or the problem is not formulated well. In order to resolve this problem, it would be a simple matter to set up a series of segments (three, for example) with associated injection ports. The center segment would then serve as the model of what happens in a real cold flow case.

Another possible shortcoming was that a grid independence study was not run to ensure that the solution generated for this sample case was independent of the grid. For this to be verified, the same case would have to be run again with a refined grid and the solution compared with the present one. It is recommended, before too many conclusions are reached from the results given that such a study be undertaken.

Finally, in spite of the shortcomings, subsequent cases with differing conditions, should be fairly easily implemented in considerable less time than it took to create this process. For a new set of conditions, the time from problem specification to grid creation and initial condition generation should be no more than a couple of hours. The Navier-Stokes code could then be executed with very little subsequent preparation. Thus the process now represents a considerable improvement over the previous inability to have knowledge of the flow in the AEDC H-3 Arc Heater.

References

1. Roach, R.L. and J. Jenkins, "Improvements to the Numerical Simulations of Flows over Delta Wings," AIAA 95-0761, AIAA 33rd Aerospace Sciences Meeting, Reno NV, Jan. 1995.

2. Shaeffer, J.F., "SWIRLARC: A Model for Swirling, Turbulent, Radiant Arc Heater Flow Fields," AIAA 78-68, AIAA Aerospcae Sciences Meeting, Jan 1978.

3. Gordon, W.J. and C.A. Hall, "Construction of Curvilinear Coordinate Systems and Applications to Mesh Generation," J Numerische Math., Vol. 1, 1973, pp. 461-477.

4. Rudy, D.H. and J.C. Strikwerda, "Boundary Conditions for Subsonic Compressible Navier-Stokes Calculations," Computers and Fluids, Vol. 9, 19981, pp.327-338.

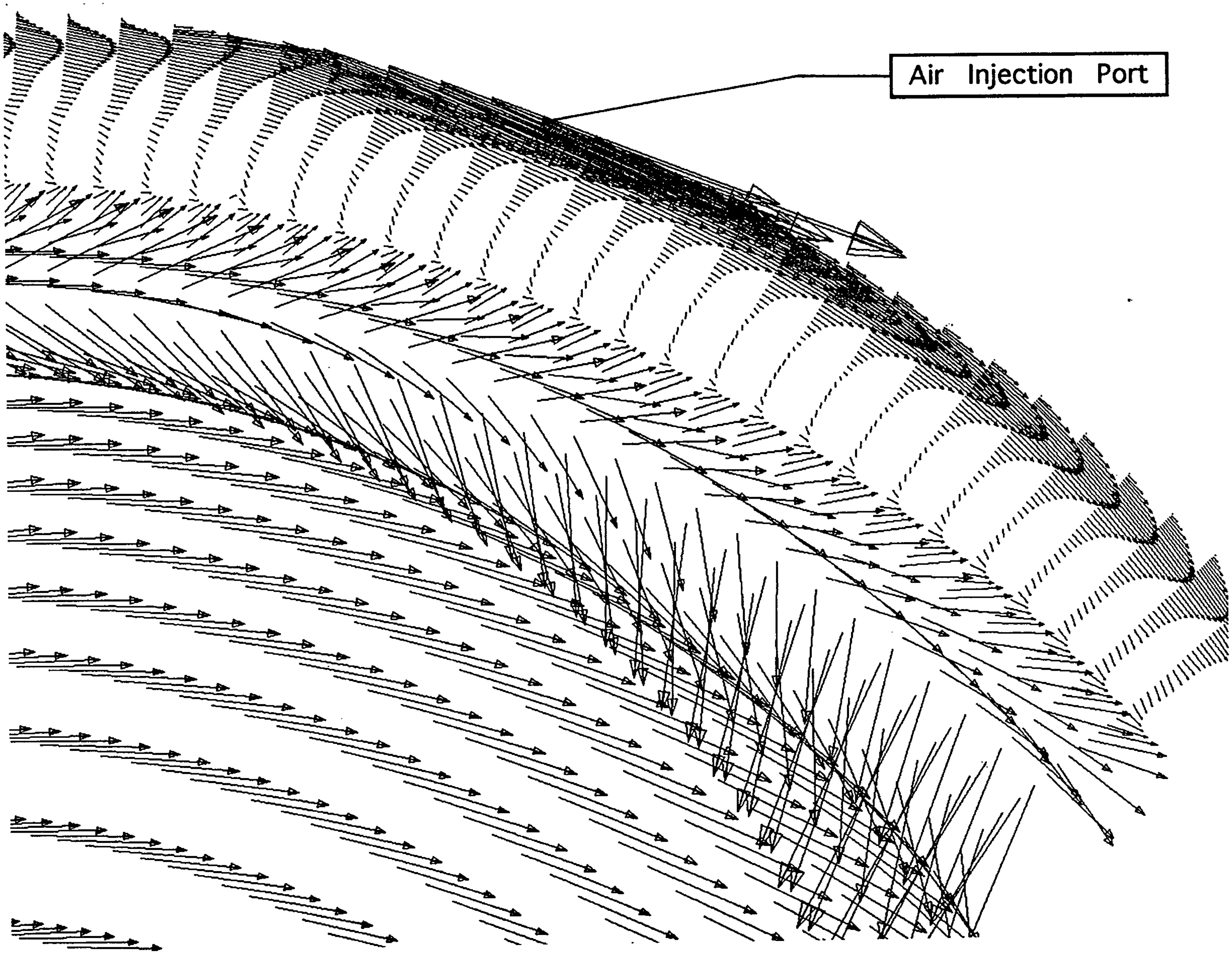


Fig. 6. Velocity vectors at a typical i-station in the AEDC H-3 Electric Arc Heater at x = 2.1 meters. p_0 = 100 atm, $p_{0_{plenum}}$ = 125 atm.

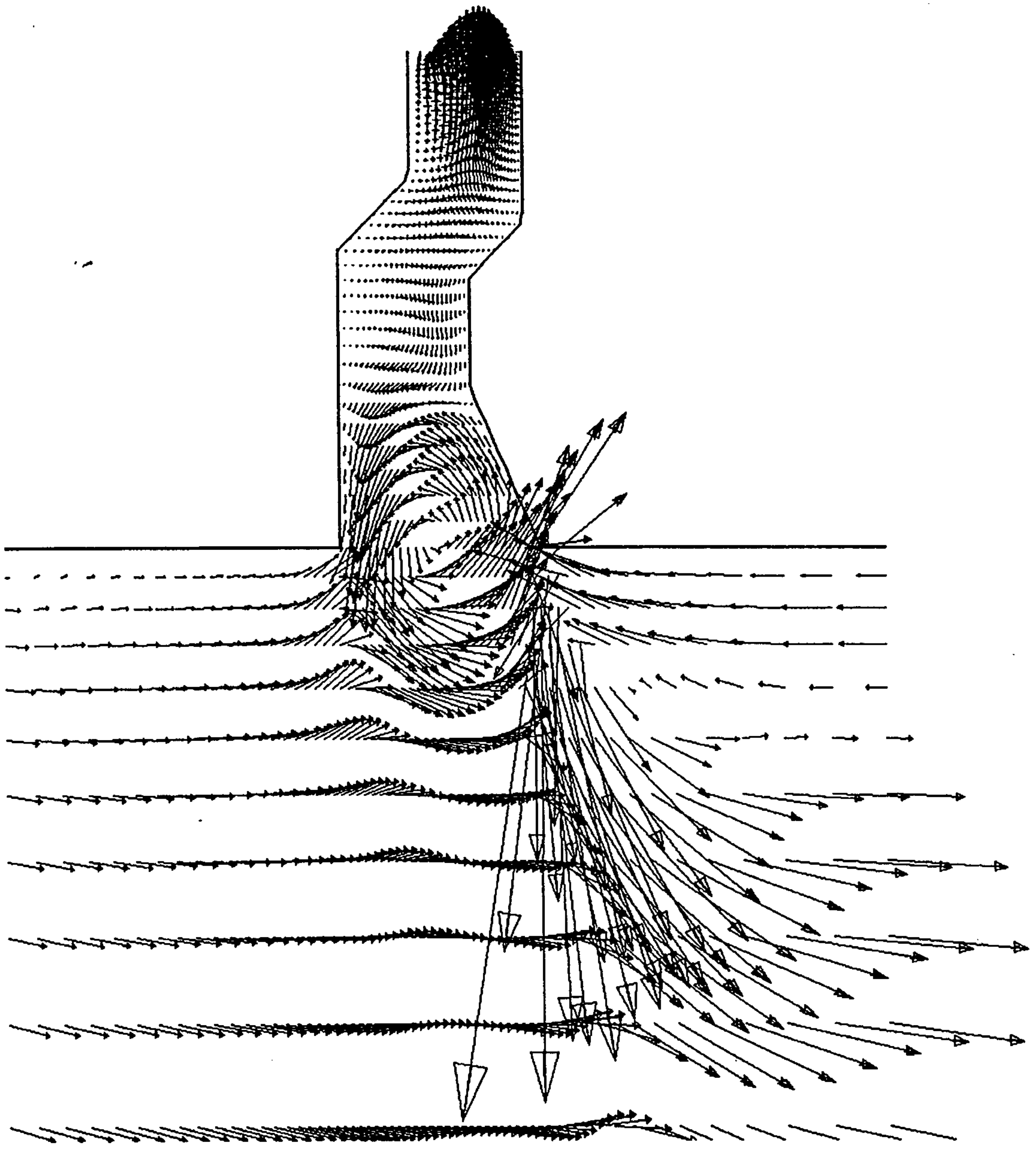


Fig. 7. Velocity vectors at a typical k-station in the AEDC H-3 Electric Arc Heater. $p_0 = 100$ atm, $p_{0_{plenum}} = 125$ atm.

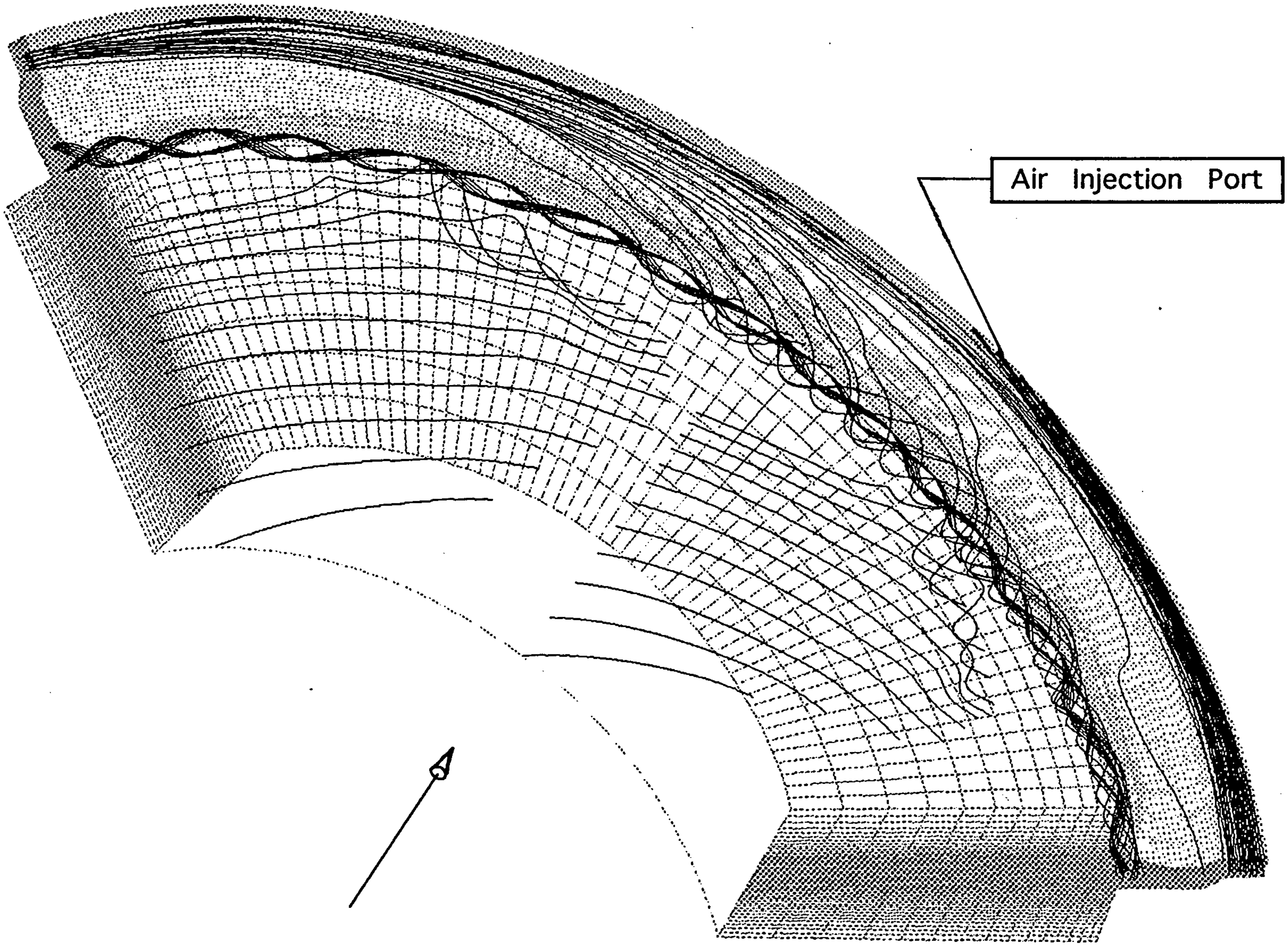


Fig. 8. Particle traces in the AEDC H-3 Electric Arc Heater
$p_0 = 100$ atm, $p_{0_{plenum}} = 125$ atm.

A Massively Parallel Ab Initio Molecular Dynamics Simulation of Polymers and Molten Salts

Ryoichi Kawai
Assistant Professor
Department of Physics

University of Alabama at Birmingham
1300 University Blvd.
Birmingham, AL 35294

Final Report for:
Summer Faculty Research Program
US Air Force Academy

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Wright Laboratory, Wright Patterson Air Force Base

August 1996

A Massively Parallel Ab Initio Molecular Dynamics Simulation of Polymers and Molten Salts

Ryoichi Kawai
Department of Physics
University of Alabama at Birmingham
1300 University Blvd.
Birmingham, AL 35294

Abstract

A new *ab initio* molecular dynamics (AIMD) simulation code based on planewave-pseudopotential local density functional method has been developed and implemented on a massively parallel computer using a high performance fortran. The new code uses significantly smaller amount of memory and disk space than the previous version. With the new code, it is possible to simulate geometric, dynamical, and electronic properties of large molecules and polymers containing more than 200 atoms from first principle. In addition, it allows both periodic and quasifree boundary conditions. Extended system such as polymers can be simulated with the periodic boundary condition whereas finite-size systems including charged systems can be calculated with the quasifree boundary condition. Various benchmark tests demonstrate the high degree of accuracy and efficiency of parallel processing. This code is applied to the simulation of semiconducting polymers with a large repeat cell and molten salts based on aluminum chlorides.

CONTENTS

1. INTRODUCTION

A major step toward the ability to calculate the properties of materials from first principles has been made possible by the appearance of very fast vector computers. Recently, we and others have developed algorithms that so efficiently implement one of the most successful methods of electronic calculations–the local density approximation–that they may be used to directly calculate the forces between atoms during a simulation of the atomic motion (*ab initio* molecular dynamics, AIMD) in systems of fairly large size ($<$ 200 atoms). This represents a significant advance, since now systems displaying very complex chemistry can be interpreted with a first principle theory. We believe that with achievable improvements, these methods can be applied to study technologically interesting problems.

In this report we demonstrate that the approach that we have been developing is remarkable in that it can be very efficiently implemented on massively parallel computers. Benchmark tests show that significant improvements can be obtained from existing similar codes. With these improvements in efficiency, the methods will be applied to electronic, geometric, and dynamical properties of conducting polymers with a large repeat unit and other complex chemical system such as molten salts.

2. THEORETICAL METHODS

2.1. **Local Density Functional Method with Planewave Basis.** The objective of this program is to develop a method of "first principles" simulation that will describe the motion of large numbers of strongly interacting atoms (N $\approx$ 500) with sufficient accuracy that reliable predictions of the properties of real systems of general composition may be made. We assume that the motion of the atomic centers may be well approximated by the classical equations of motion. However, the forces between atoms must be calculated quantum mechanically. For systems of this size the most reliable and efficient approach is based on the local density approximation (LDA) of Kohn and Sham [1]. In this approach, the total energy of electrons interacting with the nuclei is described as a functional of electron density and coordinates of the nuclei. Using the variational principle and the orbital orthogonality constraints, electronic orbital wavefunctions are given by the solutions of the generalized eigenvalue equations:

$$\mathcal{H} \mid \psi_i \rangle = \sum_j \lambda_{ij} \mid \psi_j \rangle \tag{1}$$

where the Hamiltonian is given by

$$\mathcal{H} = \left(\frac{p^2}{2m} + V_{ext} + V_H + V_{xc} \right) \tag{2}$$

and λ_{ij} is a Lagrange's multiplier which maintains the orbital orthonormality constraints, $\langle \psi_i \mid \psi_j \rangle = \delta_{i,j}$. Electron-electron interaction is included in the Hartree potential, V_H, and the exchange correlation potential, V_{xc}. Since V_H and V_{xc} are functionals of the electron density, Eq.(1) must be solved self-consistently.

In order to further implement the theory, it is necessary to expand ψ in a basis set. Usually this basis is chosen with the particular physics of the problem in mind. For example, a natural choice of basis functions would be planewaves for metallic systems, whereas an atomic like basis, e.g., Gaussian functions, would be preferred for covalently bound first row atoms, C, N, O. However, with the advent of new computer architectures, it is important to design the method of solution to match the structure of the computer. Using a planewave basis set,

$$\psi(\mathbf{r}) = \sum_{\mathbf{k}} C_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} \tag{3}$$

the Kohn-Sham equations, Eq.(1), may be implemented in a way that is well suited to vector and, even more efficiently, to SIMD parallel architecture computers. The retention of the planewave basis requires the use of very large numbers of basis functions. However, the computational cost of this is somewhat offset by the high parallelism and efficient vectorization of the algorithm.

In order to solve Eq.(1) using a planewave basis, it is necessary to remove the highly localized core orbitals and the rapid oscillations in the valence wavefunctions due to orthogonality to core orbitals. This can be done without much loss of accuracy using pseudopotentials[2] which replace V_{ext} in Eq.(1). Pseudopotentials are formed from first principles atomic calculations. They contain no empirical parameters, thereby retaining the generality of the method. Because of the differences in orthogonality conditions between the core orbitals and the valence orbitals, the pseudopotentials are dependent on the angular momentums and are semi non-local. When further modifications suggested by Kleinman and Bylander[3] are made, the pseudopotential V_p takes the form

$$V_p = V_{local} + \sum_{\ell} \mid \xi_\ell\rangle\langle\xi_\ell \mid, \tag{4}$$

where the first term is a local part of the pseudopotential and the second term is the completely non-local contribution for angular momentum ℓ.

Replacing the Coulomb potentials with the much smoother pseudopotentials greatly improves the convergence of the planewave basis set, i.e., fewer functions are needed to solve Eq.(1). Nevertheless, we have found that it may still require of the order of $10,000 - 100,000$ functions to obtain sufficient accuracy for the kinds of systems that we are interested in. With a basis set of this size, it is not possible to diagonalize the Hamiltonian matrix by traditional means. Instead methods based on iterative solutions of the eigenvalue problem for only the occupied orbitals must be used. The important feature of these methods for our algorithms is that the full matrix of the Hamiltonian is never needed. Only products of the Hamiltonian with the wavefunction, i.e., $\mathcal{H} \mid \psi\rangle$ are required. Furthermore, these products can be evaluated very efficiently on vector or parallel computers. In the momentum representation, these products are written as

$$\langle\mathbf{k} \mid \mathcal{H} \mid \psi\rangle = \frac{k^2}{2m} C_{\mathbf{k}} + \sum_{\ell}\langle\mathbf{k} \mid \xi_\ell\rangle \sum_{\mathbf{k}'}\langle\xi_\ell \mid \mathbf{k}'\rangle C_{\mathbf{k}'} + \sum_{\mathbf{k}'}\rangle V_{\mathbf{k}-\mathbf{k}'} C_{\mathbf{k}'} \tag{5}$$

where $V_{\mathbf{k}}$ is a Fourier component of all the local potentials, i.e., V_{local}, V_H and V_{xc}. The summation over $\mathbf{k}'$ in the second term of Eq.(5) is a dotproduct of the large vectors, $\langle \mathbf{k} \mid \xi_\ell \rangle$ and $C_{\mathbf{k}}$. It can be computed efficiently on vector or parallel computers. The summation in the third term is a convolution and can be evaluated by Fast Fourier Transform (FFT). There are very efficient FFT algorithms available both on vector and parallel computers.

For optimization problems, Eq.(1) must be repeatedly solved many times as the atoms move, for example, 100,000 times to simulate dynamics for 10 picoseconds. Present methods are sufficiently fast to simulate systems of 50 atoms or so. However, for larger sizes a single step on a CRAY Y-MP takes a minute of CPU time. For large systems the most time consuming parts of the present algorithm are the calculation of $V_p \mid \psi_i \rangle$ and the orthonormal constraints (λ_{ij}). The calculation of these terms involves $N_a \times N_e \times N_g$ operations. Here N_a is the number of atoms in the system, N_e is the number of electrons in the system and N_g is the number of grid points used in FFT. For systems with stronger pseudopotentials than carbon (e.g. O, and F) the size scaling may be even worse (see below). Since the algorithm scales as N^2 or higher it is clear that further improvements in efficiency must be made if systems of the order of 500 particles are to be studied. With the most advanced parallel computers, these difficulties can be partially removed.

3. AB INITIO MOLECULAR DYNAMICS SIMULATIONS

Recently Car and Parrinello[4] proposed an approach for updating the wavefunctions that gives a significant improvement in efficiency for certain systems, including some of those of interest here. In this approach the electronic wavefunction is considered to be an additional dynamical degree of freedom governed by a "fictitious Newtonian" equation,

$$\mu \ddot{\Psi}_{i\sigma} = -H\Psi_{i\sigma} + \sum_j \Lambda_{ij} \Psi_{j\sigma} \tag{6}$$

where μ is a fictitious inertial mass for the wavefunctions. The position of the atoms are governed by the regular Newtonian dynamics,

$$M_I \ddot{\mathbf{R}}_I = -\nabla_I E(\Psi, \mathbf{R}) \tag{7}$$

If μ is nearly equal to zero, the wavefunctions approximately satisfy the Kohn-Sham equation (1) during the dynamical simulation. In this case, the system evolves on the Born-Oppenheimer (BO) energy surface given by $E(\Psi, \mathbf{R})$. Eqs.(6) and (7) are governed by mathematically equivalent equations and can be solved on an equal footing. We use a Verlet algorithm to solve these equations.

4. IMPLEMENTATIONS OF AIMD

In order to take full advantage of modern computers, we must develop algorithms taking into account their hardware architectures and operating systems. We have developed AIMD codes for various different platforms taking into account their specific hardware architectures. Aspects of numerical calculation algorithms are briefly given below.

data parallel operations

Many mathematical operations in planewave based LDA calculations are data parallel operations such as $\vec{X} = \alpha\vec{X} + \vec{Y}$ where $\vec{X}$ and $\vec{Y}$ are vectors with a large number of components. Since each component can be computed independently, these types of operations are very efficient on almost any platforms. Many RISC workstations (e.g., IBM and Silicon Graphics workstations) and vector machines such as CRAY provide very efficient routines, Basic Linear Algebra Subroutines (BLAS) tuned for their architectures. Our code takes full advantage of these routines. Furthermore, data parallel operations are perfectly parallelisable without any interprocessor communication. This is why the planewave method is easily parallelized.

array reductions

Our implementation of AIMD also uses a lot of array reduction operations such as dotproduct of two large vectors. All reduction operations used in our AIMD are vectorizable on CRAY. The BLAS routines tuned for RISC and vector architectures include these routines. Therefore, these operations are performed very efficiently on RISC and vector machines. On the other hand, array reduction requires interprocessor communications on parallel machines or networks of workstations, which may cost significant cputime on distributed-memory parallel machines. Fortunately, the CM-5 has very efficient routines to carry out these operations in the high performance Fortran language.

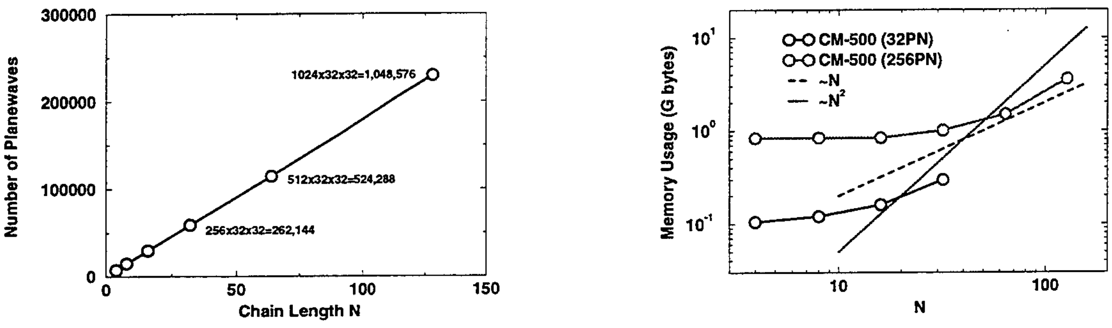
Fast Fourier Transform

The efficiency of the current AIMD code relies on 3D fast Fourier transformation (FFT) method. On serial processor machines such as RISC and vector machines, multi-dimensional FFT is essentially same as a series of 1D FFT. Very efficient FFT routines are available on these platforms. On massively-parallel distributed-memory machines, 3D FFT is inefficient due to heavy interprocessor communication. However, if the array data is properly distributed, the multi-instance multi-dimensional FFT routines on CM-5 is remarkably fast. Unfortunately, such a fast FFT is not available on other massively parallel computers at present. We are currently developing an efficient 1ent FFT for MIMD parallel computers.

4.1. **Language Consideration.** The serial version of our AIMD code was originally written in standard Fortran 77 whereas the parallel version uses a high performance fortran (HPF) (Fortran 90 + some extension). It is desirable to write a portable code that runs on any advanced vector computers, massively parallel machines and networks of workstations without major modification. High performance fortran is now available on most massively parallel platforms. Except for compiler directives that control array layout, the new code is entirely written in HPF. With simple editing of compiler directives, it should run on any platforms supporting HPF, including IBM SP2, SG Power Challenge, CM-500, etc.[5] However, an efficient 3-dimentional FFT routine is not available at present except for on the Connection Machine model 500. Therefore, we tested the new code on CM-500 at Naval Research Laboratory (NRL) and CM-5 at Army High Performance Computing Resource Center(AHCRC) using CM-fortran 2.3.

5. PERFORMANCE TESTS

The computational efficiency of the new code is tested by computing a chain of polyaccetylene using periodic boundary condition. Instead of evaluating density at many k points in the Brillouin zone, more than one repeat units are placed in a large unit cell. This system is computationally one of the worst because a large supercell is needed to isolate the chain from nearest neighbors. Furthermore, we used radix-2 FFT because mixed-radix FFT is not available on most parallel platforms at present. For a more densely packed system and with mixed-radix FFT available on certain platforms, higher performance can be obtained.



(a) Number of Planewaves

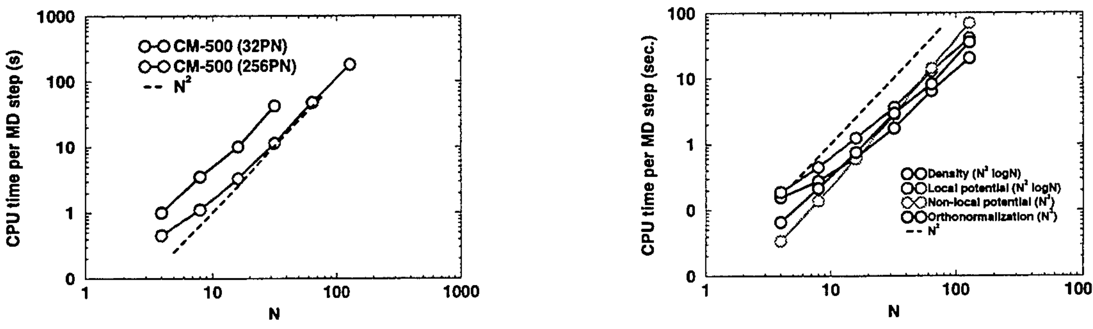
(b) Memory Usage

FIGURE 1. Size of planewave basis and memory usage of ab initio molecular dynamics simulation for $[CH]_N$ chains. The numbers in the left figure are corresponding grid sizes in real space.

5.1. **Computational Efficiency I. – Memory Usage.** The size of planewave basis used for this test is shown in Fig 1(a). The size of corresponding real space grids is given in the figure. Previously, both the wavefunction in momentum space and in real space are kept in memory. Since the real space wavefunction consumes significant amount of memory, it was not possible to calculate $[CH]_{64}$ due to memory overflow. Now, it is possible to compute up to $[CH]_{128}$, which involves as many as 360 electrons per spin.

We successfully reduced the memory usage in the new code It uses about 1/3 – 1/2 of memory previously required. Figure 1(b) displays the amount memory used on 32-PN and 256-PN CM-500 as a function of the number of repeat units. For large N, the memory usage should scale N^2 since both the number of basis functions and the number of electrons increase linearly. Fig. 1(b)(a) indicates that at N=128, the memory usage is nearly in proportion to N^2. Therefore, the memory usage becomes a major problem with planewave method above this size. However, chains of these sizes are large enough to simulate the properties of realistic polymers with various defects (e.g., kinks, dopants) We will carry out such simulation in near future.

5.2. **Computational Efficiency II. – Speed.** Reduction of memory usage increases the number of interprocessor communication, which may require additional cputime. Figure 2 shows the cputime usage per molecular dynamics time step as a function of the chain length. In order to perform a significant period of simulation the cputime per step must be less than a few minutes. The new code is fast enough to perform dynamical simulation of $[CH]_{64}$. Although further improvement in speed for $[CH]_{128}$ is desired, limited simulation can be done for this size. Using the mixed-radix FFT library routines now available on CM-500, it is possible to reduce the cputime significantly for large systems.



(a) Total cputime usage

(b) Cputime used by major routines

FIGURE 2. Cputime usage of ab initio molecular dynamics simulation for $[CH]_N$ chains.

TABLE 1. Theoretical and experimental geometries of CH_2. Bond distance(R) in angstroms, bond angle (θ) in degrees and triplet-singlet energy gap ($\Delta E = E_S - E_T$) in kcal/mol. Subscripts T and S indicate triplet and singlet states, respectively. Theoretical values except for the present calculation were taken from ref.[11]

Observable	S-VWN	B-LYP	HF	MP2	QCISD	AIMD	EXP.
R_T	1.093	1.089	1.071	1.077	1.082	1.077	1.078
R_S	1.135	1.132	1.097	1.109	1.117	1.111	1.111
θ_T	134.8	133.5	130.7	131.6	132.2	136.3	136.0
θ_S	99.1	99.1	103.0	102.1	101.5	101.8	102.4
ΔE	16.0	11.6	30.6	20.5	15.7	15.0	9.0

TABLE 2. Ionization potential of Beryllium atom. Experimental and other theoretical values are taken from Ref [12]

Experiment	HF	LSD	NL	SIC	GC	present result
0.343	0.295	0.334	0.330	0.298	0.330	0.330

Since the number of floating operations in the planewave method is proportional to N^3, the cputime also scales N^3 at large N. In principle, N^3-scaling is valid on parallel machines when N is sufficiently large. However, because of interprocessor communication and other complication in massively parallel operations, our new code does not scale N^3 but nearly N^2 for the size of systems we are interested in. This is because for small systems the routines scaled as N^3 have small prefactors and the routines scaled as $N^2 \log N$ are dominant.

5.3. **Degree of Accuracy.** The high degree of accuracy of the present method has been demonstrated in literatures. [6, 7, 8, 9, 10] Our previous code has been independently tested at Wright-Patterson Air Force Base and proved to be very accurate (See Table 1.) Here we demonstrate the accuracy of the new capability that we has added, which is ability to calculate charged species in supercell. First, we calculated the ionization potential (I) of Beryllium atom from the energies of neutral atom and positive ion. The result is compared with other calculations in Table 5.3. The present result agrees very well with other calculations and experiment very well. Next, we calculated the electron affinity (A) of Chlorine atom. We obtained $A = 3.9eV$. Agreement with the experimental value (3.6eV) is marginal but accurate enough to predict many chemical processes. Generally speaking, the calculation of negative ions is harder than that of neutral atoms or positive ions. We are currently implementing a genelarized gradient correction method which is expected to improve the electron affinity calculation.

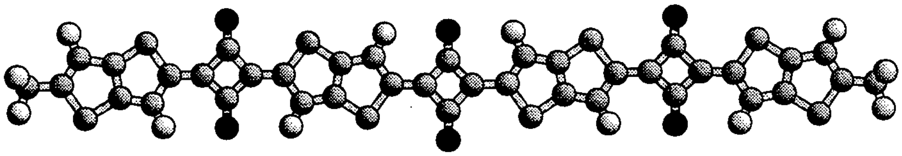


FIGURE 3. Structure of the trimer, $CH_3 - [C_{10}H_2O_2S_2]_3 - C_7H_5S_2$

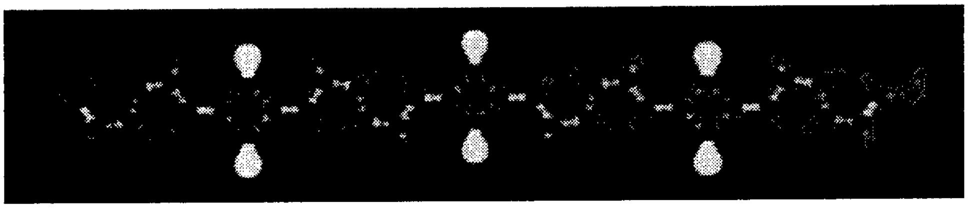


FIGURE 4. Valence electron density of the trimer, $CH_3 - [C_{10}H_2O_2S_2]_3 - C_7H_5S_2$

6. APPLICATIONS

Developing highly conducting polymers is important in Air Force Technology. We design new conducting polymers using our new AIMD. We begin with polyacetylene and a polymer based on squarelene. However, our applications will be extended to many other chemical systems such as molten salts. In the following, we show preliminary results of squarelene-based polymers and chloroaluminate melts.

6.1. **Narrow-Gap Polymers based on Squarelene.** A polymer based on a squarelene and fused thiophenes shown in Fig 3 is a candidate of narrow-gap semi-conducting polymer. An accurate band-gap as a function of polymer length is important to understand the electronic properties of this polymers. However, because of its large repeat unit, it is too expensive to apply an accurate CI method. On the other hand, the Hartree-Fock method may not provide sufficient accuracy. If the band gap is evaluated from HOMO-LUMO gap, HF overestimates the gap substantially whereas LSD significantly underestimates it. The singlet-triplet excitation energy approaches to the band gap as the size increases, if many-body effects do not play a significant role. Since LSD accurately predicts the singlet-triplet excitation energy as demonstrated in Table 1, we estimate the band gap by calculating both singlet and triplet state energy.

Oligomers up to trimer (Fig. 3) are calculated. The largest system contains 250 electrons. A supercell of $40.7\mathring{A} \times 8.5\mathring{A} \times 5.3\mathring{A}$ box and $256 \times 64 \times 32$ grid points are used. About 200,000 basis functions per orbital is needed to get converged results. The electron density on the molecular plane (Fig. 4) indicates no clear alternation in bond order. Singlet-triplet excitation energy plotted in Fig. 5 suggests that the band gap of a infinite chain will be less 0.5eV.

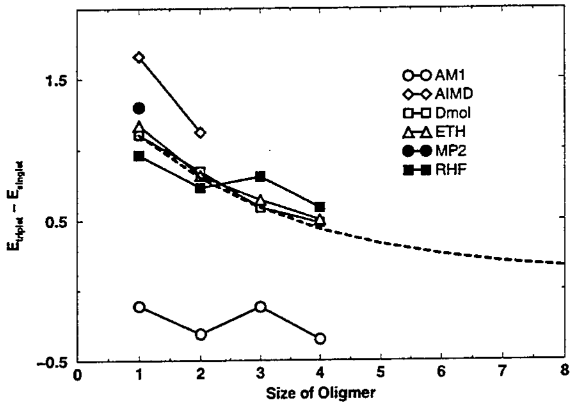
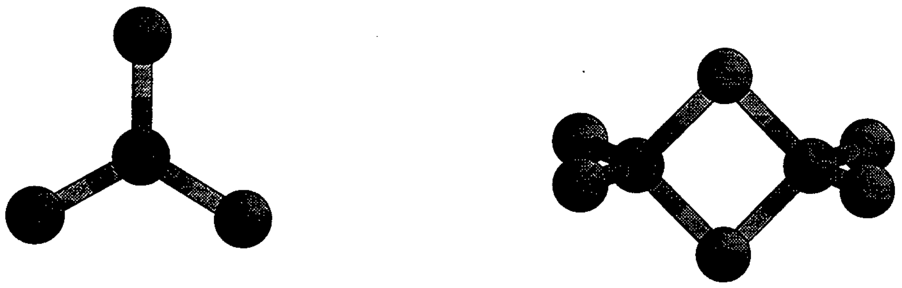


FIGURE 5. Various theoretical calculation of singlet-triplet energy gap as a function of oligomer length

6.2. **Alkali Chloroaluminate Molten Salt Mixtures.** Molten salt mixtures based on aluminum chlorides are of great importance in electrochemical engineering and have wide industrial applications. However, their physical and chemical properties are not completely understood. Since these chloroaluminate melts display strong compositional dependences, it is important to determine the structures of the melts at molecular level. Various chemical reactions between molecular species in the melts produce complicated chemical equilibria. Since interaction between them involves changes in electronic configuration, empirical force fields does not work. First principle simulation is needed. We apply our AIMD code to both static and dynamical structures of chloroaluminate melts. Before direct simulation, it is necessary to determine properties of individual molecular species. In the following, we will show our structures of various molecules involved in the melts. The geometries are fully optimized without symmetry constraint.

6.2.1. *$AlCl_3$.* We obtain a planar D_{3h} structure for $AlCl_3$ (Fig 6(a)) with the bond distance $R_{Al-Cl} = 2.04$ Å in good agreement with experimental value 2.06 Å. This molecule is not stable in the condensed phase and dimerized in the crystal and gas phases. However, the structure of pure $AlCl_3$ liquid is still under debate. Previously, main species in the liquid was believed to be dimer Al_2Cl_6. [13] However, recent experiment indicates that polymeric corner-sharing tetrahedron may be more consistent with the data.[14] If a new model is correct, a similar polymeric structure is expected for the $AlCl_3$-rich melts. We will perform ab initio molecular dynamics simulation of this liquid in near future.

6.2.2. *Al_2Cl_6.* A fluid of Al_2Cl_6 dimers has been an established model of molten $AlCl_3$. However, as mentioned above, this model is challenged by a new model based on polymeric chains of $AlCl_3$. In order to determine the structure of the melt and melting mechanism, it is necessary to understand the chemical nature of the dimer. The present method predicted a D_{2h} structure shown in Fig. 6(b). The detailed geometry is given in Table 3. When a Gaussian basis set was used, an extended basis must be carefully



(a) $AlCl_3$

(b) Al_2Cl_6

FIGURE 6. Equilibrium structure of $AlCl_3$ and its dimer.

TABLE 3. Structure of Al_2Cl_6. Cl_e and Cl_b indicate Cl at edge and at bridge, respectively. Experimental values are given in parenthesis.

	Bond length(Å)		Bond angle	
Symmetry	$Al-Cl_e$	$Al-Cl_b$	$Cl_e-Al-Cl_e$	$Cl_b-Al-Cl_b$
D_{2h}	2.03 (2.07)	2.25 (2.25)	121°(123°)	92°(91°)

chosen because the binding is relatively weak and electron density in the square ring is nearly flat.[15] The planewave method used in this report does not have such difficulty and the present result is in very good agreement with experimental data. The binding energy with respect to the monomer is 1.18eV.

6.2.3. $AlCl_4^-$. In contrast to the $AlCl_3$-rich melts, the structure of the equimolar compositions is simply consisting of $AlCl_4^-$ and alkali cations. We obtained a tetrahedron (T_d) with $R_{Al-Cl} = 2.11$ Å. (Table 6.2.3) The dissociation energy to $AlCl_3 + Cl^-$ is 4.73 eV.

6.2.4. $[Al_2Cl_7]^-$. In the $AlCl_3$-rich melt, $[Al_2Cl_7]^-$ is one of dominant species. However, its geometric structure is not well known. A HF/6-31G* calculation predicted that four isomers shown in Fig 7 are nearly degenerate.[19, 16] We obtained similar results and the energy difference between the isomers is too small to identify the most stable structure. The structure is basically corner sharing tetrahedron. Tetrahedral units

TABLE 4. Structure of $[AlCl_4]^-$.

	R_{Al-Cl}			
Symmetry	This work	MNDO[a]	HF/6-31G*[a]	Experiment[b]
T_d	2.11	2.15	2.17	2.11–2.16

[a] Ref [16]

[a] Refs [17, 18]

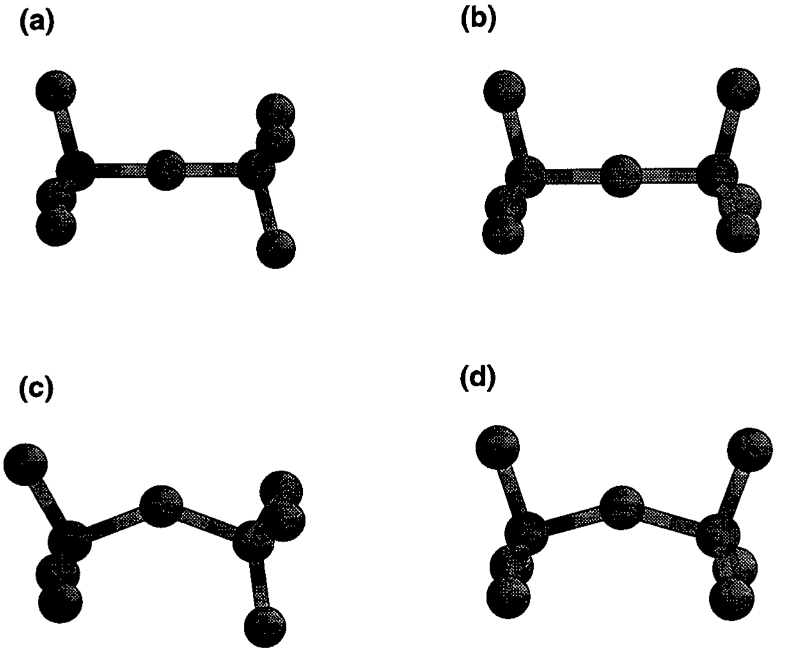


FIGURE 7. Equilibrium geometries of nearly degenerate $[Al_2Cl_7]^-$ isomers

can rotate nearly freely around the shared chlorine atom. However, the situation could be quite different under the presence of counter cations. We will investigate these isomers in the melts using our AIMD.

6.2.5. *Dissociation and Reactions.* When the composition is not equal, a complicated chemical equilibrium of various species is formed. It is a challenging problem to determine the structure of such chemical complex from first principle. As a first step toward the understanding of chemical equilibria, we investigate kinetics of possible species in gas phase. Dissociation energy and heat of formation are listed in Table 6. In order to determine realistic equilibrium structure, we need to know the kinetics and entropy in the melt. Such

TABLE 5. Energy of $[Al_2Cl_7]^-$ isomers relative to D_{3d} structure

Isomer	Symmetry	Relative Energy (eV)
Structure (a)	D_{3d}	0.0
Structure (b)	D_{3h}	0.0
Structure (c)	C_s	-0.139
Structure (d)	C_{2v}	-0.090

TABLE 6. Reaction Energies. The energy cost is defined as $\Delta E = E(products) - E(reactants)$

Reaction	$\Delta E(eV)$
$Al_2Cl_6 \Longrightarrow 2(AlCl_3)$	1.18
$[Al_2Cl_7]^- \Longrightarrow AlCl_3 + [AlCl_4]^-$	2.10
$[Al_2Cl_7]^- + Cl^- \Longrightarrow 2[AlCl_4]^-$	2.63
$[AlCl_4]^- \Longrightarrow AlCl_3 + Cl^-$	4.73
$[Al_2Cl_7]^- \Longrightarrow Al_2Cl_6 + Cl^-$	5.64

quantities can be obtained from AIMD simulation at finite temperature. With the present enhancement of our AIMD code, such simulation is feasible.

7. CONCLUSIONS

A new ab initio molecular dynamics (AIMD) simulation code based on planewave-local density functional method has been successfully implemented. The new code uses much less memory than before without increasing cputime usage. We demonstrated that with the new code, polymers with large repeat units and complicated molten salt mixtures can be simulated from first principles. The electronic structures of oligomers, $[C_{10}H_2O_2S_2]_n$ up to $n = 3$ have been calculated using the new code. This calculation involves 250 electrons and it is one of the largest LSD calculations. As another possible application, we calculated various aluminum chlorides involved in chloroaluminate melts. We obtained geometric structures and kinetics of various reaction. These results indicates that the new code is accurate enough to predicts chemistry of complex systems. We have not shown any results of dynamical simulation due to the short research period. However, the current results promise the feasibility of large ab initio molecular dynamics simulations we are going to perform.

I would like to thank Dr. Todd Yeates and Dr. John Wilkes for supporting my research in various ways and also for their hospitality. I am also thankful to AFOSR for providing me with the opportunity to work at US Air Force Academy through Summer Faculty Research Program.

REFERENCES

[1] W. Kohn and L. J. Sham. *Phys. Rev.*, 140:A1133, 1965.

[2] D. Hamann. *Phys. Rev. B*, 40:2980, 1989.

[3] L. Kleinman and D. M. Bylander. *Phys. Rev. Lett.*, 48:1425, 1982.

[4] R. Car and M. Parrinello. *Phys. Rev. Lett.*, 55:2471, 1985.

[5] S. Saini. Nas experiences of porting cm fortran codes to hpf on ibm sp2 and sgi power challenge. Technical Report NAS-95-010, 1995.

[6] R. Kawai and J. H. Weare. From van der waals to metallic bonding: The growth of be clusters. *Phys. Rev. Lett.*, 64:81, 1990.

[7] R. Kawai and J. H. Weare. Instability of the b_{12} icosahedral cluster: Rearrangement to a lower energy. *J. Chem. Phys.*, 95:1151, 1991.

[8] R. Kawai and J. H. Weare. Anomalous stability of b_{13}^{+} clusters. *Chem. Phys. Lett.*, 191:311, 1992.

[9] M.-W. Sung, R. Kawai, and J. H. Weare. Packing transition in nanosized li clusters. *Phys. Rev. Lett.*, 73:3552, 1994.

[10] R. Kawai, J. F. Tombrello, and J. H. Weare. Li_5 as a pseudorotating planar cluster. *Phys. Rev. A*, 49:4236, 1994.

[11] B. G¿ Johnson, D. M. N. Gill, and J. A. Pople. *J. CHem. Phys*, 98:5612, 1993.

[12] R. M. Dreizler and E. K. U. Gross. *Density Functional Theory.* Springer, Berlin, 1990.

[13] R. L. Harris, R. E. Wood, and H. L. Ritter. *J. Am. Chem. Soc.*, 73:3151, 1951.

[14] Y. S. Badyal, D¿ A. Allen, and R. A. Howe. The structure of liquid $alcl_3$ and structural modification in $alcl_3$ –mcl (m=li, na) molten salt mixtures. *J. Phys. Condens. Matter*, 6:10193, 1994.

[15] M. Ystense, E. Rytter, F. Menzel, and W. Brockner. Scaled quantum mechanical (sqm) vibrational analysis of $alcl_3$, al_2cl_6 and the free ion $alcl_4^-$. *Spectrochim. Acta A*, 50A:233, 1994.

[16] M. Blander, E. Bierwagen, K. G. Calkins, L. A. Price, and M. L. Saboungi. Structure of acidic haloaluminate melts: Neutron diffraction and quantum chemical calculations. *J. Chem.Phys.*, 97:2733, 1992.

[17] N. C. Baenziger. *Acta Crystallogr.*, 4:216, 1951.

[18] K. N. Semenenko, V. N. Surov, and N. S. Kedrova, 1969.

[19] L. A. Curtiss. Ab initio molecular orbital investigation of the structure of $al_2f_7^-$ and $al_2cl_7^-$. In *Proc. Joint Int. Symp. on Molten Salts*, volume 87, page 185. Electrochemical Society, 1987.

EVALUATION OF CURRENT JET ENGINE PERFORMANCE PARAMETERS ARCHIVE, RETRIEVAL, AND DIAGNOSTIC SYSTEM, AND IMPLEMENTATION OF ADVANCED TECHNOLOGY TO IMPROVE SYSTEM

Sandra A. Ashford

Adjunct Assistant Professor

Department of Aerospace and Mechanical Engineering

University of Oklahoma

865 Asp Avenue

Norman, OK 73019

Final Report for:

Summer Faculty Research Program

Oklahoma City Air Logistic Center

Sponsored by:

Air Force Office of Scientific Research

Oklahoma City Air Logistic Center

Tinker Air Force Base, OK

August 1996

EVALUATION OF CURRENT JET ENGINE PERFORMANCE PARAMETERS ARCHIVE, RETRIEVAL, AND DIAGNOSTIC SYSTEM, AND IMPLEMENTATION OF ADVANCED TECHNOLOGY TO IMPROVE SYSTEM

Sandra A. Ashford

Adjunct Assistant Professor

Department of Aerospace and Mechanical Engineering

Abstract

Engine test cell parameter data is used to determine the reliability of an engine, and the extent of any necessary repairs to the engine. This data is currently stored on a VAX11780 machine, and graphic output retrieved from a Tektronix plotter. This project successful replaced the current data archive and retrieval system with a less expensive, easier to use Personal Computer based system. Increased diagnostic capabilities were achieved through the implementation of P.C. based programs.

EVALUATION OF CURRENT JET ENGINE PERFORMANCE PARAMETERS ARCHIVE, RETRIEVAL, AND DIAGNOSTIC SYSTEM, AND IMPLEMENTATION OF ADVANCED TECHNOLOGY TO IMPROVE SYSTEM

Sandra A. Ashford

Introduction

The engineers who make critical decisions as to the performance and reliability of jet aircraft engines base their decisions on the information obtained from "test cell" runs. Currently the Engine Data System (EDS) used for storing this information contains more than 7000 runs, with up to 1216 parameters for each run. This data is currently stored in a HP1000, and is retrieved using a VAX11780. Being that the VAX11780 was purchased and its' software developed in the early 1980's, the system is limited in its capabilities as compared to modern personnel computer systems. A project was initiated to develop a new program by which the engineers in the test cell area could view not only new data, but historical data as well on IBM compatible personnel computers, by-passing the current system and allowing more freedom to the individual user.

History

Engineers have historically looked at jet aircraft engine test cell performance data to determine when a costly engine rebuild is necessary, and the extent and type of that rebuild. Trending of engine test data provides early recognition of test cell and engine problems. This data helps teams make informed decisions on the extent of maintenance required, which saves money formerly spent on unnecessary repairs. Engine testing limits can be scrutinized with this data and fine tuned if necessary to help ensure that engines meet customer expectations. Timely, accurate, and transferable data is essential to reduce not only unnecessary engine rebuilds but also minimize engineer response time.

The VAX11780 was purchased in conjunction with the AVID(Advanced Vibration Diagnostic) system in the early 1980's with a total cost exceeding 10 million dollars. This system allowed engineers a new more powerful method for analyzing their data and provided resources for the development of the EDS system. Each engine going thorough the test cell has up to 1216 performance parameters recorded on the EDS system, and this data has been kept for a period of seven years. The VAX 11780 and engine test specific software retrieves the data from the HP1000 and analyzes it as to the engineers' specific directions. Output requests are processed by a Tektronix plotting machine, which uses expensive silver-coated paper to plot its mono-color output on. As with all early computers, the VAX11780 along with the Tektronix plotter occupy hundreds of square feet of climate controlled floor space. Yearly preventive maintenance contracts alone currently exceed $7500 on the VAX11780.

The principle uses of the VAX11780 and the Tektronix are:

1. Advance Vibration Diagnostic (AVID) system
2. Engine Data System (EDS) trending information

Goal:

Develop a personnel computer program that will not only perform the same functions as the EDS program currently, but also allow customization by the engineer. Program must be user friendly.

Results and Recommendations

The replacement of the EDS system has been successful, and is more flexible and user-friendly than the current system. In addition, the EDS replacement runs on inexpensive IBM compatible machines, having a very low maintenance cost.

The new EDS system has several distinguishing features. It obtains the test cell data directly from the test cell computer, the HP1000, and restructures the data for data base input. A point and click system of buttons within the Windows environment allows easy and user flexible retrieval of data. Once the user selects the engines and data points, according to any possible criteria and sorting method, the new EDS system puts the data into a plotting environment or a compatible diagnostic spreadsheet. The new EDS system allows for printing of log sheets and gives logical names to all engine data points.

It is recommended that this project continue by moving the vibration capabilities of the VAX11780 to a less expensive and more powerful computer. This would again increase the analysis capabilities of the engineers and convert the engineering time currently spent on maintenance to analysis and jet engine corrective and preventive actions. Additionally, it would eliminate a large and expensive piece of equipment by replacing it with a smaller, less expensive, current technology piece of equipment.

ENHANCING TINKER'S RASTER-TO-VECTOR CAPABILITIES

Jeffrey M. Bigelow, Ph.D.
Assistant Professor
Department of Engineering

Oklahoma Christian University of Science and Arts
2501 East Memorial Road
Oklahoma City OK 73131

Final Report for:
Summer Faculty Research Program
Tinker AFB, OC-ALC

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, DC

and

Tinker AFB, OC-ALC

August 1996

ENHANCING TINKER'S RASTER-TO-VECTOR CAPABILITIES

Jeffrey M. Bigelow, Ph.D.
Assistant Professor
Department of Engineering
Oklahoma Christian University of Science and Arts

Abstract

This paper presents an evaluation of the AUDRE Conversion System for the ability to convert schematics from scanned raster to vector format. Particular attention is placed on the trade-off between automatic recognition versus manual insertion of electronic symbols. Text recognition is not considered. The evaluation shows that AUDRE successfully performs the raster to vector conversion. However, with the several suggested improvements, the entire process could be enhanced greatly and draw closer to the desired goal of fully automated document conversion.

ENHANCING TINKER'S RASTER-TO-VECTOR CAPABILITIES

Jeffrey M. Bigelow, Ph.D.

Introduction

The capability to convert engineering documents into a useful, revisable electronic format has become highly desirable. Electronic storage is much more convenient than paper drawings due to the reduction in physical space, the ease of retrieval and the elimination of the aging process and vulnerability of paper. Electronic forms are also more transmittable via computer than paper drawings allowing more competitive outsourcing of replacement parts. Engineering drawings can be stored electronically in two formats: raster or vector. Raster format arranges graphical data as picture cells (pixels) arranged in uniform rows and columns that recreate the drawing when viewed on a monitor or printed on paper. The density of pixels measured in dots per inch (dpi) determines the resolution or clarity of the image. In general, raster files use large amounts of computer memory and thus are expensive in terms of memory to store. A preferable format is vector. Vector format represents the graphical data as lines or other geometries drawn from point to point rather than having to store each pixel along the geometry. For this reason, vector format takes up much less memory than raster format. Text is stored as ASCII characters and thus information such as engineering notes, parts list, and drawing information can be extracted and stored in a meaningful and useful data base. Because the data is vectorized, vector formats also easily allow for scaling and editing of the image. This feature is essential to allow a design to be modified or optimized or a subset of the design to be used in a different application. Finally, computer simulation has become an essential part of the engineering design process. This is especially important for electronic circuits. Therefore, it is imperative for schematic drawings to be in the appropriate vector form for simulation to be possible.

For these reasons, it is not sufficient for engineering drawings to be scanned and stored in raster format. These documents should also be converted into the appropriate vector format. Automated Document Conversion System is a Department of Defense program to install and test new technology for Raster to Vector Conversion (RVC). The primary motivation of RVC is a reduction in cost. Tinker Air Force Base has thousands of electronic schematic drawings of avionics equipment needing reliability and maintainability improvements. RVC is an essential part of meeting these needs in a cost-efficient manner. In this paper, the AUDRE Conversion System [1] is evaluated for the ability to convert schematics from scanned raster to vector format.

Raster-to-Vector Conversion

Electrical schematic diagrams are comprised of symbols representing electronic components connected by straight lines representing wires. Each symbol provides information on the type of the component (e.g. resistor, capacitor, transistor, etc.). Text is located near the symbol that gives the part number and a unique name for each component. For example, a resistor generally has a unique name that starts with R and a value that indicates its resistance. The text near a transistor, on the other hand, must indicate the part number of the transistor as well as its name. All this information, text, symbols and wires, must be retained in order to completely define a circuit.

The RVC process is shown in Figure 1. First the schematic must be scanned into a raster form such as Tagged Image File Format (TIFF). The current standard resolutions are 200, 400 or 600 dpi.

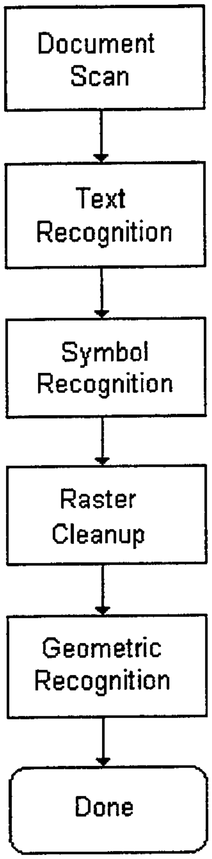


Figure 1. Raster to Vector process.

The first conversion step is to recognize and convert the text containing the part numbers, names and values of each components. The raster comprising the text must be replaced with ASCII characters in a font and size that approximates the original. This can be done using automatic character recognition or manual insertion of text. After this, the underlying raster of the text must be completely erased so that it will not be interpreted as symbols or wires later in the conversion process.

The second conversion step is to recognize and convert the symbols of the schematic. A sample of each symbol used in the collection of schematics is brought together in tabular form in a template file. Automatic Symbol

Recognition (ASR) is done by comparing the raster with this template file. Each raster symbol in the template file is linked to the vector symbol in a symbol file which is used in the final vector form of the schematic. These symbols can also be used to manually insert the correct symbol onto a component through Manual Symbol Insertion (MSI). Once again, after the symbol is recognized and extracted, the underlying raster must be erased.

The final conversion step is to convert the remaining raster to the straight lines that represent the wires of the schematic. This process is called Geometric Recognition (GR), and care must be taken to ensure all text and symbol raster has been removed or else extraneous wires will occur, possibly resulting in short circuits.

At this point the raster representation of the schematic has been converted to a vector representation, and the circuit can be stored, transferred and modified in an electronic format.

Results

The AUDRE Conversion System approaches RVC using the approach shown in Figure 1. Each of the recognition steps (text, symbol and geometric recognition) results in a separate layer being added to the vector form. C-sized schematics from avionics equipment off the B-1 were used to evaluate the performance of version 3.01 of the AUDRE Conversion System running on Solaris 2.5 [2]. A Sun SPARC 2 [3] workstation with 64 Mbytes of RAM and upgraded to a 88 MHz processor was used to perform the evaluations.

Table 1 compares the times needed to perform various tasks concerning symbol recognition. The evaluation revolved around two major points of comparison: the effect of raster resolution and the amount of ASR utilized in the evaluation. The same symbol library was utilized in all six cases and was assumed to be available. Character recognition was assumed to be completed with the text raster removed. GR included the raster cleanup following the symbol extraction.

For both resolutions (200 and 400 dpi), three tests were run. In the first, ASR was not used. Instead all the symbols were inserted manually. In the second, a template was created that included all the symbols found on the schematic. ASR was performed and MSI was utilized on those symbols not found by ASR. In the third, a partial template was created that included only those symbols AUDRE seemed to successfully and quickly find. The remaining symbols were then manually inserted. A discussion on the symbols AUDRE is and is not adept at finding is found in the next section.

TABLE 1

COMPARISON OF TIMES NEEDED FOR AUDRE TO COMPLETE VARIOUS SEGMENTS OF RASTER TO VECTOR CONVERSION

Resolution of Raster Image (dpi)	Automatic Symbol Recognition	Create Template (min)	Automatic Symbol Recognition (min)	Manual Symbol Insertion (min)	Geometric Recognition (min)	Total Time (min)
200	None	NA	NA	22	9	31
200	Full	30 min.	10	11	4	25
200	Partial	10 min.	1	12	4	17
400	None	NA	NA	18	13	31
400	Full	30 min.	10	9	8	27
400	Partial	10 min	1	12	6	19

Discussion

Resolution played a negligible role in the speed of the ASR segment. Version 3.01 has shown significant improvement in the ASR times over version 3.0. In version 3.0, the resolution had to be reduced from 400 to 200 dpi because of slowness. Resolution also does not make a difference in manual insertion. The higher resolution does, however, add two to four minutes to the GR times. The variance in the GR times for a given resolution is due to the necessity of cleaning up the raster before performing the GR. A major impediment for decreasing the overall time consumed to convert a schematic is this raster which is left when a symbol is removed. Both the automatic and manual methods leave excess raster.

A tradeoff is seen in Table 1 between the cost in time for ASR and MSI. The more the software can automatically find, the less that is needed to insert. However, all symbols do not require equal time to identify. Hollow symbols, such as amplifiers, transistors and grounds are relatively inexpensive in time to recognize. Solid symbols, such as diodes, arrows and resistors are much more time costly to find. Experience has shown that resistors and capacitors are the most difficult and time-consuming to recognize. Capacitors are a two-part symbol and AUDRE currently does not handle multiple-part symbols. Suggestions have been made to modify the raster, but one can insert the symbol more quickly. Another trick is to connect the two parts with a single line in the template. This allows a few capacitors to be found. The remaining ones are manually inserted. Resistors are difficult because they are long and narrow with the same width as the wire. Once again, the raster can be modified to aid the recognition process or the resistors can be manually inserted. The ability of AUDRE to automatically find the various symbols is summarized in Table 2. Because of the varying ability to find symbols, two template files were constructed: a full group and a partial group. The partial group included the symbols which AUDRE most successfully and efficiently recognized (transistors, diodes, grounds and amplifiers).

TABLE 2

PERCENTAGE OF SYMBOLS FOUND USING AUDRE'S AUTOMATIC SYMBOL RECOGNITION

	Resistors	Capacitors	Transistors	Diodes	Grounds	Amplifiers	Nodes	Arrows
200 dpi	2 %	11 %	100 %	100 %	85 %	73 %	61 %	10%
400 dpi	83 %	56 %	100 %	100 %	100 %	64 %	64 %	0%

For a single schematic, the cost in time of creating the template plus running the ASR is greater than manually inserting all the symbols. This is true in each of the cases of Table 1. However, a set of schematics using the same symbols quickly makes the ASR routine more valuable as the same template file can be used for all the schematics. At first glance, Table 1 seems to demonstrate that using a partial template file is the fastest method of converting the schematics: 17 minutes for 200 dpi and 19 minutes for 400 dpi. This, however, does not include the ability to run ASR in background. AUDRE allows the user to define macros that can take a set of schematics, run ASR with the complete template set, export the symbols to the vector file and erase the raster; all done without user supervision. The user can then manually insert the remaining symbols and append these to the vector file.

While AUDRE successfully aids in the RVC process for schematics, there are a few areas that could make significant improvement. As mentioned above, one major time consumer is the cleanup of the remaining raster after a symbol has been removed. This could be improved in several ways. First, the ability to move a symbol after placing it would greatly aid the process. Currently, manual insertion can be a difficult process to place the symbol directly on top of the raster. This is especially true for large symbols, such as the amplifier. Any misalignment requires greater raster cleanup. The ability to move symbols once placed would allow better alignment with the raster and allow more of the schematic to be displayed. Since most of the MSI time occurs while scrolling around the schematic, this would allow rapid placement of symbols before zooming in and later moving them precisely in place.

A second factor that leads to extraneous raster is the difficulty in designing symbols that precisely match the raster. After all, it is the image of the symbol that removes the raster lying beneath it. The ability to place a copy of the raster directly on the symbol editor would help produce a closer match between raster and template. Another helpful feature would be the ability to click pixels on or off, rather than being limited to lines, arcs, etc.

The most troublesome feature of AUDRE is in its raster editing tools. Neither the cut nor the paste functions accurately edit the raster. They are off by one pixel vertically and horizontally. In other words, the user boxes the area he wants removed but discovers the editor removed one row and column more or less than he intended. Similarly , the user uses the paste function and carefully lines up the raster in the clipboard with the raster on the screen only to see his image drop one row and one column over.

There are also several improvements in the realm of symbol recognition that would prove helpful. First, two-part symbols such as capacitors should be allowed. Second, the ASR has trouble finding three other symbols: the resistor, arrows, and nodes (the symbol used to indicate wires connect rather than cross over). All three of these symbols are very common in electrical schematics and are essential to fully automatize RVC. Finally, AUDRE's ASR cannot handle gaps in a symbol. For example, the amplifier symbol can be found quickly, unless it has a break in it. In that case, it cannot be found at all. The ability to fill in small gaps in the raster prior to ASR so that it can recognize such symbols would also speed up the process. This is also true in the GR process. Wires with one or two pixel breaks in them are never intentionally drawn that way. If those breaks could be filled in, the GR process could be more accurate.

Conclusion

AUDRE was originally designed for geological maps. Converting schematics from raster to vector has been an additional use for the software. Therefore, it should not be surprising that it is not yet optimized for schematics. Yet, a few minor improvements will make it a much more valuable product. Even so, the current version of AUDRE allows the raster to be successfully converted to vector format. The use of AUDRE was optimized by utilizing the macro capability to run ASR off hours and then completing the RVC process using MSI on the symbols ASR missed. As AUDRE continually improves its symbol recognition engine, the RVC process will approach fully automatic.

Several other steps in the entire process have not been discussed in this paper. Specifically, character recognition for the text of the schematic is vital for the RVC procedure. In addition, if the circuit schematic is to be simulated, it must be converted to a format understandable by analog or digital simulators. The standard format used by all the major electronic simulation packages is the Electronic Data Interchange Format (EDIF) and an AUDRE-to-EDIF conversion program compatible with AUDRE is available. Both of these topics, character recognition and EDIF conversion, are critical components of RVC that must be integrated into the process. Finally, documentation and a complete user's guide must be prepared to efficiently train operators with little previous engineering-specific experience so that labor can be utilized in a cost-efficient manner.

Acknowledgment

The author would like to thank Ed Kincaid of the Technology Insertion Branch of OC-ALC at Tinker AFB, and Jim Johnson and Michael Galperin of AUDRE, Inc. for their invaluable support and assistance on this project

References

[1] AUDRE is a registered trademark of AUDRE, Incorporated (San Diego, CA).

[2] Solaris is a registered trademark of Sun Microsystems, Incorporated (Mountain View, CA).

[3] SPARC is a registered trademark of SPARC International, Incorporated (Menlo Park, CA). Products bearing the SPARC trademarks are based on an architecture developed by Sun Microsystems, Inc.

A COMPUTER MODEL FOR SUSTAINABILITY RANKING

K. M. George
Computer Science Department
Oklahoma State University
Stillwater, OK 74078
(405) 744-5221
kmg@a.cs.okstate.edu

Final Report for:
Summer Faculty Research Program
OC-ALC

Sponsored by:
Air Force Office of Scientific Research
Summer Research Program

August 1996

A COMPUTER MODEL FOR SUSTAINABILITY RANKING

K. M. George
Department of Computer Science
Oklahoma State University

Abstract

A computer model is developed and implemented to compute sustainability ranking number for national stock numbers. The ranking depends on a set of critical variables spanning a multidimensional space. Values of the variables are obtained from different databases. Principal Component Analysis is used to reduce the data dimension to one. The ranking is obtained as a linear combination of the original variables. A computer program has been implemented to obtain values from the relevant databases and to compute ranking. This program accesses the databases, computes the ranking, and produces a table. The program is implemented as an embedded dynamic SQL program using C. The sustainability ranking computed by the program is expected to assist Air Force weapon systems managers in their decision making process.

A COMPUTER MODEL FOR SUSTAINABILITY RANKING

K. M. George

1. Introduction

Information systems technology have produced a "significant revolution in the opportunities for improved managerial performance" [12]. Decision support systems (DDSs) are software systems that assist managers in decision making. Elegant DSS support higher quality decision-making[12]. Well defined numerical and symbolic problems have been attractive domains for DSS [15]. Angehn and Jelasii [1] give an overview of the evolution of the DSS field. Typical characteristics of DSS include database access and user interface. This paper describes a software system - Sustainment Ranking Program (SRP) - to support Air Force weapon system mangers. Even though SRP does not provide any sophisticated user interface, it can be viewed as a special purpose DSS to address a specific need. It may also be included as part of a general DSS. SRP implements a model for ranking national stock numbers (NSNs). This ranking is called *sustainability ranking*. Sustainability depends on reliability, maintainability, supportability, and cost which are defined by specific parameters of supply and maintenance databases. The problem to select a choice from many alternatives have been addressed by researchers for a long time. Many analytical and normative models can also be found in the literature. The central idea behind all models is to use a linear function of the form:

$$f = \sum w_i x_i \text{ where } w_i \text{ are weights associated to the variables } x_i.$$

As Tavan and Banerjee [14] argue those models use weights and subjective probabilities in the evaluation process without providing any method for developing the weights and probabilities. However, the weights used in a model are the significant factors of the model. So, SRP model for ranking is based on sound principles from Statistics. Principal component analysis [9,3,8] is used for building the model. A sustainability ranking number (*rank*) is associated to every NSN. The output from SRP is a table whose columns are NSN and associated variables. The table entries are sorted on decreasing order of rank. A higher rank indicates that the NSN is "probably in a problem area."

In the following sections we describe the underlying mathematics, the mathematical model for sustainability ranking, and the software. Several tests are conducted to verify the model. Test results are described following the description of the model. In the next section we define PCA and its relation to the mathematical model for ranking.

2. Principal Component Analysis

Jolliffe [9] has traced the origin of principal component analysis (PCA) to Pearson [13] and Hotelling [5]. Two applications of PCA are data reduction and interpretation. PCA reduces the dimensionality of a data set involving a large number of variables. The original set of variables are transformed into a new set of variables called the principal components (PCs). The PCs are ordered so that the first variables represent most of the variations in the original variables. In this paper, we are concerned with finding the rank which represents the variations in a set of critical variables. So, we are only interested in the first PC and the sustainability rank is defined as a linear combination of the original variables. In other words, we use PCA as an inference method. We briefly describe how to compute the first PC.

Definition 1: Let X denote a set of raw scores of size N. Then the *average* of X (avg(X)) is defined as $\text{avg}(X) = (\sum X) / N$.

Definition 2: The *deviation-from-mean* form of $x \in X$ is defined as $x_d = x - \text{avg}(X)$. X_d denotes the set of deviation-from-mean of the scores represented by X.

Definition 3: The *correlation* between two variables X and Y is defined as

$$r_{XY} = \frac{\sum X_d Y_d}{\sqrt{\sum X_d^2}\sqrt{\sum Y_d^2}}$$

Definition 4: The *variance* of a variable X is defined as $s^2 = \dfrac{\sum X_d^2}{N}$

Definition 5: A *correlation matrix* of p variables is defined as a p x p symmetric matrix R whose diagonal elements are 1 and the (i, j)th entry is the correlation between the ith and jth variables. For example, the correlation matrix for two variables (assuming that X and Y are deviations-from-mean) is $\begin{bmatrix} 1 & \dfrac{\sum XY}{\sqrt{\sum X^2}\sqrt{\sum Y^2}} \\ \dfrac{\sum XY}{\sqrt{\sum X^2}\sqrt{\sum Y^2}} & 1 \end{bmatrix}$.

Definition 6: Let A^T denote the transpose of a matrix A. A is a *symmetric matrix* if $A = A^T$.

Definition 7: Let A be a p × p symmetric matrix. Then a real number λ is called an *eigenvalue* of A if and only if there is a non-zero vector V in $\mathbf{R}^p$ for which $AV = \lambda V$. Any such vector is called the *eigenvector* associated with the eigenvalue λ.

Definition 8: Let $X_1, X_2, \ldots, X_p$ be p variables representing p sets of data. Then *the first PC* is the eigenvector associated with the largest eigenvalue of the correlation matrix of the p variables. The correlation matrix R is a p × p symmetric matrix whose diagonal elements are 1.

There are several algorithms to find the principal components. Mahony and others provide a solution in terms of the limiting value of a continuous-time dynamical system [10]. In this research, a classical algorithm (Jacobi method) is used. The Jacobi method [2] is one of the methods available for the computation of eigenvalues and their eigenvectors. An implementation of the method is given in appendix II. According to definition 8, the eigenvector corresponding to the largest eigenvalue is the first PC. Since the correlation matrix is symmetric, all eigenvalues are real numbers. Furthermore, all eigenvectors are unit vectors.

3. Mathematical Model for Ranking

As mentioned in the introduction, principal component analysis is used as the foundation for the mathematical model for ranking. In our application sustainability depends on five critical factors. They are man hours, failures, cost, cannibalizations, and MICAP hours. The dependent variable is the sustainability rank of NSN. The sustainability ranking model associates weights to the critical factors. The weight computation is one of the important aspects which distinguishes this model from other models. Values of the critical factors are computed from the data available from databases. For the sake of convenience, let X_1 , X_2 , X_3 , X_4 , X_5 represent man hours, failures, cost, cannibalizations, and MICAP hours respectively. Let R be the correlation matrix of the variables X_i. Let $V = (V_1 , V_2 , V_3 , V_4 , V_5)$ be the eigenvector corresponding to the largest eigenvalue of R. Let S_i represent the variance of X_i. Then rank is defined as follows:

$$\text{rank(NSN)} = \sum_{i=1}^{5} (V_i / S_i) * X_i(\text{NSN}) \qquad (3.1)$$

Rank is a real valued function defined on the NSNs. Since correlation matrix is used for the computation of the first PC, the units of measurement do not unduly influence the weights. All variables are treated equally. An algorithm for rank computation is given in appendix I. It is possible that the computations yield negative V_is. In that case the values of rank may cluster on either side of the origin on V. In that case ranks of some NSNs can give misleading information. Therefore, for practical reasons algorithm I replaces V_i by $|V_i|$ in (3.1).

4. Sustainability Ranking Program

SRP is implemented in the C programming language [4]. Values of the critical variables are obtained from several databases. The database manager is DB2 and the query language used is SQL [6,7]. The general structure of SRP is shown in figure 1.

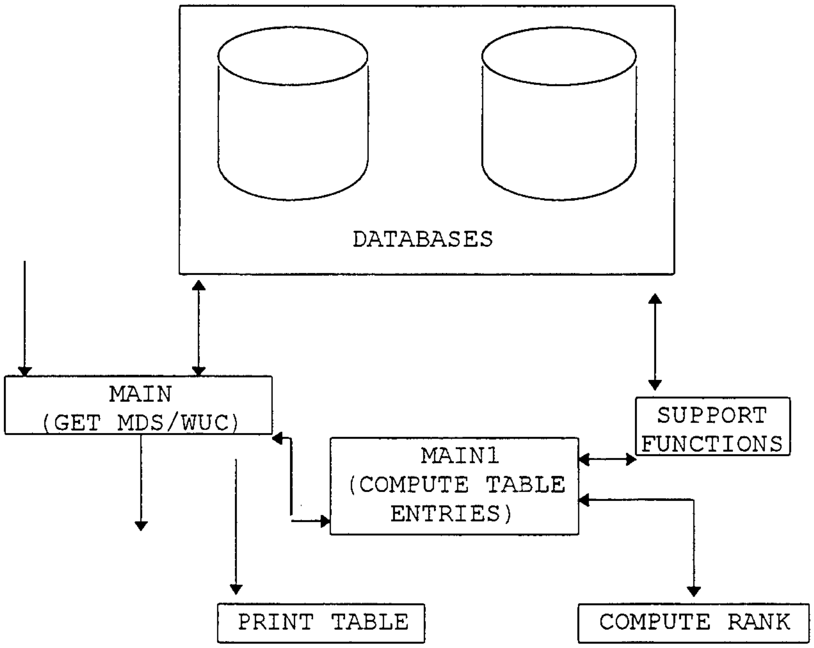


Figure 1. Software Architecture

SRP depends on a cross reference database. NSNs are associated to supply databases whereas MDS and WUC are associated to maintenance data bases. The main program retrieves all the NSN-MDS-WUC triples from the database specified for cross reference. Then for every NSN, the associated table entries are computed. A subset of the formulae used in the computations are given in appendix III. The column headings of table I shows the various computations done. SRP implements the algorithm for rank computation given in appendix I. The program also provides for sensitivity analysis. If the user of the program knows that one of the variables is more important than the others, then its weight can be changed. Sensitivity analysis is discussed in the next section.

5. Sensitivity Analysis

In general, ranking problems are concerned with ordering alternatives from "best" to "worst". Concepts such as "best" are based upon preferences [11]. The mathematical model developed in section 3 for rank computation depends only on the values of the critical variables. Therefore the computed weights of the variables are not subject to any external influence. However, there are practical situations where the values of the variables themselves may not accurately depict the influence of the variables. In such situations, the user of SRP should be given the option to influence the contribution of a variable to the rank computation. In this paper, such an action is called *sensitivity analysis*. Influencing the weight of only one of the variables is addressed in this paper. The sensitivity analysis developed in this paper and implemented in SRP is based on properties of similar matrices and thus providing a mathematical foundation for sensitivity analysis. In order to describe how sensitivity analysis is incorporated into the sustainability ranking model, the following definitions and theorem are needed:

Definition 9: A square matrix P is *nonsingular* if determinant of P is non-zero.

Definition 10: Two square matrices A and B are said to be *similar* (written as $A \approx B$) if there is a nonsingular matrix P such that $A = PBP^{-1}$.

Theorem 1: If X is an eigenvector corresponding to an eigenvalue λ of a matrix A, then PX is an eigenvector associated to the eigenvalue λ of the matrix PAP^{-1}.

Proof: Let $B = PAP^{-1}$, λ be an eigenvalue, and X an eigenvector associated with λ.

So, $AX = \lambda IX$ where I is the unit square matrix.

$$\begin{aligned} \text{Now, } B(PX) &= PAP^{-1}(PX) \\ &= PAX \\ &= P(\lambda IX) \\ &= \lambda I(PX) \end{aligned}$$

Therefore, PX is an eigenvector associated with the eigenvalue λ.

Corollary: Let P be a diagonal matrix such that the i^{th} diagonal entry is 1+d and all other entries are 1. Let E^T be an eigenvector of A where $E = [e_1, e_2, \ldots, e_i, \ldots e_p]$. Then $[e_1, e_2, \ldots, e_i(1+d), \ldots, e_p]^T$ is an eigenvector of PAP^{-1} associated to the same eigenvalue.

Thus, the corollary establishes the connection between the i^{th} weight and the i^{th} variable. Perturbation of the i^{th} column and row of the correlation matrix will induce a corresponding change in the i^{th} component of the eigenvector. So, following is the algorithm developed to perform sensitivity analysis:

```
Algorithm: -- Sensitivity Analysis
/*
    Let $w_i$ represent the weights. Assume that the weight of the $j^{th}$ variable increases by d
    percent
*/
STEP 1:    Compute the eigenvector $e = (e_1, e_2, \ldots, e_5)$ using the algorithm given in
           appendix II.
```

```
STEP 2:    if (i ≠ j) then $w_i \leftarrow e_i$ else $w_i \leftarrow e_i (1+d)$.
STEP 3:    $w_i \leftarrow |w_i| / \sqrt{w_i^2}$
STEP 4:    rank(NSN) = $\Sigma(w_i / S_i) * x_i$;
```

Another method to support sensitivity analysis is to modify the correlation matrix before eigenvector computation.

6. Discussion of Results

Tables I - III and figures 2 - 7 are used in the discussion of results. Table I shows an illustration of the typical output produced by the program.. The output is a table giving information important to the users. The columns are NSN and associated attributes specified in the heading. The table is sorted on descending order of rank. The data used to generate table I is also used as one of the test cases. All tables and figures shown in this paper is based on this particular test case. Table II is derived from table I by omitting all information unrelated to rank computation. It shows the NSNs and their associated critical variables and the rank. Table II shows NSNs and their rank order. These rank orders are used in the graphs shown figures 2 - 7. The x-axis represents the NSNs and the y-axis represents the variables. As can be observed from figures 2-7, a higher value of a critical variable is captured by the rank. Figure 2 compares rank with man hours and failures, figure 3 compares rank with cost per flying hours, and figure 4 compares rank with cannibalizations and MICAP hours. These figures show that a large value of any one of the variables will be reflected in the rank. For example, the NSN that ranked third in this data set has a very high MICAP hours even though near zero values for other variables. This demonstrates the strength of the ranking model. In other words, the model errs on the side of safety. Figure 5 is obtained by overlaying figures 3 and 4 on figure 2 and it verifies that all high values of variables are captured by the rank. Figure 6 shows the movement of the values of variables and the rank. It is obvious that the ranks behavior is the same as the behavior of the critical variables. Figure 7 relates rank to mean time between demand (MTBD). It shows that the NSNs with lower rank has higher MTBD. (The pikes in the graph is figure 7 correspond abnormal data and thus should be discounted.) Figures 2 - 7 validate the rank computation model. The program was tested on several randomly chosen test cases. The results from only one test case is shown in the paper. However, all test cases show results almost identical to the ones given here.

7. Conclusion

This paper describes a mathematical model for sustainability ranking and sensitivity analysis. The sustainability ranking model is based on principal component analysis. .A computer program is implemented based on this model. The results produced by the program is expected to be helpful to Air Force weapon system managers to identify items which may be termed problem areas. Several test cases have been used to verify and validate the model.

The program can be enhanced in several ways. They include automatic flagging of top ranking items and tracing the behavior of the NSNs over time. Current ranking scheme orders all NSNs in the specified database in descending order. However, this ranking may not serve all managers. For example, a manger may not be interested in the NSNs ranked 1-100. The 101st NSN should be ranked 1 for this manager. Such a facility is not available currently. Another enhancement will be to provide access to the table and ranking via WWW.

TABLE I - A Typical Report

ES	NSN	NSNNOUN	MAN_HR	FAILURES	MTBF	MTBD	FLYING HOURS	COST	COST / FLY_HR	CANN	CANN RATE	MIC_HRS	RANK
E1	6620012652887	CONVERTER,SIGNAL DA	6435	317	503	168	132951.5	232764.8	1.75	68	0.559	7538	29.852
EY	6620010872354	TRANSMITTER,RATE OF	3960	174	4679	9999999	141379.2	143244.1	1.01	47	0.049	1166	16.542
2M	6620012524719	SIGNAL PROCESSOR	0	0	999999	1764	132951.5	0.0	0.00	0	0.000	43530	16.061
ET	6620011787651	INDICATOR,FAN INLET	2877	186	2741	9999999	393154.2	104050.2	0.26	55	0.020	936	14.329
EX	6620011440826	INDICATOR,ELECTRICA	1051	196	2256	4348	147396.5	38000.2	0.26	11	0.030	9581	12.428
EY	6620011670874	TRANSMITTER,RATE OF	3327	91	3794	1835	345244.2	120330.4	0.35	32	0.013	171	11.731
ET	6620010613497	INDICATOR,TEMPERATU	1861	144	2672	9999999	365027.3	67323.2	0.18	49	0.019	0	10.059
EX	6620011440827	INDICATOR,ELECTRICA	1009	181	4886	6667	147396.5	36495.5	0.25	6	0.017	3464	9.500
ET	6620011569724	TRANSMITTER,PRESSUR	2520	48	15156	8264	222796.7	91137.5	0.41	9	0.013	4318	9.467
EX	6620011829763	INDICATOR,POWER LEV	1957	100	7977	7813	132951.5	70799.2	0.53	30	0.247	26	8.547
EX	6620011644913	GAGE UNIT,ENGINE	1511	104	6392	1279	132951.5	54638.4	0.41	38	0.313	1156	8.154
EY	6620010344539	TRANSMITTER,FUEL	1680	73	7604	1443	122928.0	60765.6	0.49	21	0.026	1530	7.396
EY	6620012788027	INDICATOR,RATE OF F	844	109	2536	2160	276439.1	30520.2	0.11	27	0.015	3415	7.275
UH	6620012521056	COMPUTER,ENGINE MON	1560	79	3499	453	276439.1	56407.1	0.20	21	0.011	377	6.860
UH	6620012474174	COMPUTER,ENGINE MON	1560	79	3499	9999999	276439.1	56407.1	0.20	21	0.011	0	6.721
ET	6620011297137	INDICATOR,SERVOED,T	1024	84	1638	305	68805.1	37048.9	0.54	24	0.046	0	5.622
ET	6620011805183	INDICATOR,FAN	758	51	1349	1751	68805.1	27409.6	0.40	13	0.025	4650	5.431
ET	6620011726946	INDICATOR,TORQUEMET	0	0	999999	2045	20.8	0.0	0.00	0	0.000	14643	5.403
EX	6620011297247	INDICATOR,ELECTRICA	662	84	5722	2762	412937.3	23962.6	0.06	38	0.013	333	5.144
EX	6620011670881	GAGE UNIT,ENGINE	827	58	13754	1855	132951.5	29905.4	0.22	12	0.099	60	4.103
ET	6620011347226	TRANSMITTER,PRESSUR	1368	20	20614	276	263785.2	49473.3	0.19	0	0.000	240	4.074
EY	6620010606418	INDICATOR,RATE OF F	875	51	3474	2268	88588.2	31634.3	0.36	11	0.016	25	3.970
EY	6620011484354	TRANSMITTER,RATE OF	0	0	999999	27027	147396.5	0.0	0.00	0	0.000	10716	3.954
EX	6620010081654	INDICATOR,ELECTRICA	317	38	3858	8772	73311.3	11473.1	0.16	21	0.055	3119	3.544
EX	6620012199413	INDICATOR,ELECTRICA	266	31	8917	3623	276439.1	9628.5	0.03	18	0.010	2890	3.055
EX	6620011805209	INDICATOR,ELECTRICA	266	33	2085	1661	68805.1	9639.3	0.14	11	0.021	2593	2.863
EY	6620011829489	TRANSMITTER,RATE OF	649	18	59090	7407	132951.5	23481.6	0.18	10	0.082	663	2.607
2M	6620012332583	PROCESSOR,SIGNAL	0	0	999999	1250	132951.5	0.0	0.00	0	0.000	5542	2.045
EX	6620011853016	INDICATOR,RATE OF F	430	18	36931	6803	132951.5	15567.6	0.12	10	0.082	0	1.824
EY	6620010845112	INDICATOR,RATE OF F	222	35	2095	1144	73311.3	8037.0	0.11	8	0.021	0	1.798
ET	6620011685273	TRANSMITTER,PRESSUR	669	2	206138	183	263785.2	24205.0	0.09	0	0.000	0	1.708
ET	6620010107337	INDICATOR,DISPLAY,V	223	33	1053	1290	34760.4	8062.3	0.23	1	0.020	123	1.638
2N	6620010378398	TRANSMITTER,FUEL FL	154	6	88634	4184	132951.5	5562.9	0.04	10	0.082	1684	1.392
EX	6620010258503	INDICATOR,TACHOMETE	165	28	2064	1267	21328.9	5978.9	0.28	0	0.000	72	1.302
EX	6620010182183	INDICATOR TACHOMETE	126	17	1620	828	13766.4	4564.7	0.33	0	0.000	178	0.904
EX	6620011829328	TACHOMETER	154	6	88634	4975	132951.5	5562.9	0.04	10	0.082	23	0.780
EX	6620010070103	INDICATOR,ELECTRICA	73	16	4321	8475	17283.5	2625.9	0.15	0	0.000	58	0.696
EY	6620011853017	INDICATOR,RATE OF F	90	6	22159	19231	132951.5	3255.3	0.02	0	0.000	0	0.408
ET	6620010100038	INDICATOR,ELECTRICA	13	4	4369	1258	17476.9	477.4	0.03	0	0.000	0	0.157

TABLE II Test Data - NSN 662001*

(Variables used in the rank computation and computed rank)

NSN	NSNNOUN	MAN_HR	FAILURES	MTBD	COST	CANN	MIC_HRS	RANK
6620012652887	CONVERTER,SIGNAL DA	6435	317	168	232764.8	68	7538	29.852
6620010872354	TRANSMITTER,RATE OF	3960	174	9999999	143244.1	47	1166	16.542
6620012524719	SIGNAL PROCESSOR	0	0	1764	0.0	0	43530	16.061
6620011787651	INDICATOR,FAN INLET	2877	186	9999999	104050.2	55	936	14.329
6620011440826	INDICATOR,ELECTRICA	1051	196	4348	38000.2	11	9581	12.428
6620011670874	TRANSMITTER,RATE OF	3327	91	1835	120330.4	32	171	11.731
6620010613497	INDICATOR,TEMPERATU	1861	144	9999999	67323.2	49	0	10.059
6620011440827	INDICATOR,ELECTRICA	1009	181	6667	36495.5	6	3464	9.500
6620011569724	TRANSMITTER,PRESSUR	2520	48	8264	91137.5	9	4318	9.467
6620011829763	INDICATOR,POWER LEV	1957	100	7813	70799.2	30	26	8.547
6620011644913	GAGE UNIT,ENGINE	1511	104	1279	54638.4	38	1156	8.154
6620010344539	TRANSMITTER,FUEL	1680	73	1443	60765.6	21	1530	7.396
6620012788027	INDICATOR,RATE OF F	844	109	2160	30520.2	27	3415	7.275
6620012521056	COMPUTER,ENGINE MON	1560	79	453	56407.1	21	377	6.860
6620012474174	COMPUTER,ENGINE MON	1560	79	9999999	56407.1	21	0	6.721
6620011297137	INDICATOR,SERVOED,T	1024	84	305	37048.9	24	0	5.622
6620011805183	INDICATOR,FAN	758	51	1751	27409.6	13	4650	5.431
6620011726946	INDICATOR,TORQUEMET	0	0	2045	0.0	0	14643	5.403
6620011297247	INDICATOR,ELECTRICA	662	84	2762	23962.6	38	333	5.144
6620011670881	GAGE UNIT,ENGINE	827	58	1855	29905.4	12	60	4.103
6620011347226	TRANSMITTER,PRESSUR	1368	20	276	49473.3	0	240	4.074
6620010606418	INDICATOR,RATE OF F	875	51	2268	31634.3	11	25	3.970
6620011484354	TRANSMITTER,RATE OF	0	0	27027	0.0	0	10716	3.954
6620010081654	INDICATOR,ELECTRICA	317	38	8772	11473.1	21	3119	3.544
6620012199413	INDICATOR,ELECTRICA	266	31	3623	9628.5	18	2890	3.055
6620011805209	INDICATOR,ELECTRICA	266	33	1661	9639.3	11	2593	2.863
6620011829489	TRANSMITTER,RATE OF	649	18	7407	23481.6	10	663	2.607
6620012332583	PROCESSOR,SIGNAL	0	0	1250	0.0	0	5542	2.045
6620011853016	INDICATOR,RATE OF F	430	18	6803	15567.6	10	0	1.824
6620010845112	INDICATOR,RATE OF F	222	35	1144	8037.0	8	0	1.798
6620011685273	TRANSMITTER,PRESSUR	669	2	183	24205.0	0	0	1.708
6620010107337	INDICATOR,DISPLAY,V	223	33	1290	8062.3	1	123	1.638
6620010378398	TRANSMITTER,FUEL FL	154	6	4184	5562.9	10	1684	1.392
6620010258503	INDICATOR,TACHOMETE	165	28	1267	5978.9	0	72	1.302
6620010182183	INDICATOR TACHOMETE	126	17	828	4564.7	0	178	0.904
6620011829328	TACHOMETER	154	6	4975	5562.9	10	23	0.780
6620010070103	INDICATOR,ELECTRICA	73	16	8475	2625.9	0	58	0.696
6620011853017	INDICATOR,RATE OF F	90	6	19231	3255.3	0	0	0.408
6620010100038	INDICATOR,ELECTRICA	13	4	1258	477.4	0	0	0.157

TABLE III

Translation of NSN 662001* to Integers (used in graphs)

NSN	DESCRIPTION	VALUE
6620012652887	CONVERTER,SIGNAL DA	1
6620010872354	TRANSMITTER,RATE OF	2
6620012524719	SIGNAL PROCESSOR	3
6620011787651	INDICATOR,FAN INLET	4
6620011440826	INDICATOR,ELECTRICA	5
6620011670874	TRANSMITTER,RATE OF	6
6620010613497	INDICATOR,TEMPERATU	7
6620011440827	INDICATOR,ELECTRICA	8
6620011569724	TRANSMITTER,PRESSUR	9
6620011829763	INDICATOR,POWER LEV	10
6620011644913	GAGE UNIT,ENGINE	11
6620010344539	TRANSMITTER,FUEL	12
6620012788027	INDICATOR,RATE OF F	13
6620012521056	COMPUTER,ENGINE MON	14
6620012474174	COMPUTER,ENGINE MON	15
6620011297137	INDICATOR,SERVOED,T	16
6620011805183	INDICATOR,FAN	17
6620011726946	INDICATOR,TORQUEMET	18
6620011297247	INDICATOR,ELECTRICA	19
6620011670881	GAGE UNIT,ENGINE	20
6620011347226	TRANSMITTER,PRESSUR	21
6620010606418	INDICATOR,RATE OF F	22
6620011484354	TRANSMITTER,RATE OF	23
6620010081654	INDICATOR,ELECTRICA	24
6620012199413	INDICATOR,ELECTRICA	25
6620011805209	INDICATOR,ELECTRICA	26
6620011829489	TRANSMITTER,RATE OF	27
6620012332583	PROCESSOR,SIGNAL	28
6620011853016	INDICATOR,RATE OF F	29
6620010845112	INDICATOR,RATE OF F	30
6620011685273	TRANSMITTER,PRESSUR	31
6620010107337	INDICATOR,DISPLAY,V	32
6620010378398	TRANSMITTER,FUEL FL	34
6620010258503	INDICATOR,TACHOMETE	35
6620010182183	INDICATOR TACHOMETE	36
6620011829328	TACHOMETER	37
6620010070103	INDICATOR,ELECTRICA	37
6620011853017	INDICATOR,RATE OF F	38
6620010100038	INDICATOR,ELECTRICA	39

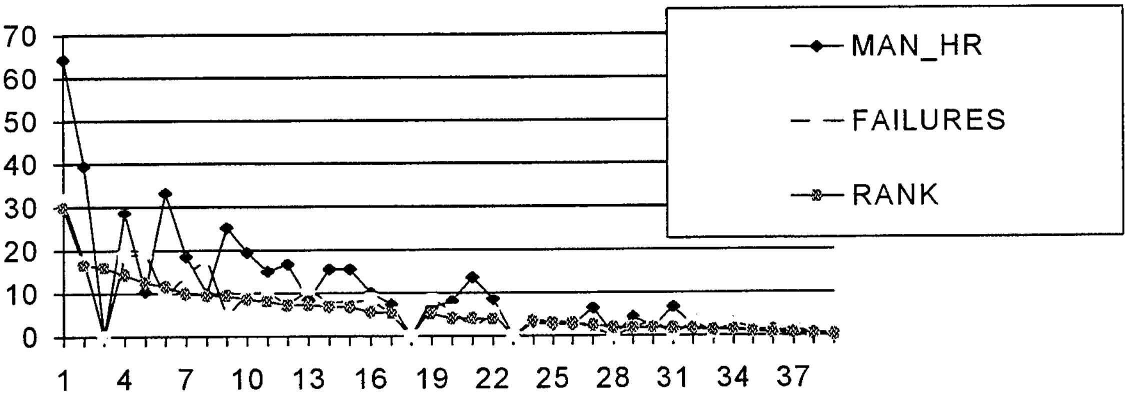


Figure 2 Ranked NSNs VS Man hour, Failures, and Rank. (MAN_HR/100, FAILURES/10, and RANK)

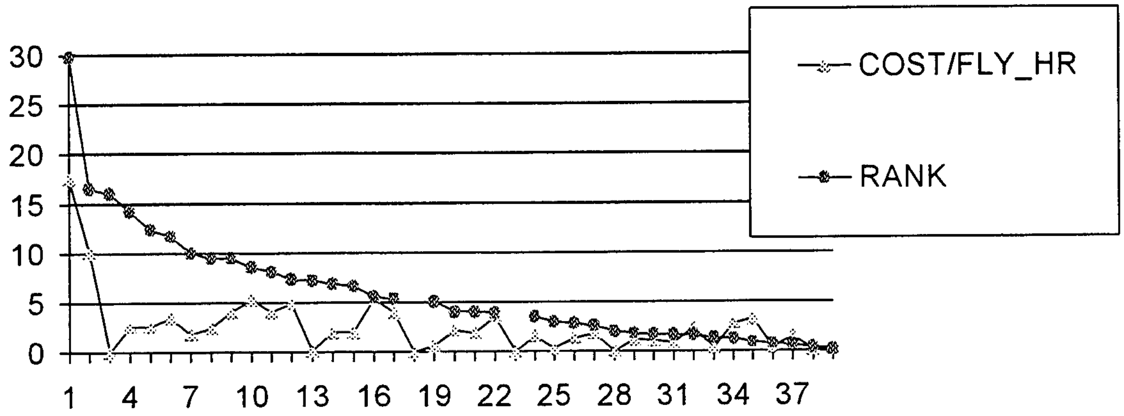


Figure 3 Ranked NSNs VS Cost per flying hour and Rank. (COST_PER_FLY_HR*10 and RANK)

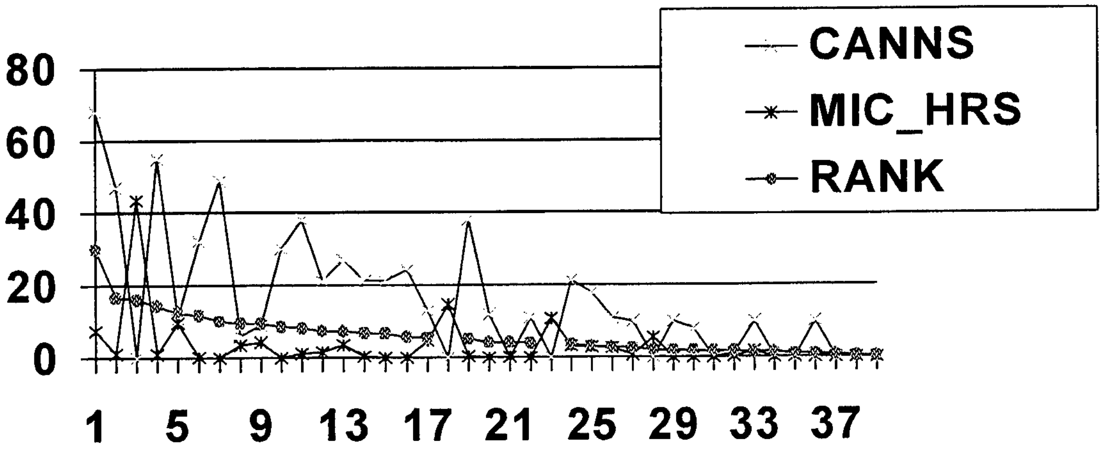


Figure 4 Ranked NSNs VS CANNS, (MIC_HRS / 1000), and RANK.

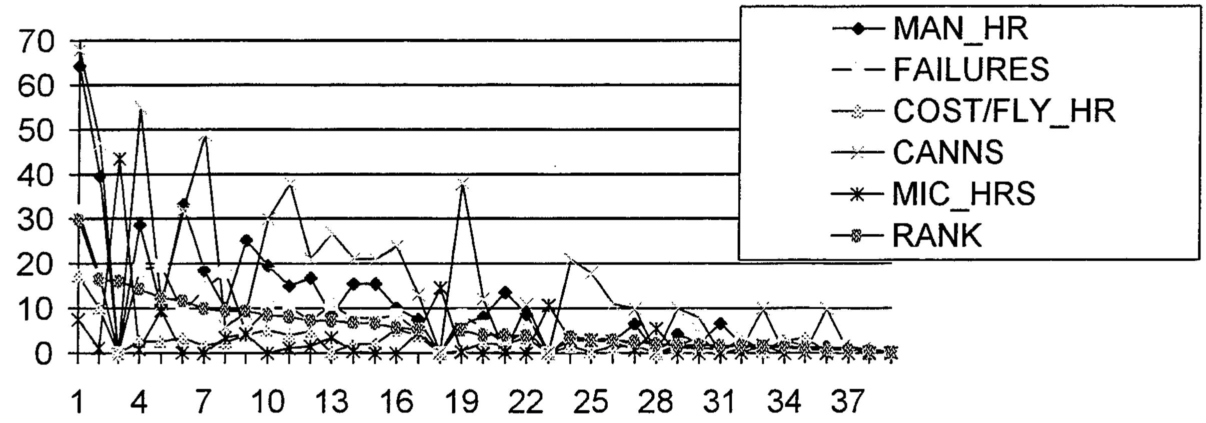


Figure 5 Ranked NSNs VS critical parameters and rank.
(MAN_HR/100, FAILURES/10, COST_PER_FLY_HR*10, CANNS, MIC_HRS/1000, and RANK)

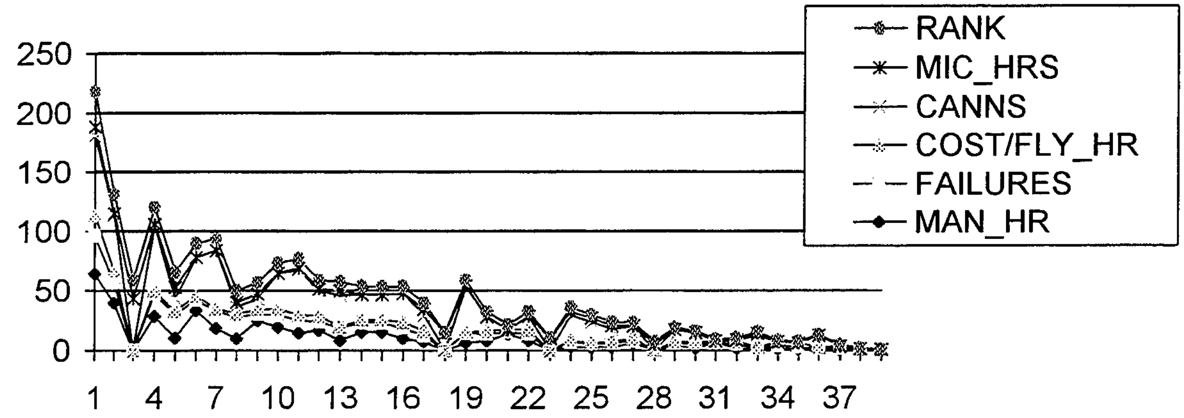


Figure 6 Ranked NSNs VS critical parameters and rank.
(MAN_HR/100, FAILURES/10, COST_PER_FLY_HR*10, CANNS, MIC_HRS/1000, and RANK)

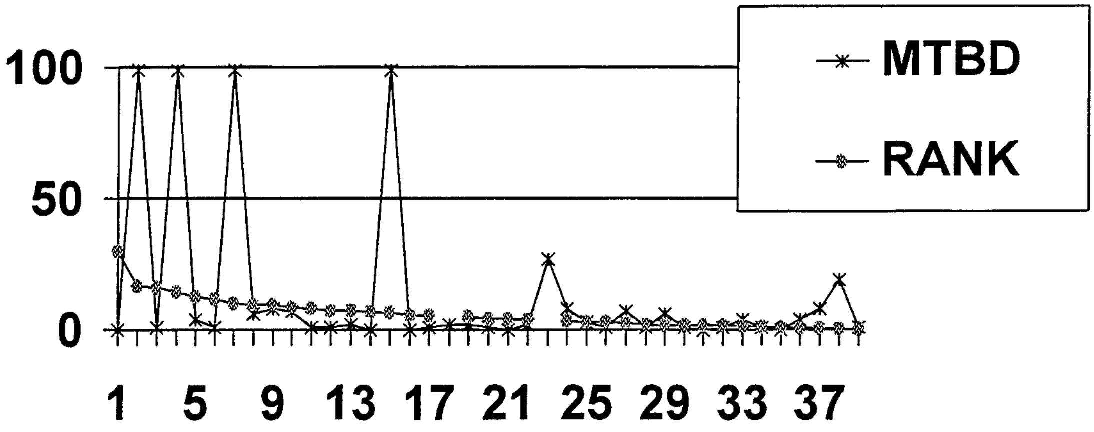


Figure 7. Ranked NSNs VS Rank and MTBD (MTBD / 1000).

APPENDIX I
Algorithm to compute rank

/* Let x1, x2, x3, x4, and x5 represent man hours, failures, cost, cannibalizations, and MICAP hours respectively */

1. Compute the average of the variables x_i.
2. Compute the variance of the variables x_i.
3. Compute the correlation matrix R of the five variables.
   /* The correlation matrix is defined in section 2. */
4. Use the procedure given in APPENDIX II to find the eigenvalues and the corresponding eigenvectors of R.
5. Let (v1, v2, v3, v4, v5) be the eigenvector corresponding to the largest eigen value.
6. rank = (|v1| * x1) / variance of x1 + (|v2| * x2) / variance of x2 + (|v3| * x3) / variance of x3
   + (|v4| * x4) / variance of x4 + (|v5| * x5) / variance of x5.

APPENDIX II
Jacobi method for computation of eigenvectors

/* Input parameters are A - the matrix, tol - error tolerance, and len - number of rows of A. When the function returns, diagonal elements of A are eigenvalues, and columns of V are the corresponding eigenvectors. */

```
jacobi(double A[][5], double tol, double V[][5], int len)
{
    double eps, c, s, t1, t2, offa, af, df, fnorm, max, abval;
    double v[5][5], J[5][5], JT[5][5];
    int i, j, p, q;

/* Initialize v as a unit matrix */
    for  (i = 0; i < len; i++)
        for ( j = 0; j < len; j++)
            v[i][j] = 0.0;
    for  (j = 0; j < len; j++) v[j][j] = 1.0;

/* Compute the norm of the matrix */
    af = 0.0;
    for (i = 0; i < len; i++)
        for (j = 0; j < len; j++)
            af += A[i][j] * A[i][j];
    fnorm = sqrt(af);
```

```
/* Compute tolerance */
    df = 0.0;
    for (j = 0;  j < len;  j ++) df += A[j][j] * A[j][j];
    offa = sqrt(af -df);
    eps = tol * fnorm;

    while (offa > eps){
        p = 0; q = 0;
        max = fabs(A[0][0]);
        for (i = 0; i < len; i++)
            for (j = 0; j < len; j++)
                if (i != j) {
                    abval = fabs(A[i][j]);
                    if (abval > max) [
                         p = i; q = j, max = abval;
                    }
                }

        if  (A[p][q] != 0) {
             t1 = (A[q][q] A[p][p]) / (2 * A[p][q]);
             if (t1 < 0) t2 = (-1) / (-t1 + sqrt(1 + t1 * t1));
             else t2 = 1 / (t1 + sqrt(1 + t1 * t1));
             c = 1 / sqrt(1 + t2 * t2);
             s = t2 * c;
         } else {
             c = 1.0; s = 0.0;
           }
         for (i = 0; i < len; i++)
              for (j = 0; j < len; j++){
                   J[i][j] = 0.0;     JT[i][j] = 0.0;
              }
         for (j = 0; j < len; j++){
                   J[j][j] = 1.0;     JT[j][j] = 1.0;
         }

         J[p][p] = c;     J[p][q] = s;      J[q][q] = c;     J[q][p] = -s;
         JT[p][p] = c;  JT[p][q] =  -s;  JT[q][q] = c;  JT[q][p] =  s;
         mat_mult(JT,  A,  A, len);                                    /* A = JT * A */
         mat_mult(A, J, A, len);                                       /* A = A * J */
         mat_mult(v, J, v, len);                                      /* v = v * J */
/* Compute the off diagonal norm of the matrix */
         af = 0.0;
         for (i = 0; i < len; i++)
              for (j = 0; j < len; j++)
```

```
                af += A[i][j] * A[i][j];
        df = 0.0;
        for (j = 0;  j < len;  j ++) df += A[j][j] * A[j][j];
        offa = sqrt(af -df);
    }
    for (i = 0; i < len; i++)
        for (j = 0; j < len; j++)  V[i][j]  =  v[i][j];
}
```

APPENDIX III

Cross reference and formulae

Cross refernce:

nsn <-> mds/wuc

Formulae:

flying hours = sum of flying hours from avisurs
man hours = man hours from on_eq + man hours from offeq
cost = depot_cost + shipping_cost + base_cost
A = condemns * unit_cost
B = (NRTS -condemns) * unit_cost * .086
depot_cost = A + B
shipping_cost = B * .025
base_cost = if WDS <> 'S' or TMD <> 'R' then manhours * 36.17 else 0

References

[1] Angehn, Albert A., and Jelassi, Twafik, "DSS research and practice in perspective", *Decision Support Systems,* Vol. 12, No. 4/5, November 1994 pp. 257-265.

[2] Golub, Gene H., and van Loan, Charles F., *Matrix Computations*, The Johns Hopkins University Press, Baltimore, 1993.

[3] Green, Paul E., and Carroll, J. Douglas, *Mathematical Tools for Applied Multivariate Analysis*, Academic Press, New York, 1976.

[4] Harbison, Samuel P., and Steele JR, Guy L., *C Areference Manual*, fourth edition, Prentice-Hall, Englewood Cliffs, 1995.

[5] Hotelling, H., "Analysis of a complex of statistical variables into principal components", *J. Educ. Psychol.*, 24, 1933, pp. 417-441, 498-520.

[6] *IBM DATABASE 2 Application Programming Guide - for common Servers,* Version 2, IBM Corporation, 1995.

[7] *IBM DATABASE 2 SQL Reference - for common servers*, version 2, IBM Corporation, 1995.

[8] Johnson, Richard A., and Wichern, Dean W., *Applied Multivariate Statistical Analysis*, Prentice-Hall, Inc., Englewood Cliffs, 1982.

[9] Jolife, I.T., *Principal Component Analysis*, Springer-Verlag, New York, 1986.

[10] Mahony, R.E., Helmke, U., and Moore, J.B., "Gradient Algorithms for Principal Component Analysis", *J. Austral. Math. Soc. Ser. B*, 37(1996), pp. 430-450.

[11] Meszaros, Cs., and Rapcsak, T., "On sensitivity analysis for a class of decision systems", Decision Support Systems,Vol.16, No. 3, March 1996.

[12] Palvia, Shailendra C., and Chervany, Norman L., "An experimental investigation of factors influencing predicted success in DSS implementation", *Information & Management*, 29 (1995), pp. 43-53, July 1995.

[13] Pearson, K., "On lines and planes of closest fit to systems of points in space", *Phil. Mag.* (6), 2, 1901, pp. 559-572.

[14] Tavana, Madjid, "Strategic Assessment Model (SAM): A Multiple Criteria Decision Support System for Evaluation of Strategic Alternatives", *Decision Sciences*, (26), 1, January/February 1995.

[15] Wagner, Christian, "Decision support for "messy" problems", *Information & Management*, 28(1995), pp. 393-403, June 1995.

OPTIMAL GROUNDWATER MANAGEMENT USING GENETIC ALGORITHM

Jagath J. Kaluarachchi, Ph.D., P.E.
Associate Professor
Department of Civil and Environmental Engineering and
Utah Water Research Laboratory
Utah State University
Logan, UT 84322-8200

Final Report for:
Summer Faculty Research Program
Hill Air Force Base

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington DC
and
Hill Air Force Base, Ogden, Utah

August, 1996

ABSTRACT

Genetic algorithm (GA) is considered to be a robust technique for solving groundwater optimization problems. These problems are often nonconvex, nonlinear and discontinuous, and they are difficult to solve using traditional gradient-based techniques. In this study, a linear groundwater management problem is solved using GA. The global-optimal solution obtained using linear programming (LP) is known, and the LP solution is used to find the robustness of GA. The effect of ignoring infeasible solutions is studied, and a string-length defining method (SLDM) is outlined. Also, the problem is subjected to a sensitivity analysis and attempts have been made to standardize GA parameters. Finally, a general guideline for solving such problems using GA is suggested. The guideline reduces GA to the selection of population size only; the size refers to the number of trial solutions GA considers during a standard generation.

INTRODUCTION

Groundwater management problems are encountered during allocation of groundwater among competing demands in a groundwater system, and optimization analysis proves to be a powerful tool for solving such problems. However, such an optimization problem is often nonlinear, nonconvex and discontinuous, and a global-optimal solution cannot be guaranteed by traditional gradient-based techniques (Willis and Yeh, 1987). In addition, these techniques require certain continuity-differentiability of the objective function, and these techniques cannot accommodate objective functions with discrete decision variables. Alternatively, GA is considered to be a robust technique capable of accommodating the nonconvexity, nonlinearity and discontinuity (McKinney and Lin, 1994).

GA is a heuristic, probabilistic, combinatorial, search-based optimization technique patterned after the biological process of natural evolution (Holland, 1975). Although GA can theoretically find the global optimal solution, it is computationally limited to searching near-optimal solutions. As such, GA finds several distinctly different, but equally acceptable near-optimal solutions while offering sufficient flexibility in decision making. Recently, several researchers have robustly solved groundwater management problems using GA (McKinney and Lin, 1994; Rogers and Dowla, 1994; Huang and Mayer, 1995). However, their pursuit for robustness has led to problem-specific selection of GA parameters, and a general guideline for selecting these parameters is not adequately addressed. In this manuscript, attempts are made to develop such a guideline. SLDM is outlined for obtaining the string length and standardized GA parameters are suggested for general application to groundwater problems.

GROUNDWATER MANAGEMENT PROBLEM

The optimization-based groundwater management problems used in this study is obtained from McKinney and Lin (1994). Figure 1 shows a homogeneous and isotropic unconfined aquifer with ten wells. The aquifer is under steady flow condition. The aquifer domain is discretized into 19x21 nodes, and the head datum is taken at the base. The management objective is to maximize the groundwater yield from the wells in the aquifer while restricting individual yield within a specified range. The corresponding optimization model can be stated as

$$\text{Maximize} \quad Q(\mathbf{r}) = \sum_{i=1}^{10} q_i \qquad \mathbf{r} = [q_1, q_2, q_3, \ldots, q_{10}]^T \tag{1}$$

$$\text{subject to} \quad h_i(\mathbf{r}) \geq 0 \qquad i = 1, 2, 3, \ldots, 10 \tag{2}$$

$$q_i \in [0, 7000] \qquad i = 1, 2, 3, \ldots, 10 \tag{3}$$

where **r**=trial solution; Q=objective value; q_i=yield from the i-th well; and h_i=head at the i-th well. **r** belongs to a search space of trial solutions from which the optimal solution is searched. The governing flow equation for the aquifer is written as

$$\nabla^2 \frac{K h^2}{2} = \sum_{i=1}^{10} q_i \, \delta(x_i, y_i) - R \tag{4}$$

where h=head at (x, y); (x_i, y_i)=coordinates of the i-th well; $\delta(x_i, y_i)$=Dirac delta function; K=hydraulic conductivity; and R=areal recharge. In this study, K=50 m/day and R=0.001 m/day are taken from McKinney and Lin (1994). Equation (4) is solved using the finite difference method to generate the simulation model, and Equations (1), (2), (3) and (4) form the groundwater optimization model.

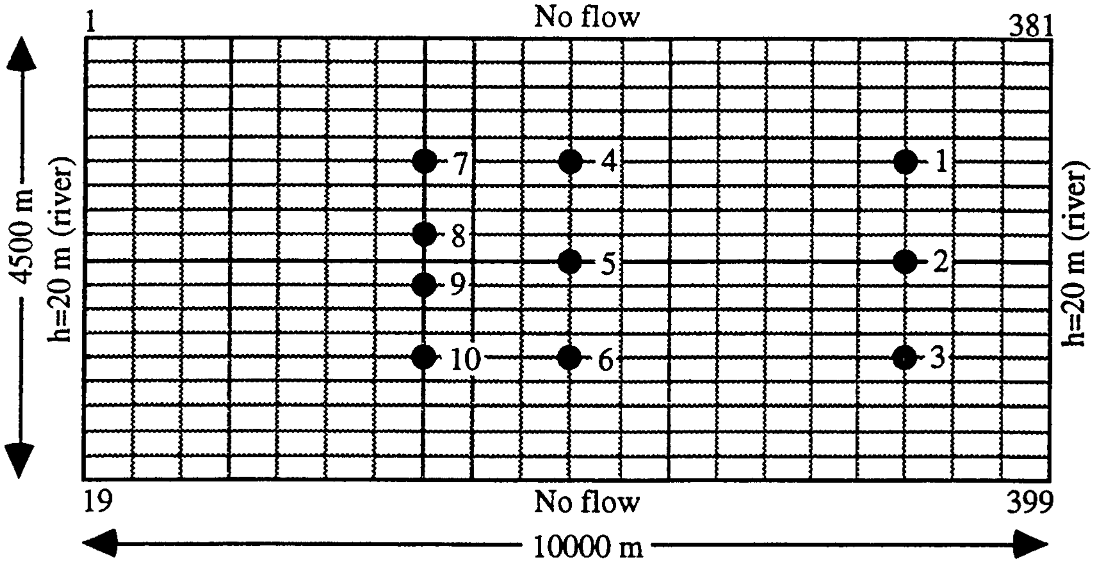


Figure 1. Two-dimensional unconfined aquifer used in this analysis (from McKinney and Lin, 1994).

GENETIC ALGORITHM

Figure 2 presents a flowchart for simple GA. GA encodes an initial population of N trial solutions using binary strings and then propagates this initial population through G generations. A standard generation includes three operations: selection, crossover, and mutation. Selection is performed using fitness values, and strings with high fitness are selected with high probabilities. Crossover and mutation are performed with probabilities of p_c and p_m, respectively, and these are implemented using a uniform random generator initiated by a seed, ρ. Finally, the near-optimal solutions are searched from N(G+1) solutions. A broad discussion on the topic is provided by Goldberg (1989).

```
1.  Acknowledge input: L, N, G, $\rho$, $p_c$ and $p_m$.
2.  Find $\Delta$ using L.
3.  Initialize population.
        • Encode the N binary strings bit-by-bit.
        • Decode the strings into decision variables.
        • Calculate the strings' fitnesses.
4.  Perform G standard generations.
        For i=1 to G
        • Selection  : Select the fittest strings.
        • Crossover  : Mate two strings with probability $p_c$.
        • Mutation   : Mutate each bit with probability $p_m$.
        Next i
5.  Provide output: Optimal solutions searched from $N \cdot (G+1)$ solutions.
```

Figure 2. Flowchart for simple GA.

GA selects a string based on its fitness. Fitness provides a measure of a string's performance, and high fitnesses are provided to feasible strings closer to the global optimum. Fitness is defined by combining the objective function with the constraints, and it provides GA with the capability of handling the constraints. Often, penalty method is used to define the fitness and this method is considered to be problem specific and difficult to implement (McKinney and Lin, 1994). In order to be general, a simple approach is adapted here. A feasible solution is provided with a fitness equal to the objective value, and an infeasible solution is provided with a fitness equal to zero. As such, the present approach completely ignores infeasible solutions, and appears good only provided GA encounters fewer infeasible solutions during the search.

STRING-LENGTH DEFINING METHOD

GA encodes **r** using binary strings. For the present problem, each string has ten substrings corresponding to ten wells, and the i-th substring decodes as

$$q_i = \Delta \sum_{j=1}^{L} B_{ij}\, 2^{j-1} \tag{5a}$$

$$B_{ij}=\{0,\ 1\} \tag{5b}$$

$$\Delta = \frac{w}{2^L - 1} \qquad (w = 7000 - 0) \tag{5c}$$

where L=substring length; B_{ij}=i-th substring's j-th bit value; and Δ=search step. Thus, GA search space contains 2^{10L} trial solutions. A large L increases the search space, and it induces a higher accuracy at the cost of increased computation. As such, there appears to exist an optimal L corresponding to an optimal accuracy-cost trade-off.

SLDM is an analytical method developed for finding an optimal L. Differentiating Equation (1),

$$dQ = \sum_{i=1}^{10} dq_i \tag{6}$$

Assuming the errors to be of the same order,

$$dQ = 10\,\Delta \tag{7}$$

where $\Delta = |dq_i|$. Now, dQ is defined as

$$dQ = (1 - \eta)\, Q_{max} \qquad 0 < \eta < 1 \tag{8}$$

where η=desired accuracy; and Q_{max}=maximum objective value. Taking η=0.99 and Q_{max}=70,000 (from Equations 1 and 3), Equation 8 yields dQ=700. Consequently, Equation 7 yields Δ=70, and Equation 5 yields L=7.

RESULTS AND ANALYSIS

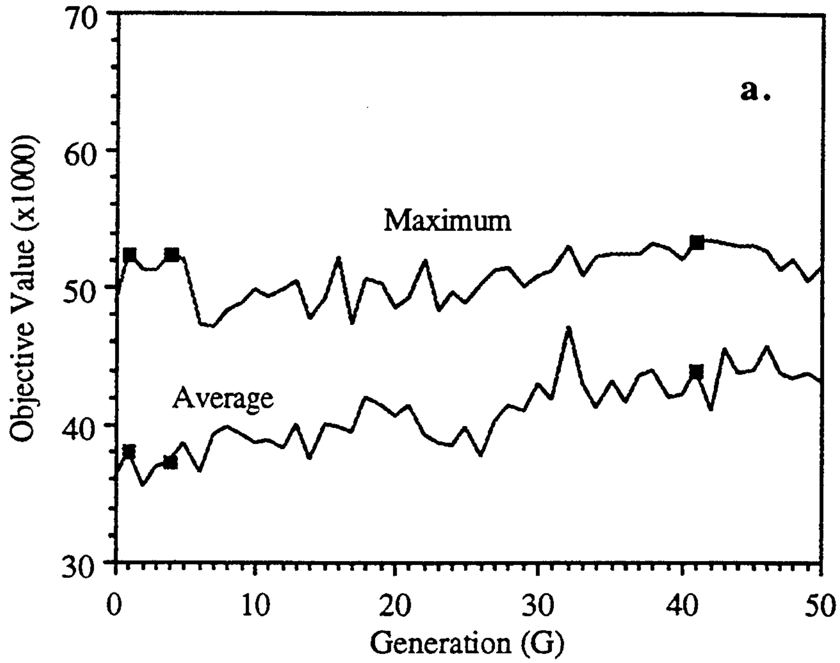
McKinney and Lin (1994) solved the present problem using both LP and GA, and Table 1 summarizes the LP and GA solutions in the second and third columns respectively. LP found Q_*=59,300 m^3/day (Q_*=global-optimal solution), and GA found Q_{10}=58,000 (Q_g=near-optimal

Table 1. Solutions for the base problem.

	McKinney and Lin (1994)		Present study	
Method	LP	GA	GA	GA
L	-	3	7	7
N	-	64	30	30
G	-	10	1	41
Well, i	Rate, q_i (m^3/day)	Rate, q_i (m^3/day)	Rate, q_i (m^3/day)	Rate, q_i (m^3/day)
1	7000	7000	4300	4000
2	7000	7000	5700	7000
3	7000	7000	4100	6200
4	6000	7000	6100	1900
5	4500	2000	6000	6400
6	6000	6000	5300	5800
7	6800	7000	6100	4600
8	4100	4000	4000	5700
9	4100	4000	4500	6000
10	6800	7000	6200	5800
Total (Q)	59,300	58,000	52,300	53,400
η_G	1.00	0.98	0.88	0.90

solution obtained at the g-th generation) attaining $\eta_{10}=0.98$ (η_g=accuracy attained at the g-th generation). GA was implemented with the tournament selection, L=3 and N=64, and no reason for the selection of L and N was discussed.

In this study, the problem is critically studied. The goal is to test SLDM and to standardize GA parameters. Here, GA is implemented with the roulette-wheel selection (Goldberg, 1989). Also, $p_c=0.6$ and $p_m=N^{-1}$ are adapted from DeJong (1975), and a base problem is defined with L=7, N=30, G=50, and $\rho=1$. L is estimated using SLDM, and N and G are arbitrarily selected. Figures 3a and 3b show the overall performance of GA in solving the base problem. Figure 3a shows the maximum and average objective values of the population at any generation. The maximum is observed to be much higher than the average, and the population has been able to maintain sufficient diversity for preventing premature convergence. Also, three solutions are marked in the figure. The best solution yields $Q_{41}=53{,}400$ attaining $\eta_{41}=0.90$ for the 50 generations, and the other two comparable solutions obtained during the early part of the run yield $\eta_1=0.88$ and $\eta_4=0.88$. Figure 3b shows the number of feasible strings contained in the population at any generation from a reference of 30, the maximum number of feasible strings corresponding to N=30. The figure shows the occurrence of few infeasible solutions throughout the 50 generations, and the simple approach to ignore these infeasible solutions appears reasonable. Table 1 summarizes two of the three solutions marked in Figure 3a. These two solutions correspond to G=1 and G=41, and are summarized in the forth and fifth columns of the table, respectively. Compared to the solutions of McKinney and Lin (1994), these two solutions are noted to be of slightly lower accuracy, and no attempts are made to improve their accuracy for generality. At this stage, it is interesting to examine the solution corresponding to G=1 critically. This solution is obtained with double the string length and one-twentieth the computational effort of McKinney and Lin (1994). Thus, for solving the groundwater optimization problem, GA appears to accommodate a larger L and less computation than that suggested by McKinney and



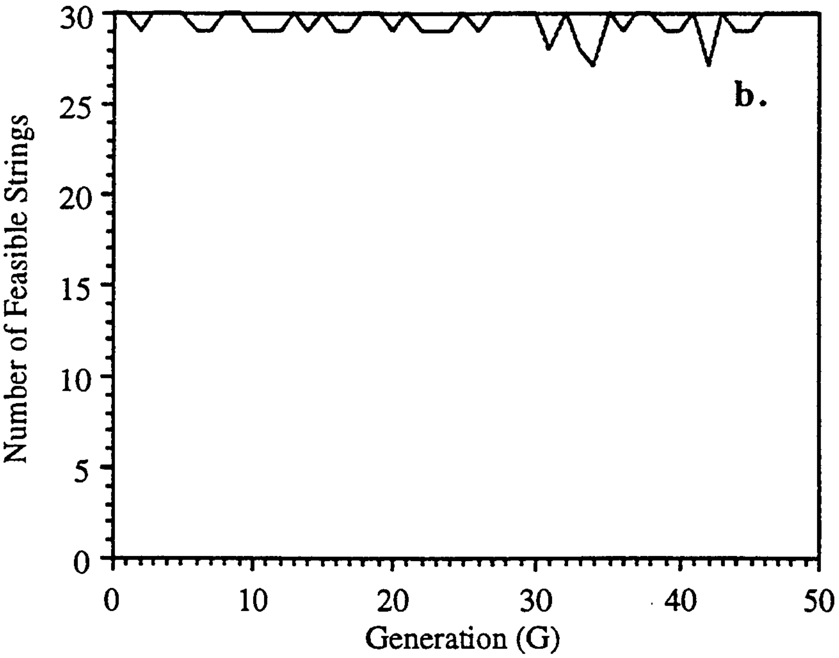


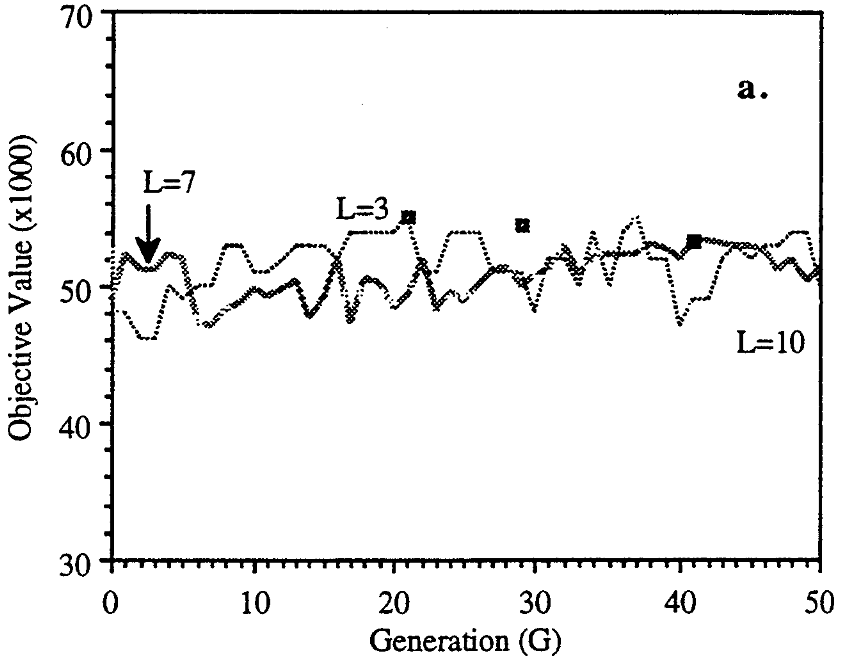
Figure 3. GA solution for the base problem with p_c=0.6, p_m=0.033, L=7, N=30, G=50, and ρ=1. (a) Maximum and average objective values over generations. (b) Number of feasible strings over generations.

Lin (1994). Having studied the overall behavior, the problem is now subjected to a sensitivity analysis.

First, the problem is solved for L={3, 10}. Figure 4a shows the effect of L on the maximum objective value of the population at any generation. The solid line shows the effect for the base problem with L=7, and the light- and heavy-dotted lines show the effects for L=3 and L=10, respectively. Although the effect of L on the performance of GA cannot be generalized, Figure 4a shows L=7 to be a reasonable trade-off between L=3 and L=10. GA finds Q_{21}=55,000 attaining η_{21}=0.93 for L=3, Q_{41}=53,400 attaining η_{41}=0.90 for L=7, and Q_{29}=54,500 attaining η_{29}=0.92 for L=10, and these solutions are marked in Figure 4a. As such, GA may be considered not to be sensitive around L=7 over the 50 generations, and SLDM appears to estimate L reasonably.

Second, the problem is solved for ρ={0, 5}. Figure 4b shows the effect of ρ on the maximum objective value of the population at any generation. The solid line shows the effect for the base problem with ρ=1, and the light- and heavy-dotted lines show the effects for ρ=0 and ρ=5, respectively. Theoretically, ρ should not affect the performance of GA; yet, Figure 4b shows some bias of the performance towards ρ. GA finds Q_{44}=55,900 attaining η_{44}=0.94 for ρ=0, Q_{41}=53,400 attaining η_{41}=0.90 for ρ=1, and Q_{18}=54,400 attaining η_{18}=0.92 for ρ=5, and these solutions are marked in Figure 4b. As such, the performance of GA may be considered not to be sensitive around ρ=1. At this stage, it is interesting to note the solution Q_{44}=55,900 for ρ=0; it suggests that the present solutions shown in Table 1 could have been improved simply by using a different ρ.

Third, the problem is studied for N and G. Obviously, T=N(G+1) (T=total number of fitness evaluations required in propagating N strings through G generations). As fitness evaluation



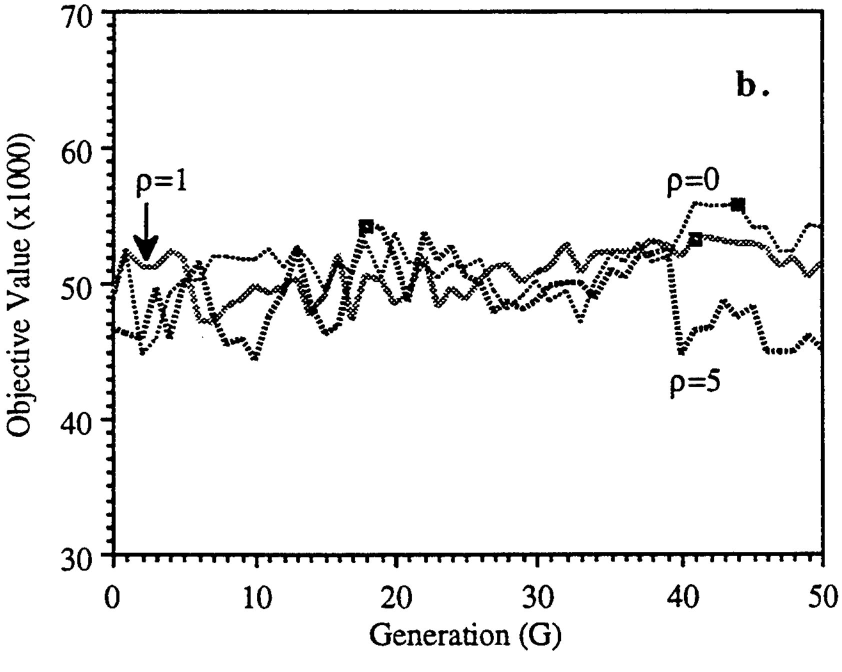


Figure 4. Sensitivity results for the base problem with p_c=0.6, p_m=0.033, N=30, and G=50. (a) Effect of L for ρ=1. (b) Effect of ρ for L=7.

involves the major computational component of GA, T may be used as an index to monitor the performance of GA. Figure 5a shows the minimum G required at any N to achieve $\eta_G \in [0.90,\ 0.93]$. The G-N relation is observed to be erratic, and G=5 for $N \in [50,\ 100]$ appears a good starting point. Figure 5b shows the minimum T required at any N corresponding to the G in Figure 5a for that N. Again, the T-N relation is observed to be erratic, and T=500 for $N \in [10,\ 200]$ appears a good starting point. Thus, an increase in N does not guarantee the convergence of GA, and a small N taken through more G provides GA with an opportunity to find the optimal solutions during the early part of the run. As such, GA is reduced to the selection of N only. The problem-specific response of GA may be embedded in the selection of N only, and the difficult task of finding an appropriate combination of GA parameters involving L, N, G, ρ, p_c, and p_m may be circumvented.

At this stage, the potential of GA in solving practical problems may be studied. Pumps are available only at certain discrete capacities, and their yields cannot be considered to be continuous. In order to restrict the yields to a set of discrete values, Equation (3) is replaced by the following equation

$$q_i^* \in \{0,\ 1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 7\} \qquad i = 1,\ 2,\ 3,\ \ldots,\ 10 \tag{9}$$

where $q_i^* = q_i/1000$. In general, GA requires minor modification to solve this discrete problem. GA should not decode the binary string codings using Equation (5a), but should decode these codings using a tag-table tagging each coding to a single discrete value of q_i. In Equation (9), q_i has eight possible values, and L=3 with 2^3 possible codings suffice for implementing GA. In fact, McKinney and Lin (1994) solves this discrete problem as an approximation to the original continuous problem. In contrast, linear programming (LP) needs extensive modifications to solve this discrete problem. LP requires a mixed-integer-programming (MIP) technique such as the

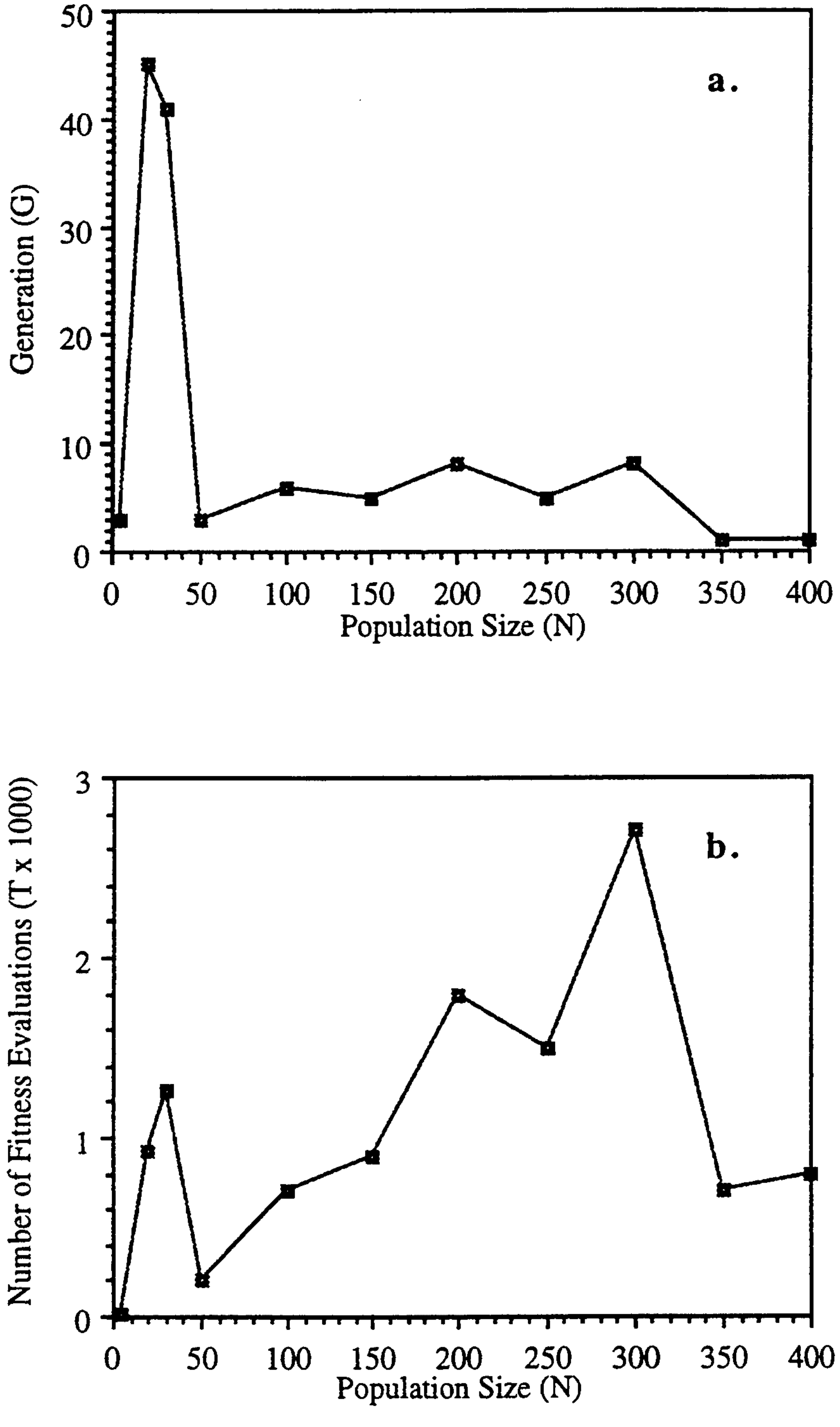


Figure 5. GA solution for the base problem with p_c=0.6, p_m=0.033, L=7, and ρ=1 observed at $\eta_G \in [0.90, 0.93]$. (a) Effect of G over N. (b) Effect of T over N.

branch-and-bound method. This method may lead to a complete enumeration of all possible combinations of the q_i-values to find the optimal solution, and the task can be gigantic. Thus, GA appears to be more robust than linear or nonlinear programming.

SUMMARY AND CONCLUSIONS

GA is considered to be robust for solving groundwater optimization problems with certain advantages and disadvantages. The advantages are: (i) GA imposes no continuity requirement on the objective function, and it can solve problems with discrete decision variables; (ii) GA usually does not become locked at a local-optimal solution, and it finds near-optimal solutions with high probability; and (iii) GA generates several near-optimal solutions, and it provides sufficient flexibility in decision making. In contrast, the disadvantages are: (i) GA does not guarantee global-optimal solution, and (ii) it requires too many fitness evaluations.

GA is found to perform robustly provided the constraints are properly handled and the parameters are properly estimated. Several conclusions may be drawn regarding the constraint handling and parameter estimation from the present study. These are: (i) If the probability of encountering infeasible solutions is small, the simple approach to ignore the infeasible solutions may be adapted, and the penalty method may be circumvented; (ii) SLDM may be used to define L. It automatically adjusts L depending on the accuracy specified; (iii) recommendations given by DeJong (1975) for $p_c=0.60$ and $p_m=N^{-1}$ may be used; and (iv) a large N does not guarantee improved performance of GA. As such, N and G satisfying T=500 and $N\in[10, 200]$ may be taken as a good starting point. If needed, problem-specific response of GA may be embedded in the selection of N only, and the difficult task of finding an appropriate combination of GA parameters involving L, N, G, ρ, p_c, and p_m may be circumvented. Thus, GA parameters are standardized reducing GA to the selection of N only, and a general guideline for solving certain groundwater optimization problems using GA is presented.

GA can only solve unconstrained maximization problem with positive fitness. As such, a minimization or a constrained problem needs to be modified as an equivalent unconstrained maximization problem before implementing GA. In Appendix A, these modifications are systematically discussed. In Appendix B, SLDM is generalized for solving groundwater management problems. In future, attempts will be made to extend the present guidelines for solving problems with tight constraints and fixed costs using GA.

REFERENCES

DeJong, K. A. 1975. An Analysis of the Behavior of a Class of Genetic Adaptive Systems. Ph.D. Dissertation, University of Michigan, Ann Arbor.

Goldberg, D. E. 1989. Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley, Reading, Massachusetts, 1989.

Holland, J. H. 1975. Adaptation in Natural and Artificial Systems. University of Michigan, Ann Arbor, Michigan.

Huang, C., and A. S. Mayer. 1995. Dynamic optimal control for groundwater remediation management in using genetic algorithms. Models for Assessing and Monitoring Groundwater Quality, Proceedings of a Boulder Symposium, IAHS Publication 227: 149-155, July.

McKinney, D. C., and M. D. Lin. 1994. Genetic algorithm solution of groundwater management models. Water Resources Research 30(6): 1897-1906.

Rogers, L. L., and F. U. Dowla. 1994. Optimization of groundwater remediation using artificial neural networks with parallel solute transport modeling. Water Resources Research 30(2): 457-481.

Willis, R. L., and W., W-G. Yeh. 1987. Groundwater Systems Planning and Management. Prentice-Hall, Englewood Cliffs, New Jersey.

A STATISTICAL PROCESS CONTROL SUCCESSFUL APPLICATION: DESIGN, IMPLEMENTATION, AND PROCESS IMPROVEMENT IN SEVERAL AREAS AT A CLOSING BASE

Roger G. Ford, Ph.D., P.E.
Associate Professor of Engineering
Department of Engineering

St. Mary's University
One Camino Santa Maria
San Antonio, Texas 78228-8534

Final Report for:
Summer Faculty Research Program
THE SA-ALC, Kelly AFB, TX 78241

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

THE SA-ALC, Kelly Air Force Base, TX

August, 1996

A STATISTICAL PROCESS CONTROL SUCCESSFUL APPLICATION: DESIGN, IMPLEMENTATION, AND PROCESS IMPROVEMENT IN SEVERAL AREAS AT A CLOSING BASE

Roger G. Ford, Ph.D., P.E.
Associate Professor of Engineering
Department of Engineering
St. Mary's University

Abstract

The repair/refurbish/remanufacture environment of the repair depot at Kelly Air Force Base is a non-traditional manufacturing operation aimed at the maintenance of the Air Force's aircraft. Since maintenance of used aircraft entails the disassembly of the aircraft and components of the aircraft, inspection of the parts for wear replacement or repair, refurbishment and reassembly, testing, and return to the users, the variability of the functions necessary to repair the aircraft is high. With small lot sizes, high variability of the work load, and long cycle times, traditional use of statistical process control is very difficult or inappropriate. The assessment of statistical process control in this unique manufacturing environment was the aim of this project.

The use of Statistical Process Control in the San Antonio Air Logistics Center was investigated as to the present use of the discipline as of the summer of 1995. The summer of 1996 was an extension of that effort with the additions of two more area in one Directorate and another area within an additional Directorate.

Results are presented for the original Statistical Process Control study on-going until the present. The results of that study are quantified and analyzed. Personnel response to the study are added. The new areas that were analyzed with this year's study have data presented. Conclusions and recommendations are included.

A STATISTICAL PROCESS CONTROL SUCCESSFUL APPLICATION: DESIGN, IMPLEMENTATION, AND PROCESS IMPROVEMENT IN SEVERAL AREAS AT A CLOSING BASE

Mark L. Smith
Aerospace Equipment Directorate Quality Advisor
Kelly Air Force Base

Roger G. Ford, Ph.D., P.E.
St. Mary's University
Air Force Office of Scientific Research (AFOSR)

INTRODUCTION

The Air Force has introduced the philosophy of Quality Air Force (QAF) (1) which is based upon the principles of Total Quality Management (TQM) (2). Many corporations have enjoyed a great deal of success using the TQM philosophy (3), so the military is attempting to repeat the success of these private corporations. The Commodity Production Division (CPD) within the San Antonio Air Logistics Center has been following the QAF philosophy for the past three years. On the 22nd of June, 1995, the San Antonio Air Logistics Center (SA-ALC), which employs over 11,000 civilian people (15), was targeted for closure by the Base Realignment and Closure (BRAC) Commission. On the 11th of July, 1995 the President of the United States approved the base closure list recommended by the BRAC commission which included Kelly Air Force Base. This will be the largest base closure in Air Force history. SA-ALC is the largest employer in San Antonio and south Texas. With the advent of closure looming ahead, the people assigned to SA-ALC face many challenges.

According to Major General Lewis Curtis III, the Air Force is planning to privatize the maintenance workload currently being performed at Kelly Air Force Base beginning in early 1997 (6). He said the goal is to place as many current SA-ALC maintenance workers into jobs with private defense contractors as possible. The CPD division was looking for ways to improve efficiency and effectiveness, so that a private defense contractor would want to hire the present workers with the current workload rather than replace these workers (16). All of the maintenance work that is performed by the workers at the SA-ALC is ongoing and still needs to be supported, even though the base is facing closure. All of the customers that receive services from SA-ALC still require the services. The CPD intends to use SPC tools as a means of achieving these goals. The CPD wants to introduce Statistical Process Control (SPC) tools popularized by Deming (8). The purpose of these tools in the CPD division is to highlight potential problem areas in the process of assembling aircraft parts. The goal is to improve this process and provide the best quality and price of parts possible to the customer.

This study deals with the situation of implementing SPC in four maintenance shops - the Jet Fuel Starter (JFS), Central Gear Box (CGB), the Aircraft Mounted Accessory Drive (AMAD), and the Gas Turbine Engine (GTE) - and the beginning of the implementation of SPC in the Engines Directorate on the F-100 engine. The maintenance shops are under the Secondary Power and Starter sections of the Secondary Power and Gas Turbine Engine Branch employing 232 personnel. The value of this study is in the insight it can provide to top management officials at the SA-ALC on whether SPC could be used in other maintenance organizations at the logistics center (5). The annualized savings of the JFS and CGB application begun in the summer of 1995 and continuing at the present is over $280,000.

SPC USE AT SA-ALC

The use of SPC has been successful in many organizations (12). It is crucial that SPC tools measure real problems and the organization act upon the results. For TQM to work, it needs to be fully supported by top management (4). The literature discusses several elements that needs to be in place for TQM to work, and several reasons why TQM programs fail.

There have been several studies done about the implementation of TQM in the military(3)(9)(11)(13). The military by its very nature is bureaucratic and chain of command oriented. There are many challenges for a military organization to face when implementing TQM principles. There have also been many military organizations going through downsizing and closure. Several studies in the literature were cited that address the effects of downsizing on department of defense personnel (7)(14). However, no study was found where an attempt was made to implement SPC at a military base that is closing.

STEPS IN IMPLEMENTATION OF SPC

The steps in the implementation of the SPC study can be traced through the following table:

Steps	**Results**
Step 1. Meet with Management	**Gain approval for study from the top**
Step 2. Design data gathering tools	**Forms and procedures developed for data gathering purposes**
Step 3. Begin collecting data in the shops	**Mechanics record data for the study on a daily basis**
Step 4. Gather data from shops	**Data to be assembled on a monthly basis**
Step 5. Assemble reject analysis team	**Determining who needs preparation for the team meetings**
Step 6. Prepare data for the team meeting	**Present data so that problems identified can be discussed**
Step 7. Team meeting on a monthly basis	**Problems given to the responsible people for resolution**

By following these steps, the problems identified as to kit nonconformities and possible reject causes were brought to the attention of the reject analysis team. By opening discussions on these problems, resolution was achieved on most of them.

In addition to the SPC analysis done in the Summer of 1995 - and continued through to the present - to determine if the JFS and CGB assembly operations were in control or not, four additional sets of data were collected for this study: the average number of rejects per kit per month for the JFS and CGB and the number of end item rejects per month for the JFS and CGB. The JFS and CGB were looked at separately to determine whether significant improvements had occurred in each of the areas. The additional data were tracked, from September 1995 to February 1996, to determine if there was a positive trend in the reduction of defects in a kit and in the number of end item rejects.

NEW APPLICATIONS OF SPC FOR THE SUMMER OF 1996

The AMAD assembly is similar to the JFS and CGB with the same supervision. The AMADs were added this summer to the analysis by SPC in the same manner as the JFS and CGB because of the success of their analyses and cost savings. The GTE assembly area was also added this summer to the analysis for the same reasons. All four of these areas are within the same building and lend themselves to simultaneous scrutiny and data gathering.

There were several opportunities to report on the successful implementation of SPC analysis in the JFS and CGB operations during the summer of 1996. The presentations were made to Quality Councils of two Directorates, before both Directorates' management, and in front of the commanding General, Major General James Childress, of Kelly Air Force Base. As a result of these meetings and General Childress' desire to expand the SPC effort elsewhere at the base, the Engine Directorate was targeted for implementation of the same SPC analysis successful in the Secondary Power Division. Care had to be taken in the choosing of the many different possible areas where SPC could be used. There is an inherent danger in the forced application of SPC in an area where it may not have a significant impact. That would result in the loss of reputation of the technique and diminish its effectiveness in the future at Kelly AFB.

The task of choosing an appropriate area in which to implement SPC in the Engines Directorate was a difficult task. During the summer of 1995, several areas were identified as potential application areas for SPC within the Engines Directorate. That list was the beginning point to determine the first installation of SPC in the Engines Directorate. Exhaustive interviews and evaluations were accomplished in each of the areas identified. The F-100 engine, which powers the F-16 and F-15 aircraft, was chosen for the study. The engines come to Kelly as components for refurbishing. The main engine component, the compressor core and high pressure turbine, is disassembled, inspected, cleaned, refurbished, replacements made, and reassembled. If parts or components are not available when needed for reassembly, that constitutes a nonconformance. These nonconformities can then be tracked and charted using a "c chart" as well as the causes for the nonconformity. The causes can be examined with Pareto analysis which identifies the most egregious problems first or those that are causing the most trouble are identified and solved first. Through this on-going process of tracking nonconformities and solving Pareto analysis identified problems, over time the system will gain in efficiency and throughput time will decrease.

DATA RESULTS

The CGB and JFS statistical process control data can be compared from the Summer of 1995 through until the Summer of 1996. Figures 1,2,3 and 4 show a comparison of the data for the CGB and JFS and the improvement seen from the meetings with pertinent personnel.

Figure 1

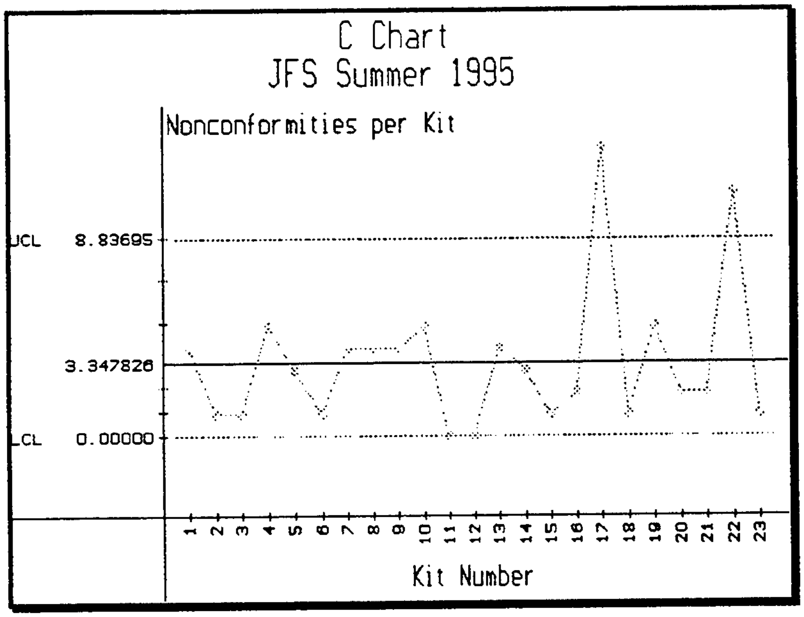


Figure 2

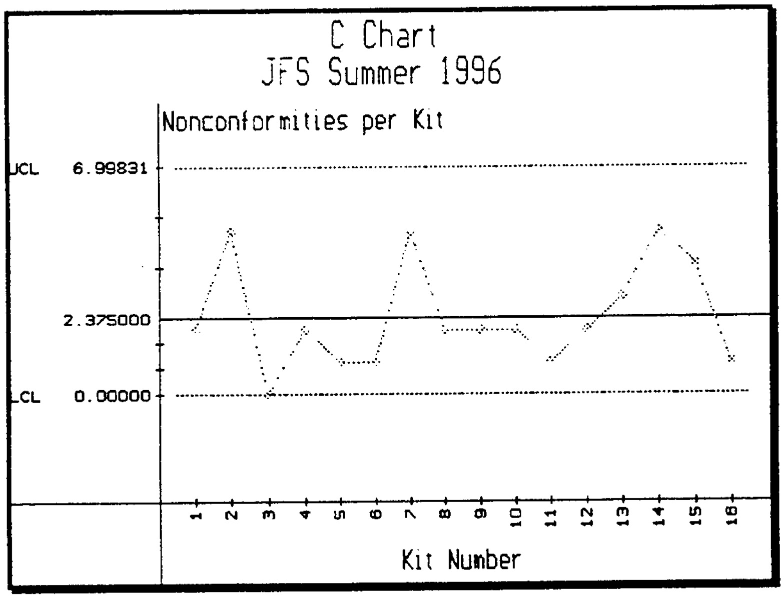


Figure 3

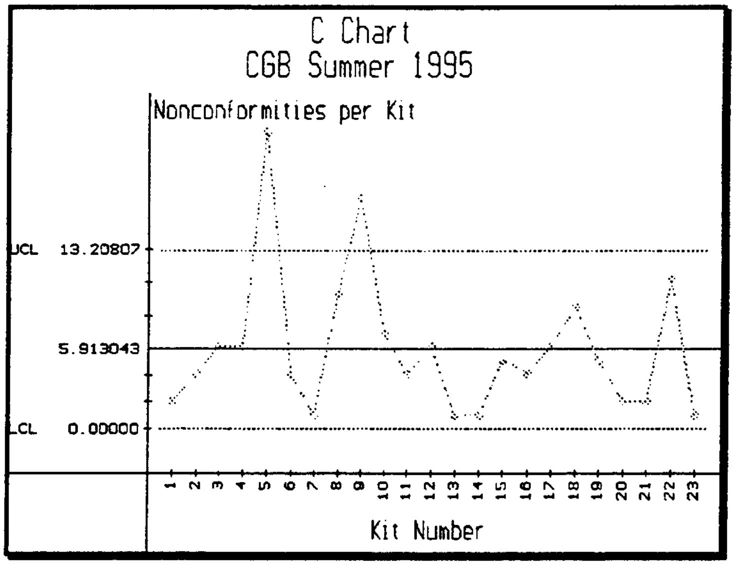
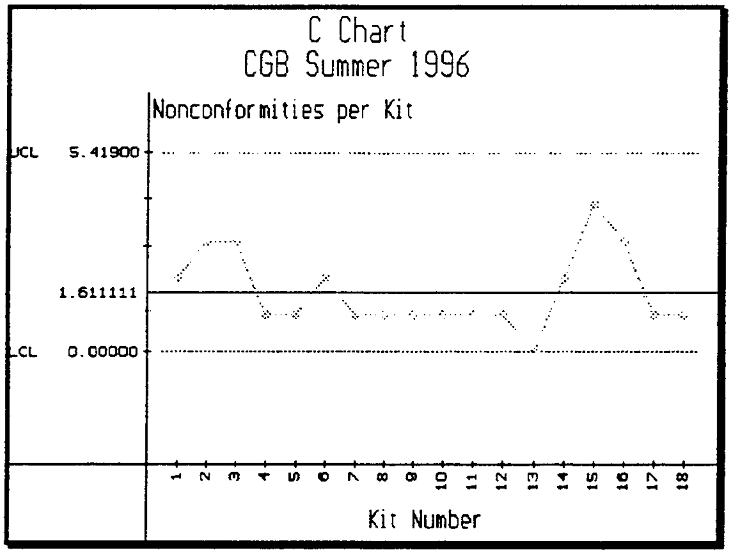


Figure 4



The Pareto analysis data shown in Figures 5 and 6 was used to solve the problems of the nonconformities at the source instead of dealing with the same problems each time that they arose. As can be seen from the figures, a significant amount of control has been gained in one year.

Figure 5

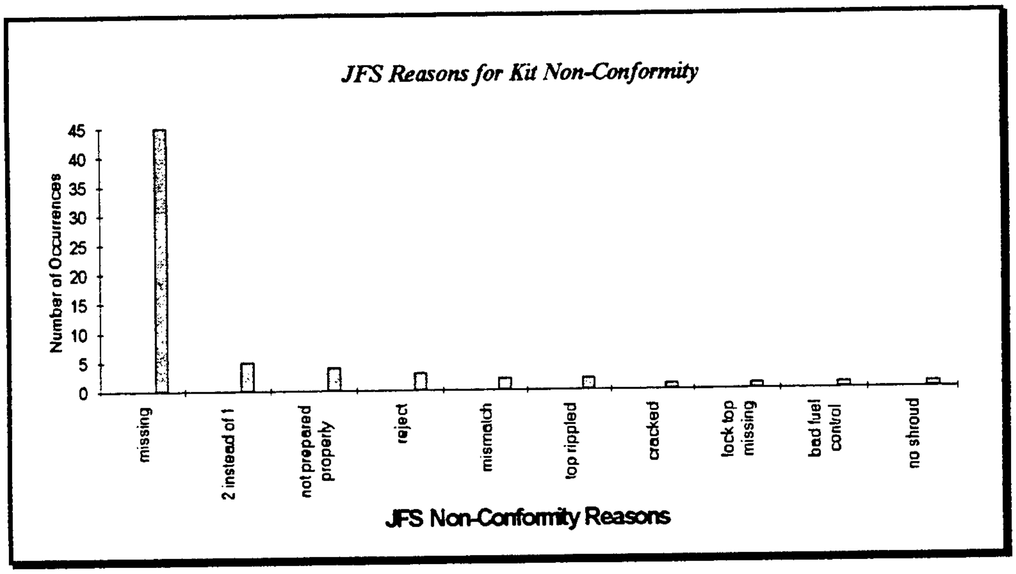
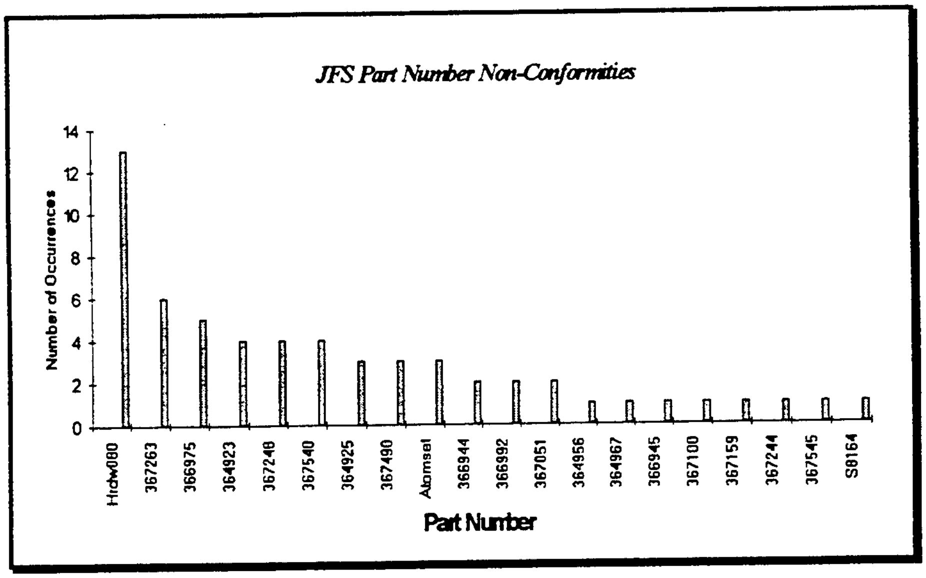


Figure 6



The additional data on the CGB and JFS kits that was monitored was the total number of rejects per month and the number of rejects per kit. For both the JFS and CGB there were significant decreases between the first and last month of the study. The number of rejects per kit for the JFS was reduced from 4.04 per kit to 1.47

per kit. The number of rejects per kit for the CGB was reduced from 3.68 per kit to 2.38 per kit with one month having only 1.58 per kit. The percentage of parts rejected per kit for the JFS reduced an average of 16.58 percent per month. The percentage of parts rejected per kit for the CGB was reduced an average of 2.89 percent per month. A detailed analysis was conducted as to the amount of money has been saved through these reductions in rejects. When time savings, wage rates, overhead rates, and benefit rates are taken into consideration, savings amount to over $280,000 annually.

The initial SPC analysis for the GTE and AMAD assemblies are presented in Figures 7 and 8.

Figure 7

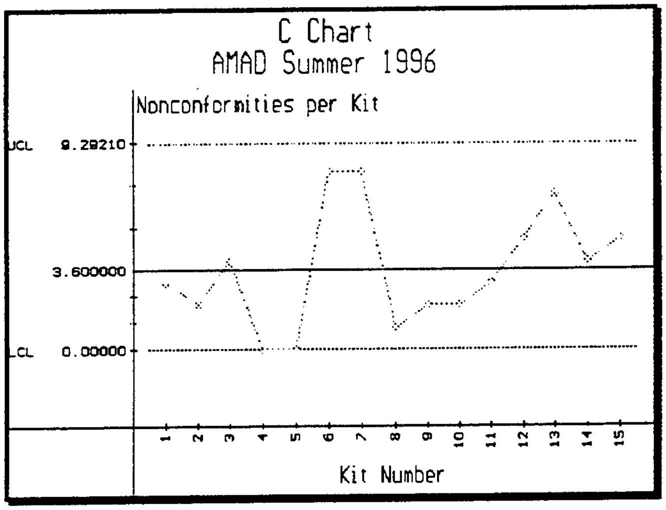
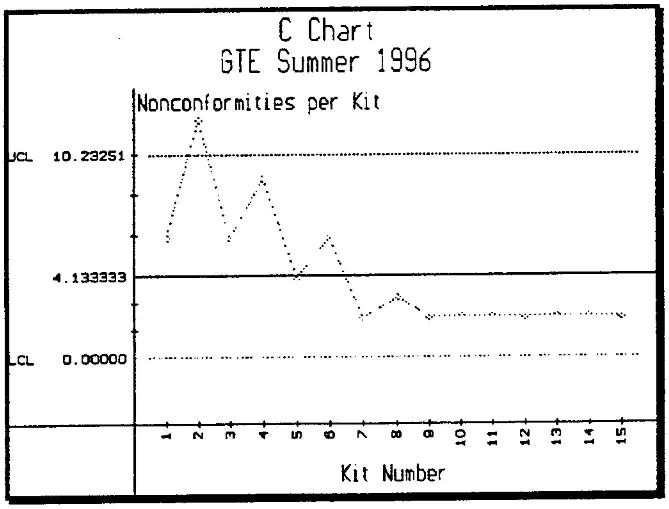


Figure 8



As can be seen, the AMAD data is in control probably due to the fact that the same supervision is in place for the AMAD area as for the CGB and JFS. But, the on-going analysis of the AMAD data is now in place, so further improvements are expected. The GTE area showed a marked decrease in nonconformities per kit after a short time of data gathering. This could be attributed to the unfamiliarity of the mechanics to the process and a possible reluctance to reporting what they may think are problems that they will be held accountable for. With future diligence, these problems can and will be worked out.

MECHANICS RESPONSE TO STUDY

An analysis of the CGB and JFS mechanics opinion to the SPC study was conducted as a way of evaluation of the impact of the technique on those closest to the work being performed. The majority of mechanics felt that they had noticed a difference in their jobs as a result of measuring rejects. There has been less robbing of parts, and fewer parts are being rejected. Most of them felt that the kits still need to be improved, but that there has been significant improvement. All of the people that were interviewed felt that this type of effort using SPC would be beneficial for other areas in the division, as well as the rest of the maintenance activities on base. Many of them felt that measuring reject rates was a good indicator of the maintenance areas performance, and that it was possible to solve these problems. Three of the mechanics thought it was beneficial to have a person assigned outside of the shop area to do the data collection because of objectivity. Another theme among the people interviewed was that management needs to support such an effort and act upon the data when it is presented. Many felt that was what helped this effort and would be crucial in any other attempts to implement this process.

People were divided in whether the base closure process had an affect upon the results of this study. Some of the mechanics felt that most people still had the sense of pride to do a good job and work hard despite the base closure. Another common theme was that as a result of the closure many people are leaving and/or getting moved between organizations necessitating that these people be retrained (14). They felt that this factor could have a negative impact on this study.

CONCLUSIONS

In the CPD, the specific intent was to determine if the implementation of SPC in the JFS and CGB shops could be used to reduce the number of parts kit rejects and end item rejects. If there was a significant reduction in both JFS and CGB kit and end item rejects, the CPD wanted to expand the use of SPC tools to other areas of its maintenance operations. There was a 62.38 percent reduction in the number of parts kit rejects per month for the JFS and a 33.4 percent reduction in the number of parts kit rejects per month for the CGB. There were a few reasons for these reductions.

First, this was the first time that these problems were ever tracked and brought to the attention of branch level in the CPD. In this study, problems were brought directly to the individual supervisor over the area involved, during the process of kit assembly. Thus, real problems were brought to the attention of people who could resolve them.

Second, there was tremendous support for this effort from the CPD management involved in this study. When the authors presented data that pointed out problems with the kits, the supervisors over the areas responsible for those problems took responsibility and worked hard to get the problems resolved.

Third, the team leader held all team members accountable for resolving all problems identified in the data. Without the acceptance of responsibility for the problems pointed out in the data by the reject analysis team leader and all members of this team, the improvements that occurred during this study would not have been possible.

Management support is crucial to the success of implementing SPC in any area. Many of the JFS and CGB mechanics interviewed expressed the belief that this effort was successful because of management support and the willingness of the supervisors in the affected areas to act upon the information. The data that was presented during the reject analysis team meetings served as a catalyst for open communication between the supervisors on the reject analysis team. Many problems were resolved as a result of this open communication which did not take place in the past.

There was also a significant reduction in the number of end item rejects. The average number of end item rejects for the JFS decreased from 11.6 per month before the study to 8.13 per month during the study, a 29.91

percent improvement. The average number of rejects for the CGB decreased from 17.7 per month before the study to 7.9 per month during the study a 55.36, percent improvement. There were several reasons for these improvements.

First, there was an improvement in the parts in the kits, resulting in better parts going into the end items. Second, the end item tester who was interviewed stated that he noticed an improvement in how the mechanics put the end items together. When the mechanics received better parts they spent more of their time putting end items together, versus correcting problems. Also, another point borne out by the mechanics during the interviews was that it was important to have a person outside the CPD division to collect and present the data. They said that in the past, when they would bring up problems, the were met with denial and resistance by the people who could have corrected the problems. They said having a neutral party collect the information eliminated infighting within the CPD and helped all people involved in putting the JFS and CGB kits together to focus on correcting and solving problems. Additionally, having a neutral party prevented the mechanics from spending time collecting and presenting all of the data, allowing the mechanics to focus on correcting and solving problems. Also, when several of the mechanics were interviewed they expressed the desire to perform an outstanding job. There is still a deep sense of pride among most of the mechanics and a desire to make improvements to their work and put together quality end items. This is despite the hardships many of them were going through as a result of the logistics center being closed.

As a direct result of these very positive results, the SPC effort was extended to the GTE and AMAD areas within CPD. Also, because of the results gaining praise all the way to the top, the SPC effort was expanded into the Engines Directorate. At the present time, the study is being implemented there with results to come. The management meeting members are being determined in order to evaluate the Pareto analysis and SPC charts for the F-100 Core.

RECOMMENDATIONS

There are seven recommendations as a result of this study. First, there are some valuable insights and lessons learned that could be applied across the entire SA-ALC as well as the other ALCs. Many of the other maintenance activities at the SA-ALC have processes where SPC could be applied. These were identified by Ford (10). The successful implementation of SPC tools would depend heavily upon management support in these areas to take on a similar effort to this one.

Second, in the future implementation of efforts similar to this one, each reject should be weighted as to its' significance and importance. During the study, each reject was counted as one. The authors discovered that some rejects were more significant than others. For example a bolt with paint on it counted as a reject even though the mechanics had plenty of bolts at their workstation to use. This type of reject created minimal problems for the mechanic in putting the end item together. However, a part that had a large crack posed a major problem for the mechanics, especially if it was a difficult part to replace. Yet it was counted the same as a minor problem. A system where a weight is assigned to each reject could help a team, using SPC to solve problems, focus on the more significant problems first.

Third, that in the future implementation of this type effort, all of the actions taken be communicated to all of the people involved in the process. This was not done until the third month of this study. This is important because the mechanics putting the end items together and the people involved putting the parts kits together need to know that their efforts are valuable and that inputs given by the mechanics are being acted upon. This will help continue their "buy-in" and support for this type of effort. Without their hard work and support an effort like this can not work.

Fourth, is that when a team like this makes significant improvements in the future, these efforts need to be recognized. The CPD division was the only division within the Aerospace Equipment directorate that did not recognize an employee or team of the quarter. The reject analysis team made some very significant improvements, but there was no formal process in place in the CPD to recognized these efforts. Future recognition of efforts where improvements are made will go a long way toward ensuring future improvements.

Fifth, an article should be drafted for proposed publication in the Kelly Observer, which is the weekly newspaper of Kelly Air Force Base. The results of this study were very positive and it would be a means for people in other organizations to learn about these efforts. Additionally, it would provide positive publicity about work being done at the SA-ALC.

Sixth, prior to a team meeting using SPC tools, all members should be trained on how to use and apply these tools. This training needs to be tailored to help the team members solve problems that it is working on. This training needs to have clear and specific terminal performance objectives and enabling objectives, and the learning from the training needs to be applicable to the work the team members will perform. This would help ensure the success of other future attempts to improve processes.

Seventh, is to let all parties know at the outset of this type of effort that there will be a great deal of hard work and that improvements will take time. There are very few overnight fixes, and that with patience and hard work positive results will follow. During this study, many of the JFS and CGB mechanics expected instant improvement. But once they understood that improvements take time, they had more support for this effort.

RECOMMENDATIONS FOR FURTHER STUDY

First, study the expansion of this effort to other areas in the CPD as well as other Directorates such as the Engines Directorate. Second, this effort should be replicated and studied in other Department of Defense maintenance organizations to determine where this type of effort is transferable. Third, this effort should be replicated and studied in the private organizations that take over the JFS and CGB workload. The results of such a study would be beneficial to private organizations attempting to use SPC tools after taking over a maintenance workload from the government.

REFERENCES

1. Air Force Quality Institute (1993). A guide to implementing quality air force principles. Government Printing Office.
2. Air Force Quality Institute (1989). Air force facilitator training manual. Government Printing Office.
3. Bailey, Marshall H. III (1993). Public administration efficiency through total quality management (quality management), George Mason University: Fairfax, VA.

4. Boyle, Ruth F. (1993). Attitudes and commitment of the top management officials toward total quality leadership (quality management), University of La Verne, San Diego CA.
5. Coates, James E.(1988). Statistical Process Control: Ignore it at your risk, Productivity Management, Vol. 19 No. 3 Jul-Aug 1988 pp. 3-5.
6. Day, Ross (1995). General Curtis addresses Kelly air force base workers, Kelly Observer, July 7, 1995 Vol. 30 No.31.
7. Demarie, Samuel M. (1996) Strategic issue interpretation and organizational action: An investigation of organizational downsizing (organizations, restructuring), Arizona State University: Tempe, AZ.
8. Deming, W.E. (1982). Quality without tears: The art of hassle free management, New York: McGraw Hill.
9. Denega, Peter (1992). Total quality management: are the principles associated with participatory management being internalized and practiced by United States Air Force organizations?, University of Alabama: Tuscaloosa, AL.
10. Ford, Roger (1995). Use of statistical process control in a repair/ refurbish/remanufacture environment with small lot sizes. Air Force Office of Scientific Research.
11. Hammons, Charles and Maddux, Gary A. (1995). Total quality management in the public sector, Quality Progress. pp. 41-45.
12. Kumar, Jamal and Gupta, Ron (1993). Statistical process control at Motorola's Austin assembly plant, Interfaces. pp. 84 Vol. 7 No.3.
13. Marshall, Terry S. (1993). The impact of quality management on a United States Air Force Strategic Logistics Group: Gonzaga University: Spokane, WA.
14. Owen, Sharon L. (1994). An assessment of senior-grade civil servant perceptions of an organizational reduction within the defense department (downsizing, layoffs), Walden University: Washington, DC.
15. Sanchez, Robert (1995). Base realignment and closure commission closes Kelly, San Antonio Express News. June 23, 1995.
16. Shiang, Jing (1995). Successful privatization: Contracting policies and performance of public service contractors, Ohio State University: Columbus, OH.

VOC Emission Control by A Gas Turbine Combustion System
-Economic and Technical Evaluations

Garng M. Huang
Professor
Department of Electrical Engineering

Texas A&M University
College Station, TX 77843

Final Report for:
Summer Faculty Research Program
Kelly Air Force Base, San Antonio

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Directorate of Environmental Management
Kelly Air Force Base, San Antonio

August, 1996

VOC Emission Control by A Gas Turbine Combustion System -Economic and Technical Evaluations

Garng M. Huang
Professor
Department of Electrical Engineering
Texas A&M University
College Station, TX 77845

Abstract

On July 31, 1995, the Aerospace NESHAP was officially released and was officially published on Sept. 1, 1995. The rule is that the industry has to either use compliant coatings or use a compliant facility that can reduce the VOC emissions by 81% by Sept. 1, 1998. This implies that by Sept. 1, 1998, not a single plane can be painted unless the hangar meets the definition of a compliant facility or other coating techniques that emit VOCs less than the NESHAP standards are employed. It is then urgent that the industry find a cost effective way to solve this problem.

The traditional and commercially available emission control technologies for VOCs include process modification, solvent reformulations, and add-on control systems that incinerate or trap the VOCs. When cost effective process and solvent changes are not available, more costly add-on controls are the only alternative. These systems can be a significant economic burden since the flow rate of the VOC-laden air stream can be quite large, and with the concentration of the VOCs being quite small, the economic value of recovering the VOC solvents is negligible.

This study concludes that:

1) Gas turbine cogeneration is the most effective and efficient VOC control technology since:

- it destroys the pollutants. Many other technologies, such as carbon adsorption, only transform the pollutants into other phases. To eliminate the pollutants, it takes additional energy to either recover or consume them.
- it generates power and steam in the process of destruction of VOCs. The other technologies use extra energy to destroy the VOCs.
- when combined with recirculation, the VOC concentration can be increased to further reduce the energy consumption by 15%.
- the initial capital cost is higher than the other technologies. However, the generated power and steam produce positive investment returns. Note that other technologies involve additional energy consumption and waste disposal, which can be as high as their installation costs. However, the capital cost is associated with energy production; the cost associated with VOC control is actually negative due to the contributed VOC heat value.

- larger cogeneration units are more efficient but require a much larger capital investment. If the unit is oversized for the energy load requirements one is faced with selling excess power or operating at part load and reduced operating efficiencies. For areas that need additional power this can be an advantage; however in areas where excess power is available this can be a disadvantage.

2) One should use predicted demands of electricity and steam as major factors to decide the gas turbine size. VOC destruction is a convenient byproduct of cogeneration, but it should not be used as a deciding factor for cogenerator size.

3) Gas turbines cannot tolerate chlorinated pollutants, which will shorten the engine life cycle. This concern is predominately for the methylene chloride based paint strippers that have been the industry standard. This may be an issue for paint facilities that strip and paint in the same facility and will need to continue using methylene chloride base strippers. However, this restriction also applies to many other control technologies.

4) Gas turbines cannot tolerate a high density of particulates, which will also reduce the engine life cycle. Based on the data from Acurex for Armstrong Laboratory, this does not seem a problem for a standard aircraft paint facility due to the particulate filter system. Note that this restriction also applies to almost all other technologies.

5) The only other control technology that may destroy the pollutants as effortlessly is biofiltration. However, the technology is still not yet mature. Its cost and effectiveness need to be further investigated.

VOC Emission Control by A Gas Turbine Combustion System
-Economic and Technical Evaluations

Garng M. Huang

I. Introduction

Volatile organic compounds (VOCs) are organic compounds that participate in atmospheric photochemical reactions to form ozone. Many VOCs are also toxic and/or are listed hazardous air pollutants (HAPs). The Clean Air Act Amendments of 1990 has a substantial impact on all painting facilities, including aerospace coating facilities. On July 31, 1995, the Aerospace NESHAP was officially released (Lord, 1996) and it was officially published in the Federal Register on Sept. 1, 1995 (NESHAP, 1995). The rule is that the industry has to either use compliant coatings or use a compliant facility that can reduce the VOC emissions by 81% by Sept. 1, 1998. This implies that by Sept. 1, 1998, not a single plane can be painted unless the hangar meets the definition of a compliant facility or other coating techniques that emit VOCs less than the NESHAP standards are employed.

To meet environmental regulations for VOCs, the aerospace coating industry must rely on available control technologies. Currently, thermal and catalytic incineration and adsorption are three commonly used controls in surface coating operations. However, all these add-on controls are an extra expense to the industry. In addition, they have no economic payback as they either consume extra energy to burn the VOCs and/or only transform the pollutants into another phase that results in wastes that must be processed or disposed. They have high capital and operating costs and can become major economic components of the surface coating operation.

In this project, an alternative VOC emission control is explored, one which uses the combustor of a revenue-producing energy system for VOC thermal destruction. Note that the solvent-laden air (SLA) has a low heating value due to the requirement that the VOC concentration cannot exceed a fraction (below one quarter) of the lower explosive limit (LEL). However, the capital and operating costs of the thermal destruction can be justified by the revenues from the energy output from the system. In addition, gas turbine generating units have been manufactured in large quantities all over the world due to their efficiency, clean combustion process and stable low priced fuel. Also the power deregulation act entitles power producers to have the right to access the power transmission lines for power wheeling. This implies the cogeneration unit can produce revenues by selling excess power.

There has been little operating experience using gas turbine equipment as a VOC control device. Technical and economic issues need to be investigated to demonstrate its feasibility. Technical issues include airstream (paint booth exhaust) cleanup requirements to avoid deposition or erosion to the turbine, and the ability of the turbine combustor to destroy VOCs efficiently. Economic issues include the capital and operating costs compared to traditional VOC control techniques and the effect of gas and electricity rate structures on the economic feasibility of the cogeneration approach. This study will clarify these issues and identify issues for further investigation. In section 2, we will sum up the benefits and disadvantages of traditional VOC control technologies. In section 3, the focus will be on the technical feasibility of the gas turbine approach. Section 4, will investigate the split-flow recirculation approach and its potential to reduce the costs of add-on controls. Investigation of the economic issues in section 5 followed by a few examples that will demonstrate the overall approach for cost estimates in section 6 with conclusions in section 7.

II. Some Traditional VOC Control Technologies

1. Thermal Oxidizers:

Process Description:

VOC-laden gaseous vapors pass through the oxidizer, which operates at a temperature of 1400F-1500F with a residence time of 0.3 - 0.5 seconds. For most solvents, a thermal oxidizer with a residence time of 0.5 seconds at 1500F will achieve a destruction rate in excess of 98%. The products of combustion are H_2O, CO_2, CO and NOx. As with all other control technologies, the incoming air stream must be filtered to remove particulates, such as paint overspray that has escaped from the exhaust duct of the spray booth. Failure to remove the particulates will shorten the life cycle of the thermal oxidizer.

Advantages:

Probably the most efficient system (>98%) for destroying VOCs.

Direct flame operation ensures destruction of longer chain solvent molecules.

Low maintenance.

Exceptionally long operating life.

Considered by the South Coast Air Quality Management District as the most efficient BACT (Best Available Control Technology) device.

Disadvantages:

High capital and operating (fuel and electricity) costs. Operating cost can be reduced by adding a heat exchanger or heat recovery feature; however, this raises the capital cost.

Produces NOx (around 45 ppm at 1450F).

The products of combustion of halogenated compounds can be extremely toxic. These include phosgene, hydrochloric and hydrofluoric acid, dioxins and furans. Hot halogenated compounds can be extremely corrosive and can corrode construction material such as the metal shell and ducting of the incinerator.

2. Catalytic Incinerators

Process Description:

Like the thermal oxidizers, the vapors pass through the system at an elevated temperature at which combustion takes place. However, the catalyst initiates and accelerates the combustion at much lower temperature, 500-800 F; it is flameless. The residence time remains the same as 0.3-0.5 seconds. The combustion products are the same though vary in their ammounts. The catalyst consists of either a precious metal (Platinum, Rhodium, etc.) or a base metal (active alumina), which is dispersed on a high surface area ceramic. It can have a honeycomb structure to increase available surface area, or can be in the form of pellets. Particulates will also shorten the life cycle of the catalysts.

Advantages:

Probably the second most efficient method (>95%), second next to thermal oxidizer.

Lower operating costs and shorter startup time because of the lower temperature.

Lighter weight and smaller space requirements.

Lower NOx emissions (5 ppm at 650F).

The recuperation or recovery of waste heat is allowed since the combustion reaction is exothermic.

Catalysts that can handle air streams with halogenated compounds are available.

Disadvantage:

Lower operating cost is at the expense of higher capital cost and maintenance cost.

Catalyst replacement is expensive. The expense can reach more than half of the capital cost since it involves waste disposal of the poisoned catalyst and the replacement of a precious material and labor. The catalyst usually needs to be replaced every three to five years. The catalyst can easily be poisoned if the incoming vapor stream contains contaminants that react with the catalyst metal. If the catalyst should even partially fail, noxious odorous emissions may be emitted, and these can be a source of public nuisance and complaint.

The palletized catalysts tend to move against each other in the air stream. With time they wear away and must be replaced. This is not the case with for honeycomb catalyst. The catalyst bed (pellets) is subject to channeling, which will reduce conversion efficiency.

If the inlet air stream is preheated by a recuperative heat exchanger and the air stream contains a high VOC concentration, there is a danger that the excess heat will raise the temperature of the catalyst too high. The higher the operation temperature, the shorter the life of the catalyst.

Accordingly, monitors and alarms may be needed to control the temperature of the catalyst and to warn against excessively high exhaust temperatures, high or low natural gas fuel pressures, process blower malfunctions and excessive pressure drop across the system. Also, some air quality agencies require the installation of continuous monitoring devices to ensure compliance. The same care must be given to halogenated compounds.

3. Adsorption Technologies

Process Description:

Adsorption technologies rely on the condensation of organic vapor in the exhaust stream onto the surface of adsorption media, whose function is caused by the polarities of the solvent molecule and the adsorbing media. Adsorption is an exothermic process; thus heat energy is released as the solvent vapor deposits on the active sites. This generally causes a 10-20 F increase in temperature for a fixed bed system. The condensed solvent on the absorbing material must be removed, or desorbed, for subsequent recovery and reuse. Accordingly, either a desorption process follows or the saturated media is removed and shipped to an offsite solvent recovery system, or disposed of as hazardous waste. The first case is classified as a regenerated bed, while the other as a non-regenerated bed. The regeneration process can be cycled or continuous.

The adsorption media includes carbon, zeolite and other materials. The choice depends on the chemical constituents, the operating characteristics and duty cycle. The desorption technologies include steam stripping, hot air desorption, inert gas stripping and vacuum desorption. After desorption, the recovered solvent can be refined for recovery if it has sufficient economic value or it can be destroyed by oxidation or disposed as waste.

In the past 5-10 years, adsorption has found widespread application as a concentrating technology to increase the solvent level present in a process stream, which enhances the economics of downstream treatment. It has been successfully used in low and high flow rate concentration application. Detailed advantages and disadvantages depend on the choice of different technologies. A few examples are given below:

- Advantages and disadvantages of a non-regenerated bed:

Advantages:

Relatively lower (less than 35% of a thermal oxidizer) cost on system installation and operation, also lower maintenance requirement.

Disadvantages:

The advantages are at the expense of a vey high waste treatment/disposal cost. It is also generally limited to low flow rate applications.

- Advantages and disadvantages of a fixed bed with on-site cycled regeneration:

Advantages:

The modular operation allows the system to process highly variable flow rates; thus multiple emission sources are easily integrated.

Disadvantages:

Parallel systems must be sized appropriately to process the full flow rate to avoid saturation. This and on-site regeneration causes higher capital costs. If the flow rate or concentration are highly variable, a breakthrough monitor should be used.

The fixed nature of these systems makes them susceptible to migration of lighter solvents deeper into the bed, particularly if heavier solvents are present in the process stream. This preferential adsorption tends to widen the mass transfer zone (MTZ), and results in solvent breakthrough before saturation.

- Advantage and disadvantages of a continuous adsorption/desorption configuration such as rotor concentrator systems

Advantages:

It achieves a higher flow reduction of 10:1 or better for most applications. This significantly reduces treatment cost. It can be applied to medium and high flow rate applications. Lower fuel cost to destroy the recovered solvent due to the higher heat content form the higher concentration. The reduced amount can be approximated by the ratio of the concentration over the LEL concentration.

Disadvantage:

Higher initial capital and installation cost.

The rotor may be too hot to collect solvent vapors at the exit of the desorption zone, thus cooling air may be required to lower the rotor temperature and avoid solvent breakthrough.

Since the designed rotor speed and adsorption bed depth are based on the maximum process flow rate and solvent concentration, the rotor system is not optimized for all operating conditions. This can be overcome to some extent by monitoring solvent concentration in the exhaust, and adjusting rotor speed accordingly; or by replacing an adsorbing prefilter in front of the rotor to smooth out these concentration spikes.

III. An Alternative Approach: Using Gas Turbine Cogeneration as a VOC emission Control Device

1. Process Description:

Use the SLA stream for combustion air in a gas turbine. The high temperature of the gas turbine will destroy VOCs and capture their heat value to produce electricity and steam.

2. Technical Feasibility:

Little operating experience on using a SLA stream for combustion air in a gas turbine is available. To decide the feasibility of this approach, three major technical issues need to be resolved: Does the process really destroy VOCs efficiently as expected? Does the SLA stream satisfy the requirements to avoid deposition on or erosion of the turbine blades? Are there incremental emissions resulting from such a practice?

2.1 Thermal Destruction Efficiency of VOCs

An important issue of using SLA for gas turbine combustion air is the degree of destruction of VOCs. Many of these compounds are EPA listed pollutants with known or suspected toxic and carcinogenic properties. Combustion of these pollutants in the turbine must be high enough to meet control efficiencies mandated by regulatory agencies, which should be at least comparable to other conventional technologies as summarized in Section II. Also byproduct emissions resulting from combustion of these pollutants must have minimum health effects on the surrounding population.

The fundamental principle of VOC destruction in the gas turbine combustion process is the same as thermal oxidizers described in earlier sections. The emission byproducts are also the same although in varying amounts. As long as the gas turbines meet the temperature and residence time requirements, the VOCs would be expected to be destroyed with equal efficiency. Typical gas turbine generators are designed for high efficiency running usually above 1700 F. At this high temperature, which is higher than 1400-1500 F of thermal oxidizers, the residence time can be reduced to less than 0.05 seconds to achieve the same level of thermal destruction. Accordingly, we can expect as high a degree of destruction as thermal oxidizers (>98%) for the proposed approach.

2.2 Associated Emissions during Combustion

NOx, CO and total unburned hydrocarbons (TUHC) will be emitted from the gas turbine. TUHC and CO emissions from natural gas-fired turbines comes from the unburned fuel or partially burned products that escape the reaction zone and are emitted in the exhaust. For the newly designed high efficientcy gas turbines, the combustion efficiency typically exceeds 99 percent at base load. Accordingly, these emissions are typically low, less than 100ppm for CO and 5ppm for TUHC. Also, these new gas turbines have NOx control to meet the stringiest requirements on NOx emission (10 to 25 ppm mandated in some states and locatities).

The unreacted VOCs or hydrocarbon byproducts from uncomplet combustion of VOCs, which commonly referred to as products of incomplete combustion (PICs), can be emitted. However, the formation of PICs does not imply that they will be emitted in the turbine exhaust. In the highly oxidative environment and high temperature of the gas turbine, these PICs will be destroyed.

2.3. Gas Turbine Air Filtration Requirements

Industrial Gas turbine applications require a clean inlet air supply with minimum pressure loss to minimize adverse effects on engine life. An inlet air filtration system is normally installed to control the concentration of airborne dust, sand or salt to a set of air quality criteria that are specific to the engine design and its application.

Particulates can cause high temperature damage from deposition, plugging of cooling passages, and errosion of compressor and hot section blades. However, the standard requirements vary. Westinghouse and United Technologies estimate that the allowable particle mass concentration passing through a gas turbine is 4.6 mg/Nm^3 and 2.8 mg/ Nm^3 (2 and 1.2 grains/1000 SCF) (Sverdrup, 1977 & Robson, 1976) respectively. The limit also varies with the size of the particulates. For example, General Electric has estimated that a heavy duty industrial gas turbine can tolerate particle concentration in the range of 100 mg/Nm^3 (43 grains/1000SCF), provided 98 percent or more of the particles are under 10 μm in diameter (Boericke, 1980). On the other hand, aircraft derivative engines have a much lower tolerance limit. G.E. Aircraft Engine Business Group indicates solid particulate limits of 0.17 to 0.017 mg/Nm^3 (.07-.007 grains/1000SCF) (Sailer, 1989). According to the data from Armstrong Laboratory, the particulate concentration of a standard paint booth is well below the tolerance (Acurex, 1994). The concentration can be further reduced by turbine inlet filtering methods.

The size of the particles and the overall dust loading entering the turbine are not the only considerations for the feasibility study of the engine life. As for the thermal or catalytic oxidizers, care must be taken with the chlorine and sodium related pollutants, which are corrosive. For painting facilities that use corrosive chemical stripping, the same cautions are required as for the thermal and catalyst oxidizers.

3. Advantages and Disadvantages of the Gas Turbine approach

Advantages:

Due to the steady and declining natural gas price and its clean burning properties, many light weight and compact gas turbine generators have been designed for commercial use as power cogeneration and steam producers. These units are easily available and have been operating for many years. The equipment cost has been declining due to the large demand from all over the world. Compared to the thermal oxidizers, these gas turbine units are better designed at more reasonable prices. They are easy to install and to operate. Note that the combustion chamber cost of a cogenerator is only 12% of total cost. For the same oxidizing capability, in many cases, the cogenerator combustion chamber is 50% cheaper than the thermal oxidizers.

All the advantages of thermal oxidizers apply to the gas turbines.

It generates power and steam at high efficiency and low cost. In most cases, this produces net positive capital investment returns.

Disadvantages:

The added equipment for electricity and steam generation increase the capital cost. The capital cost is increased for larger power and steam generations.

If more than needed power is generated, power wheeling is needed to sell the excess power. More sophisticated economic analysis is needed to decide the optimum generation size and the associated investment risk.

A cogenerator is a much more complicated system than oxidizers. The system includes many roating moving parts to generate electricity, which results in higher maintenance cost. For aircraft derived gas turbines, major maintenance/overhaul is required every four to six years. However, the cogenerator has a life expectancy of more than 20 years when properly maintained.

IV. Split-Flow Ventilation and Recirculation to Reduce the Exhaust Air Flow Rate

To process the low VOC concentration but high volume air flow is rather expensive due to the low heat value of such a condition: external energy sources are required, such as gas to oxidize the VOCs in the air stream. Accordingly, many oxidizers are combined with a concentrator as described in section 3. Usually 10 to 1 concentration can be achieved by using some carbon adsorption technologies. Higher concentration capability to 100:1 is also available, according to Mr. James Nester from REECO, using polymer technology. Then the concentrated flow has higher heat value and can be oxidized with much less external energy input. The fuel savings can be approximated by the same way as for concentrators.

Another mean to concentrate the VOC flows is by split-flow ventilation and recirculation. In addition to the advantage of a lower emission control cost for the smaller amount of exhaust air, it also reduces the heating and cooling cost of the painting process. The disadvantage is a higher concentration of VOCs in the working environment. Therefore, care must be taken to ensure the VOC concentration is within the health tolerance limit in the working environment.

The split-flow circulation can be summarized from the report (Acurex, 1994). The idea consists of three basic ducts, one duct draws outside fresh air and mixes the fresh air with the circulated air, one duct from the top of the painting booth draws the air to recirculate, and the last duct from the bottom portion of the booth draws the air for treatment. The ducts' arrangement is based on observation of the law of gravity wherein the air in the bottom portion of the booth usually contains the majority of VOC vapors which are heavier than air. The above observation is also verified by measurements from Booth 2 at Travis AFB.

V. Economic Analysis

1. Cost Effectiveness of the Split-flow Ventilation/Recirculation Approach

Clearly the main payback from the split-flow ventilation/recirculation is from the heating and cooling cost savings as well as the reduced oxidization fuel cost during operation. Additionally, there will be some reduced capital cost for the emission control due to the smaller required control device size. The smaller the exhaust flow rate, the higher the payback due to the more circulated air saving heating and cooling costs. Accordingly, no matter what the emission control technology used, one should use as much recirculation as allowed to maintain a safe and healthy working environment. The incremental payback is decreasing as the circulation increases. Thus, the small economic gains must be weighed against health and safety concerns. Our analysis shows that usually the payback period is less than three years for a 90% recirculation and a reasonably priced modification rate at $2 per scfm. The assumed electricity rate is 3.5 cent/kW and gas rate of $2.5/kcf, which is very low. The higher the energy cost, the shorter the payback period.

2. Cost Effectiveness of Concentrators and Oxidizers

For safety and health reasons, the working environment for the paint facility is expected to have a VOC concentration below a few hundred ppm, whose heat content is still well below the flammable point, which is around 4,000 ppm. Accordingly, it takes extra energy to oxidize these VOCs. To improve the oxidization efficiency, concentrators are usually used to further increase the VOC concentration before feeding into the thermal or catalytic oxidizers. For a cogeneration unit, this is not usually necessary unless the flow rate is too large for the unit to handle. When the cogeneration unit is large enough (more than a few mega watts), the burner is able to handle more than 25,000 scfm. One should view a cogeneration unit as an energy source for electricity and steam, optimizing its operation accordingly. VOC destruction should only be viewed as a convenient byproduct.

Commercially available concentrator gains range from a single digit to thousands. For example, a rotary carbon concentrator has a gain around 10, a polymer based concentrator can achieve 100 and a fluidized bed concentrator can achieve a gain around thousands. The concentrated flow has a lower volumetric flow rate with higher VOC concentrations. The flow rate enables the system designer to choose smaller oxidizers to lower the capital cost. The higher concentration density contains higher heat value and can reduce the operating cost of the oxidizers.

However, normally oxidizers generate neither electricity nor steam, and it requires energy to oxidize the VOCs. This is one major economic difference between oxidizers and cogeneration units. Also the capital cost of a concentrator is expected to be paid back in a few years by reducing the capital on a smaller oxidizing unit and the operational savings.

3. Cost Effectiveness of Gas Turbine Cogenerators as VOC Destruction Devices

The fundamental idea of cogeneration is to utilize the waste heat energy produced from the generation of electricity to generate steam. The heat energy is normally rejected to the atmosphere through cooling towers, cooling lakes, and other means. For example, according to the statistics of the US Energy Information Agency, more than 67% of the primary energy from the production of electricity is lost. Common ways to utilize heat is to produce steam for further electricity generation, for heating the facility or to drive chillers to cool the facility. With the combined electricity generation and heat usage, cogeneration units can achieve up to 85% energy efficiency. Accordingly, if one wants to use gas turbine cogenerators solely as VOC thermal oxidizers without using the heat energy (such as the incorrect approach presented in the model analysis of Creax program 1996), the chance of payback is small and it is against the fundamental principle of cogeneration usage.

Despite the high potential for energy savings from cogeneration usage, there are a variety of psychological, social, technical, economic, and environmental factors which prevent more widespread use of cogeneration. Typical barriers include: financing; lack of expertise; economic uncertainty; siting issues; institutional complexity and the trade-off between electricity, thermal energy production and VOC destruction. In a cogeneration system, heat or cooling energy from the power plant is most often provided to consumers through a network of distribution pipes. The cost of constructing and maintaining this network tends to favor the siting of cogeneration systems in areas with high energy use density, such as the downtown sections of major cities, or multi-building facilities such as hospitals, university campuses and shopping centers. Air Force bases certainly satisfy the profile as high energy use density areas. The potential energy saving, economic gains and environmental benefits are definitely worth investigating.

A reasonable approach for a big paint facility is to begin with the past utility records to determine the base and peak electricity, steam and cooling demands to determine the optimal size of a cogenerator. Find the associated electricity rate, gas rate, steam water rate and chilling water rate. This data will allow the claculations to be performed to determine how much recirculation/concentration would be required to use the cogenerator as a control device. Detailed algorithms are still needed for this analysis. For example, if a paint facility is located in a high electricity rate area, one may want to generate as much power as possible to sell to generate positive cash flow. However, higher generation capacity implies higher capital cost and higher investment risk. One may need to find a financial partner to co-own the cogeneration unit. Otherwise, one may just cover the base load and try to purchase the peak load power to reduce the risk. On the other hand, if the facility is at an electricity surplus area, one may want to choose a smaller unit just to cover the steam need. A cost analysis is needed to decide the choice. Other complications, such as rising heat rate (lower efficient) due to a smaller unit, need to be worked out for more detailed analysis.

Based on the flow rate, one can design the recirculation system to closely match the exhaust air with the inlet air. If the chosen cogenerator is larger than 4 MWatts, the VOC concentration in the resultant recirculation system usually will be below 400 ppm, which is well below the safety and health upper constraints. Otherwise, a larger exhaust flow rate is needed so that it will satisfy the safety and health constraints. For these cases, a concentrator is needed to further concentrate the exhaust flow rate for the cogenerators.

VI. Some Practical Case Study

Example 1: Ford Motor Company

On July 24, 9:30 am, Dr. Huang called Dr. Byung Kim from Ford Research Institute, Dearborn, Michigan to discuss about the adopted VOC control approach in Ford paint Booths for their automobiles. One acquired example is as follows: The involved paint facility flow rate is about 1,000,000 scfm with a VOC concentration of 10-20 ppm. To reduce the VOC control cost, recirculation is used to reduce the exhaust air flow rate. The exhaust flow is reduced to 100,000 scfm and the VOC concentration is increased to 100-200 ppm range. Then a carbon absorption concentrator is used to further increase the concentration to 1000-2000 ppm range. The condensed flow is then followed by a thermal oxidizer to destroy the VOCs. Natural gas is used for the thermal oxidizer, where the heat value of VOCs is estimated around 20%. No electricity or steam was generated in the process for use. The operating cost is in the range of 2 million dollars per year.

Example 2: Anderson Lithograph

This example does not use a gas turbine to deal with the painting VOC problems. However, they use it to deal with other VOC sources and the benefit is clear. The following is quoted from http://www.andlitho.com/ALcogen.html Website address.

In 1991 Anderson Lithograph started investigating and evaluating viable alternative solutions to meet shortfalls in its manufacturing process and environmental chilling capacity. The ideal solution would provide sufficient capacity to ensure a stable and controllable temperature and humidity range in the sheetfed and web press manufacturing areas of 72F, 5F, and 55%, 5%; as well as, meeting future process chillwater requirements associated with increased press manufacturing capacity. Several system design and financial arrangement alternatives were evaluated over the past eighteen months. These included, but were not limited to, additional "high-efficiency" electrically powered chillers, absorption chilling utilizing waste heat from the existing Katec thermal incineration pollution control device, and absorption chilling utilizing a cogeneration system. The cogeneration system was selected as the most viable alternative.

The cogeneration alternative selected, and now installed and operational, is based on a natural gas fueled, combined cycle, combustion turbine based power generation facility capable of producing a total maximum of 5.1 mega watts of power. The natural gas fired combustion turbine generator set is capable of producing 3.9 mega watts of electrical power. The waste heat from firing the combustion turbine is utilized to produce steam via a heat recovery steam generation boiler unit (HRSG). The steam is then utilized to drive a back-pressure type steam turbine powered generator set capable of producing 1.2 mega watts of electrical power. The combination of the gas fired and steam powered turbine generator sets is where the term "combined cycle" is derived. The low pressure steam output of the steam turbine is then used to generate in excess of 1,300 tons of chilled water capacity via two 1-stage absorption chillers. The use of the waste heat from the exhaust of the gas fired combustion turbine to generate steam and then the steam used for generation of additional electrical power via the steam turbine, and chilled water for HVAC and process cooling via an absorption chiller is the basis for the operational term "cogeneration".

The cogeneration system is capable of meeting all current, and future, equipment/process and building environmental chilled water requirements. The cost of acquiring, operating and maintaining the cogeneration facility will be offset by eliminating the cost of electrical power, previously purchased from Southern California Edison Company. The first year of the project resulted in an incremental profit contribution and positive cashflow, with each of the succeeding nine years of the project showing further increases in the incremental profit and cashflow contribution. There are other significant benefits associated with the project that can dramatically improve the profit impact noted, and the environmental impact of the Company's manufacturing operations. Most significant is the Emissions Reduction Credits (ERC) to be obtained from elimination of fugitive VOC emissions of 350 lbs/day associated with ducting the building HVAC exhaust directly into the combustion turbine for thermal destruction. The innovative system design for the cogeneration plant incorporates the capability to capture and destroy all fugitive VOC emissions generated from all manufacturing and maintenance operations. Based on daily VOC emissions reporting, Anderson generates approximately 350 lbs/day of fugitive VOC emissions. This would reduce current average daily VOC emission levels by 63% - a dramatic positive impact on air quality. This will enable Anderson Lithograph to not only meet, but exceed, air quality VOC emission levels planned for the year 2010 by SCAQMD -within the next twelve months. With the installation of this system, Anderson will be able to generate an "ERC" for these emissions and place them in the SCAQMD bank. These emissions are currently selling at an average price of $3,500 per pound. This would generate and asset worth over $1,225,000 that could be either sold and/or used as offsets against any future regulatory mandated reductions. Further, they could be utilized for new press expansions to accommodate additional sales volume growth.

Note that Anderson Lithograph is a printing company located in Los Angeles, a non-attainment region which has tight air pollution regulations. This example demonstrates that NOx emissions from gas turbine cogeneration can be controlled to satisfy the strictest regulation requirement, it also demonstrates the feasibility of the cogeneration concept as a VOC destruction device. According to Frank Bornett, the system has been operating for one year and produces positive cash flow as predicted. The system was designed by BPR Energy.

Example 3: Kelly Air Force

Total airflow rate: Designed at 740,000 scfm, provided by four induced draft squirrel cage blowers of 125 horse power each. However, it is operated at 500,000 scfm.

Total concentration: 10-40 ppm

Chemical components: Some standard paint components, primary ones are MEK, Toulene, Xylene, etc.

The operating temperature is around 75F.

The humidity is around 80%.

Load: Each C-5 receives about 300 gallons of paint (including 100 gallons of primer), applied continuously over a 16-24 hour period. These 300 gallons of paint contain about 1,500 lbs of organic solvents.

Building height: 33ft, width 248ft.

NESHAP requirement (has to be satisfied by September 1998): To reduce the above VOC emissions by 81%. However, we are expecting removal efficiency of at least 90% of the acquired control.

Available utilities: Boilers and gas pipeline nearby.

Exhaust airflow rate: (airflow needs to be treated) 25,000scfm with a VOC concentration of 400 ppm.

Note that none of the following analysis considers the split flow recirculation capital cost and the associated operating cost, since all these cost are the same for all cases; and the pay back period is in 3 years as mentioned earlier (Creare program, 1996). It costs about one million dollars for the recirculation system. The exhaust air flow is designed at 25,000 scfm with concentration of 400 ppm. We also assume the currently available rate of 3.5 cent/KW for electricity and $2.5/kcf for gas in the following estimates. A very conservative overall system efficiency, 70%, is assumed for the cogenerator.

Case 1: We can use a concentrator and a thermal oxidizer (which uses no fuel saving technology) similar to the Ford example, $550,000 for the concentrator and thermal oxidizer, and it costs about $300,000 per year to operate.

However, many fuel saving technologies are availabe to reduce the operating cost. We thus consider the following three fuel saving technologies.

Case 2: One can also use a regenerative thermal oxidizer directly, then the capital cost is reduced to $400,000, and the operating cost is around $70,000 per year. (Adwest Technologies, Inc. 1996, which uses Retox and Katox Oxidizers.). The regenerative thermal oxidizer does not need gas as long as the exhaust air exceed 3% of the LEL concentration. For example, for Toulene, the LEL is around 4,000 ppm. 3% of the LEL is only 120 ppm. For 400 ppm, this condition is mostly satisfied. It only produces NOx during startup (it takes about one hour) when the burner is on. The maintenace will be more than case 1 since valves are used for the regenerative bed, and the valves need to be maintained regularly. Note that the above quote may be unusually low. For example, another quote from Smith Engineering Company costs about $625,000.

Case 3: Considering recuperators as another means to destroy the VOCs, we need a concentrator to increase the VOC concentration. Then the capital cost is increased to $850,000, but the operating cost is reduced to $33,600 according to Mr. Merrel F. Schilling from Renew Inc. using Durr Oxidizers and concentrators. The recuperators apparently have a lower operating cost, and it is much faster (minutes instead of hours) to shut down and start up.

Case 4: One can also use a concentrator and a catalyst oxidizer as our VOC control device. Accoding to Wheelabrator, the capital cost is about about $750,000 with the opearting cost reduced to $20,000 due to its low temperature (around 500F) operation. Accordingly, it is the fastest to shut down and start up. However, the catalyst needs to be replaced every three to five years, which increases its maintenance cost. The replacement cost can reach two thirds of the capital cost.

Case 5: One can also use a high gain concentrator to reduce the fuel cost. According to REECO, a FluiSorb Concentrator System cost about $875,000. The gain is about 800 for the assumed scenario. After such a high concentrator, there is no need for energy saving oxidizers since the concentration is high enough to fire the oxidizer itself. The operating cost is reduced to only $5,000 since no more natural gas is needed after start-up, the used oxidizer is much smaller and it needs much smaller fans.

Note that all the concentrators may need to have their adsorption medium reactivated every one to five years. The cost varies from technology to technology, but it will not be as high as the catalyst replacement. The cost ranges from $5,000 to $20,000 per year.

Case 6: Another possibility is to use a gas turbine cogenerator to treat the exhaust air. To limit the investment risk, say we limit ourselves to 5 MW generators, which can handle the VOC contaminated exhaust air of 25,000 scfm. The heat rate is about 12,000 BTU/KW-hr for a 5MW cogenerator. For the gas price of $2.5/kcf, it is equivalent of an electricity rate of 3 cent/KW. It costs about 2 million dollars for a 5 MW generator. On the other hand, the generated power and steam (assumed around 25,000 lbs/hr) will produce revenues. The net revenue from electricity (electricity revenue minus the gas cost) is about $200,000 per year, and the net revenue (saved steam fuel) from the steam about $500,000. Supposing that the maintenance cost and other material cost are about $300,000, then it takes about 5 years to payback.

It is then clear that we need a higher initial capital cost for the cogenerator project (eg. $2,000,000 vs. 650,000), but it produces positive cash flow (eg. $400,000 vs. -$300,000) and has a payback in five years.

According to the steam demand record, Kelly Air Force Base has a base steam demand of 50,000 lbs/hr. We can add the next smallest available equipment, which can generate total steam up to 64,000 lbs/hr with higher efficiency (86%), but we need to add extra fuel to burner at 40.1 MMBTU/hr. Then the saving from the steam will be increased to $600,000. The cash flow will be increased to $500,000. The extra capital cost is $300,000. The payback time is slightly less than 5 years.

Case 7: Assuming we do not have the constraint on the initial capital cost, but we do not want to become involved with power wheeling, then it is reasonable to decide to cover the base load as much as we can. The base load is around 25 MW, which is 72% of Kelly's average demand and 45% of the peak demand. Additional power demands will be picked by City Public Service. Note that the local power company generates electricity with an efficiency of 30%, and we are using 70% (compared to 85% as estimated by most cogenerators) as a conservative estimate of our cogeneration plant, and we assume that we only use 50,000 lbs steam/hr out of the steam byproduct of 120,000 lbs/hr). For such a scenario, it takes $12M initial capital cost. For 25MW cogenerators, the heat rate can be reduced to 8800 BTU/KW (more efficient). To make a conservative estimate, we use 9500 BTU/KW. Then the generated electricity rate is less than 2.5 cent/KW. Accordingly, the net saving from the electricity is $2.2 M. The steam fuel saving is about $1M. However, we expect a higher electricity rate from City Public Service due to the dramatic reduced demand, and the demand at peak hours, estimated at $.5 M and the other items and maintenance are estimated at $1M. Accordingly, it takes slightly less than 5 years to payback. Note the steam potential is more than twice of the needed steam. The more the steam used, the faster the returns. Here, 50,000 lbs is used for a conservative estimate.

By comparison with a 5MW cogenerator, it is clear that the initial capital cost is much higher, but the potential payback is slightly better due to the higher electricity production efficiency. However, the payback is limited by the steam demand. Unless higher steam demand is expected, it may not worth the trouble to choose the 25MW design.

Case 8: Suppose that we want to further restrict the initial capital to less than $1M, then we need to choose a 1MW cogenerator, which costs around $650,000. However, the machine can only handle the airflow rate of 8,000 scfm. A concentrator is needed, which increases the capital to about $1M. However, the heat rate for such generators is increased to around 14,000 BTU/KW-hr, which reduces the electricity saving to zero. The only potential saving is from the steam produced typically at 8,000 lbs/hr. The saving from the steam is around $150,000 per year. Assuming the maintenance cost is around $80,000 per year, then the payback is about 15 years. However, it is still better than the other oxidizers, which need extra expense to operate. Note that heat content of the condensed exhaust air will be increased to around 10%, which will reduce the fuel cost of the cogenerator; with this consideration, the payback will be reduced to 12 years.

If we buy extra addon devices to generate extra steam up to 37,000 lbs/hr with higher efficiency (90%) with extra fuel input at 28.5 MMBTU, then the saved steam fuel will be increased to $200,000. The extra capital cost is $600,000. Then the payback period is reduced to around 12 years without considering the heat content of VOC, and 11 years when the heat content is considered.

From the above analysis, we conclude 5MW is the best choice in terms of meeting the steam demand, investment risk and overall performance. Based on the obtained insights, we may want to find a generator that can meet the steam demand without using extra add-on devices since the steam is a 'free' byproduct. In contrast, the added devices require additional fuel even though consumed at a higher efficiency. To elaborate the idea, we look at a case that can meet the steam demand closely, which is the 10 MW case.

Case 9: Consider the 10 MW generator, which can produce 50,000 lbs/hr steam to meet our demand. Clearly, the saved steam fuel will be increased to $1 Million dollars. The electricity net revenue will be $480,000 with the assumed heat rate 11,000 BTU/KW. The capital cost is around $4,500,000. Assume the maintenance cost is around $480,000, then the payback period is reduced to 4.5 years. We can see this is the optimal choice in terms of payback period and meeting the steam demand. However, the initial capital cost is more than twice the 5 MW design.

All the above eight cases have more than 95% VOC destruction rate. The following carbon adsorption technology can only achieve 81-90% destruction rate.

Case 10: Another possibility is to use a non-regenerated carbon adsorption system for the VOC emission control. According to Unisorb Corporation, this system costs about $170,000. The maintenance cost is about $10,000 per year. The operating cost is about $20,000 per year. However, the waste disposal cost is about $7,500 per airplane (per 300 gallons of paint). It is clear then that the initial capital cost is very low (less than one third of the others), but the waste disposal cost is very high. Supposing the hangar paints 50 airplanes per year, the waste disposal cost reaches more than $375,000 dollars per year. The high cost is associated with shipping and reactivating the saturated carbon. In addition, according to the operating experience from the Tinker Air Force Base, Oklahoma, one should be aware of the potential fire hazard problem for such a system. The saturated carbon adsorption system can easily catch fire.

Example 4: United Airline aircraft painting hangar

According to Mr. C. Castaldini from Acurex, United Airline is considering using their existing cogenerator for VOC destruction. However, since there is some distance between the cogenerator and the painting facility, the whole idea is still under evaluation.

One should be aware that when an on-site cogenerator is used for VOC destruction, the associated VOC control cost is actually zero (or even negative due to the added energy value) since the cogenerator cost to supply the energy need remains the same.

VII. Conclusion

Technical and economic feasibility of gas turbine cogenerators as VOC destruction control devices for painting facilities have been evaluated. Conventional VOC control technology is summarized for comparison. We conclude that the gas turbine cogenerator has a high potential for VOC destruction; however, barriers such as financial, siting, complex organization issues do exist. To follow the fundamental idea of cogeneration, one should analyze the potential economic benefit from the predicted demand for electricity and steam if no cogenerator already exists; and VOC destruction should be only treated as a convenient byproduct. If a cogenerator already operates at the facility, then modification of the facility for VOC destruction becomes a much simpler issue. A few studied cases were used to demonstrate our overall approach.

VIII. Acknowledgment

Dr. Huang appreciates the inputs from numerous contributors, including but not limited to: Captain Mike Blank, Bob Chabot, Mike Haas, Greg Vallery from Kelly Air Force, and Dr. Joe Wander from Armstrong Laboratory, numerous discussions with engineers from Adwest, Ford, GM, Acurex, Acreax, Reeco, Durr, Solar Turbines, Smith engineering, etc. Without their generous support, this report may have taken much more time and effort to finish.

References

[1] Jeffrey R. Lord, The Aerospace NESHAP: How & When the Rule May Affect You, Envirscope, July 1996.

[2] National Emission Standard for Aerospace Manufacturing and Rework Facilities, Rules and Regulations, Federal Register, Sept. 1, 1995.

[3] M. K. Sink, Handbook: Control Technologies for Hazardous Pollutants, EPA/625/6-91/014 (NTIS PB92-141-373) U.S. Environmental Protection Agency, Center for Environmental Research Information, Cincinnati, OH, June 1991.

[4] Acurex Environmental Corporation, Demonstration of Split-Flow Ventilation and Recirculation as Flow-Reduction Methods in an Air Force Paint Spray Booth, AL-EQ-TR-1993-0002, Final Report of EPA Contract No. 68-D2-0063, 1994.

[5] Acurex Environmental Corporation, Paint Spray Booth VOC Thermal Destruction by a Gas Turbine Combustion system, Final Report of Acurex Project 8059, 1990.

[6] F.L. Robson and W. A. Blencher, "Particulate Control Requirements for Gasified Coal, Combined-Cycle Power Plants," NTIS/PB 260499, Proceedings of Particulate Control in Energy Processes Symposium, San Francisco, Ca. 1976.

[7] E.F. Sverdrup, et. al., "The Tolerance of Large Gas Turbines to Rocks, Dust, and Chemical Corrodents," EPA Department of Energy Symposium on High Temperature, High Pressure Particulate Control, EPA-600/9-78-004, Sept. 1977.

[8] R. Boericke, R. Hartman, and J. Kuo, " Assessment of Gas Turbine Erosion by PFB Combustion Products," paper presented at the 6th International Conference on Fluidized Bed Combustion, Atlanta, Georgia, April 1980.

[9] E.D. Sailer, GE Aircraft Engine Product Division, Private Communication with C. Castaldini, August 25, 1989.

[10] Creare Incorporated, System Specification for Abatement of Emissions at Kelly Air Force Base, San Antonio, Texas, attached with an energy calculation program, Project 7791, April 1996.

[11] Adwest Technologies, Inc., RETOX Regenerative Thermal Oxidizer System, Proposal No. 96-2256, August, 1996.

[12] Smith Environmental Corporation, Proposal # BO7-96-147-0, August, 1996.

[13] Renew Inc., Proposal for VOC Emission Control, August, 1996.

[14] Wheelabrator Clean Air Systems Inc., VOC Emission Control System, WCAS/Amcec Reference A0570, August, 1996.

[15] REECO, REECO Proposal No. P-40-6999-103, FluiSorb Concentrator/Oxidizer System for VOC Control, August, 1996.

[16] Ed. Biedell and H. Cowles, "A Novel Pre-concentration Technology for VOC and Air Toxics Control," presented at the Air & Waste Management Association's Specialty Conference on Emerging Solutions to VOC & Air Toxics Control, Feb. 28-March 1, 1996, Clearwater Beach, Florida.

[17] Unisorb Corp., VOC Emission Control Proposal, August, 1996.

RISK ANALYSIS OF AGING AIRCRAFT: A COMPARISON BETWEEN CURRENT METHODS AND STATE OF THE ART PROBABILISTIC DESIGN TECHNIQUES

Dr. Lidvin Kjerengtroen
Associate Professor
Department of Mechanical Engineering

South Dakota School of Mines & Technology
501 E St. Joseph Street
Rapid City, SD 57701

Final Report for:
Summer Research Program
San Antonio Air Logistic Center
Kelly Air Force Base

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC
and
San Antonio Air Logistic Center

August 1996

RISK ANALYSIS OF AGING AIRCRAFT: A COMPARISON BETWEEN CURRENT METHODS AND STATE OF THE ART PROBABILISTIC DESIGN TECHNIQUES

Dr. Lidvin Kjerengtroen
Associate Professor
Department of Mechanical Engineering
South Dakota School of Mines & Technology

Abstract

Probabilistic crack growth analyses are currently used in the risk assessment of individual aircraft and for aircraft fleets (Berens et. al., 1991). The software package PROF is, for example, in use at SA-ALC. Although PROF may be a helpful tool, there are many simplifying assumptions associated with the usage of this software. These assumptions are necessary in order to use the "exact" probabilistic procedures implemented in PROF. Some of these simplifying assumptions are explored and compared with results obtained using state of the art technology.

Through numerical examples it is demonstrated that PROF in general will underestimate the risk associated with a fatigue crack. The results are verified through Monte Carlo simulation.

RISK ANALYSIS OF AGING AIRCRAFT:
A COMPARISON BETWEEN CURRENT METHODS
AND STATE OF THE PROBABILISTIC DESIGN TECHNIQUES

Lidvin Kjerengtroen

1. Introduction

Fatigue has been a problem in the design of many structures in aeronautical engineering (e.g., aircraft structures, jet engines, etc.); in civil engineering (e.g. bridges, offshore structures, power plants, etc.); and in mechanical engineering (e.g., engines, shafts, pressure vessels, piping, etc.) (Yao et. Al. 1986). Because of the inherent variability in fatigue strength in materials as well as the statistical nature of structural loads, a probabilistic approach in the analysis and design of fatigue critical structures and components is important.

There exist many theorems on fatigue and fatigue damage. The theory of fracture mechanics has been well established. Nevertheless, the damage analysis of redundant structural systems remains a challenging task.

In recent years considerable research efforts have been reported on probabilistic modeling of fatigue crack growth. Uncertainties in loading conditions, crack geometry, and material properties have been included. Numerical algorithms for the calculation of probability of failure / reliability have been developed and the first and second order reliability methods are recognized as being both very accurate and efficient (Madsen 1985, Yao et. al. 1986). Inspection and maintenance programs for new and aging aircraft have been developed and implemented. Crack sizes are measured either directly or indirectly through non-destructive inspection methods. When the crack is measured indirectly the signal must be interpreted as a crack size. This process is clearly associated with uncertainty and is random in nature. Likewise, a visual inspection has a certain probability of non-detection. Based on in service inspection results, the estimate of the reliability can be updated and the uncertainty in the initial modeling of the basic variables can be reduced. This report look at some of the aspects associated with probability based inspection and maintenance. Current practice and procedures are compared with state of the art reliability methods.

2. Structural Fatigue

2.1 Background

The two important phases of the physical process of fatigue are (a) crack initiation and (b) crack propagation. In engineering crack initiation refers to the formation of a crack that is reliably detectable with the use of nondestructive evaluation techniques. The process of crack initiation is complicated since it involves the development of a crack at the micro-scale level. There is still a gap between micro-mechanics and engineering applications.

In the US Air Force the conservative assumption is made that there is an existing flaw before the very first load cycle. For low stress fluctuations at a fatigue-critical location the initiation period can be a significant fraction of the total fatigue life. When stress fluctuations are high and/or notches are present, the initiation of fatigue cracks may occur early and a large portion of the service life of the component may be spent in propagating the crack to a critical size. Sources of uncertainty in the fatigue process include:

1) The location, shape and size of a crack may not be clearly defined.
2) Fatigue test data are subject to large scatter.
3) Mathematical relationships to describe fatigue behavior are empirical in nature
4) Loading processes causing fatigue damage are often not well defined.
5) The internal forces and moments produced by a given load may not be precisely known.
6) From idealized stress analysis the state and magnitude of stresses are uncertain.
7) Effects from corrosion, temperature, etc. are not precisely known.

Nevertheless engineers must make decisions regarding the design and integrity of components and structures, for these reasons probabilistic and statistical methods can be very helpful.

2.2 Fracture Mechanics - Fundamentals

It is assumed that the readers of this report are at least somewhat familiar with the basic theory of fracture mechanics and the fracture mechanics approach to fatigue. In fracture mechanics the stress intensity factor K is the "modules" and the fracture toughness K_C is the critical strength parameter Collins (1993). Failure is predicted to occur if

$$K \geq K_C \qquad 2.1$$

For most applications the stress intensity factor K is of the form

$$K = \beta(a) S \sqrt{\pi \cdot a} \qquad 2.2$$

where β(a) is a geometry correction factor, S is the far-field stress, and a is the characteristic crack depth. Although the details of calculating K and determining K_C for some cases may be difficult, the basic concept of predicting failure by brittle fracture is no more complicated than this.

Fracture mechanics is a method of characterizing fracture behavior in structural and mechanical components. It has evolved as a result of attempts to understand and prevent structural and mechanical failures and to provide fracture criteria. Fracture mechanics can be used to describe quantitatively the trade-off's among factors, such as stress, material toughness, flaw size, and flaw shape. The conventional approach in design and during a strength analysis is to make conservative assumptions, or assumptions based on a worst-case scenario. The maxima of loading, crack size, and crack growth rate, and the minima of strength are often treated as representative design situations, and of simultaneous occurrence. A probabilistic approach, where frequency of occurrence of all of the above mentioned parameters is considered is much more realistic.

For a growing crack which is subject to a cyclic load, the crack growth can be described by the well known Paris equation,

$$\frac{da}{dN} = C(\Delta K)^m \qquad 2.3$$

where a is the crack size, N is the number of load cycles, ΔK is the stress intensity factor range, and C and m are experimentally determined material properties.

The use and applications of crack growth equations are also straightforward. In general numerical integration is required.

Probabilistic crack growth analysis are currently used in the risk assessment of individual aircraft and for aircraft fleets (Berens et. al., 1991). The software package PROF is for example in use at SA-ALC. Although PROF may be a helpful tool, there are many simplifying assumptions associated with the usage of this software. These assumptions are necessary in order to use the "exact" probabilistic procedures in PROF. Some of these simplifying assumptions are explored and compared with results obtained using state of the art approximate, but consistent, technology.

3. Probabilistic-Design-Methods

3.1 Introduction

Structural and mechanical design, typically, consists of choosing components in such a way that various criteria involving safety, serviceability, and durability under the action of loads are satisfied. Unfortunately, strength is not a single variable, nor is load or stress a single variable.

Numerous sources of uncertainty are involved and their relationship is often non-linear. For this reason, classical probability theory is often of academic interest only. Approximate solution methods governs modern probabilistic design theory and often probabilities should be considered notional.

The first step toward evaluating the reliability of a structural or mechanical component is to decide on the relevant load and strength parameters. Theses variables are called basic variables X_i. The second step is to formulate a mathematical relationship among these variables. This relationship can be expressed as

$$Z = g(X_1, X_2, \ldots, X_n) \tag{3.1}$$

Failure occurs when the performance function or limit state function reaches the limit condition of Z = 0. So the third step is to find the minimum of the function Z with the constraint Z = 0, this then represents the most probable failure point. The boundary Z = 0 separates the safe and the unsafe regions in the design parameter space. When Z = 0 the component can no longer fulfill the function for which it was designed.

For example for a simple rod in tension z = strength - stress = R - Q/A = 0

Q is the axial load and A is the cross sectional area. If the strength equals to stress, failure will occur.

The limit state function or performance function plays a key role in probabilistic design. A limit state function can be an explicit or implicit function of the basic variables. Failure occur when Z <0 and the probability of failure p_f can be expressed in terms of the integral of the joint probability distribution,

$$p_f = \int_{\Omega} f_X(x_1, \ldots, x_n) dx_1 \ldots dx_n \tag{3.2}$$

In general this joint probability distribution is not known, and even if it was known a solution would not be feasible.

Methods approximating the equation for p_f have been developed. These methods can be grouped into three types, first- and second-order reliability methods (FORM and SORM respectively) and Monte Carlo simulation. The limit state functions can be either linear or nonlinear functions of the basic variables. For linear failure functions, with normal basic variables, FORM gives an exact probability of failure. In the case of a nonlinear limit state function a first order approximation can be made. The SORM estimates the probability of failure by using a second order approximation to the limit state function.

3.2 Fundamental Concepts (Cornell 1969)

The original formulation by Cornell defined a performance function,

$$Z = R - S \tag{3.3}$$

where R and S are independent variables and defined as strength and stress respectively. From fundamentals of statistics the mean value μ_Z and the variance σ_Z^2 of Z is then

$$\mu_Z = \mu_R - \mu_S \tag{3.4}$$

and

$$\sigma_Z^2 = \sigma_R^2 + \sigma_S^2 \tag{3.5}$$

The event of failure is defined as $Z < 0$ or $R < S$. If R and S are normally distributed, the probability of failure is

$$p_f = P(Z < 0) = \Phi(-\frac{\mu_Z}{\sigma_Z}) \tag{3.6}$$

where Φ is the cumulative distribution function for a standard normal variable. Cornell named the ratio μ_Z/σ_Z the ***safety index*** or the ***reliability index*** and it was denoted by the symbol β

$$\beta = \frac{\mu_Z}{\sigma_Z} \tag{3.7}$$

Probability of failure is then,

$$p_f = \Phi(-\beta) \tag{3.8}$$

The safety index β can be looked upon as a probabilistic factor of safety, and it has gained wide acceptance as a fundamental measure of safety. In general, for a non-linear performance function with n independent random variables the safety index β can be approximated from a first order Taylor series approximation, if

$$Z = g(\bar{X}) = g(X_1, X_2, \ldots, X_n) \qquad 3.9$$

then a Taylor series expansion of the performance function about the mean values gives

$$Z = g(\bar{X}) + \sum_{i=1}^{n} \frac{\partial g}{\partial X_i}(X_i - \bar{X}_i) + \frac{1}{2}\sum_{i=1}^{n} \frac{\partial^2 g}{\partial X_i^2}(X_i - \bar{X}_i)^2 + \cdots \qquad 3.10$$

The derivatives are evaluated at the mean values of the random variables X_i. Also note that $\bar{X}_i$ is labeled as the mean value of the variable. Then truncating after the linear terms, the first order approximation of the mean value of the performance function is

$$\mu_Z \approx g(\bar{X}_1 \ldots, \bar{X}_n) \qquad 3.11$$

If the basic variables are statistically independent then the variance is

$$\sigma^2{}_Z \approx \sum_{i=1}^{n} \left(\frac{\partial g}{\partial X_i}\right)^2 \sigma^2{}_{X_i} \qquad 3.12$$

3.3 Hasofer-Lind Method (Hasofer and Lind, (1974))

The Hasofer-Lind (HL) method defines a reduced variable

$$X'_i = \frac{X_i - \mu_i}{\sigma_i} \qquad 3.13$$

where X'_i is also a random variable with zero mean and unit standard deviation. The original limit state function $g(\bar{X}) = 0$ is transformed to a reduced limit state $g(\bar{x}') = 0$. Where **X'** are the reduced coordinates.

Example (Demonstration of fundamental principles)

Given the limit state function

$$g(X) = X_1 - X_2^2 \qquad 3.14$$

Let X_1~(100, 10) and X_2 ~(5,1). Where the notation, ~ (i,j), indicates a random variable with mean value i and standard deviation j. The probability distributions of the basic variables are not given, or if known they are not used. Using equation 3.13

$$X_i = X'_i \sigma_i + \mu_i \qquad 3.15$$

Substituting 3.15 into 3.14, the reduced failure function is

$$g(X') = X'_1\sigma_1 + \mu_1 - (X'_2\sigma_2 + \mu_2)^2 = 10X'_1 + 100 - (X'_2 + 5)^2 \qquad 3.16$$

The Hasofer-Lind safety index β_{HL}, defined as the minimum distance from the origin of the reduced coordinate system to the failure surface g(**X'**) = 0, is

$$\beta_{HL} = \sqrt{X'^2_1 + X'^2_2} \qquad 3.17$$

Where X'_1 and X'_2 in equation. 3.17 are defined as the reduced coordinates at the design point. Using a simple optimization algorithm the safety index β_{HL} is,

$$\beta_{HL} \approx \sqrt{2^2 + 3.9^2} = 4.4 \qquad 3.18$$

The safety index β_{HL} and other key concepts are illustrated in Fig. 3.1.

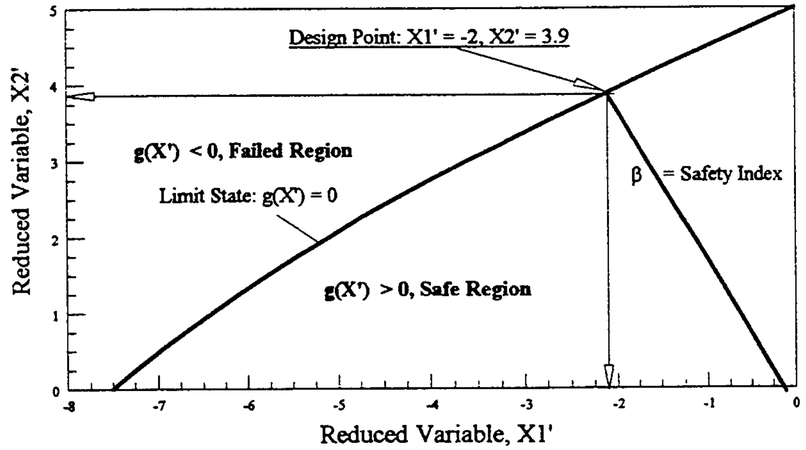


Figure 3.1, Definition of Safety Index and Design Point.

The design point is in part defined by the direction cosine's α_i (Eqn. 19), they are the so-called **sensitivity factors**. The sensitivity factors are a measure of the relative importance of each random variable at the design point. In many applications the sensitivity factors gives valuable information on the relative importance of different design parameters. A sensitivity factor is defined from,

$$\alpha_i = \frac{\partial g(X')/\partial X'_i \cdot \sigma_i}{\sqrt{(\sum_{i=1}^{n}((\partial g(X')/\partial X'_i)\cdot\sigma_i)^2)}} \qquad 3.19$$

Also,

$$\sqrt{\sum_{i=1}^{n}\alpha_i^2} = 1 \qquad 3.20$$

So if a variable has a large α_i, close to one, its relative importance is significant. A small change in that variable would result in a relative large change in the safety index. If, on the other hand, the sensitivity factor is small, close to zero, the random variable could be treated as deterministic. For our example, $\alpha_1^2 = 0.242$ and $\alpha_2^2 = 0.758$, indicating that X_2 is the variable contributing the most to the total uncertainty at the design point.

3.4 Normal Tail Approximation (Rackwitz and Fiessler, 1978)

Many problems in probabilistic design involve a limit state function with non-normal random variables. The probability distribution of these variables may be known, or they may be assumed known. In general there is no exact solution to such a problem, however, reasonable results are often found using the normal tail approximation.

The key assumption in the normal tail approximation is, that at the design point, the basic variable has the same cumulative cdf and probability density function pdf as an equivalent normal distribution as illustrated in Eqn 3.21 and in Fig 3.2.

$$\begin{aligned} F_{X_i}(x_i^*) &= \Phi_{y_i}(x_i^*) \\ f_{X_i}(x_i^*) &= \varphi_{y_i}(x_i^*) \end{aligned} \qquad 3.21$$

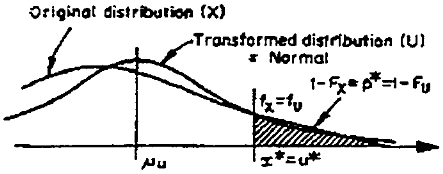


Figure 3.2. Equivalent Normal Distribution

where $\Phi_y(\,)$ and $\phi_y(\,)$ denote the standard normal cdf and pdf respectively.

Algorithm's for the normal tail approximation are well documented in the literature, for example, Madsen, Krenk, and Lind (1986).

4. A Look At The Assumption Of A Deterministic *Geometry Correction Factor* β(a)

4.1 Background

The geometry correction factor β(a) is a key parameter in the calculation of the stress intensity factor K. Considerable efforts has been made towards the derivation of an explicit form of β(a), and for many practical applications this is readily available. When no explicit solution for β(a) is available, it is typically approximated from a linear elastic finite element solution.

In the risk assessment of a structural component or of a structural system it is tempting to treat β(a) as deterministic. The functional form in which β(a) appear does not lend it self to a closed form probabilistic analysis. The computer program PROF (Berens, 1991) does indeed assume β(a) to be deterministic. Intuitively this is not a reasonable assumption, and the effect of treating β(a) as deterministic deserves a closer look.

In the following sections it will be shown a) how the uncertainty in the geometry correction factor can be accounted for using FORM, and b) the effect of treating β(α) as deterministic. The accuracy of the FORM solution will be investigated using Monte Carlo simulation.

4.2 Analysis of Non-Statistical Uncertainty

Suppose the true state of nature is β(a), whose actual realization is unknown. Prediction or estimation of β(a), therefore is necessary; for this purpose, a model denoted as $\hat{\beta}$ may be used. As $\hat{\beta}$ is a model of the real world, imperfection in the model may be expected; the resulting predictions, therefore, will contain an error, and a correction B may be applied. Consequently the true state of nature may be represented by (Ang 1973)

$$\beta(a) = B \cdot \hat{\beta} \qquad 4.1$$

B is then a random variable accounting for the uncertainty in β.

4.3 Probabilistic Fracture Mechanics Using A First Order Reliability Method FORM

4.3.1 Definition Of A Failure Function

Failure is predicted to occur if the stress intensity exceeds the fracture toughness,

$$K \geq K_C \qquad 4.2$$

or

$$g = K_C - K = K_C - \beta(a) \cdot S \cdot \sqrt{\pi \cdot a} \qquad 4.3$$

and the associated probability of failure is

$$p_f = P(g \le 0) \qquad 4.4$$

4.3.2 A Cracked Under Plate In Tension - An Example

For the sake of illustration an axially loaded thin plate is used, as shown in Fig. 4.1 the nominal stress is denoted S, crack depth is a, plate width is b, and thickness (not shown) is t.

The material is 7075-T651 Al and the probability distributions associated with this example are;

$K_{IC} \sim LN(28, 3.4)$ ksi*in$^{0.5}$

a ~ LN(μ_a, σ_a) inch (μ_a will be varied from 0.1 to 0.4 inches. The coefficient of variation $C_a = \sigma_a/\mu_a$ is constant)

S~N(10, 3) ksi

B~LN(1, σ_B) (Accounts for the uncertainty in β(a))

It was shown (not included in the report due to space limitations) that the geometry b and t can be treated as deterministic, i.e. the sensitivity factor $\alpha_b \cong 0$.

LN ( ) and N( ) denotes the lognormal and the normal probability distributions respectively.

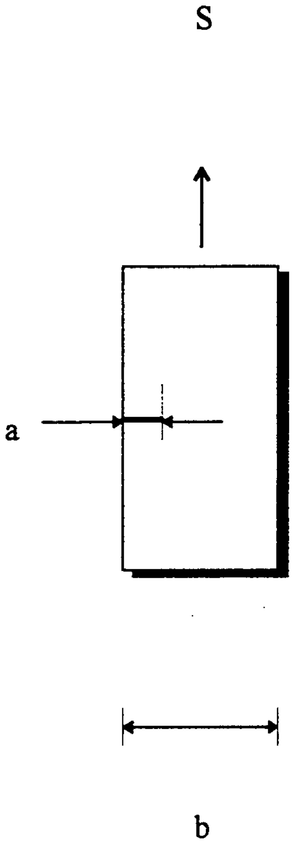


Figure 4.1 Plate Geometry

For the geometry shown in Fig. 4.1 Dowling (1995) proposed the following expression for β(α)

$$\beta(a) = 0.265 \cdot (1-\alpha)^4 + \frac{0.857 + 0.265 \cdot \alpha}{(1-\alpha)^{1.5}} \qquad 4.5$$

where α = a/b

4.3.2.1 The Effect On The Overall Probability Of Failure From The Geometry Correction Factor $\beta(\alpha)$ Being Treated As A Random Variable

From Fig. 4.2 it is shown that uncertainties in the geometry factor does indeed have a significant effect on the overall component probability of failure. A sensitivity study was carried out by assuming coefficients of variations of 0% (i.e. deterministic), 10%, and 20%. For example, at a crack depth of 0.1 inches, it is seen that there are almost four orders of magnitude difference in probability of failure, between treating $\beta(a)$ as deterministic versus treating it as being lognormally distributed with a coefficient of variation of 20%

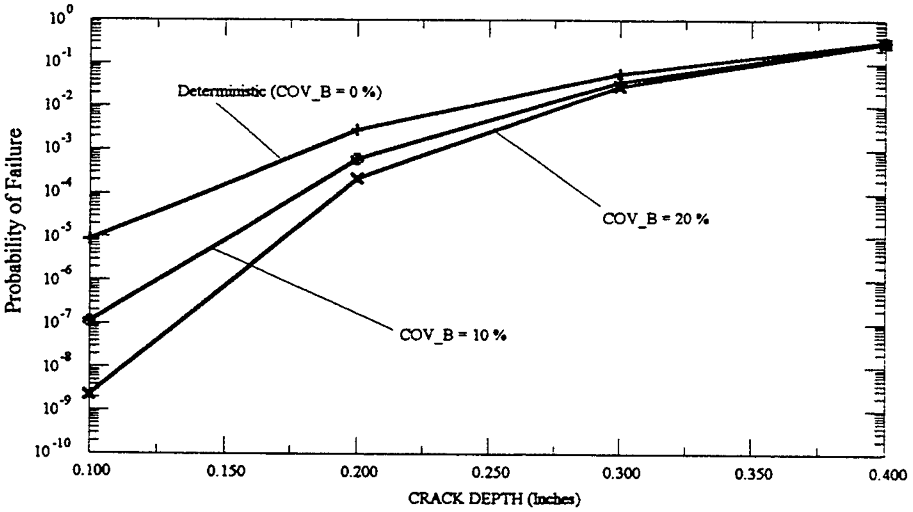


Figure 4.2 Effect Of The Uncertainty In The Geometry Factor

As discussed in section 3.3 the sensitivity factors α_i are very powerful in describing the relative importance of the individual random variables. Recall that the sum of all the individual sensitivity factors is one.

Assuming the geometry correction factor $\beta(a)$ is a random variable, the relative importance of the random variables, fracture toughness K_{IC}, nominal stress S, crack depth a, and geometry correction factor $\beta(a)$ (represented by the bias factor B) is shown in Figure 4.3. At a crack depth of 0.1 inches, it is seen, that the uncertainty in $\beta(a)$ accounts for as much as 43% of the <u>total</u> uncertainty at the design point. It is also seen that crack depth a only account for 7% of the uncertainty. At a crack depth of 0.4 inches, β(a) accounts for 29% of the total uncertainty.

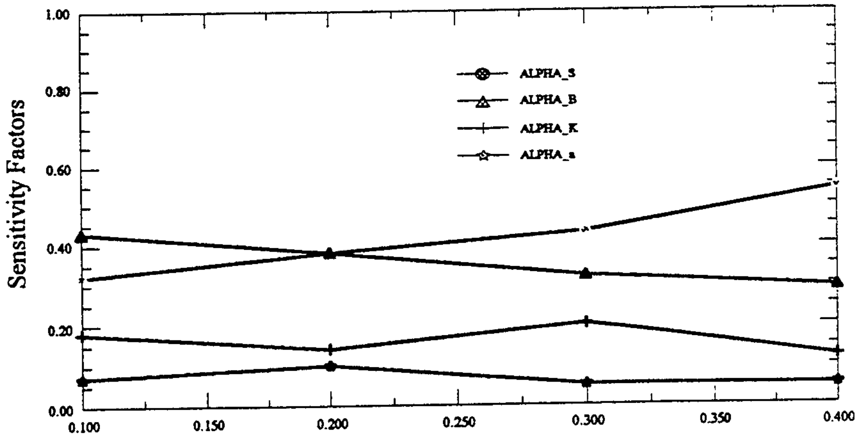


Figure 4.3 Sensitivity Factors For COV_B = 20%

4.3.2.2 Accuracy Of The Probability Estimates.

The accuracy of the normal tail approximation was checked using Monte Carlo simulation, one such check is shown in Fig. 4.4. In all case there was a good agreement between the Monte Carlo simulation and the normal tail approximation. The "exact" probabilities using Monte Carlo simulation was found from a sample size n of two million simulations. The accuracy of the Monte Carlo simulation can be found from the following equation, Ang and Tang (1985),

$$\%error = 200 \cdot \sqrt{\frac{1-p_f}{n \cdot p_f}} \qquad 4.6$$

So for a probability of failure $p_f = 0.0002$ and $n = 2*10^6$ trials, the percent error in the Monte Carlo estimate of p_f is 10.0 %. Or in other words, for $n = 2*10^6$ simulations and the estimated $p_f = 0.0002$ is then in the range; $0.00018 < p_f < 0.00022$. In all cases, the Normal Tail Approximation Method was shown to give excellent estimates of the probability of failure.

4.4 Discussion

A fairly extensive study on the significance of assuming a deterministic β(a) was carried out (not all included in this report). It was demonstrated that β(a) significantly influences the

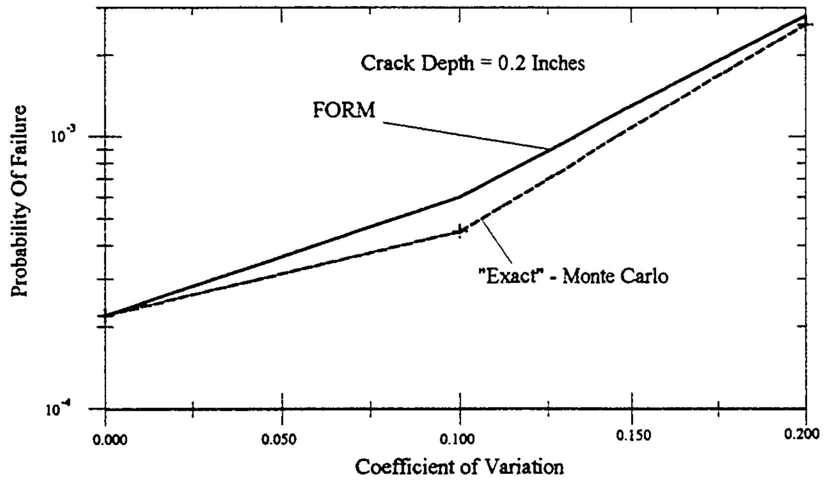


Figure 4.4 Comparison Between Form and Exact (Monte Carlo)

probability of failure, and unless there is strong evidence of β(a) truly being deterministic an underestimation of the probability of failure will result. In other words, to treat β(a) as deterministic is non-conservative.

5 Probabilistic versus Deterministic Crack Growth

5.1 Introduction

In recent years considerable research efforts have been reported on probabilistic modeling of fatigue crack growth. Uncertainties in the loading conditions, the initial crack geometry, and the material properties have been included. Numerical methods for reliability calculations have been developed during the past two decades and the first and second order reliability methods (FORM and SORM) are recognized as both accurate and efficient. Crack sizes are measured either directly or indirectly through a non destructive inspection method. Crack measurements and corresponding crack sizes are also associated with uncertainty and consequently random in nature. Based on the inservice inspection results the estimate of the reliability can be updated and the uncertainty in the initial modeling of the basic variables can be reduced.

In this section a probabilistic crack growth model will be compared to the approach currently used by the software package PROF.

5.2 Crack Growth Model

Paris Equation is often used to describe crack growth,

$$\frac{da}{dN} = C(\Delta K)^m \tag{5.1}$$

where a is the crack size, N is the number of stress cycles, C and m are material parameters and K is the stress intensity factor. K is computed by linear elastic fracture mechanics as

$$K = \beta(a) \cdot S \cdot \sqrt{\pi \cdot a} \tag{5.2}$$

and the various parameters are described in section 2.1. The model applies to both constant and variable amplitude loading. Sequence effects are not accounted for. A probabilistic crack growth model can be expressed as (Ortiz (1987), Wirsching and Ortiz (1987)),

$$\frac{da}{dN} = \frac{C_1}{C_2(a)}(\Delta K)^m \tag{5.3}$$

C_1 is a random variable modeling variation in C from specimen to specimen, while $C_2(a)$ is a stationary log-normal process modeling variation in C within each specimen. The expected value of $C_2(a)$ is taken as one.

Using separation of variables, the differential equation, da/dN, can be solved

$$\int_{a_o}^{a} \frac{C_2(a)}{\beta(a)^m(\sqrt{\pi \cdot a})^m} da = C_1 \int_0^N S^m dN = \begin{Bmatrix} C_1 S^m N \\ C_1 \sum_{r=1}^{N} S_r^{\,m} \end{Bmatrix} \tag{5.4}$$

where the far-right parenthesis represents constant and variable amplitude loading respectively. The variable S represents far-field stress, a_0 is initial crack length, and a is the crack depth at N cycles. C_1, C_2, a, β, m, and S can all be treated as random variables. The material parameters m and C_1 are in general correlated, however, they can be de-coupled using well known transformations from linear algebra.

5.3 Crack Growth Example

The purpose of this example is to (a) demonstrate the difference between treating the crack growth phase as probabilistic and treating it as deterministic and (b) the relative significance of the different variables will be investigated. The accuracy of the analysis will be checked using Monte Carlo simulation.

For the sake of illustration and simplicity, consider a wide panel with a center crack, in this case the geometry correction factor $\beta(a) = 1.12$ = constant, also let $C_2(a) = 1$, then $C = C_1$. Define the failure condition as $a_c < a$ and the corresponding performance function $g = a_c - a$. The probability of failure $p_f = P(g < 0)$. Relevant properties and loads are,

Property	Mean Value	Stand. Deviation	Comment
a_o initial crack length	0.05 in	0.025	Exponential
K_{IC} Fracture Toughness	95 ksi*in$^{0.5}$	10 ksi*in$^{0.5}$	Log-normal
C crack growth parameter	$3.6*10^{-10}$	$1.44*^{-10}$	Log-normal
m slope	3.0	0.3	Log-normal
$\beta(a)$ Geometry factor	1.12		Deterministic
N_T Applied load cycles	Variable		Deterministic
S Applied Load	7.25 ksi	2.9 ksi	Normal

Table 5.1 Variables and associated probability distributions.

The parameters C and m are in general highly correlated. If the degree of correlation is known, a de-coupling using either the Rosenblatt transformation or methods from linear algebra are straightforward.

There is a significant difference between treating the crack growth phase as deterministic and on the other hand treating a_o, C, and m as random variables. This is clearly illustrated in Table 5.2 and in Figure 5.1. At low cycles / small cracks the difference between the two approaches is enormous, however, it is also seen that as the crack length increases and as it approaches the critical crack length a_c, the two methods yield similar results. As a matter of fact, for a completely deterministic crack growth analysis, with $a_0 = 0.05$ in., brittle fracture is predicted to occur at $N \cong 1.1*10^7$ cycles. From Fig. 5.1 it is seen that both the deterministic and the probabilistic crack growth models predict approximately the same safety index / probability of failure at $N = 1.1*10^7$ cycles.

5.5 Discussion

Two possible approaches for the evaluation of the probability of failure of a component with a growing fatigue crack were investigated:

Deterministic Crack Growth			Probabilistic Crack Growth		
Load Cycles N_T	Safety Index β	Probability of Failure p_f	Safety Index β FORM	Safety Index β Monte Carlo	Probability of Failure p_f FORM
$5*10^5$	15.450	$3.772*10^{-54}$	3.503		$2.300*10^{-4}$
10^6	11.117	$5.185*10^{-29}$	2.837	2.847	$2.277*10^{-3}$
$2*10^6$	10.901	$5.700*10^{-28}$	2.414	2.456	$7.889*10^{-3}$
$5*10^6$	9.319	$5.873*10^{-21}$	2.129	2.160	$1.663*10^{-2}$
10^7	4.377	$6.016*10^{-6}$	2.075	2.060	$1.899*10^{-2}$
$1.1*10^7$	1.577	$5.740*10^{-2}$	2.001		$2.270*10^{-2}$

Table 5.2 The difference between deterministic and probabilistic crack growth.

a) Deterministic crack growth

From a given initial crack length a_0, the crack was allowed to grow N number of cycles assuming the crack growth parameters, C, m, and β(a) are all deterministic. Then at N cycles corresponding to a crack length a_F, the probability of fracture was calculated,

$$p_f = P(a_F \ge a_c = K_{IC}\beta(a)\sqrt{\pi \cdot a}) \tag{5.5}$$

b) Probabilistic crack growth

In this case it is recognized that the crack growth rate da/dN is random in nature, the parameters C and m along with the initial crack length a_0 are all treated as random variables. The problem of probabilistic crack growth is difficult and can not be solved using classical methods, however modern approximate methods such as FORM can easily be used. Monte Carlo simulation was used for verification, and as shown in Table 5.2 there was good agreement between the approximate FORM and the “exact” Monte Carlo simulation. The sample size for the Monte Carlo simulation was 10^6.

There is a clear difference between methods a) and b),where method b), probabilistic crack growth, is the more accurate approach. For small cracks the probability of failure is significantly under estimated using approach a).

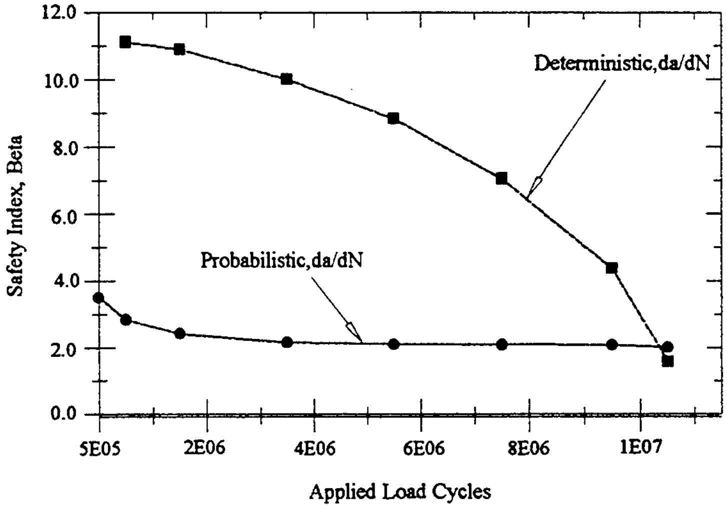


Figure 5.1 Safety Index vs. Load Cycles

6. Model Up-Date Through Inspection

Inspection and maintenance programs for new and aging aircraft have been developed and implemented. Crack sizes are measured either directly or indirectly through non-destructive inspection methods. When the crack is measured indirectly the signal must be interpreted as a crack size. This process is clearly associated with uncertainty and is random in nature. Likewise, a visual inspection has a certain probability of non-detection. Based on in service inspection results, the estimate of the reliability can be updated and the uncertainty in the initial modeling of the basic variables can be reduced.

Although it is relatively straight forward to do, an up-date on the probability of failure resulting form in-service inspection was not considered in this study.

7. Project Summary and Recommendations

Risk assessment methods associated with fatigue and fracture failures in aging aircraft was compared with state of the art methods. It was shown that current methods can significantly underestimate the above mentioned risks. Due to the “exact” mathematical formulations in the methods currently used, such as in the software package PROF, simplifying assumptions are made. These assumptions include a deterministic crack growth process, initial crack length, and geometry correction factor. More modern methods, such as First Order Reliability Methods FORM, although approximate in nature, need not make such assumptions. Methods such as FORM are easy to implement and use.

It is recommended that the ALC's, as a part of their risk assessment of aging aircraft consider the implementation First Order Reliability Methods.

8. References

1. Ang, A,H-S., (1973), "Structural Risk Analysis And Reliability Based Design", *J. Of Structural Division,* ASCE, Vol. 99, No. ST9, September 1973
2. Ang , A.H-S.and Tang, H.T., (1984) "*Probability Concepts In Engineering Planning and Design*", John Wiley & Sons, 1984
3. Berens, B. P., et. al, (1991) "*Risk Analysis For Aging Aircraft*", Final Report Sept. 1987-Jan. 1991, University of Datyon Research Institute, Dayton, OH. 45433
4. Collins J.A. (1993), *"Failure Of Materials In Mechanical Design"*, 2nd Ed., John Wiley & Sons 1993.
5. Cornell, C.A., "*A Probability Based Structural Code*", ACI-Journal., Vol 66, 1969
6. Dowling, N.E., (1995) "*Mechanical Behavior of Materials*", Prentice Hall, Englewood Cliffs, New Jersy 07632, 1995
7. Hasofer, A.M., and Lind N.C., (1974), "*An Exact and Invariant First Order Reliability Format*", Proc. ASCE, J. Eng. Mech. Div., 1974
8. Madsen H.O., Krenk S., Lind N.C. (1986), "*Methods of Structural Safety*", Prentice Hall, Inc., Englewood Cliffs, NJ 07632
9. Rackwitz, R. and Fiessler B., (1977), "*An Algorithm for Calculation Of Structural Reliability Under Combined Loading*", Berichte zur Sicherheitstheorie der Bauwerke, Lab. F. Konstr. Ingb., Munchen, 1977
10. Wirsching P. H., Ortiz K., and Chen Y.N., (1987), "*Fracture Mechanics Fatigue Model In A Reliability Format*",OMAE, Houston TX, March 1987
11. Yao, J.T.P, et. Al., (1986), "*Stochastic Fatigue, Fracture, and Damage Analysis*", Structural Safety, 3, 1986, 231-267

A Feasibility Study of Re-Engineering and Re-Manufacturing Aircraft Structural Components Using Laser Scanning

Joe G. Chow
Associate Professor
Department of Industrial and Systems Engineering
Florida International University
Miami, FL 33199

Final Report for:

Summer Research Program

Robins Air Force Base

Sponsored by:

Air Force Office Of Scientific Research

Bolling Air Force Base, DC

and

Warner Robins Air Logistics Center

September 1996

A Feasibility Study of Re-Engineering and Re-Manufacturing Aircraft Structural Components Using Laser Scanning

Joe G. Chow
Associate Professor
Department of Industrial and Systems Engineering
Florida International University

Abstract

Last summer a feasibility study was initiated to study the applicability of laser scanning to aircraft structural components manufacturing. A sample part, F-15's leading edge rib, was scanned by laser scanners and the scan data was used to create a CAD model at Laser Design Inc. and generate toolpaths at Sharnoa Corp, respectively. A part was then produced by machining an aluminum block using toolpaths generated by Sharnoa. This study results clearly indicate that laser scanning had matured to a stage that they could capture and reproduce intricate surface details typically present in the aircraft structural components. However, in spite of the promising results, the data obtained was not convincing enough for the WR-ALC to implement this new technology on their production floor.

A follow-up study was conducted this year. To broaden this study's scope, two more sample parts, a F-15's canopy fitting and a C-141's forward latch fitting, were also included in this study. A long term plan that can demonstrate laser scanning is capable of quickly and accurately duplicating an existing part was also made. Because of time constraints, this summer's effort was concentrated only on the canopy fitting since it is one of the most difficult parts that WR-ALC had to manufacture. Based on the initial scanning and surfacing results for this part, one can further conclude that laser scanning can significantly reduce the model creation time as well as the skill levels required to create the CAD model. The question of its accuracy when used in the aircraft structural components manufacturing will be answered later by the final results from the canopy fitting and the other two sample parts.

A Feasibility Study of Re-Engineering and Re-Manufacturing Aircraft Structural Components Using Laser Scanning

Joe G. Chow

Introduction

WR-ALC is one of the five Air Logistics Centers of the United States Air Force. One of its primary responsibilities is to manufacture spare and replacement structural parts for military aircrafts. Manufacturing a part by machining at WR-ALC typically starts with a 3-D CAD surface model for the part. Though all the parts that need to be machined are accompanied by 2-D part drawings and, in most cases, by a copy of the actual part, a 3-D CAD model of the part is rarely provided. The CAD model, therefore, needs to be created manually from the 2-D drawings or by directly measuring the part. This modeling process is very labor intensive and time consuming, and it can only be performed by a highly skilled modeler.

Once a 3-D CAD model is created, subsequent activities can benefit greatly. The model can be used during process planning to determine the machining operations required and to select the machine tools, cutting tools and machining parameters for each operation. The NC toolpaths for each machining operation can be generated using the CAD model and a commercial CAM software. The fixtures for holding the part during each machining operation are then designed and manufactured. Fixturing elements can be displayed along with the CAD model to verify fixtures designed and toolpaths generated. The CAD model also facilitates part inspection after it is built.

Currently, at WR-ALC, all the activities in the manufacturing process are performed manually with the aid of computers. Due to the extent of human involvement, it requires on the average six weeks to manufacture parts of medium complexity. In the case of high complexity parts, the turnaround time may take as long as several months. To increase combat readiness and sustainability, it is essential for WR-ALC to reduce the part turnaround time and costs. This requires automation and elimination of some of the steps in the manufacturing process.

Last summer, a feasibility study was initiated to study the applicability of laser scanning to aircraft structural components manufacturing [1]. Two different approaches, re-engineering and re-manufacturing, have been studied. For re-engineering approach, a F-15's leading edge rib, was digitized using a laser scanner at Laser Design Inc. (LDI), Minneapolis MN and the point cloud data obtained was used to create a 3-D CAD model. A comparison between the reconstructed model and the part's blue prints indicated that the errors of the reconstructed model were on the order of 0.005". The reconstructed model could then be used to create toolpaths using a CAM software. Due to time constraints, toolpathing and machining were not carried out for this part of the study.

For the re-manufacturing approach, the model creation step is skipped and the NC toolpaths are generated directly from the point cloud data. By eliminating model creation from the manufacturing process, it allows a part to be duplicated quickly. To verify the accuracy and efficiency of this approach, the same test part, a leading edge rib, was scanned and machined at Sharnoa, Plymouth MI. The duplicated part, as shown in Figure 1, was then compared to the original part using a dial caliper. The comparisons, as shown in Table 1, indicate that, except a few locations, the dimensional discrepancies between the two are within the desired tolerances (0.010"). This study also demonstrated that laser scanning not only dramatically reduces the part turnaround time, but also the skill levels required to reproduce the part.

While the results of both approaches showed great promise of the laser scanning in manufacturing aircraft structural components, additional studies would be required before it can be implemented on the WR-ALC's production floor. The goal of this summer research is, therefore, to continue last year's feasibility study until more conclusive results are obtained.

Based on the experiences learned last year and recent technological developments in laser scanning, some modifications were made to last summer's work in this year's study. These modifications are discussed as follows.

1. During the last year's feasibility study, we mainly concentrated on one sample part: a leading edge rib. While leading edge rib is a good representation of aircraft structural parts, its model creation by

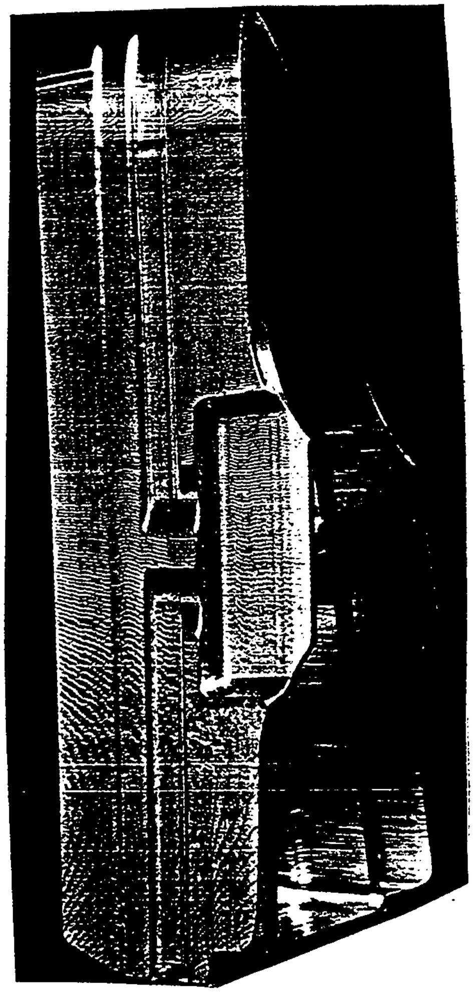


Figure 1. The leading edge rib machined at Sharnoa

Table 1. Comparison of the leading-edge rib machined at Sharnoa to the original leading edge rib

1. Top side (the surface with lofted data)

Overall Width	Original	Duplicated	Error
near the nose	2.794	2.782	-0.012
in the middle	3.267	3.251	-0.016
near the end	2.788	2.776	-0.012

2. Shallow side

Overall width	Original	Duplicated	Error
near the nose	0.885	0.879	-0.006
in the middle	2.985	2.976	-0.009
near the end	4.144	4.135	-0.009

Wall Height	Original	Duplicated	Error
near the nose (left)	0.905	0.898	-0.007
in the middle (left)	0.894	0.886	-0.008
near the end (left)	0.888	0.878	-0.010
near the nose (right)	0.916	0.912	-0.004
in the middle (right)	0.922	0.918	-0.004
near the end (right)	0.904	0.900	-0.004

3. Deep Side

Wall thickness	Original	Duplicated	Error
near the nose (left)	0.131	0.132	0.001
in the middle (left)	0.130	0.131	0.001
near the end (left)	0.381	0.388	0.007
near the nose (right)	0.076	0.081	0.005
in the middle (right)	0.091	0.090	-0.001
near the end (right)	0.081	0.080	-0.001

Overall Width	Original	Duplicated	Error
near the end	4.174	4.161	-0.013

4. Bottom Side

Overall Width	Original	Duplicated	Error
near the nose	2.794	2.779	-0.015
in the middle	3.362	3.347	-0.015
near the end	2.788	2.774	-0.014

All the units are in inches

a skilled operator may not require a significant amount of time. Therefore, the results from this part alone may not be sufficient to prove that laser scanning is capable of significantly reducing the lead time for aircraft components manufacturing. For this summer's work, it was proposed that two other components, a C-141's forward latch fitting and a F-15's canopy fitting, as shown in Figures 2 and 3, be also tested. By including these two new sample parts in this study, most of the common geometrical features found in the aircraft structural components can be captured.

2. In last summer's study, we only interacted with two laser scanner vendors: LDI and Sharnoa. Those two vendors were selected from the list provided by a previous study conducted by Concurrent Technologies Corporation [2]. After contacting all the vendors in the list, LDI was selected for re-engineering study and Sharnoa was selected for re-manufacturing study. However, since last year's study, more scanning vendors, especially those from overseas, have emerged on the US market. In addition, new improvements have been made to laser scanning hardware and software since last summer. We, therefore, re-investigated the scanners offered by the vendors we contacted last year and also studied the ones manufactured by new vendors in this study.

3. Last year, more emphasis was placed on the re-manufacturing approach, and less on re-engineering method. Due to flexibility of design modifications and dimensional verification, re-engineering approach was considered more appropriate for aircraft structural components. Also this approach is more compatible with the manufacturing method currently practicing at the WR-ALC. Therefore, in this year's study, it was decided to first concentrate on the re-engineering approach and, after it is completed, the same set of data will be used to study the re-manufacturing approach.

The objective and long-term plan of this year's feasibility study are further summarized as follows.

Objective, Approach, and Schedule

The objective of this work is to characterize the accuracy and efficiency of laser scanning systems for aircraft structural components manufacturing. The three sample parts, a leading edge rib, a forward latch fitting and a canopy fitting, will be scanned and reproduced. Both re-engineering and re-manufacturing approaches will be studied; first re-engineering and then re-manufacturing.

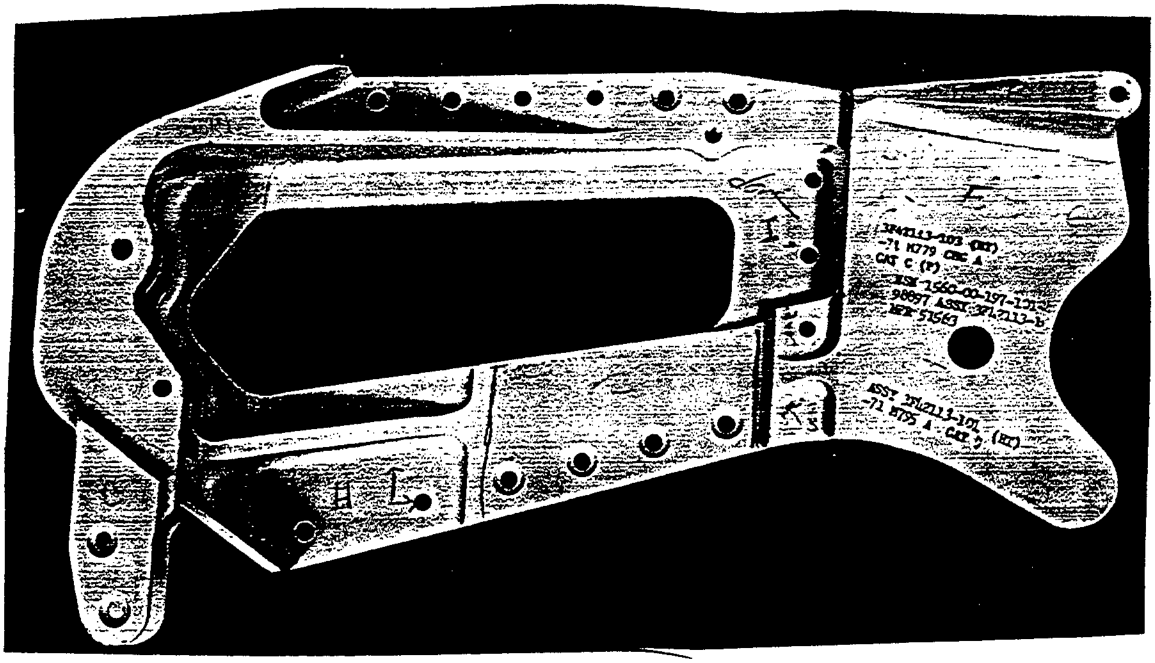


Figure 2. C-141's forward latch fitting

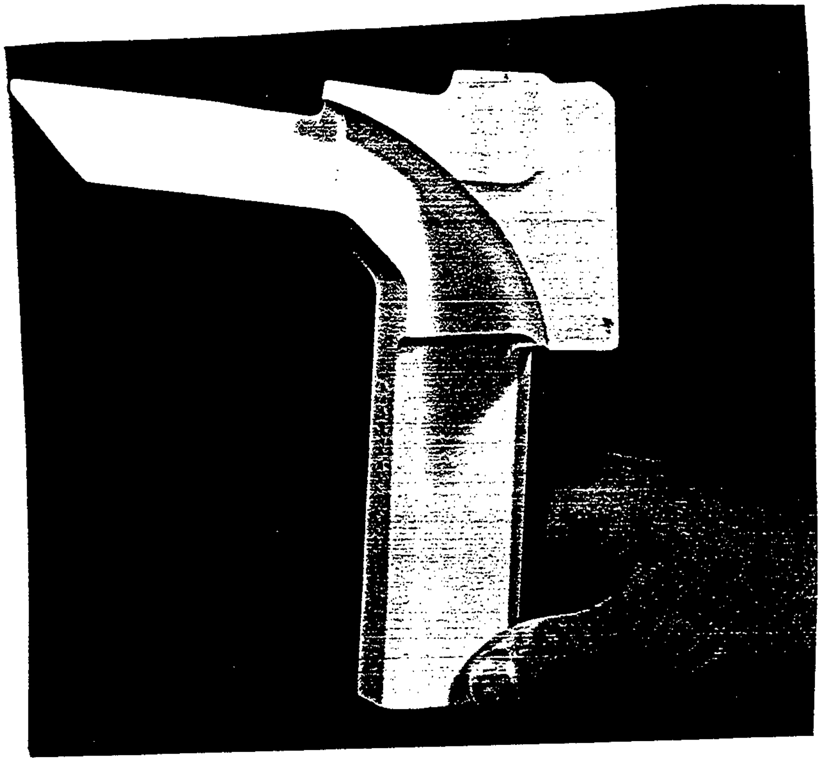


Figure 3. F-15's canopy fitting

For the re-engineering approach, all the sample parts will be scanned by a laser scanner and the scan data for each part will be converted into a CAD surface model by commercial reverse-engineering software. A part will be machined using toolpaths created from each reconstructed surface model. A comparison between the duplicated and the original part will reveal the accuracy of the scanning and modeling errors.

After the re-engineering study is completed, the laser scan data for the sample parts will be used to generate toolpaths without going through model creation. Machining will be performed and the duplicated parts will be compared to the original parts. Also, by comparing the parts produced by the re-manufacturing approach to the ones made by the re-engineering approach, one can determine the effect of model creation on part accuracy and quality.

The results will be summarized and stored in a table which displays the accuracy and consistency of the laser scanning systems for various geometric features. Results will also include records of time utilization for each phase, such as setup, scanning, editing, model reconstruction, and toolpathing from the scan data. Estimates for the savings achieved through new improved processes over the traditional method will also be made.

It is estimated that this work will be completed by the end of 1997. After this summer, this research will be continued at Florida International University (FIU) until its completion. At this moment, the scanning of all three sample parts have been completed and they are being converted into surface models by the STRIM software at FIU. To produce and inspect machined parts, a 3-axis milling machine and a Brown and Sharp CMM, expected to arrive at FIU in September '96, will be used.

For this project, the most important tools are: a laser scanner and reverse engineering software. These tools are very different from those traditional equipment and software. Consequently, a great deal of efforts has been spent in investigating their capabilities and characteristics since last summer. The investigation results and findings are what follows.

Hardware and Software Tools

The most critical hardware for this research is a laser scanner. The main purpose of the scanner is to acquire XYZ dimensional data from a planar or complex free form surface. The main criteria for selecting a laser scanner are accuracy, speed, ease of operation, and price. To make an in-depth evaluation of different scanners, the principal investigator visited six scanner vendors and conducted a scanning experiment for one of the WR-ALC's sample parts at their sites. Based on the experimental procedure used and the scanning data obtained, a comparison of different scanners, as summarized in Table 2, has been made. The main difference between the scanners listed in the Table 2.a and 2.b is that the ones in the Table 2.a are stand-alone machines while the ones in Table 2.b contain only the laser sensor and should be integrated with an existing coordinate measuring machine.

The main software tool required for this work is a reverse engineering software, which will convert scan data into a surface model. The reverse engineering software differs from the traditional CAD software in that it is capable of processing a huge number of points. The ASCII file of a laser scan point data is typically very large - up to 50 mega bytes. The traditional CAD software, such as ProEngineering and Autocad, are designed to handle entities such as lines, curves, surfaces, and solids, instead of a huge volume of point data. Therefore, they are not suitable for this research.

Currently there are two software that are most popular for reverse engineering applications. The first one is the *Surfacer* developed by Imageware Corp, Ann Arbor MI. Surfacer was designed specifically for reverse engineering applications. Thus, it can handle any size of data points and possesses very powerful point processing capabilities. Other important features the Surfacer has are that it can compare scan points to a CAD model and a CAD model to another CAD model. These features allows it to be used as an inspection tool and to verify the accuracy of a reconstructed CAD model by comparing it to the original CAD model.

Another popular reverse engineering software is the STRIM marketed by Matra Datavision, Andover MA. The STRIM software, designed originally for surface modeling and styling, is one of the most powerful software for creating surface models. Recently, the software developer added a new module

Table 2.a Comparison of Stand-Alone Laser Scanners

Manufacturer	**Digibotics**	**Laser Design**	**3D Scanners**	**Sharnoa**
Model	Digibot II	Surveyor 1000	Replica 500	HPM 55
Total Price	$50,500	$50,000	$52,500	$240,000
Accuracy	0.002"	0.001"	0.002"	0.001"
Work Volume (Max. Part Size)	18"x18"x18"	12"x12"x12"	20"x20"x18"	55"x40"x28"
Light Source	Point	Stripe	Stripe	Point
Scanning Speed (points/second)	100	2,000	1,000	200
Manufacturer's Location	US	US	UK	Israeli

Table 2.b Comparison of Laser Probes

Manufacturer	**Kreon Industries**	**Hymarc**
Model	KLS 25	Model 25
Total Price (Inc. installation)	$40,000	$100,000
Accuracy	0.001"	0.001"
Work Volume (Max. Part Size)	Same as the CMM's work volume	Same as the CMM's work volume
Light Source	Stripe	Flying spot
Scanning Speed (points/second)	1,200	10,000
Manufacturer's Location	France	Canada

“Surface Reconstruction” for converting the point data into a surface model. Even though its point processing capabilities are not as good as the Surfacer, it does tend to create better quality surface models. A distinct advantage of the STRIM software is that it has “digital point machining” capability, which allows it to generate toolpaths directly from the scan data without ever creating a CAD model.

1996 Summer Research

Since the entire study can’t be completed during the summer of ‘96, it was decided to concentrate on just one of the sample parts: the canopy fitting. This component is one of the most complex ones that WR-ALC’s NC Machine Shop ever had to manufacture. It took their NC programmers several months to model the part and several more months to machine the part. If this part can be modeled and machined successfully using laser scanning, its results will be most convincing to the Robins’ managers and NC programmers.

Because of its complex geometry and limited time during the summer, it was determined to conduct this part of the study in a most efficient fashion, i.e., each task required to duplicate a canopy fitting would be performed by the most qualified people/site. By doing this, the best possible results can be obtained in a shortest amount of time. The tasks and the site that will perform each task for the canopy fitting are listed in the following :

1. Scanning at Hymarc Inc.
2. Surfacing (modeling) at Imageware
3. Manufacturing a plastic part at a selected service bureau
4. Compare the plastic part to the original part at WR-ALC/TIELE
5. Machine a metal part at the WR-ALC’s NC machine shop
6. Compare the duplicated (metal) part to the original part at WR-ALC/TIELE

The most qualified place to machine a canopy fitting is the WR-ALC’s NC machine shop since they possess the right machine tool, fixtures, and experiences. However, to machine a canopy fitting at the WR-ALC’ NC shop, it requires approximately 320 hours after a reconstructed surface model is provided to them. To convince WR-ALC's managers to invest this amount of time for this study, a solid evidence that laser scanning technology will work would be required. The third task in the above list, manufacturing a plastic part using stereolithography process, is included in this study to serve this purpose. A comparison between the plastic part and the original metal part will indicate the accuracy of scanning and modeling processes. If the comparison is very favorable, the actual machining will be conducted at the WR-ALC's NC shop.

At the time of this writing, scanning for the canopy fitting has been completed, surfacing is being performed at Imageware, and the other tasks are being investigated or planned. The current status of each activity is provided in the following sections.

Scanning at Hymarc

The scanning of a canopy fitting was performed at Hymarc Limited on July 10, 1996 by the following procedure:

1. Part Setup. Based on the geometry and the capabilities of their scanner, Hymarc engineers decided that two different setups would be required to capture all the geometric features of the canopy fitting. For the first setup, the part, supported by the modeling clay, had similar orientation as the one shown in Figure 3, while in the second setup the part was rotated 180 degree. After scanning, the geometries captured during these two setups would be merged into one point cloud data for the entire part. To ensure the accuracy of this merging process, three tooling balls were attached to the part and were scanned as part of the canopy fitting's geometry.

2. Scanning The Hymarc scanner is capable of capturing surface data at a continuous rate of 10,000 points per second and delivering 0.001" accuracy. After examining the geometric features of the canopy fitting, a 0.1 mm x 0.7 mm grid was chosen. The scanning probe was moved in two rotational axes to enable data capture from different views around the object. Data from any number of multiple views was automatically integrated into a common reference coordinate system and was mapped directly to a large SGI monitor for detailed on-line manipulation and inspection. At the end of scanning, a point model was obtained as high resolution cloud data for that particular setup. The total scanning time for the entire part was approximately 5 hours: 3 hours for the first setup and 2 hours for the second one.

3. Postprocessing After scanning, the files were brought into the Surfacer software and were further examined for any missing data. Within Surfacer, the tooling balls and the modeling clay were edited out and the empty regions under tooling balls were filled in manually. By merging the top and bottom scans, a clean complete point model, as shown in Figure 4, was obtained.

- **Surfacing at Imageware**

After the part was scanned at Hymarc, the scan data was sent out to the Imageware for surfacing. At Imageware, the scan data is being converted into a surface model using the Surfacer software and a STL file will be created for stereolithography process. The estimated time for surfacing a canopy fitting is three to four days. A reconstructed CAD model will allow one to compare some of the dimensions in the model to the corresponding dimensions in the physical part. The comparison will reveal scanning and modeling errors.

Manufacturing a plastic part by Stereolithography (SL) process

The stereolithography process can translate a CAD model quickly and accurately into a solid object by directing ultraviolet laser radiation onto a vat of liquid photopolymer. Where the polymer is sufficiently exposed to the laser's beam, it hardens down to a regulated depth. After each pass of the laser, the platform on which the exposed polymer sits is lowered, submerging the hardened material under the surface of the liquid in the vat. Another pass of the laser then hardens the next layer of polymer. Layer by layer, the so-called "SLA" part is built up through repeated passes of the laser. By comparing the reconstructed CAD model to a SLA part, one can determine the errors caused by the rapid prototyping process.

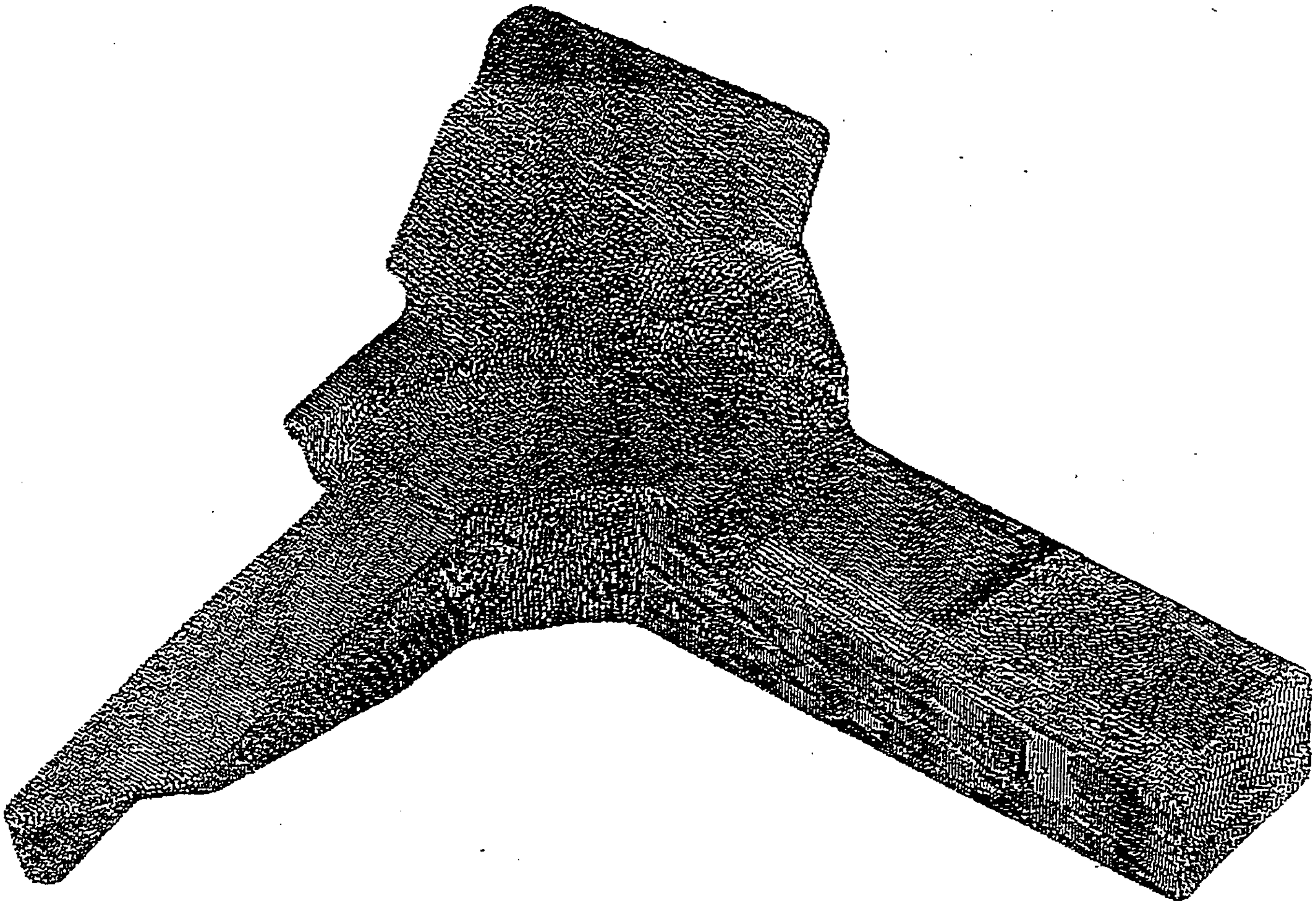


Figure 4. Canopy fitting's point cloud data

Part Characterization at WR-ALC/TIELE

Two different methods are available for checking the dimensions of two identical parts: using a manual device, such as dial caliper, and a coordinate measuring machine (CMM). With a CMM, measurements can be taken from all the part surfaces, especially those complex free form surfaces. The dimensional verification lab (TIELE) at WR-ALC has extensive experiences in measuring complex parts using a coordinate measuring machine. They agreed to provide assistance that WR-ALC/TILAA needs in comparing a physical part to another. At this moment, they are attempting to establish a reliable procedure for comparing a canopy fitting to another one using two metal canopy fittings. This procedure, once developed, will be used in Tasks 4 and 6, as discussed in the '96 summer research tasks outline.

Error analysis for the SLA part

The SLA part will consist of several types of errors: scanning, modeling, and rapid prototyping error. Our main interest is to determine the magnitude of the scanning and modeling errors that exists in the reconstructed CAD model. By comparing the reconstructed model to the original part, one can get some idea about the scanning and modeling errors. However, this type of comparison will be limited to some dimensions that are easy to measure. The best comparison one can make is to compare a physical part to another one using a coordinate measuring machine. For an SLA part, it contains the rapid prototyping errors (~ 0.005") which will not appear in a machined part. Therefore, the following three scenarios are possible when a plastic canopy fitting is compared to a metal canopy fitting:

1. If the SLA part tolerances are less than 0.010", it means that the scanning and modeling errors are very likely less than 0.005". When the estimated machining error (0.002") is included, a machined component may be well within the desired tolerance (0.010"). This is the most desirable case and its result will be most convincing to the WR-ALC's managers.

2. If the worst error in the SLA part is between 0.010" and 0.015", the scanning/modeling errors are likely to be between 0.005" and 0.010". When the machining error is added, the machined part may

still be within the tolerance.

3. If the SLA part error is larger than 0.015", it could well mean that both the scanning/modeling errors and the SL process errors are larger than the vendors claimed. For this case, one can't tell whether the excessive errors in the part come from the scanning/model or rapid prototyping.

Conclusions

Based on last year and this year's work, one can clearly see that laser scanning can, indeed, reduce the lead times and skill levels required to duplicate aircraft structural parts, especially the complex ones. For parts with almost any kinds of geometries, scanning can usually be done in less than one day and surfacing in less than one week. Thus, one of the advantages in using a laser-based reverse engineering system is that the model creation time is basically independent of the part geometry. Also, the skill levels required for scanning and modeling are also fairly low, compared to those required to manually create a CAD model. For scanning, two weeks of training will be sufficient to run a scanning test for a very complex part. For surfacing, a longer training time, such as three to six months, would be required, which is still much less than five to ten years' experiences necessary for creating a CAD model manually.

The key question remaining for laser scanning is the accuracy of the aircraft structural components it produce. This question will be first answered by the canopy fitting result and later by the results from the other two sample parts.

Bibliography

1. J. Chow and Ramesh M. C., "A Feasibility Study for Re-Manufacturing Aircraft Structural Components Using Laser Scanning And X-Ray Computed Tomography", Technical Report to Air Force Office of Scientific Research, September 1995.

2. "Re-Engineering of Aircraft Structural Components Using Laser Scanning", Concurrent Technologies Corporation, 1994.

AN ASSESSMENT OF THE CURRENT TREND CAPABILITIES OF THE MH53J VIBRATION MONITORING SYSTEM

S. A. McInerny
Associate Professor
Department of Mechanical Engineering

The University of Alabama
P.O. Box 870276
Tuscaloosa, AL 35487-0276

Final Report for:
Summer Faculty Research Program
Warner Robins Air Logistic Center

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC
and
Warner Robins Air Logistics Center

August 1996

AN ASSESSMENT OF THE CURRENT TREND CAPABILITIES OF THE MH53J VIBRATION MONITORING SYSTEM

S. A. M^c^Inerny

ABSTRACT

The findings of a review of the vibration monitoring program on the MH53J Special Operations Forces aircraft is summarized. The new data acquisition unit permanently installed on the MH53J aircraft, called the VMS, performs three different functions: track and balance; acquisition of vibration trend data; and acquisition of monitor data. This review focused on the vibration trend capabilities of the program, as success of envisioned enhancements to the system rests on trending capabilities. Shortcomings in the trending aspects of the current system are identified. These include a lack of clearly defined expectations for the vibration trending program, against which its performance might be judged. Furthermore, while trend data is regularly acquired on aircraft at Kirtland Air Force Base, this is true only for selected aircraft at Hurlburt Field. There is almost no trend data being acquired on aircraft at Osan and Mildenhall. On the technical side, the locations of the vibration sensors are not optimal and no spectral averaging is performed. The latter two factors are suspected of being the major contributors to the large variation seen in trend levels measured under nominally identical mechanical and operating conditions. Due to sensor type, location, and mounting, as well as the lack of zoom analysis capabilities, H53 gearbox and bearing diagnoses cannot be adequately performed with the VMS as currently configured. Improvements in effective aircraft maintenance and aircraft sustainability can be expected if these system shortcomings are addressed.

In order to improve the trend capabilities of the vibration monitoring program, it is recommended that clear, specific, realizable program goals be developed. These should include a detailed list of those mechanical components and fault conditions that are to be monitored. A stepped plan for improving the quality of the trend data is presented. Regardless of changes that might be made in hardware (e.g., sensors, data acquisition unit), software (e.g., the database program that tracks the trend data), or personnel, the effectiveness of the vibration monitoring program ultimately depends on comprehensive program oversight. Responsibility and, by implication, culpability for cost effective program growth and performance should assigned to an appropriate Air Force representative. The program must have the support of impartial technical expertise that remains abreast of current developments in the field. This is especially critical in a funding environment where open bids are difficult to issue and the incorporation of new program capabilities into existing contracts is the only practical option

AN ASSESSMENT OF THE CURRENT TREND CAPABILITIES OF THE MH53J VIBRATION MONITORING SYSTEM

S. A. M^{c}Inerny

I. INTRODUCTION

The Air Force MH53J helicopters used by the Special Operations Forces are outfitted with a vibration monitoring system whose genesis can be traced back to early work performed by the Navy (Pensacola). The aircraft system consists of 12 velocimeters with cabling to a signal conditioning / data acquisition unit. Originally, the velocimeters had to be installed and cables run each time measurements were to be made. The velocimeters are now permanently installed on standard brackets. The data acquisition unit has recently been upgraded to one that is also permanently installed on the aircraft; data is downloaded from the new unit via a data transfer unit (DTU) and imported into a database program called Vibralog. Fewer than a quarter of the Pave Low fleet have been outfitted with the new, permanent data acquisition unit, known as the vibration monitoring system (VMS). The remainder of the fleet still utilizes the older 8500C unit, which must be carried out to the aircraft and temporarily installed in order to acquire data. Both the 8500C and VMS data are imported into the Vibralog database.

Continued developments in the Pave Low vibration monitoring system have been championed by personnel at Kirtland Air Force Base (KAFB), with substantial funding provided by the LUJE program out of Warner Robins Air Logistics Center (WR-ALC). The Vibration Program Management Office (VPMO) at Kirtland provides support for vibration monitoring efforts on many different Air Force aircraft in addition to the MH and TH53 helicopters. The VPMO staff include two Air Force Engineering Technical Support personnel and 3 contractor personnel. The office provides assistance in the installation of vibration monitoring systems and in interpretation of vibration data. They are also tasked with maintaining the fleet wide MH53J database. One of the most important services provided by the VPMO is training: in track and balance; in the use of Vibralog; and in the use of the 8500C and VMS systems.

Several enhancements to the current vibration monitoring system have been envisioned. These include an "expert shell" or "neural network" that would digest the vibration data and provide concise instructions to maintenance crews for any required remedial actions. Such a system would make it unnecessary for maintenance personnel to review and interpret the vibration data, allowing for less highly trained maintenance crews and faster, more timely fault indications. A second major development would be the incorporation of flight regime recognition software into the flight data recorders that are to be installed on the MH53J fleet. Information on flight regimes could then be correlated with the vibration data continuously recorded by the new VMS unit, allowing for improved estimates of structural fatigue exposure. Phase II of the VMS contract provides for the investigation of modest extensions to the existing VMS system, with the addition of sensors on several tail rotor drive shaft bearings. The VMS unit has the capacity for an additional 16 channels, which could include sensors on bearings and gearboxes.

In view of the potential transition of the MH53J vibration monitoring system from what was originally an informal Naval program to a sophisticated structural monitoring system linked to the flight data recorder, a comprehensive review of the existing system was requested. This report summarizes the findings of an evaluation of the current MH53J vibration monitoring program performed during the period 20 May 1996 to 9 August 1996 under the auspices of the Air Force Office of Scientific Research (AFOSR) Summer Faculty Research Program.

II. SCOPE OF STUDY

It is important to make a clear distinction between three very different tasks that are currently performed with the VMS: Track and Balance; Acquisition of Vibration Trend Data; and Acquisition of Monitor Data.

Track and balance. The present study does not include a review of the track and balance procedures. According to both KAFB and WR-ALC personnel, the procedures using the 8500C (which utilize two velocimeter signals, a tachometer signal for balancing and a strobe for blade tracking) have provided a substantial improvement in the time required to accurately track and balance the main and tail rotors. In early testing, a recently developed optical tracker has proven to be even more effective, but the necessary software algorithm has not, yet, been incorporated into the VMS unit. (The 8500C and VMS track and balance procedures require a limited number of vibration sensors and the acquisition and storage of a relatively small set of data.)

Acquisition of Vibration Trend Data. Vibration trend data is obtained from 12 velocity sensors (integrating accelerometers) mounted at the locations and in the orientations indicated in Table A-I of Appendix A. Trend data consists of vibration spectra acquired during three different flight regimes: Ground, Hover, and 120 knot Forward Flight. These spectra are recorded by the 8500C or the VMS and then transferred to a PC where they are imported into a database called Vibralog. The levels of the peaks in these spectra, at frequencies which correspond to shaft rotation rates and rotor blade passage frequencies, amongst others, are then available for long term trending and comparison to *trend* alarm levels. Table A-II of Appendix A provides the current trend alarm levels applicable to the Ground regime. Similar tables for Hover and Forward Flight can be found in Reference 1.

Acquisition of Monitor Data. The new VMS unit, but not the 8500C, has numerous capabilities for the acquisition of random spectra and "monitor" data. As currently configured, the monitor mode works as follows:

1. Two alarm severity levels are set for specified frequency bands and sensor locations. An alarm can be either an overall root mean square (rms) level or a peak level (corresponding to a single frequency bin of width equal to the frequency resolution) in the vibration spectrum.
2. From the time the main rotor comes up to speed until the end of the flight, monitor data in the form of peak and rms levels in each of the defined alarm bands, as well as the RPM (frequency) to which the peak level corresponds, are stored in the VMS unit. These values are recorded every monitor "sweep cycle," or roughly every two minutes. The three values (for each defined alarm band) can be later plotted in Vibralog, in what is essentially a plot of value versus time.
3. During a flight, if a measured level exceeds the designated alarm level, for a period of time longer than a pre-set alarm "trigger time," an exceedance advisory is recorded. This advisory indicates the number of the alarm band in which the exceedance occurred (which identifies the sensor and frequency band), the maximum level measured in the band during the time that the band remained in alarm, but after the elapsed trigger time, and the spectrum (also stored) in which this maximum level was measured.

4. Alarm status is maintained until the level in the band has remained below the alarm level for a duration equal to the "release time."

This study focuses primarily on the trending capabilities of the vibration monitoring system. There are several reasons for this. First, there appears to be widespread satisfaction with the track and balance procedures, particularly the new optical blade tracker. Secondly, the success of the envisioned system enhancements rests on the consistency and reliability of the system's trending capabilities. Lastly, the characteristics of the trend data must be more clearly understood before reasonable expectations can be established for the monitor data.

III. METHODOLOGY

To begin, a clear definition of the goals of the vibration system was sought from WR-ALC and KAFB personnel involved in the funding, management, and implementation of the system. One of the specific questions asked was "What are the components and mechanical conditions on the MH53J aircraft that the vibration monitoring program is intended to track?" The phrasing of this question reflected the author's impression that the vibration monitoring program constituted part of a proactive maintenance program. This impression can be attributed to frequent references to data "trending" and to hopes, expressed by H53 program personnel, of implementing bearing and gear condition monitoring techniques.

To obtain hands on experience with the system, a trip was made to Kirtland AFB to look at aircraft, system hardware and data in the database. A copy of the database was not, then, available at WR-ALC. A preliminary review of the archived H53 data was made at that time and arrangements were made for a copy of the database to be installed at WR-ALC at the conclusion of the trip. Also during this trip, numerous questions on technical details of the VMS and 8500C data acquisition units were addressed to manufacturers representatives (who were present for tests on the new optical tracker).

Back at WR-ALC, the "trend" data in the database was reviewed in detail and compared to the expected behavior of data used to trend developing faults in machinery. Evidence of changes in vibration levels consistent with maintenance actions was also sought. Following this, a second trip was made to KAFB. The purpose of this trip was twofold. First, an explanation was sought of exactly how the trend data is interpreted by the VPMO and what the resulting maintenance actions are. Secondly, instrumentation including a two channel spectral analyzer, an impact hammer, accelerometers, and a portable accelerometer calibrator (all of which were sent ahead) were used to investigate a number of technical questions on the acquisition and processing of the vibration data.

IV. RESULTS

The goals of a proactive maintenance program include extended service life, reduced downtime and reduced maintenance costs. The first step in an effective proactive maintenance program is the identification of those components considered most critical, either due to purchase and / or installation cost or potential downtime due to failure. The costs of implementing the necessary monitoring methods must be weighed against the benefits. An ideal proactive maintenance program requires coordination, cooperation and feedback between management, engineering and maintenance staffs.

The H53 vibration monitoring program is not part of a true proactive maintenance program. It is unreasonable to expect that such a program, requiring close coordination between management, engineering and maintenance and detailed tracking of parts and maintenance actions, could be implemented on a fleet of heavily utilized, mission critical, Air Force helicopters located at geographically diverse bases in New Mexico, Florida, England and Korea. Nonetheless, improvements in aircraft sustainability can be expected if system shortcomings are addressed. *Note that track and balance is a large and important part of the MH53J maintenance program and the benefit of the 8500C and VMS track and balance capabilities is not questioned.*

A. MANAGEMENT ISSUES

There is no clear definition of the expectations of the trending aspects of the existing vibration monitoring system. Impressions of and enthusiasm for the trending capabilities of the system vary widely from person to person and base to base. Lack of comprehensive tracking of parts and maintenance costs makes it difficult to identify which components would be the best candidates for vibration. (The remedial measures listed in Table 5-2, page 5-5, of Reference 2 indicate a number of mechanical problems that are thought to be reflected in the currently acquired trend data.)

The vibration monitoring system has never been part of a formal Air Force program and the success of the system's track and balance functions are widely recognized. But, no one has been tasked with responsibility for the effectiveness of the system's trend capabilities and the vibration monitoring program has matured to the point where further beneficial developments cannot be insured without knowledgeable program oversight.

There is a very high level of enthusiasm for the H53 vibration monitoring system at Kirtland Air Force Base. After Kirtland, the H53 base with the highest level of participation in the vibration monitoring effort is Hurlburt. Even so, only selected Hurlburt aircraft are consistently trended and there is, presently, an extremely low level of system utilization at Osan and Mildenhall. The VPMO staff believe that the need to carry the 8500C out to the aircraft and hook it up before data acquisition can occur contributes to the lack of timely, periodic acquisition of trend data. There is no doubt that the acquisition of trend data is simpler with the VMS unit. However, it is still necessary for crews to maintain the aircraft in the three trended flight regimes and someone (currently a crew chief) is still required to initiate the data acquisition process.

B. TECHNICAL ISSUES

The changes in the mechanical condition of a variety of machinery components can be assessed by trends in vibration data when

1. data is acquired in a consistent manner, i.e. using the same sensors, at the same locations, in the same orientations and under nominally equal operating conditions;
2. data acquired in a consistent manner shows a reasonable level of repeatability for nominally equivalent mechanical health and machine operating conditions;
3. the data displays a discernible pattern indicative of the development of a specific mechanical fault condition.

As a general rule, based on experience in industry, an imbalance condition can be detected by a vibration measurement taken anywhere on a machine. Accurate diagnoses of shaft misalignment may require more than one vibration measurement, in radial and axial orientations, made as close as possible to the load path (on or near bearings) on either side of a coupling. Measurements to detect developing gear and bearing faults should be made with accelerometers, as opposed to velocimeters, because of the greater emphasis that acceleration measurements place on high frequency vibrations. Attention must be paid to the mounting of the sensors in order to avoid a reduction in their useable frequency range. Data should be acquired as close as possible to the source to minimize any attenuation or resonant amplification effects. If there is an extended distance between the source and sensor, large changes can occur in monitored vibration levels that have nothing to with source conditions, but are due to changes in the structural transmission path. (For more information, see References 6 and 7.)

1. Sensors Locations, Type and Mounting

With the exception of the sensors on the main gearbox and tail rotor gearbox, the velocimeters on the H53 aircraft are located on the interior airframe or skin (nose). The origin of these sensor locations dates back to the early days of the program at the Navy base in Pensacola and to tests performed at KAFB in 1985.[2] Easy access and a desire to measure and verify vibrations reported by the crew may have influenced the selection of some of these locations. At each sensor location, the frequencies selected for monitoring appear to those that display peaks in the measured spectra. The sensor locations on the H-53 aircraft do not satisfy the criteria that they be located on a structural load path as close as possible to the source being monitored.

The sensors themselves are velocimeters. The specifications sheet for the velocimeters[3] states that they have a minimum resonance frequency of 20,000 Hz, but the frequency response curve extends only up to 100 Hz (6000 RPM). As a general rule, sensors that are hard mounted (stud mounted) to a structure have a frequency response curve that is flat up to one-third of their resonance frequency. For the H53 velocimeters, this is ostensibly 7000 Hz (420,000 RPM). However, the H53 velocimeters are not directly stud mounted to the structure. They are mounted on brackets. In the case of the velocimeters in the nose and on the airframe, the distance of these sensors from the vibration sources being monitored is likely to place a greater limitation on the useful frequency range of their data than do their bracket mounts. The velocimeter on the main rotor gearbox is located on a bracket attached to a flange at the top of the gearbox housing. This is not the optimal location, mounting, or sensor type if one intends to monitor gearmesh frequencies. The velocimeter mounting bracket on the tail gearbox is particularly poor. Tap tests indicate that the velocimeter loaded bracket on the tail rotor gearbox has its first resonance frequency at 1275 Hz (76,500 RPM), see Figure A-1. Curiously, a spectral peak corresponding to the gearmesh frequency of the intermediate gearbox (82,000 RPM) is picked up and monitored at the tail rotor gearbox.

2. SIGNAL PROCESSING

Details of the processing of the velocimeter signals by the data acquisition units differ between the 8500C and VMS. Some of the issues of interest are the spectral frequency ranges, the number of lines of resolution, the type of windowing applied to the time domain data before transformation to the frequency domain, the dynamic range and noise floor, and the number of averages performed in the spectral calculations. Sets of low and high frequency trend spectra are acquired and archived in the Vibralog database. The information that follows has been pieced together from a variety of documents, word of mouth, and inferences drawn from the format of the data in the database. Some inaccuracies may be expected.

Frequency Range and Resolution of Trend Spectra. The 8500C has selectable frequency ranges extending from 0 Hz up to a maximum of 10,000 Hz (600,000 RPM). The number of lines in the spectra calculated by the 8500C is 461. Both the 8500C and VMS data acquisition units utilize a 2 Hz (120 RPM) high pass filter. After the high pass filter, the 8500C calculates a 461 line spectrum extending from 0 to the selected upper frequency limit (if the upper frequency limit were to exceed 7.5 kHz, a 400 line spectrum would be calculated).[4] When this data is imported into Vibralog, it is truncated at the high frequency end due to the fact that the database accepts only 400 of the 461 spectral lines. As presently configured,[5] the VMS calculates a 400 line low frequency spectrum, which does not require truncation when imported into Vibralog. Depending on the "high pass filter" (an inaccurate label) setting in Vibralog, the values of a certain number of low frequency spectral lines (both the 8500C and VMS) are, essentially, set to zero upon importation into the database.

The earliest data in the Vibralog database includes low frequency spectra from 120 RPM to 10,200 RPM, with a frequency resolution of 25.05 RPM. This was later changed to a range of 150 RPM to 12,000 RPM with a resolution of 30 RPM This frequency range and resolution are maintained in the new VMS low frequency spectra. The frequency range and resolution used in the low frequency trend spectra are appropriate for monitoring the drive shaft RPMs and blade passage frequencies below 12,000 RPM.

The earliest 8500C high frequency spectra in the Vibralog database extend from 675 RPM to 90,000 RPM with a frequency resolution of 250 RPM. This was later changed to 3000 RPM to 300,000 RPM with a resolution of 750 RPM. This frequency range and resolution are maintained in the new VMS high frequency spectra. The peaks monitored in the high frequency spectra are primarily those associated with gearmesh frequencies. H53 gearbox diagnoses cannot be adequately performed with a spectral resolution of 750 RPM, particularly when the signals being analyzed are acquired by sensors (velocimeters rather than accelerometers) that are not hard mounted to the gearbox housing.

Windowing. To minimize or reduce what is known as leakage errors in fast Fourier transform estimates of spectral levels, windowing is applied to the time domain data before transformation to the frequency domain. The 8500C unit uses a Kaiser-Bessel window which provides good frequency resolution at the cost of some amplitude accuracy. The amplitude error is repeatable, allowing for consistent data trending. The new VMS unit uses a flat-top window. This window was selected for use in the balancing algorithm, because it provides the best amplitude accuracy for spectral peaks. It was then carried over to the trend spectra calculations. The flat top window does not provide good frequency resolution, as spectral leakage smears closely spaced spectral peaks. The flat-top window currently used in the VMS does a good job of resolving the widely spaced peaks that are currently monitored in the low frequency (0-12,000 RPM) spectra. It does not provide good resolution of closely spaced or relatively low level peaks in vibration spectra; thus, it is not well suited to bearing and gearbox diagnoses. Compare, for example, the acceleration spectra measured with Hanning and Flat-top windows on one of the tail drive shaft frequencies in Figure A-2 of the Appendix.

Dynamic Range and Noise Floor. The 8500C has a 10 bit analog to digital converter. The unit has a 60 dB dynamic range with 4080 automatically set input gain settings that optimize the amplitude of the input signal.[4] The VMS unit has a 16 bit analog to digital converter with a theoretical maximum dynamic range of about 95 dB. Given the size of this dynamic range, it was decided that adjustable input gains were unnecessary. All of the velocimeter input channels have a fixed range of 33 ips (peak); the reasoning behind this particular ips value is unknown. Because levels below 0.01 ips are considered inconsequential for the spectral peaks currently monitored, all spectral values below 0.01 ips are set 0 in the VMS. For an input of less than 3 ips, this results in an effective dynamic range of less than 50 dB (and some very peculiar looking spectra, see Figure A-3 in the Appendix). The VPMO considers levels in excess of 3 to 5 ips to be extreme, so that the 33 ips range setting is too large. The advantage of setting spectral levels below 0.01 ips to zero is unclear; in any case, this feature makes the currently configured VMS ill suited to the investigation of mechanical conditions that generate relatively low level vibrations (e.g. bearing faults).

Number of Spectral Averages. Most real processes are subject to some random fluctuations, so that the level of a spectral peak in a single fast Fourier transform (FFT) spectrum can be expected to have a large variance. What is sought in a vibration trend spectrum is a measure of the *average* vibration as a function of frequency under fixed operating conditions. For measurements on fixed base machines in industry, Reference 6 recommends that estimates of vibration levels be based on 8 to 10 averages. This author was under the impression that a 8500C trend spectrum is a 4 spectrum average, but Table 1-4 of Reference 4 suggests that no averaging is performed in the unit's spectral mode. No averaging is performed in the calculation of a VMS trend spectrum. Particularly on an aging helicopter, the level of a peak in a single shot spectrum is likely to show wild variations that are not reflective of changes in mechanical condition.

3. QUALITY OF TREND DATA

To assess the reliability of currently trended levels as indicators of mechanical faults, the data in Vibralog (both 8500C and VMS data) were inspected and the following observations made. Short term variations in band levels are large relative to differences between advisory and investigation alarm levels. There are very few notes on maintenance actions that might explain the large variations in short term levels. In many cases, the "trends" indicated in the different flight regimes are inconsistent. See, for example, the data of Figure A-4 in the Appendix.

Based on the data in the database, it is difficult to see how any sort of trending is done; although, upper limits on levels are undoubtedly appropriate. This is, indeed, how the data is interpreted by the VPMO staff. If a monitored level exceeds an alarm level, it is flagged and brought to the attention of the maintenance crew. Nonetheless, the trend data currently acquired, in conjunction with existing alarm levels, do not appear satisfactory even for use in this limited sense.

Consider the outcome of a test at KAFB, wherein Ground regime spectra were acquired five times over a two day period. During this time period, the tail rotor *may have been* balanced, but the mechanical condition of the other components of the aircraft remained essentially unchanged. Table A-III summarizes the alarm states indicated by each of the five measurements for levels corresponding to the main rotor RPM (185 RPM), the main rotor blade passage frequency (1110 RPM) and the tail rotor drive shaft RPM (3011 RPM). Depending on which of the five measurements is used, monitored levels are either below the advisory alarm level, above the advisory level but below the investigation level, or in excess of the investigation level.

In short, trend levels acquired on a single aircraft under nominally identical conditions vary widely. It was noted by Jim Ramsey of the VPMO office that the variability seen in data acquired on the H53 aircraft is much larger than that seen on other aircraft. The question that must be addressed is this, "How much of this variability is due to the less than optimal signal processing and sensor locations, and how much is inherent to this aircraft?"

C. MONITOR DATA

When monitor data was first acquired by the VMS, an alarm trigger time of 10 minutes was used and large numbers of band alarms were generated during every flight. It was decided that the use of a 30 minute trigger time, rather than 10 minutes, would eliminate alarms generated by high level vibrations occurring during transitions between flight regimes. Now, a band must exceed its alarm level for 30 minutes before an alarm exceedance is logged and a spectrum is recorded. Then, only one single shot spectrum corresponding to the maximum (peak or rms) level measured during the entire time that the band remained in alarm (exclusive of the 30 minute trigger time) is recorded. A single spectrum corresponding to the maximum level measured at a single frequency during a possibly extended time period, no matter what the 'trigger time," is as likely as not to be statistically meaningless.

Documentation is lacking on the derivation of monitor alarm levels and there is, at best, only limited data available on what can be considered excessive in a single shot spectrum measured during an arbitrary flight regime. With a few exceptions, the levels of the monitor alarms appear to be based on the largest of the trend alarms from the Ground, Hover, and 120 knot Forward regimes. The rms and peak levels, and the RPM corresponding to the peak level are recorded for every 2 minutes. This data could be used by the VPMO to establish bounds on expected flight levels. At present, users of the VMS monitor mode are acquiring prodigious amounts of data of unknown value.

V. RECOMMENDATIONS

A. MANAGEMENT

1. Responsibility should be assigned for the cost effective growth of the MH53J vibration monitoring system, including
   - formulation of a clear and detailed definition of the short and long term program expectations;
   - provision of adequate technical support to the program (see paragraph B-2, below);
   - for any new development or system modification, provision to the contractor(s) of a clear and detailed definition of requirements, consistent with long term, cost effective program goals;
   - periodic evaluation of costs and benefits of all program elements;
   - publication and circulation of system success stories;
   - ensurance of the fullest possible participation and "buy in" by the maintenance staff at all of the MH53J bases;
   - establishment of a mechanism that insures relevant maintenance actions are recorded in the vibration database.
2. A document detailing the technical specifications of the VMS unit was not included in the purchase contract and is needed. Note that changes to the present default monitor settings or signal processing options (e.g., number of averages, type of window) must be done by the manufacturer.

B. TECHNICAL

1. Although the track and balance capabilities of the vibration monitoring system are not the subject of this report, it is clear that the top priority at this moment is the incorporation of software to support the optical blade tracker into the VMS units.
2. Continued system growth and improvement requires ongoing technical oversight by an appropriately trained engineer. Suggested background for such an individual is provided in the Appendix. Impartial expertise that remains abreast of current developments in a rapidly developing field is especially critical when open bids for research and development projects are difficult to issue and incorporation of these efforts into existing contracts is the only practical option.
3. The long term enhancements to the vibration monitoring program that were discussed in the introduction to this report rely on the ability of the system to provide data that consistently and reliably reflect the mechanical and operational conditions of the aircraft. It stands to reason that if such data cannot be obtained under a set of standard operating conditions (e.g., Ground, Hover, and 120 knot Forward Flight), it will not be possible to do so in flight under arbitrary and constantly varying flight conditions. Thus, it is imperative that improvements in the quality of trend data be pursued. One possibility is to designate one or more aircraft at Kirtland (perhaps, TH53 aircraft) as experimental testbeds. No changes in the vibration monitoring system would be made to the remainder of the fleet until the benefit of such changes is proven. With minimal accommodation on the part of

the crews and minimal impact on schedules, it should be possible to determine if the trend data can be improved on the test aircraft. The following steps are suggested.

- decide which mechanical components and conditions (e.g., tail rotor drive shaft imbalance and alignment) might be effectively monitored with a relocation of existing velocimeters (velocimeters used in the test and balance procedures or otherwise deemed critical would not be moved);
- relocate velocimeters to bearings or other load paths appropriate to the components and conditions to be monitored;
- using a spectral analyzer and one of the velocimeter power supply units (the units that provide the necessary charge amplification to the velocimeters and a BNC signal output connection), determine if increasing the number of spectral averages increases the consistency of data acquired, by one of the relocated sensors, under nominally identical operating conditions.

If these measures prove successful in improving the quality of the trend data, the VMS contractor will have to be tasked with modification of the unit software to accommodate multiple spectral averages.

4. Diagnoses of gearbox conditions can be investigated by the VPMO. These investigations cannot be made using the VMS as currently configured. Periodic measurements on the intermediate, tail rotor and main rotor gearboxes could be made with accelerometers and the DI PL302 analyzer and imported into Vibralog with the aid of an Entek driver recently ordered by the VPMO. The number of spectral lines used should be 400, compatible with the Vibralog database structure. Several baseband and zoom frequency ranges may be required. The diagnoses of faults in gearboxes can be subtle, even when there are only two gears. Thus, expectations for diagnosing faults from the complex spectra measured on the main gearbox should be modest. A literature search and review of related work, e.g. that in Ref. 8, should be performed.
5. Diagnoses of bearing faults can, also, be investigated by the VPMO. These investigations cannot be made using the VMS as currently configured. A study by the Navy concluded that faults in tail rotor drive shaft bearings could be detected using band limited, rms vibration levels.[9] A comparison of spectra measured on tail drive shaft bearings on a test bench at the Air Force Academy and those measured on the ground in an aircraft at Kirtland indicates substantial qualitative differences.[10] Any decision on the feasibility of monitoring the MH53J tail drive shaft bearings must be based on data measured on a loaded drive shaft in an aircraft.
6. The VPMO should be tasked with investigating improvements in drive shaft alignment methods. Although Sikorsky maintains that sections 1 through 6 of the tail rotor drive shaft are designed to accommodate considerable misalignment, the higher the level of misalignment, the higher the level of bearing wear and the greater the level of maintenance required. Section 7 of the tail rotor drive shaft, from the intermediate gearbox to the tail rotor gearbox, does not have any mechanism for alignment adjustment; the two ends of the shaft are set in fixed housings. If it becomes apparent, from reliable trending measurements on the test aircraft, that misalignment of this shaft is generating adversely high levels of vibration, engineering modifications might be considered.
7. Ultimately, alarms and trend patterns should be tailored to individual aircraft.

VI. REFERENCES

1. "Section V, Rotors, Titanium Spar Main Rotor Blades, and Aluminum Tail Rotor Blades," TO 1H-53(M)J-2-4.
2. "MH-53J Vibration Program Plan," informal report prepared by the Vibration Program Management Office, kirtland AFB, NM, for WR-ALC/LUJE.
3. "Engineering Specification Model 7310 Velometer," drawing no. 8291, Rev. A, Chadwick-Helmuth Co., Inc., El Monte, CA, 19 September 1984.
4. "8500C System Maintenance Manual," Manual No. 10902-2, Chadwick Helmuth, El Monte, CA, January 1993.
5. "Prime Item Development Specification for The Vibration Processor Unit, P/N 12000-1 and Display/Data Transfer Unit and Printer Kit, P/N 12852-1," Co. No. 13422, Chadwick Helmuth, El Monte, CA, 6 October 1995.
6. Wowk, V., Machinery Vibrations, Measurement and Analysis, McGraw-Hill Company, N. Y., N. Y., 1991
7. Taylor, J. I., The Vibration Analysis Handbook, A Practical Guide for Solving Rotating Machinery Problems, Vibration consultatnts, Inc., Tampa, FL, 1994.
8. Technology Showcase, Integrated Monitoring, Diagnostics and Failure Prevention, Proceedings of a Joint Conference, Machinery Failure Prevention, Mobile, AL, 23-26 April 1996, published by Joint Oil Analysis Program Technical Support Center, Univesrity of Wales, Swansea.
9. Hollins, M., Miller, D., Weil, M, and Green, B., "H-53E Tail Rotor Drive Shaft Bearing Analysis Program," Reprot No. RW-153R-91, Navail Air Warfare Center Aircraft Division, Patuxent River, MD, 3 June 1992.
10. Vibration spectra on a tail drive shaft bearing in an aircraft in the Ground trend regime were measured in July 1996 at KAFB. These were compared to spectra calculated from data recorded on a TEAC DAT recorded in 1995 during a recent study by Chadwick-Helmuth and the University of New Mexico for Lt. Bui of Phillips Laboratories. A draft of the final report on this study was read by the author while at KAFB in June 1996.

VII. APPENDIX

In house technical oversight should be provide by someone with a background in most if not all of the following areas:

Vibrations theory- natural frequencies and mode shapes, resonant amplification, damping, random vibrations, frequency descriptions, attenuation of traveling waves in complex structures, transfer functions.

Vibration measurements - sensors, spectral analyzers, dynamic range and frequency response, suitability of sensors and analysis for specific applications, power spectral densities, transfer functions, coherence.

Statistics of random processes - mean (DC) and rms values, random and bias errors in statistical estimates.

Digital signal processing - anti-alias filtering, sampling, quantization, windowing, spectra, Fourier transforms, frequency resolution, sinusoidal vs. narrowband vs. broadband.

Machine condition monitoring basics - optimal sensor locations, fault recognition patterns, database structures, data trending.

Table A-I. Current MH53J Velocimeters.

Location	Orientation	VMS Channel No.
Sta. 182	Vertical (Vert)	2
"	Lateral (Lat)	16
"	Fore / Aft (F / A)	27
Sta. 322	Vert	20
"	Lat	6
"	F / A	9
Sta. 522	Vert	13
"	Lat	28
"	F / A	21
Tail Rotor Gearbox (TRGB)	Vert	14
Main Rotor Gearbox (MRGB)	Lat	7
Nose	Vert	23

Table A-II. Trend Alarm Levels for Ground Regime (in ips-peak).

	Sta	RPM	Advisory			Investigation Required		
			Vert	Lat	F / A	Vert	Lat	F / A
1x MR	182	185	0.32	0.23	0.16	0.39	0.28	0.19
6 x MR		1110	0.07	0.26	0.06	0.09	0.34	0.08
MGB SG		4544	0.06	0.03	0.05	0.08	0.03	0.06
IDS / AGBDS		6023	0.10	0.03	0.06	0.13	0.05	0.08
1 x MR	322	185	0.34	0.20	0.21	0.42	0.24	0.26
6 x MR		1110	0.15	0.24	0.10	0.18	0.30	0.13
12 x MR		2220	0.06	0.06	0.04	0.08	0.08	0.05
IDS / AGBDS		6023	0.17	0.12	0.13	0.22	0.15	0.18
MGBSPGM		24600	0.06	0.07	0.14	0.08	0.08	0.18
MGBGM		82800	0.15	0.23	0.13	0.21	0.38	0.18
1 x TR	522	791	0.28	0.17	0.03	0.36	0.22	0.04
6 x MR		1110	0.27	0.20	0.07	0.35	0.24	0.09
#7 DS		2298	0.08	0.05	0.03	0.09	0.06	0.03
TRDS		3011	0.92	0.14	0.24	1.16	0.19	0.31
4 x TR		3163	0.40	0.09	0.07	0.55	0.12	0.08
1 x TR	TRGB	791	0.18	--	--	0.23	--	--
6 x MR		1110	0.12	--	--	0.16	--	--
#7 DS		2298	0.10	--	--	0.12	--	--
TRDS		3011	0.17	--	--	0.22	--	--
4 x TR		3163	0.24	--	--	0.29	--	--
TGBGM		48251	0.26	--	--	0.34	--	--
IGBGM		87324	0.22	--	--	0.28	--	--
1 x MR	Nose	185	0.40	--	--	0.47	--	--
6 x MR		1110	0.19	--	--	0.23	--	--
1 x MR	MRGB	185	--	0.29	--	--	0.34	--
6 x MR		1110	--	0.21	--	--	0.25	--
TRDS		3011	--	0.03	--	--	0.04	--
IDS / AGBDS		6023	--	0.12	--	--	0.16	--
OCDS		7070	--	0.03	--	--	0.03	--
MGBGM		82800	--	0.19	--	--	0.25	--

Table A-III. Alarm States Indicated by Five Trend Data Sets (Ground Regime) Acquired with the VMS on an H53 Aircraft under Nominally Identical Mechanical and Operating Conditions.

White indicates no alarm, gray indicates advisory alarm, black indicates investigation alarm.

Frequency in RPM	Sta	7/19/96 5:00:58 PM	7/19/96 5:06:56 PM	7/19/96 6:07:30 PM	7/20/96 3:36:26 PM	7/20/96 4:05:08 PM
185	Sta 182 Vert	■	▒	□	□	■
	Sta 182 Lat	□	□	□	□	□
	Sta 322 Vert	□	▒	□	□	■
	Sta 322 Lat	▒	□	□	▒	■
	Nose	□	□	□	▒	▒
1110	Sta 182 Vert	■	■	▒	■	■
	Sta 182 Lat	□	□	□	▒	▒
	Sta 322 Vert	■	□	■	■	□
	Sta 322 Lat	■	■	■	■	□
	Sta 522 Vert	■	□	■	■	□
	TRGB	■	■	■	□	□
	Nose	□	□	■	■	■
3011	Sta 522 Vert	□	□	□	□	□
	Sta 522 F/A	▒	▒	▒	□	▒
	TRGB	□	□	□	■	□

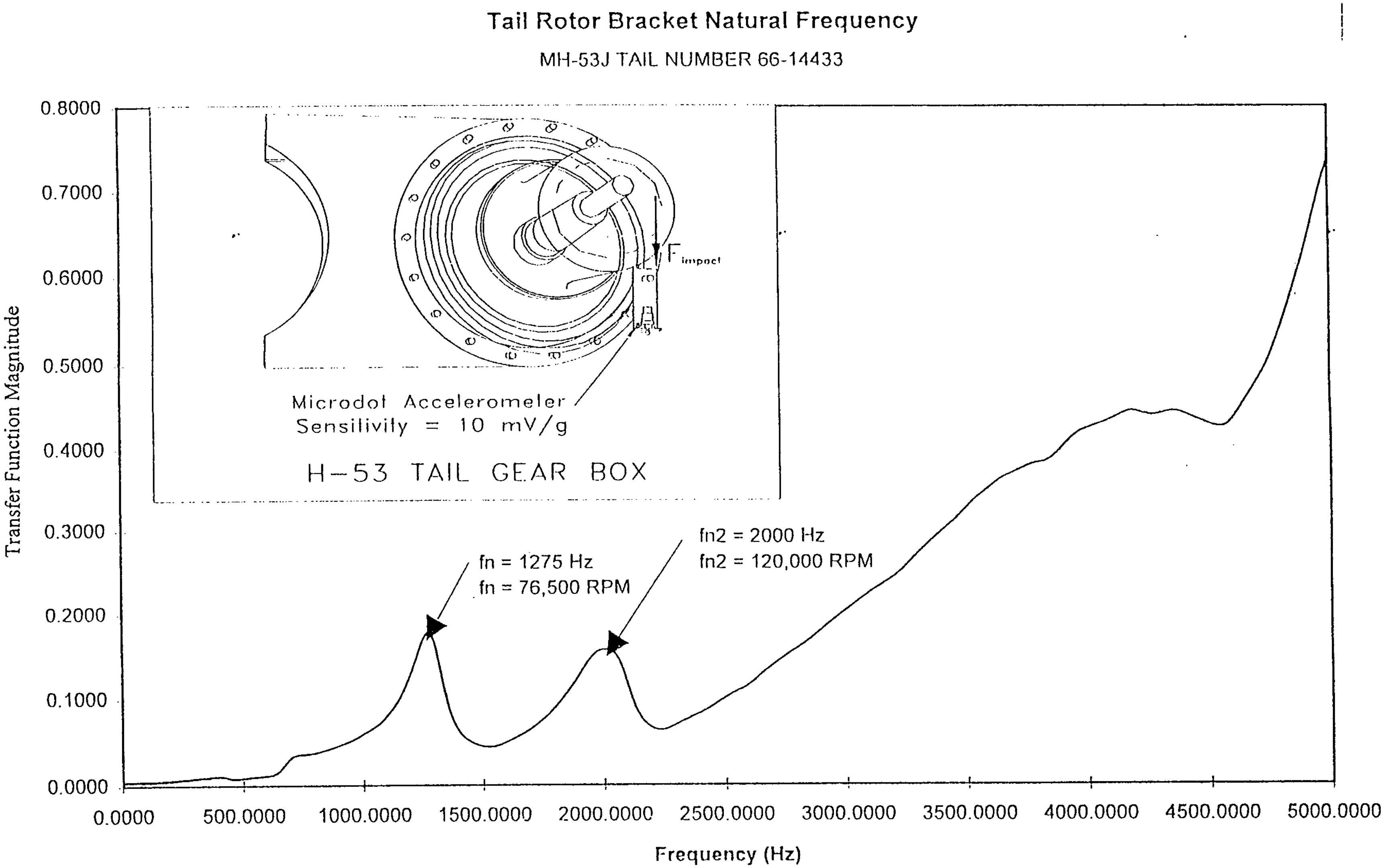


Figure A-1. Transfer Function Between Acceleration and Force on Tail Rotor Gear Box Bracket.

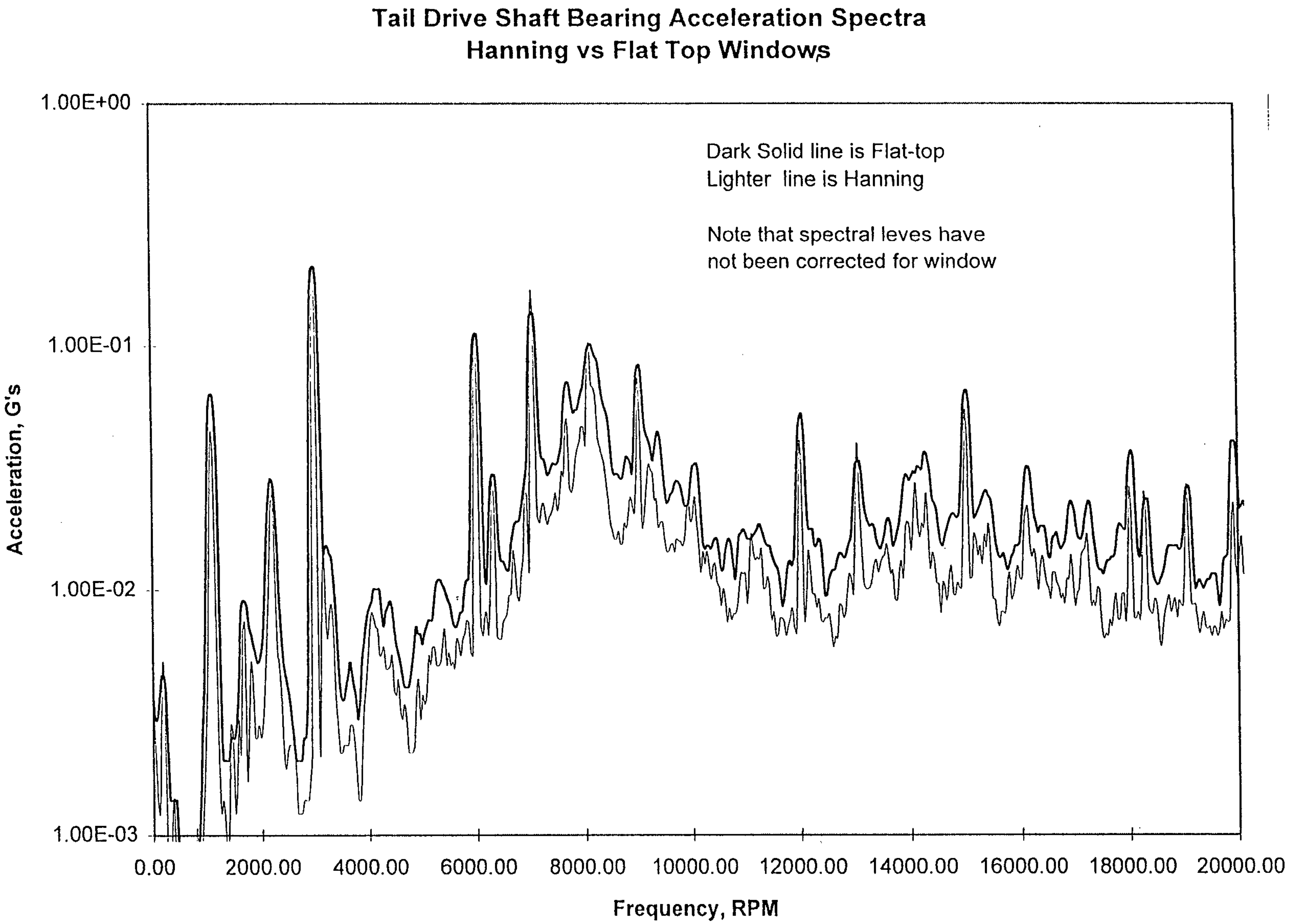


Figure A-2. Comparison of Spectra Calculated with Hanning and Flat Top Windows. 20 Spectral Averages.

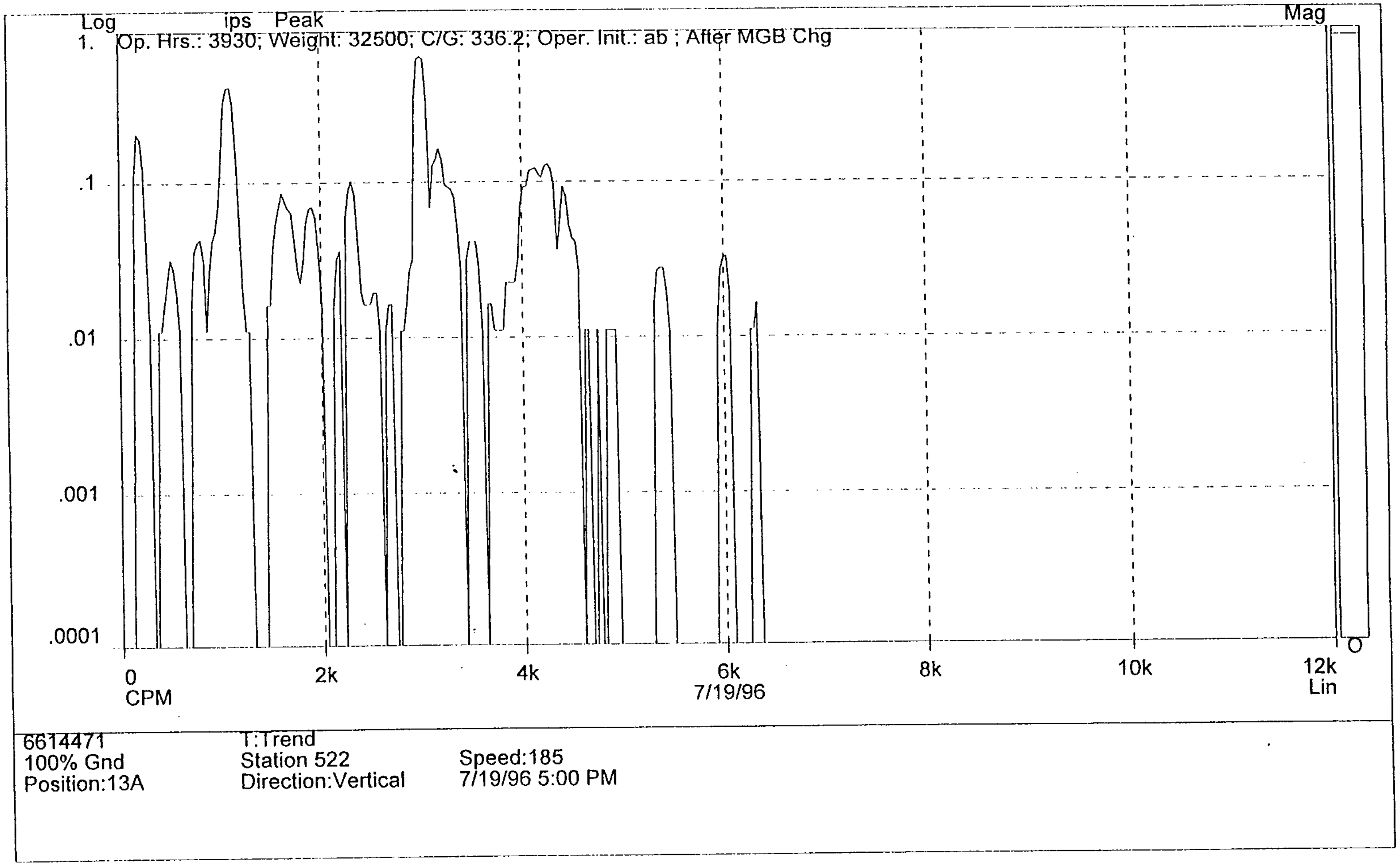


Figure A-3. Low Frequency Spectrum Acquired with VMS.

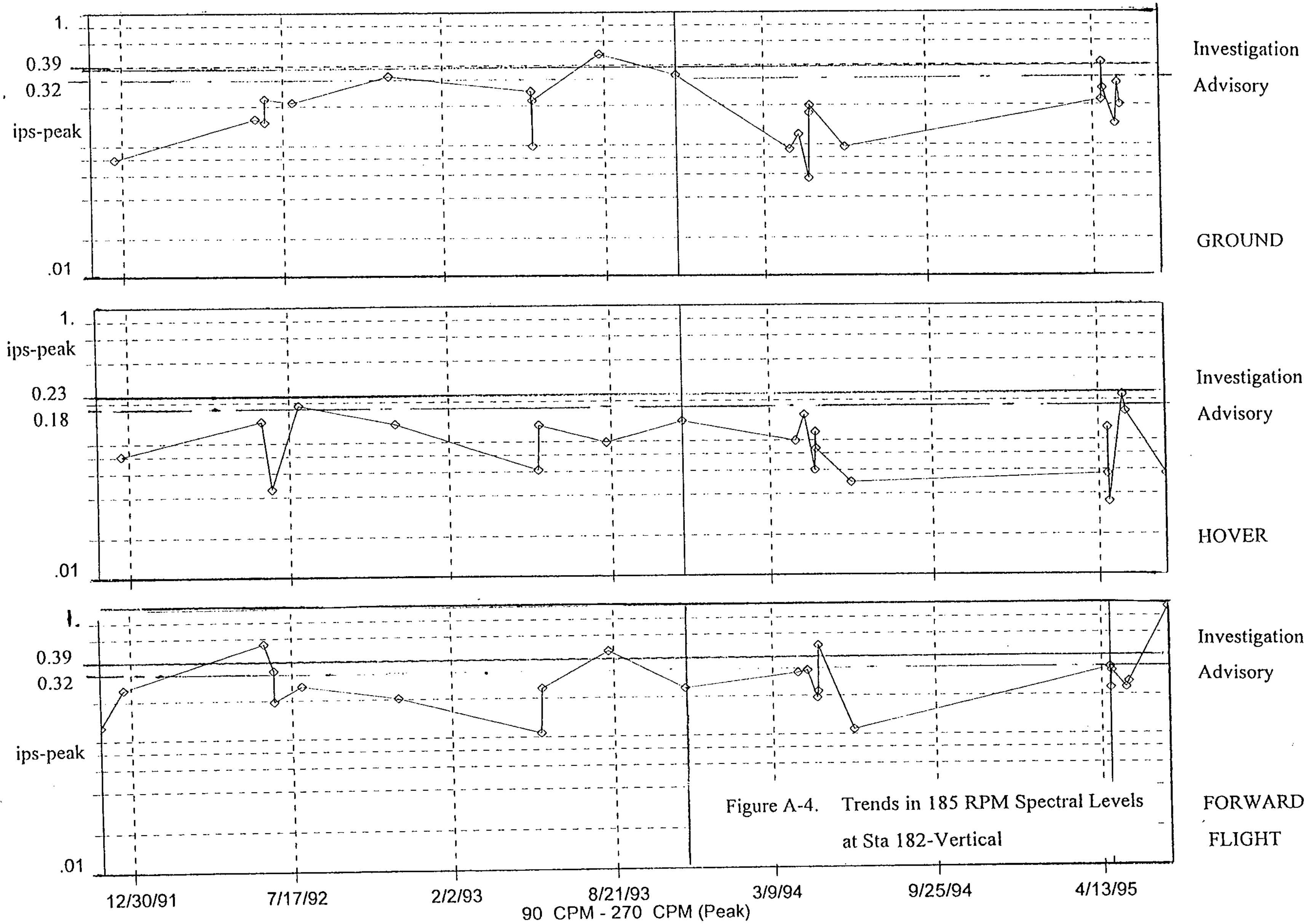


Figure A-4. Trends in 185 RPM Spectral Levels at Sta 182-Vertical