

1999 RESEARCH REPORTS

NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

JOHN F. KENNEDY SPACE CENTER
AND
UNIVERSITY OF CENTRAL FLORIDA



1999 RESEARCH REPORTS

NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

JOHN F. KENNEDY SPACE CENTER

UNIVERSITY OF CENTRAL FLORIDA

EDITORS:

Dr. E. Ramon Hosler, University Program Director
Department of Mechanical, Materials and Aerospace Engineering
College of Engineering
University of Central Florida

Mr. Gregg Buckingham, NASA/KSC Program Director
University Programs Office
John F. Kennedy Space Center

NASA Grant #NAG10-260

Contractor Report No. CR-2000-208586

November 2000

PREFACE

This document is a collection of technical reports on research conducted by the participants in the 1999 NASA/ASEE Summer Faculty Fellowship Program at the John F. Kennedy Space Center (KSC). This was the fifteenth year that a NASA/ASEE program has been conducted at KSC. The 1999 program was administered by the University of Central Florida (UCF) in cooperation with KSC. The program was operated under the auspices of the American Society for Engineering Education (ASEE) and the Education Division, NASA Headquarters, Washington, D.C. The KSC program was one of nine such Aeronautics and Space Research Programs funded by NASA Headquarters in 1999.

The basic common objectives of the NASA/ASEE Summer Faculty Fellowship Program are:

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

The KSC Faculty Fellows spent ten weeks (May 25 through July 30, 1999) working with NASA scientists and engineers on research of mutual interest to the university faculty member and the NASA colleague. The editors of this document were responsible for selecting appropriately qualified faculty to address some of the many research areas of current interest to NASA/KSC. A separate document reports on the administrative aspects of the 1999 program. The NASA/ASEE program is intended to be a two-year program to allow in-depth research by the university faculty member. In many cases a faculty member has developed a close working relationship with a particular NASA group that had provided funding beyond the two-year limit.

TABLE OF CONTENTS

	<u>PAGE</u>
1. ALLEN, John R. <i>Determination of the Components of a Simulated Atmosphere of Mars</i>	1
2. ANDRAWIS, Alfred S. <i>Effects of Restricted Launch Conditions for the Enhancement of Bandwidth-Distance Product of Multimode Fiber Links</i>	11
3. BAREISS, Catherine C. <i>STEP: What Is It and Should It Be Used for KSC's ISE/CEE Project in the Near Future?</i>	23
4. BLUM, Linda K. <i>Measures of Water Quality in Merritt Island National Wildlife Refuge Impoundments and Adjacent Indian River Lagoon</i>	33
5. CALLE, Luz M. <i>Characterization of Molybdate Conversion Coatings for Aluminum by Electrochemical Impedance Spectroscopy</i>	43
6. CHRISTENSEN, Kenneth J. <i>Performance Evaluation of the NASA/KSC Transmission System</i>	53
7. ERICKSON, Lance K. <i>Evaluating Education and Science in the KSC Visitor Complex Exhibits</i>	67
8. HARVEY, Bruce A. <i>Full Duplex, Spread Spectrum Radio System</i>	73
9. HOUSHANGI, Nasser <i>Mars Umbilical Technology Demonstrator</i>	85
10. KOZEL, David <i>Noise Reduction Using Frequency Sub-Band Based Adaptive Spectral Subtraction</i>	97
11. LAVELLE, Jerome P. <i>Strategic Project Management at the NASA Kennedy Space Center</i>	107

12.	MILLS, Aaron L. <i>The Effect of Dilution on the Structure of Microbial Communities</i>	119
13.	MOYNIHAN, Gary P. <i>Integrated Workforce Modeling System</i>	129
14.	OJHA, Anand K. <i>Performance Monitoring of Distributed Data Processing Systems</i>	139
15.	POZO DE FERNANDEZ, Maria E. <i>Control for NOx Emissions from Combustion Sources</i>	149
16.	RAWAT, Banmali S. <i>Optimization of Connector Position Offset for Bandwidth Enhancement of a Multimode Optical Fiber Link</i>	159
17.	RUIZ-TORRES, Alex J. <i>Operations Assessment of Launch Vehicle Architectures Using Activity Based Cost Models</i>	169
18.	SHU, Wei W. <i>Performance Evaluation of Reliable Multicast Protocol for Checkout and Launch Control Systems</i>	179
19.	STANSIFER, Ryan D. <i>A Visual Editor in Java for Jview</i>	189
20.	WHITLOW, Jonathan E. <i>Modeling and Analysis of the Reverse Water Gas Shift Process for In-Situ Propellant Production</i>	197
21.	COREY, Kenneth A. <i>Testing Plant Responses to Rarified Atmospheres for Inflatable Greenhouses</i>	207

**1999 NASA/ASEE SUMMER FACULTY FELLOWSHIP
PROGRAM**

JOHN F. KENNEDY SPACE CENTER

**Determination of the Components of a Simulated
Atmosphere of Mars**

**Dr. John Allen
Instructor of Chemistry
Dept. of Chemistry and Physics
Southeastern Louisiana University
Dr. Cole Bryan**

ABSTRACT

The purpose of this project was to evaluate a possible detector for a Mars chamber. Due to theoretical considerations a Mass Spectrometer unit was decided to be the best choice. The detector had to be design for the new role and characterized. To do this, the original mass spectrometer of a bell jar vacuum system was used. The original schematic had to be abandon and a new design was formed. Gas leaks and other problems were addressed and solved. Due to the difficulties in troubleshooting this particular system several recommendations are put forth. The main recommendation is to replace the old mass spectrometer unit with a new instrument.

1. INTRODUCTION

Landing on the moon, one of Man's most celebrated achievements, was a culmination of a steady wave of missions starting with the Mercury mission to finally the Apollo mission. Before someone could be safely set foot on the moon, the environment of the moon and space and effect of this environment on the instruments and design had to be known. This was the reasons for Mercury and Gemini missions and the unmanned satellites.

Now, a new goal is dawning, to put the first man or woman on Mars, a more changeling achievement than any taken before. One that promises to bring technological advances in communications, computers, life sciences, and material sciences in addition to gaining a broader understanding of the solar system. The trip will last years instead of days along with living on a hostile planet. Also, the astronauts will need to make use of the resources of the planet itself to survive and to return home. The spacecraft will not be able to carry all the food, oxygen, and fuel necessary for the complete mission. (1) (2) This challenge is being attacked now, and is in the "Mercury Mission" stage of development.

The Mars atmosphere is very different from the atmosphere here on Earth. The atmospheric pressure on the surface of Mars is less than 1% of Earth's atmosphere. The measured composition of the Martian atmosphere is given in Table 1. Carbon dioxide makes up the bulk of the atmosphere and there is very little oxygen. (3) Therefore it is expected that materials will react and wear differently in the Mars atmosphere.

Table 1: Composition of the Martian Atmosphere

Constituent	Composition
carbon dioxide	95.3 %
nitrogen (N ₂)	2.7 %
argon	1.6 %
oxygen (O ₂)	0.13 %
water vapor	0.03 % (variable)

To overcome the before mentioned problems, experiments using the Mars Chamber are necessary. A Mars Chamber is actually a sealed vacuum chamber pumped to a low pressure (millitorr range) and backfilled with the appropriate gas mixture to about 7 torr. The chamber comes in a variety of sizes and styles from glass bell jars to large heavy steel chambers depending on the applications. Mars chambers can address such problems as the production of plants for food at low pressures to the wear and tear of materials.

2. DETECTOR EVALUATION

How is the composition of the gases in the Mars chamber known? Knowing the composition of the gas mixture in the inlet valve is not enough. The composition slowly changes with time due to leakage. With all vacuum chambers there are leaks, but the leakage should only give small error. Or a serious leak may occur greatly changing the composition. Finally, the gas mixture composition may not be accurately known. Detectors allow the monitoring of the Mars chamber for leaks and can accurately analyze the gas mixture.

One focus of this project was to determine which type of detector would best suited to the task of monitoring the atmosphere in a Mars chamber. The theoretical advantages and limitations will be discussed for infrared spectroscopy, Raman spectroscopy, and mass spectroscopy and from these discussions the technique to be further evaluated will be chosen.

Infrared spectroscopy is useful for determining such atmospheric components as carbon dioxide and water vapor.(4,5) Both have strong absorption peaks in the infrared region. However, other atmospheric gases such as oxygen and nitrogen are silent in the infrared region due to symmetry. Determination of noble gases is impossible with infrared as well.(6) Therefore since nitrogen and oxygen are indicators of leaks, infrared spectroscopy was not considered for general use in the Mars Chamber. However, if only the carbon dioxide levels are needed to be known in a Mars Chamber this method can be used since the carbon dioxide levels would be higher in the chamber than outside in Earth's atmosphere.

Raman spectroscopy occurs from an inelastic scattering of light. The amount of light scattered gives a much smaller signal than absorption techniques and therefore uses a laser to increase to initial amount of light used. This technique is actually a compliment to infrared spectroscopy. Raman spectroscopy gives good signals for oxygen and nitrogen, but not as intense signals for water and carbon dioxide.(6) However, since the signals are low to begin with Raman is not a choice technique for the analysis of gases at low pressures.

Mass spectroscopy looks at the mass to charge ratio by ionizing the molecule or atom than introducing the ion to either an electric field or a magnetic field. If the ion is of the right mass to charge ratio the ion will safely navigate the field to the detector. The field can be altered to produce to scan the masses. Mass Spectrometers used in vacuum systems are called "residual gas analyzers" or RGAs and have been used since 1960s. The RGAs can monitor any gas, molecule or atom, and therefore do not have the restrictions the former techniques have. One major drawback for the RGAs is they need a very lower pressure to operate than the pressure of the Mars Chamber.(7) This drawback can be overcome and this technique was one to be further investigated.

3. DESIGN

This project is to incorporate a mass spectrometer as an RGA system to a bell jar vacuum system. The mass spectrometer was originally attached to the bell jar, but currently was used independently with its own vacuum system. Another group had used the mass spectrometer and had had some results with the instrument.

The original schematic is shown had the mass spectrometer separated from the turbopump by two solenoid valves. This system also had an inlet for external gases in order to calibrate the instrument. Since the bell jar was designed to be used at low pressures, the RGA unit could provide good detection of leaks and problems with the vacuum system or the turbo pump.

However, now the needs were different. Since the bell jar would be filled with a Martian atmosphere at about 7 mbars only the vane pump would be needed. And the mass spectrometer would not be monitoring the turbo pump but instead would be monitoring the bell jar. These requirements called for a new design allowing the monitoring of the bell jar and calibration of the mass spectrometer. The new schematic is shown in Figure 1.

The mass spectrometer is separated from the turbo pump by only one solenoid valve (V5). A turbo pump brings the pressure down to the necessary level (10^{-5} to 10^{-6} torr). The turbo pump and the bell jar are well separated, therefore the mass spectrometer either could be close to the turbo pump or the bell jar containing the sample. The pathlength for the sample was minimized in order to obtain faster response time and improve accuracy. Therefore, a vacuum hose runs from Valve 5 to the mass spectrometer.

The mass spectrometer is attached to the bell jar through two manual valves (VA and VB). The Valve A serves to separate the bell jar from the sample line. The Valve B is a precise microvalve providing the orifice to step down the pressure for the mass spectrometer. This is the only stepping down that occurs because the pressure inside the bell jar is less than 7 torr much less than atmospheric pressure.

The RGA system has two gauges. One for the bell jar attached to the cross tube above Valve A (PA) and the second after the mass spectrometer (PB). The first gauge (PA) is for a capacitance gauge for the relatively higher pressure in the bell jar. This gauge ranges down to 1 mtorr. The other gauge is an ion gauge meant to monitor the pressure of the mass spectrometer. This gauge is activated when the pressure falls below 10^{-4} torr in the initial pump down of the system. Since this is an ion gauge it should not be left on when the mass spectrometer is on and the gauge is turned off before engaging the mass spectrometer.

There have been difficulties in implementing the inlet system for calibration of the mass spectrometer. The inlet is designed to be place on the cross tube. The inlet

primarily consists of a needle valve in order to allow a small stream of gas into the cross tube. Current inlets have large leaks. This is one of the future projects. The procedure for calibration involves using a calibration gas with similar composition to the Martian atmosphere. The gas is let in the cross tube with the valves to the bell jar and the microvalve in the off position. Only 7 torr of the gas is let into the cross tube as monitored by the pressure gauge PA. The microvalve is opened to allow a small orifice for molecular flow into the hose to the mass spectrometer sample port. The mass spectrometer than can be calibrated and the accuracy of the instrument can be determined.

4. RESULTS

4.1 Leak Detection

After assembling the vacuum system and the mass spectrometer, leak detection was performed on sections of the RGA. Initial experiments indicated large leaks. Several modifications enacted were to replace the inlet and vent ports on the cross tube with seal flanges. This greatly improved the performance of the RGA system. Also an elbow joint between the large vacuum hose and valve 5 was replaced with a straight tube to relieve strain on the vacuum hose. However, the ion gauge nor the mass spectrometer was operable. When the ion gauge was replaced with another ion gauge the pressure was shown to decreased to 2×10^{-5} torr (figure 3). For these experiments both pressure gauges (PA and PB) were inline between the mass spectrometer and the large vacuum hose.

The leakage rate is the rate at which air from the room leaks into the vacuum system. This rate is calculated by pumping the system down to a desired pressure (in this case about 7 torr), shutting off the system under investigation, and recording the pressure versus time. The pressure increase versus time (Leakage Rate) for the Bell Jar and the cross tube (P1), the Bell Jar only (P2), and cross tube only (P3) is plotted in Figure 4. Both the Bell Jar by itself and cross tube by itself increases approximately 2 torr per hour. Together, the rate is larger 6 torr per hour. Each trial was performed once which could account for the much higher rate for both combined. The rates are higher than expected and methods for decreasing the rate need to be investigated.

4.2 Mass spectrometer

Still the mass spectrometer was inoperable. Since the pressure was in range for the operation of the mass spectrometer, the mass spectrometer itself must be at fault. One reason for the mass spectrometer not to come online is a blown filament. The filament is the source for the electron beam responsible for the production of ions. The mass spectrometer was disassembled and the filaments replaced with a new set of filaments. Both sets of filaments were tested and found to have not blown. Therefore the hypothesis about the filament was wrong.

Another source of error can come from the computer used to control the mass spectrometer. There are several voltages and currents controlled by the computer such as the current amount flowing through the filament. However, this current was found to be correct. Other voltages and currents also were in tolerance range. Afterwards a fuse was found to be blown. The 0.1 amp, 250 volt fuse was replaced. The mass spectrometer came online after the pumps again lowered the pressure to 4×10^{-6} torr. However, the mass spectrometer scan did not show a peak for water vapor or nitrogen, only background noise. This experiment was repeated later with the same results. This may be due to the system having too much water vapor and other volatile chemicals. If these results persist after bake out, then there are more serious problems with the computer or the mass spectrometer.

5. FUTURE EXPERIMENTS AND RECOMMENDATIONS

This project is far from being finished. While the mass spectrometer is still the best type of detector, a new instrument is needed. A portable mass spectrometer, which has been ordered, can be incorporated replacing the old spectrometer. This should correct many of the problems this project has seen. The portable mass spectrometer has ports for calibration, eliminating the need for an inlet port in the cross tube. Some modifications of the vacuum line will be needed to simplify the design.

Once the modifications are accomplished, the system can be bakeout eliminating any contamination in the vacuum lines, pumps, or the mass spectrometer. The result of the bakeout is a lower background noise in the signal. After the bakeout, the system will be calibrated with a gas mixture similar to the composition of the simulated Martian atmosphere. Particular attention will be paid to monitoring of carbon dioxide, nitrogen, and oxygen depending on needs of the experiment. Finally, experiments will use the bell jar system with the RGA detector monitoring the atmosphere inside the chamber for leaks and changes in the composition of the gases in the chamber.

REFERENCES

- (1) Raeburn, Raul; Uncovering the Secrets of the Red Planet Mars; National Geographic Society; Washington, DC; 1998.
- (2) Zubrin, Robert; Wagner, Richard; The Case for Mars; Touchstone; New York, NY; 1996.
- (3) Encyclopedia of Planetary Sciences; ed. Shirely, James; Chapman & Hall; London, UK; 1997.
- (4) Nondestructive Testing Handbook; 3rd ed.; ed. Moore, Patrick; American Society for Nondestructive Testing; 1998.

- (5) Sensors, A Comprehensive Survey; Vol. 6 Optical Sensors; ed. Gopel, W.; Hesse, J; Zemel, J. N.; VCH; Weinheim, FRG, 1992.
- (6) Willard, Hobart; Merritt, Lynne, Jr.; Dean, John; Settle, Frank; Instrumental Methods of Analysis; 7th ed.; Wadsworth Publishing Company, Belmont, CA; 1988.
- (7) O'Hanlon, John; A User's Guide to Vacuum Technology, 2nd ed.; Wiley & Sons; New York, NY, 1989.

Figure 1. New Schematic for Bell Jar System

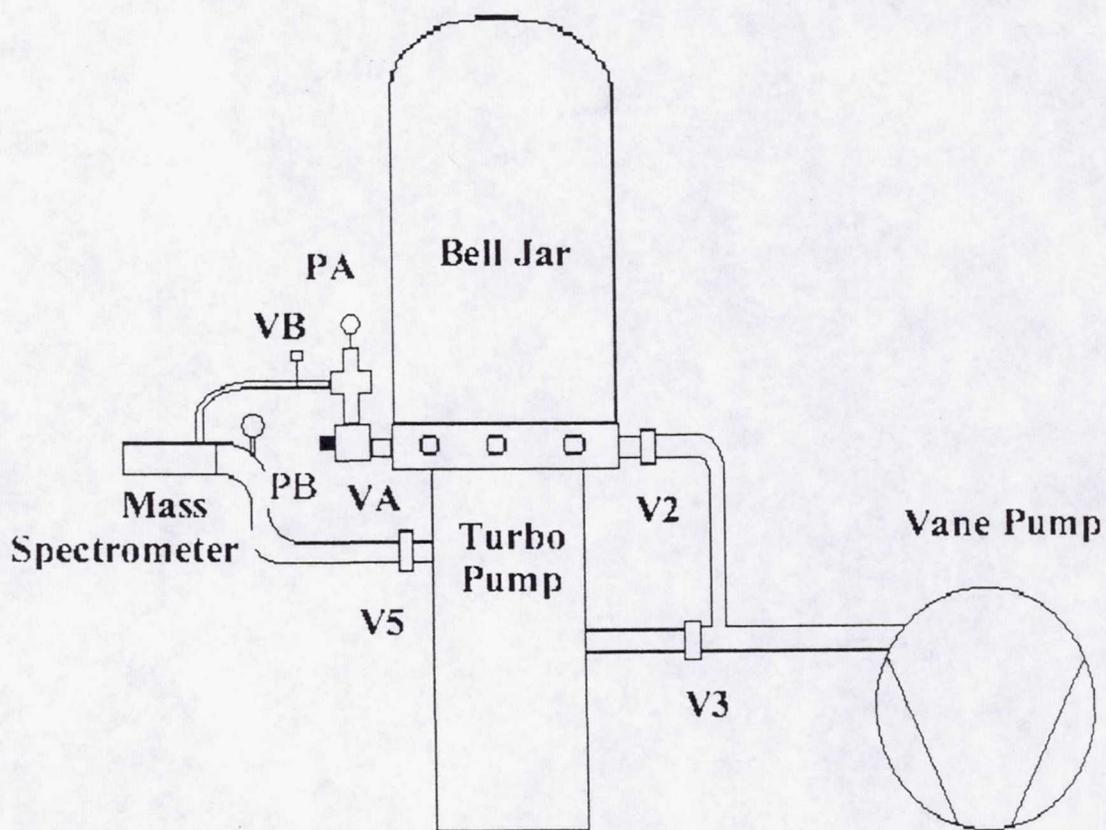


Figure 2. Vacuum Test for Mass Spectrometer and Vacuum Line

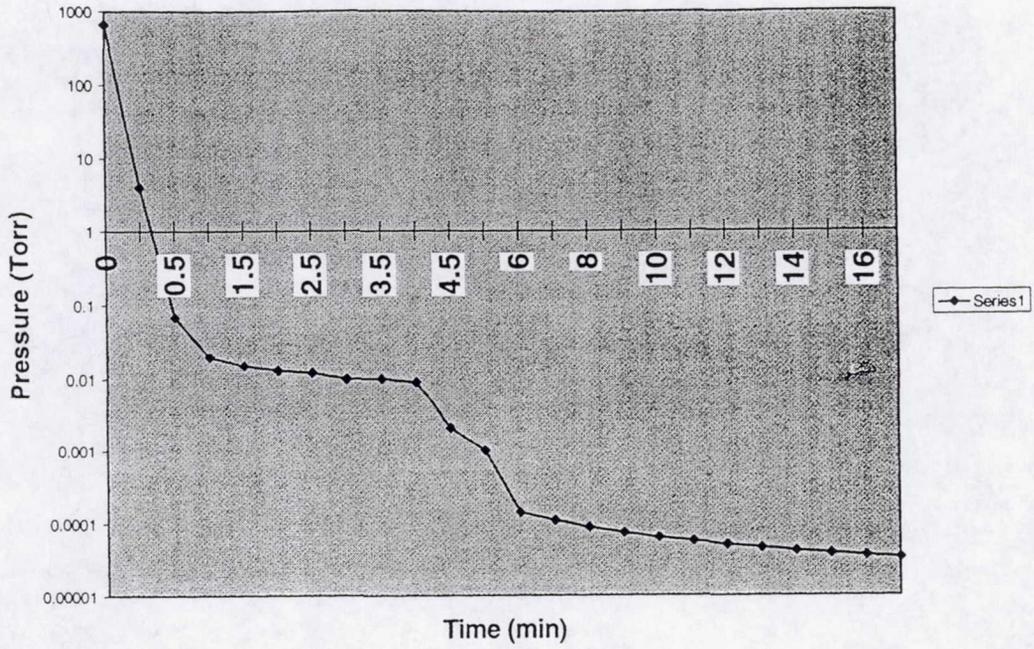
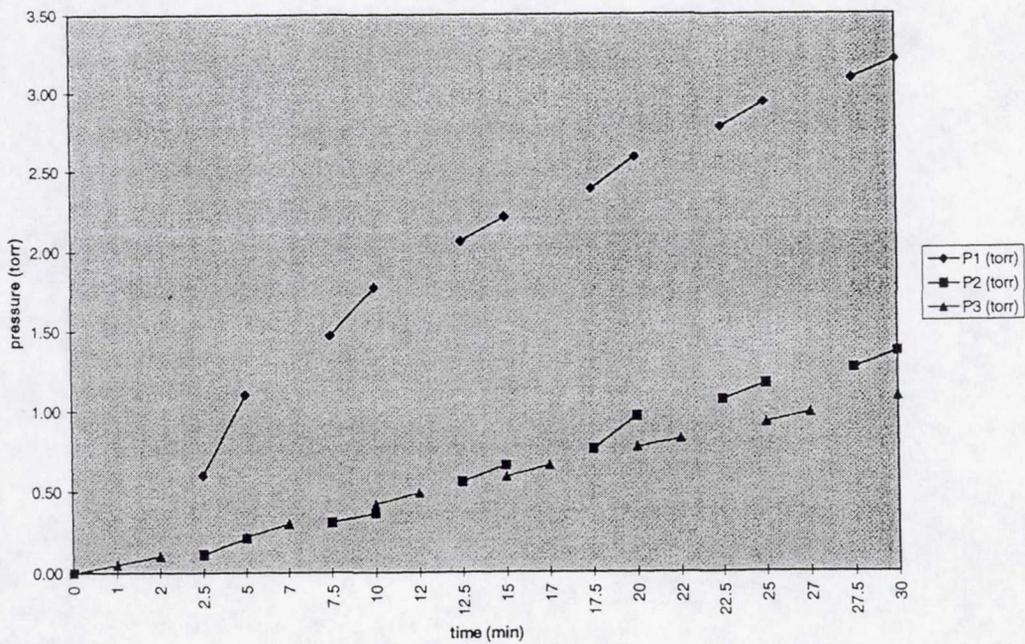


Figure 3. Leakage Rates



P1 is the leakage rate for the Bell Jar and Cross Tube
 P2 is the leakage rate for the Bell Jar
 P3 is the leakage rate for the Cross Tube

ASEE/NASA Summer Faculty Fellowship Program

*JOHN F. KENNEDY SPACE CENTER
&
UNIVERSITY OF CENTRAL FLORIDA*

**EFFECTS OF RESTRICTED LAUNCH CONDITIONS FOR THE
ENHANCEMENT OF BANDWIDTH-DISTANCE PRODUCT OF
MULTIMODE FIBER LINKS**

Alfred S. Andrawis

Associate Professor
Electrical Engineering Department
South Dakota State University
Brookings, SD 57007
E-mail: andrawia@mg.sdstate.edu

KSC Colleague

Po Tien Huang

MM-G1-C
Kennedy Space Center, Florida 32899
E-mail: PoTien.Huang-1@ksc.nasa.gov

ABSTRACT

Several techniques had been proposed to enhance multimode fiber bandwidth-distance product. Single mode-to-multimode offset launch condition technique had been experimented at Kennedy Space Center. Significant enhancement in multimode fiber link bandwidth is achieved using this technique. It is found that close to three-fold bandwidth enhancement can be achieved compared to standard zero offset launch technique. Moreover, significant reduction in modal noise has been observed as a function of offset launch displacement. However, significant reduction in the overall signal-to-noise ratio is also observed due to signal attenuation due to mode radiation from fiber core to its cladding.

EFFECTS OF RESTRICTED LAUNCH CONDITIONS FOR THE ENHANCEMENT OF BANDWIDTH-DISTANCE PRODUCT OF MULTIMODE FIBER LINKS

Alfred S. Andrawis

1. INTRODUCTION

The introduction of digital television and the increase popularity for Internet services are continually increasing the demand for higher data rate transmission on local area networks (LANs). LAN standards have been recently upgraded to rates up to 622 Mb/s and the standardization of a gigabit per second is currently under consideration. Such high-speed networks are expected to be essential for the Kennedy Space Center (KSC) backbone networks. Since the dominant fiber base currently in KSC is multimode fiber (MMF), its modal bandwidth imposes an upper limit on the achievable transmission speed and link distance. For such fiber, the maximum bandwidth distance product is limited to 500 MHz.km for over-filled-launch (OFL) conditions. Transmission rates greater than 1 Gb/s over distances beyond 1 km is not feasible on the basis of this specification. Although alternatives, such as using single mode fiber (SMF) links exist, it would be more economically feasible to develop high speed links using the installed MMF LAN.

There have been several attempts to overcome the OFL bandwidth limit by selective excitation of limited number of modes extending the bandwidth distance product [1-4].

This project is to explore the possibility of selective excitation of limited number of modes using a technique simpler than the ones reported in literature. This technique consists of SMF-to-MMF offset launch conditions such that only few modes are selectively excited resulting in extending the bandwidth distance product. Single mode fiber patch cord (10 meter long) terminated with a standard FC connector is coupled to a standard ST connector of the multimode fiber link. The offset is in the connector-to-connector coupling. This technique is expected to be more rugged than reported attempts and hence it may be more suitable for KSC applications.

2. PARAMETERS UNDER INVESTIGATION

The offset launch relies on exciting only a subset of all propagating modes of the MMF at the launch, so that the pulse broadening due to modal dispersion is lower and hence the bandwidth is larger. However, offset connector coupling loss increases due to the change in refractive index profile of the graded index fiber core. Furthermore, offset launch excites higher order modes. Higher order modes radiate off fiber adding to the nominal attenuation of the fiber. This effect will decrease signal's power at the receiving end. Another side effect of selective mode excitation is reducing modal noise [5].

The combination of reduction of, signal's power and modal noise results in changing the over all signal-to-noise ratio (S/N) affecting the quality of service (QoS) of the LAN. This change could be reducing or improving the overall S/N depending on other noise sources in the link.

Detailed performance studies for overall noise floor, modal noise, signal-to-noise ratio, connector and fiber attenuation, and bandwidth all as function of offset displacement are

investigated. Optimum connector offset position can then be determined for a desired application according to signal-to-noise or/and fiber attenuation.

3. EXPERIMENTAL SET-UPS

Offset launch condition is accomplished in the laboratory using connectors lateral misalignment. Single mode fiber terminated with a standard FC connector is coupled to a standard ST connector of the multimode fiber link under consideration. The offset is in the connector-to-connector coupling. Three experimental set-ups are shown in Figure 1. Short description for each of the three experiments is given in the following sections.

3.1. Noise Measurement Set-up

This experiment, as shown in Figure 1 (a), is to measure overall system noise and to calculate modal noise. The lightwave component analyzer is used to generate 100 MHz sinusoidal signal. The Overall signal-to-noise ratio (S/N) is initially measured for 8 spools of Corning graded index fiber. Each spool is ~1 km long (LDF 50/125, BW 3-dB= ~1214 MHz-km, and NA= ~ 0.2). The measured S/N includes noise due to laser, modal, shot, thermal, and amplifier. The measurement is repeated for five SMF-to-MMF connector lateral offsets: 0 μm , 5 μm , 10 μm , 15 μm , 20 μm , and 25 μm . The experiment is then repeated without the 8 spools of MMF and the offset connector. An optical attenuator (Anritsu MN934F) is then inserted instead, using two SMF patch cords. The attenuator is used to attain the same received optical powers as in the previous parts. The measured S/N in this case does not include modal noise however, it includes all other noise sources (laser + receiver). The S/N due only to modal noise is then calculated by substituting in the equation:

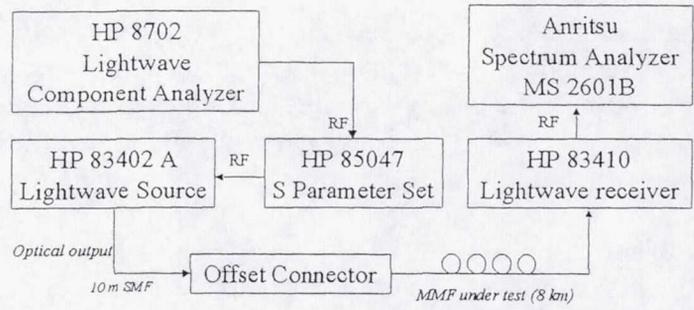
$$\left(\frac{S}{N}\right)_{\text{Modal}}^{-1} = \left(\frac{S}{N}\right)_{\text{Total}}^{-1} - \left(\frac{S}{N}\right)_{\text{Laser+Receiver}}^{-1}$$

3.2. Attenuation Measurement Set-up

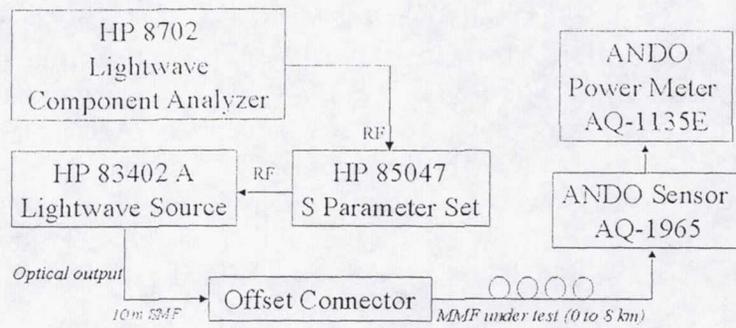
This experiment, as shown in Figure 1 (b), is to measure offset connector attenuation as well as the attenuation of the fiber as function of the length. Similar to the previous experiment, the lightwave component analyzer is used to generate 100 MHz sinusoidal signal. The signal emitted by the lightwave source (HP 83402 A) has a wavelength of 1.3 μm and optical power equal to 0.0 dBm. The optical power is then measured immediately after the offset connector and after each one of the eight spools of fiber.

3.3. Bandwidth Measurement Set-up

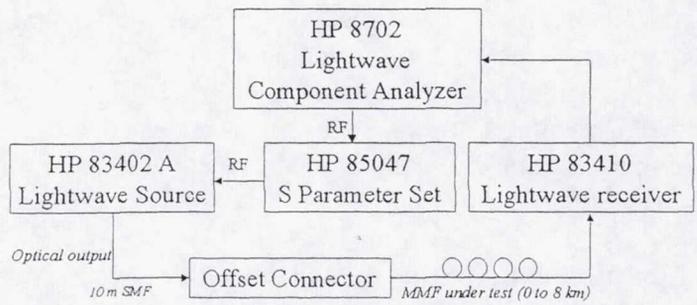
This experiment, as shown in Figure 1 (c), is to measure the bandwidth of the MMF after each connector. The HP S-Parameter set sweeps the spectrum starting at 300 kHz and ending at 3 GHz. The lightwave component analyzer is then used to measure the electrical 3-dB bandwidth.



(a) Noise Measurement Set-up

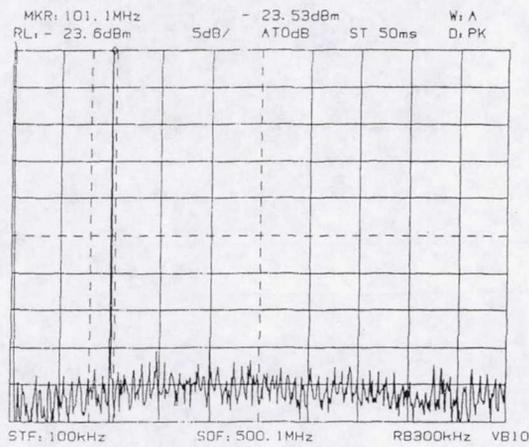


(b) Attenuation Measurement Set-up

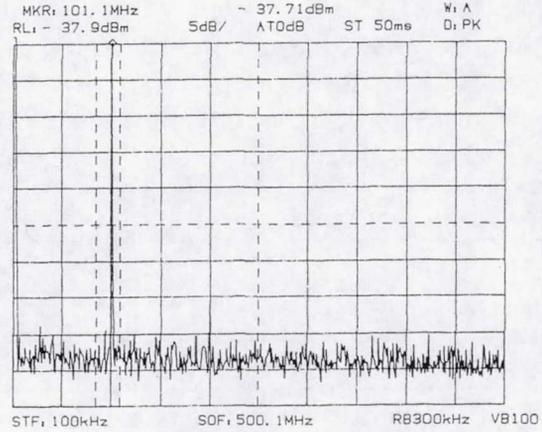


(c) Bandwidth Measurement Set-up

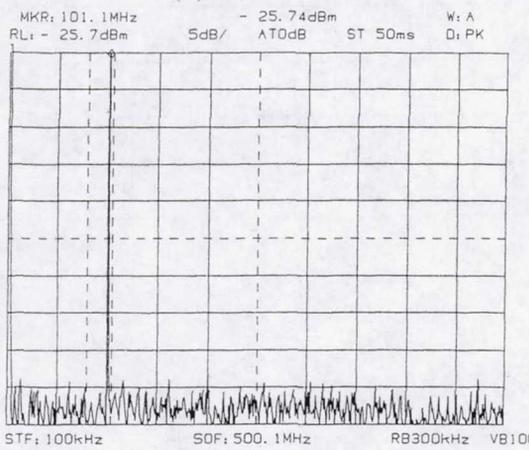
Figure 1: Experiments Set-up



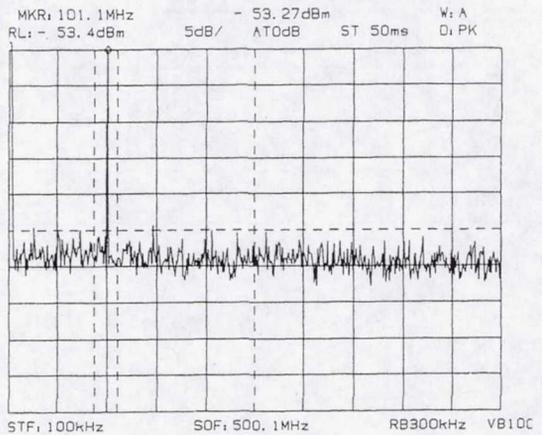
(a) 0 μm Offset



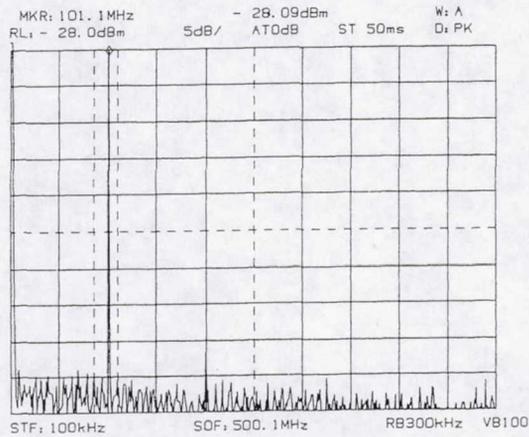
(d) 15 μm Offset



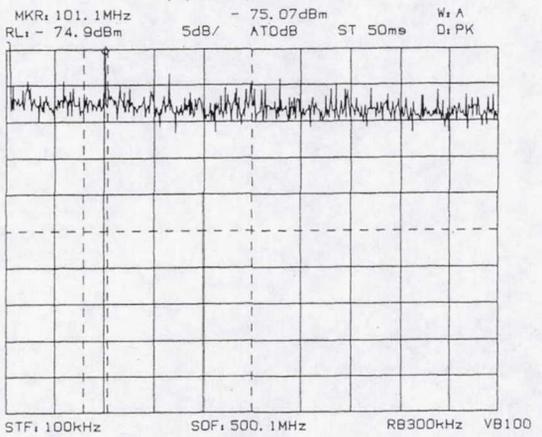
(b) 5 μm Offset



(e) 20 μm Offset



(c) 10 μm Offset



(f) 25 μm Offset

Figure 2: Signal-to-noise measurements for 8 km of MMF for different connector offsets. Signal frequency is 100 MHz for all six cases

4. RESULTS

4.1. Noise Measurement Results:

Output signal spectra for different offset connectors are shown in Figure 2. Similar measurements are repeated without including the 8 km MMF. Using spread sheet, the S/N due to modal noise is calculated as shown in Table 1.

Offset (μm)	0	5	10	15	20	25
Noise MMF (measured in mW)	-64.5	-73.0	-78.0	-81.0	-82.0	-82.9
Noise SMF (measured in mW)	-73.0	-76.0	-78.0	-82.0	-83.0	-83.2
P_{signal} (measured in mW)	-23.4	-25.6	-28.2	-38.0	-53.4	-74.0
$(S/N)_{\text{laser} + \text{receiver}}$ (calculated in dB)	49.6	50.4	49.8	44.0	29.6	9.2
$(S/N)_{\text{total}}$ (calculated in dB)	41.1	47.4	46.8	43.0	28.6	8.9
$(S/N)_{\text{modal}}$ (calculated in dB)	41.7	50.4	49.8	49.8	35.4	20.6
Modal Noise (calculated in dBm)	-65.1	-76.0	-78.0	-87.8	-88.8	-94.6
Modal Noise (calculated in μW)	3.5×10^{-4}	2.5×10^{-5}	1.5×10^{-5}	1.6×10^{-6}	1.3×10^{-6}	3.4×10^{-7}

Table 1: Calculations for modal noise and $(S/N)_{\text{modal}}$

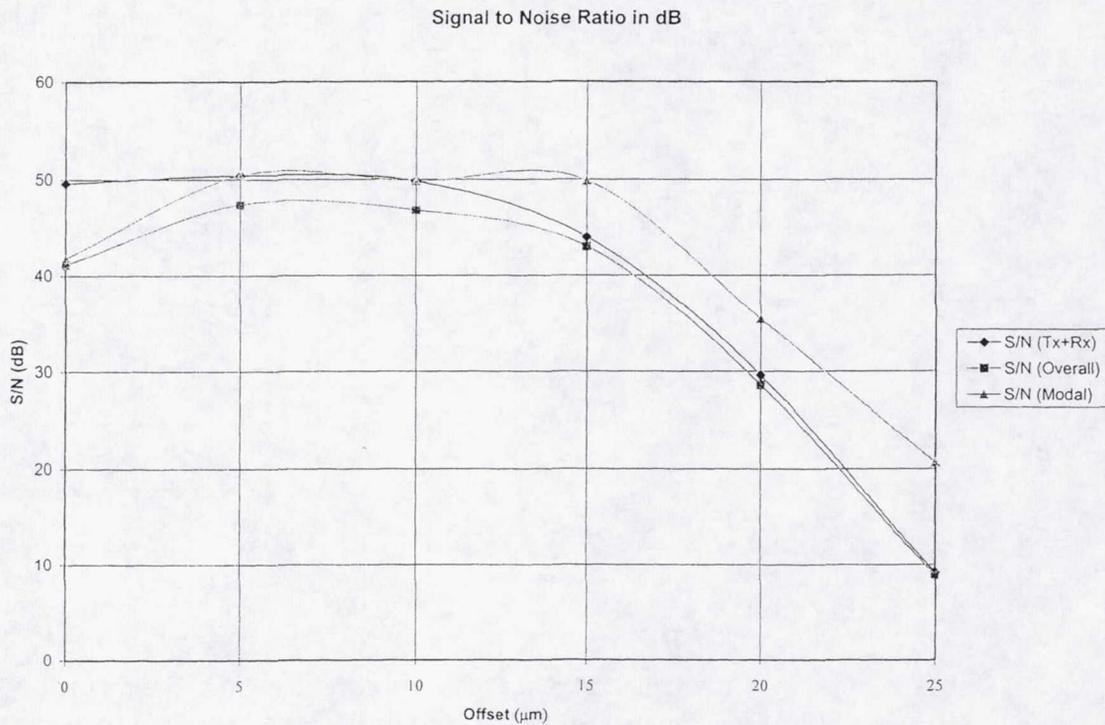


Figure 3: $(S/N)_{\text{dB}}$ vs. lateral offset in μm due to: laser + receiver, shot, and total noise.

We conclude from Table 1 and Figure 3, that shot noise decreases with increasing connector lateral offset. However, connector lateral offset increases signal attenuation as well. The combination of the two effects gives the curve shown with the symbol Δ . For 0 μm offset, we notice that the dominant noise is modal noise. For offsets larger than 15 μm the dominant noise is due to laser + receiver. Between 5 and 10 μm , modal noise is in the same order of magnitude of the sum of other noises. Hence, based only on S/N, the optimum operating point should be between 5 and 10 μm . However, there are other factors to be put into consideration such as attenuation and bandwidth.

4.2. Attenuation Measurements

Optical power P_o is measured immediately after the offset connector and after each segment (each segment is 1 km) up to 8 segments. The measured values are shown in Table 2.

Offset (μm)	0	5	10	15	20	25
P_o @ 0 km (dBm)	-0.4	-1.0	-1.0	-2.4	-2.4	-15.5
P_o @ 1 km (dBm)	-1.1	-1.5	-1.6	-4.3	-10.0	-20.5
P_o @ 2 km (dBm)	-2.0	-2.2	-2.8	-6.6	-13.5	-23.1
P_o @ 3 km (dBm)	-2.7	-3.2	-3.9	-8.0	-15.1	-24.6
P_o @ 4 km (dBm)	-3.4	-4.0	-5.2	-9.7	-17.0	-26.7
P_o @ 5 km (dBm)	-4.1	-4.6	-6.1	-10.7	-18.3	-27.9
P_o @ 6 km (dBm)	-5.1	-5.8	-7.4	-12.0	-19.7	-29.4
P_o @ 7 km (dBm)	-5.8	-6.5	-8.2	-12.8	-20.6	-30.1
P_o @ 8 km (dBm)	-6.7	-7.4	-9.2	-13.8	-21.7	-31.2

Table 2: Optical power P_o in dBm as function of connector lateral offset at different lengths of MMF.

Table 2 is graphed in Figure 4 from which we notice that for 0 μm offset the attenuation is linear vs length i.e. the attenuation of the first km of fiber is equal to the attenuation of the 8th km of the fiber. However, for large connector (such as 25 μm) offsets the attenuation at the beginning of the fiber is much larger than the attenuation at the end of the fiber (more than five folds). This is because most of the modes excited by lateral offset launch are “radiation modes,” these are the modes that are **not** confined in the core [6]. However, zero connector-to-connector offset mainly excites “meridional and helical modes”. These modes are guided inside the core. We Also notice that connector loss jumps suddenly from 2.4 dB to 15.5 dB when offset is incremented from 20 μm to 25 μm . This is due to the high radiation loss occurring at the first 10 m of MMF connecting the offset connector with the Optical Power Meter.

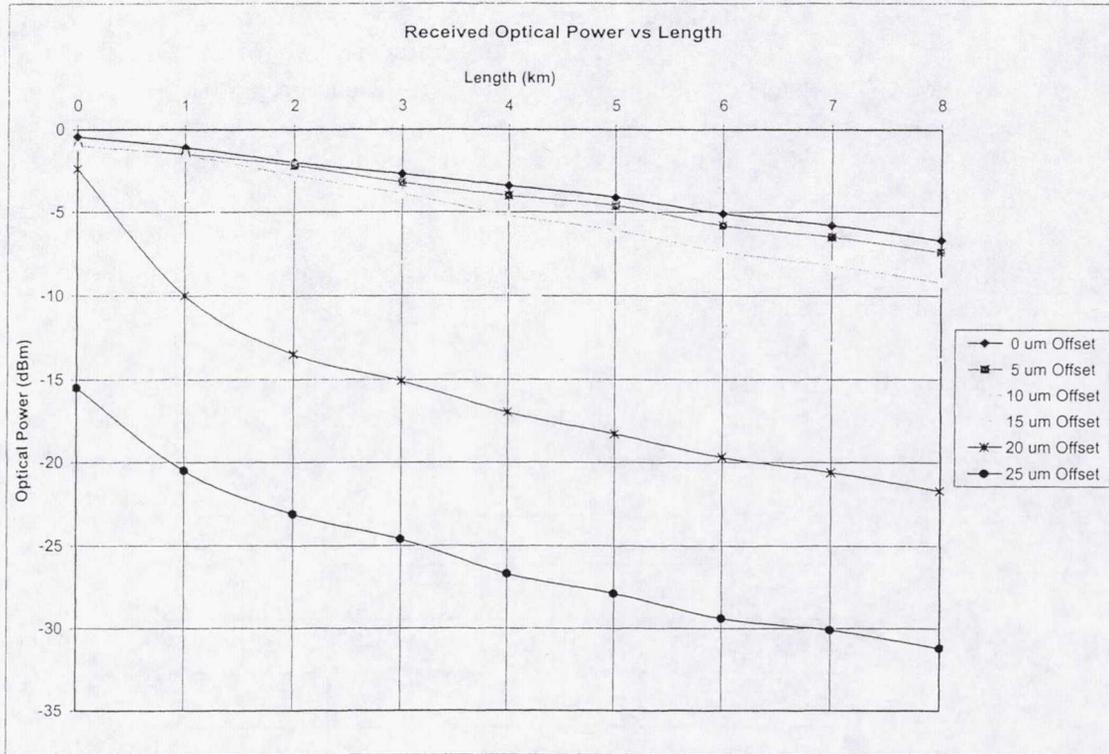


Figure 4: Received optical power in dBm vs. MMF length in km for different connector offsets.

4.3. Bandwidth Measurements

Electrical bandwidth is measured after each segment (each segment is 1 km) up to 8 segments. The measured values are shown in Table 3.

Offset (μm)	0	5	10	15	20	25
BW @ 1 km (MHz)	694	890	1100	1200	1300	1500
BW @ 2 km (MHz)	400	536	675	676	715	810
BW @ 3 km (MHz)	382	450	580	630	665	675
BW @ 4 km (MHz)	346	370	441	520	535	540
BW @ 5 km (MHz)	308	330	406	428	440	510
BW @ 6 km (MHz)	279	309	389	396	405	460
BW @ 7 km (MHz)	193	238	328	346	375	395
BW @ 8 km (MHz)	189	225	300	310	335	360

Table 3: Bandwidth in MHz as function of connector lateral offset at different lengths of MMF.

The plot of this table is shown in Figure 5.

Bandwidth vs. Number of MMF Segments

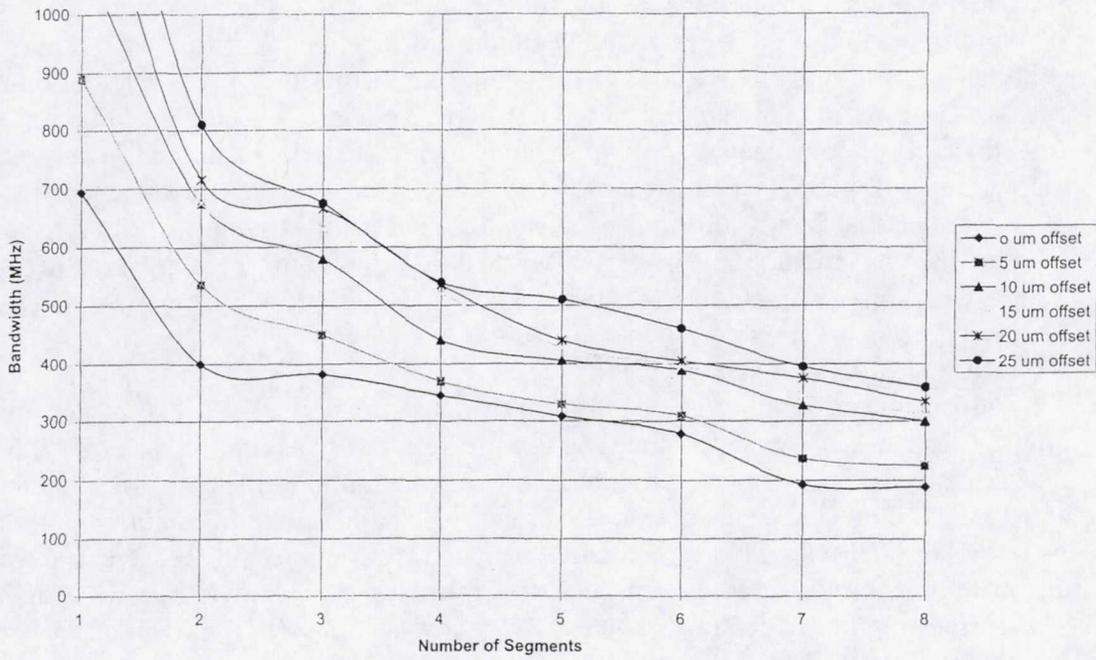


Figure 5: Bandwidth in MHz vs. MMF number of segments (1 km each) for different connector offsets.

Bandwidth vs Offset

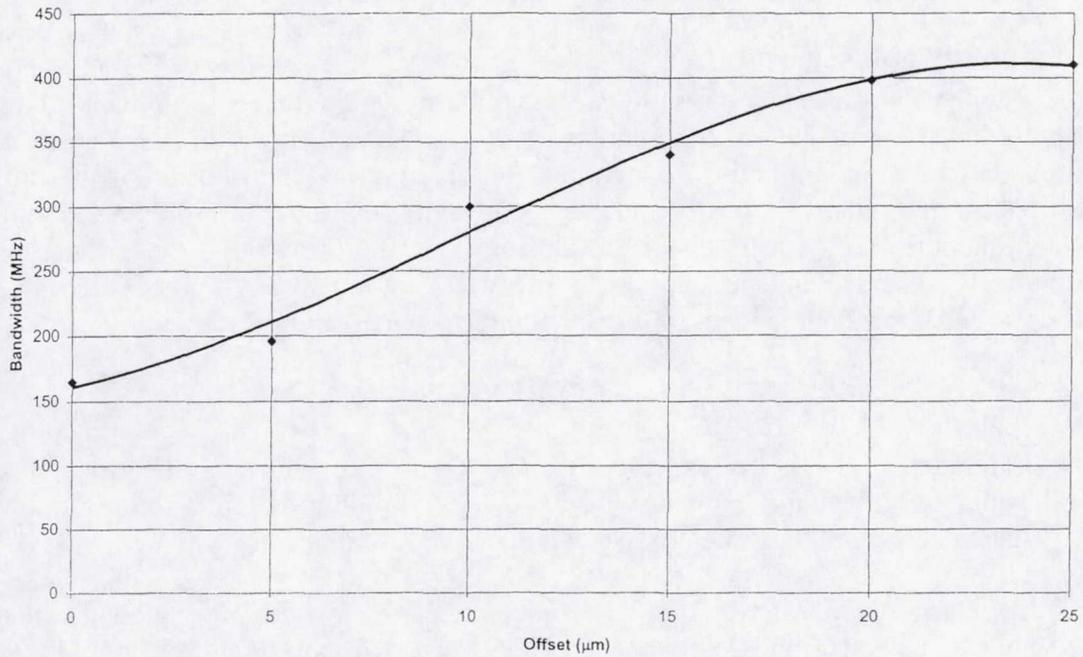


Figure 6: Bandwidth in MHz vs. lateral offset in um for an 8 km MMF link.

We observe here that the dispersion is large at the beginning of the MMF link and it level off after few segments. This could be explained by the fact that modal dispersion, which is dominant in MMF, is proportional to the number of modes. The number of modes is larger at the beginning of the fiber link than at the end of the fiber link. Propagating modes at the beginning of the fiber include radiation modes as well as meridional and helical modes. After couple of kilometers, radiation modes radiate off the core leaving only meridional and helical modes in the fiber core.

We obtain Figure 6 when the length is fixed to 8 km and the experiment is repeated. We notice that bandwidth improvement for 25 μm offset is almost three fold better than the bandwidth at on-line SMF-to-MMF coupling. Explanation of this phenomenon is well studied previously [1-6]

5. CONCLUSION

It is evident that extending the bandwidth-distance product of a MMF is possible. A simple lateral offset connector could be used to selectively excite higher order modes to extend the bandwidth-distance product. However, the optimum lateral offset depends on signal modulation, and transmission distance. Depending on the minimum required S/N, the optimum connector offset is initially estimated from Figure 3. Receiver's sensitivity and maximum bit rate could then be found from Figure 4 and Figure 5. For binary digital, transmission is achievable with S/N as low as ~ 20 dB. From Figure 3 we find that the offset could be as large as 22 μm for a link of 8 km. Receiver's sensitivity is then estimated from Figure 4 for an 8 km link to be equal to ~ -23 dBm (for an 8 km link). Depending on the distance, the maximum bit rate transmission could then be determined from Figure 5 (maximum bit rate $\approx 2 \times$ bandwidth). Following the same procedure for analog transmission, connector offset is estimated to be 15 μm . However, a significant enhancement in S/N (more than 5 dB) is achievable with a SMF-to-MMF 5-10 μm offset.

ACKNOWLEDGMENTS

The author would like to express his appreciation to Po T. Huang for allowing him the opportunity to participate in his ongoing project. The author would like also to thank Robert Swindle for his help in experiment set-ups and his contribution to the project. Moreover, the author is greatly indebted to: David Johnson for providing computer-related support; Professor Banmali Rawat, from The University of Nevada, Reno, for his stimulating discussions and support; Professor Ramon Hosler, Judie Gilliam, Jane Hodges, and Greg Buckingham for organizing the program and its activities.

REFERENCES

- [1] L. Raddatz, et al, "An experimental and theoretical study of the offset launch technique for the enhancement of the bandwidth of multimode fiber links," *Journal of Lightwave Technology*, Vol.16, No. 3, pp 324-331, March 1998.
- [2] M. Hackert, "Characterizing multimode fiber bandwidth for gigabit ethernet applications," Corning White paper, WP 4062, Corning Inc. Telecom. Products Division, November 1998.

- [3] Z. Hass and M Santoro, "A mode-filtering scheme for improvement of bandwidth-distance product in multimode fiber systems," Journal of Lightwave Technology, Vol.11, No. 7, pp 1125-1131, July 1993.
- [4] Z. Hass and M Santoro, "Lightwave transmission systems using selected optical modes," US Patent no 44318, 1993.
- [5] G. Papen and M. Murphy, "Modal noise in multimode fibers under restricted launch conditions," Journal of Lightwave Tech., Vol.17, No. 5, pp 817-822, May 1999.
- [6] D. Marcuse, "Principals of optical fiber measurement," ISBN 0-12-470980-X, Academic Press Inc., 1981.

1999 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

**JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA**

**STEP: WHAT IS IT AND SHOULD IT BE USED
FOR KSC'S ISE/CEE PROJECT IN THE NEAR FUTURE?**

Dr. Catherine C. Bareiss
Associate Professor
Computer Science
Olivet Nazarene University

Mike Conroy
NASA/KSC

ABSTRACT

The ability to exchange information between different engineering software (i.e. CAD, CAE, CAM) is necessary to aid in collaborative engineering. There are a number of different ways to accomplish this goal. One popular method is to transfer data via different file formats. However this method can lose data and becomes complex as more file formats are added. Another method is to use a standard protocol. STEP is one such standard. This paper gives an overview of STEP, provides a list of where to access more information, and develops guidelines to aid the reader in deciding if STEP is appropriate for his/her use.

STEP: WHAT IS IT AND SHOULD IT BE USED FOR KSC'S ISE/CEE PROJECT IN THE NEAR FUTURE?

Dr. Catherine C. Bareiss

1. Introduction

A. Different methods for information exchange

This paper looks into one specific way to exchange information between different engineering software packages. This problem is similar to the problem of information exchange between different office productivity software (i.e. word processors, spreadsheets, presentation software, databases). Even though the data for engineering models is much more complex than that of office productivity software, much can be learned by looking at the different methods of office productivity software information exchange.

The easiest way to exchange office documents is by using the same version of the same software. This is often done within a company that has chosen to standardize on one package. When using one vendor (i.e. Microsoft, Lotus, Corel, etc.) another method for exchange of information is found between different types of software (i.e. spreadsheets, word processing, etc.). Platform developers have designed methods to make it very easy to access information in one document (for example, a spreadsheet) in another (for example, a paper created in a word processor). A third way to exchange information is used when working with different platforms. The developers often include routines that allow their package to read data created by other packages. A fourth way involves total transparent information interchange and can be seen in database packages. Platform independent standards (i.e. SQL and ODBC) have been developed, and most database packages have implemented these standards.

Options for information interchange in engineering packages are similar. The easiest option is for an organization to standardize on one package. However, as in office software, it is common to need different packages to work on different aspects of engineering modeling (i.e. word processor and spreadsheets, mechanical and electrical engineering). The next easiest option would be to standardize within an organization on software developed by the same developer that allows for information interchange. However, such software is rare, and the politics are such that it may be difficult to require all engineers associated with an organization to use software from the same developer. The next option would be to use packages that include routines to read other data formats. But as in the case of office software, it is usual that detailed information can be lost. Also, these can get to be a very large package because as each new file format is added to the conversion routines, all previous file formats must be allowed to convert to and from the new one. Very quickly, this becomes too complex and large. In addition, upgrading versions can often cause problems with transferring the more complex data that the newer versions allow.

The fourth option (and the one presented in this paper) is to use a platform independent standard for information interchange. The three most common are ORB, CORBA, and STEP. The first two are generic standards for any type of communication protocol. STEP is an international standard designed to facilitate the exchange of information specifically for engineering and product design, and is the one covered in this paper. While this option seems the best, it is important to keep in mind that STEP is still not finished and not completely implemented commercially. This may require additional on-site programming to support the exchange between different software packages depending on the needs of an organization.

B. Overview of STEP

To better understand if STEP is worth any additional required effort, STEP must be better understood. STEP (STandard for Exchange of Product model data) is an ISO standard (ISO 10303 industrial automation systems - product data representation and exchange). STEP is a very large and complex standard. While it is very difficult for any one person to completely understand all of STEP, the components that are necessary to determine if it can meet KSC's ISE/CEE immediate needs can be understood by a single individual.

There are four major components of STEP that are of interest: integrated resources (IRs), application protocols (APs), application modules (AMs) and EXPRESS (a description method). An AP defines the protocol for describing all the information associated with a given area (i.e. electrical) of engineering. APs are referred to as parts 201-299. Each AP is made of one or more IRs. The IRs are used to describe "components" that might be of interested in one or more AP's (for example, the shape of an object). IRs are further divided into 2 categories. Parts 41-99 are integrated generic resources whereas parts 101-199 are integrated application resources. Generic resources are independent of the applications and can reference each other. Application resources can reference generic resources and are used to add additional constructs that might be used by a group of similar applications. AMs are used to help the sharing of information between APs and the speed up the development of APs. They are a new concept are still under development. EXPRESS is the data definition language for STEP and is used to define that structures of APs, AMs, and IRs. It is the first of the description methods (referred as parts 11-19). The other parts of STEP include Part 1 (which is the overview), parts 21-29 (implementation methods), parts 31-39 (conformance testing methodology and framework), parts 301-399 (abstract test suites), and parts 501-599 (application interpreted constructs) which are components of application protocols.

C. Guide to the rest of the document

The next section will introduce STEP by covering its history. The second section explains the contents of STEP that might be of importance to KSC's ISE/CEE project. It includes a list of acronyms and terms are given to help understand the paper and information to help understand integrated resources, application protocols, application modules, and EXPRESS and implementation methods. The third section talks about what is involved in PDM (product data management) software and how STEP fits in that picture. The fourth section (which is the conclusion) is designed to help the reader decide if STEP would help meet the immediate and short term needs of the KSC's ISE/CEE project.

2. History of STEP

Work on STEP was officially started in 1984. It was proceeded in the 1970s by three specifications: IGRES (Initial Graphics Exchange Specification) in the USA, SET (Standard D'Exchange et de Transfert) in France, and VDA-FS (Verband der Automobilindustrie-Flachen-Schnittstelle) in Germany.

STEP started with the following objectives.

1. The creation of a single international standard, covering all aspects of CAD/CAM data exchange.
2. The implementation and acceptance of this standard in industry, superseding various national and de facto standards and specifications.
3. The standardization of a mechanism for describing product data, throughout the life-cycle of a product, and independent of any particular system.
4. The separation of the description of product data from its implementation, such that the standard would not only be suitable for neutral file exchange, but also provide the basis for shared product databases, and for long-term archiving.

The initial release of STEP occurred in 1995 with the following twelve parts.

Part 1	Overview and Fundamental Principles
Part 11	The EXPRESS Language Reference Manual
Part 21	Clear Text Encoding of the Physical File Exchange Structure
Part 31	Testing Methodology and Framework: General Concepts
Part 41	Fundamentals of Product Description and Support
Part 42	Geometric and Topological Representation
Part 43	Representation Structure
Part 44	Product Structure Configuration
Part 46	Visual Presentation
Part 101	Draughting
Part 201	Explicit Draughting
Part 203	Configuration Controlled Design

3. Overview of STEP

A. TERMS

The following are a list of acronyms and terms that will be of help for understanding the rest of the paper.

AAM	Application Activity Model
AIC	Application Interpreted Construct
AIM	Application Interpreted Model
AM	Application Module - an extension of an AIC
AP	Application Protocol
API	Application Program Interface
ARM	Application Reference Model
CEE	Collaborative Engineering Environment
ER	Entity-Relationship
EXPRESS	A data definition language defined in STEP
EXPRESS-G	A graphical standard (defined in STEP) to represent a portion of EXPRESS
IR	Integrated resource
IAR	Integrated Application Resource
IGR	Integrated Generic Resource
ISE	Intelligent Synthesis Environment
ISO	International Standards Organization
PDM	Product Data Management
SDAI	STEP Data Access Interface
STEP	STandard data Exchange Protocol

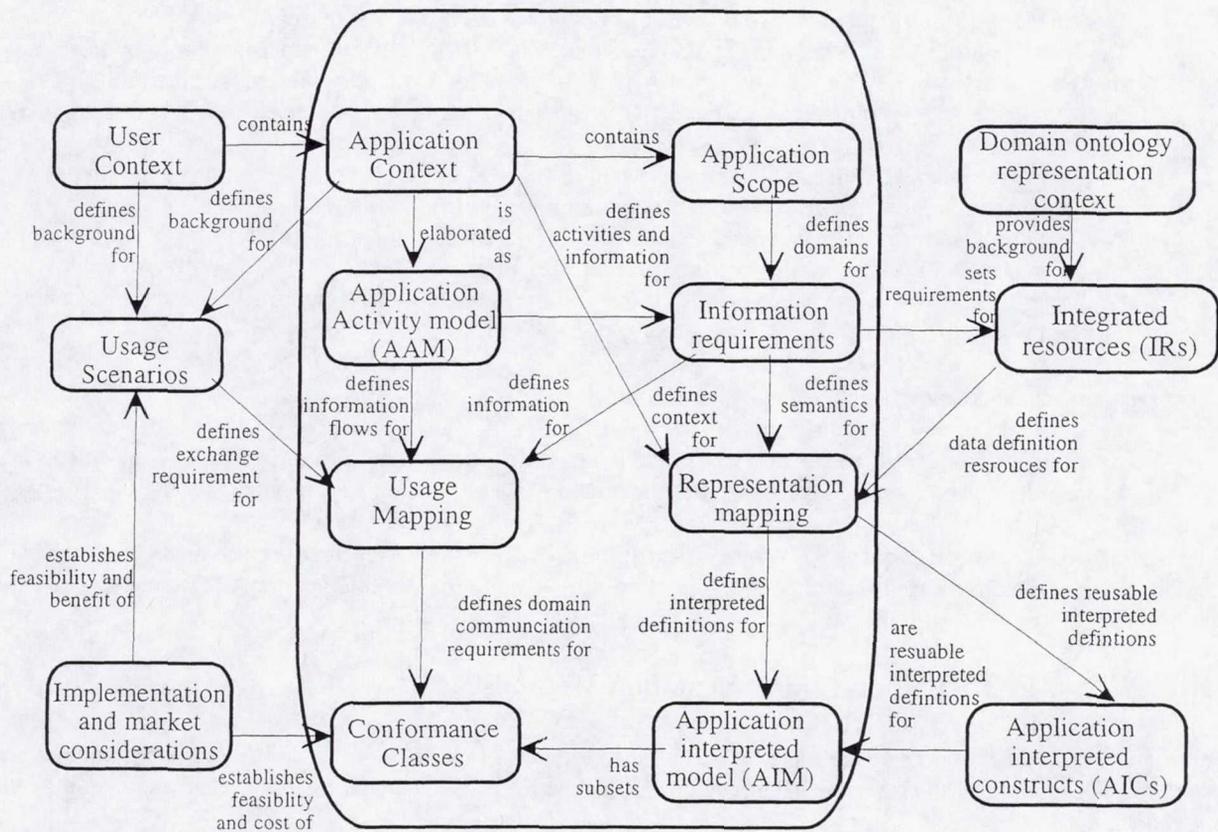
B. Integrated Resources

Integrated resources are used to describe components that might be of interest to one or more APs. There are two categories of integrated resources. Generic resources are application independent, and application resources support the needs of a group of applications. The integrated generic resources support common themes such as geometric and topological representation, materials, visual presentation, and process structure and properties. Integrated application resources include draughting, finite element analysis, and kinematics. While IRs are used to develop APs, they are not sufficient by themselves to meet the needs of an application and are therefore not used by themselves but are only used as part of an application protocol.

C. Application Protocols

Application protocols are the primary way to interface with STEP for most users. APs define the requirements for a specific application of product data related to a particular industry. They do this by combining both integrated generic resources and integrated application resources. The numbering systems allows for ninety nine different APs with thirty two being worked on, finalized, or retired. As industry needs expand, more APs will most likely be developed with each AP possibly requiring over one year to be developed.

APs are more complex than just identifying which integrated resources to use. They start with one or more application contexts which contain the descriptions of the functionalities, technologies, types of product, disciplines, industry sectors, and life cycle stages that comprise the background knowledge of users. The AAM (application activity model) elaborates the activities and flow of information between activities for given application context. These activities and flow of information are identified by the application scope. Both the application scope and AAM are used to define



the information requirements of the users of a particular AP. These requirements include the natural language definitions of the application objects and application assertions using the terminology of the users. These requirements can be displayed graphically via an application reference model (ARM).

The AAM, information requirements, and usage scenarios are combined to define usage mappings. These mappings identify which activities, information flow, and requirements are needed by each usage scenario. The usage mappings are then used to define conformance classes which group specific information requirements based upon usage scenarios that describe how the information is expected to be used. These classes are also influenced by implementation and market considerations.

A representation mapping combines the application contexts, information requirements, and IRs to create an application interpreted model (AIM) which is the formal specification for communication when satisfying a conformance class of an AP. Groups of AIM constructs may represent a common semantic in more than one AIM. These are combined into reusable application interpreted constructs (AIC). An AIC satisfies information requirements that are common to more than one AP.

D. Application Modules

Application modules (AMs) extend the AIC (application interpreted construct) concept by including the relevant portions of the AP application reference model. This extension was done for five reasons.

1. To reduce the high cost of developing an application protocol. (It currently takes a team of four or more people about a year and a half to produce an initial AP specification for a typical engineering function.)
2. To ensure the ability to implement a combination of subsets of multiple APs or to extend existing APs to meet a business need.
3. To ensure the ability to reuse application software developed to support one AP in the development of an implementation of another AP with the same or similar requirements.
4. To avoid duplication and repeated documentation of the same requirements in different application protocols leading to potentially different solutions for the same requirements.
5. To ensure the ability to reuse data generated by an implementation of one or more APs by an implementation of one or more different APs.

There are three major components to an application module.

1. the scope and functional requirements;
2. the application reference model as a representation of the application domain information requirements; and
3. the module interpreted model that specifies the required use of the common resource.

The use of AMs is relatively new to STEP and is just starting. No standards have yet to be produced using AMs but they are being worked on (esp. in terms of PDM applications). Any future development in APs should also develop and use AMs.

E. EXPRESS and Implementation Methods

i. EXPRESS

Express is a data definition language that describes the structure of the IR's, AP's, and AM's (including the data structure and constraints) and is used in a number of standardization projects including STEP. It does not contain the actual data, nor is it executable.

EXPRESS supports the following modeling capabilities:

- definition of data entities, attributes, and relationships,
- the specification of local and global constraints on these, and

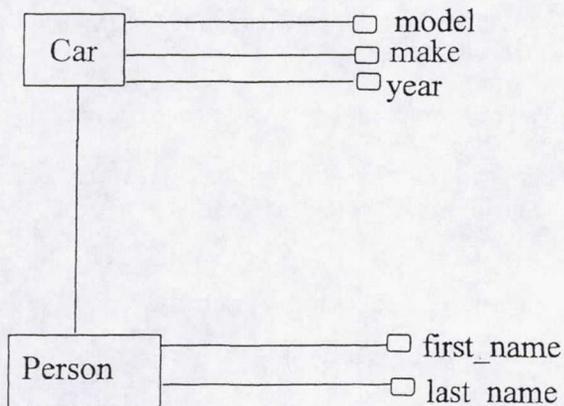
- the collection of data definitions and constraints in separate schemata, supporting modular development of data models.

An example of a data model in EXPRESS might include the following:

```
ENTITY car;  
  make   : STRING;  
  model  : STRING;  
  year   : INTEGER;  
  owner  : person;  
END_ENTITY;  
ENTITY person;  
  first_name : STRING;  
  last_name  : STRING;  
END_ENTITY;
```

EXPRESS has seven constructs: schema, type, entity, constant, function, procedure, and rule. A complex EXPRESS schema (which is very typical) can be very hard to follow, so EXPRESS-G was developed and displays a subset of EXPRESS graphically.

EXPRESS-G is similar to ER diagrams in that it has definitions (similar to entities), relationships, and compositions (which allow a diagram to span more than one page). It does not support functions, procedures, or rules. It supports simple data types, named data types, relationships, cardinality, and one or more schemas. See following diagram for an example.



ii. Implementation Methods

Currently there are six implementation methods defined. The first is a physical file exchange structure. This method uses EXPRESS-I (instantiation language) and records this is a plain ASCII file. Below is the data portion for an instance of a car.

```
#1 = CAR ("Saturn", "SL", 1996, #2);  
#2 = PERSON ("John", "Doe");
```

While this method is sufficient (can support all file exchange), it does have a number of disadvantages.

- It can be very slow for accessing specific information.
- It does not support concurrent access.
- It does not support distributed storage.
- It can be difficult to share information between different AP's (since the data may be structured differently and software would be needed to converted to the structure of the other AP).

Part 22 defines SDAI (STEP Data Access Interface) which defines how to interface with advanced data storage methods (i.e relational databases, object-oriented database, knowledge bases, etc.). It defines the methods used by an application to access the data. It is then up to the data storage method to support these methods. The last four parts (23, 24, 25, 26) of the implementation methods define bindings for languages. Part 23 defines the C++ language binding to the SDAI, while part 24 defines the C language binding and part 25 defines the late FORTRAN binding. Part 26 defines the interface definition language binding to SDAI. These are used to support language access to the STEP data. They could be used when developing necessary code to meet any needs that cannot be met via commercially available products.

4. PDM

The purpose of a Product Data Management system(PDM) is to help an organization track the information associated with a product throughout its entire life-cycle. This can include multiple versions, views of parts and/or products, managing the hierarchy of a product composition, tracking the status of parts and/or products, and/or managing the information associated with different variations of a part and/or product.

The field of PDM software is still be refined and researched. Some of the still open questions include:

- 1) What is the theoretical required information to be managed?
- 2) Since different organizations have different life-cycles for products, what is the best way to design a system to allow for these differences and still ensure that it will support all needs?
- 3) How should PDM systems be integrated with design software (i.e. CAD, etc.)?

Since the problem of PDM systems have not been completely solved, if one wants to use such a system, there are a number of questions to ask.

- 1) How does my organization manage the life-cycle of a product?
- 2) What information is important to be tracked?
- 3) What software can best meet answer one and two?
- 4) What software can best interface with other software (i.e. CAD) being used to developed the product?

The fourth question is where STEP plays a role. Within the STEP community there is the PDM Implementor Forum. This form is made up of software developers and testers who are working on PDM systems and translators based on STEP. This form is called PDES, Inc. consortium and was started in February of 1995.

The STEP PDM Schema (as of May 1999) contains the intersection of AP 203, AP 212, AP 214, and AP 232. It includes the following units: part identification, part classification, part properties, part structure and relationships, document identification, document classification, document and file properties, document structure and relationships, external files, document and file association to product data, alias identification, authorization, configuration and effectivity information, and work management data.

5. Conclusions

A. What options exist to meet KSC's ISE/CEE's immediate and short-term needs?

When trying to decide if STEP should be used for the immediate needs of the ISE/CEE environment at KSC, there are a number of issues to look at. For each alternative one needs to consider the costs and benefits. The options are to:

1. Standardize on one software package
2. Use existing protocols to convert to standard file formats
3. Develop own information exchange protocols for limited needs
4. Use STEP and develop own software to fill in the gaps.

It is assumed that options 1 and 2 will not meet the immediate requirements of ISE/CEE. If either one does, then there is no cost and that option should be used. This leaves two options: total internal development or the use of STEP.

B. Why one might not use STEP.

There are three primary disadvantages with using STEP: training, development costs, and complexity with the complexity being the primary cause of the other two. Because of this complexity, the system developers will need to invest a significant amount of time in becoming proficient in the use and development of APs along with the remaining parts of STEP. Also, if the existing software and/or APs do not meet all the current requirements, there will be a lot of time and effort necessary to develop these. Because one would be developing generalized software and APs instead of ones designed to meet the exact needs of KSC's ISE/CEE, more effort would be required.

C. Why one might use STEP.

There are four major advantages of using STEP: already developed application protocols, investment in the future, a large amount of international commitment, and the existence of commercial products.

With STEP, the application protocols have either already been developed or are in the process of being developed. The process of developing APs (either for STEP or the equivalent of own software) is a large, time-consuming task. It involves correctly understanding the application domain, identifying all requirements in that area, creation of application protocols, and testing the final results. There are many places in this process where errors can develop. By using STEP, either the application protocol is finished (and "error-free") or is being developed by a "large" group of people investing in this process. This would tend to make the development of the application protocols more error-free in STEP than doing it by own self given the complexity of the task. In addition, the cost of develop of APs should be decreasing as the development and use of AMs increase. A presentation by Julian Fowler from PDT Solutions in 1997 also indicated

that "AP interoperability" is still perceived as a problem. This is where a good deal of current effort in STEP is being focused.

In May of 1999, the NASA CIO officially approved NASA-STD-2817, which includes the requirement for CAE/CAD/CAM systems used by NASA to have interchange tools that support STEP. This includes making sure that tools to enable data interchange (which are compliant with STEP, currently AP 203, AP209, AP210, AP 225, and AP 227)) be available to all CAE/CAD/CAM. users at each NASA center. In addition NASA is a member of PDES, Inc., which is a joint industry/government consortium specifically formed to accelerate the development and implementation of STEP. It appears as if NASA's future includes a strong commitment to STEP. Therefore any investment in STEP now (both learning and development of systems) is an investment in the future.

There is a large international commitment to STEP (over 25 nations), in addition to a significant commitment from vendors and industry (a total of more than 200 companies involved). This also indicates that the initial investment in time and energy by KSC would not be wasted.

D. How to decide.

So with these advantages and disadvantages, should KSC use STEP to meet the immediate and short term needs of the ISE/CEE? How should KSC go about answering that question?

Step 1: Identify the needs. This includes identifying the CAx software used and the related APs in STEP.

Step 2: Determine the status of the APs. Are they still under development?

Step 3: Identify if there is existing software that implement the necessary APs and what software exists that can help with data interchange that is not based on STEP. These steps will identify how much can be achieved via "off-the-shelf" vs. internal development, and how much work is involved in using STEP vs internal development vs other commercially available methods.

Step 4: Make the decision. After determining how much would need to be developed on-site, the decision on using STEP can be made based upon the one wishes to invest in the future vs. the need for immediate solutions.

References

Danner, William F., "Developing Application Protocols (APs) Using the architecture and methods of STEP (STandard for the Exchange of Product Data). Fundamentals of the STEP Methodology", NISTIR 5972, Jan. 1997.

"Guidelines for application module development, Revision 0.4: Proposed ISO TC 184/SC4 Standing Document", 1997, ISO TC 184/SC4 Nxxx

Kindrick, Jim and Markus Hauses, "Usage Guide for the STEP PDM Schema", May 27, 1999, <http://www.pdm-if.org>.

"Product Data Exchange Using STEP (PDES) Part 1 - Overview and Fundamental Principles", ANS US PRO/IPO-200-001-1994

"Product Data Exchange Using STEP (PDES) Part 11 - The EXPRESS Language Reference Manual", ANS US PRO/IPO-200-011-1994

1999 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA

MEASURES OF WATER QUALITY IN MERRIT ISLAND NATIONAL WILDLIFE
REFUGE IMPOUNDMENTS AND ADJACENT INDIAN RIVER LAGOON

Linda K. Blum
Research Associate Professor
University of Virginia
Kelly Gorman

ABSTRACT

The goal of this project was to conduct preliminary investigations to determine appropriate sampling strategies to measure the flux of dissolved nutrients (specifically, NH_4^+ , NO_3^- , NO_2^- , and PO_4^{3-}) and suspended particulate matter (TSS) between impoundments and the IRL in preparation for an intensive 3-yr monitoring program. In addition to nutrients and TSS, a variety of common water quality indicators were also measured during these preliminary studies. Six impoundments and a single restored marsh selected for study. Over a month long period, water samples were collected weekly at selected impoundment culverts. Water was collected in duplicate as independent grab samples from both the lagoon side and within the perimeter ditch directly adjacent to the culverts. Water quality indicators inside and outside the marsh impoundments were different. Ammonium, salinity, bacteria and chlorophyll-*a* were higher inside the impoundments as expected possibly as a result of the great affect of evaporation on impoundment water. Water quality indicators responded rapidly both inside and outside the impoundments as exemplified by the increase in NH_4^+ -N concentrations during a horseshoe crab die-off. Water quality indicators were high variable during the month in which water samples were collected. Because the impoundments are widely spaced it is logistically unrealistic to sample each of the impoundments and associated seagrass beds on a single day, sampling must be stratified to allow patterns of material movement and the annual flux of materials to and from the impoundments to be determined.

MEASURES OF WATER QUALITY IN MERRITT ISLAND NATIONAL WILDLIFE REFUGE IMPOUNDMENTS AND ADJACENT INDIAN RIVER LAGOON

Linda K. Blum

1. INTRODUCTION

In much of the coastal United States, large areas of salt marshes have been impounded to control mosquitoes and sea-level intrusion into agricultural land. At Merritt Island National Wildlife Refuge (MINWR) marsh impoundment began as early as 1956 and was accelerated when NASA funded the mosquito control districts in the late 1950's and early 1960's. Most of the impoundments were completed by 1966. In 1963, NASA transferred management authority for many of the impoundments to the U.S. Fish and Wildlife Service (USFWS). The MINWR wetlands are presently managed by USFWS biologists under a multispecies management program with focused on, but not limited to, fish, wildlife, plant communities, and consumptive and non-consumptive public uses (Epstein, 1995).

To enhance biodiversity and provide better management of the MINWR wetlands restoration of impoundment T-10-K was begun when it was drained in 1969. T-10-K was fully restored (the dikes were removed and the perimeter ditches filled) in 1977 (Leenhouts 1982). Restoration efforts continued when refuge managers began installing culverts to reconnect impoundments to the adjacent lagoons in the 1970's. In addition to enhancing refuge biodiversity and restoring wetland interaction with the lagoons, reconnection of the impoundments has created unique opportunity for scientists to study the response of the lagoon-marsh ecosystem to restoration efforts. Reconnection of the impoundments to the lagoon has restored the flow of materials, both dissolved and particulate, between the marsh impoundments and the lagoon. Although Day (1989) stresses that the flux of materials between estuaries and their adjacent wetlands is a critical process in estuarine ecosystems because of the impact on system biodiversity and production, little is known about how reconnection of impoundments will affect material movement of materials to and from the lagoon and the impoundments.

As a result, SJRWMD, NASA, USFW, and a number of universities have been awarded funding from EPA to examine a variety of questions focused on the response of the impoundments and lagoon to reconnection. One of the primary objectives of the EPA-funded project is to examine material flux between the impoundments and lagoon. The goal of my NASA/ASEE project during summer 1999 was to conduct preliminary investigations to determine appropriate sampling strategies to measure the flux of dissolved nutrients (specifically, NH_4^+ , NO_3^- , NO_2^- , and PO_4^{3-}) and suspended particulate matter (TSS) between impoundments and the IRL. In addition to nutrients and TSS, a variety of common water quality indicators were also measured during these preliminary studies.

2. MATERIALS AND METHODS

A. Impoundments: The criteria used in impoundment selection included: proximity to the IRL, presence of culverts installed through the impoundment dike directly to the IRL (i.e., the impoundments were reconnected to the IRL), and hydrologic management regime (Fig. 1). The six impoundments and single restored marsh selected for study vary in size (Fig. 1) and in plant community composition (Table 1).

Although a variety of structural marsh management techniques are used at MINWR, the two most commonly employed methods are wildlife aquatic management (WAM) and open. The combination of these techniques is used by MINWR biologists to maximize biodiversity and ecosystem productivity (M. Epstein, personal communication). The primary objective of the

WAM technique is to create habitat that favors utilization by ducks and wading birds by encouraging the growth of submerged aquatic vegetation, especially *Ruppia maritima* an important food source for herbivorous water birds. *R. maritima* growth is promoted by flooding the marsh surface with IRL water for approximately 8-10 months of the year. When the WAM impoundments are flooded, exchange of water with the IRL is limited to times when sufficient rain fall raises the level of the impoundment above that necessary for *R. maritima* growth or water level in impoundments falls below the water level in the lagoon and water is flow is from the lagoon into the impoundment. Water flow is from the impoundment to the lagoon during drawdown when the impoundment is allowed to drain allowing for consolidation of the sediment to provide substrate suitable for reestablishment of *R. maritima*.

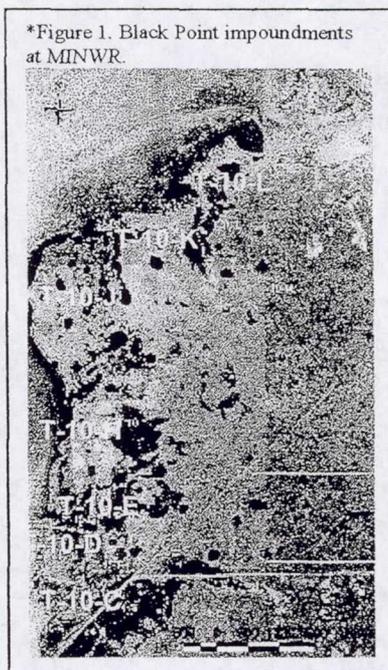


TABLE 1. Description of impoundments selected for water quality study during summer 1999.

Marsh Impoundment	Management Type	Months Flooded per Year	% Emergent Plant Cover
T-10-J	WAM	8-10	90
T-10-L	WAM/RIM	8-10/4-6	50
T-10-H	Open	2-3	35
T-10-C	Open	2-3	5
T-10-D	WAM/RIM	8-10/4-6	2
T-10-E	WAM	8-10	2

The primary objectives of open impoundment management are to provide estuarine organisms with access to the marsh surface and to allow movement of suspended materials to and from the lagoon. Open management is critical to some commercially and ecologically important fishes such as tarpon, snook, red and black drum, spot, and blue crabs (Gilmore 1999). Furthermore, in many lagoon-marsh systems, overall lagoon production and biodiversity are dependent on the export of particulate organic material from the marshes to the lagoon (Nixon, 1980) while marshes are maintained by import of mineral particles from the lagoon (Day and Templet 1989). The movement of materials (including organisms) to and from the lagoon is accomplished by leaving the culverts open to natural changes in water level so that

TABLE 2. Sampling dates and the direction of water flow in the culverts during preliminary water quality sampling, summer 1999.

Date	Impoundment Sampled	Direction of Water Flow
June 22	T-10-C	no flow
	T-10-D	no flow
	T-10-H	no flow
	T-10-E	to lagoon
June 29	T-10-J	no flow
	T-10-H	to lagoon
	T-10-L	to lagoon
	T-10-C	no flow
July 6	T-10-D	to lagoon
	T-10-H	from lagoon
	T-10-L	to lagoon
	T-10-E	no flow
July 13	T-10-H	no flow
	T-10-J	no flow

open impoundments are flooded approximately 2-3 months of year.

B. Sampling: Over a month long period, water samples were collected weekly at selected impoundment culverts. Water was collected in duplicate as independent grab samples from both the lagoon side and within the perimeter ditch directly adjacent to the culverts.

Logistical constrains allowed for sampling only 3 or 4 impoundments each week (Table 2). All samples were collected between 0830 and 1130 in clean, 4-L cubitainers and held in ice chests until return to the lab for sample processing.

C. Water Quality Measures: In the field, the direction of water flow was determined visually (Table 2). Measurements of dissolved oxygen (DO), salinity, and water temperature were made (Table 3). Immediately upon return to the laboratory sample pH was determined in each cubitainer and subsamples for determination of bacterial abundance were preserved in 2% formaldehyde (Table 3). Within 4 h or less of return to the laboratory, samples were filtered for determination of NH_4^+ , NO_3^- , NO_2^- , PO_4^{3-} , TSS, particulate organic matter (POM), mineral matter, and chlorophyll-*a* (chl-*a*) (Table 3). Samples for nutrient analysis were preserved by filtration through 0.45 μm pore size Gelman supor membrane filters. Filtered nutrient samples were frozen until analysis, which occurred with 9 days or less from sample collection (Table 3). Samples for TSS, POM, mineral matter and chl-*a* were filtered collected on Whatman GF/C filters and processed immediately upon collection (Table 3).

Table 3. Major Water Quality Measures and Method of Analysis

VARIABLE	DESCRIPTION	APPROVED METHOD
Total suspended solids	gravimetric	EPA 160.2
Organic and mineral particulates	gravimetric	EPA 160.4
NO_2^- and NO_3^-	Cd reduction	EPA 353.1 & EPA353.2
NH_4^+ ortho- PO_4^{3-}	Phenol hypochlorite	EPA 351 EPA 365.1 & EPA 365.2
salinity		
Chlorophyll- <i>a</i>	acetone-DMSO extraction	SM10200H
Bacteria	Direct microscopic count	Hobbie et. al. 1977
pH	Orion pH meter	
Dissolved oxygen	YSI DO meter	
Water temperature	YSI meters	

3. RESULTS

A. Impoundment Comparisons:

Examination of the entire water quality indicator data set by PCA showed that PC1, PC2 and PC3 accounted for 36.2%, 18.5%, and 14.9%, respectively, of the variance in the data set or a total variance of 69.6%. A clear pattern was observed when PC1 (comprised of TSS, POM and mineral particles) and PC2 (comprised of chlorophyll-*a* and bacterial abundance) and PC2 and PC3 (comprised of NH_4^+ -N) were compared to one another (Fig. 2).

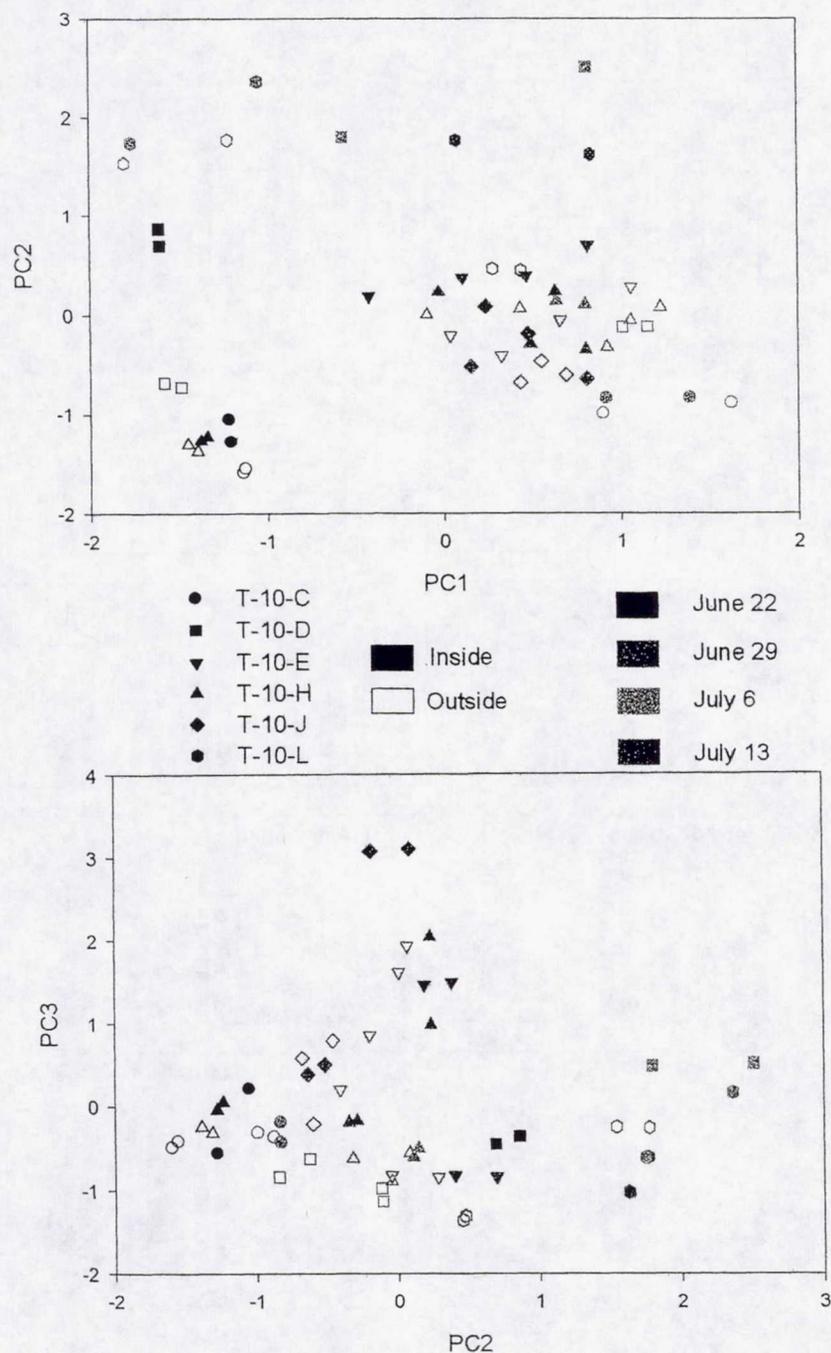
Water samples collected inside the impoundment consistently separated from those collected outside the impoundment. This pattern was confirmed to be statistically significant by one-way ANOVA of inside vs. outside values for salinity, pH, ammonium, bacterial abundance, and chlorophyll-*a* ($\alpha = 0.05$). Values for each of these variables were always lower outside the impoundment than they were inside the impoundment. There were no significant correlations among any of the water quality variables measured ($\alpha = 0.05$).

In addition to the inside vs. outside pattern, the PC2 vs. PC3 comparison revealed that those samples collected on July 13 (blue symbols) separated along PC3 (the nitrogen component) from all other sampling times (Fig. 2 bottom panel). On July 13, we observed a significant increase in the number of dead *Limulus polyphemus* (horseshoe crab) around the impoundment culverts. Horseshoe crab carcasses were abundant both inside and outside the impoundments where samples were collected. Decay of the crab carcasses was the likely cause of the 2- to 12-fold increase in NH_4^+ -N concentrations measured at T-10-E, T-10-H, and T-10-J observed for the July 13 samples.

A second PCA was done without the July 13 data to ascertain if these high N samples were the underlying cause of the patterns obtained when the entire data set was subjected to PCA. The results of the second PCA without the July 13 data were nearly identical to the first analysis, which included the July 13 data (data not shown). The variables comprising PC1, PC2, and PC3

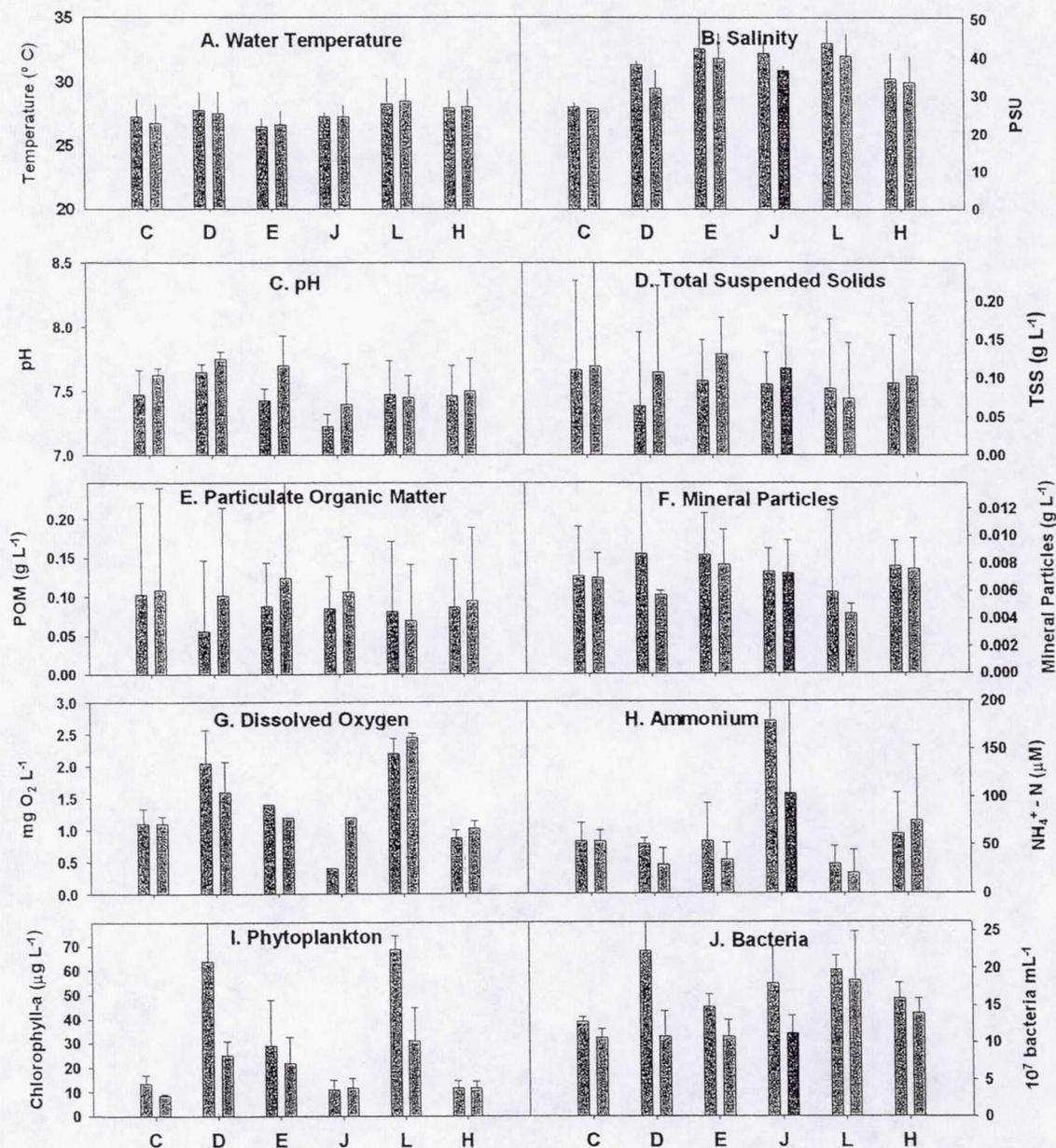
for the second analysis were identical to those in when all the data were analyzed. For the second PCA, the variance explained by the first 3 PCs was even greater (PC1=38.3%, PC2=23.2%, PC3=13.2%, total = 74.7%) than that explained when all the data were used. In contrast to the first PCA, the inside vs. outside separation was the only obvious pattern seen when PC1 vs. PC2 or PC2 vs. PC3 were compared.

*Figure 2. Results of principal component analysis using all data for collected on all dates (top panel) and without data from July 13, 1999.



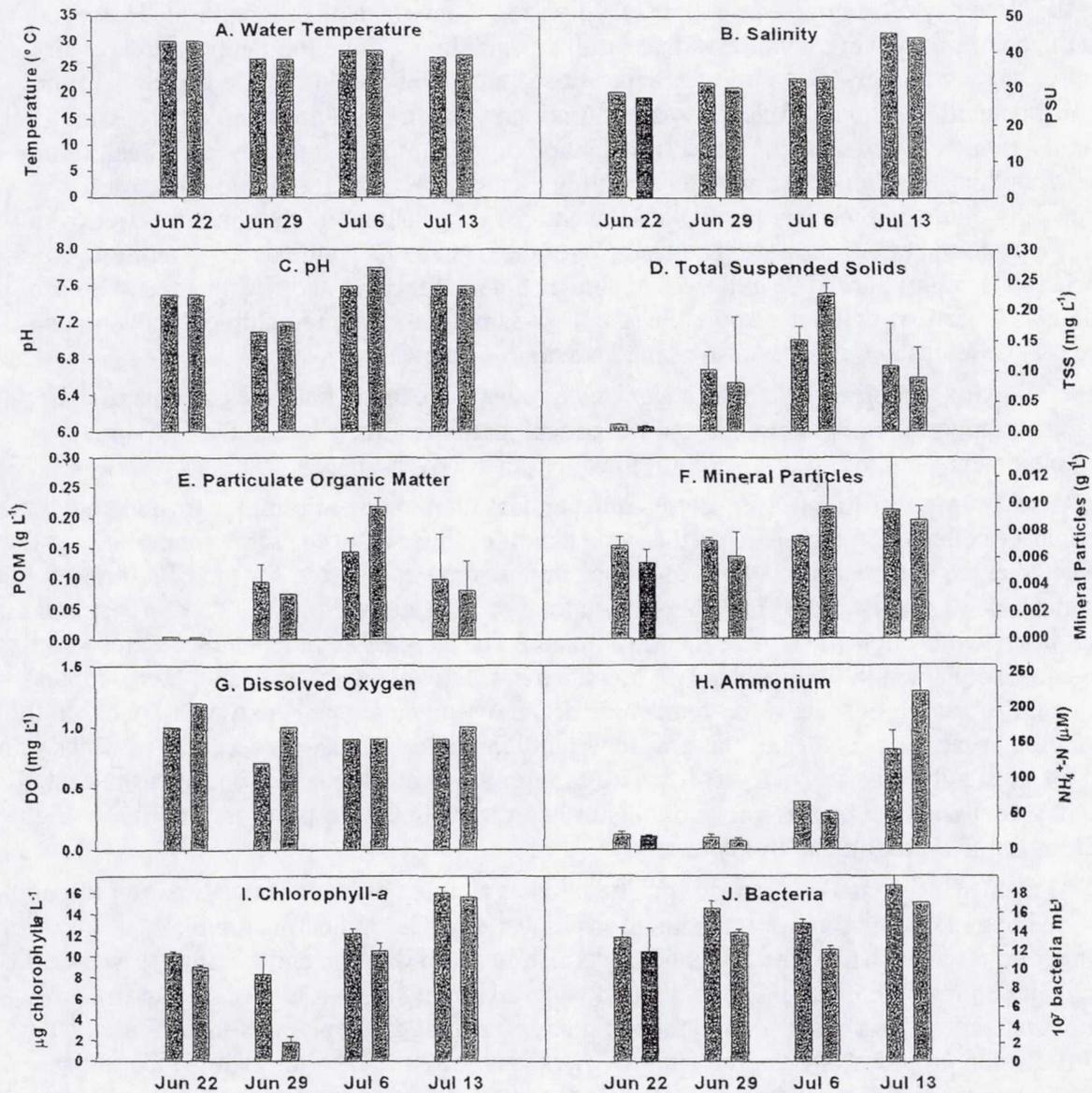
No clear patterns are evident among the impoundments (Fig. 3). This is not surprising given the limited nature of the current data set.

*FIGURE 3. Mean concentrations (μM) of NH_4^+ for all samples collected from each of the impoundments during summer 1999 in panel (a) where $n = 2$ for all impoundments except T-10-H where $n = 4$ and (b) where $n = 2$ for T-10-C, T-10-D, T-10-L, $n = 1$ for T-10-E and T-10-J, and $n = 3$ for T-10-H. Green bars = inside the impoundment. Blue bars = outside the impoundment.



B. Temporal variation among samples: Impoundment T-10-H was the only site where samples were collected each of the weeks of the experiment. Water quality measures were highly variable across the times sampled at T-10-H (Fig. 4). The only consistent pattern during these 4 weeks was that chlorophyll-a, and bacterial abundance were significantly different ($\alpha = 0.05$) and always greater inside the impoundment than outside the impoundment. No significant correlations were detected among the variables measured for impoundment T-10-H.

*FIGURE 4. Temporal variation of water quality indicators for T-10-H on the four sampling dates. Green bars = inside the impoundment. Blue bars = outside the impoundment.



4. DISCUSSION AND RECOMMENDATIONS

The primary objective of the work completed in summer 1999 was to determine an appropriate water quality sampling scheme for work to be conducted with EPA funding in 1999-2001. One of the primary concerns was and is the feasibility of sampling 7 impoundments and the associated seagrass beds in the IRL within a time frame to allow for direct comparison among impoundments. This concern is based on the EPA-funded project's long-term goal to compare the effect of WAM and open water management on the impoundments, which requires knowledge of both the annual flux and the patterns of material (particulate and dissolved) movement between the IRL and the impoundments. An additional goal of the EPA-funded work is to determine what effect material flux from the impoundments has on seagrass beds adjacent to the impoundment culverts. Ideally the impoundments and the seagrass beds should be sampled on the same day to minimize temporal variation for comparison of flux patterns. However, based on the preliminary work completed this summer, sampling 7 impoundments would require approximately 5 hours. Sampling the adjacent seagrass beds would require additional time – probably another 4 hours. An additional limitation is that imposed by sample processing time and the time between collection and stabilization of the samples once they have been returned to the laboratory. Further constraints on sampling include diel variation in biological activity, especially photosynthesis by phytoplankton and SAV in shallow impoundment water, which can have a profound effect on water chemistry over the course of a single day. To minimize diel variation, samples should be collected at similar times of the day and within several hour of one another. Clearly, it will not be possible to collect samples at the impoundment culverts and over the seagrass beds and process those samples within a single day.

The following sampling scheme is suggested as a way to capture both the patterns of materials loading among the impoundments and the annual materials flux between the individual impoundments and the IRL as well as the water quality over adjacent seagrass beds.

A. Weekly Impoundment Sampling: Independent impoundment samples (inside and outside) should be collected in duplicate on the same day of each week, at the same time of day at the selected culverts for each of WAM and open impoundments selected for the EPA-funded study (6 impoundments, described in Table 1) or a total of 24 water samples. T-10-K, a restored marsh that has no culverts, will serve as a control marsh. For this marsh, independent water samples should be collected in duplicate at two locations in shallow lagoon water (4 water samples). It is critical that sampling time, water temperature, and dissolved oxygen content are noted in the field log because the values of these water quality indicators are highly variable on diel cycles. At all sites, but especially at T-10-K where sampling will need to be conducted from a small boat, great care must be taken to avoid disturbing the sediments to prevent abnormally high values for all water quality indicators.

B. Monthly Seagrass Bed Sampling: The relatively large volume of the IRL water as compared to the impoundments suggest that alterations in water quality indicators are likely to occur on a longer time scales than in the impoundments as a result of dilution and mixing. However, D. Scheidt (Blum & Scheidt, 1999) found that water overlying a seagrass bed 150 m from an impoundment culvert responded to the rapid drawdown of the impoundment within one day. The EPA-funded project study impoundments are not subjected to rapid drawdown like those examined by Scheidt where large differences in hydraulic head created rapid mixing between impoundment and lagoon water. Unless rapid draw down of WAM impoundments is planned or an extreme storm tide or rainfall event occurs creating large head differences, monitoring of water quality over seagrass beds can occur less frequently than weekly sampling at impoundment

culverts. A better integrated indicator of the effect of water quality on the seagrass beds will be obtained from monitoring seagrass distribution and production (Virnstein 1990, 1995). Thus, in association with ongoing water quality sampling in the IRL, water overlying seagrass beds near to each of the impoundments and T-10-K should be sampled once each month. A minimum of triplicate, independent grab samples should be collected over each bed along a transect centered on the impoundment culvert (to the extent possible) and be oriented parallel to the impoundment dike (21 water samples). To allow for direct comparison with impoundment sampling, seagrass and impoundment water quality should be sampled within one day of one another.

C. Intensive spatial sampling:

1. Perpendicular to impoundment dikes through the culverts. Work by D. Scheidt (Blum & Scheidt, 1999) found that a transect extending from 50 m inside the impoundment culvert to 150 m into the lagoon showed a detectable gradient in water quality characteristics when the impoundment was closed and a steeper gradient at least 96 h after rapid drawdown. Water was collected on the marsh surface about 50 m from the culvert, in the perimeter ditch and in the lagoon at the culvert mouth, and 150 m from the culvert. In both WAM and open impoundments, perpendicular transects similar to those used by Scheidt should be sampled to determine if there is a gradient in water quality indicators from the marsh surface, to the perimeter ditch, through the culvert, to nearest seagrass bed. Collecting additional samples 50 and 100 m from the culvert into the lagoon is also recommended to better define any gradients that might exist. If independent, duplicate samples are collected at each interval along the transect, a total of 12 water samples per impoundment would result from sampling in this manner. Ideally, WAM-managed impoundments would be sampled prior to culvert closure, one month after closure, just prior to drawdown, and daily during drawdown. Open impoundments should be sampled at times coincident with sampling of the WAM impoundments to determine if gradients occur when water exchange is allowed to occur freely between the impoundment and the lagoon. Similarly, T-10-K should be sampled from the intertidal region to 150-m off-shore.

2. Parallel to impoundment dikes in the perimeter ditch and lagoon. During the preliminary sampling conducted this summer (1999), impoundment (inside) and IRL (outside) water were collected at the ends of culverts installed through impoundment perimeter dikes. Sampling in this way provides the most conservative estimate of differences in the water quality indicators: differences would have been minimized between the impoundment and the lagoon, particularly when water was flowing through the culverts. Greater differences in water quality indicators are likely to occur away from the culverts. To determine if this assumption is true, at least once during the first year of the EPA-funded project each impoundment should be sampled along a transect centered on the culverts and extending perpendicular to the impoundment dikes. Duplicate, independent grab samples should be collected along the lagoon shoreline and in the perimeter ditch at 50-m intervals over a total distance of 400 m (i.e. 200 m in each direction from the culvert). This intensive sampling will generate 108 water samples from the six impounded marshes and 18 water samples from the restored marsh. These transects would be sampled at the same time as the perpendicular transects. Because of the logistical limitations associated with sampling both types of transects, it may be necessary to establish impoundment pairs that would be sampled on the same day. For example, T-10-D might be paired with T-10-J, T-10-E with T-10-H, and T-10-C with T-10-L so that the entire set of impoundments would be sampled over a 3-day period.

4. CONCLUSIONS

Water quality indicators inside and outside the marsh impoundments were different. Ammonium, salinity, bacteria and chlorophyll-*a* were higher inside the impoundments as expected possibly as a result of the great affect of evaporation on impoundment water. Water quality indicators responded rapidly both inside and outside the impoundments as exemplified by the increase in NH_4^+ -N concentrations during a horseshoe crab die-off. Water quality indicators were high variable during the month in which water samples were collected. Because the impoundments are widely spaced it is logistically unrealistic to sample each of the impoundments and associated seagrass beds on a single day, sampling must be stratified to allow patterns of material movement and the annual flux of materials to and from the impoundments to be determined.

** Figures 1-4 of the original document are in color. To obtain a copy of this report in color, please contact Dr. Linda Blum at LKB2e@virginia.edu or (804) 924-0560.*

5. REFERENCES

- Blum, L.K. and D. M. Scheidt. 1999. Impact of marsh impoundment water management on Indian River Lagoon water quality. Abstracts of the 6th Symposium on Biogeochemistry of Wetlands, Ft. Lauderdale, FL, July 11-14.
- Day, J.W., Jr., et. al. 1989. Estuarine ecology. Wiley, New York, 558 p.
- Day, J.W., Jr. and P. H. Templet. 1989. Consequences of sea level rise: implications from the Mississippi Delta. *Coastal Management* 17:241-257.
- Epstein, M.B. 1995. Managing wetlands for multiple objectives: Atlantic Flyway. Pages G1-G35 In W.R. Whitman, et. al. (Eds.). *Waterfowl habitat restoration, enhancement and management in the Atlantic Flyway*, 3rd Ed. Environmental Management Committee, Atlantic Flyway Council, Tech. Sect., and Del. Div. Fish and Wildl., PO Box 1401, Dover, DE.
- Gilmore, G.R., Jr. 1999. Wetland ecosystem management: Indian River Lagoon, Florida, USA. A comprehensive review of the Indian River Lagoon wetland ecosystems and human influence on these systems. Final Rep. Natl. Estuarine Prog. And St. John's River Water Man. Dist. Contract No. 98W230.
- Leenhouts, W.P. and J. L. Baker. 1982. Vegetation dynamics in dusky seaside sparrow habitat on Merritt Island National Wildlife Refuge. *Wildl. Soc. Bull.* 10:127-132.
- Nixon, S.W. 1980. Between coastal marshes and coastal waters – review of twenty years of speculation and research on the role of salt marshes in estuarine productivity and water chemistry. p. 437-525. In. Hamilton, P. and MacDonald, K.B. (eds.). *Estuarine and Wetland Processes*, Plenum, New York, NY.
- Virnstien, R.W. 1990. The large spatial and temporal biological variability of Indian River Lagoon. *Florida Scientist* 53:249-256.
- Virnstien, R.W. 1995. Seagrass landscape diversity in the Indian River Lagoon: the importance of geographic scale and pattern. *Bull. Mar. Sci.* 57:64-74.

1999 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

**JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA**

**CHARACTERIZATION OF MOLYBDATE CONVERSION COATINGS FOR
ALUMINUM ALLOYS BY ELECTROCHEMICAL IMPEDANCE SPECTROSCOPY**

Luz Marina Calle
Professor
Chemistry Department
Randolph-Macon Woman's College
NASA Colleague: Louis G. MacDowell

ABSTRACT

Electrochemical impedance spectroscopy (EIS) was used to investigate the corrosion inhibiting properties of newly developed proprietary molybdate conversion coatings on aluminum alloy 2024-T3 under immersion in aerated 5% (w/w) NaCl. Corrosion potential and EIS measurements were gathered for six formulations of the coating at several immersion times for two weeks. Nyquist as well as Bode plots of the data were obtained. The conversion-coated alloy panels showed an increase in the corrosion potential during the first 24 hours of immersion that later subsided and approached a steady value. Corrosion potential measurements indicated that formulations A, D, and F exhibit a protective effect on aluminum 2024-T3. The EIS spectra of the conversion-coated alloy were characterized by an impedance that is higher than the impedance of the bare alloy at all the immersion times. The low frequency impedance, Z_{lf} (determined from the value at 0.05 Hz) for the conversion-coated alloy was higher at all the immersion times than that of the bare panel. This indicates improvement of corrosion resistance with addition of the molybdate conversion coating. Scanning electron microscopy (SEM) revealed the presence of cracks in the coating and the presence of cubic crystals believed to be calcium carbonate. Energy dispersive spectroscopy (EDS) of the test panels revealed the presence of high levels of aluminum, oxygen, and calcium but did not detect the presence of molybdenum on the test panels. X-ray photoelectron spectroscopy (XPS) indicated the presence of less than 0.01 atomic percent molybdenum on the surface of the coating.

ACKNOWLEDGEMENT

I would like to express my deepest appreciation to Louis MacDowell of the Materials Science Division and to NASA/ASEE for providing me with the opportunity to participate in the 1999 Summer Faculty Fellowship Program. My professional development and my appreciation for the space program have been enhanced considerably during the ten weeks of the fellowship. Gregg Buckingham, Dr. E. Ramon Hosler, Dr. Jane Hodges, and Ms. Judie Gilliam provided the leadership and support that makes this program such a rich and wonderful experience for the faculty fellows. Jane, with her enthusiasm and interest in the program, provided us with unique opportunities to learn about NASA and the Kennedy Space Center. Ray and Judie were always there for us when we needed them.

I would also like to acknowledge Claudia Amorim, a student intern from Brazil, for performing the EIS measurements. Her professionalism and enthusiasm were an asset to this project. Scott Murray, Jennie Cummings, Peter Marciniak, Patrick Faughnan, Steve McDanel, and Don Parker in the Materials Engineering Section of the Materials Science Division were always there to help and provided me with a warm and friendly environment during long hours of research. Peter contributed with his photographic skills to the project. Jennie contributed her expertise in SEM and EDS. Stan Young of the Analytical Chemistry section also performed SEM and EDS on the exposed areas of the test panels. Dr. Mike Hampton of the University of Central Florida performed the XPS analysis.

CHARACTERIZATION OF MOLYBDATE CONVERSION COATINGS FOR ALUMINUM ALLOYS BY ELECTROCHEMICAL IMPEDANCE SPECTROSCOPY

Luz Marina Calle

1. INTRODUCTION

Chromate conversion coatings have been used for the protection of aluminum alloys for over 70 years. Although their efficiency in minimizing corrosion attack is excellent, there are health and safety concerns over their use due to their toxicity and carcinogenic nature. Despite an extensive research effort over the past decade, a completely satisfactory replacement for chromate conversions coatings has yet to be identified.

Hexavalent molybdenum has received attention recently as a replacement for chromate conversion coatings. Molybdate mimics chromate in a variety of applications but it exhibits significantly lower toxicity.¹ Several proprietary molybdate conversion coatings have been developed for NASA. Six proprietary formulations of the coatings on aluminum alloy 2024-T3 were available for characterization by Electrochemical Impedance Spectroscopy (EIS).

2. EXPERIMENTAL PROCEDURE

Test Samples

Six proprietary formulations (labeled A-F) were provided by Lynntech, Inc. (College Station, Texas). Panels of aluminum 2024-T3 alloy (with average composition of 4.4% Cu, 0.6% Mn, 1.5% Mg, and 93.5%) were cut to a size of 10.16 cm × 15.24 cm (4" × 6") and treated with a formulation of the molybdenum conversion coating.

Corrosion Potential Measurements

Corrosion potential measurements were performed using a system manufactured by EG&G Princeton Applied Research Corporation. The system used includes: (1) the Model 273 Computer-Controlled Potentiostat/Galvanostat, (2) the Model 5210 Computer-Controlled Lock-In Amplifier, and (3) the Model 352 SoftCorr™ III Corrosion Measurement Software. The electrochemical cell included a saturated Calomel electrode, SCE, a platinum counter electrode, the sample working electrode, and a bubbler/vent tube. Three electrochemical cells were used. The flat specimen holders exposed a surface area of 1 cm², 13 cm², or 32 cm² depending on the cell that was used. Corrosion potential values were gathered for one hour in aerated 5% (w/w) NaCl before the first set of EIS measurements was obtained. Subsequent corrosion potential values were collected before each set of EIS measurements which were collected at several time intervals up to a maximum of 168 hours. All solutions were prepared using deionized water.

Electrochemical Impedance Measurements

A Model 378 Electrochemical Impedance system manufactured by EG&G Princeton Applied Research Corporation was used for all EIS measurements. The system includes: (1) the Model 273 Computer-Controlled Potentiostat/Galvanostat, (2) the Model 5210 Computer-Controlled Lock-In Amplifier, and (3) the Power Sine Electrochemical Impedance Software. Data were gathered in the frequency range from 100 kHz to 0.01 Hz. The AC amplitude used for the experiments was 10 mV. Each sample was studied at various immersion times in aerated 5% (w/w) NaCl for up to one week.

Bode magnitude plots of the data (showing the logarithm of the modulus of the impedance, $\log |Z|$, as a function of the logarithm of frequency and phase angle, α in degrees, as a function of the logarithm of frequency) as well as Nyquist plots (showing the negative of the imaginary component of impedance as a function of the real component of the impedance) were obtained for each formulation.

Scanning Electron Microscopy (SEM)/Energy Dispersive Spectroscopy (EDS)

SEM examinations were carried out on a JEOL JSM-6100 scanning electron microscope at 5.0 to 25.0 kV. Phase identification was done using EDS in SEM with a Kevex SI(LI) detector.

X-ray Photoelectron Spectroscopy (XPS)

Samples were analyzed with a Kratos XSAM800 X-ray photoelectron spectrometer with argon etching between analyses (for a total of 40 minutes).

3. RESULTS AND DISCUSSION

Corrosion Potential Measurements

Corrosion potential values gathered during the hour immediately preceding the first set of EIS measurements for each molybdate formulation indicated that an hour was sufficient to allow the potential to equilibrate. Subsequent values of the corrosion potential were obtained immediately before each set of EIS measurements was collected. Figure 1 shows the change in corrosion potential with immersion time for all formulations of the molybdate conversion-coated panels and for a bare aluminum alloy panel. The coated panels showed a steady increase in the corrosion potential towards the corrosion potential of the aluminum alloy during the first 24 hours of immersion. A gradual increase in corrosion potential with immersion time is expected if the coating encourages slow but uniform (i.e., pit free) growth of the oxide layer. This behavior was observed for all formulations. The initial increase in the corrosion potential subsided and after 48 hours of immersion a steady value of the corrosion potential was obtained. Panels coated with formulations A, D, and F achieved a corrosion potential slightly higher than the corrosion potential of the aluminum alloy while panels B, C, and E achieved a corrosion

potential that was lower. A shift of the corrosion potential of the alloy in the positive direction is indicative of the presence of a protective effect on the aluminum alloy. Visual inspection of the panels revealed that panels A,D, and F had the fewest number of corrosion spots per cm².

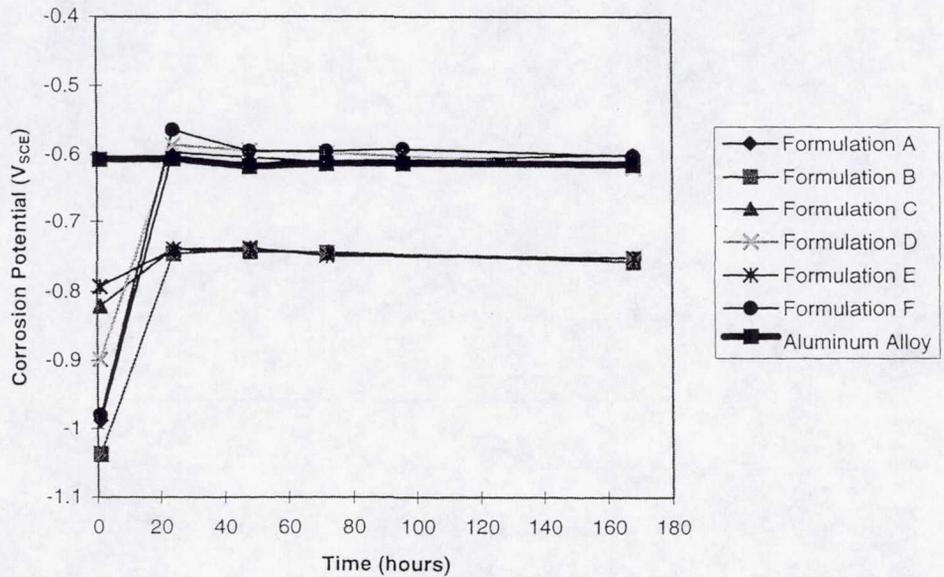


Figure 1. Corrosion potential as a function of immersion time in 5% (w/w) NaCl

Electrochemical Impedance Spectra

Figure 2 shows representative impedance spectra (Bode format: The phase angles have been omitted for clarity) for formulation A and for aluminum 2024-T3 at various immersion times. The impedance of the coated alloy is greater than the impedance of the bare alloy at all frequencies. The higher value of the impedance is indicative of the improvement of corrosion resistance with the addition of the molybdate conversion coating.

The low-frequency impedance, Z_{lf} (determined from data at 0.05 Hz), has been proposed as the optimal EIS parameter to evaluate the performance of corrosion inhibiting coatings on aluminum.² The impedance at this frequency includes the response of the coating as well as part of the response of the oxide and/or corrosion product in the pores at the metal interface. Z_{lf} reflects the condition of the substrate as well as that of the coating. Although the Z_{lf} values do not give information about how a coating degrades (i.e., water uptake, blistering, etc.), they correlate with coating performance. Figure 3 shows the variation of Z_{lf} with immersion time for the molybdate-coated as well as the bare aluminum alloy. Z_{lf} values for the molybdate coated

panels are in the order of 10^3 - $10^5 \Omega\text{cm}^2$ while those for the bare panel are of the order of $10^3 \Omega\text{cm}^2$. Low frequency impedance for chromate conversion coatings on aluminum alloy 2024 after 3 hours of immersion in aerated 0.5 M (3%) NaCl have been reported³ as being in the order of $10^4 \Omega\text{cm}^2$. This value is lower than $10^{10} \Omega\text{cm}^2$ which is the value expected of an excellent barrier coating.⁴

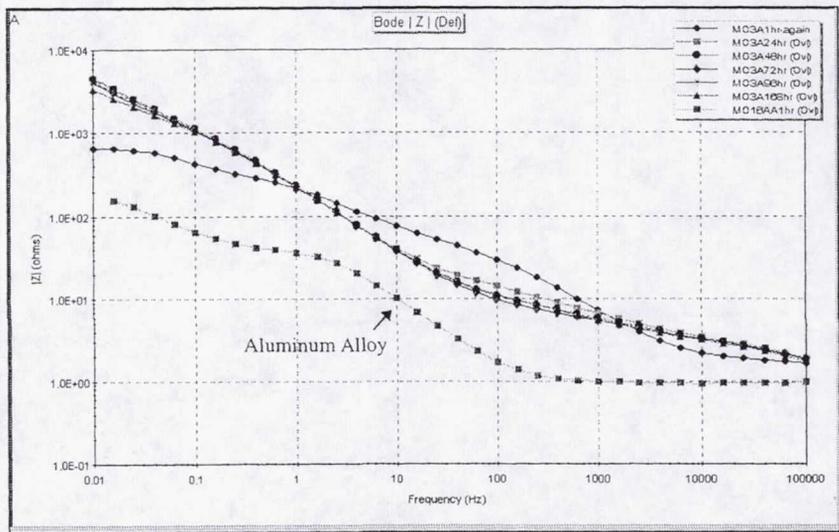


Figure 2. Bode plots for formulation A and aluminum alloy at different immersion times

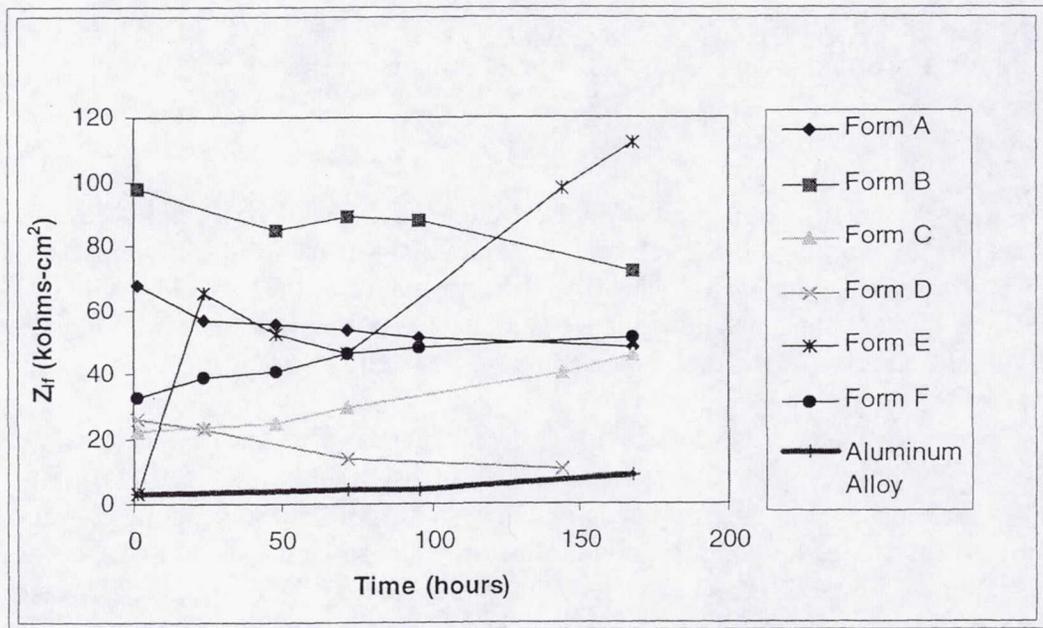


Figure 3. Low frequency impedance Z_{lf} as a function of immersion time in NaCl 5% (w/w) NaCl

Impedance data in the Nyquist format were analyzed using the simplified equivalent circuit shown in Figure 4. This circuit gives a coating resistance (R_c) which can be used as a figure of

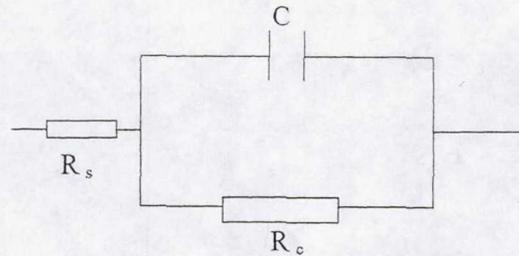


Figure 4. Simplified circuit for conversion-coated aluminum alloys⁵

merit to assess coating performance. Values for R_c for each conversion-coated and for the bare panel are shown in Figure 5. The conversion-coated panels achieved R_c values in the 10^4 - 10^5 Ωcm^2 range while the bare aluminum alloy panel had R_c values in the order of 10^3 Ωcm^2 .

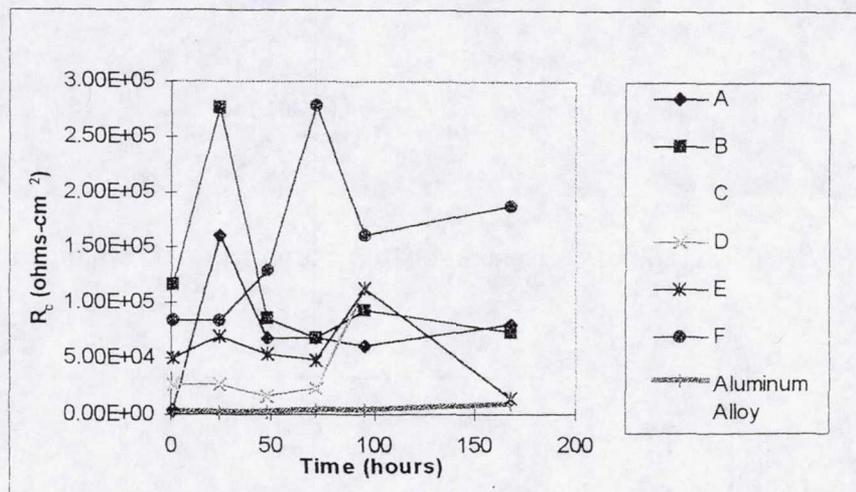


Figure 5. Apparent integrated resistance as a function of immersion time

The threshold R_c value⁵ proposed to define the minimum value for which a given coating can be expected to attain a passing result in a 168-hour salt spray test for aluminum 2024-T3 is 2×10^6 Ωcm^2 to 5×10^6 Ωcm^2 . None of the conversion-coated panels attained these values despite their successful performance in salt fog testing. However, the reliability of salt spray results as well as the use of EIS data for prediction of salt spray performance has been challenged by Mansfeld.⁶

SEM/EDS

SEM of the of one of the conversion coated panels (formulation A) revealed the presence of cracks on the coating as well as the presence of cubic crystals (see Figure 6). EDS analysis of the crystals indicated the presence of high concentrations of calcium as well as oxygen (see Figure 7).

This information can be interpreted as evidence of the presence of calcium carbonate on the coating.

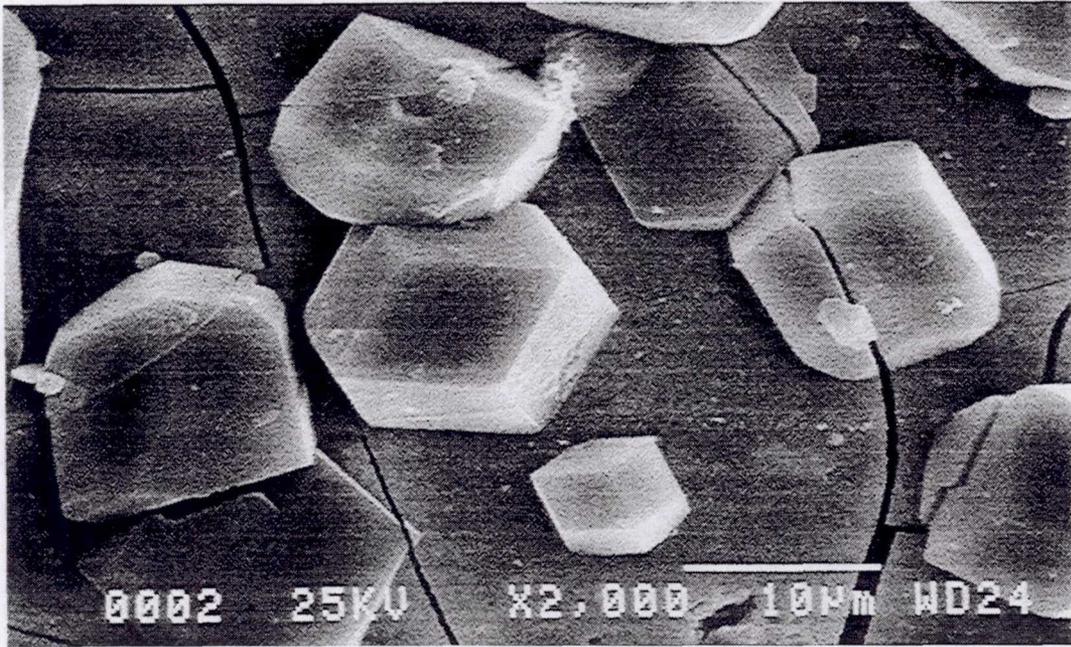


Figure 6. SEM micrograph of conversion-coated (formulation A) aluminum alloy

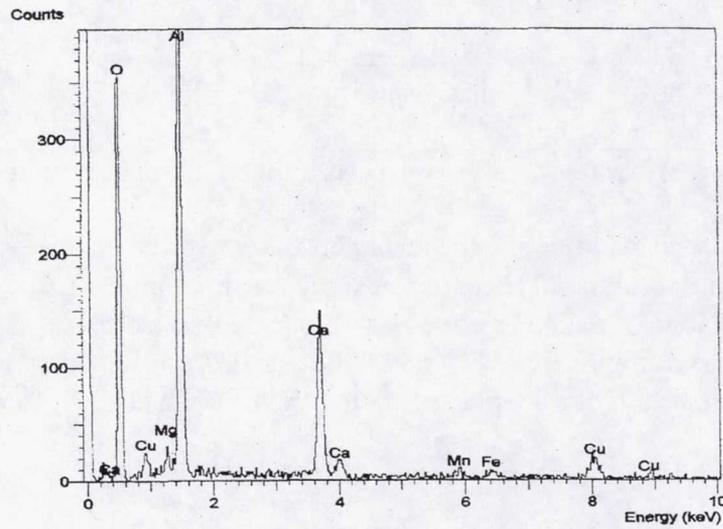


Figure 7. EDS spectrum of conversion-coated (formulation A) aluminum alloy

EDS analysis did not detect the presence of molybdenum in the coating. This finding can be attributed to the fact that EDS is sensitive to most elements if they are present above 0.4% by weight.

XPS

The conversion-coating was further analyzed by x-ray photoelectron spectroscopy (XPS) with argon etching between analyses. The sample showed a small peak in the XPS spectrum that corresponded with the expected location of a molybdenum peak. This would amount to less than 0.01 atomic percent molybdenum on the surface of the sample.

4. CONCLUSIONS

Corrosion potential measurements indicated that formulations A, D, and F exhibit a protective effect on aluminum 2024-T3.

EIS results for the six formulations tested showed that there is an improvement in the corrosion resistance of aluminum 2024-T3 with the addition of the molybdate conversion coating.

SEM/EDS revealed the presence of cracks in the coating. Cubic crystals of what it is believed to be calcium carbonate were also observed on the surface of the coating. SEM/EDS failed to detect the presence of molybdenum in the coating.

XPS analysis of the coating indicated that there is less than 0.01 atomic percent molybdenum on the surface of the sample.

Further analysis of the EIS data as well as data gathered from atmospheric exposure of the panels are suggested before more definite conclusions can be drawn.

REFERENCES

-
- [1] B.A. Shaw, G.D. Davis, T.L. Fritz, and K.A. Oliver, *J. electrochem. Soc.*, 137 (1990): p. 359.
 - [2] J.A. Grandle and R.S. Taylor, *Corrosion*, 50 (1994): p. 792.
 - [3] J.E.O. Mayne, and D.J. Mills, *J. Oil & Colour Chemists Assoc.*, 58 (1975) p. 155.
 - [4] R. G. Buchheit, M.D. Bode, and G.E. Stoner, *Corrosion*, 50 (1994): p. 205.
 - [5] R.G. Buchheit, M. Cunningham, H. Jensen, M.W. Kendig, and M.A. Martinez, *Corrosion*, 54 (1998): p. 61.
 - [6] F. Mansfeld, *Corrosion*, 54 (1998): p. 595.

1999 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA

Performance Evaluation of the NASA/KSC Transmission System

Kenneth J. Christensen
Assistant Professor
Department of Computer Science and Engineering
University of South Florida

KSC Colleagues: John Schnitzius (NASA) and Chris Kerios (Dynacs Engineering)

ABSTRACT

NASA-KSC currently uses three bridged 100-Mbps FDDI segments as its backbone for data traffic. The FDDI Transmission System (FTXS) connects the KSC industrial area, KSC launch complex 39 area, and the Cape Canaveral Air Force Station. The report presents a performance modeling study of the FTXS and the proposed ATM Transmission System (ATXS). The focus of the study is on performance of MPEG video transmission on these networks. Commercial modeling tools - the CACI Predictor and Comnet tools - were used. In addition, custom software tools were developed to characterize conversation pairs in Sniffer trace (capture) files to use as input to these tools. A baseline study of both non-launch and launch day data traffic on the FTXS is presented. MPEG-1 and MPEG-2 video traffic was characterized and the shaping of it evaluated. It is shown that the characteristics of a video stream has a direct effect on its performance in a network. It is also shown that shaping of video streams is necessary to prevent overflow losses and resulting poor video quality. The developed models can be used to predict when the existing FTXS will "run out of room" and for optimizing the parameters of ATM links used for transmission of MPEG video. Future work with these models can provide useful input and validation to set-top box projects within the Advanced Networks Development group in NASA-KSC Development Engineering.

Performance Evaluation of the NASA-KSC Transmission System

Kenneth J. Christensen

1. Introduction

NASA-KSC currently uses three bridged 100-Mbps FDDI segments as its main transmission system, or backbone, for data traffic. The FDDI Transmission System (FTXS) connects the KSC industrial area, KSC launch complex 39 and the Cape Canaveral Air Force Station. NASA-KSC is in the middle of developing a consolidated networking infrastructure based on Asynchronous Transfer Mode (ATM) as an upgrade to the FTXS. This new transmission system is the ATM Transmission System, or ATXS. The KSC Development Engineering (DE) directorate is accountable for this project which is described in [6]. Within the DE directorate, the Advanced Networks Development (AND) group is examining new networking technologies and developing the design specification. The goal of the consolidated networking infrastructure is to reduce the number of disparate networks at KSC and thus achieve cost savings by 1) increased efficiency in the use of the existing in-ground cabling plant, and 2) reduced operations and maintenance requirements. Current transmission of shuttle launch video is via the KSC WideBand and BCDS analog (cable TV) networks. A goal of the shuttle upgrade program is to generate MPEG-2 encoded video and transmit this video via the ATXS.

A key component in the planning, design, and deployment of any large network is performance evaluation, or capacity planning. The goal of capacity planning is to deliver the most cost-effective network designs while ensuring that performance levels will meet specified service requirements (see [8]). While very valuable, building experimental test networks in a laboratory setting is expensive and limited in size and scope. Due to its size and complexity, the entire KSC transmission system cannot be duplicated in a laboratory. Thus, a performance model of the existing and proposed systems needs to be developed and experiments conducted on the model. The results from the experiments can give guidance on the amount of "head room" available on the existing FTXS and the anticipated performance of possible ATM upgrade designs.

During the summer of 1999 (funded as a second-year NASA/ASEE faculty fellowship) a capacity planning study of FTXS and ATXS was conducted. The CACI suite of modeling tools - Predictor and Comnet (see [1]) - was used for performance evaluation. The purchase of the CACI tools was the direct result of a requirements definition and procurement completed in the summer of 1998 (see [2]). The goals of this capacity planning study were:

1. Baseline the topology and data traffic of the existing FTXS
2. Build models of the FTXS and proposed ATXS including video transmission subsystems
3. Collect or acquire representative MPEG video traffic traces and characterize them
4. Conduct experiments on the FTXS and ATXS models to determine:
 - Expected lifetime (or "headroom") of the systems given anticipated traffic growth rates
 - Predicted performance of video traffic
5. Evaluate shaping of video traffic as a means of improving the efficiency of use of network resources

The key deliverables from this study were:

1. CACI Predictor models for the FTXS and ATXS
2. Tools to characterize and import traffic into the Predictor models
3. CACI Comnet models for video transmission on FDDI and ATM networks
4. Characterization of some packetized MPEG-1 and MPEG-2 video clips

These deliverables forward the goal of consolidating video services on existing and future networks at NASA-KSC.

The remainder of this report is organized as follow. Section 2 is a brief overview and motivation for capacity planning. Section 3 is an overview of the FTXS and ATXS network designs. Section 4 describes characterization of MPEG video streams. Section 5 describes the CACI Predictor models and results for FTXS and ATXS based on a "normal day" baseline traffic. Section 6 presents a traffic characterization from the July 22nd scrubbed shuttle launch day. Section 7 describes the CACI Comnet models for video transmission on ATM networks and presents the results of a study of shaping of video traffic. Finally, section 8 is a summary and describes future work.

2. Overview and Motivation for Capacity Planning

Capacity planning of computer systems is the process of predicting the future performance of a system. Specifically, capacity planning intends to predict when performance will drop below a desired level of service. For any moderately large or complex system it is not possible to physically build and test design options, hence modeling is required. Capacity planning uses both simulation and analytical modeling methods. Analytical methods use "math formulas" to model and solve for the performance measures of a system. Analytical methods are rooted in queueing theory and require simplifying assumptions to build tractable models. Reference [5] is the classic reference on queueing theory. Simulation methods use discrete event simulation programs executed on a computer to find statistical approximations for the performance measures of a system. Reference [7] is the classic reference on discrete event simulation. Simulation models require much longer execution time than analytical models. However, simulation models do not require many of the simplifying assumptions of analytical models. Another trade-off is correctness, given that the underlying assumptions are accepted, an analytical model can be formally proved correct. However, a simulation model consists of thousands to millions of lines of code and is hence almost certain to contain "bugs", some of which may affect the accuracy of the output. The accuracy of output (i.e., the results) for any model is no better than the accuracy of the input. Given first order approximations for input, the output of both an analytical and simulation model will be poor. This last fact is very important, if many assumptions must be made on the input to a model then faster analytical modeling may be better than simulation modeling since one can study many design options with an analytical model in the same amount of time that only one option can be studied with simulation methods.

Computer networks are modeled as queueing systems. The buffers in network components such as routers and switches are packet (or cell for ATM) queues. Delay in a network occurs when packets are forced to wait for transmission in these buffers. Packet loss occurs when these buffers overflow due to packet arrival rates temporarily exceeding the service rate. Packets arrive with probabilistically distributed interarrival times, often in bursts. A packet is serviced when it is transmitted on an outgoing link. Packet service time is a function of the packet size and the data rate of the outgoing link. For example, a 1500 byte packet has a 1.2 millisecond service time on a 10-Mbps Ethernet. Both packet interarrival times and packet sizes can be modeled with probability distributions. Packet streams are modeled and analyzed as time series and their characteristics determined by time series analysis.

Burstiness can be informally defined as how much the arriving packets are clumped together. A more formal definition of the burstiness of a packet arrival stream is the Coefficient of Variation (CoV) of the packet interarrival times (IAT),

$$\text{CoV of packet IAT} = \frac{\text{Standard deviation of packet IAT}}{\text{Mean of packet IAT}}, \quad (1)$$

where the mean of packet IAT is simply the reciprocal of the packet arrival rate. The utilization of a network is the ratio of measured arrival rate divided by the maximum rate ("bandwidth") of the link. Determining the variance of packet interarrival times requires analyzing packet traces of network traffic. Packet traces contain time-stamps for received packets. It is from these time-stamps that the variance of packet interarrival times is determined. For deterministic traffic the CoV = 0, for random (Poisson) traffic the CoV = 1, and an increasing burstiness is indicated with a higher value of the CoV. As the arrival process increases in burstiness, the saturation knee occurs at lower utilization levels. The knee effect is common to all queueing systems. The challenge of capacity planning is to determine at what utilization level the knee will occur.

3. Overview of the FTXS and ATXS Designs

The current data communications infrastructure of NASA-KSC is the FTXS (see [6]). The underlying technology of the FTXS is FDDI, a 100-Mbps shared-medium token ring technology. FDDI is an ANSI standard (see [3]) developed for reliable and efficient data communications based on variable length packets. In an FDDI LAN, all of the attached hosts share the common bandwidth of a 100-Mbps fiber optic ring. In FTXS, three FDDI rings are interconnected via bridges. Organizations, often identified by a KSC building, attach to the FTXS via a router. Figure 1 shows a graphic of the FTXS taken from a Predictor model (the model is described in Section 4). The ATXS (see [6]) is a complementary network to the FTXS with a plan to consolidate some, or all, of the FTXS traffic. The ATXS is based on ATM technology with standards largely set by the ATM Forum (see [2]). The

ATXS is especially intended to carry video traffic that is currently carried on the separate analog BCDS and WideBand systems. The purpose of this study is to evaluate the ATXS for this video traffic consolidation. Figure 2 shows the ATXS network (as a Predictor graphic).

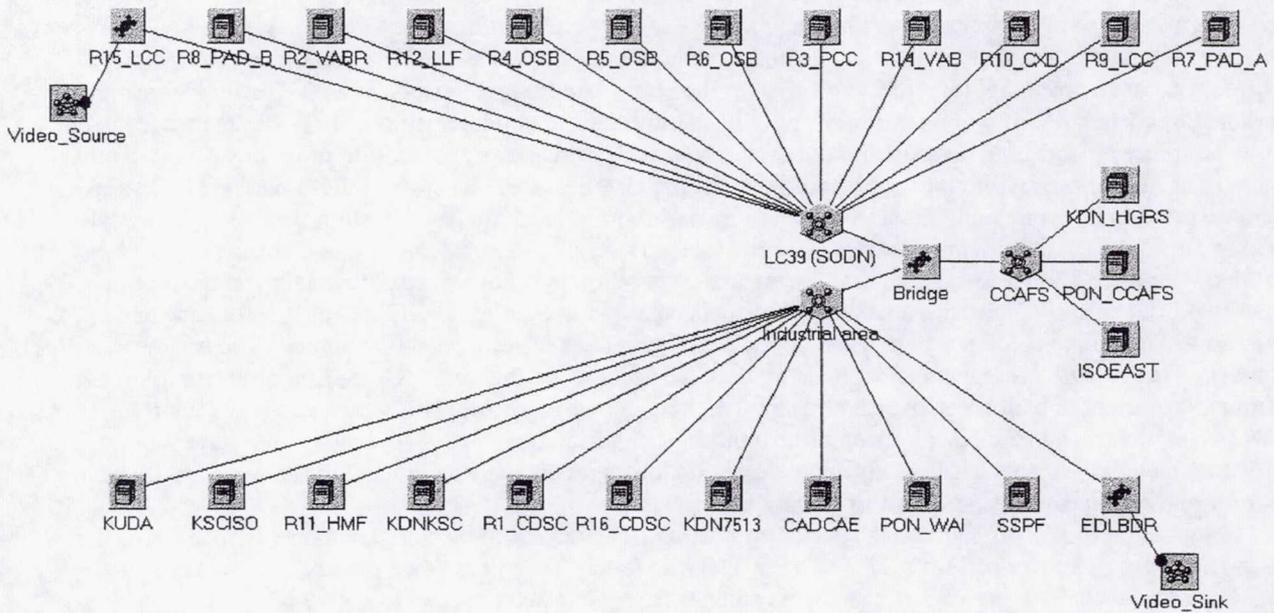


Figure 1 - Predictor graphic of the FTXS

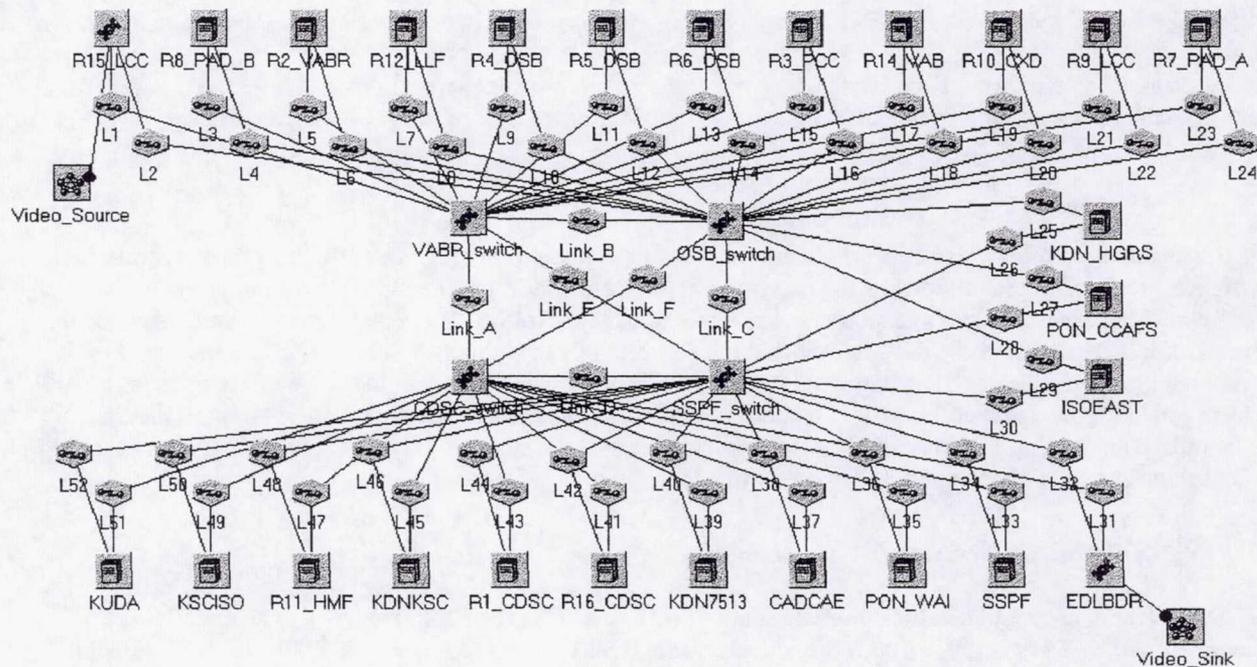


Figure 2 - Predictor graphic of the ATXS

4. Characterization of MPEG Video

MPEG-1 and MPEG-2 are video encoding standards used for digital video, see [9]. A video encoder generates a stream of compressed frames that can be stored as digital data (e.g., as a file on a hard disk) or sent across a network for video transmission. A matching decoder uncompresses the encoded frames and generates a "video out" signal as NTSC or composite video. The encoding and decoding incurs a delay - the amount of delay depends on the implementation of the encoding and decoding hardware and/or software. Additional delay may be incurred in a playout buffer used to compensate for delay variance (called "delay jitter") caused by the variable queuing delays in the network. For video transmission on a network, MPEG frames are packetized into one, or more, packets per frame. MPEG defines three types of frames, I (interpolated), P (predictive), and B (bi-directional). An I frame is an independently compressed video frame, P and B frames are inter-frame compressed in reference to the previous I frame (and to each other). I frames are typically generated every half-second (with 14 P and B frames between each I frame for 30 frames per second video). The loss of an I frame prevents the subsequent B and P frames from decoding correctly and typically results in a half-second "freeze" in the video. The loss of a B or P frame has less effect on the video quality. Since I frames are independently compressed, they are typically larger in size than B and P frames making their loss more likely in a congested network. In addition, for most video encoders the size of the frames is highly dependent on the complexity and amount of motion in the scene on which the camera is focused. For example, a talking-head in front of a solid color background will encode into very small frames (and hence low mean and variance in bandwidth required to transmit the frames on a network), an automobile race with frequent scene changes will result in the opposite. For most encoders, the compression rate is a settable parameter with a higher compression ratio resulting in a lower transmission bandwidth requirement (or less storage), but lower video quality. For MPEG-2 encoding, an average rate of about 6-Mbps results in VCR quality video. Constant bit-rate, variable video quality encoding is also possible where the output rate of the encoder is constant, but video quality varies as a function of the amount scene complexity and motion. Figure 3 shows a video source, encoder, network, and decoder (and assumes one video frame per packet). Note that video frames entering the network are deterministically spaced (typically one every 1/30-th of a second for 30 frames per second video) and that this same timing must exist at the decoder if the decoder is not to be over-run or under-run. An end-to-end network delay of a few milliseconds per packet is considered to be acceptable to most video applications.

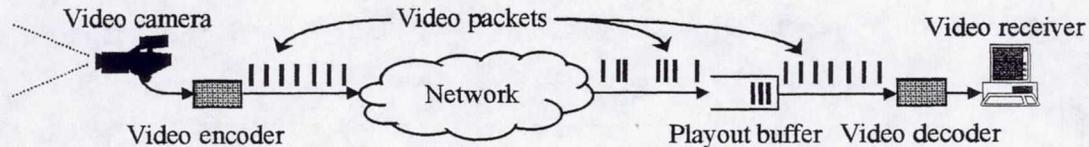


Figure 3 - Digital video on a network - encoding, playout buffering for jitter removal, and decoding

Understanding the characteristics of digital video is necessary in order to predict the network resources required to transmit the video. Significant research has been done in the area of video characterization and traffic modeling. The results in [11] show that the characteristics of a video stream are highly dependent on the content of the video. At the time of this work, there were no MPEG-2 encoders available at KSC. Thus, to study the time series characteristics of MPEG-2 video streams, traces of MPEG-2 encoded 1996 Olympic Games events were obtained (see [13]). Figures 4, 5, and 6 show the bandwidth requirements for 30 seconds of MPEG-2 encoded Olympic Games waterpolo. Increasing the time increment aggregates the time series. At a 30 second increment, the mean bandwidth requirement is 5.94-Mbps. At the smaller time increments the large spikes (one every half-second) correspond to the I-frames, with a peak bandwidth of over 15-Mbps. At the larger time increments, the spikes are not visible, this can result in under-allocation of network bandwidth for video transmission.

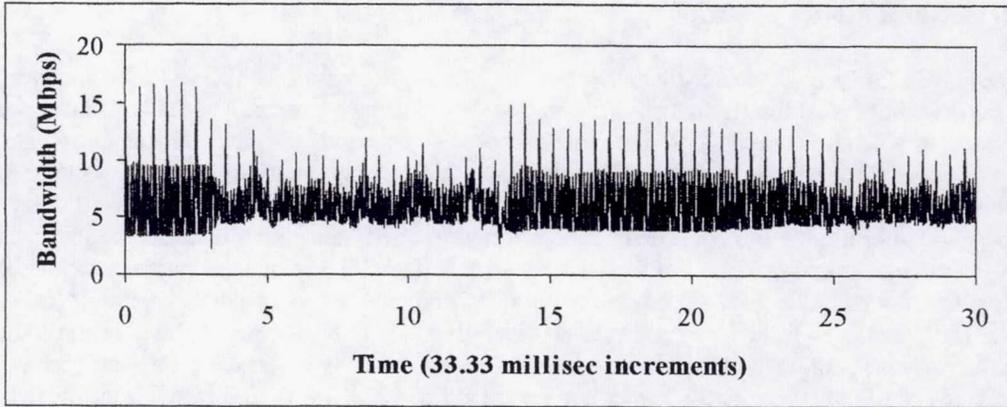


Figure 4 - Bandwidth for 30 seconds of MPEG-2 Olympic games waterpolo (33.33 millisecond increment)

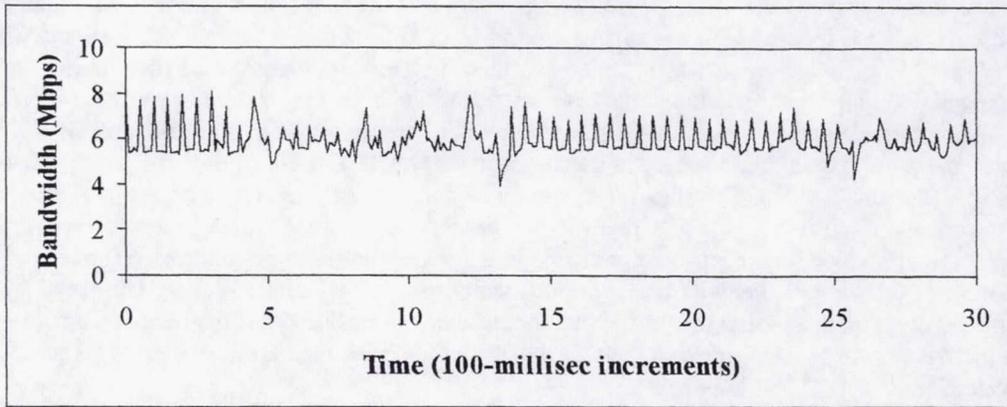


Figure 5 - Bandwidth for 30 seconds of MPEG-2 Olympic games waterpolo (100 millisecond increment)

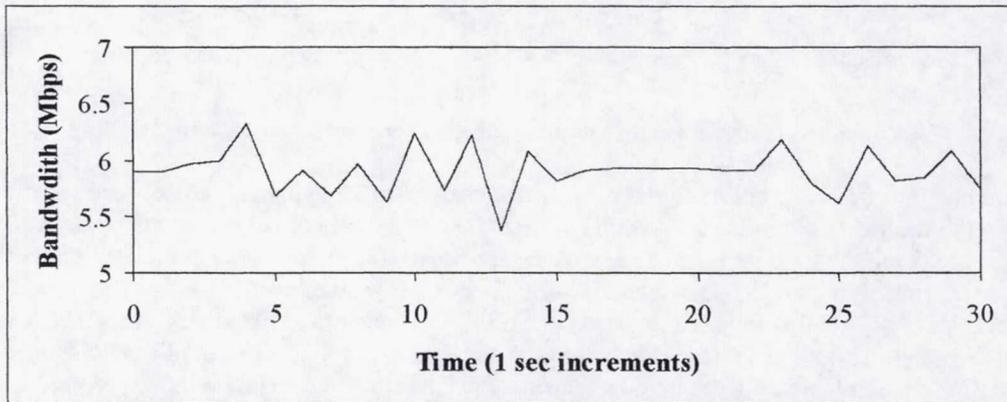


Figure 6 - Bandwidth for 30 seconds of MPEG-2 Olympic games waterpolo (1 second increment)

Table 1 summarizes two key characteristics of the set of Olympic Games MPEG-2 frame size traces. The mean bandwidth required for a video stream is directly proportional to the mean frame size. The burstiness of the video stream is directly proportional to the Coefficient of Variation (CoV) of frame size, which is the ratio of the standard deviation of frame size to the mean of frame size (similar to equation (1) for CoV of packet IAT). It can

be seen in Table 1 that the mean frame size varies very little between video clips (ranging from 24200 to 24552 bytes, a 1.5% difference), however the CoV varies by 41% (ranging from 0.4858 to 0.6861).

Sporting event	Full video clip (40 minutes)		First 10 seconds of video clip	
	Mean frame	CoV of frame	Mean frame	CoV of frame
Archery	24498 bytes	0.6816	24203 bytes	0.6584
Badminton	24200	0.6861	24252	0.6398
Boxing	24504	0.5592	24500	0.6005
Diving	24438	0.5748	24457	0.6651
Sailing	24552	0.4858	24612	0.5145
Swimming	24498	0.6551	24565	0.5855
Table tennis	24363	0.6409	24522	0.5329
Waterpolo	24507	0.5058	24558	0.4617

Table 1 - Key characteristics of Olympic Games MPEG-2 frame traces

5. Predictor Modeling of the FTXS and ATXS

A CACI Predictor model can give a first order measure of the capability of the existing FTXS and the future ATXS for transporting video streams. A layer-2 model of the FTXS was manually built as shown in Figure 1. Each of the routers attaching to the FTXS is modeled as a traffic source. A video source is modeled in the LCC building and a video receiver in the EDL building. The FTXS is modified to model the ATXS by replacing the FDDI core with a switched core. The Predictor models compare a shared medium (i.e., that of the shared 100-Mbps FDDI) with a dedicated medium performance of a switched infrastructure. The more advanced functions of ATM including virtual circuits, traffic classes, and traffic shaping cannot be modeled with Predictor.

5.1 Traffic collection and import into the Predictor model

Predictor classifies network traffic as packet flows where each flow has a mean and standard deviation of packet size and mean and Squared Coefficient of Variation (SCV) of packets per second. The SCV is defined for a flow of packets in terms of the variance and mean of packet interarrival time (IAT) as:

$$\text{SCV of packet IAT} = [\text{CoV of packet IAT}]^2 = \frac{\text{Variance of packet IAT}}{[\text{Mean of packet IAT}]^2} \quad (2)$$

The SCV of IAT is a measure of the burstiness (or "clumping" tendency) of a packet flow, as described in Section 2. A deterministic packet flow (i.e., packets always have a constant IAT) would have SCV = 0. A Poisson packet flow (i.e., packets have exponentially distributed IAT) would have an SCV = 1. Using Network Associates Sniffer (see [10]) trace files as input, a program named `crunch.c` was written to identify the conversation pairs (between layer-2 MAC addresses) and characterize each conversation for mean and standard deviation of packet size and mean and SCV of packet IAT. The output of the program is in a format suitable for import directly into a Predictor tabulated packet flow. Reference [4] is a NASA technical disclosure filed for this software program. The program has two possible output modes - verbose and CSV (for input to CACI Predictor as a tabulated packet flow). Predictor will import an Expert Sniffer file as traffic input. However it was found via experimentation (and confirmed by CACI technical support) that the Expert Sniffer is "behind" the actual capture. In one experiment, the FDDI Expert Sniffer identified about 10-Mbps of traffic, however the capture file showed approximately 26-Mbps.

The process of collecting and characterizing traffic over a period of many weeks is called "baselining". A baseline is the starting point for a network model. Growth rates, additional traffic sources, changes in network topology, and so on are then made to the baseline to predict future performance of the network. The `crunch.c` program applied to Sniffer capture files can characterize small snapshots of traffic on a LAN. For a longer term view, RMON probes can be used to record the layer-2 mean packet size and packets per second for all conversations in a network. The use of RMON for long term monitoring of shared FDDI bandwidth usage needs to be further investigated. For a fully switched network (e.g., switched Ethernet or ATM), RMON cannot be used.

For a router-centric view of traffic, Cisco NetFlow provides session level information (a session is a TCP connection) for all packet flows through a monitored router. The use of NetFlow traffic collection for input to Predictor models deserves further study.

Over a period of several days in mid June 1999 it was found that the industrial area FDDI ring of the FTXS ranged from about 15% to 25% utilization. The Network Associates FDDI Sniffer monitor (with a 100-millisecond sampling time) was used for these measurements. A single capture of "typical" non-launch day traffic on the industrial area FDDI was made at an approximate 25% utilization level and characterized. The results from the characterization are shown in Figure 7 as `crunch.c` verbose output where all conversation pairs with less than 1% relative utilization have been removed for purposes of clarity. It can be seen that 21 conversation pairs generate a relative 81% of the traffic for this snapshot. One MAC address could not be identified (labeled as "unknown") and is suspected to be a RMON probe or some other monitoring device on the industrial area FDDI ring.

```
-----
Media data rate           =      100 Mbps
Total time of trace      =      9.3434 sec
Total number of pkts / bytes = 93526 pkts / 30429145 bytes
Total (absolute) util    =      26.0541 %
Total number of pairs    =      355 pairs
-----
```

to/from Addr	count	util (a)	util (r)	m_pkt_len	sd_pkt_len	pkt/sec	scv
R9_LCC KSCISO	1600	0.55	2.09	398.25	584.34	171.24	0.97
ISOEAST KDN7513	2597	0.67	2.59	303.04	397.66	277.95	0.61
R16_CDSC R14_VAB	532	0.41	1.58	903.36	662.26	56.94	22.07
R4_OSB KSCISO	548	0.30	1.14	631.59	475.12	58.65	1.38
R14_VAB KSCISO	1201	0.93	3.57	905.22	656.27	128.54	0.88
R15_LCC KDN7513	509	0.40	1.54	917.71	630.73	54.48	19.59
KDNKSC KDN7513	1535	0.33	1.27	251.43	250.29	164.29	1.86
SODN R6 KSCISO	1675	0.60	2.30	418.47	440.45	179.27	1.14
KDN7513 ISOEAST	2528	0.35	1.35	163.00	182.58	270.57	0.52
KDN7513 R14_VAB	503	0.41	1.59	962.92	620.32	53.83	23.24
KDN7513 KDNKSC	1529	0.48	1.86	369.73	331.12	163.65	1.95
KDN7513 KSCISO	1465	1.05	4.04	838.38	619.88	156.80	1.74
SSPF KSCISO	1071	0.52	1.98	562.07	499.73	114.63	1.36
SSPF PONWAI	3756	2.94	11.27	913.12	646.42	402.00	2.40
unknown PONWAI	411	0.45	1.72	1270.64	433.16	43.99	4.24
KSCISO R9_LCC	2423	1.39	5.35	672.16	566.75	259.33	2.37
KSCISO PONWAI	1966	1.12	4.30	666.07	623.48	210.42	2.78
PONWAI SSPF	3514	1.25	4.80	415.91	569.45	376.09	2.05
PONWAI KSCISO	2714	0.93	3.57	399.89	443.89	290.47	1.83
PONWAI CADCAE	4981	5.78	22.19	1355.68	416.60	533.10	5.56
CADCAE PONWAI	3002	0.30	1.17	118.21	154.34	321.30	2.94

Figure 7 - Characterization of the industrial area FDDI ring FTXS traffic (6/15/99 - baseline 26% utilization)

5.2 Experiments and results for FTXS and ATXS Predictor models

The FTXS and ATXS Predictor models were used to predict the level of performance that could be expected for four video MPEG-2 video streams transmitted from the LCC to the EDL. The MPEG-2 waterpolo video clip frame size trace data was packetized into Ethernet (1500 byte maximum size) packets where frames of size larger than 1500 bytes are packetized into multiple back-to-back maximum size packets. The time series characteristics were found to be mean packet size = 1455 bytes, standard deviation of packet size = 208 bytes, packet rate = 506 pkts/sec, and SCV of packet IAT = 13.99 using the tools available at [12]. The data traffic characterized in Figure 7 was imported as the base data traffic representing a high 26% utilization of the FTXS industrial area FDDI ring. Figure 8 shows the mean and 90th percentile video packet delay for one through four video streams and with both no background data traffic and the baseline 26% utilization traffic. Figure 8 shows video packet delay for a growth in baseline data traffic. Figure 9 shows that video packet delay is fairly constant through a 150% growth, of about 65% utilization of data traffic of the 100-Mbps FDDI ring (the four video streams consume about 24% utilization for a total FDDI ring utilization of about 90%). This result shows that FDDI is "stable" for very high (up to 90%)

utilization levels after which packet delay increases sharply. The ATXS model predicts sub-millisecond mean and 90th percentile packet delay for growth in excess of 1000% (an 11 fold increase) in data traffic from the 26% baseline. This result shows that dedicated bandwidth can result in very large aggregate bandwidths for carrying many-to-many traffic. Buffer loss cannot be predicted with Predictor and is studied using Comnet in Section 7.

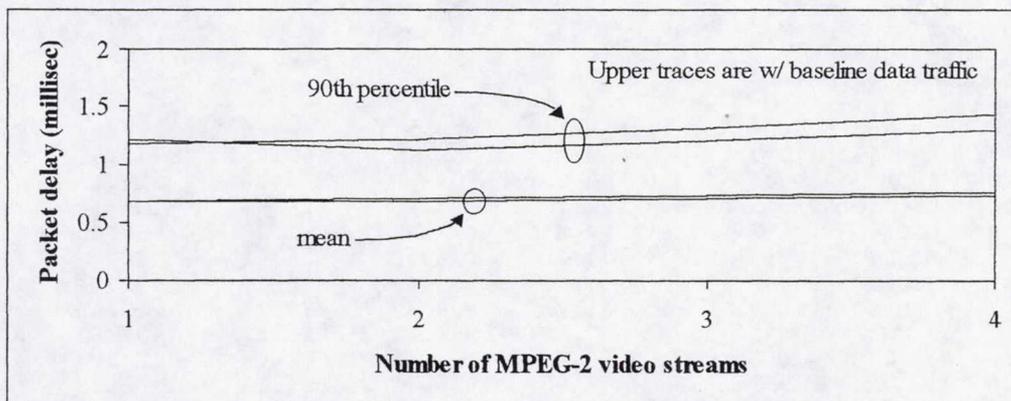


Figure 8 - Video packet delay for FTXS with one through four video streams

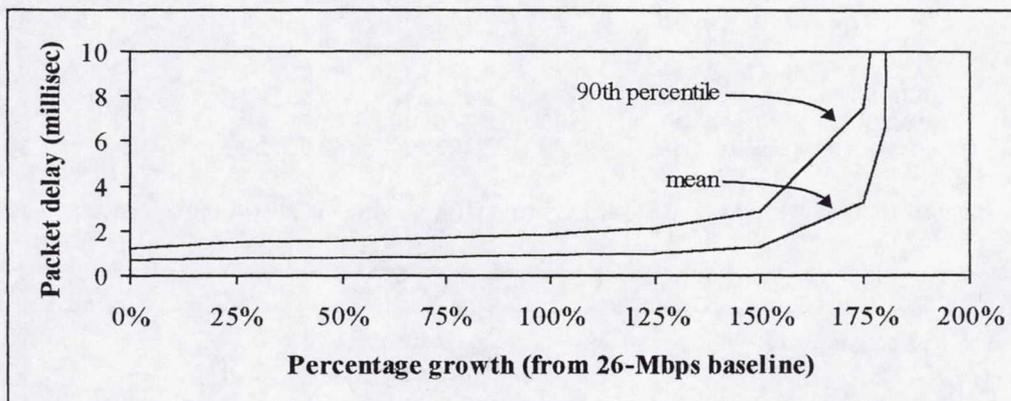


Figure 9 - Video packet delay for FTXS with four video streams and growth in data traffic

6. Launch Day Traffic Characterization

The KSC Web server is attached to the FTXS. On launch days the number of queries to this server is greater than on non-launch days. It is important that the internal NASA-KSC network perform "well" on launch days. Traffic data was collected for the 3 hours preceding the scrubbed July 22nd shuttle launch and characterized using the `crunch.c` program. Traffic data was also collected during the day of the scrubbed July 20th launch. An unusually high peak of almost 50% utilization of the FTXS FDDI ring was observed on 11:00am on July 20. Figure 10 shows characterization (`crunch.c` verbose output) of this capture where all conversation pairs with less than 1% relative utilization have been removed for purposes of clarity. It can be seen that 21 conversation pairs generate a relative 82% of the traffic for this snapshot. Figure 11 shows utilization during the night of the scrubbed July 22nd launch. It can be seen that traffic coming from and to the KSCISO router comprises the majority of the traffic on the FDDI ring. Traffic passing through the KSCISO router is also monitored with MRTG. The Sniffer data consists of approximately 10 second snapshots (10 seconds results in a full buffer on the Sniffer tool) taken every 10 minutes. Table 2 shows a breakdown of the protocols passing through the KSCISO router based on analysis (again, with `crunch.c`) of a 1:05am Sniffer trace. The large component of UDP traffic is believed to be from a new (first time used for this launch) streaming video application. It can be observed that

the burstiness of this streaming video is higher than that of the TCP traffic, most of which is presumed to be "normal" HTTP Web traffic.

```

-----
Media data rate           =      100 Mbps
Total time of trace      =      7.2778 sec
Total number of pkts / bytes = 122334 pkts / 45045982 bytes
Total (absolute) util    =      49.5158 %
Total number of pairs    =      322 pairs
-----

```

to/from Addr	count	util (a)	util (r)	m_pkt	len	sd_pkt	len	pkt/sec	scv
ISOEAST KDN7513	1241	1.30	2.63	955.95		659.15		170.52	5.25
R16_CDSC SODN_R6	1634	1.61	3.26	897.77		628.06		224.52	4.56
R15_LCC KSCISO	1355	0.98	1.97	656.45		616.54		186.18	1.57
SODN_R6 KSCISO	4045	2.18	4.40	490.48		518.27		555.80	2.21
KDN7513 KDNKSC	657	0.60	1.20	824.75		608.81		90.27	2.27
KDN7513 KSCISO	1402	0.90	1.81	581.44		535.36		192.64	1.98
SSPF KSCISO	1587	0.85	1.72	488.87		478.29		218.06	2.21
SSPF PONWAI	3578	3.74	7.56	951.32		695.01		491.63	6.34
KSCISO R5_OSB	625	0.81	1.63	1174.82		522.85		85.88	4.81
KSCISO R2_VABR	974	0.98	1.98	916.42		694.10		133.83	1.72
KSCISO R15_LCC	1468	1.37	2.76	846.38		709.44		201.71	1.88
KSCISO KDNKSC	966	0.96	1.95	908.41		571.51		132.73	2.78
KSCISO SODN_R6	3547	1.46	2.94	373.52		515.45		487.37	2.22
KSCISO PONWAI	4287	3.10	6.26	657.66		610.23		589.05	3.53
PONWAI KDNKSC	664	0.61	1.22	829.10		679.84		91.24	18.81
PONWAI KDN7513	906	0.75	1.51	748.64		554.55		124.49	2.30
PONWAI SSPF	3593	3.48	7.04	882.14		711.52		493.69	4.73
PONWAI KSCISO	4291	1.58	3.18	334.21		405.07		589.60	2.62
CADCAE EDL_BDR	2734	4.43	8.94	1472.42		219.03		375.66	7.20
CADCAE KDNKSC	2807	3.91	7.89	1266.50		509.70		385.69	5.94
CADCAE <new>	2956	4.85	9.80	1493.04		192.20		406.16	6.27

Figure 10 - Characterization of the industrial area FDDI ring FTXS traffic (7/20/99 - launch day 49% utilization)

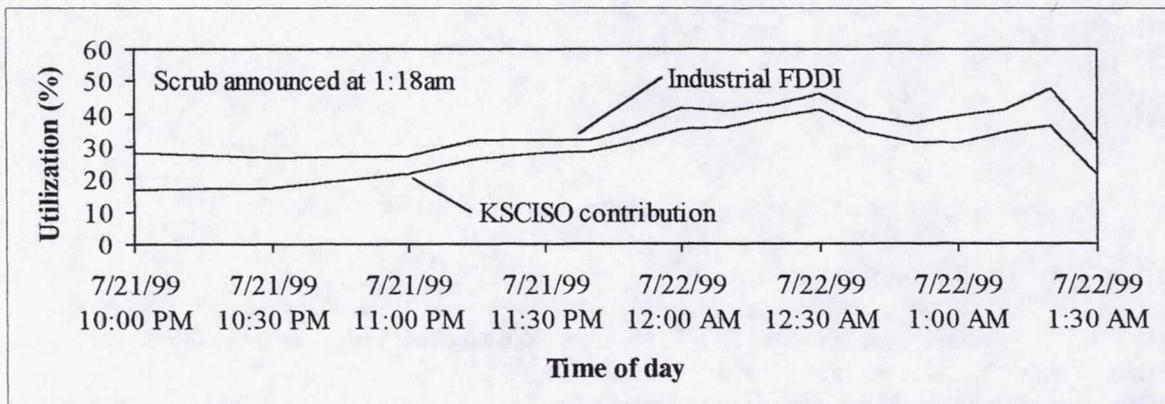


Figure 11 - Industrial FDDI ring utilization for night of July 22nd scrubbed launch (data from crunch.c)

Component	FDDI utilization	Weighted average of SCV
All traffic on FDDI	37.9%	10.19
KSCISO only	26.0	11.39
TCP to FTXS	1.0	3.47
TCP from FTXS	6.4	6.23
UDP to FTXS	2.0	5.71
UDP from FTXS	13.6	21.62

Table 2 - Breakdown of protocols at 1:05am from the July 22nd scrubbed launch (data from crunch.c)

7. Comnet Modeling of Video Transmission on ATM

A CACI Comnet model can give a detailed, or “low level”, view of network behavior for a small snapshot of time. This low level view includes an understanding of packet or ATM cell loss due to buffer overflows in switches and routers. A Comnet model was built to study the delay and loss performance of video traffic in a simple ATM network. In particular, the use of shaping of video traffic is to be studied as a means of improving both video and data traffic performance. Figure 12 shows the Comnet model developed to study video traffic performance. Since no MPEG-2 encoded shuttle launch video was available, Olympic Games MPEG-2 video traces were used. Two experiments were defined:

Experiment #1 - Four streams of video with high CoV (archery, badminton, boxing, and diving)

Experiment #2 - Four streams of video with low CoV (sailing, swimming, table tennis, and waterpolo)

The video sources are Video_node_1, etc. and the video sinks are Video_sink_1, etc. The link speeds are all OC-3 and the switch buffer size is set to 512 cells. Shaping (and SAR to/from ATM cells) of the video traffic at the source is done using the Comnet external messages feature. Shaping rates of 6-Mbps through 40-Mbps are used.

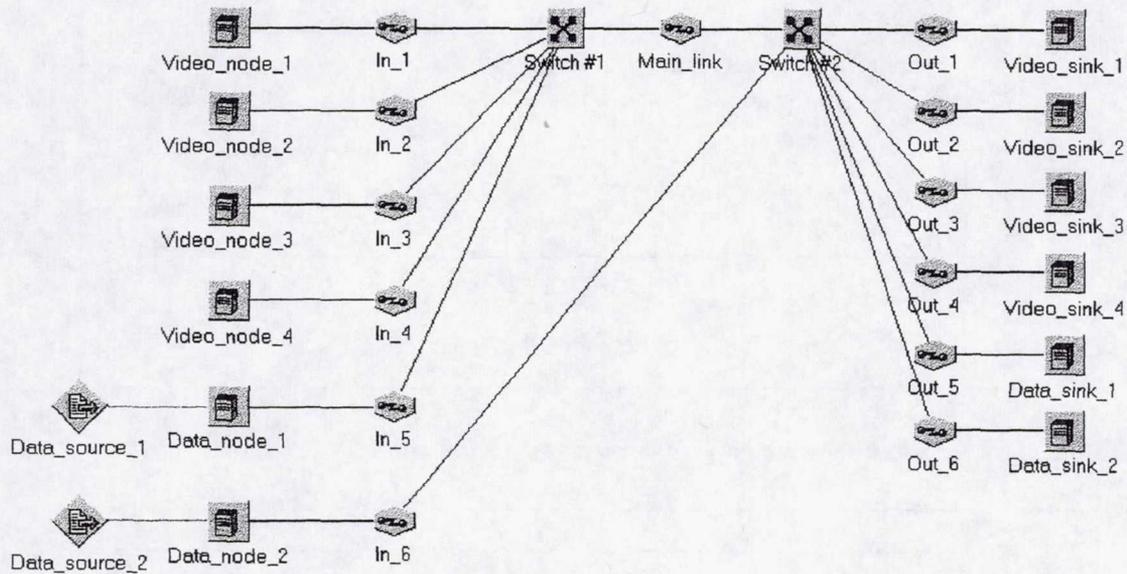


Figure 12 - CACI Comnet model for studying video traffic performance in an ATM network

Figure 13 shows the frame delay of MPEG-2 video for **Experiment #1** as a function of shaping rate. Results are only shown for the case where all frames are successfully delivered. A shaping rate of 6-Mbps is “too slow” and shaping at a rate higher than about 35-Mbps results in cell loss, and thus frame loss, at the switch. Figure 14 shows switch buffer use and its sharp increase at the higher rates of shaping. At a shaping rate of 40-Mbps, very few MPEG-2 frames are successfully delivered. This is not surprising since four times 40-Mbps exceeds the OC-3 line rate. Figure 15 shows frame delay for **Experiment #2**. It can be seen that the lower CoV video streams result in a significantly lower maximum frame delay and slightly lower mean and standard deviation of frame delay.

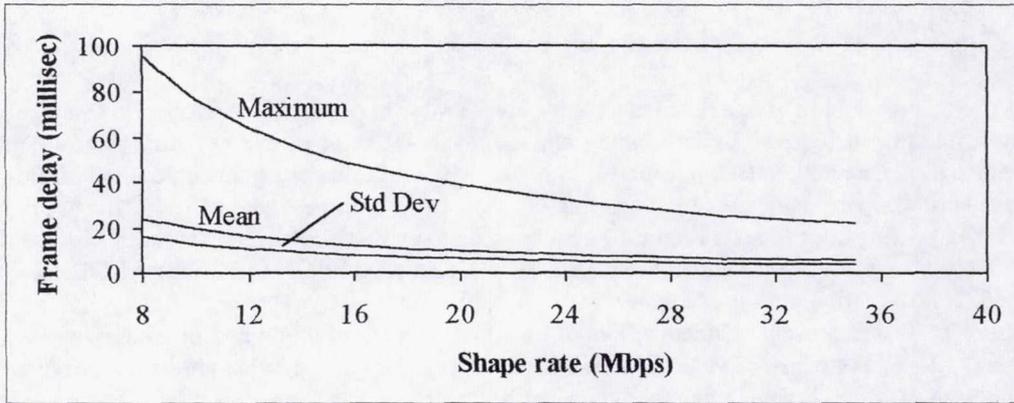


Figure 13 - Frame delay results from shaping experiment #1 for high-CoV video streams

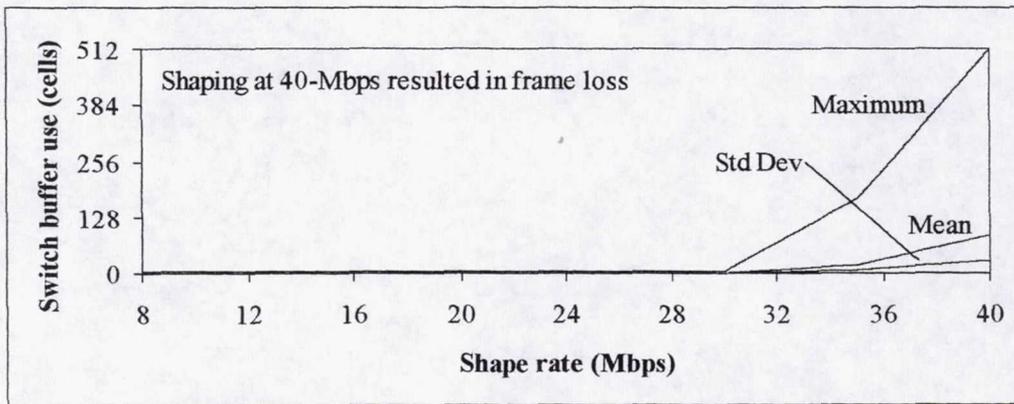


Figure 14 - Buffer use for switch #1 high-CoV video streams (for experiment #1)

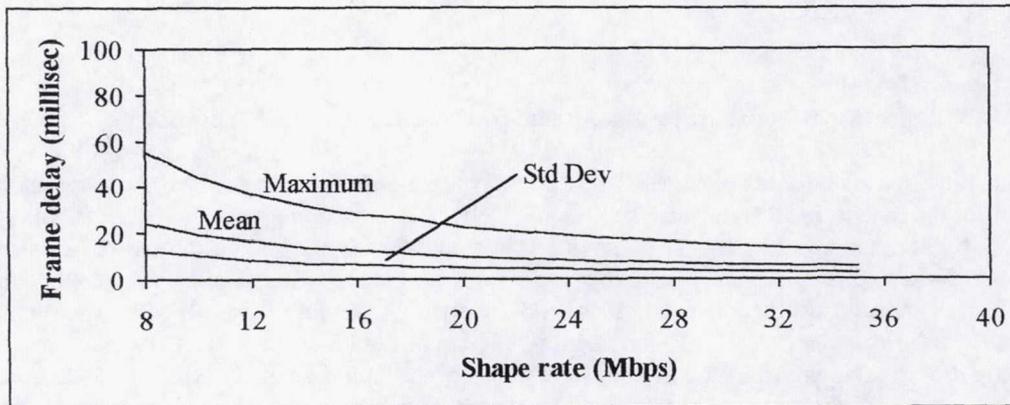


Figure 15 - Frame delay results from shaping experiment #2 for low-CoV video streams

8. Summary and Future Work

This work has demonstrated the use of performance modeling and the use of the CACI Predictor and Comnet tools. Modeling can be used to predict network performance and tune parameters for optimum performance. It

was shown that the current FTXS can support "a few" video streams with current load levels and traffic mixes. ATXS with its switched infrastructure offers, at least, a ten fold improvement for transport of video compared to FTXS. It was shown that the characteristics of a video stream has a direct correlation on its performance in a network. It was also shown that shaping of video streams is necessary to prevent overflow losses. Such overflow losses would both destroy video quality and affect performance of data application (i.e., by possibly inducing data traffic to also be lost). The Comnet simulation model can be used to explore parameter setting for video shaping that optimize the trade-off between video frame delay, allocated bandwidth (for video), and switch buffer utilization.

Future work should focus on fully validating the models against "real networks". This could be done by injecting, measuring, and comparing probe traffic in the real FTXS and doing the same in the model of FTXS. The developed models can be used to predict when the FTXS when "run out of room". To do this, long term trending of traffic must be done. This could be done using RMON, or some other traffic collection, probes and careful collection, archiving, and trending of traffic data over many months. In addition, the adoption of new technologies, such as streaming video, need to be fully understood. Finally, parameter settings for ATM links for transmitting shuttle launch video (e.g., to set-top MPEG-2 decoder boxes) can be optimized using the developed Comnet model. To do this, the "specs" on the actual MPEG-2 encoders to be used at NASA-KSC and sample video streams from these encoders will be required. Thus, this future work can provide useful input and validation to future set-top box projects within the AND group.

Acknowledgments

The author would like to acknowledge Gregg Buckingham, Jane Hodges, Ray Hosler, and Judie Gilliam of the NASA-KSC University Programs Office, John Schnitzius of NASA-KSC Development Engineering, and Chris Kerios and Paul Kaun of Dynacs Engineering. The FTXS network topology was generated by Chris Knoell of Dynacs Engineering.

References

- [1] CACI Products Company, Simulation Products, 1998. URL: <http://www.caciasl.com/>.
- [2] K. Christensen and C. Kerios, "Network Modeling and Planning Tool Specification," NETX1-0081 (Kerios, Chris, Subject 1540 Modeling and Planning Tool) EDC Purchase Requisition, August 1998.
- [3] Fiber Distributed Data Interface, Media Access Control (MAC), ANSI X3.139-1987, ISO 9314-2:1989.
- [4] P. Kaun and K. Christensen, "Crunch - A Software Program to Characterize Conversation-Pair Traffic Data from a Local Area Network (LAN) Sniffer Trace File," *Disclosure of Invention and New Technology (Including Software)*, NASA-KSC, July 28, 1999.
- [5] L. Kleinrock, *Queueing Systems, Volume 1: Theory*. John Wiley and Sons, New York, 1975.
- [6] KSC Communications Systems Upgrade Preliminary Engineering Report (PER). *KSC-DL-3700*, September 1996.
- [7] A. Law and D. Kelton, *Simulation Modeling and Analysis*. McGraw-Hill, New York, 1982.
- [8] D. Menasce, V. Almeida, and L. Dowdy, *Capacity Planning and Performance Modeling*. Prentice Hall, Englewood Cliffs, New Jersey, 1994.
- [9] The MPEG Home Page, 1999. URL: <http://drogo.cselt.stet.it/mpeg/>.
- [10] Network Associates, Sniffer Network Analyzer, 1999. URL: <http://www.networkassociates.com/>.
- [11] O. Rose, "Simple and Efficient Models for Variable Bit Rate MPEG Video Traffic," *Performance Evaluation*, Vol. 30, No. 1-2, pp. 69 - 85, July 1997.
- [12] Tools Page for Ken Christensen, 1999. URL: <http://www.csee.usf.edu/~christen/tools/toolpage.html>.
- [13] Traffic (Version 2.1), Knoltex Corporation, 1999. URL: <http://www.knoltex.com/>.

1999 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

**JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA**

**Evaluating Education and Science in the KSC
Visitor Complex Exhibits**

Lance K. Erickson
Professor
Applied Aviation Sciences
Embry-Riddle Aeronautical University

ABSTRACT

The continuing development of exhibits at the Kennedy Space Center's Visitor Complex is an excellent opportunity for NASA personnel to promote science and provide insight into NASA programs & projects for the approximately 3 million visitors that come to KSC annually. Stated goals for the Visitor Complex, in fact, emphasize science awareness and recommend broadening the appeal of the displays and exhibits for all age groups. To this end, this summer project seeks to evaluate the science content of planned exhibits/displays in relation to these developing opportunities and identify specific areas for enhancement of existing or planned exhibits and displays.

To help expand the educational and science content within the developing exhibits at the Visitor Complex, this project was structured to implement the goals of the Visitor Center Director. To accomplish this, the exhibits and displays planned for completion within the year underwent review and evaluation for science content and educational direction. Planning emphasis for the individual displays was directed at combining the elements of effective education with fundamental scientific integrity, within an appealing format.

Evaluating Education and Science in the KSC Visitor Complex Exhibits

Lance Erickson

1. Introduction

Preliminary discussion of the goals of this project were used to define the scope of the work for the summer. That effort is to assist with the improvements in the KSC Visitor Complex exhibits in the areas of general science and education. This was to include assisting in the communication of NASA's operations and interests to visitors at the complex, and to help introduce science and technology to the visitors through the displays and exhibits being planned.

2. Project Description

Organization of this project began with the identification of the galleries and exhibits that were undergoing development and that would be completed within one year. The primary focus was on these exhibits because of the need for input during the development stages. Suggestions and/or recommendations for educational and scientific content in these exhibits were needed before the exhibit designs were essentially fixed. Other displays and exhibits could be reviewed during the summer project as time provided, and were, as outlined in the appendix.

Review procedures for the exhibits began with a discussion of the planning documents with the Director and Assistant Director. A summary of recommendations/suggestions was composed for several of the exhibits and discussed. At this point, the elements of educational content were introduced and recommendations made for the first series of exhibits.

The first exhibit/gallery review was made at the beginning of the project on the "Exploration in the New Millennium." The content and character of the exhibits and displays was evaluated for accuracy and continuity, with several suggestions outlined (see Appendix A). Some modification of the displays was recommended, and an addition of a night sky display suggested. More detailed discussion on the gallery continued through the summer program, with some adjustments and modifications being integrated into the gallery under the direction of the Visitor Complex Director.

An important exhibit area under redesign during the coming year was the Rocket Garden, a collection of mostly early manned space flight rockets and propulsion

engines. Since many of the rockets on display are historically important in NASA's evolution, a recommendation was made to expand the exhibits to include information on the background of space flight propulsion, and the importance of the individual launch vehicles (see Appendix B).

During a paper presentation on this NASA-ASEE project in early summer, I had the opportunity to visit to the Adler Planetarium, which provided some insight into similar display and exhibit development. Another astronomer had help direct the exhibit development at the Planetarium and described the educational and scientific components within the planning, design, and installation of the newer displays. Those fundamental educational elements were adopted, in part, as the recommended development procedures for the KSC Visitor Complex future exhibits.

Discussions with the education office of the American Astronomical Society indicated the need for education specialists within the review and recommendation process if the exhibits are to provide significant educational benefit. Professional advice on educational aspects of exhibit development was also suggested because of the broad age groups targeted for the future displays. These suggestions for education specialists are included in the conclusions of this report. Consultation with the NASA-HQ Education Office also provided several evaluation principles for the educational content within the planning and implementation process.

Subsequent exhibit/gallery reviews at the Complex included the Apollo/Saturn V Center and the International Space Station Center. Both of these visitor sites were already completed, but were considered useful for educational review because of their large impact on the visitors. Suggestions for additions or alterations are listed in the appendices.

3. Results

A series of reviews for the various galleries, exhibits and displays at the KSC Visitor Complex was made during the summer program as part of the project. Those reviews were incorporated into recommendations, originally appearing in several memos to the Visitor Complex Director. The recommendations cover specific and general objectives, primarily within the exhibits of focus. Other recommendations and suggestions cover other exhibits/displays, and can be found within the appendix.

The highest level recommendations developed during this summer project were related to educational principles and planning. Recommendations from the staff at the Adler Planetarium, The American Astronomical Society, and the NASA-HQ Education Office indicate a need for planning at all stages for exhibits with educational intent. Although some of these displays and/or exhibits must also

provide an interesting and imaginative character, the primary content should adhere to the fundamentals of learning. To accomplish this ideal, is recommended that the following structure be adopted for exhibit/display development.

- A. To identify both specific and general interests of the visiting public, one or more surveys should be conducted with the educational and information objectives of the Visitor's Complex in mind. The sampling/surveys could be conducted at the Center, or through mail-in questionnaires. Although this step is not necessary for individual exhibit development, it would be useful for identifying program interests, or for planning multiple exhibit projects.
- B. One of the fundamental steps in exhibit planning & development is to identify the basic concepts to be covered within the planned exhibits. An outline of the concepts should include the goals and objectives of each of the concepts that make up the exhibit/display. This serves as the foundation of the educational content review.
- C. Once the objectives and goals have been established, an educational consultant should be used to verify the information and/or learning content and efficacy of the planned exhibits. This is important for evaluating the educational displays, and may also be useful for a review of any amusement-oriented displays.
- D. An evaluation of each of the exhibits by a survey team before it is placed on permanent display is recommended. Several styles of evaluation could be used for the exhibits/displays including; 1) a survey of children and/or adults interacting directly with the individual exhibits; 2) a broad survey of several exhibits evaluating the direct interest and comparative interest in the exhibits.
- E. Follow-on surveys were highly recommended to evaluate the public interest in the exhibits, which could also be very useful for estimating the public's interest in a number of space and related-science areas.

A more general suggestion on increasing the effectiveness of exhibits and displays at the Visitor Complex includes increasing the continuity between exhibits throughout the Complex to increase the synthesis of information for the visitors. An example of linking/integrating exhibits at the Complex would be a listing of the uses for and current flights of launch vehicles/rockets at KSC/CCAFS. This could be a listing within the Rocket Garden. Long-duration research and laboratories at KSC/CCAFS could also be included in the International Space Station Center exhibits to assist the public in understanding the KSC role in space exploration.

It is also strongly recommend that direct coordination of the exhibit & display development be made with the NASA Education Offices at Headquarters, and at

other centers. Those resources can help establish the educational benefits and content of the displays intended for that purpose. Second-year efforts in the project would include the establishment and maintenance of contacts between the Visitor Complex management and the Education offices of NASA.

4. Conclusions

A complete planning process for public education exhibits, including establishing goals and objectives, was emphasized by the educational staff affiliated with the Adler Planetarium, the American Astronomical Society and the NASA Education Office. Because of the emphasis on planning within these organizations, it is recommended that the planning activities in exhibit & display development include goals and objectives of the individual displays be adopted. Similarly, it is recommended that visitor surveys be conducted before the development of new exhibits at the complex, and that surveys follow the installation of new exhibits for evaluation of the effectiveness. Evaluation of existing exhibits could also be conducted with survey instruments before modifications/additions are developed.

Because one of the major goals and current efforts of the Visitor Complex management includes broadening the appeal of the exhibits for all age groups, emphasis on the addition to many of the exhibits for children is strongly supported.

Integration of the visitor exhibits is also suggested to increase the awareness of the visitors to the exhibit elements, and to help the visitor understand some of the technologies and difficulties associated with the space flight and space launch programs. An attempt was made in these recommendations/suggestions to help integrate the various visitor sites and exhibits.

Follow-on efforts to this project would concentrate on the coordination of the planning stages of exhibits while defining the goals and objectives for the exhibits/displays, and helping to integrate the exhibits to other NASA visitor sites and space flight efforts at KSC, and throughout NASA.

Appendices

A summary of the suggestions to the Director of the Kennedy Space Center Visitor Complex were included in the final report, but have been edited from this document due to space limitations. These suggestions and recommendations are part of earlier reports and reviews made as part of the summer project. To obtain a copy of the appendices, please contact Dr. Lance Erickson:
erickson@db.erau.edu

1999 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

**JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA**

Full Duplex, Spread Spectrum Radio System

Bruce A. Harvey, Ph.D.

Assistant Professor

Department of Electrical Engineering

Florida A&M University – Florida State University College of Engineering

KSC Colleague: Richard Nelson

ABSTRACT

The goal of this project was to support the development of a full duplex, spread spectrum voice communications system. The assembly and testing of a prototype system consisting of a Harris PRISM spread spectrum radio, a TMS320C54x signal processing development board and a Zilog Z80180 microprocessor was underway at the start of this project. The efforts under this project were the development of multiple access schemes, analysis of full duplex voice feedback delays, and the development and analysis of forward error correction (FEC) algorithms. The multiple access analysis involved the selection between code division multiple access (CDMA), frequency division multiple access (FDMA) and time division multiple access (TDMA). Full duplex voice feedback analysis involved the analysis of packet size and delays associated with full loop voice feedback for confirmation of radio system performance. FEC analysis included studies of the performance under the expected burst error scenario with the relatively short packet lengths, and analysis of implementation in the TMS320C54x digital signal processor. When the capabilities and the limitations of the components used were considered, the multiple access scheme chosen was a combination TDMA/FDMA scheme that will provide up to 8 users on each of three separate frequencies. Packets to and from each user will consist of 16 samples at a rate of 8,000 samples per second for a total of 2 ms of voice information. The resulting voice feedback delay will therefore be 4 – 6 ms. The most practical FEC algorithm for implementation was a convolutional code with a Viterbi decoder. Interleaving of the bits of each packet will be required to offset the effects of burst errors.

Full Duplex, Spread Spectrum Radio System

Bruce A. Harvey, Ph.D.

1. INTRODUCTION

The primary focus for this summer's efforts was the development of a prototype full-duplex, spread spectrum, digital voice radio system. This system is intended to provide open voice communications between at least 4-7 persons in large rooms (i.e. high bays) or in open outdoor areas (i.e. landing strip). The persons using the system are expected to be able to conduct normal conversation as if they were located in close proximity. This differs from push-to-talk systems where there is only half-duplex operation (only 1 speaker at a time). The spread spectrum operation of the system allows license-free operation in the 2.4 GHz ISM band and alleviates much of the fading and signal loss due to signal multipath.

The goals for the prototype communication system included:

- 1) Minimum of 4 – 7 users,
- 2) Full loop audio feedback to ensure transmit and receive operation,
- 3) Feedback delay much less than 20 milliseconds to avoid speech interference,
- 4) 8000 samples/second with 8 or more bit pulse code modulation (PCM), and
- 5) Operating ranges from a few hundred feet to approximately 3 miles.

The hardware for the design of the prototype devices had already been selected and were in the process of being assembled and debugged at the start of the effort. The Harris PRISM radio set (HWB1151) was chosen to provide the spread spectrum modulation, RF devices and baseband processing. The Texas Instruments TMS320C54x signal processor and evaluation board was selected to provide analog-to-digital and digital-to-analog conversion, signal processing and error control. A Zilog Z80180 microprocessor provides overall control of the system. The analysis and designs under this effort were to be used in this existing prototype design.

The primary tasks under this effort were trade-off analysis of multiple access protocols, analysis of communication link performance and development of forward error correction (FEC) algorithms for the prototype.

2. MULTIPLE ACCESS PROTOCOL DEVELOPMENT

A digital communications system can use a number of techniques or protocols to allow multiple users access to the communications channel. The three general techniques are frequency division multiple access (FDMA), time division multiple access (TDMA) and code division multiple access (CDMA). FDMA, TDMA and CDMA may be used individually or in combinations to meet the requirements and conditions of the communication system. Typically, very high bandwidth channels use a combination of FDMA with either TDMA or CDMA to obtain the benefits of TDMA or CDMA while allowing a narrower bandwidth receiver design.

The PRISM chip set provides the baseband to RF components necessary to implement a spread spectrum communications system in the 2.400 – 2.483 GHz band. It is capable of transmission rates of 1, 2, 5.5 and 11Mbps in a null-to-null bandwidth of 22 MHz. It is designed for packet communication and includes a baseband processor that controls preambles and headers. It uses direct sequence spread spectrum (DSSS) technology. Media access control (MAC) is provided by a Z80180 microprocessor.

The HFA3860B baseband processor controls the preamble and header generation, modulation and demodulation, spreading and de-spreading, and digital interfacing of the PRISM system. The preamble and header are always transmitted using differential binary phase shift keying (DBPSK) at 1 bit per symbol while the data can be transmitted in one of several different formats with up to 8 bits per symbol. One of the most significant factors concerning the preamble and header is that they total 146 to 322 symbols and are always transmitted at 1 Msymbols/sec regardless of the symbol rate of the data. The data symbol rate for the PRISM is 1 Msymbols/sec. for the 1 and 2 MBPS data rates, and 1.375 Msymbols/sec. for the 5.5 and 10 Mbps data rates. Thus increasing the data symbol rate has no impact on the transmit time required for the preamble and header.

The PRISM chip set can have an 11- to 16-bit spreading sequence depending on the version used and the user settings. However, 16 bits is insufficient to provide adequate isolation between CDMA users. The PRISM system is thus not capable of implementing a useful CDMA protocol for this application. The PRISM is capable of operating in 3 non-overlapping 22 MHz channels within the total 83 MHz operating band. This allows the implementation of a 3-channel FDMA scheme. The radio is designed to operate as a packet radio and thus is also well suited for the implementation of a TDMA protocol. A combination TDMA/FDMA protocol will provide even more flexibility and was chosen for implementation in the prototype system. [1]

Several protocols for establishing a TDMA system have been proposed or implemented. Many of these protocols are wholly inappropriate for this application. These protocols include ALOHA or Slotted ALOHA, token ring, and carrier sense multiple access (CSMA) or ethernet. For continuous communication with fixed data rate requirements, the most efficient TDMA protocol will include fixed and allotted time slots. Each user (transceiver) will be allocated a specific slot in a TDMA time frame in which to transmit its data. For reliability, monitoring and interconnection with other voice systems, the TDMA system will include a Central Controller through which all communications will be conducted. The Central Controller (CC) will be a set of up to three (3) transceivers each operating in its own frequency band. The CC will receive transmission from every user transceiver, combine the voice signals and transmit the combined voice communications back to each transceiver. Also, the transmitted frame from the CC will be used to establish the starting time for each TDMA frame.

The most obvious first approach to TDMA implementation for voice applications is for each transmitter to take each voice sample and form a packet that is then transmitted during a prescribed timeslot. Since sampling is intended to have an 8 kHz rate, then there is 125

microsecond (μs) between each sample. However, the minimum preamble/header size for a packet is 146 μs . Therefore, this method will not work.

Multiple samples of the audio input must be sent in a single packet to increase the time between packets and the amount of data sent per preamble/header. Storing multiple samples for transmission introduces delay into the communications channel. A delay of 20 – 30 milliseconds (ms) or even longer is generally not noticeable in most voice communication systems. However, one of the goals for this system is full feedback of the audio signal to the users ear for confirmation of operation. Delayed feedback of auditory information to the speaker of over 20 – 50 ms has been shown to cause speech and concentration problems for the speaker. Thus, the feedback delay will be kept well under 20 ms to avoid these problems.

A maximum auditory feedback delay of 5-6 ms was chosen for the prototype system; longer delays may be tested in the future. Since the transmission and reception of the data will each introduce a delay, then the delay due to storage of samples must be no greater than 2.5 ms. To allow for processing delays, 2 ms was chosen as the storage time for each packet of data. A storage time of 2 ms is equivalent to storing 16 samples at an 8 kHz rate for each packet of data.

Sixteen samples using 8-bit pulse-code modulation (PCM) results in 128 bits of data for each packet. If rate $\frac{1}{2}$ FEC is used, then the data will be 256 bits (512 bits for 16-bit PCM). At the 1 Mbps data rate the number of bits per symbol is 1 and therefore a packet with preamble, header and 256 bits of data is a minimum of 402 symbols or 402 $\mu\text{seconds}$. There is 2 ms in each frame and thus 4 TDMA time slots can be implemented in a single frequency channel. Table 2-1 shows the packet size and maximum number of users possible for each transmit data rate. Preamble synchronization lengths of 80 and 128 symbols (146 and 194 symbol preamble & header) are considered. The 128 symbol synchronization length is the minimum suggested for use with antenna diversity and is the required length for the IEEE 802.11 standard. The data packet is assumed to be 256 bits. The results in parenthesis assume 512 bits per packet (16 samples of 16-bit PCM with rate $\frac{1}{2}$ FEC). The actual prototype uses 14-bit PCM padded to 16 bits using the 14-bit analog-to-digital converters included on the DSP card used.

Data Rate (Mbps)	Symbol Rate (Msym/sec)	Synch. Length (symbols)	Packet Length (symbols)	Packet Time (μsec)	# TDMA Slots / Freq.	
					Minimum	50 Symbol Guard Band
1	1	80	402	402	4 (3)	4 (2)
		128	450	450	4 (3)	4 (2)
2	1	80	274	274	7 (4)	6 (4)
		128	322	322	6 (4)	5 (4)
5.5	1.375	80	210	192.55	10 (8)	8 (7)
		128	258	240.55	8 (6)	7 (6)
11	1.375	80	178	169.3	11 (10)	9 (8)
		128	226	217.3	9 (8)	7 (7)

() – Data packet assumed to be 512 bits = 16 samples \times 16 bits/sample \times 2

Table 2-1. TDMA Slot Analysis

The TDMA slots in operation require guard bands (time) in order to assure no overlapping transmitted signals. The guard band was chosen to be 50 symbols to allow for the propagation (maximum of 3 mile range), timing and shut down periods. The addition of these guard bands increase the TDMA time slot requirements and consequently reduces the number of allowable users in each frequency band. The numbers of available TDMA slots with guard bands are listed in the final column of Table 2-1.

The HFA3860B Baseband included in the most recent PRISM chipset accomplishes the DSSS acquisition. The HFA3860B uses a parallel correlator to detect the 11-bit PN sequence (Barker code) used to spread the signal in the header and preamble. The timing and phase of the received signal is detected and used to synchronize the serial correlators used in data decoding. To improve the signal-to-noise ratio (SNR) for detection the acquisition process uses the average of 15 received symbols. The de-spreading by the correlators improve the SNR by $10\log(11) = 10.4$ dB. The averaging improves the SNR by another 11.7 dB to ensure reliable PN detection and accurate timing and phase information. Detailed information on the HFA3860B can be found in the Harris data sheet "HFA3860B Direct Sequence Spread Spectrum Baseband Processor." [2]

The preamble SNR is improved by 10.4 dB by the de-spreading of the PN code, the same as for data demodulation. However, the preamble SNR is improved during acquisition by an additional 11.7 dB above that of the data demodulation by averaging. If the received signal level is sufficient to provide a high enough SNR for reliable data demodulation, then the SNR will be more than 10 dB higher for acquisition and thus reliable acquisition will be achieved. Therefore, the reliability of the communication and link budget analysis will be determined by the bit error rate (BER) of the data demodulation.

3. LINK BUDGET AND OPERATING RANGE ESTIMATES

The initial link budget for this implementation of the full duplex, spread spectrum system relies heavily on the receiver sensitivity specifications in the Harris PRISM MACless DSSS Radio HWB1151 User's Guide [1]. The receiver sensitivity is given for each possible data rate and are listed in Table 3-1. The specification lists the minimum received signal level to achieve 8% packet error rate (PER) assuming a packet size of 1000 bytes. Assuming random bit errors (not bursty), this is equivalent to the minimum received signal level for a bit error rate (BER) of 10^{-5} .

Transmitter Data Rate and Modulation	Receiver Sensitivity
1 Mbps DBPSK	-89 dBm
2 Mbps DQPSK	-86 dBm
5.5 Mbps CCK	-86 dBm
11 Mbps CCK	-83 dBm

Table 3-1. MACless PCMCIA PRISM Receiver Sensitivity (8% PER)

According to the Harris data sheet for the HFA3860B Baseband Processor (used in the newest versions) [2], the observed errors occur in groups (i.e. burst errors). The 1 and 2 Mbps modes

had errors in groups of 4 and 6 bits due to the differential decoding and de-scrambling error propagation. The 5.5 and 11 Mbps CCK modes had errors in symbols of 4 and 8 bits, respectively. These burst errors will be extended if the de-scrambler is used. An 8% PER over a 1000 byte packet translates into a 4×10^{-5} and 8×10^{-5} symbol or burst error rate for bursts of 4 and 8 bits, respectively.

The receiver sensitivities given in Table 3-1 can be used to estimate the range of operation for the communications system. The version of the PRISM currently being used has a transmit power P_t of 18 dBm. An external amplifier is being considered which will bring the transmitter power up to 27 dBm (1/2 Watt). If all values are expressed in decibels (dB), then the transmit power P_t and receive power P_r can be related by $P_r = P_t + G_t + G_r - L_p - FM$ where $G_{t,r}$ = transmitter, receiver antenna gain in dB, L_p = propagation loss in dB, and FM = the designed fade margin in dB.

There is no set level for the fade margin that will compensate for all signal variability in signal level. Rather, the fade margin provides a measure of protection. A larger fade margin improves the likelihood of compensation for the variability and hence the reliability of the link. In a Harris application note "Tutorial on Basic Link Budget Analysis" [3] the authors suggest the fade margin be 20 – 30 dB, even for this spread spectrum system. However, in a clear, unobstructed outdoor environment, the fade margin may be reducible to 10 – 15 dB.

The propagation loss L_p is highly dependent on the environment in which the system is operating. In free space,

$$L_p = L_{fs} = 20 \log \left(\frac{4\pi D}{\lambda} \right)$$

where D = the transmit distance and λ = the wavelength. However, if the antennas are outdoors close to the ground, then the plane earth model for propagation is more appropriate.

$$L_p = L_{pe} = 20 \log \left(\frac{D^2}{h_t h_r} \right)$$

where $h_{t,r}$ = the height of the transmit, receive antenna. The plane earth model is more accurate whenever

$$D > \frac{4\pi h_t h_r}{\lambda}$$

Models for indoor communications are widely varied and highly dependent on room sizes, building construction and orientation of the antennas. The authors of the Harris application note [3] suggest a simple and conservative estimate for indoor communications. Their model uses free space communication for up to 20 feet and then 30 dB attenuation for each additional 100 feet. This model probably assumes an office environment consisting of individual offices and conference rooms. Thus much of the propagation loss is likely to be walls, furniture and other obstructions.

The environment for indoor communications required at NASA will likely include large processing room or high bays. While the room will be larger than a typical office, a high-bay will likely contain many large metallic objects and structures which will translate into increased multipath and attenuation. Therefore, the model proposed in the Harris application note will suffice for initial designs. Tests of the prototype system in a working environment will be required to ascertain the multipath effects and propagation models to be used.

The maximum operating ranges for outdoor and indoor environments can now be calculated using the receiver sensitivity, link budget equation and propagation models. A relatively modest fade margin of 10 dB will be assumed for this prototype analysis. Maximum operating ranges are listed in Table 3-2 as a function of antenna heights and transmitter power. The 18 dBm transmit power of the Harris PCMCIA version of the PRISM and the 27 dBm output of a ½ Watt amplifier are depicted in the table. An antenna height of 5 feet is assumed for the mobile unit carried by a person. The base or central controller antenna height is varied.

Antenna Height (ft.)		Maximum Range (ft.) +18 dBm			Maximum Range (ft.) +27 dBm		
Mobile	Base	Outdoor	Model	Indoor	Outdoor	Model	Indoor
5	5	942	Plane	138	1879	Plane	178
5	10	1158	Free	138	2658	Plane	178
5	15	1158	Free	138	3255	Plane	178
5	20	1158	Free	138	3758	Plane	178
5	25	1158	Free	138	4202	Plane	178

Plane = plane earth outdoor model, Free = free space outdoor model

Table 3-2. Estimated Operating Ranges for Prototype System

The results in Table 3-2 indicated several important facts. First, outdoor maximum operating range will ultimately be limited by plane earth propagation. Increasing the base or central controller antenna height will increase the range of operation, but it will limit the mobility of the system. The ½ Watt transmit power and 25 ft. base antenna resulted in a maximum range of 0.8 miles. The indoor propagation model used resulted in an operating range of less than 200 ft. even for ½ Watt transmit powers. This model is expected to be overly conservative for operation in a high bay and testing is suggested.

4. FORWARD ERROR CORRECTION

Many voice communications use a minimum or no forward error correction. Speech communications is rather tolerant to random and burst errors and thus FEC is not implemented. However, there are two factors that make an analysis of FEC for this system necessary. First, this system may be used in mission critical or safety-related operations where clear communications is required. Second, one possible extension of this system is the overlay of data on the voice network that requires higher reliability than voice. For these reasons, an analysis of error correction techniques and capabilities was conducted.

The TDMA protocol selection described previously resulted in voice data packets of 128 to 256 bits depending on whether 8-bit or 16-bit PCM encoding is used. These relatively small data packets will be subject to burst errors due to the modulation, scrambling and detection of the transmitted symbols. The 11 Mbps mode of the PRISM chipset is the likely operating mode of the prototype system to maximize the number of users. This mode is characterized by 8-bit burst errors; longer if scrambling is used. A single burst error in a 256-bit packet lowers the average bit error rate of the packet to about 0.016. This burst error will result in a very brief noise burst in voice communication, but is unacceptable for data communications.

The TMS320C54x signal processor has been designed to facilitate the decoding of the convolutional codes using the Viterbi decoder. Efficient implementation algorithms have been designed and estimates of the processing time required to decode blocks of data using the Viterbi decoder have been provided in application notes. Block codes can provide the burst error protection desired, but these codes generally have fixed lengths that limit the flexibility and reduce the efficiency of the system. Therefore, convolutional coding with Viterbi decoder was selected for the prototype system.

Convolutional codes are well suited to provide increased reliability in the presence of random bit errors during transmission. Interleavers will be used to distribute the burst errors to make them appear more like random bit errors. Interleavers require less processing and memory storage, and are more flexible than other approaches such as concatenated codes or specially designed burst error correcting convolutional codes. [4] Interleavers for this effort were assumed to distribute burst errors over the entire length of the coded data packet.

Encoding a packet of data using a convolutional code with k inputs, n outputs and memory length m requires the encoding of $k \times m$ additional input zeros in order to "complete the code". [4,5] The number of added output bits to the packet after completing the code is $n \times m$ bits and is often referred to as the "tail" of the encoded packet. For example, if 100 bits were encoded using a rate $2/3$, memory length 4 convolutional code ($m = 4$, $k = 2$ inputs and $n = 3$ outputs) then the number of bits in the encoded packet will be $100 \times 3/2 + 3 \times 4 = 150 + 12 = 162$ bits. This makes the effective rate of the code $100/162 = 0.62$.

For analysis purposes, it is assumed that the burst errors are randomly distributed over the entire packet by the interleaver and thus are effectively modeled as an increase in bit error rate for the packet. For example, a burst error of length 8 in a 162-bit packet will result in an average of 4 bit errors in the packet. The simulation models the burst error after interleaving as random probability of bit error (bit error rate) or $4/162 = 0.025$.

Simulations were conducted to quantify the effects of the burst errors on short packets with convolutional codes. A Viterbi decoder was simulated to decode convolutionally encoded packets with randomly occurring burst errors. The length of the burst errors was chosen to be 8 bits to correspond to characteristics the 11 Mbps mode of the Harris PRISM without data scrambling. The simulations were conducted on packets of 128 and 256 data bits (before encoding) to bound the expected range of data bits for 16 samples of voice data (8 to 16-bit

PCM). For easy comparison to decoding with random bit errors, the simulations were conducted as a function of the average bit error rate of the system.

The results of the simulations for rate $\frac{1}{2}$ convolutional codes with a memory length of 2 (constraint length K of 3) are shown in Figure 4-1. The abbreviations in the legends are defined as follows: NB = no burst errors (random only), 8B = 8-bit burst errors, Est = analytical estimate of 8-bit burst errors assuming only 1 burst error per packet, and BK = the length of data (uncoded) in the blocks. The simulation results demonstrate that the burst errors significantly degrade the performance of the Viterbi decoder for short packets. Longer packet lengths have better (lower decoded BER) performance. However, the convolutional code and Viterbi decoder still provide a significant BER improvement even at short constraint lengths. Longer codes provide lower decoded BER at the cost of greater processing and memory requirements.

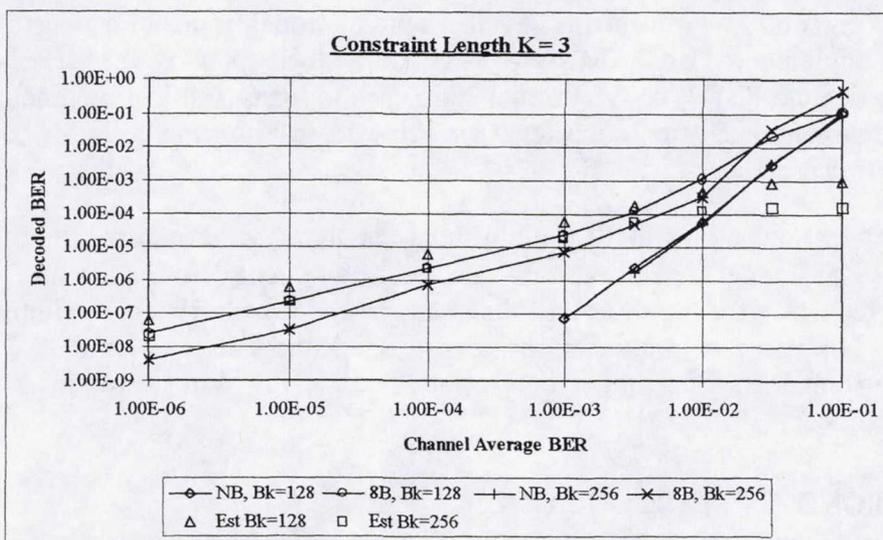


Figure 4-1. Viterbi Decoder Performance vs. Burst Errors and Short Blocks

Implementing the Viterbi decoder on the TMS320C54x series DSP is simplified through a special function designed especially for this decoder. Key algorithms are found in a TI application report "Viterbi Decoding Techniques in the TMS320C54x Family." [6] Further explanation is available in the "TMS320C54x DSP Reference Set, Volume 4: Application Guide" in Section 7.2. [7] Also, a complete Viterbi decoder assembly code and application report [8] is available for the earlier TMS320C5x processor. The 'C5x processor does not have the 'C54x special instruction, but the memory handling and processing algorithms are similar. All the TI application notes and source listings discussed here are available on the TI web pages at "<http://www.ti.com/sc/docs/apps/dsp>".

The TI application report "Viterbi Decoding Techniques in the TMS320C54x Family" [6] also provides benchmarks for determining the processing speed (in MIPS) required to implement a Viterbi decoder. This benchmarks considers, the code rate R , constraint length K , code puncturing rate PR , frame size FS before coding (# data bits) and the number of frames per

second *FR*. The benchmarks for rate $\frac{1}{2}$ convolutional codes with various constraint lengths were calculated. Recall that the frame period was calculated as the time required for 16 samples at 8,000 samples per second. Therefore the frame rate is 1/(2 ms.) or 500 frames per second. For 16-bit PCM, the frame size will be 256 bits. The calculated benchmarks are listed in Table 4-1.

Code Rate, R	Constraint Length, K	Required MIPS
$\frac{1}{2}$	3	3.9
$\frac{1}{2}$	5	7.84
$\frac{1}{2}$	7	24.0
$\frac{1}{2}$	8	45.0

Table 4-1. Benchmarks for Viterbi Decoder on TMS320C54x

The benchmarks in Table 4-1 reveal that even a convolutional code with constraint length as high as 8 can be implemented on a TMS320C54x with a clock speed of 60 MHz. Note that the $K = 8$ code is the defacto (2,1,7) convolutional code used in many applications including satellite and mobile communications, often concatenated with a Reed-Solomon code for combined burst and random bit error protection.

For voice communications alone the FEC algorithms may be used to increase the operating range and reduce noise at longer operating ranges. A rate $\frac{1}{2}$ convolutional code can provide significant error reduction for the short packets in the presence of burst errors (implemented with an interleaver). For voice communications, constraint lengths as low as $K = 3$ will provide significant improvements, while longer constraint length codes will be required for high data reliability applications.

5. CONCLUSIONS AND DISCUSSIONS

The prototype full duplex, spread spectrum voice communication system specifications are:

- Data Rate: 11 Mbps
- Voice Coding: 14-bit PCM (padded to 16 bits) at 8,000 samples per second
- Packets: 16 samples = 256 bits
- Frame Rate: 1 / 2 ms = 500 Hz
- TDMA Slots: 9 slots (8 users, 1 CC) per frequency (0.22 ms per slot)
- Frequencies: 3 frequency bands
- Maximum Users: 24

The prototype system will initially be implemented without forward error correction. If test results warrant or data transmission is added to the system, the FEC can be added at a later time. If rate $\frac{1}{2}$ FEC coding is added, then the packets size will increase to 512 bits, the number of TDMA slots per frequency will reduce to 8 (7 users, 1 central controller), and the maximum number of users will reduce to 21.

The prototype system currently under construction can provide full duplex voice communication for groups of up to 24 people. The system will have 3 frequency channels that can be configured for 3 separate talk groups of up to 8 users or combined into a single 24-person talk group. The system can be operated with open microphones, push-to-talk or combinations of open and push-to-talk users. Potential applications include Ad Hoc wireless communications, remote communication where other radio systems do not exist, and transportable communications. The full duplex system is currently configured to be a standalone communication system.

This system can be extending by interfacing the central controller to other existing communication systems such as the Operational Intercommunication System – Digital (OIS-D), or T1 (or T3) wired or wireless links. The wireless operating range can be extended by increasing the transmit powers or raising the height of the central controller's antenna. More time slots can be made available by using rapid acquisition techniques such as SAW filters [9], or by reducing the data rates required for voice communication through vocoders such as [10]. Future versions of this system can employ advanced channel management techniques to allow more users to share the existing time/frequency slots and thus increase the number of users.

REFERENCES

- [1] Abrahams, Richard L., "2.4 GHz 11 Mbps MACless DSSS Radio HWB1151 User's Guide," Application Note 9835.1, Harris Semiconductor, 1999.
- [2] "HFA3860B Direct Sequence Spread Spectrum Baseband Processor," Data Sheet, File Number 4595, Harris Semiconductor, December 1998.
- [3] Zyren, Jim and Petrick, Al, "Tutorial on Basic Link Budget Analysis," Application Note AN9804.1, Harris Semiconductor, June 1998.
- [4] Lin, Shu and Costello, Jr., Daniel J., *Error Control Coding: Fundamentals and Applications*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1983.
- [5] Wicker, Stephen B., *Error Control Systems for Digital Communication and Storage*, Prentice-Hall, Inc., Upper Saddle River, N.J., 1995.
- [6] Hendrix, Henry, "Viterbi Decoding Techniques in the TMS320C54x Family," Application Report, Literature Number SPRA071, Texas Instruments, June 1996.
- [7] "TMS320C54x DSP Reference Set, Volume 4: Application Guide," Literature Number SPRU173, Texas Instruments, October 1996
- [8] Chishtie, Mansoor A., "U.S. Cellular Error-Correction Coding Algorithm Implementation on the TMS320C5x," Application Report, Literature Number SPRA137, Texas Instruments, October 1994. (Also Viterbi Decoder assembly program listing.)
- [9] "STEL-2000A Digital, Fast Acquisition, Spread Spectrum Burst Processor," Stanford Telecom, Web Site: www.stelhq.com/index.htm.
- [10] Preliminary Data Sheet for the AMBE-2000 Vocoder Chip, Digital Voice Systems, Inc., Burlington, MA, Web Site: www.dvsinc.com.

1999 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

**JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA**

MARS UMBILICAL TECHNOLOGY DEMONSTRATOR

Nasser Houshangi, Ph.D., P.E.
Associate Professor of Electrical Engineering
Department of Engineering
Purdue University Calumet

NASA/KSC Colleague: Tom Lippitt
Lead/ Automated Ground Support Systems Lab
Robotics and Automation

ABSTRACT

The objective of this project is to develop a autonomous umbilical mating for the mars umbilical technology demonstrator. The Mars Umbilical Technology Demonstrator (MUTD) shall provide electrical power and fiber optic data cable connections between two simulated mars vehicles . The Omnibot is used to provide the mobile base for the system. The mate to umbilical plate is mounted on a three axis Cartesian table, which is installed on the Omnibot mobile base. The Omnibot is controlled in a teleoperated mode. The operator using the vision system will guide the Omnibot to get close to the mate to plate. The information received from four ultrasonic sensors is used to identify the position of mate to plate and mate the umbilical plates autonomously. A successful experimentation verifies the approach.

1. INTRODUCTION

The goal of this project is to design and implement a smart umbilical with a connect and re-connect capability with remote interface verification. The ability to quickly and reliably mate and de-mate umbilical connections under automated control is needed in flights to mars [1]. After review of mars reference mission [1], three umbilicals (i) electrical power/fiber optic data cable umbilical; (ii) mars rover cryogenic servicing umbilical; and (iii) contingency in-situ propellant production water supply umbilical were identified. The electrical power/fiber optic data umbilical is mission critical since it must provide immediate power to the hab/crew vehicle and was chosen for a technology demonstration project.

The Mars Umbilical Technology Demonstrator (MUTD) shall provide electrical power and fiber optic data cable connections between two simulated mars vehicles (cargo vehicle and hab/lab vehicle) spaced 100 meters apart. The MUTD shall deploy 100 meters of electrical and fiber optic cables from the simulated hab/lab to the simulated cargo vehicle. The MUTD shall align the surface and the vehicle umbilical plates and parallel mate the connectors in an accurate manner.

The Omnibot mobile base system [12] at automated ground support system laboratory at Kennedy Space Center will provide the mobile base for the MUTD. The Omnibot is controlled in a teleoperated mode. The umbilical plate is mounted on a three axis Cartesian table on the Omnibot. The table is used to mount the two umbilical plates.

The problem of automated umbilical mating can be divided into two subproblems, global and local positioning. The global positioning is defined as when the two umbilical plates do not have any vertical common projection on each other, and the problem is to bring them close enough such as they do. If two plates have any common vertical projection on each other, the problem to mate them from now on is referred to as the local positioning.

For global positioning, a teleoperated mode of control is used to move the Omnibot such that the umbilical plates are in local positioning zone. The operator uses vision feedback to command the Omnibot during global positioning and bring the mate in plate to about 14 inches from mate to plate. The work explained in this report will concentrate on the local positioning problem.

The umbilical mating is performed autonomously during local positioning. During local positioning, the position and orientation of mate to plate needs to be determined. Ultrasonic sensors are used to find the position of mate to umbilical plate. At current stage, the project requirement is to have a system functional at Kennedy Space Center. Therefore, the mars environment is not considered.

To identify the position of the plate to mate to, four ultrasonic sensors are installed in the four corner of mate in plate as shown in figure 1. The sensors are numbered clockwise.

Ultrasonics sensors are used extensively in mobile robots [2-7] applications. Sonar sensing, is cost effective, relatively quick in response, processing is not time consuming, and can cover wider distance range. The edge 1 (E1) is defined across sensors 1 and 2 as shown in figure 1. The following steps outline the procedure for local positioning of autonomous umbilical mating:

1. Obtain raw distance data $d_1, d_2, d_3,$ and d_4 from the four ultrasonic sensors
2. Filter and separate the data according to constraints
3. Engage and apply least square estimation procedure to obtain the line segments
4. Based on the obtained edge position, the manipulator on the Omnibot is autonomously moved to align two umbilical plates and mate them successfully (ie. all four edges of the two plates line up and connect with each other).

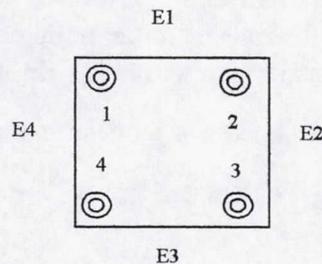


Figure 1. Location of four ultrasonic sensors

This report summarizes the work completed during the summer of 1999. The use of ultrasonic sensors is detailed in Section 2. Section 3 discusses the determination of plate position using the four ultrasonic sensors. The experimental set up and results are presented in Section 4. The discussions and conclusions are presented in Sections 5 and 6, respectfully.

2. ULTRASONIC SENSORS

The ultrasonic sensors work on the principle of sound. Sound is a sequence of waves of pressure which propagates through compressible media such as air or water. Sounds can propagate through solids as well, but there are additional modes of propagation. During their propagation, waves can be reflected, refracted, or attenuated by the medium. A relationship between density and pressure, affected by temperature, determines the speed of sound within medium. The viscosity of the medium determines the rate at which sound is attenuated. For many media, such as air and or water, attenuation due to viscosity is negligible. More information about ultrasonic sensors and their properties are found in [9,10].

The ultrasonic sensors could be used to measure and detect distances to moving objects, impervious to target materials, surface and color; solid -state units have virtually unlimited, maintenance-free lifespan; detects small objects over long operating distances;

and resistant to external disturbances such as vibration, infrared radiation, ambient noise and EMI radiation . Ultrasonic sensors are not affected by dust, dirt or high-moisture environments.

The effect of temperature change [11] that should be considered in calibration is indicated by

$$ADC = 0.0019 \times D \times \Delta F \quad (2)$$

where ADC is the apparent distance change, D is the distance from transducer to target in inches, and ΔF is change in temperature in Fahrenheit.

The ultrasonic waves transmitted by the piezoelectric ring have approximately a cone shaped band [11]. The distance is calculated based on time of flight. To calculate the distance more accurately the sensor needs to be modeled. An approximate mathematical models for the ultrasonic sensors are shown in figure 2. The desired distance D_i depends on the incident angle θ and sensor beam angle β . Assume $\theta \geq \frac{\beta}{2}$ as shown in figure 2a. The desired distance is calculated by

$$D_i = \frac{d_i \cos(\theta - \frac{\beta}{2})}{\cos \theta} \quad (3)$$

The angle θ can be calculate by the law of cosines and sine's by

$$x^2 = s_d^2 + \delta_d^2 - 2s_d\delta_d \cos(\frac{\pi}{2} - \frac{\beta}{2}) \quad (4)$$

and

$$\frac{\delta_d}{\sin \theta} = \frac{x}{\sin(\frac{\pi}{2} - \frac{\beta}{2})} \quad (5)$$

Otherwise if $\theta < \frac{\beta}{2}$, as shown in figure 1b , the incident angle θ needs to be determined by

$$\frac{\delta_d}{\sin \theta} = \frac{s_d}{\sin \frac{\pi}{2}} \quad (6)$$

The desired D_i is then calculated by

$$D_i = \frac{d_i}{\cos \theta} \quad (7)$$

Based on the distance information, it is determined which sensors are engaged with the plate during local positioning by

$$D_i - D_s \leq D_t \quad (8)$$

where D_i is the distance calculated from the i th sensor, D_s is the shortest distance read by any sensor, and D_t is the threshold distance. If equation (8) holds for the i th sensor, the i th sensor is considered to be engaged.

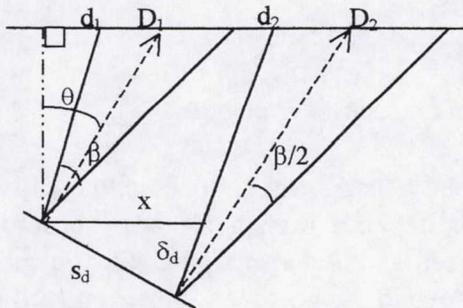


Figure 2a. Ultrasonic mathematical model: $\theta \geq \frac{\beta}{2}$

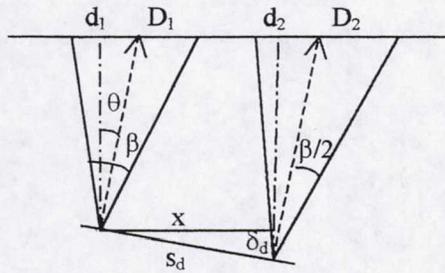


Figure 2b. Ultrasonic mathematical model: $\theta < \frac{\beta}{2}$

3. POSITION DETERMINATION OF THE UMBILICAL PLATE

As was indicated earlier the position and orientation of mate to plate needs to be determined. It is assumed that the orientation is known and fixed. The position of the plate to mate to is calculated based on identifying the edge feature of the plate. The procedure to determine the edges of the plate depends on which sensor is engaged as shown in the flowchart in figure 5.

Possible situations when one sensor is engaged are shown in figure 3. For this scenario the motion depends on which sensor is engaged. For example, if sensor 1 is engaged as shown in figure 3a, the strategy to engage four sensors are to move in the positive z-

direction(up) to engage sensor 4. Then, the plate is moved in the negative x-direction (left) until all four sensors are engaged.

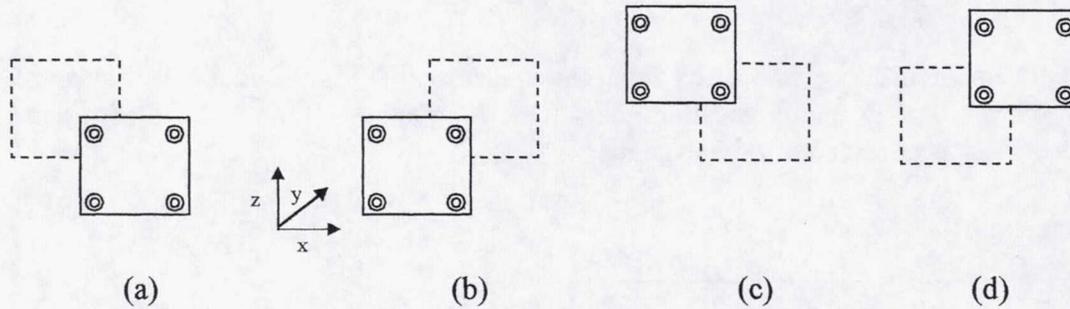


Figure 3. Possible situations when one sensor is engaged

As shown in figure 4, four possibilities also exist for two sensors to be engaged. The strategy to move such as to engage four sensor depends on which two sensors are engaged. For example if sensors 1 and 2 engaged as shown in figure 4a, the motion command to the 3-axis Cartesian table will be to move in the positive z-direction (up) until all four sensors are engaged.

With the assumption that there is not any orientation mismatch around the y-axis between the two plates around the y-axis, the scenario where three sensors engaged should not happen. In case that happens the strategy is to move the plates closer to each other in the y-direction and recheck the situation.

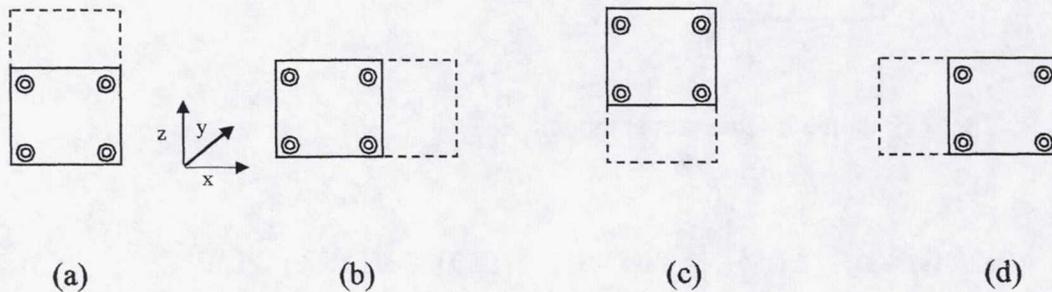


Figure 4. Possible situations when two sensors are engaged

The objective of the approach is to engage all four sensors engaged and move the plates to about four inches apart in the y-direction as indicated in the flow chart shown in figure 5. At this point the algorithm to find the edges of the plates are initiated.

To identify vertical edges vertical edges E2 and E4 shown in figure 1, the manipulator is moved in the x-direction until sensors 2 and 3 are disengaged. This position is marked as

$x_{pos_{23}}$. Then the manipulator is moved in the negative x-direction monitoring the sensors until sensors 1 and 4 miss the plate. This position is also marked as $x_{pos_{14}}$. All positions are determined with respect to the home coordinate frame of the table. With the obtained information one point on both vertical edges can be estimated. The same procedure is repeated in the y-direction to estimate a point on the top and bottom edges (E1 and E3) of the plates.

The least square algorithm is used to find the best fit based on the estimated positions. To estimate a line uniquely, with orientation uncertainty, at least two points are needed. To obtain the second point the mentioned procedure is repeated.

4. EXPERIMENTAL SET UP AND RESULTS

An experiment conducted to verify the feasibility of the proposed approach. The objective of this experiment is to align two plates autonomously using ultrasonic sensors during local positioning. The experimental set up is shown in figure 6. In this experiment, four Migatron ultrasonic sensors installed in the four corner of mate in plate are used. The sensors analog output is connected to a multifunction I/O card with 16-bit analog input. The location of receiving plate is fixed. The dimensions of both plates are 12"×12" with the thickness of 1/8". The plates are made of aluminum. The other umbilical plate is mounted on a three axis Cartesian table on the Omnibot mobile base.

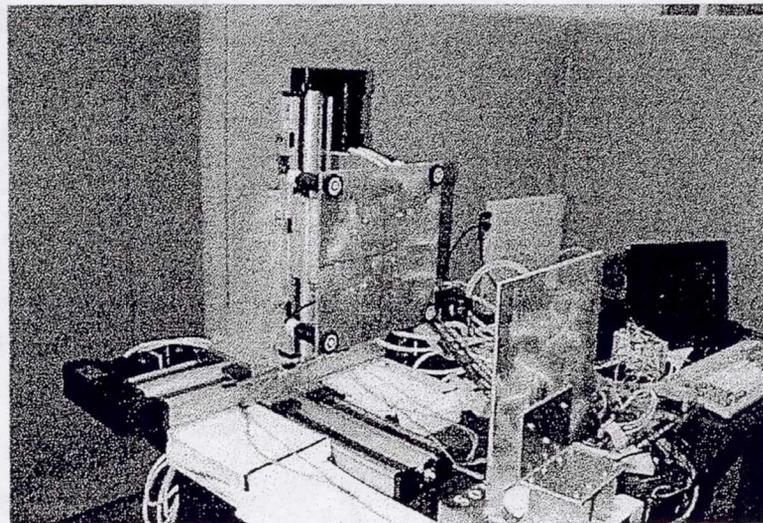


Figure 6. Experimental Set Up

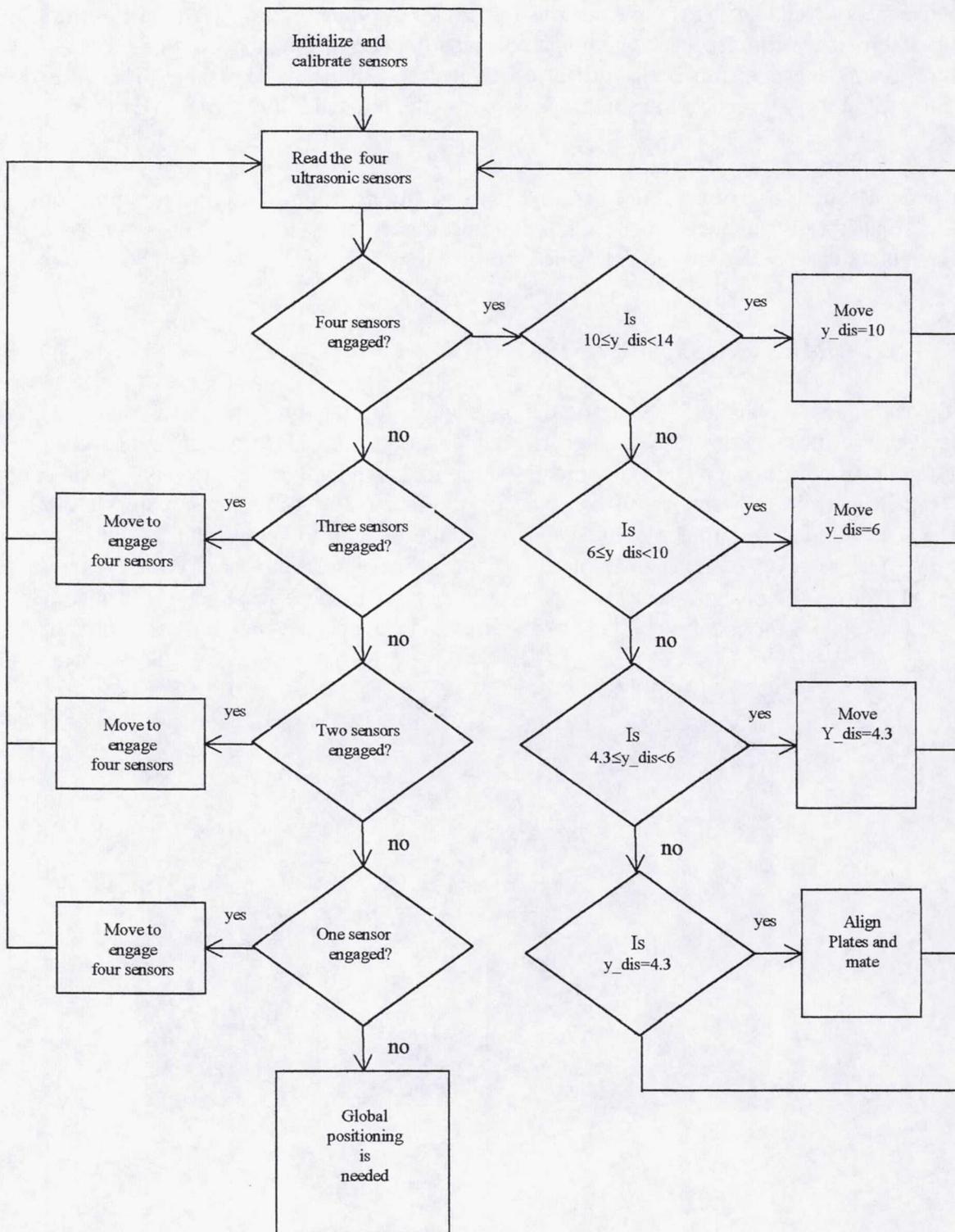


Figure 5. Local positioning software flowchart

A Migatron narrow beam band analog ultrasonic sensor mounted in a self-contained barrel housing is used in this project. It is powered by 20-30 VDC and is reverse polarity protected. The frequency of operation of the sensor is 212 kHz for the 4" to 40" range. The 0-10 VDC analog output of the sensor is connected to 16-bit resolution Versallogic I/O card. The sensor output can easily be scaled over range of 4"-40" by adjusting the zero and the span control.

The Omnibot mobile base [12] is controlled remotely by a operator using a control box. A joystick, an RF transmitter, and number of switches (power, low/high speed, E-stop, and a deadman) make up the control box. The communication between the control box and the mobile base computer is established using an RF modem. The STD-32 star multiprocessing computer system is used to control the wheels of the mobile base and the 3-axis Cartesian table in conjunction with a motion controller. The software program is written using C language on a DOS operating system platform. The operator using the vision feedback [12] will bring the Omnibot within 14 inches of the umbilical plate. Then, the operator will switch the system to the autonomous mode from the command base.

The program is written to perform the autonomous mating based on the flow chart shown in figure 5. The four sensors are engaged and the plate edges determined as described in Section 3. The edges are estimated within ± 0.05 inches by using only one point on each edge. By repeating the edge finding for 10 times, the edges can be estimated within ± 0.042 inches. We considered the improvement to be minimal and decided to perform the autonomous aligning based on one point.

5. DISCUSSIONS

Other sensors like vision and laser range finder could have been used for this project. The objective of the visual sensing is to locate the receiving umbilical plate using appropriate features. Edge detection techniques could be used to determine plate location in 2-D space. Problem with using vision system is that natural lighting and shadows can distort the information received. The alternative to overcome this problem is active techniques of acquiring images. Images may be acquired by active techniques like ultrasound and laser projection. The ultrasonic sensors are simple to use, provide wider view, and more cost effective. For the future work, considering mars environment, the same approach can be implemented by using the laser sensors.

In the final stage of the project, there will be a passive compliance build into umbilical plate with guided cones. This arrangement will insure to avoid excessive forces during umbilical plates mating and insure the proper alignment. As experimentation showed the umbilical plates are aligned within 0.05 inches. Being able to align the plate this accurately

means the guided cones will be small. The passive compliance will also be designed to take care of couple of degree of orientation misalignment in the system.

6. CONCLUSIONS

As was mentioned earlier, the positioning problem of mars umbilical technology demonstrator was divided into global and local positioning control problems. The teleoperated controlled Omnibot [12] provided the solution for the global positioning. My main emphasis was autonomous local positioning which is the main contribution made during the summer of 1999. The system hardware and software platform for local positioning was designed and implemented. The approach verified successfully by experimentation and achieved the goal of autonomous local positioning within specification. As was indicated in the introduction, the ultrasonic sensors are extensively used in mobile robot application for collision avoidance. To the best of the author knowledge, there is not any previous work using the ultrasonic sensors for the fine positioning as explained in this report.

For future work, the following tasks are to be completed for this project: (i) Integrating the three axis Cartesian table hardware and software with Omnibot mobile system hardware and software; (ii) Checking the overall positioning system; (iii) Applying experience gained and concept to the mars environment. Another possible application of this work is automated umbilical mating for the shuttle launch. Umbilicals represent a significant portion of the infrastructure and operational costs associated with launch vehicle turn around. The ability to quickly and reliably mate umbilical connections under automated control would reduce the time and labor hours required preparing for launch.

ACKNOWLEDGEMENTS

I like to express my sincere appreciation to my NASA-KSC colleague Tom Lippitt for his help and useful suggestions during this summer and Denise Mandez for her contribution on writing the software for the project. I also like to thank Ramon Hosler, Greg Buckingham, Jane Hodges, and Judy Gilliam for their contributions to a successful Summer Faculty Program. I learned a lot about Kennedy Space center facilities, operations, and goals. The people in automated ground system Bill Jones, Kim Ballard, Andy Bradley, Mike Hogue, Robert Morrison, and Felix Soto-Toro provided much help when it was needed. Many thanks to all of you for helping to make this summer both enjoyable and professionally interesting.

REFERENCES

- [1] NASA Mars Reference Mission Version 3.0, Kennedy Space Center

- [2] K. Song, and W. Tang, "Environment Perception for a Mobile Robot Using Double Ultrasonic Sensors and a CCD Camera", IEEE Transactions on Industrial Electronics, Vol. 43, No. 3, June 1996.
- [3] J. Borenstein, and Y. Koren, "Obstacle Avoidance with Ultrasonic Sensors", IEEE Journal of Robotics and Automation, Vol. 4, No. 2, April 1998.
- [4] J. L. Crowley, "World Modeling and Position Estimation for a Mobile Robot using Ultrasonic Ranging", Proceedings of the IEEE International Conference on Robotics and Automation, pp. 674-680, 1989.
- [5] J. J. Leonard, and H.F. Durrant-Whyte, "Directed Sonar Sensing for Mobile Robot Navigation", Kluwer Academic Publishers, 1992.
- [6] O. Wijk, P. Jenesfelt, and H. Christensen, "Triangulation Based Fusion of Ultrasonic Sensor Data", Proceedings of IEEE International Conference on Robotics and Automation, pp. 3419-3424, 1998.
- [7] B. Ayrulu, B. Barshan, and S. Utete, "Target Identification with Multiple Logical Sonars using Evidential Reasoning and Simple Majority Voting", Proceedings of IEEE International Conference on Robotics and Automation, pp. 2063-2068.
- [8] R. C. Luo, and M. G. Kay, "Multisensor Integration and Fusion in Intelligent Systems", IEEE Transaction on System, Man, and Cybernetics, Vol. 19, No. 5, pp. 901-931, September/October 1989.
- [9] J.A. Kleppe, "Engineering Applications of Acoustics", Artech House, 1989.
- [10] H. Kuttruff, "Ultrasonics Fundamentals and Applications", Elsevier Applied Science,
- [11] Ultrasonic Sensors, Migatron Corp Catalog #500, 1999.
- [12] N. Houshangi, and T. Lippitt, "Omnibot Mobile Base for Hazardous Environment", 1999 IEEE Canadian Conference on Electrical and Computer Engineering, Edmonton, Alberta, Canada, pp. 1357-1361, May 1999.
- [13] N. Houshangi, "Mars Umbilical Technology Demonstrator", Final report delivered to Tom Lippitt, 1999 NASA/ASEE Summer Faculty Fellowship Program, NASA Kennedy Space Center, 1999.

1999 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA

NOISE REDUCTION USING FREQUENCY SUB-BAND BASED ADAPTIVE
SPECTRAL SUBTRACTION

Dr. David Kozel
Associate Professor of Electrical Engineering
Purdue University Calumet
Department of Engineering
Hammond, IN

NASA/KSC
NASA Colleague: Frederick McKenzie
Communications/RF and Audio

July 31, 1999

Abstract

A frequency sub-band based adaptive spectral subtraction algorithm is developed to remove noise from noise-corrupted speech signals. A single microphone is used to obtain both the noise-corrupted speech and the estimate of the statistics of the noise. The statistics of the noise are estimated during time frames that do not contain speech. These statistics are used to determine if future time frames contain speech. During speech time frames, the algorithm determines which frequency sub-bands contain useful speech information and which frequency sub-bands contain only noise. The frequency sub-bands, which contain only noise, are subtracted off at a larger proportion so the noise does not compete with the speech information. Simulation results are presented.

1. Introduction

It is desired to incorporate adaptive noise suppression into the communications equipment on the Emergency Egress Vehicle and the Crawler-Transporter. In the case of the Emergency Egress Vehicle, the spectral content of the noise source changes as a function of the speed of the vehicle and its engine. In the case of the Crawler-Transporter, the noise a person hears will vary with his location relative to the Crawler-Transporter and if the hydraulic leveling device on the Crawler-Transporter is being used. Due to the varying nature of the noise, an adaptive algorithm is necessary for both applications. Furthermore, the noise frequencies produced by both applications are in the voice band range, so standard filtering techniques will not work. To remove noise from a noise-corrupted speech signal; a frequency sub-band based adaptive spectral subtraction algorithm is developed. In the following sections, a brief overview of spectral subtraction and its limitations is given, the frequency sub-band based adaptive spectral subtraction algorithm is described in detail along with the advantage to using frequency sub-bands, and simulation results are presented and discussed.

2. Spectral Subtraction

Spectral subtraction assumes that noise-corrupted speech is composed of speech plus additive noise.

$$x(t) = s(t) + n(t) \quad (1)$$

Where:

$x(t)$ = noise-corrupted speech

$s(t)$ = speech

$n(t)$ = noise

Taking the Fourier Transform of equation (1),

$$|X(f)|e^{j\theta_x} = |S(f)|e^{j\theta_s} + |N(f)|e^{j\theta_n} \quad (2)$$

When no reference microphone is used, the magnitude and phase of the noise are unavailable when speech is present. The phase of the noise-corrupted speech is commonly used to approximate the phase of the speech. This is equivalent to assuming that the noise-corrupted speech and the noise are in phase. The average magnitude of the

noise, $|\hat{N}(f)|$, is usually used to approximate the magnitude of the noise. Since the noise spectrum will in general have sharper peaks than the average noise spectrum, a multiple, μ , of the average noise spectrum is subtracted. This is done to reduce "musical-noise" which is caused from these random peaks. Solving for the estimated speech spectrum,

$$|\hat{S}(f)|e^{j\theta_x} = (|X(f)| - \mu|\hat{N}(f)|)e^{j\theta_x} \quad (3)$$

The inverse Fourier Transform yields the estimated speech:

$$\hat{s}(t) = \mathcal{F}^{-1}\{\hat{S}(f)\} \quad (4)$$

2.1 Limitations of Spectral Subtraction

When using any algorithm, it is important to understand its limitations and restrictions. Since the noise and speech have no physical dependence, the assumption that the noise and speech are in phase at any or all frequencies has no basis. Rather, they can be thought of as two independent random processes. The phase difference between them at any frequency has an equal probability of being any value between zero and 2π radians. Thus, the noise and speech vectors at one frequency may add with a phase shift while simultaneously at a different frequency may subtract with a different phase shift. Thus, subtracting an assumed in-phase noise signal from the noise-corrupted speech has the same probability of reducing the particular frequency component of the speech even further as it does of bringing it back to its proper level. Furthermore, it is almost certain to cause some distortion in the phase. The amount of error produced at each frequency depends upon the relative phase shift and the relative magnitudes of the speech and noise vectors. As noted in [1], for each spectral frequency that the magnitude of the speech is much larger than the corresponding magnitude of the noise, the error is negligible. For the consonant sounds of relatively low magnitude, the error will be much larger. This is true even if the magnitude of the noise at each frequency could be exactly determined during speech.

3. The Value of Sub-bands

For a given range of frequencies, say zero to six kilohertz, each speech sound is only composed of some of the frequencies. No sound is composed of all of the frequencies. If the spectrum is divided into frequency sub-bands, the frequency sub-bands containing just noise can be removed when speech is present. Furthermore, during speech the power level of the frequency sub-bands that contain speech will increase by a larger proportion than the power level of the entire spectrum. Thus, speech will be easier to detect by looking at the sub-band power change than by looking at the overall power change. This is especially true of the consonant sounds, which are of lower power, but are concentrated in one or two frequency sub-bands. By dividing the signal into frequency sub-bands, frequency bands that do not contain useful information can be removed so that the noise in those frequency sub-bands does not compete with the speech information in the useful sub-bands.

3.1 Adaptive Spectral Subtraction Algorithm

Details of the frequency sub-band based adaptive spectral subtraction algorithm are described in this section. The signal is sampled, windowed with a hamming window, and zero padded by the same procedure described in [1]. Each time frame of signal overlaps the previous time frame by 50 percent. An "m" point Fast Fourier Transform is taken, and the magnitude of the frequency response is separated from the phase angle. The magnitude response is partitioned into frequency sub-bands as shown in Table 1. The

range of frequencies in each sub-band is chosen in accordance with the bark scale [2] to account for the hearing characteristics of the human ear.

Sub-band	Start Bin	Stop Bin	Number of Bins	Beginning Frequency (Hz)	Ending Frequency (Hz)
1	1	16	16	0	388
2	17	21	5	388	505
3	22	26	5	505	622
4	27	32	6	622	763
5	33	38	6	763	904
6	39	45	7	904	1069
7	46	53	8	1069	1257
8	54	62	9	1257	1468
9	63	72	10	1468	1703
10	73	84	12	1703	1985
11	85	98	14	1985	2314
12	99	114	16	2314	2690
13	115	133	19	2690	3136
14	134	156	23	3136	3676
15	157	186	30	3676	4381
16	187	224	38	4381	5273
17	225	256	32	5273	6025

Table 1. Frequency Ranges of the Frequency Sub-bands

To key into the communication system, the user is required to press and hold a push-to-talk button while speaking into the microphone. Thus, it is assumed that speech is not present when the push-to-talk is not pressed. For each time frame, L , when the push to talk is not pressed, the signal is just noise.

$$|X_L(kf)| = |N_L(kf)| \quad \text{for frequency bins } k = 1, \dots, m \quad (5)$$

While the push-to-talk is not pressed, the statistics of the noise are determined, and the algorithm is initialized. The statistics of the noise are updated every n_A time frames until a push-to-talk occurs. n_A is chosen large enough to provide reliable noise statistics and small enough to be updated before each push-to-talk. The average noise magnitude for each frequency bin is determined using the sample mean.

$$\bar{N}(kf) = \frac{1}{n_A} \sum_{L=1}^{n_A} |N_L(kf)| \quad \text{for frequency bin } k = 1, \dots, m \quad (6)$$

The power in frequency sub-band v for time frame L is

$$P_{Lv} = \sum_{k=\beta_v}^{\xi_v} |X_L(kf)|^2 \quad (7)$$

Where β_v and ξ_v are the beginning and ending frequency bins for sub-band v . The average power in frequency sub-band v over the n_A time frames is estimated using the sample mean.

$$P_{Av} = \frac{1}{n_A} \sum_{L=1}^{n_A} P_{Lv} \quad \text{for sub-band } v = 1, \dots, \eta \quad (8)$$

The standard deviation of the power in frequency sub-band v over the n_A time frames is estimated using the square root of the sample variance.

$$\sigma_v = \sqrt{\frac{1}{(n_A - 1)} \sum_{L=1}^{n_A} (P_{Av} - P_{Lv})^2} \quad \text{for sub-band } v = 1, \dots, \eta \quad (9)$$

The threshold proportions for duration and burst speech in each frequency sub-band are dependent on the standard deviation of the power in that frequency sub-band and externally adjustable proportions, α_d and α_b .

$$\tau_{dv} = (1 + \alpha_d \sigma_v) \quad \text{for sub-band } v = 1, \dots, \eta \quad (10)$$

$$\tau_{bv} = (1 + \alpha_b \sigma_v) \quad \text{for sub-band } v = 1, \dots, \eta \quad (11)$$

Once an average value for the noise is determined, the maximum ratio of noise to average noise over the sub-band

$$MR_{Lv} = \max_{k=\xi_v, \dots, \beta_v} \left(\frac{|N_L(kf)|}{|\bar{N}(kf)|} \right) \quad \text{for sub-bands } v = 1, \dots, \eta \quad (12)$$

and the running average of MR_{Lv}

$$AMR_v = (1 - \mu)AMR_v + \mu MR_{Lv} \quad \text{for sub-bands } v = 1, \dots, \eta \quad (13)$$

are determined.

When the push-to-talk is pressed, the algorithm must determine if speech is present during that particular time frame. For each time frame, L , the noise flags for the sub-bands, γ_v , the noise flag counter, γ_C , and the noise flag record vector, γ_R , are initialized to the following values:

$$\gamma_v = 1 \quad \text{for sub-band } v = 1, \dots, \eta \quad (14)$$

$$\gamma_C = 0 \quad (15)$$

$$\gamma_R(1) = 0 \quad (16)$$

Then, for sub-band v ,

$$\text{if} \{ [\text{all } P_v(L, \dots, L+\delta_d) > \tau_{dv} P_{Av}] \text{ or } [\text{all } P_v(L-\delta_d, \dots, L) > \tau_{dv} P_{Av}] \text{ or } [\text{all } P_v(L-\delta_c, \dots, L+\delta_c) > \tau_{dv} P_{Av}] \text{ or } [P_v(L) > \tau_{bv} P_{Av}] \} \quad (17)$$

set

$$\gamma_v = 0 \quad (18)$$

$$\gamma_C = \gamma_C + 1 \quad (19)$$

$$\gamma_R(\gamma_C) = v \quad (20)$$

Equations (17) through (20) are repeated for sub-band $v = 1, \dots, \eta$. In equation (17), the time frame shifts, δ_d and δ_c , required for duration speech are based upon the minimum time duration required for most speech sounds [3, p.62]. The time frame shift, δ_d , is used to detect the beginning and ending of speech sounds. The frame shift, δ_c , detects isolated speech sounds. The burst speech threshold proportion, τ_{bv} , should be larger than the duration speech threshold proportion, τ_{dv} ; but the time required shorter since bursts generally have more energy but don't last as long. Equation (17) looks into the future (i.e., $P_v(L, \dots, L+\delta_d)$) by processing frames of data but holding back decisions on them for δ_d time frames.

After using equation (17) to check all of the sub-bands, if $[(\gamma_C > 1) \text{ or } (\gamma_R(1) > 14)]$, the frame is considered to be a speech frame. During speech frames, the ratio of the sum of noise-corrupted speech to sum of average noise

$$R_{L\nu} = \frac{\sum_{k=\beta_\nu}^{\xi_\nu} |X_L(kf)|}{\sum_{k=\beta_\nu}^{\xi_\nu} |\bar{N}_L(kf)|} \quad \text{for frequency sub-bands } \nu = 1, \dots, \eta \quad (21)$$

is updated. Then, the speech estimate is determined using

$$\begin{aligned} |\hat{S}_L(kf)| &= |X_L(kf)| - \min[R_{L\nu}, \text{AMR}_\nu] (1 + \alpha_p \sigma_\nu) (1 + \alpha_f \gamma_\nu) \bar{N}(kf) \\ &\quad \text{for } \nu = 1, \dots, \eta \text{ and } k = \xi_\nu, \dots, \beta_\nu \end{aligned} \quad (22)$$

If the magnitude of the estimated speech is less than zero for any frequency, it is set equal to zero. In equation (22), the proportion of the average noise subtracted is weighted by the minimum of $R_{L\nu}$ and AMR_ν . $R_{L\nu}$ is large during strong vowel sounds, but small during weaker consonant sounds. AMR_ν is the running average of the proportion needed to remove all of the noise. This proportion will remove too much speech information during weaker consonant sounds. The above weights are multiplied by σ_ν to account for the variation in the noise. The noise flag, γ_ν , increases the proportion subtracted when speech is not present in a frequency sub-band.

If the time frame is not a speech frame, it is a noise frame. During noise frames,

$$|N_L(kf)| = |X_L(kf)| \quad \text{for frequency bins } k = 1, \dots, m, \quad (23)$$

and the following values are updated. The maximum ratio of noise to average noise over each frequency sub-band

$$\text{MR}_{L\nu} = \max_{\text{over } k=\xi_\nu, \dots, \beta_\nu} \left(\frac{|N_L(kf)|}{|\bar{N}(kf)|} \right) \quad \text{for frequency sub-bands } \nu = 1, \dots, \eta. \quad (24)$$

The running average of $\text{MR}_{L\nu}$

$$\text{AMR}_\nu = (1 - \mu)\text{AMR}_\nu + \mu\text{MR}_{L\nu} \quad \text{for } \nu = 1, \dots, \eta. \quad (25)$$

The running average of the power

$$P_{A\nu} = (1 - \mu)P_{A\nu} + \mu P_{L\nu} \quad \text{for frequency sub-bands } \nu = 1, \dots, \eta, \quad (26)$$

and the running average of the noise at each frequency

$$\bar{N}(kf) = (1 - \mu)\bar{N}(kf) + \mu |N_L(kf)| \quad \text{for } k = 1, \dots, m. \quad (27)$$

Also, the estimated speech signal is set to zero.

$$|\hat{S}_L(kf)| = 0 \quad \text{for } k = 1, \dots, m \quad (28)$$

At this point the algorithm checks to see if the push-to-talk is still being pressed. If it is, the process is repeated starting at equation (14). If it is not, the algorithm goes back to the initialization stage, equation (5), to update the statistics of the noise and obtain new threshold proportions.

4. Results and Discussion

The algorithm developed in Section 3 was tested using noise-corrupted speech collected at 12.05 K Hz from the Emergency Egress Vehicle [4]. To generate each time frame, the data was windowed with a hamming window of length 256 points and zero padded to 512 points. Each frame of data overlapped the previous frame of data by 50 percent. The section of data contained the words, “pond”, “key”, “so”, and “wren” chosen from the list given in the Diagnostic Rhyme Test (DRT) [5]. Spectrograms of the original signal containing the noise-corrupted speech and the signal after frequency sub-band based adaptive spectral subtraction are shown in Figure 1.

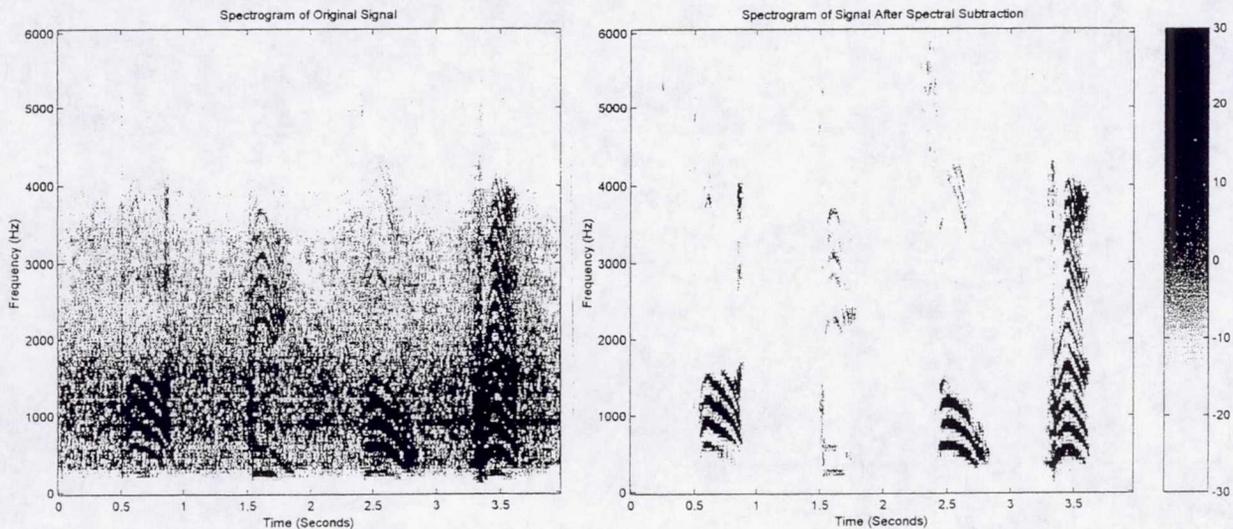


Figure 1. Spectrogram of Original Signal and Signal After Frequency Sub-band Based Spectral Subtraction

The original signal was pre-filtered [6] to compensate for the effects of the anti-aliasing filter, which was required for the A/D converter and the power reduction in speech at higher frequencies [7, p.238]. The ratio of power to average power in the noise-corrupted speech signal for frequency sub-bands 7, 13, and 17 is shown in Figure 2 along with the corresponding long and short term speech power thresholds, the sub-band noise flag, and overall noise flag for each time frame of the data sequence. As can be seen by the power of the signal relative to the power thresholds in each frequency sub-band, one frequency sub-band may contain speech information during a given time frame, while another does not. Frequency sub-band 7 contains the “on” sound of the word “pond”, the “k” sound of the word “key”, and the “o” sound of the word “so” in time frames approximately 50 - 75, 138 - 145, and 230 - 250, respectively. The noise flag for frequency sub-bands 13 and 17 for the same time frames indicate that these frequency sub-bands do not contain speech information during these time frames. Frequency sub-band 13 contains the “d” sound of the word “pond” and the “e” sound of the word “key” in time frames approximately 75 - 85 and 150 - 165, respectively. The noise flag for frequency

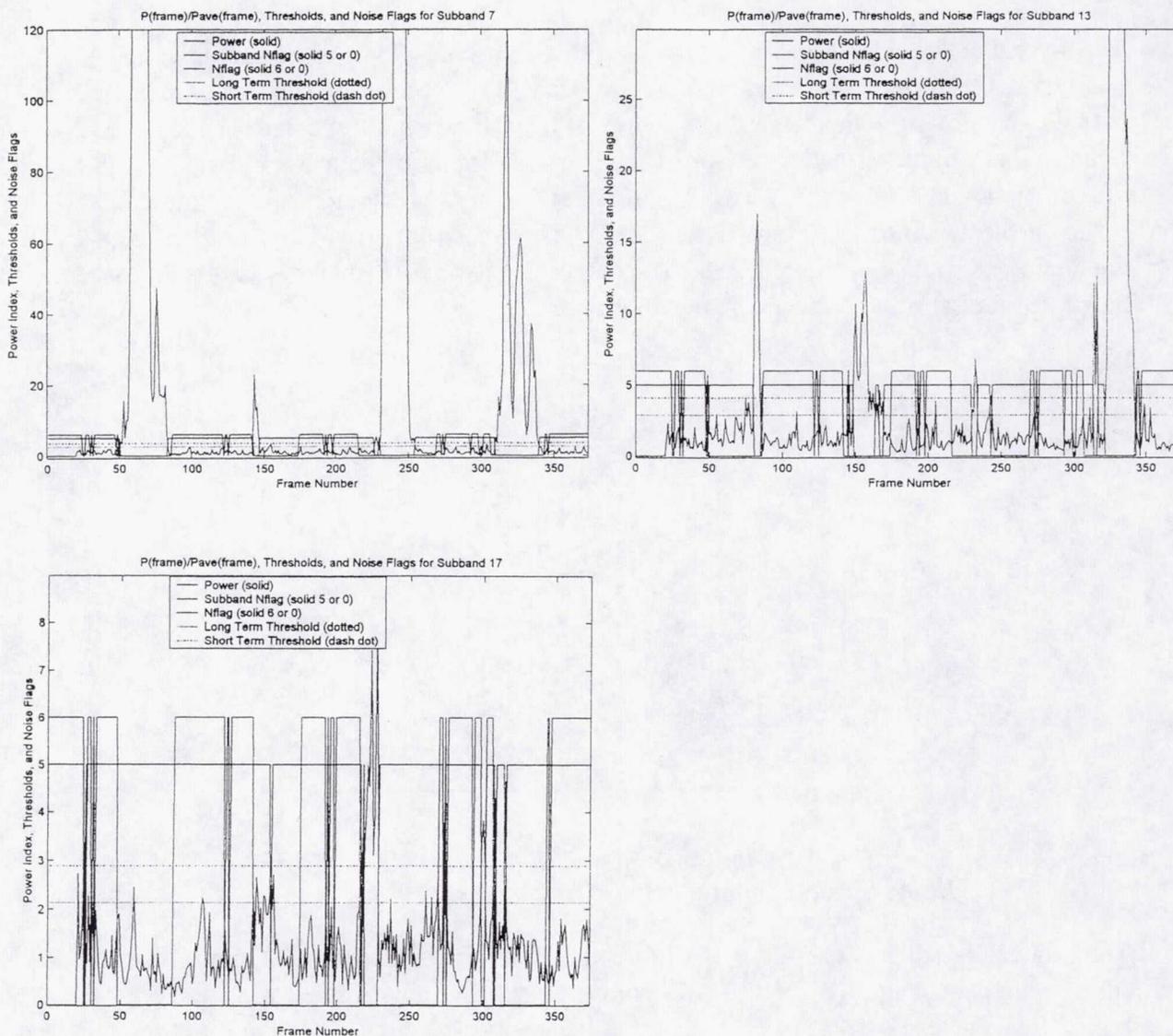


Figure 2. Power/(Average Power), Speech Power Thresholds, and Noise Flags for Frequency Sub-bands 7, 13, and 17, Respectively.

sub-bands 7 and 17 for the same time frames indicate that they do not contain speech information during these time frames. Finally, frequency sub-band 17 contains the “s” sound of the word “so” in time frames approximately 220 - 230. The noise flag for frequency sub-bands 7 and 13 for the same time frames indicate that they do not contain speech information during these time frames. According to equation (17), the noise in the frequency sub-bands that do not contain speech information will be subtracted off at a much greater proportion than the noise in the frequency sub-bands that contain speech during that particular time frame. This is done to essentially remove all noise in frequency sub-bands that do not contain speech information while preserving as much speech information as possible when removing noise from frequency sub-bands that

contain speech information. Comparing the magnitude scales for the different sub-bands in Figure 2, it is apparent that a very small overall relative power increase occurs for some of the consonant sounds such as the “s” in the word so. These power increases would be difficult to detect if sub-bands were not used.

A plot of the noise and average noise as a function of frequency for the final time frame is displayed in Figure 3. It is apparent that a multiple of the average noise must be subtracted from the noise in order to remove the spectral noise peak values. Due to the nature of the noise being considered, these spectral peaks vary in frequency and magnitude from time frame to time frame. When speech is present, the amount of over subtraction for frequency sub-bands containing speech information must be limited or too much of the speech information will be removed with the noise. Figure 4 displays R_{Lv} , MR_v , and AMR_v as a function of time frame for frequency sub-band 13.

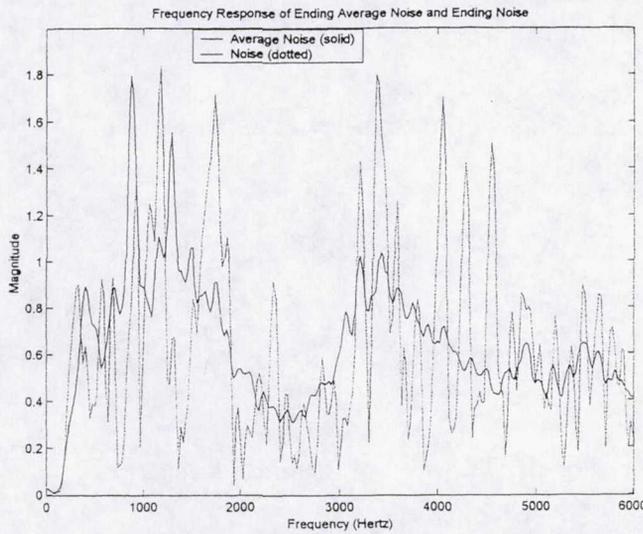


Figure 3. Frequency Response of [Noise] and Average [Noise] for Final Time Frame

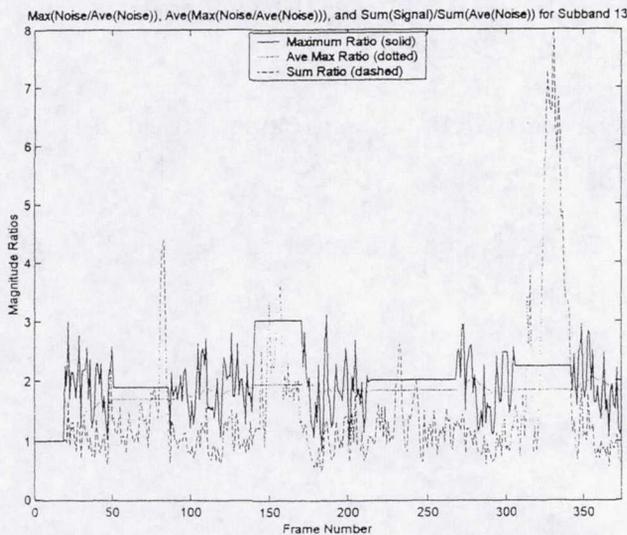


Figure 4. Maximum(|Noise|/(Average |Noise|)), Average Maximum(|Noise|/(Average |Noise|)), Sum(|Signal|)/Sum(Average |Noise|) for Frequency Sub-band 13

The minimum of R_{L_v} and AMR_v is used to limit the amount of noise removed in a frequency sub-band when speech is present.

5. Conclusion

Figure 1 demonstrates that the algorithm removes noise from the frequency sub-bands that do not contain speech information, while preserving the speech information in the frequency sub-bands that contain speech. Places for improvement in the algorithm include an estimate of the ratio of noise power to speech power so that the user would not have to set the parameter, α ; the use of feedback to estimate MR_v so that it does not need to be calculated; and a better estimate of the instantaneous noise when speech is present. All of these goals can be achieved by using multiple microphones.

References

- [1] Kozel, David, NASA/ASEE Summer Faculty Fellowship Program Research Reports: NASA CR-202756; 1996, p143-157.
- [2] E. Zwicker and H. Fastl, Psychoacoustics Facts and Models, Springer-Verlag, 1990.
- [3] Digital Signal Processing Applications with the TMS320C30 Evaluation Module: Selected Application Notes, literature number SPRA021, 1991.
- [4] Kozel, David, NASA/ASEE Summer Faculty Fellowship Program Research Reports: NASA CR-1999-208546; 1998, p103-112.
- [5] Voiers, William, D., Evaluating Processed Speech using the Diagnostic Rhyme Test, Speech Technology, January/February, 1983, p30-39.
- [6] Kozel, David, NASA/ASEE Summer Faculty Fellowship Program Research Reports: NASA CR-207197; 1997, p113-122.
- [7] Davis, Don, and Davis, Carolyn, Sound System Engineering Second Edition, Macmillan Publishing Co., New York, NY, 1987.

1999 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

**JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA**

STRATEGIC PROJECT MANAGEMENT AT THE NASA KENNEDY SPACE CENTER

Jerome P. Lavelle
Associate Professor
Industrial & Manufacturing Systems Engineering
Kansas State University

Jean Flowers, NASA-KSC, PZ-QA

ABSTRACT

This paper describes Project Management at NASA's Kennedy Space Center (KSC) from a strategic perspective. It develops the historical context of the agency and center's strategic planning process and illustrates how now is the time for KSC to become a center which has excellence in project management. The author describes project management activities at the center and details observations on those efforts. Finally the author describes the Strategic Project Management Process Model as a conceptual model which could assist KSC in defining an appropriate project management process system at the center.

STRATEGIC PROJECT MANAGEMENT AT THE NASA KENNEDY SPACE CENTER

Jerome P. Lavelle

1. INTRODUCTION

In the past twenty years many types of organizations have felt the pressure to be more responsive to those parties to whom they are accountable. In the early 1980's it was the Japanese and Germans who forced American industries to look at their attitude toward customer (internal and external) satisfaction. In the late 1980's service industries and product producers in nearly all industrial sectors recognized this need. Later it was the healthcare industry as well as government organizations that heard the call to more effectively manage the processes that produce utility for their customers (again internal and external).

In the case of the National Aeronautics and Space Administration (NASA) the organizational "call to action" was clearly heard in the early 1990's. Drivers such as the new congress, vice-president Gore's re-inventing government initiative, voter expectations for increased accountability and the Government Performance and Results Act all meant that it was time for NASA to act — and act it did. From the early beginnings of the strategic planning process right up to today, NASA has answered the call, and has firmly rooted itself as the key government entity ready to lead the nation's space program into the next millennium.

This paper briefly develops NASA and Kennedy Space Center's (KSC) strategic planning process from the early 1990's to today. It discusses what impact that process has had on the agency and specifically the center and illustrates how project management (PM) is a natural result of that process at KSC. It discusses current PM activity at the center and finally concludes with a description of the Strategic Project Management Process Model (SPMPM) and its uses in the previously developed strategic environment and context.

2. STRATEGIC PLANNING AT NASA AND KENNEDY SPACE CENTER

Strategic planning is a process whereby an organization evaluates its mission, guiding principles and overarching *raison d'être* and establishes its objectives, goals and plans. This process focuses organizations on their internal strengths and weaknesses and the opportunities and threats that exist in their environment. From this, strengths are capitalized on and weaknesses improved, and external opportunities and threats are recognized. Strategic planning focuses an organization on short term, medium term and long term goals and plans, and produces action with the organization's resources to be successful in each of those time frames. When done correctly strategic planning provides a framework from which every activity that the organization engages in ties directly to achieving its goals. In this way all decision makers, all workers, and all people in the organization understand how their own work connects, and has meaning in, the big picture of what the organization is reaching to accomplish.

Strategic planning is a very important organizational tool that has been used by most every major company in most every industrial and service sector. It is a tool that has been used in hospitals, schools and school systems, churches, universities, non-profit and community service and many other types of organizations.

The Government Performance and Results Act of 1993 (GPRA-93) and the National Performance Review were the federal legislative catalysts that precipitated NASA's movement in the direction of agency-wide strategic planning. The Act itself was a formal proclamation to all federal government entities that they need to become more accountable for the nation's resources which they manage. Table 1, from the GPRA-93, illustrates the findings and basic purposes behind the Act. *Findings* in the Act include items such as: waste and inefficiency in programs, insufficient articulation of goals, and insufficient attention to performance and results; and *Purposes* include: improve accountability,

initiate performance reform, focus on results/service quality/customer satisfaction, improve congressional decision making and overall management of the Federal Government.

Table 1

Government Performance and Results Act 1993, Findings and Purposes [1]

SEC. 2. FINDINGS AND PURPOSES.

(a) FINDINGS- The Congress finds that:

1. Waste and inefficiency in Federal programs undermine the confidence of the American people in the Government and reduces the Federal Governments ability to address adequately vital public needs;
2. Federal managers are seriously disadvantaged in their efforts to improve program efficiency and effectiveness, because of insufficient articulation of program goals and inadequate information on program performance; and
3. Congressional policymaking, spending decisions and program oversight are seriously handicapped by insufficient attention to program performance and results.

(b) PURPOSES- The purposes of this Act are to:

1. Improve the confidence of the American people in the capability of the Federal Government, by systematically holding Federal agencies accountable for achieving program results;
2. Initiate program performance reform with a series of pilot projects in setting program goals, measuring program performance against those goals, and reporting publicly on their progress;
3. Improve Federal program effectiveness and public accountability by promoting a new focus on results, service quality, and customer satisfaction;
4. Help Federal managers improve service delivery, by requiring that they plan for meeting program objectives and by providing them with information about program results and service quality;
5. Improve congressional decision making by providing more objective information on achieving statutory objectives, and on the relative effectiveness and efficiency of Federal programs and spending; and
6. Improve internal management of the Federal Government.

Federal Government agencies across the board have been affected by this legislation and there has been a true “re-inventing of government” because of it. Table 2 lists federal programs that have developed case studies from their GPRA-93 initiated strategic planning processes.

Max Weber developed the bureaucratic model of organizational design in the 1940s as a means for standardization and structure in communication, authority, and chain of command [2]. These principles were meant to lead to a more focused, efficient and effective organization, and at that time were new and innovative from a organizational design perspective. However, the word *bureaucratic* itself has been used most recently in a derisive manner to refer to officialism, red tape, proliferation and grid lock in organizational systems. The GPRA-93 was meant to address the very worst of that negative definition.

2.1 STRATEGIC PLANNING AT NASA

NASA’s reaction to the strategic planning call was swift and immediate and has been the process through which the agency has transformed itself under the guidance of Mr. Daniel Goldin, NASA Administrator. With a budget that is now 0.7% of the federal budget (compared to a flush 5.7% during the Apollo days) and under 18,000 employees (versus 31,000 in the Apollo days) Mr. Goldin and the NASA management team has molded today’s NASA into a diverse, results and customer oriented modern organization.

Table 3 shows the journey that NASA has been through to get to today's "better, faster, cheaper" new NASA.

Table 2

Federal Programs that have developed case studies from their GPRA-93 Initiatives [3]

- Inspector General of U.S. Army Audit Agency
- Research and Development Function of the Army Research Laboratory
- U.S. Coast Guard
- Internal Revenue Service
- Department of Veteran Affairs, National Cemetery System
- National Park Service's Denver Service Center
- National Aeronautics and Space Administration
- Department of State Personnel Security/Suitability Division
- Energy Information Administration
- InterAmerica Foundation
- Public Health Service Healthy People Program
- Pension Guarantee Corporation
- National Science Foundation
- Office of Child Support Enforcement
- National Highway Traffic Safety Administration
- Environment Protection Agency
- Bureau of Land Management
- Defense Logistics Agency
- Department of Education
- Social Security Administration
- National Weather Service

Table 3

Journey to the New NASA [4]

- 1915 Congress establishes National Advisory Committee for Aerospace
- 1958 NASA established as part of the National Aeronautics and Space Act
- Developed reputation as a "can do" Agency; successes include: John Glenn's earth orbits, Neal Armstrong's first steps on the moon, landing two Viking spacecraft on Mars, Skylab successes, Shuttle development & missions
- Challenger accident, Post Cold-War drift, Hubble problems, lost Mars Observer
- 1993 Government Performance and Results Act, National Performance Review
- Senior Management Group, Strategic Management Working Group
- NASA Strategic Plan, 25-Year Roadmap
- Strategic Enterprises and Center of Excellence concepts
- Emerged as the New NASA; focused on: development not operations, new frontiers not dead ends, leveraging resources, partnering, and developing value for all stakeholders

Two central concepts that describe the agency from a strategic perspective are Strategic Enterprises and Centers of Excellence. Strategic Enterprises were established to manage the programs and activities of the agency that will implement the mission and "be responsible for answering specific fundamental questions, and satisfy the requirements of NASA's customers." [5]. As given in [5]:

NASA's Strategic Enterprises identify at the most fundamental level what we do and for whom. They focus us on the ends, not the means, of our endeavors. Each of our Strategic Enterprises is analogous to a strategic business unit, employed by private-sector companies to focus on and respond to its customers' needs. Each Strategic Enterprise has a unique set of goals, objectives, and strategies that address the requirements of its primary external customers. However, each Enterprise must ensure synergy with and support of the Agency's common goals and the strategies of the other Enterprises.

NASA has established four Strategic Enterprises as follows: Space Science, Mission to Planet Earth (MTPE), Human Exploration and Development of Space (HEDS), and Aeronautics and Space Transportation Technology (ASTT).

Centers of Excellence are tied to NASA's physical facilities throughout the United States that implement the programs of the agency. These Centers have been established to improve the effectiveness and efficiency of the program and to reduce duplication, overlap and administrative overhead in achieving the goals of the agency. Each physical facility is assigned to lead one or more areas of key competency within the agency. Together the Centers of Excellence describe the body of core competency that the agency possesses for use in achieving its goals and leveraging with customers and suppliers. As in [5]:

Each Center of Excellence represents a focused, Agency-wide leadership responsibility in a specific area of technology or knowledge. Centers of Excellence are chartered with a clear definition of their capabilities and boundaries. They are charged to be preeminent within the Agency, if not worldwide, with respect to the human resources, facilities, and other critical capabilities associated with the particular area of excellence. Each Center of Excellence must maintain or increase the Agency's preeminent position in the assigned area in line with the program requirements of the Strategic Enterprises and the long-term interests of the Agency. The capabilities to support a Center of Excellence can be distributed across multiple Centers. These capabilities are available to all of the Strategic Enterprises.

Figure 1 below describes the eleven Centers of Excellence at NASA and their physical location.

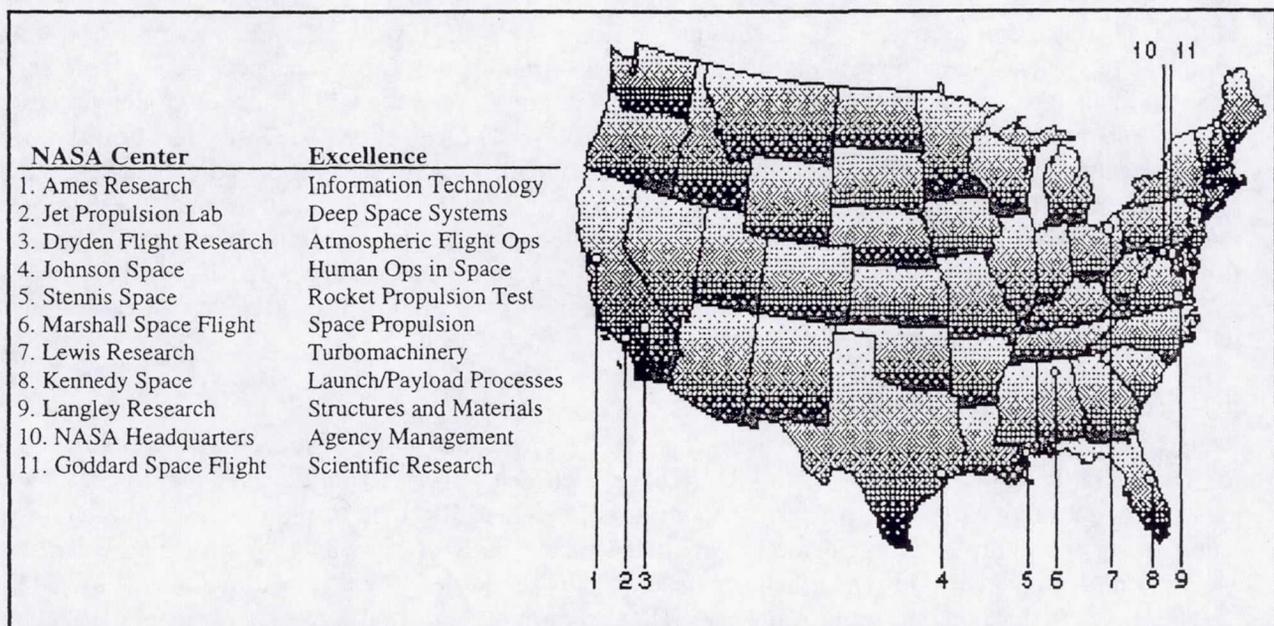


Figure 1: NASA's Centers of Excellence

In general, the strategic planning process at the agency level has been directed toward making NASA: more customer focused, more accountable and more involved with development and resource activities than with sustaining activities. In short NASA seeks to be the developer and creator of space technologies and competency. They seek to be the agency that sets and leads the agenda regarding the how, when, where and why of mankind's exploration and use of space. But at the same time NASA realizes that this must be done in such a fashion that NASA's constituents are satisfied that the agency is a productive and efficient organization worthy of taxpayers investments. By getting out of the space technology management domain and into the space technology innovation domain NASA is taking on that leadership role.

2.2 STRATEGIC PLANNING KENNEDY SPACE CENTER

Kennedy Space Center's role within NASA has been defined through the agency's strategic planning process described above. KSC is assigned to the HEDS Strategic Enterprise which means that all resources, processes, customers and beneficiaries of the output of the center are aligned with human exploration and development of space. Kennedy Space Center is also NASA's *Center of Excellence for Launch and Payload Processing* and the lead center for *Acquisition and Management of Expendable Launch Vehicles, Payload Carriers, Payload Processing and Support*. Thus, from a strategic planning perspective it is clear that Kennedy Space Center exists primarily to be NASA's organizational business unit that maintains knowledge and expertise in space vehicle (and payload) launch and processing competency in support of human exploration and use of space.

Strategic Planning at the agency level has precipitated subsequent planning at the center levels of NASA. These center level plans are required to align with the agency plans in order to achieve overall agency goals. Center plans, by definition will to be more specific and task oriented because they are more focused within the overall agency strategic planning process. Kennedy Space Center has responded to agency plans by developing the KSC Road Map and KSC Implementation Plan [6, 7]. These documents form the basis of the what, when, where, how and why of how Kennedy Space Center will achieve its roles within the agency for the next 25 year period.

As management at KSC looks around it finds itself in a very different environment when compared to 10 years ago. The new environment brings with it "better, faster, cheaper" government and a stronger focus on accountability, efficiency and research and development. The use of center resources must to planned, justified, engaged and measured to support of center goals. Tasks must be defined, budgets must be combed and performance, cost and quality attributes measured whenever center resources are utilized. Kennedy Space Center's current and future environment is one where project management (PM) principles can be applied very effectively. Project management seeks first and foremost to: plan and define work; prioritize how resources are used at the organizational level; schedule and cost work content; control project parameters to support goal attainment; and create an environment where success is no accident.

3. PROJECT MANAGEMENT EXCELLENCE

Project management as a body of knowledge has only very recently been organized and promulgated to the masses, although components of this body have been in use for several decades (centuries). There exist many definitions of what project management is [e.g. 8, 9, 10]. One of the most recent attempts at categorizing and defining an aggregate project management body of knowledge is found in *A Guide to the Project Management Body of Knowledge* developed by the Project Management Institute (PMI) [11].

Ketzner develops an interesting chronological progression of the use of project management which ranges from Traditional Project Management (1960-1985), to Renaissance Project Management (1985-1993), to Modern Project Management (1993-present) [12]. In this chronology, the body of

project management knowledge and its use has matured simultaneously based on the needs of organizations using project management, technology and other driving forces. The Traditional period was dominated by large organizations whose focus was on technical completion and cost, where vast resources were applied to projects. Firms in the aerospace, defense and construction industries dominated the use of project management in this period. Massive programs run on mainframe computers dominated in this period. During the Renaissance period companies from many industries began to understand the power of project management principles for directing resources and achieving their goals. Project management was applied to both small and large projects alike. Multidisciplinary teams were common during this period and more focus was placed on company decisions versus project decisions. Personal computer-based project management software gave ready access to sophisticated planning, scheduling and controlling tools during this period. Lastly, during the Modern Project Management period there has been use of increasingly sophisticated tools and techniques in project management by many types of organizations. Companies using project management today do so at the organizational and qualitative levels and want their clients, employees and all stakeholders to see the value of project management in doing business this way. This approach has spawned the term *modern project management* as a way of differentiating it from the way that project management was perceived in the past [12].

What is excellence in project management? In the text *In Search of Excellence in Project Management* excellence in project management is defined as "a continuous stream of successfully managed projects." [12] But what is success? As in our personal lives success is defined individually and interpreted independently. For organizations, success is the same. Kertzner [12] reports that:

A brewery in Venezuela defines a successful project as one that falls within its predetermined time, cost, quality and scope limitations. Disney decides project success is fulfilling its time, cost, and safety requirements, with safety the most important requirement. Brian Vannoni of General Electric's Plastic Group defines success this way: "The technical aspects, timing and costs [in the past] were the three critical areas of performance measurement for our project managers. In today's world, that is not sufficient. We have to also be concerned with environmental and safety regulations, quality, customer satisfaction and ... productivity [of] manufacturing operations. So a project now has at least eight measurables and critical parameters that we gauge success around."

Kertzner stresses that success in project management, like quality in the production of goods and services, must be defined by the customer. His list of factors for project success includes: (1) completed on time, (2) completed within budget, (3) completed at the desired level of quality, (4) accepted by the customer, (5) resulting in customer allowing contractor to use customer as a reference, (6) with minimal scope change, and (7) without disturbing the ongoing business of the company.

Table 4 illustrates the phases of project management maturity that organizations may progress through on their way to creating a project management culture that can produce successful projects.

4. PROJECT MANAGEMENT AT KENNEDY SPACE CENTER

Kennedy Space Center on the whole has had a very mixed use of project management principles. From Table 4 the organization exhibits characteristics from each of the phases of maturity. Yet has not completely progressed through any of them. At the center today there is an increased awareness and visibility of what project management as a discipline and body of knowledge can do to assist KSC in achieving their goals. In moving toward development and research activities and away from sustaining the center: will be more task oriented; will require a higher accountability of their fixed resources; and will need to priority and control driven at the task level. This is an environment where project management principles can be very effectively applied. Mr. Roy Bridges, KSC Center Director, in his

July 27th 1999 KSC Rollout presentation entitled "Vision, Progress and Challenges" alluded to the fact that project management can help KSC at the center level prioritize program opportunities given the current environment of fixed budget and employment resources from the agency. Yet as a center KSC is just beginning to look at project management from a strategic perspective and to in-culture the need for excellence in the PM skills throughout the organization.

Table 4
Phases of Organizational Project Management Maturity [12]

Embryonic Phase	Executive Management Acceptance Phase	Line Management Phase	Growth Phase	Maturity Phase
Recognize need	Get visible executive support	Get line management support	Recognize life cycle changes	Develop a management cost & schedule control system
Recognize benefit	Achieve executive understanding of project management	Achieve line management support	Develop a project management system	Integrate cost and schedule control
Recognize applications	Establish project sponsor	Provide line management education	Make the commitment to planning	
Recognize what must be done	Become willing to change way of doing business	Become willing to release employees to project management	Minimize creeping scope Select a project tracking system	Develop an educational program to enhance project management skills

In looking at the project management activities at the center today several key observations can be made. There is a need for the center to have visibility of all resources engaged in project (or sustaining) activities in order to react to current opportunities and to plan for future opportunities (at the program level). It is required then to be able to roll-up the various project activities to provide this visibility. At the same time it is not completely necessary to establish a single project management office responsible for managing the life-cycle activities of all of center projects. Project management processes that are appropriate for the various types of projects undertaken at KSC can be developed and managed at the office level (where the action is). As an example of the diversity of projects at the center: the LSE budget involves physical launch site equipment enhancements; new special program projects would involve feasibility and exploration activities; CLCS and other on-going projects need support and close-out from a PM perspective; new R&D projects involve proof of concept, basic research, product development and technology creation phases; MM projects involve detailed design and build phases; and FF projects involve grounds, construction and maintenance activities. Each of these types of projects has its own activities, phases and requirements, each has varying involvement with contractors, subcontractors and internal/external units, each has a different involvement of Project Managers (PM) and Project Management Administrators (PMAs).

Thus as a center it may be appropriate to allow for variability in the PM processes depending upon the class of project, its over-arching goals and the phases involved in completing the project. From a strategic perspective, as the center transitions from an operations and sustaining environment to a research and development environment there is a high likelihood that project diversity will exist.

However, if after studying the projects conducted at the center a single PM processes can be developed and administered within a single unit which can roll-up projects for macro-management and also provide for their day-to-day management then this approach may be appropriate. To get to that point the following activities may be undertaken:

1. Investigate the various types of projects currently undertaken and planned for the near term.
2. Classify them in terms of their distinguishing characteristics, including most importantly the overarching goal of the project and likely phases of its life cycle.
3. Be sure to include all potential parties in developing this list of projects and project types.
4. Be sure to include a life cycle perspective of the projects under consideration including the specific characteristics of the life cycle (who is involved, what is done, etc.).
5. Develop a master list of appropriate local processes for each project classification, be sure to include all stakeholders (NASA and contractors) in this activity including the ultimate customer of the project work.
6. Evaluate the relative merits of a single center-wide PM process or a process that allows local control with a visibility activity at the macro-level.

In following these steps the Strategic Project Management Process Model described in the next section would be of great value to center.

5. THE STRATEGIC PROJECT MANAGEMENT PROCESS MODEL

The Strategic Project Management Process Model (SPMPM) is a conceptual model to be developed to assist KSC as it builds appropriate project management processes at the center. On a larger scale the model will incorporate a scope beyond specific uses at KSC that should add to the body of knowledge in the general field of project management and thus add to the understanding and effective use of this important (and growing) field. The model involves three distinct phases as illustrated in Figure 2 and is developed upon the assumption that not all project management processes are appropriate for all classes of projects in all environments and for all types of objectives. In the end the model will define and describe the project management processes which are appropriate for specific classes of projects.

The Characterization phase of the model will delineate all possible differentiating project variables. These variable will be identified from tactical, strategic, and environmental perspectives (as well as other), and will be developed from published literature and general specifications of the domains and characteristics in which project management tools are used. The Capture phase of the model classifies these general characteristics into meaningful discriminating factors from which groups of projects can be distinguished. Data from PMI and other discipline resources as well as case data, best practices and theory will be applied in this phase. Once captured the Prescriptive phase will offer an output of the particular project management processes and key variables most appropriate for a given class of projects. In going from the Classification to the Prescriptive phases a rules or knowledge engine will be developed and integrated. This engine will serve to process the specific requirements of the captured class of projects to produce the Prescriptive output. The SPMPM would be very appropriate for organizations developing or establishing new project management processes. It also could be used by organizations who have currently operational processes as a tool to evaluate those existing PM processes and offer recommendations for improvement for a higher level of organizational goal attainment.

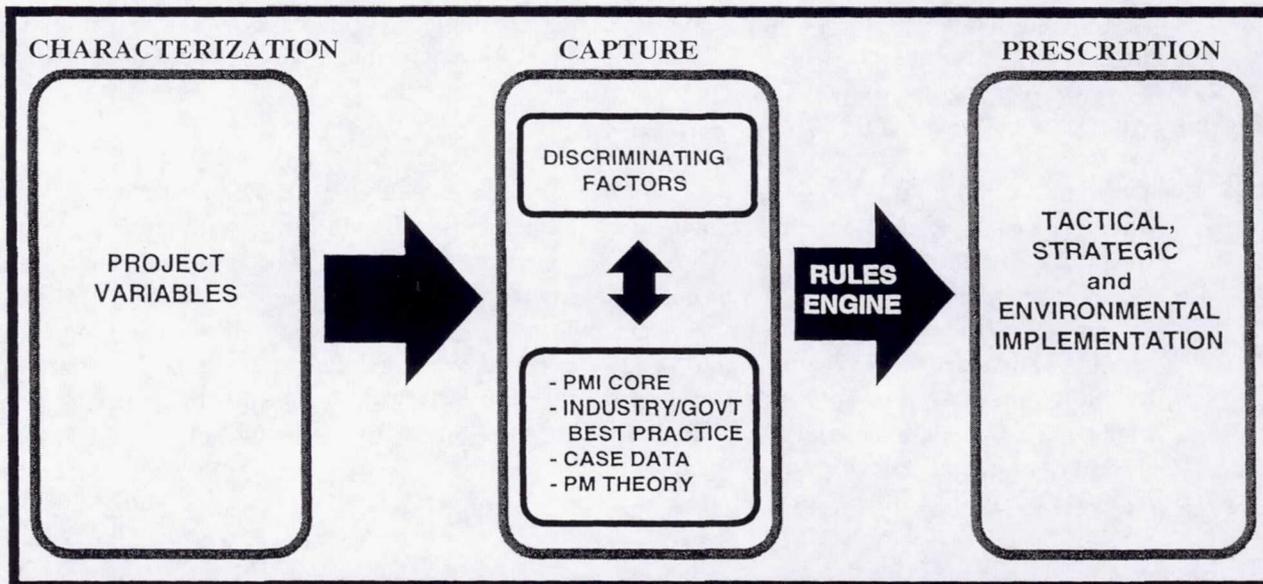


Figure 2: Strategic Project Management Process Model

6. CONCLUSION

NASA's strategic planning process has caused a profound change within the agency over the period of the last 6 years. This process has forced the agency and all of its centers to be more focused, more accountable, more customer driven, and to do more with less in a "better, faster, cheaper" environment. At Kennedy Space Center the strategic planning process has led to the center positioning itself as the spaceport of choice for government and industry in the United States (and world). Being one of only two sights from which mankind has launched humans into space the center has great knowledge and expertise in the checkout, processing and launch of space vehicle crews and payloads. The challenge for KSC is to build on that knowledge and expertise as space flight changes over the next 50 years. KSC must also develop the physical infrastructure to support such changes.

As KSC moves to a research and development environment to support its goals, expertise in and use of project management principles become critical. Now is the time for KSC to look at the project management processes that will be required to support their overarching goals, to put into place the organizational structure to support these processes, and develop the necessary skills in the workforce so that Kennedy Space Center can claim success in project management. The SPMPM will be developed to provide a prescriptive PM process recommendation when project characteristics are provided. In this way the model will aid and assist organizations as they develop PM processes to support their environment and goals. The model itself is currently at the conceptual design stage and will be further defined and developed as time and interest dictates. This model should be of great use to NASA KSC and other organizations who are using PM as a vehicle toward organizational excellence.

As a final thought, it is this author's opinion that KSC should aspire to be NASA *Center of Excellence in Project Management*. This would involve creating a system for the development of appropriate project management processes, an organizational structure model to support projects for visibility at the macro-management and day-to-day management levels, a training/re-training education/curriculum model for transitioning current workforce skills to include "as needed" project management skills, and an integrated model that illustrates how PM integrates with the strategic planning process to create value and success for the organization. Becoming NASA's PM Excellence Center (or Lead) establishes PM as a priority at the center and seizes agency competency of this critical skill.

REFERENCES

- [1] Government Performance and Results Act of 1993, File s20.enr, S.20, s.20, One Hundred Third Congress of the United States of America, @ <http://www.npr.gov/library/misc/s20.htm>
- [2] Weber, M., *The Theory of Social and Economic Organization*, New York: Free Press, 1924.
- [3] National Partnership for Reinventing Government, internet website @ <http://www.npr.gov/library/studies.html>
- [4] Strategic Planning and Strategic Management Within NASA, NHTSA Case Study, @ <http://www.npr.gov/library/studies/nasa.html>
- [5] NASA Strategic Plan 1998, NASA Policy Directive (NPD)-1000.1 @ <http://www.hq.nasa.gov/office/nsp/cover.html>.
- [6] Kennedy Space Center Roadmap, KDP-KSC-S-2001, Rev: Basic 3/98, @ http://wit.ksc.nasa.gov/BusinessWorld/html/strategic_planning.html
- [7] KSC Implementation Plan, Kennedy Space Center Implementing NASA's Strategies, KDP-KSC-S-2000, Rev: Basic 3/98 @ http://wit.ksc.nasa.gov/BusinessWorld/html/strategic_planning.html
- [8] Rosenau, M.D., *Successful Project Management: A Step-by-step approach with Practical Examples*, 2nd edition, Van Nostrand Reinhold Publishers, 1992.
- [9] Culp, G. and Smith, A., *Managing People (Including Yourself) for Project Success*, Van Nostrand Reinhold Publishers, 1992.
- [10] Kertzner, H., *Project Management: A Systems Approach to Planning, Scheduling and Controlling*, 6th Edition, John Wiley & Sons, Copyright 1998.
- [11] Project Management Institute, *A Guide to the Project Management Body of Knowledge*, PMI Publishing Division, Copyright 1996.
- [12] Kertzner, H., *In Search of Excellence in Project Management*, Van Nostrand Reinhold Publishers, Copyright 1998.

**1999 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM
JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA**

***THE EFFECT OF DILUTION ON THE
STRUCTURE OF MICROBIAL COMMUNITIES***

Aaron L. Mills

Professor

Laboratory of Microbial Ecology

Department of Environmental Sciences

University of Virginia

Colleagues: John Sager and Jay Garland

ABSTRACT

To determine how dilution of microbial communities affects the diversity of the diluted assemblage a series of numerical simulations were conducted that determined the theoretical change in diversity, richness, and evenness of the community with serial dilution. The results of the simulation suggested that the effects are non linear with a high degree of dependence on the initial evenness of the community being diluted. A series of incubation experiments using a range of dilutions of raw sewage as an inoculum into sterile sewage was used for comparison to the simulations. The diluted communities were maintained in batch fed reactors (3 day retention time) for 9 days. The communities were harvested and examined by conventional plating and by molecular analysis of the whole-community DNA using AFLP and T-RFLP. Additionally, CLPP analysis was also applied. The effects on richness predicted by the numerical simulations were confirmed by the analyses used. The diluted communities fell into three groups, a low dilution, intermediate dilution, and high dilution group, which corresponded well with the groupings obtained for community richness in simulation. The grouping demonstrated the non-linear nature of dilution of whole communities. Furthermore, the results implied that the undiluted community consisted of a few dominant types accompanied by a number of rare (low abundance) types as is typical in unevenly distributed communities.

THE EFFECT OF DILUTION ON THE STRUCTURE OF MICROBIAL COMMUNITIES

Aaron L. Mills

Introduction

Although microbes are capable of growing anywhere even a small amount of liquid water exists, habitats that contain organic carbon that can be used as an energy source by the microorganisms will support abundant microbial populations. The inclusion of bioregenerative life support (BLS) elements (i.e., plant growth systems and bioreactors for waste processing) will significantly increase the total abundance of microbes in extraterrestrial facilities. For example, bacterial numbers on the roots of plants within prototype hydroponic systems can be as high as 10^{11} cells g^{-1} dry wt(3). The potential exists for the communities associated with these systems (e.g., biofilms attached to plant roots or hardware surfaces, or mixed populations in the suspended phase of reactors) to harbor microbes pathogenic to humans or to the plants (14). Management of microbial communities to minimize the potential for risk to the crew and to the plants to be used for supporting the crew is an essential component of successful BLS systems.

Recommended approaches to the management of microbial communities in ALS range from strict decontamination and control to "seeding" the system with a diverse group of microorganisms (Fig. 1). One extreme would be the use of axenic plants in subsystems "bio-isolated" from the human habitat module in order to prevent the proliferation of human-associated organisms (including opportunistic human and plant pathogens) in plant growth systems. Alternatively, inoculation of plant systems with soil from fertile, disease-suppressive soils has been proposed to provide a stable microbial population that reproduces the balance in population dynamics found on Earth (1). An integrated approach involving quarantine,

sanitation, compartmentalization, and construction of microbial communities has also been recommended (13). Another approach to community management would be to use gnotobiotic communities, defined assemblages in which all of the members are known. Generation of defined mixed cultures to carry out all of the essential processes necessary for operation of BLS systems is a daunting task, and maintenance of such defined cultures without contamination from the crew or from other systems is likely impossible. Regardless of the starting mixture, evolution of the community will eventually produce something quite different from the initial assemblage. The alternative to the gnotobiotic approach invokes the ecological paradigm that diverse communities tend to be more stable and more resistant to invasion than counterparts with

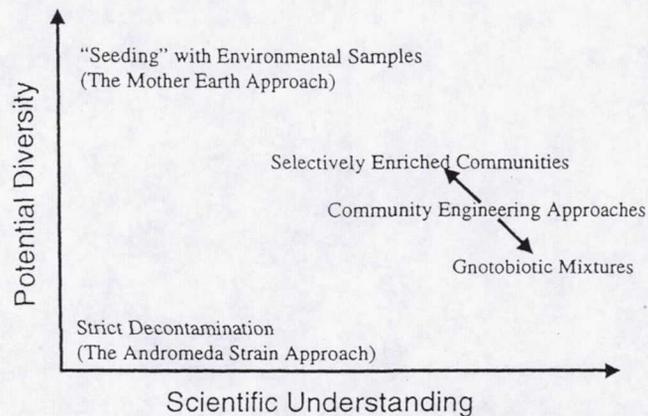


Fig. 1. Approaches to inoculum preparation for hydroponic plant growth systems and bioreactors for long term space habitation. The community engineering approaches should provide substantially greater protection of the communities against invasion by human associated bacteria or plant pathogen as compared with the strict decontamination approach (which cannot be reliably maintained for extended periods), while reducing the threat of inclusion of potentially harmful microbes as part of the inoculum. Figure reproduced from Garland, et al. (6).

few types of organisms present (e.g., 2, 7, 9, 11, 12, 15). Selection of highly diverse communities should result in narrower niches which will be difficult for pathogens or other unwanted microbes to fill. Although this concept is generally accepted by ecologists (the current emphasis on biodiversity in terrestrial ecosystems is a direct result of that acceptance), it has never been explicitly examined for microbial (i.e., bacterial) communities.

Tests of the effects of diversity on any community property or system function require a knowledge of the diversity of the community under examination, but there is no method currently available that allows diversity of a microbial community to be measured. Despite the inability to measure diversity directly, Garland (6) and Morales et al. (10) successfully used dilution to manipulate diversity for several applications. There is an inherent assumption that dilution is a linear process, even though dilution of environmental samples for microbiological analysis often produces non-linear results. The present work sought to define the relationship between dilution and resultant diversity in a series of numerical simulations accompanied by appropriate incubation experiments.

Dilution Simulations

To examine the effect of dilution on community composition, a series of numerical simulations were done. Communities were constructed by assigning each of 1×10^6 individuals a random species identification based on a normal distribution of integers from 1 to 1000 with a mean of 500 and a preset variance. By adjusting the variance, communities of different evenness could easily be simulated; low variance resulted in communities with a large number of rare species, whereas high variance yielded a community of more even distribution (Fig. 2). For the simulations described here, the total abundance and richness were held constant, and only the evenness was allowed to vary. Evenness was altered by setting the variance level to 100, 250, 1,000, and 20,000. An even distribution was simulated by creating 1000 types, each containing 1000 individuals. The initial communities were then diluted mathematically by selecting 1/10 of

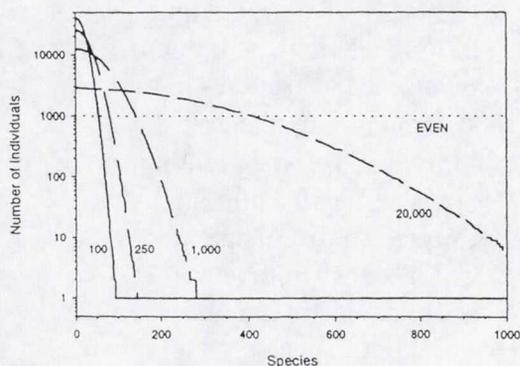


Fig. 2. Distribution of individuals among 1000 types or species as produced. These distributions are those for the initial communities used in the simulations reported here. Note that both abundance and richness are constant in the initial communities

the individuals from the array that represented the undiluted community. The species identification was placed in another array which then served as the initial community for the next dilution in the series. All serial dilutions were done in powers of 10, as is typically done for microbiological analysis. For each community and at each dilution level, the richness, evenness and diversity were calculated. Richness was taken to be the number of types in the community, diversity was expressed as the Shannon-Wiener index

$$H' = \sum_1^i p_i \ln p_i \quad (1)$$

where p_i is the fraction of the individuals in the i^{th} species.

Evenness is calculated as

$$E = \frac{H'}{H'_{\max}} \quad (2)$$

$$H'_{\max} = \ln S \quad (2a)$$

where S represents the total number of species present.

A sample dilution series for the most dominant community (var = 100) and the most even community (var = 20,000) show the type of changes in the community generated by the simulated dilutions due to difference in initial evenness (Fig. 3).

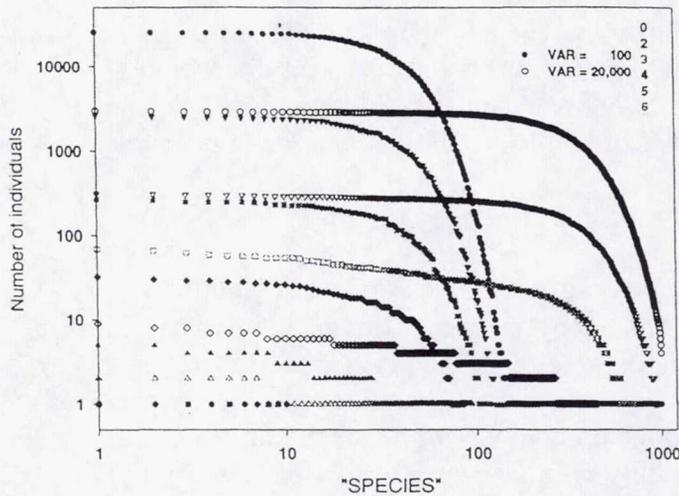


Fig. 3. Effect of dilution on species distribution in the most even (var = 20,000) and most dominant (var = 100) communities.

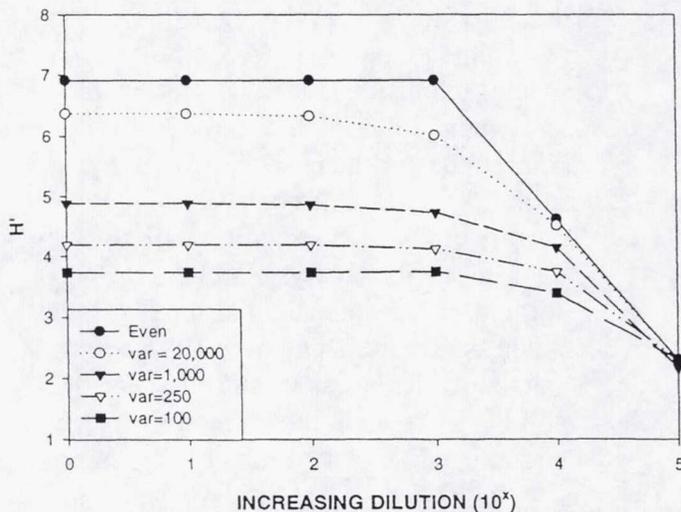


Fig. 4. Value of the Shannon-Wiener index for initial communities (0) of differing evenness and serial dilutions made from each of the communities.

created by setting var = 1000. In that case, the value of H' at the 10⁻⁵ dilution was 2.1640. The lower value was obtained because only 9 species were recovered in that particular simulation whereas 10 were recovered in all the others.

The serial dilution made corresponded to typical 10-fold dilutions used in microbiology, and extended from 0 to 10⁻⁵. Given an initial cell concentration of 1 × 10⁶ mL⁻¹, the dilutions were made by selection of 1 × 10⁵ cells from the initial community, then 1 × 10⁴ cells from the first dilution, 1 × 10³ cells from the second dilution, 1 × 10² cells from the third dilution and 1 × 10¹ cells from the fourth dilution. Ten cells represents the smallest number of cells in a 10-fold dilution series for which H' can be meaningfully determined (H' for 1 individual is 0).

Diversity values for the range of initial communities were from 3.7324 for the least even community (var = 100) to 6.9078 for the perfectly even community (Fig. 4). Note that for the even community, dilution caused no change in H' until the number of individuals in the community reached the number of species in the community, in this case 1 × 10³ individuals, corresponding to 1 × 10³ types at the 10⁻³ dilution. Upon dilution, little change in H' was observed when the number of individuals greatly exceeded the number of species; however, H' decreased to the theoretical value of 2.3026 (based on H' _{max} for the even distribution) at a dilution of 10⁻⁵. The sole exception was in the community

Evenness of the communities not initially perfectly even increased with dilution and reached the theoretical maximum of 1.0 at the 10^{-5} dilution (Fig. 5). In the community obtained by setting $\text{var} = 20,000$, there was a change in evenness that was not consistent with the patterns displayed by dilution of the other communities. Because the communities were formed by probabilistic sampling, it is possible that the anomaly is simply a variant, although it is not understood why the evenness obtained did not seem to influence subsequent dilutions.

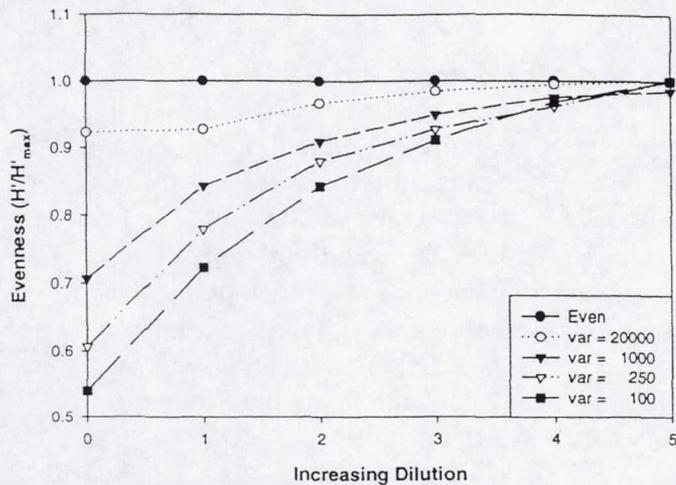


Fig. 5. Changes in evenness with dilution in communities of differing initial evenness.

species lost in the first dilution decreased substantially and approached the theoretical value of no species lost in the first dilution. Furthermore, in the treatment $\text{var} = 20,000$, loss of species in subsequent dilutions was also low, comparing favorably with the perfectly even state.

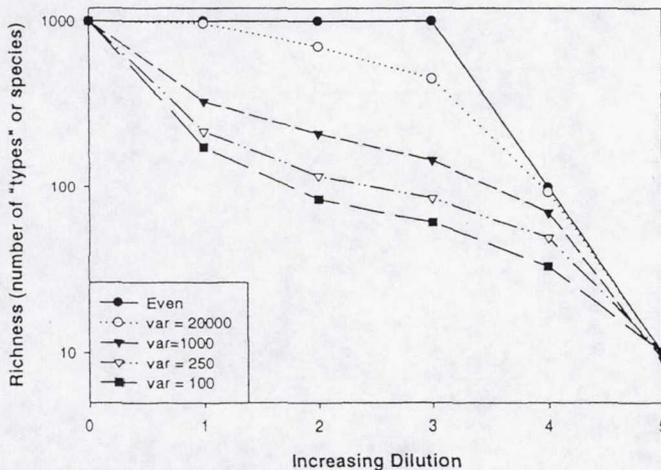


Fig. 6. Changes in richness with dilution in communities of differing initial evenness

Richness of communities changed differently with dilution depending on the initial evenness (Fig. 6). For communities of low initial evenness, richness dropped rapidly with the first dilution. As the initial community became more even, the number of

Incubation Experiments

Raw sewage was collected from the Cape Canaveral Air Station Waste Water Treatment Facility (Kennedy Space Center, FL). Samples were allowed to settle for ~2 hours (to remove large particles) and a serial dilution (through 10-6) was prepared from the supernatant to use as the inoculum for the batch culture experiments. Probability suggests that, because the dilution of the original community should remove rare organisms, these inocula should systematically differ in diversity (both richness and evenness).

Seven treatments (undiluted (100) through 10-6) were established by adding

1 ml of inoculum to 60 ml of sterile sewage in a 125-ml flask. For each treatment, three replica flasks were maintained; all flasks were kept on a shaker table (150 rpm) to ensure aerobic conditions. Each day, 20 ml of liquid were removed from each flask and replaced with 20 ml of sterile sewage.

After 9 days (3 retention times), flasks were harvested and samples collected for total cell counts (Acridine Orange direct counts), culturable counts on R2A agar, richness, evenness, and

diversity of colony morphologies on R2A, community level physiological profiling (CLPP (5)), dilution-extinction analysis of CLPP (4), molecular analysis of whole community DNA, (including AFLP (16) and T-RFLP (8))

Richness, evenness, and diversity of colony morphologies on R2A

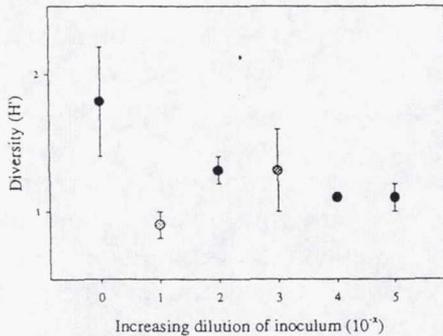


Fig. 7. A decrease in diversity (based on colony morphology on R2A agar) was observed with increasing dilution. The X-axis shows the exponent of the dilution factor used in the original inoculum (e.g. "4" is the regrown community derived from the 10^{-4} dilution inoculum). The Y-axis is diversity (Shannon-Wiener index), and error bars shown are \pm one standard deviation.

For each treatment, a serial dilution of the regrown community was plated onto R2A agar and incubated at room temperature. After 72 hours, 25 colonies were randomly chosen from each countable plate and colony morphology described. Richness (S) was then determined as the total number of colony types encountered (Fig. 8); evenness (E) (Fig. 8) and diversity (H') (Fig. 7) were calculated using the Shannon-Wiener diversity index.

T-RFLP analysis of bacterial 16S rRNA genes

Terminal restriction fragment length polymorphism (T-RFLP) was used to compare the overall diversity of the bacterial communities from these flasks. The bacterial 16S rRNA gene was amplified using two primers (1392 Reverse (5' ACGGGCGGTG TGTRC) and 8 Forward (5' AGAGTTTGATCCTGGCTCAG (labeled)));

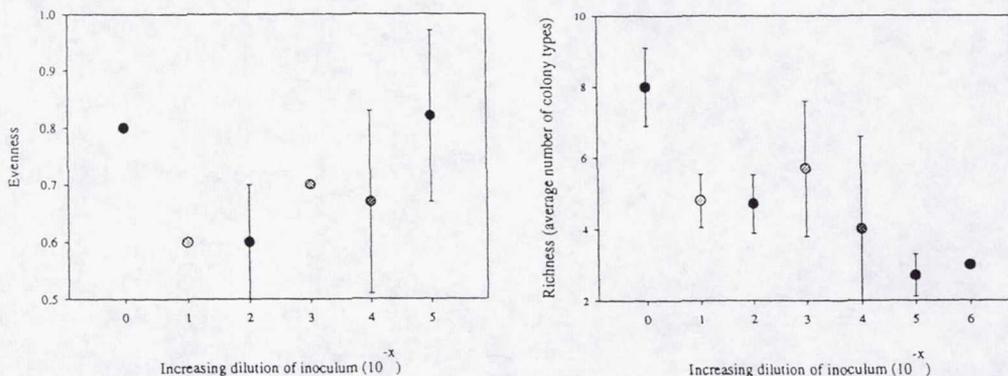


Fig. 8. An increase in overall community evenness and a decrease in richness (number of colony types) was observed with increasing dilution. The X-axis shows the exponent of the dilution factor used in the original inoculum, and error bars shown are \pm one standard deviation.

aliquots of the PCR product were then digested using either the HhaI or MspI restriction enzyme. Using an automated DNA sequencer, the length of the terminal restriction fragments was then compared across the different communities. HhaI generated 43 different T-RFLP fragments (the average sample contained 16 fragments) and MspI generated 42 different fragments (the average sample contained 12). The number of fragments observed across the dilution/diversity gradient did not differ (Table 1), though one might have expected a decrease corresponding to a loss in

diversity (richness). Additionally, principle components analysis (PCA) was performed on the combined (MspI and HhaI) datasets (Fig. 9).

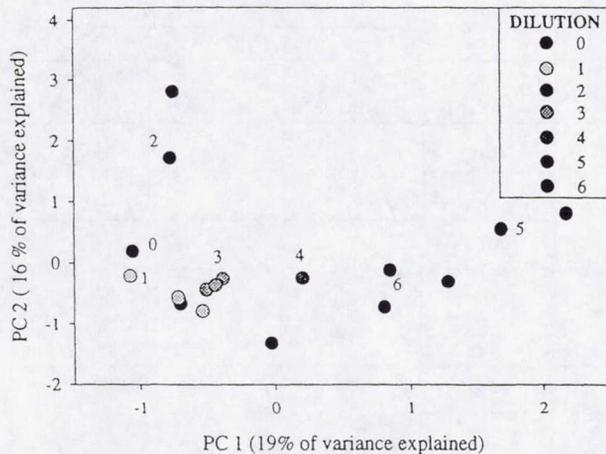


Fig. 9. PCA diagram of the T-RFLP fingerprints. Each "dilution" identified in the legend is the exponent of the dilution used in the original inoculum (e.g. "4" represents the regrown community inoculated with the 10^{-4} dilution). On the graph, the high dilution (theoretically low diversity) treatments 5 and 6 separated from all the other treatments on the first principle component.

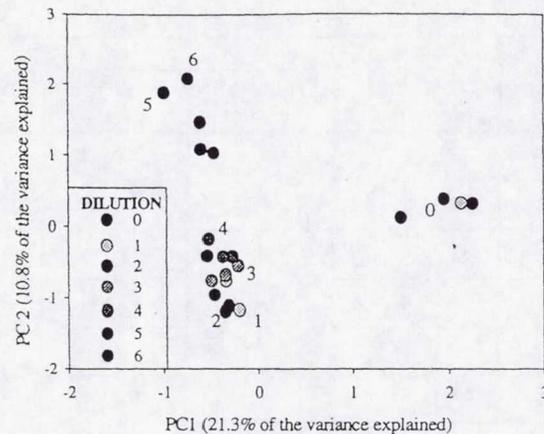


Fig. 10. PCA of the AFLP fingerprints. Each "dilution" identified in the legend is the exponent of the dilution used in the original inoculum (e.g. "4" represents the regrown community inoculated with the 10^{-4} dilution). This graph shows that the different dilution/diversity inocula separated into three main groups based on overall genetic structure. The undilute inocula ("0") were most distinct, separated on PC1, and the very dilute inocula ("5" and "6") were distinct from all others on PC2. Cluster analysis was also performed on the AFLP data and the same major divisions were observed (data not presented).

Fingerprinting of whole-community DNA using AFLP

AFLP (Amplified Fragment Length Polymorphism) was used to profile overall community structure in this experiment. AFLP is based on amplification of DNA restriction fragments, after ligation of an adapter sequence, using fluorescently labeled PCR primers (4). It is similar to other "arbitrarily primed" PCR based fingerprinting techniques because it requires no *a priori* sequence information for primer design (unlike T-RFLP); it differs in that it is more reproducible and provides a greater resolution with fewer primers. In this research, three different primer sets were used in the amplification, and the data pooled (85 bands). The presence or absence of each band in each sample was recorded and analyzed using principal components (Fig. 10) and cluster analyses (data not shown).

The average number of AFLP bands per treatment decreased along the dilution/diversity gradient (Table 1). Because AFLP fragments are, in principle, simply RFLPs resolved by selective PCR amplification of the DNA restriction fragments, the number of fragments produced can be expected to correlate with the overall diversity of the communities. Furthermore, most of the bands encountered in the undilute ("0") treatment were unique (not found in any further treatments), corresponding to a large change in community structure (perhaps richness) after the first dilution.

Community level physiological profiling

Community level physiological profiling (CLPP) was used to compare overall functional potential among the different dilution/diversity treatments based on each community's ability to metabolize 95 different sole carbon sources (1). Principle components analysis was used to

analyze the data and the first two PCs were plotted as a means of visualizing the relationships among the different samples (Fig 11).

Table 1. Summary of diversity analysis for plates and molecular techniques

Dilution	R2A plates		AFLP		T-RFLP			
	Avg. # of colony types	Proportion unique ¹	Avg. # of bands	Proportion unique ²	# bands w/ MspI	Proportion unique ³	# bands w/ HhaI	Proportion unique
10 ⁰	8	0.4	26	0.9	7	0.21	19	0.07
10 ⁻¹	4.8		24		9		15	
10 ⁻²	4.7	0	17	0.6	11	0.21	20	0.23
10 ⁻³	5.7		17		12			
10 ⁻⁴	4		14		13			
10 ⁻⁵	2.7		20		21			
10 ⁻⁶	3	1.0	15	0.5	8	0.17	10	0.19

1 - Number of unique colony types in a treatment group/total number of colony types encountered in each treatment group. "Unique" refers to a band (colony type, etc.) that is present in a particular treatment group (e.g. 10⁰ and 10⁻¹) that is not present in either of the other two treatment groups. It is interesting to note that all of the colony morphologies observed in the 10⁻⁶ treatment were unique (not encountered in any of the higher dilution/diversity treatments). Nearly half (40%) of the colony morphologies encountered in the low dilution/high diversity treatment were unique while none of the colony types encountered in the mid-dilutions were unique (all of those colony types were encountered in these treatments were also observed in the 10⁰ and 10⁻¹).

2 - Number of unique AFLP bands in a treatment group/total number of AFLP bands encountered in each treatment group. Nearly all of the bands encountered in the low dilution/high diversity treatments were unique suggesting many organisms types were present in this treatment and diluted out of others.

3 - Number of unique T-RFLP bands in a treatment group/total number of T-RFLP bands encountered in each treatment group. The majority of the T-RFLP peaks were not unique, and were found across the entire dilution series. This suggests a large bias in the T-RFLP/PCR procedure towards certain organisms. If we assume these organisms to be dominant, then these results suggest that the same types of organisms were dominate throughout the dilution series. It is also interesting to note that number of T-RFLP peaks (number of ribotypes) was significantly greater than the number of colony morphologies on R2A agar.

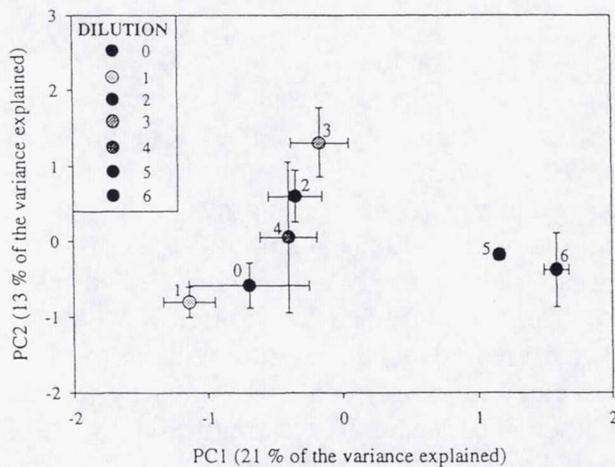


Fig. 11. PCA of the CLPP. Each "dilution" identified in the legend is the exponent of the dilution used in the original inoculum (e.g. "4" represents the regrown community inoculated with the 10⁻⁴ dilution). Error bars represent 95% confidence intervals around the centroid for each treatment (average of 3 independent replica flasks). Two distinct groups can be distinguished on this plot; treatments 0 through 4 (undiluted through 10⁻⁴) form one group separate from the 10⁻⁵ and 10⁻⁶ dilution communities.

Dilution-extinction analysis of CLPP profiles

Garland and Lehman (4) used the extinction of functional characters in CLPP assays to make inferences about diversity of microbial communities. Plotting the number of positive tests obtained in CLPP analyses of dilutions made of each regrown community against the number of cells inoculated into each well of the BIOLOG plate yields curves that resemble those obtained in saturation plots. Fitting a rectangular hyperbola to the data produces two parameters that are analogous to those obtained with saturation plots:

$$F = \frac{F_{\max} \times K_F}{1 + K_F} \quad (3)$$

where F is the number of positive tests at a given inoculum density (referred to by Garland and Lehman as the functional richness), F_{\max} is the estimated maximum number of tests that would be positive for the community, and K_F is the cell density at which $F = F_{\max}/2$. Parameter estimates and the curves representing each dilution treatment are shown in Fig. 12).

Differences in microbial community structure were detected throughout the dilution/diversity gradient with every measure employed. Major groups formed as follows: the undilute, regrown communities ('0' (high diversity?)), the very dilute (10^{-5} and 10^{-6} (very low diversity?)), and all others (10^{-1} through 10^{-4}).

Conclusions

- Different measures (which focused on different elements of "diversity" and had different inherent methodological biases) showed different levels of resolution. For example:

- comparing culturable diversity (richness and evenness) of colony types of R2A highlighted the difference between the undilute inocula and all others
- PCA analysis of CLPP showed that the 10^{-5} and 10^{-6} treatments were different from all others, but did not separate the undilute community. Dilution/extinction analysis of CLPP divided the communities into the three dilution/diversity groups.
- PCA of the T-RFLP profiles showed that the 10^{-5} and 10^{-6} treatments were different from all others, though no change in diversity (measured by the relative number of peaks) was observed.
- AFLP divided the communities into three dilution/diversity groups (using PCA) based on overall structure.

Interpreting all of these results in the context of the numerical simulations, it seems that the original community was highly dominant (e.g. var = 100 or var = 250). For these curves, the greatest difference in diversity was observed after the initial dilution (loss of richness) and at extreme dilutions (10^{-5}); these theoretical divisions agree well with the experimental results. Assuming a highly dominant initial community, measures that showed the undilute regrown community to be unique (R2A colony morphology, AFLP and dilution/extinction of CLPP) probably were better at detecting differences in richness. Measures that showed the 10^{-5} and 10^{-6} treatments to be unique (T-RFLP, AFLP, PCA of CLPP, dilution/extinction of CLPP) probably reflected differences in overall diversity (richness and evenness) better. It is interesting to note

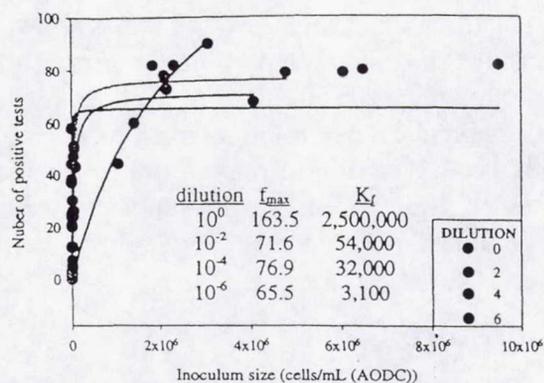


Fig. 12. Treatments 0, 2, 4, and 6 were profiled using dilution/extinction analysis. The raw data are presented (number of positive characters versus cell concentration) and fit with a rectangular hyperbola; parameter estimates for this fit are given in the legend. The general shapes of each dilution/extinction curve were used by Garland and Lehman to suggest differences in communities based on the dominance of generalists vs specialists. In this figure, the curve for the undiluted inoculum ("0") suggests a community dominated by specialists, where rapid loss of strains also means a rapid loss of characters. In the higher dilution treatments, the lower values of F_{\max} and the lower values of K_f imply a community composed of generalists in which there are fewer functions, but in which those functions are conserved throughout the community. An alternate explanation is that more even communities tend to lose strains (and by implication, functions) more slowly through dilution.

that nearly all measures (except R2A) showed the 10^{-5} and 10^{-6} treatments to be unique, not all measures separated the undiluted regrown community.

Although the use of dilution to produce communities of different (lower) diversity is a reasonable means of manipulating diversity, the effects are non-linear. When dilution is coupled with the additional non-linear effects of regrowth of mixed cultures, a very complex result is obtained. Clearly much further work is needed to determine the extent to which the effects can be predicted and controlled for experimental and applications purposes.

Literature Cited

1. **Alexander, D. B., D. A. Zuberer, and D. Hubbell.** 1989. Microbiological considerations for lunar derived soils. *In* D. W. Ming and D. L. Henninger (ed.), *Lunar Base Agriculture: Soils for Plant Growth*. American Society of Agronomy, Madison, WI.
2. **Case, T. J.** 1990. Invasion resistance arises in strongly interacting species-rich model competition communities. *Proceedings of the National Academy of Science*. **87**:9610-9614.
3. **Garland, J. L.** 1994. The structure and function of microbial communities in hydroponic systems. *Advances in Space Research*. **14**:383-386.
4. **Garland, J. L., and R. M. Lehman.** Submitted. Dilution/extinction of community phenotypic characters to estimate relative structural diversity in mixed communities. *FEMS Microbiol. Ecol.*
5. **Garland, J. L., and A. L. Mills.** 1991. Classification and characterization of heterotrophic microbial communities on the basis of patterns of community-level-sole-carbon-source utilization. *Appl. Environ. Microbiol.* **57**(8):2351-2359.
6. **Garland, J. L., A. L. Mills, A. Morales, and C. Cook.** 1999. Presented at the ICES 99, Denver, CO.
7. **Huston, M. A.** 1997. Hidden treatments in ecological experiments: re-evaluating the ecosystem function of biodiversity. *Oecologia*. **110**:449-460.
8. **Liu, W.-T., T. L. Marsh, H. Cheng, and L. J. Forney.** 1997. Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. *Appl. Environ. Microbiol.* **63**(11):4516-4522.
9. **McGrady-Steed, J., P. M. Harris, and P. J. Morin.** 1997. Biodiversity regulates ecosystem predictability. *Nature*. **390**:162-165.
10. **Morales, A., J. L. Garland, and D. V. Lim.** 1996. Survival of potentially pathogenic human-associated bacteria in the rhizosphere of hydroponically grown wheat. *FEMS Microbiol. Ecol.* **20**(3):155-162.
11. **Naeem, S., L. J. Thompson, S. P. Lawler, J. H. Lawton, and R. M. Woodfin.** 1994. Declining biodiversity can alter the performance of ecosystems. *Nature*. **368**:734-737.
12. **Naeem, S., L. J. Thompson, S. P. Lawler, J. H. Lawton, and R. M. Woodfin.** 1995. Empirical evidence that declining species diversity may alter the performance of terrestrial ecosystems. *Phil Trans. R. Soc. Lond.* **347**:249-262.
13. **Nelson, B.** 1987. The role of plant pathology in development of controlled ecological life support systems. *Plant Disease*. **71**:580-584.
14. **Schuerger, A. C.** 1998. Microbial contamination of advanced life support (ALS) systems poses a moderate threat to the long-term stability of space-based bioregenerative systems. *Life Support and Biosphere Science*. **5**:325-337.
15. **Tilman, D., and J. A. Downing.** 1994. Biodiversity and stability in grasslands. *Nature*. **367**:363-365.
16. **Zabeau, M., and P. Vos.** 1993 patent European Patent Application 92402629.7.

1999 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

JOHN F. KENNEDY SPACE CENTER

UNIVERSITY OF CENTRAL FLORIDA

INTEGRATED WORKFORCE MODELING SYSTEM

Gary P. Moynihan
Associate Professor
Department of Industrial Engineering
The University of Alabama
Ronald Kent HM-E

ABSTRACT

There are several computer-based systems, currently in various phases of development at KSC, which encompass some component, aspect, or function of workforce modeling. These systems may offer redundant capabilities and/or incompatible interfaces. A systems approach to workforce modeling is necessary in order to identify and better address user requirements.

This research has consisted of two primary tasks. Task 1 provided an assessment of existing and proposed KSC workforce modeling systems for their functionality and applicability to the workforce planning function. Task 2 resulted in the development of a proof-of-concept design for a systems approach to workforce modeling. The model incorporates critical aspects of workforce planning, including hires, attrition, and employee development.

INTEGRATED WORKFORCE MODELING SYSTEM

Gary P. Moynihan

1. INTRODUCTION

Workforce planning, alternatively referred to as workforce forecasting or manpower planning, refers to predicting the size and mix of a pool of workers for some future time period [1]. According to Gass [2], it is determining "the number of personnel and their skills that best meets the future operational requirements of an enterprise".

There are three primary workforce planning strategies: transaction, event-driven, and process-driven workforce forecasting [3]. The transaction-based strategy is the most widely accepted. It monitors the flow of human resources through the organization over time. This strategy utilizes one or more views of the future that are driven by choices under the control of the organization. (Niehaus [3] refers to these as internal demand-driven business strategies.) The analytical tools associated with transaction-based workforce planning are categorized into two groups. The bottoms-up review reflects the perspective of project or functional department managers, regarding their human resource needs for a given planning horizon. These managerial inputs are summed into a consolidated workforce projection for the entire organization. These bottoms-up transaction forecasts are useful when incremental organizational growth is expected. However, top-down strategic organization changes are often required to address declining workforce situations [3]. These strategic resizing situations are usually estimated by using mathematical "flow" models, incorporating transition rate data, manpower requirements, and an existing population profile.

The objective of workforce planning is to define the future human resource needs of an organization in relation to the availability of suitable personnel, organizational goals, and budgetary constraints. A number of approaches are available for management to achieve this objective [4]. Traditionally, workforce planning has been accomplished based on judgement and experience. However, with the increase in workforce composition, complexity, and organizational competition, many mathematical models and computer-based information systems have been applied to this problem area. Decision support systems (DSS) are information systems that incorporate such sophisticated mathematical models. They can provide the opportunity to project workforce requirements across a defined planning horizon, as well as evaluate the alternatives in order to find the optimal solution to sub-problems.

2. TASK 1

There are several computer-based systems currently in various phases of development at Kennedy Space Center. These systems may offer redundant capabilities, and/or have incompatible interfaces. A systems approach to workforce modeling was requested by the client. The initial task in this approach was to assess current and proposed KSC workforce modeling systems for their functionality and applicability to workforce planning. Systems

identified for review include the Integrated Management System (IMS), the Project Resource Management System (PRMS), the Integrated Financial Management System (IFMP), the Expert Workforce Modeling System, and Expert Seeker (for corporate knowledge). It was also requested that this review provide recommendations for possible integration of these initiatives.

2.1 DESCRIPTION OF THE CURRENT PROCESS

The Workforce Planning group (designated HM-E), within the KSC Administration Office (HM) directorate, is responsible for Center-wide workforce planning. The overall processes associated with KSC workforce planning, and subsequent workforce development, are consistent with ISO documentation requirements. This process of workforce planning follows an overall transaction-based strategy, such as that described in the Introduction section of this research report. This transaction-based strategy combines aspects of both the bottoms-up and top-down approaches described.

Workforce Planning receives inputs with regard to future needs over the next five year planning horizon, from each of the individual programs/projects, as well as the functional departments, at KSC. These inputs frequently take the form of an Excel spreadsheet. Five-year KSC workforce thresholds, normally based on Agency budgetary limits, are provided to Workforce Planning by Mr. Roy Bridges (KSC Center Director). This guidance may also contain specific skill level targets.

These future needs and constraints are then compared to the existing baseline. The current workforce profile is identified at the center-level as well as at the program/project/functional department level. Again, these are represented as Excel spreadsheets. The Center-wide skills inventory describes present employee competencies and subordinate skills. The skills inventory was compiled during Fall 1998, via the Project Resource Management System (PRMS). The skills inventory is available as an Access file downloaded from PRMS. In addition to comparing these existing baseline profiles with the objective requirements and constraints, other factors (e.g. hires and attrition) are considered. The output of this effort is a time-phased workforce plan across the Center, for a five-year horizon, showing both prospective FTE (Full Time Equivalents) and headcounts. These outputs are also generated in the form of Excel spreadsheets, with breakdowns available at the program/project/functional department level.

One important aspect of the workforce plan is to identify the skills mix required to meet Center missions and roles, over the given forecasting horizon. Any projected skills imbalances are determined by manual inspection. Employee development plans are then established in order to relieve any skills mix constraints.

To-date, this process at KSC has been accomplished on a primarily manual basis, via the judgement and expertise of the Workforce Planning group. This approach has been sufficient in a relatively stable environment, or one experiencing incremental changes [1]. The Kennedy Space Center workforce is undergoing a major restructuring. This process involves a transition from operations to a focus on research and development [5]. This changing environment requires new computer-based tools to more effectively support KSC workforce planning.

2.2 CONCLUSIONS AND RECOMMENDATIONS OF REVIEW

The functionality of several current/proposed computer-based systems was reviewed. Based on this functional assessment, neither IMS nor Expert Seeker are either applicable or even relevant for Center-wide workforce planning. Both systems are merely concepts at this point, with far different objectives than those of the KSC Workforce Planning group.

As noted previously, the Expert Workforce Modeling System also does not meet these requirements. Although it could conceivably roll up workforce projections and associated skill requirements to the Center level, the level of detailed data required to conduct this analysis would be prohibitive. The skills identified in the system would have to be recrafted to align with those currently in the KSC skills inventory. IFMP focuses on the correct organizational level, and plans to conduct workforce analyses consistent with the current HM-E processes. However, due to schedule slippages, IFMP is still only in the early stages of design. Its Budgeting module is particularly ill-defined at this point.

PRMS has the most applicable functionality of the systems investigated. It has the potential to provide a consistent source of workforce requirements, in much the same way that it was utilized to compile the skills inventory. These future requirements could be input by the functional departments as well as the programs/projects. (However, procedures would have to be carefully established to insure consistent use across the Center.) Utilization of PRMS might provide more structure than the present Excel-based inputs.

However, PRMS has significant limitations with regard to the use of these future requirement inputs. It supports the bottoms-up aspect of the transaction-based strategy, but cannot apply the top-down constraints. Although new hires may be considered, there is no facility to address attrition. Finally, PRMS treats each project's set of inputs in isolation, although it appears to be able to roll-up or summarize across multiple input sets. There is no capability to avoid potential "double-counting" the same project, or identifying which KSC projects have not provided inputs. Some types of commercially-available project management software provide a facility for linking roles (using the PRMS term) to a given organization. These may then be related in a hierarchy to reflect the overall organizational structure. (The application software "Project Scheduler", marketed by Scitor, is one such example.) Incorporating analogous functionality into PRMS would address this last limitation.

Since these systems do not adequately support Center-wide workforce planning, Task 2 of this research must be initiated. This task will provide a functional systems design to better meet the needs of the Management Planning group. This proof-of-concept design may then be implemented either on a stand-alone basis or incorporated into either IFMP or PRMS.

3. TASK 2

Due to the cited limitations of the current/proposed systems at KSC, the second major task within this project was to develop a proof-of-concept design for a systems approach to workforce planning.

3.1 LITERATURE REVIEW:

Based upon the documented requirements and nature of the data sources, an extensive literature search was conducted. Keyword searches of the internet (Using the Net Search, Yahoo, and Alta Vista search engines) were conducted, as well as engineering and business databases. The literature documents a variety of mathematical modeling techniques applied to workforce planning. However, Edwards [6] notes that beyond a certain level, "the models currently available are sophisticated enough for the needs of most organizations, and that greater use of simple models would be more effective than looking for ever-increasing (mathematical) refinement". He further notes that the current trend is to use variations on one of three widely accepted modeling approaches: Markov chains, network flow models, or optimization models. Surveys conducted by both Verhoeven [7] and Gass [2] confirm this opinion.

According to Verhoeven [7], the purpose of a Markov chain model for workforce planning is to forecast the number of employees in each of the specified categories at equidistant points in time (i.e., $t = 0, 1, 2, \dots$). In a Markov chain model, the number of employees who make a transition from one category to another in a given period is assumed to be a constant fraction of the size of the first mentioned category at the beginning of the period. These fractions are called transition fractions. The actual number of employees in all categories, the transition fractions, and the future recruitment in all categories have to be given. Formulas exist to estimate the transition fractions [7]. The future number of employees in all categories (i.e. the future workforce distribution) is forecasted.

Network representations of personnel problems are extensions of the classical operations research assignment problem. For these problems, the basic form of the network is that of the more general minimal cost transshipment (or flow) network [2]. In the network, the arcs represent the flow of personnel, the source nodes represent initial personnel inventories, and the sink nodes represent the final inventories. Intermediate nodes may be established to force inventories to meet specific grade and skill goals.

Optimization models (e.g. linear programming, goal programming) are the third category of mathematical models. The model is formulated to either maximize (or minimize) the objective function (depending upon the purpose of the model) within the context of available resources and constraints.

3.2 APPROACH SELECTED FOR THIS PROJECT

NASA KSC provides a unique environment in which to conduct workforce planning. It comprises both project-oriented elements as well as those associated with functional departments. This situation complicates the selection of the mathematical modeling approach. For example, network flow models have proven to be effective forecasting tools, but *only in a project-oriented environment* [2]. Further, the proposed transition from operations to R&D negates the usefulness of such tools as linear regression, which are dependent on the relevance of historical data.

In general, the Markov model is preferred due to its flexibility [7]. In a Markov chain model, the categories can be defined such that several characteristics of the employees are included (e.g. critical skills). Attrition is reasonably assumed to be a push flow (i.e. the future workforce distribution is determined by the transitional fractions, thus employees are "pushed through the system as a result of attrition). Although optimization models can be crafted to include these additional characteristics, the size of these models becomes very large very quickly. This impacts the computational efficiency of the resulting DSS.

Like NASA, the U.S. Department of Defense has been undergoing a period of organizational restructuring, as well as a reorientation of mission, with the end of the Cold War. Case studies associated with military workforce planning were found to be very applicable to KSC's situation. Although both network models [8] and optimization models [9] have been used, the case study by Gass et al. [10] appears to be the most relevant. Here, a combination Markov chain/linear programming-based system was constructed to address Army long-range manpower planning. This integrated approach appears to support the KSC Management Planning (HM-E) group's requirements. and will be used as the basis for designing this Integrated Workforce Modeling System.

3.3 PROPOSED SYSTEM REVIEW

3.3.1 SYSTEM OVERVIEW

The purpose of the Integrated Workforce Modeling System (IWMS) is to support Center-wide workforce planning with regard to headcount, FTE, and skills requirements. A combined bottom-up/top-down transaction-based approach to workforce planning is utilized, consistent with existing methods. It will be capable of calculating the workforce requirements per year, over a five year planning horizon, at both the Center-level and for individual programs and projects. The skills profile will be optimized within these projections. Markov chain and linear programming mathematical modeling techniques are incorporated into the system's processing to provide these results. The capability to conduct "what-if" analysis, in order to facilitate the decision-making processes, is also included.

The nature of the inputs (i.e. Excel and Access) to the current Center-wide workforce planning process infers a microcomputer-based approach to developing and deploying the system. One objective of this functional systems design is to utilize existing KSC hardware, software, and communications assets to the extent possible. Continuing the use of Excel and Access inputs also suggests the use of a Microsoft Windows-based operating environment to support software compatibility and interoperability.

The recommended Markov chain and linear programming analyses may be conducted using Excel. An initial design of the IWMS Markov module was developed by employing a combination of Excel's Solver and Growth functions. Further investigation indicated that the volume of input data was too large, and the necessary mathematical models were too complicated, for Excel to adequately support.

A more powerful modeling solver engine is needed. Lindo version 6.1, marketed by Lindo Systems Inc., is such a software application. Lindo provides an interactive

modeling environment for building various types of optimization problems, including linear, integer, and quadratic programming problems [11]. It provides extensive documentation and help facilities, and can be used to construct Markov chain models. Lindo is a relatively easy software package to learn, and is frequently used in undergraduate operations research classes. Due to this functionality, as well as the fact that it executes in an MS Windows microcomputer environment, it is recommended for this project.

However, one drawback to Lindo is that it provides a limited user interface. In order to better customize the IWMS system to HM-E needs, a Visual Basic (VB) shell is suggested. This VB interface will guide the user through the system by using a series of customized displays and facilities. VB can seamlessly integrate with the Lindo solver engine into a Windows-based application [11]. In this project, VB will be used to retrieve input from the data sources, prepare Lindo model files, activate lindo.exe files, and import initial data and constants into the Lindo model. Further, it can facilitate report generation through the use of the Crystal Reports module, incorporated in the VB development software.

Projected workforce requirements by the programs/projects/functional departments, the current workforce profile, and the current skills inventory are read from their current media by the VB interface, and merged. Center workforce and skills thresholds, as well as factors for hiring and attrition, are entered via VB input screens. The Lindo software utilizes this combined input to support the mathematical modeling processes. Results are stored in an Access database, for later use by VB's Crystal Reports facility. Access was selected as the storage medium since it allows for easier manipulation by Crystal Reports, to support subsequent formatting of the system output displays and reports.

Three modules are envisioned for the IWMS system: System Inputs, Workforce Quantity Planning, and Skills Optimization. Upon system initiation, the user will be able to identify the source of the input data (or in the case of user-specified inputs, data can be entered directly). These data are then read, merged, and stored in a temporary Access file (TMP1). TMP1 provides the input to the Markov chain analysis, upon initiation of the Workforce Quantity module. TMP1, like the other temporary files, is overwritten the next time that the supplying module executes. The results stored in TMP2 (i.e. adjusted workforce requirements in terms of headcount and FTE). TMP1 similarly provides the input to the linear programming algorithms in the Skills Optimization module. Here, the profile of critical skills is optimized based on the constraints provided by the previous module. Output is sent to the TMP3 Access file. TMP1, TMP2, and TMP3 are then accessed by Crystal Reports to provide the user-requested output. Upon completion of processing, the user will be provided with the opportunity to either exit the system, or to initiate another processing iteration. This additional iteration will support "what-if" analysis, where the user may modify certain input criteria (e.g. changes in hires, or attrition) to determine the impact on the workforce projections.

4. CONCLUSIONS

According to Blanchard and Fabrycky [12], "Every system is developed in response to a need or to fulfill some anticipated function. The effectiveness with which the system fulfills this function is the ultimate measure of its utility and its value to the customer". In Task 1 of this research, a series of current and proposed systems at Kennedy Space Center were reviewed with regard to their applicability to Center-wide workforce planning.

Due to the cited limitations of these systems, Task 2 was initiated. Task 2 encompassed the development of a functional design specification for the Integrated Workforce Modeling System, which is intended to better meet the needs of the KSC Management Planning group (HM-E). Inputs, processing considerations, and resulting outputs are identified, as well as the necessary hardware and software assets. These specifications are consistent with the KSC Chief Information Officer's guidelines.

It is recommended that the IWMS be constructed and implemented as designed, as a stand-alone system. This will allow the opportunity to prove out its functionality, as well as providing the Workforce Planning group with a tool to support KSC's transition to an R&D-based organization over the near-term. It is further recommended that, in subsequent years, IWMS should be incorporated into either IFMP or PRMS discussed in Sections 2.3 and 2.4, respectively). These two "supersystems" are currently under development at KSC. They are independently developing their own standardized linkages to other Center information systems and databases (and sometimes replacing them). Assimilation will reduce the number of redundant/conflicting systems and interfaces, as well as improving long-term maintenance of the IWMS functionality.

REFERENCES

- [1] Atwater, D.M., "Research Update: Workforce Forecasting", Human Resource Planning, Vol. 18, No. 4, April 1995, pp. 50-53.
- [2] Gass, S.I., "Military Manpower Planning Models", Computers and Operations Research, Vol. 18, No. 1, January 1991, pp. 65-73.
- [3] Niehaus, R.J., Computer-Assisted Human Resource Planning, Wiley Interscience, New York, 1979.
- [4] Wessels, J. and van Nunen, J.A.E., "A Decision Support System for Manpower Planning", in Quantitative Methods in Management: Case Studies of Failures and Successes (C.B. Tilanus, O.B. de Gans, and J.K. Lenstra, eds.), John Wiley & Sons, New York, 1986.
- [5] Bridges, R.D., Kennedy Space Center Implementing NASA's Strategies, Kennedy Space Center, FL, 1998

- [6] Edwards, J.S., "A Survey of Manpower Planning Models and Their Application", *Journal of the Operations Research society*, Vol. 34, No. 11, November 1983, pp. 1031-1040.
- [7] Verhoeven, C.J., Techniques in Corporate Manpower Planning, Kluwer Publishing, Boston, 1982.
- [8] Durso, A. and Donahue, S.F., "An Analytical Approach to Reshaping the United States Army", *Interfaces*, Vol. 25, No. 1, January/February 1995, pp. 109-133.
- [9] Collins, R.W., Meinhardt, C.W., Lemon, D.M. and Gillette, M.D., "The Army Manpower Long-Range Planning System", *Operations Research*, Vol. 36, No. 1, January/February 1988, pp. 5-17.
- [10] Gass, S.I., "A Process for Determining Priorities and Weights for Large-Scale Linear Goal Programmes", *Journal of the Operations Research Society*, Vol. 37, No. 8, August 1986, pp. 779-785.
- [11] Lindo Systems Inc., "Lindo 6.1", <http://www.lindo.com/lindo.html>, June 10, 1999.
- [12] Blanchard, B. and Fabrycky, W., Systems Engineering and Analysis (3rd ed.), Prentice Hall, Englewood Cliffs, NJ, 1998.

1999 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

**JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA**

PERFORMANCE MONITORING OF DISTRIBUTED DATA PROCESSING SYSTEMS

Anand K. Ojha
Associate Professor
Department of Engineering Technology
University of Arkansas at Little Rock
KSC Colleague: Craig Jacobson, NN-J2, KSC

ABSTRACT

Test and checkout systems are essential components in ensuring safety and reliability of aircraft and related systems for space missions. A variety of systems, developed over several years, are in use at the NASA/KSC. Many of these systems are configured as distributed data processing systems with the functionality spread over several multiprocessor nodes interconnected through networks. To be cost-effective, a system should take the least amount of resource and perform a given testing task in the least amount of time. There are two aspects of performance evaluation: monitoring and benchmarking. While monitoring is valuable to system administrators in operating and maintaining, benchmarking is important in designing and upgrading computer-based systems. These two aspects of performance evaluation are the foci of this project.

This paper first discusses various issues related to software, hardware, and hybrid performance monitoring as applicable to distributed systems, and specifically to the TCMS (Test Control and Monitoring System.) Next, a comparison of several probing instructions are made to show that the hybrid monitoring technique developed by the NIST (National Institutes for Standards and Technology) is the least intrusive and takes only one-fourth of the time taken by software monitoring probes. In the rest of the paper, issues related to benchmarking a distributed system have been discussed and finally a prescription for developing a micro-benchmark for the TCMS has been provided.

PERFORMANCE MONITORING OF DISTRIBUTED DATA PROCESSING SYSTEMS

Anand K. Ojha

1. INTRODUCTION

The Kennedy Space Center has several test and checkout systems to ensure safety and reliability of spacecraft and equipment for space missions. Some of the commonly used systems are: CITE (Cargo Interface Test Equipment), PPCU (Partial Payload Checkout Unit), TCMS, (Test, Control, and Monitor System), and CMU (Control and Monitor Unit.) These test and checkout systems are essentially distributed data processing systems with the computational functionality distributed over several processors or clusters of processor interconnected through custom or commodity network. Several TCMS units are in use, and because of its interesting and open architecture, the performance of this system is the focus of this project.

System performance is one of the key concerns to designers, administrators, and users of a computer system. Information about system performance can be obtained by two methods: system monitoring and system benchmarking. Monitoring tools help in the operation and maintenance of an existing system and are indispensable tools for computer system administrators. These monitoring tools provide valuable information such as system utilization and workload statistics to help efficiently manage and upgrade the system. Benchmarking, on the other hand, applies a fixed workload to a system to determine latency, throughput, and resource utilization and is helpful in comparing the performance of two different systems under identical workload. While there are several de-facto benchmarking test-suites for CPU, no comprehensive test-suite has achieved this status in the arena of I/O benchmarking. This has been primarily due to the inherently diverse nature of the I/O systems, and has led to the development of system-specific and in-house benchmarking test-suites. A similar undertaking seems to be a viable solution for the TCMS as well. This paper covers the issues in monitoring and benchmarking of distributed systems, especially as they apply to the TCMS.

The next section provides a brief overview of the TCMS. Performance monitoring choices and issues are discussed in Section 3, and the quantitative results presented in Section 4 indicate that hybrid monitoring takes only one-fourth of the time compared to software monitoring. Section 5 provides a synopsis of existing benchmarks, analyzes several benchmarking options, and outlines the details of a comprehensive micro-benchmarking scheme for the TCMS to test not only the CPU but also I/O and operating system.

2. BACKGROUND ON TCMS

TCMS has evolved from PPCU to support the activities at SSPF (Space Station Processing Facility), launch pads, SLF (Shuttle Landing Facility), and the OPF (Orbiter Processing Facility.) It is based on UNIX/X-Windows, and is built around SGI Origin 200 multi-processor computers with the functionality spread over the internal processors and Motorola's MVME167 cards linked though Ethernet. TCMS is essentially a distributed data acquisition, processing, and control system, and supports a mix of inputs including direct discrete and analog measurements and various types of serial and telemetry formats. The TCMS was originally built around Nighthawk computer system, but to benefit from the advances in the technology, the Nighthawks

have been replaced by the Origin 200 containing R10000 dual processors. A detailed schematic of the TCMS is shown in Figure 1 (attached as the end of this paper.) Since several TCMS units are now in use, this is the system chosen for this project.

3. ISSUES IN PERFORMANCE MONITORING

There are three approaches for implementing a performance monitor: software, hardware, and hybrid. These three techniques and their pros and cons are discussed in this section.

3.1. Software Monitoring

Software monitoring is also known as software instrumentation. In the most basic software monitoring scheme, execution time of a program module is obtained by inserting system-timer calls at appropriate points in the application program. The following three such system calls of interest are available in Unix:

- `time(NULL)`: Wall clock time in seconds since 1/1/1970
- `gettimeofday(&tv, NULL)`;
 `Time_seconds = tv.tv_sec; // Integer number of seconds since 1/1/70`
 `Time_microseconds = tv.tv_usec; // Fraction seconds in microseconds`
 `// since 1/1/70`
- `Clock()`: The system call returns the CPU time in ms or μ s depending on the implementation. One needs to be very careful in using this system call, because it provides information only about CPU utilization and not the wall clock time. Instructions, such as `sleep(int seconds)`, that make the processor wait for a specified amount of seconds are not accounted for by the `clock()` instruction.

Software monitoring has three obvious limitations. First, the 1- μ s resolution of the system timer is not very good for present-day computers whose clock period is on the order of 1-ns. To obtain results with satisfactory resolution, the execution time of a program module (or for that matter of any individual instruction) is determined by executing the module a large number of times and then obtaining the average execution time; clearly, a time-consuming process.

The second problem is a result of what is known as the "probing effect" in which the inserted monitoring instructions add to the execution time of the code under measurement. The probing effect is sometimes tolerable for a standalone uniprocessor system. However, for a multi-user-multi-tasking system that has a node with multiprocessors (such as the Origin 200 used in TCMS), the probing effect may significantly degrade the system performance due to loss of synchronization of the cooperating or multi-threading processes. Probing effects are particularly detrimental to parallel programming environments.

Thirdly, the software approach provides only a partial monitoring capability, because it cannot efficiently monitor cache hits, or events caused by external conditions, such as exceptions, interrupts, etc.

3.2. Hardware Monitoring

Hardware monitoring requires an external dedicated hardware in the form of either a plug-in card or a dedicated box to monitor the signals on the bus at a processing node [1]. The dedicated hardware contains its own local memory and functions like a logic analyzer in that it

records the events of interest based on the signals captured from the address, data, and control bus of the system being monitored. Every event of interest is encoded as a unique bit pattern called an *event token* that uniquely specifies *what* external event occurred and *where* it occurred. An event token is like an ID number of an event. For every event of interest, an event record is created that contains the event token and its timestamp so that the samples could be correlated in time. Event records are first stored in the local memory of the monitoring hardware, and later downloaded on to a host for trace-generation and post-analysis.

On the positive side, hardware monitoring is least intrusive, and in addition to interrupts, it can also monitor designated events during program-execution. However, hardware monitoring is impractical for systems employing virtual memory because while the compiler generates virtual addresses, the monitoring hardware must be designed to be triggered by the physical addresses and signals. This knowledge of this virtual to physical address translation mechanism makes the design of hardware monitoring systems impractical for virtual memory.

3.3. Hybrid Monitoring

Hybrid monitoring combines the desirable features of both software and hardware techniques. The embedded software probes log the events caused by the program itself, while external events are recorded by the hardware. It is generally agreed that hardware monitoring provides an efficient solution to monitoring distributed systems.

Since there were numerous ways to design a hardware monitor, each research group initially came up with its own custom hardware monitor [2]. Later, in an effort to standardize hybrid monitoring, the National Institutes of Standards and Technology (NIST) developed a plug-in board known as the Multikron Interface Board (MIB) for PCI, S-Bus, and VME bus [3, 4]. Universities and research organizations can obtain these MIBs on loan from the NIST. Otherwise, these can be bought from the NIST for about \$750 at the time of the writing of this paper in 1999. Salient features of the MIB are described in the following paragraph.

The MIB is a memory-mapped device and can monitor up to eight CPUs sharing a bus. It has a 56-bit timestamp counter with a resolution of 100ns. In contrast, the best resolution obtainable from software system calls is only 1 μ s. Two types of samples can be collected by the MIB: *trace* sample and *resource* sample. A 20-byte trace sample contains CPU ID, process ID, user-written identifier, and the timestamp. The 84-byte resource sample, in addition, contains the contents of the 16 32-bit counters that can be started, incremented, and stopped under hardware or software control. Hence, these counters can be used to count the occurrence of events such as cache hits, duration of memory bus cycles, etc. The user simply writes an integer identifier value through a single assignment (memory write) statement; all other data are non-intrusively provided or captured by the board itself. The MIB has 16 MB of on-board RAM, and so it can store 800K trace samples in its local memory before needing a download to a host for post-analysis. A RESET signal is available on the MIB to accurately synchronize the timestamp counters of all of counters in a situation where multiple nodes are monitored by as many boards. Thus, the error in time-synchronization is limited only by wire-distance and is far superior to the software time-synchronization. The next section compares the results from MIB with those obtained from software instrumentation.

4. RESULTS FROM MULTIKRON MONITORING SYSTEM

To determine the intrusiveness of various monitoring techniques, a test platform was set up using Linux 2.2-15 (Redhat 6.0) on a 200 MHz Intel Pentium-Pro processor (8KB Instruction cache and 8KB data cache) with 256 KB L2 cache and 96 MB DRAM. GNU's EGCS 1.1 compiler that comes along with Redhat 6.0 Linux, was used. Average execution time from one million iterations are listed in Table 1 below.

Table 1. Average Execution Time of Monitoring Techniques.

Technique	Instruction	Average Execution Time
Software	time(NULL)	1.77 μ s
	gettimeofday(&tv, NULL)	2.38 μ s
	clock()	2.41 μ s
Hybrid	MIB Assignment Statement	0.36 μ s

As noted before, `gettimeofday(&tv, NULL)` has the best resolution of 1 μ s among all software instructions, whereas as mentioned before, MIB's timestamp counter has an order of magnitude better resolution. In addition, it can be observed from the above table that in terms of execution time the hybrid monitoring is more than four times better than the best software technique.

5. ISSUES IN BENCHMARKING DISTRIBUTED SYSTEMS

A benchmark suite provides a figure of merit for system-performance. This section provides a synopsis of the issues related to various benchmarks. The performance of a distributed system is impacted by the following components [5].

- CPU
- Memory
- Operating System
- I/O
- Network and Communication

Based on the level of details the test-results contain, there are two classifications of benchmarks: macro and micro [6]. A *macro-benchmark* contains minimum amount of detail in its result; generally, a number indicating how better or worse a given computer system is compared to some other system. Whereas the result from a macro-benchmark is a useful measure in comparing two different computer systems, it does not answer the question as to why a system performs better or worse than the other one. The usefulness of macro-benchmark is limited only to a user in making a decision about which system to use. However, to identify bottlenecks, integrators and designers of computer systems need detailed information. These details provide a

better insight into how the components fit and interact with each other in determining the system performance. Benchmarks that generate such detailed results are known as *micro-benchmarks*.

Another benchmark-classification is based on the system component they test. Earlier benchmarks were developed to test only the CPU, and so they consisted of computation-intensive programs to test the computational prowess of a computer system. Some others were then developed to test the file transfer or I/O capability of a computer system. A few others were primarily intended for transaction processing for a client-server system. However, no benchmark currently exists that can test all of the components of a distributed system listed above. The following sections discuss the issues and tools available to benchmark some of these individual components of a distributed system.

5.1. CPU Benchmarking

CPUs have been the focus of most computer system performance evaluation, and a major concern of such studies has been the interaction between the processor and the memory, which in turn depends on the following factors [5]:

- Instruction Set
- Hardware architecture: caching, pipelining, superscaling
- Clock rate
- Compiler
- Memory Management
- Operating System

A CPU benchmark, as the name suggests, primarily measures the computational prowess of a CPU and its interaction with the memory system. Therefore, CPU benchmark programs tend to be computation-intensive. A number of CPU benchmarks are available but the de-facto standard for CPU benchmarks is SPEC95 which consists of two suites: Cint95 and Cfp95. CInt95 is a collection of eight C programs, while Cfp95 is a collection of 10 Fortran-77 programs [7]. The output of SPEC95 benchmark suites is a number that provides a comparison of the performance of the system under test with the performance of a Digital Equipment Corporation's VAX-11/780 mini-computer of the late 70's.

5.2. File System and I/O Benchmarking

In Unix system, all I/O devices are treated as files. The performance of a computer system that includes I/Os is greatly influenced by the architecture used to interface the three components: CPU, memory, and I/O. In addition, the impact of operating system is very pronounced. Therefore, systems containing I/O devices possess a much higher level of complexity in benchmarking. The confusion in this area is further compounded by the fact that depending on the system and application, there are several performance-measures for systems consisting of I/O devices, and improving one measure might degrade the other. Thus, in contrast to processor performance benchmarking, the performance criterion for systems consisting of I/Os is not as unambiguous. The state of I/O benchmarking is still quite primitive compared to the processor benchmarking and above all is application dependent [5]. This is because of the widely varying schemes as a result of tradeoffs. Some of the notable file system benchmarks that appear to be promising for Unix systems are discussed below for the sake of completeness. Readers may consult the indicated references for further details.

The Andrew benchmark for Unix system is one of the oldest Unix file system benchmarks consisting of a suite of 70 files that take up 200 KB and run in five phases - MakeDir, Copy, ScanDir, ReadAll, and Make [8].

5.3. Operating System Benchmarking

System performance is impacted by the algorithm the OS uses to create, schedule, and switch processes, especially in multi-user or multi-tasking environment. Memory management techniques, if implemented under the operating system, also impact the performance of the system.

The resources for operating system benchmarking are almost non-existent. Nevertheless, the benchmark, LMBench [9], needs mentioning here. LMBench, is a synthetic micro-benchmark to test the operating system as well as the I/O capabilities of a computer system. LMBench relies on determining the execution time of the system and function calls needed to mimic the working of the operating system. For example, consider the problem of determining the system latency in creating a new process and running a new program. Unix performs this task by executing `fork()` and `execve()` system calls. Incidentally, this is how the Unix command-interpreter, *shell*, works in the inner loop. Thus, the latency in creating a new process and running a new program can be obtained by determining the execution time of `fork()+execve()` system calls. Clearly, a better operating system should have smaller execution time..

5.4. Prescription for a Comprehensive Micro-Benchmark Suite for TCMS

A test and checkout system is a complex distributed data processing system, and in addition to the components discussed above, consists of heterogeneous analog and discrete signals, multiplexers/demultiplexers, and networked systems with multiprocessors.

It should nevertheless be observed from the discussion in the preceding sections that due to the specific application and often competing requirements, a macro-benchmark is impractical in evaluating distributed systems. Instead, a better insight into all these issues can be obtained by quantifying the parameter that show how the components belonging to the system fit and interact with each other in determining the system performance. In other words, a micro-benchmark test suite that can test the components that impact the performance of a distributed system seems to be the best alternative. The idea here is to have a benchmark that can help the design and upgrade of the system. For example, consider the upgrade of the R10000 processors in Origin 200 by a later version. The benchmark should not only provide the gains for computation intensive tasks, it should also shed light on the new performance results for operating system and I/O operations as well. As a second example, consider replacing the optical drives of the Origin 200 system by faster drives. One would expect the file-transfers to be faster, but a CPU-bound system might not show a commensurate improvement in the file transfer rate because of the processor's inability to cope up with the more frequent interaction with the faster drives. These examples highlight the significance of a micro-benchmark in system design and upgrade. A comprehensive micro-benchmark being proposed here should have the following capabilities.

5.4.1 Computational Capability: This should consist of a set of computation-intensive programs to determine the computational capability of the host. In a multiprocessor system, the benchmark should provide information on the utilization of individual processors. A set of computation-intensive programs, such as LINPACK, should suffice [6].

5.4.2 Memory Performance: The following tests should reveal performance of cache and DRAM.

- Reading a block of data from resident memory (Test for different block sizes)
- Writing a block of data to resident memory (Test for different block sizes)
- Reading a block of resident memory and writing to a different block (Test for different block sizes)

5.4.3 File System and I/O Capability: Since I/O devices are treated like files in Unix, the following tests can benchmark both file systems and I/O devices.

- Time taken in creating/deleting files
- Time taken in writing to a file using *putc()*
- Time taken in reading from a file using *get()*
- Time taken in sequentially writing a block of data to a file or I/O device (Test for different file sizes): *write(int handle, void *buf, unsigned len)*
- Time taken in sequentially reading a block of data from a file or I/O device (Test for different file sizes): *read(int handle, void *buf, unsigned len)*
- Time taken in reading/writing random files: In this case, the buffer pointer is randomized by *rand()* in the function *lseek(int handle, long offset, int fromwhere)*. It should be noted that this randomized I/O function can be used to determine the time it takes for a disk drive to position its head at a random location.

5.4.4 Operating System: The following tests should reveal how efficiently the OS can create, schedule, and switch processes. Networking as well as the operation of many other tasks can be synthesized and the system benchmarked.

- Time to create a process
- The latency in context switching
- Data transfer rate through a Unix pipe
- Opening and closing TCP/IP connection

5.4.5 Network and Communication: It should be observed from the schematic of the TCMS that the user interacts with the AP or ARS through a number of X-Terminals connected to a common Ethernet. Since the user interaction is random, it is not realistic to obtain a deterministic benchmark value or a figure of merit. A possible approach would be the installation of a network monitor capable of providing important network statistics such as number of packets transferred, number of collisions, throughput, etc. Inexpensive COTS (commercial-off-the-shelf) network monitors are available. Conclusion about how the network is performing can then be drawn from these statistics. For example, if the number of collisions is high, the network is saturated either because of increased user requests or due to the bandwidth limitation of the network. The same strategy may be applied to the other Ethernet link that interconnects the FEPs (Front End Processor) of the TCMS to the AP and ARS in the upper half of the figure.

Unfortunately, the X-terminals are not intelligent terminals; they have display capabilities but they do not have any kernel. If they had a kernel, a benchmark program running on either the AP or the ARS could have been designed to emulate the user interaction by implementing a client-server model to account for the communication between AP/ARS and the user terminals. This aspect, however, needs more investigation.

The micro-benchmark suggested here would not only be useful for the test and checkout systems, but would also have the capability to test the CPU, operating system, and I/O capability of any generic computer system.

6. CONCLUSION

This paper discussed monitoring and benchmarking issues for performance evaluation of distributed systems with special emphasis on the test and checkout systems. Details of a prototype Linux-based hybrid monitoring system, that employs NIST's PCI-MultiKron card, have been provided. Results obtained from the monitoring system indicate that compared to software monitoring, the hybrid monitoring reduces probing effect by at least a factor of four. Finally, issues related to benchmarking a distributed system have been discussed and the details on how to develop a micro-benchmark for the TCMS have been provided.

7. ACKNOWLEDGEMENT

Thanks to Craig Jacobson and Robert Yascovic from the NASA Kennedy Space Center for their support. Thanks also to Wayne Salamon, NIST, Gaithersburg, MD, for his assistance in upgrading the Multikron code and for his prompt clarifications on several operational aspects of the Multikron interface board.

8. REFERENCES

- [1] R. Hoffman, et al., "Distributed Performance Monitoring: Methods, Tools, and Applications," *IEEE Trans. Parallel Distributed Systems*, Vol. 5, No. 6, pp. 585-597, 1994.
- [2] J. C. Harden, D. S. Reese, M. B. Evans, S. Kadambi, G. J. Henley, C. E. Hudnall, and C. Alexander, "In Search of a Standards-Based Approach to Hybrid Performance Monitoring," *IEEE Parallel and Distributed Technology*, Vol. 3, No. 4, pp. 61-71, Winter, 1995.
- [3] A. Mink and W. Salamon, *Operating Principles of the PCI Bus MultiKron Interface Board*, NIST Report No. NISTIR 5993, U S Department of Commerce, Gaithersburg, MD, March 1997.
- [4] A. Mink, et al, "Performance Measurement using Low Perturbation and High Precision Hardware Assists," *Proc IEEE Real-Time Systems Symposium, Madrid, Spain, pp 379-388, December 1998.*
- [5] D. A. Patterson and J. L. Hennessy, *Computer Organization and Design: The Hardware/Software Interface*, Fifth Edition, Morgan Kauffman, San Francisco, CA, 1998
- [6] K Hwang, and Z. Xu, *Scalable Parallel Computing*, pp. 2631, WCB McGraw Hill, Boston, MA, 1998.
- [7] SPEC95 Benchmark, SPEC, Santa Clara, CA. (<http://www.spec.org/osg/cpu95/>).
- [8] J. H. Howard, M. L. Kazar, S. G. Menees, D. A. Nichols, M. Satyanarayanan, R. N. Sidebotham, and M. J. West, "Scale and Performance in a Distributed File System," *ACM Transactions on Computer Systems*, 6(1), February 1988.
- [9] L McVoy and C. Staelin, "lmbench: Portable Tools for Performance Analysis," *Proc. Of the 1996 USENIX Technical Conference*, pp. 279-295, San Diego, CA, January 1996. (<http://www.bitmover.com/lmbench/>)

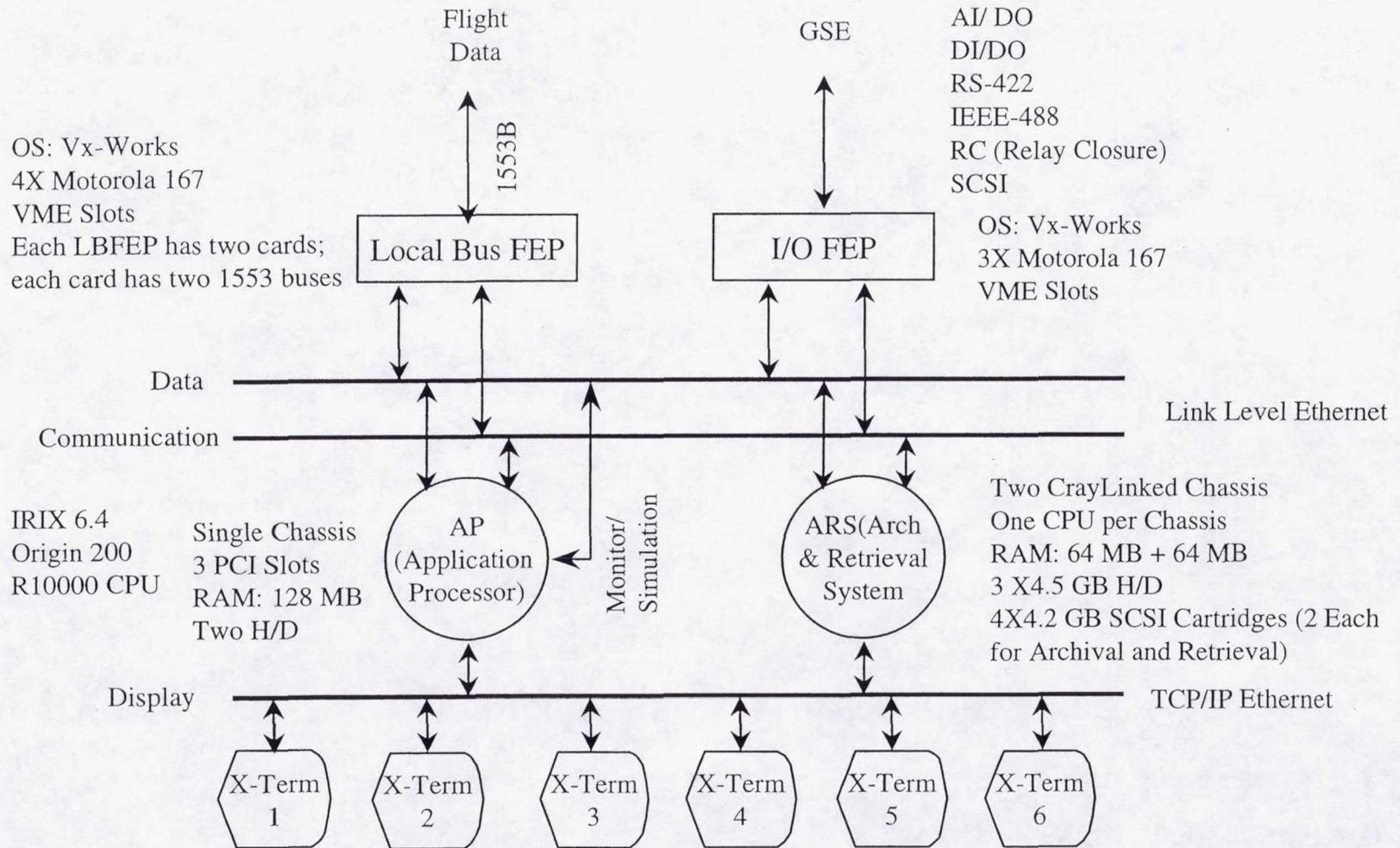


Figure 1. Schematic of TCMS

1999 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

**JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA**

CONTROL FOR NO_x EMISSIONS FROM COMBUSTION SOURCES

Maria E. Pozo de Fernandez, Ph.D.
Assistant Professor
Chemical Engineering
Florida Institute of Technology

Michelle M. Collins, Ph.D., P.E.
NASA-KSC
Environmental Program Office

ABSTRACT

The Environmental Program Office at the Kennedy Space Center is interested in finding solutions and to promote R&D that could contribute to solve the problems of air, soil and groundwater contamination. This study is undertaken as part of NASA's environmental stewardship program. The objective of this study involves the removal of nitrogen oxides from the flue gases of the boilers at KSC using hydrogen peroxide. Phase I of this study have shown the potential of this process to be used as an alternative to the current methods of treatment used in the power industry.

This report summarizes the research done during the 10-week summer program. During this period, support has been given to implement the modifications suggested for Phase II of the project, which focus on oxidation reactions carried at lower temperatures using an ultraviolet source. The redesign and assembly of the modifications for the scrubbing system was the main objective of this research.

CONTROL FOR NO_x EMISSIONS FROM COMBUSTION SOURCES

Maria E. Pozo de Fernandez, Ph.D.

1. INTRODUCTION

Over the years, the nature of the operations at the Kennedy Space Center including the use and handling of hazardous propellants and toxic substances have contributed to air, soil and groundwater contamination. Today better knowledge in the handling of hazardous and toxic materials and stringent regulations about their use and disposal has minimized the harmful impact and the risk that these substances pose to the environment. The Environmental Program Office through the Remediation Group at the Kennedy Space Center is interested in finding ways to obtain immediate and permanent solutions to these problems. Nitrogen Oxide (NO_x) emissions are a primary criteria pollutant regulated by the USEPA. KSC is permitted to emit approximately 60 tons of NO_x annually under their Title V air permit. NASA is pursuing R&D to minimize the detrimental environmental effects of KSC operations on the environment. This study is undertaken as part of NASA's environmental stewardship program.

This study involves the removal of nitrogen oxides from the flue gases of the boilers at KSC. The primary focus of the research is the conversion of NO_x to nitrogen acids for the purpose of scrubbing them from the gas stream. NO is virtually non-soluble and NO₂ is only slightly soluble; however, nitrogen acids (HNO₂ and HNO₃) are highly soluble and can be removed via scrubbing. Phase I of this study was completed in December 1998 and consisted of oxidation of the NO_x at high temperatures (~930 °F). Phase II began in January 1999 and is expected to be completed by December 1999. Phase II will focus on oxidation at lower temperatures utilizing an ultraviolet (UV) light source. Key to the success of this study is optimizing the scrubber operation. The redesign included the following: 1. Conversion of batch to continuous reservoir flow. 2. Lower temperatures in the scrubber recirculation system. 3. Continuous and Batch caustic feed. 4. Providing for sampling and analysis in-line. 5. Increase recirculation flow rate.

Support was also provided in the coordination, implementation, assembling, logistics and follow-on procedures for the changes needed in the Reaction Zone of the Unit and in other areas of the project. This report summarizes the research done during the 10-week program at NASA-KSC. During this time the whole unit was assembled, with the exception of the UV lamps. Hydraulic tests were carried out at the scrubbing system.

2. BACKGROUND

Laboratory studies have demonstrated that hydrogen peroxide (H₂O₂) when injected under proper conditions into hot gases of the exhaust stream it oxidizes nitric oxide (NO) into NO₂, HNO₂ and HNO₃. The formation of nitrogen acids allows for more inexpensive methods of post treatment for nitrogen removal, such as, scrubbing. Sulfur

dioxide can also be added to the combustion source to simulate the exhaust of an industrial power plant.

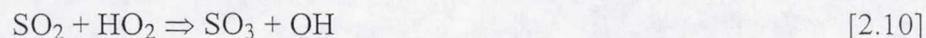
The primary reactions that are expected to occur during this process are as follows:



Hydroxyl-Radical Reactions:



Hydroperoxyl-Radical Reactions



To simulate the kinetics of the process, several parameters are needed, such as, activation energies, reaction temperatures, reaction intermediates and reactor flow regime, just to mention a few. Analysis of the reactions through reaction modeling determined that these reactions should take place within 0.1-0.3 seconds. The model used did not consider the effect of sulfur dioxide added to exhaust gases.

To demonstrate the viability of such process, researchers at NASA-KSC and the University of Central Florida (UCF) joint efforts to implement this process from a laboratory scale to a pilot plant scale. The experimental setting is located at the KSC Central Heat Plant (CHP).

Phase I of this project was performed on a 35 mmBTUH natural gas boiler. The fuel used consist mainly of methane and varying amounts of ethane, propane, butane, a sulfur-containing mercaptan added to natural gas, and inerts, such as, nitrogen, carbon dioxide and helium. When burning natural gas the major pollutant in the exhaust gases is mainly NOx. The percentage of NOx presence depends on the temperature of the combustion chamber as well on the fuel/oxygen ratio.

Part of the flue gas was diverted to the experimental apparatus consisting of different sections: Injection Zone, Reaction Zone, Quenching and Scrubbing. At the Injection Zone hydrogen peroxide (H_2O_2) was injected to the system. Also NO and SO_2 was added to the exhaust gases to simulate the composition of a flue gas from an industrial power plant. Once passed the Injection Zone the mixture of flue gas, NO, SO_2 and (H_2O_2) went through an auxiliary gas burner to bring the slip stream gases to the desired reaction temperature. From this point, the gases went through the Reaction Zone, where the conversion of NO to NOx and nitric acids took place. The Reaction Zone consisted

of a 12-inch diameter, 8-ft long stainless steel pipe, having sampling ports and thermocouples distributed along its length. Reactions could be carried out up to 500 °C. The reaction products and by-products passed then through the Quenching system, where water was added to cool off the gas mixture before entering the Scrubbing System. The scrubber was a packed column consisted of an 8-ft tall acrylic column with a 6-ft bed depth. 1-inch Hyflow 25-7 polypropylene rings were used as packing material. The scrubber had a 30x60x30 inch³ reservoir, a 1¼ - inch PVC recirculation line with a Hayward diaphragm valve, a by-pass valve, a ½ HP submersible pump and reservoir drain and sampling port. Sodium bicarbonate was used to control the pH of the fluid at the reservoir.

Data Acquisition System (DAQ) using LabView software from National Instruments was used to obtain instant readings of the key parameters in the process. Figure 2.1 shows the schematics of the experimental system as the computer monitor displays it.

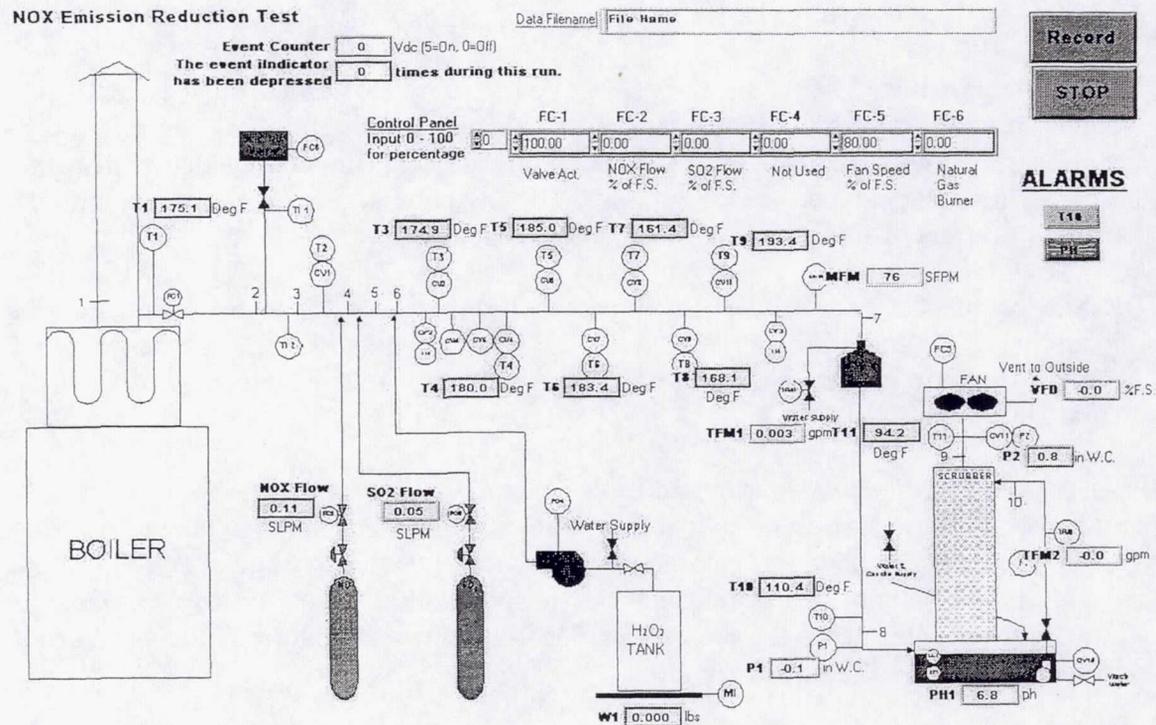


Figure 2.1. Schematics of the Experimental Process (Phase I)

Several runs were made with this experimental configuration changing the following variables: Reaction temperature, NO, SO₂ and H₂O₂ concentration, NO/ H₂O₂ ratio, SO₂/H₂O₂ ratio and residence time.

The results obtained during Phase I were encouraging but some modifications would be required in order for the system to perform to the optimum expected values. The conversion from NO to NO_x worked very well, getting conversions greater than 96% for NO/ H₂O₂ ratios of 1:1. Conversion from NO₂ to nitric acids was not as high as expected. Operating the reactor at high temperatures and having the surface of the stainless steel reactor without any type of coating could have been part of the problem. Some authors believe that iron can act as a catalyst leading to reverse reactions. The removal of the nitric acids by the scrubber was not as efficient as expected. The scrubbing system was never optimized for the kind of removal required by the process. The submersible pump provided was under-specified, since the rates of liquid flow recommended for optimal scrubbing operation were never achieved. The implementation of these modifications led to Phase II of the project.

3. EXPERIMENTAL

Based on the results obtained during Phase I, several modifications were proposed for Phase II of the project. The sections to be re-designed are the Reaction Zone and the Scrubbing System.

Reaction Zone: The modifications proposed for the reaction zone were as follows:

1. Coating of the reaction zone pipe:

It was proposed to coat the interior of the 12-inch diameter, 8-ft long stainless steel pipe with a boron-nitride paint to prevent the release of iron during the reaction. From findings in the literature it has been suggested that the use of this kind of paint will help in the NO₂ to nitric acids conversion. The paint selected was high temperature ceramic boron nitride paint from Carborundum.

Before the pipe could be sent to the paint shop, it was necessary to dismantle the apparatus previous configuration. In order to do that, several groups within KSC needed to perform several tasks in an orderly manner. First the "insulators" needed to take care to remove the insulation in order to have access to the pipe. Once the insulation was out, the bolts that held the pipe flanges in place, the thermocouples and sampling ports needed to be removed. This turned out to be a big problem, since some of the bolts were corroded and "frozen" in place. Once this task was accomplished, personnel from "Heavy Equipment" needed to move the pipe to the paint shop. To carry out the whole process until the pipe finally arrived to the paint shop took several weeks.

The process of painting the pipe with the boron-nitride paint was not as simple as one could think. On a first attempt the pipe was painted without any prior surface treatment. It turned out that the paint did not adhere well to the surface of the pipe and it did not have a smooth finish as expected. The paint could be easily scratched from the surface. Also, there were brown spots (pitting) all over the interior surface of the pipe, which were not expected to appear. It was proposed to re-paint the pipe but this time the surface to be painted was sandblasted in order to provide a better surface preparation for adhesion.

Thin coats of paint were added to the surface, letting them dry well before each application. This time, the pipe interior surface looked smoother and the paint showed good adhesion. Still the brown spots appeared this time. After some debate, it was demonstrated that the "brown spots" were signs of corrosion occurring at the surface of the pipe. The facility where the pipe was painted did not have a "controlled environment" room, meaning that the pipe was exposed to humidity and heat, which are quite high during the summer months in Florida. To prove this point, two identical smaller pipes were painted and one was let dry outdoors and the other one was placed inside in an air-conditioned room. The first pipe showed signs of corrosion the very first day, whereas the second pipe seemed fine. By talking to the vendor, pitting should not occur at all. He also suggested that the pipe should be heated at a high temperature prior painting the surface. This would ensure that the surface would have the "enamel" type finish desired. It was concluded that even with the corrosion spots, the pipe was ready to be put back in place. Figure 3.1 shows surface finish of the reaction zone prior to be taken back to the Central Heat Plant (CHP).

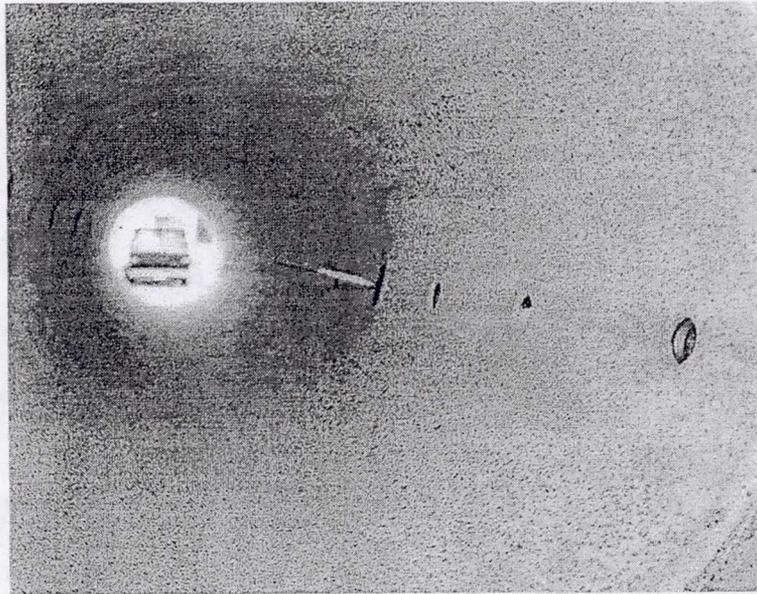


Figure 3.1. Reaction Zone surface

2. Use of UV lights:

It was suggested to use UV light as a heat source to carry the reaction at lower temperatures. Two UV lamps were bought from Calgon in Canada. In order to accommodate the lamps, a spool was designed and plates were welded to ensure the proper alignment of the lamps. The spool was made of a 12-inch diameter, 2-foot long stainless steel pipe having end slotted flanges to be bolted between the Injection Port and the Reaction Zone. It took several weeks before the spool was ready to send to the paint facility due to a delay on shipment of the plates that needed to be welded to support the lamps. Once the welding was finished, the spool was sent to the paint facility where its interior was sandblasted and coated with the ceramic boron nitride paint as it was done

for the reaction zone. Since the spool was made of a new pipe, it did not show as much pitting as the reactor zone did. The method for painting the spool was the same as the one used to paint the other pipe. Once the spool was painted, both the spool and reaction zone were delivered to the CHP plant to be mounted in place. This time, special precaution was taken to use anti-seize compound in the bolts used, to prevent rusting. Figure 3.2 shows the spool and reaction zone in place. The picture also shows the thermocouples in place, the new sampling ports and part of the system with the new insulation.

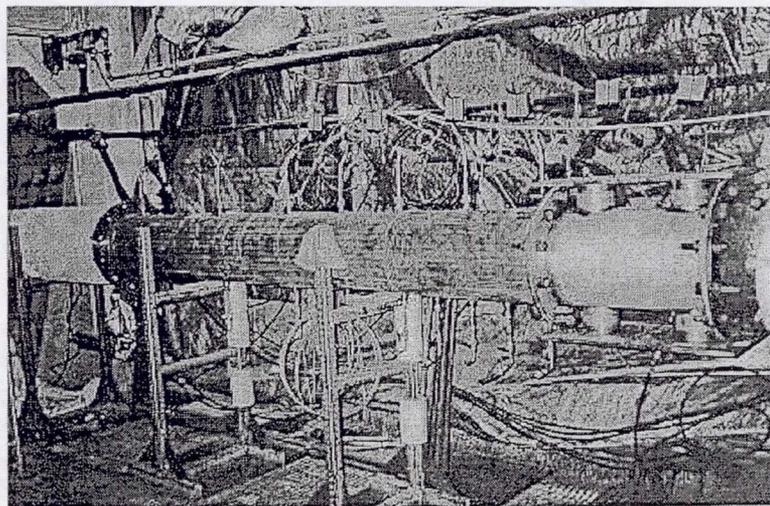


Figure 3.2. Reactor zone and spool in place

Once the whole unit was in place, new insulation was placed over the reaction zone. The delivery of the lamps was delayed for almost two months. Several problems in the fabrication of the lamps itself and its power supply system delayed the shipment. The lamps and power supply were delivered during the final week of the Summer Program. Figure 3.3 shows the reaction zone and spool with the UV lamps in place.

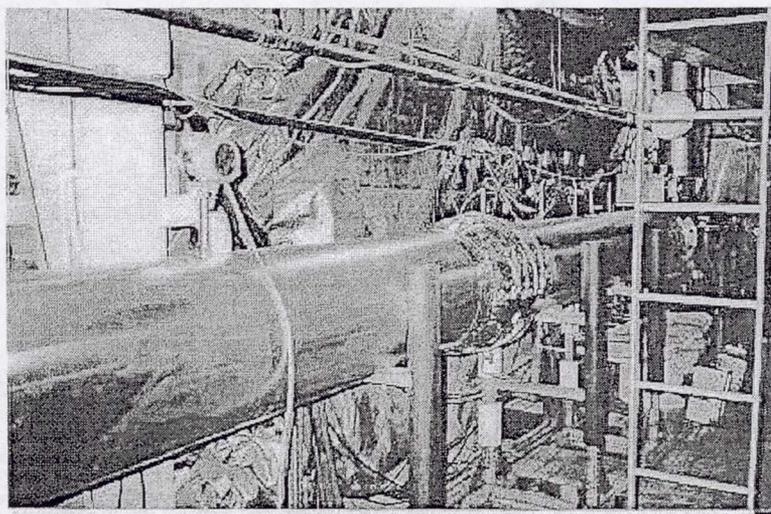


Figure 3.3. System with insulation and UV lamps in place

Scrubbing System: Gas absorption is a mass-transfer operation where a soluble vapor is absorbed by means of a liquid in which the gas is more or less soluble. This operation is commonly known as scrubbing. The apparatus used in gas absorption is the packed tower. The packed tower has gas inlet and distributing space at the bottom, liquid inlet and distributor at the top, gas and liquid outlets at top and bottom, and it is filled with solids of different shapes called *packing material*. The purpose of the packing material is to provide a surface for good contact to favor the mass transfer between the liquid and the gas. The packing could be placed inside the tower in a random manner or in order. The packing should be inert to the fluids used in the process.

A configuration of the tower previously mentioned is *counter-current* since the liquid and the gas flow in opposite directions. A good contact between the liquid and the gas is important to obtain the desired efficiency of the process. The optimum liquid and gas flow rates through the column depend on many factors, such as, the type of packing, the depth of the packing, number of packing stages, etc. Two common problems that are present in packed towers are *channeling* where the packing surface could be dry or covered by an stagnant film of liquid, and *flooding* when the column is filled with liquid.

Every packed column has its own *hydraulics*, meaning the points of loading and flooding which limits the flow rates within the tower. To characterize a scrubber several factors has to be taken into consideration, such as, components to be separated, type of packing, characteristics of the packing, mass transfer properties of the components to be separated, temperature of operation, gas and liquid flow rates within the column, just to mention a few.

The scrubber used during Phase I of the project was designed and manufactured by Rauschert Industries, Inc. Based on the parameters supplied, they estimated the size, material, packing material, optimum gas and liquid rate within the column and the operation temperature. The scrubber specifications were given in section 2. The packing material selected by the vendor was 1-inch Hyflow 25-7 polypropylene rings. This packing material is proprietary, meaning that its characteristics and hydraulic behavior can not be found in the open literature. Once the vendor specified the unit, it recommended 500 scfm for the gas rate and 35 gpm for the liquid rate. Since the recirculation pump provided with the unit was under-specified the maximum liquid flow achieved was approximately 15 gpm, which proved to be too low for the scrubber operation.

An attempt was made to get information from the vendor in relation to the hydraulics of the packing material to perform calculations to obtain the optimal relationship for the gas and liquid rates at the temperature of separation. For a given system and packing material a characteristic curve relating the HTU (height of transfer unit) and gas flow rate can be obtained. With the gas flow rate and mass transfer data the optimum liquid flow rate can be obtained. Unfortunately, the system used in the scrubber is quite complex, since there are several components present and also sodium bicarbonate is added to the reservoir to maintain a favorable pH. The changes made in the scrubber were based on the original design but providing some improvements. (See Figure 3.4)

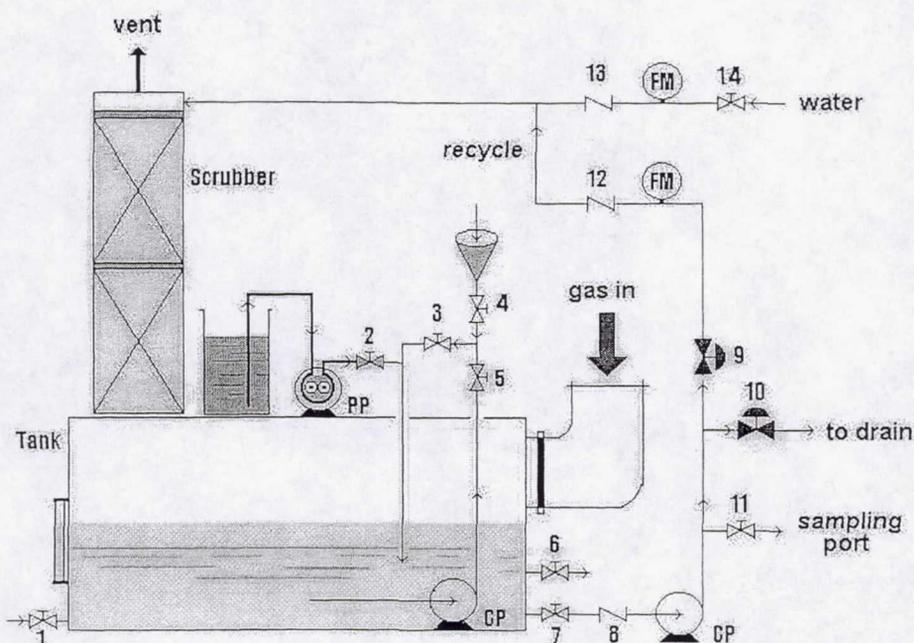


Figure 3.4. Schematics of the Scrubber Modifications

First, the scrubber was re-designed for continuous operation having inlet, outlet and recirculation lines. Fresh water was introduced in the system through the inlet line. Two Hayward diaphragm valves control the flow through the outlet line and the recirculation line in the system. Mass flowmeters hooked up to the DAQ provided accurate flow measurement at the inlet, outlet and recirculation lines. The lines were made of 1¼ - inch PVC pipe and fittings. Check valves were placed in each line to prevent back-flow. Liquid recirculation was provided through a new 1-HP pump, able to deliver 35 gpm if desired. The system was tested for leaks and it worked fine.

A continuous and batch caustic feed was also added to the system, to provide versatility in the operation mode. For continuous operation, a reservoir for the caustic solution was connected to a peristaltic pump to deliver the solution at a given rate. If it is needed, a funnel was provided to introduce sodium bicarbonate to the reservoir in large quantities. The submersible pump at the reservoir provides continuous mixing of the tank fluid and the caustic solution added to the tank. Figure 3.5 shows the actual scrubber with the modifications.

Variables: Once the UV lamps are connected and the whole system is ready for operation, several parameters will be varied. The parameters to be varied are as follows:

- 1) Reaction temperature: Test will be carried at three temperatures 350, 400 and 450 °F.
- 2) Location of Injection point: It is very important during the first test to locate the optimum position for the injection of H₂O₂.
- 3) NO₂/H₂O₂ ratio.
- 4) SO₂/H₂O₂ ratio.
- 5) pH of the scrubber reservoir solution.
- 6) Recycle flow at the scrubber.

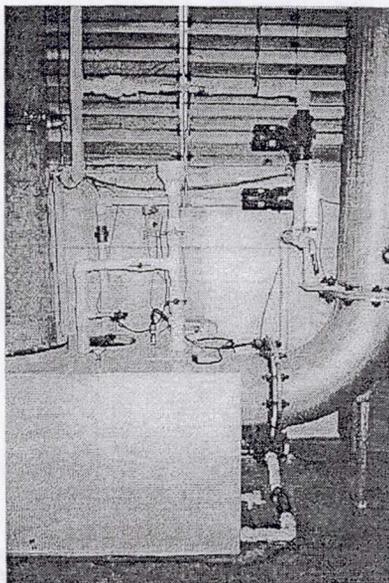


Figure 3.5. Actual Scrubber with Modifications

4. CONCLUSIONS

The objectives during this study were accomplished. The modifications suggested for Phase II of this project were implemented. Once the UV lamps are properly connected, the whole unit will be ready for testing. The scrubber modifications are in place and the unit will be able to deliver the required liquid flow rate of 35 gpm. A continuous/batch caustic feed was added to the system.

5. REFERENCES

- [1] Collins, M. M. (1998) Ph.D. Dissertation. University of Central Florida.
- [2] Collins, J. G. (1999) M.Sc. Thesis. University of Central Florida.
- [3] Coulson, J. M. and J. F. Richardson (1991) *Chemical Engineering*, Vol.2, Pergamon Press, USA.
- [4] McCabe, W. L.; J. C. Smith and P. Harriot (1985) *Unit Operations in Chemical Engineering*, 4th Edition, McGraw-Hill, Inc., USA.

6. ACKNOWLEDGMENTS

I would like to thank my NASA colleague Dr. Michelle M. Collins for providing the opportunity to work in this project. Special thanks to the members of the BES Group: Christal Banks, Debra Erving, John Collins, Dr. Chris Clausen, Dr. David Cooper, Dr. John Dietz and the UCF graduate students for their help and support. My sincere appreciation to the personnel of the Central Heat Plant, especially Gary Relay, Tom Petelle, George Broyles, and Tommy Sizemore. Without their help, support and expertise it would have been impossible to accomplish the tasks of this project. Special thanks to Dr. Jackie Quinn, Rosaly Santos-Ebaugh, Lu Richards, Harrold Williams, John Adkinson and Janice Everet of the Remediation Group at KSC for providing a friendly working environment. Special thanks to Dr. Ramon Hosler, Dr. Jane Hodges, Gregg Buckingham and Judie Gilliam from the NASA/ASEE Summer Program for all their efforts. This has been a wonderful research experience.

1999 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

**JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA**

***OPTIMIZATION OF CONNECTOR POSITION OFFSET FOR BANDWIDTH
ENHANCEMENT OF A MULTIMODE OPTICAL FIBER LINK***

Dr. Banmali S. Rawat
Professor
Department of Electrical Engineering
University of Nevada
Reno, Nevada

KSC Colleague- Po T. Huang
Optical Fiber Communications Laboratory

ABSTRACT

The multimode fiber bandwidth enhancement techniques to meet the Gigabit Ethernet standards for local area networks (LAN) of the Kennedy Space Center and other NASA centers have been discussed. Connector with lateral offset coupling between single mode launch fiber cable and the multimode fiber cable has been thoroughly investigated. An optimization of connector position offset for 8 km long optical fiber link at 1300 nm with 9 μm diameter single mode fiber (SMF) and 50 μm diameter multimode fiber (MMF) coupling has been obtained. The optimization is done in terms of bandwidth, eye-pattern and bit pattern measurements. It is simpler, is a highly practical approach and is cheaper as no additional cost to manufacture the offset type of connectors is involved.

OPTIMIZATION OF CONNECTOR POSITION OFFSET FOR BANDWIDTH ENHANCEMENT OF A MULTIMODE OPTICAL FIBER LINK

Banmali S. Rawat

1. INTRODUCTION

Recently there has been a demand for increased data transmission rate of local area networks (LAN) in the form of new Gigabit Ethernet standards. These high-speed optical links are required for offices, buildings and campus backbones. One of the most important requirements of these high-speed links is low cost, which can not be achieved by replacing the existing narrow band MMF by large bandwidth single mode fiber links. Unfortunately the bandwidth – distance product of these multimode fiber links is limited due to modal dispersion. It has been observed that even at 1300 nm wavelength for low dispersion MMF, the maximum bandwidth capacity for 62.5 μm diameter multimode fiber is only 500 MHz.km for over-filled- launch (OFL) conditions which is far short of the required 1 GHz.km. The modal dispersion in the MMF is caused due to various propagation paths for different modes. The higher order modes travel farther away from the fiber axis thus taking longer time to travel the same distance as the lower order modes traveling close to the fiber axis. This time difference in traveling modes results into pulse broadening or dispersion, which reduces the fiber capacity. Haas and Santoro [1] in 1991 reported a method to overcome this problem in MMF by using single-mode launching to multimode to single –mode reception. The idea was to launch only fundamental mode into multimode fiber and filter out the fundamental mode on reception. This splicing provides larger bandwidth but results into heavy signal losses especially at the receiver splice due to energy losses in higher order modes. Later these authors proposed another scheme of selective excitation of higher-order modes by using single mode fiber launch into multimode fiber at an angle as shown in Fig. 1, [2]. This method provides fairly good improvement in the bandwidth but maintaining the angular offset under vibrations is not easy. In an another SMF to MMF launch technique, a small launching spot is radially offset from the MMF core, Fig. 2, [3]. In this technique the authors have experimented three methods of launching light from SMF into MMF. The first one uses two lenses, an initial collimating lens followed by a focusing lens so that the spot size at MMF can be varied. In the second method a fiber lens is placed close to the MMF end resulting in a very small spot. In the third method, the SMF end is placed against the MMF with an offset. These methods provide up to four times bandwidth enhancement even with large number of fiber modes being propagated. However, focusing the spot on MMF under vibrations and other ambient conditions may be difficult. The method under investigation at KSC uses the SMF to MMF launch using connector position offset rather than simple fiber offset. The connector position offset is simple, highly practical, sturdy, cheaper as no additional cost to manufacture the offset connector and is easy to fabricate and maintain even under adverse conditions. The main drawback of offset connector launch is the attenuation of the signal and introduction of noise. Therefore in order to maintain proper signal to noise ratio, the

optimization of offset is very important. The main objective of this project is to do an experimental study of the optimization of connector position offset for bandwidth and attenuation of optical signal at higher bit rate.

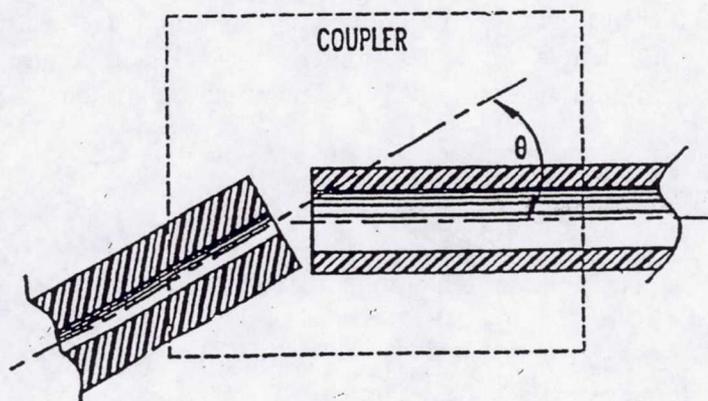
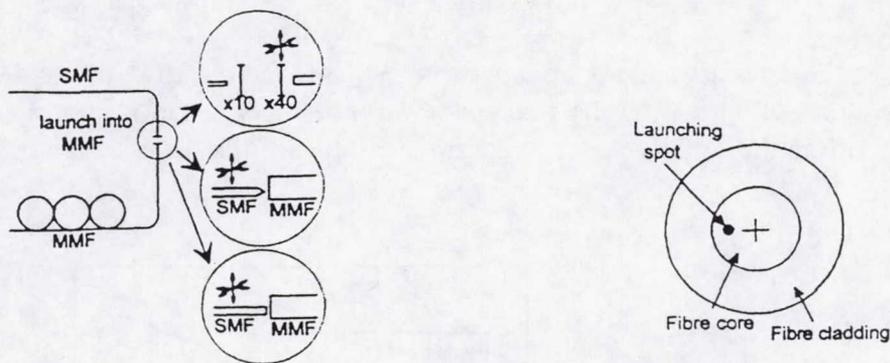


Fig.1. Angular offset coupler for SMF to MMF launch



(a) Three different launching schemes

(b) Realization of offset launch

Fig.2. SMF to MMF coupling using radially offset launching spot

2. BASIC CONCEPT AND ANALYSIS OF OFFSET CONNECTOR LAUNCH TECHNIQUE

The basic concept of offset launch is to excite only a group of all the modes of the MMF at the launch thus reducing the pulse broadening due to lower modal dispersion. This lower modal dispersion results into increase in bandwidth. It is observed that the relative propagation delay is dependent on the refractive index profile of MMF and mode group order in the fiber as shown in

Fig. 3. It is evident that the resulting pulse broadening can be reduced by considering a selective excitation of modes as the modal group delays change linearly with mode number [5]. Also the higher order mode groups contain more modes than the lower order ones. It means the selective mode excitation still propagates larger number of modes resulting into reduced modal noise along with reduced dispersion. However, the noise is certainly higher compared to OFL condition, therefore the optimization of connector position offset is necessary. The modal propagation constants and delays can be determined for MMF with a pure "power law" refractive index profile and through analytic solutions [4]. The effective number of excited fiber modes are given as

$$m = (\sum a_i)^2 / \sum a_i^2 \quad (1)$$

where a_i is the power in the i^{th} mode. The mode dependent losses are obtained from loss coefficient of the type [3],

$$P_{v\mu} = \exp(x \cdot R_{v\mu}^{2y}) \quad (2)$$

for $LP_{v\mu}$ th mode with x and y as fitting constants and $R_{v\mu}^2$ as the relative propagation constant given as

$$R_{v\mu}^2 = \frac{1}{2} \cdot [n_{\text{core}} / (n_{\text{core}} - n_{\text{clad}})] [1 - \beta_{v\mu}^2 / k_0^2 n_{\text{core}}^2] \quad (3)$$

where n_{core} is the refractive index at core center, n_{clad} is the cladding refractive index, $\beta_{v\mu}$ is the propagation constant and k_0 is the free space wave number. The fitting constants x and y are obtained by fitting the function to the experimental data of the dependence of detected power on the launch position.

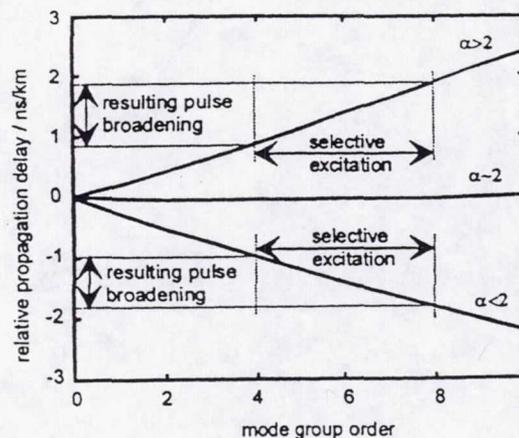


Fig.3. Relative propagation delay for various mode groups as a function of refractive index profile parameter α , with fiber diameter of 50 μm and 1300 nm wavelength (Ref. [3]).

If all the modes of a MMF are not excited perfectly, the bandwidth varies due to changes in the launch power distribution. The tuning or excitation of the modes in the fiber is determined in terms of differential mode delay (DMD) measurement [5]. The DMD measurement is important for determining and improving the performance of MMF as any variation in the refractive index profile from the optimal refractive-index profile is easily determined by measuring DMD compared to measuring refractive-index profile. At the same time the noise for offset connector condition increases due to reduced number of modes. Under OFL condition, the noise is reduced due to noise phase cancellation of various modes while for offset condition this cancellation does not take place due to reduced number of modes.

3. EXPERIMENTAL STUDY OF OFFSET CONNECTOR UNDER INVESTIGATION AT KSC

The offset connector under study at KSC consists of a standard FC type connector at the end of input SMF coupled to a standard ST type connector at the end of output MMF. This connector assembly is housed in a 12.4-cm long cylindrical metallic container. The input SMF (yellow color) in the form of a cable has a diameter of 9 μm and is about 10 m long while the output MMF (orange color) also as a cable has a diameter of 50 μm and length of 8 km. The complete offset connector assembly on a positioning system is shown in Fig. 4. The input power from the source is kept at about -0.35 dBm and the output power at the end of 8 km long cable for no-offset condition is about -8.3 dBm i.e. a loss of 7.95 dB in the fiber. For offset conditions the power at the end of 8 km long MMF cable depends on the size of offset. The measurements of bandwidth, bit-pattern and eye-pattern were conducted using offset connector coupler. For comparison purpose all the measurements were conducted for 500 and 700 Mbps signals at 1300 nm wavelength. It is to be noted that the measurements in the Gbps range could not be conducted as the signal generator in this range was not available in the KSC- Optical Fiber Laboratory at the time of these experiments. The block diagrams of attenuation, bandwidth and data-rate/eye-pattern measuring systems are shown in Figs. 5 (a), 5 (b) and 5 (c).

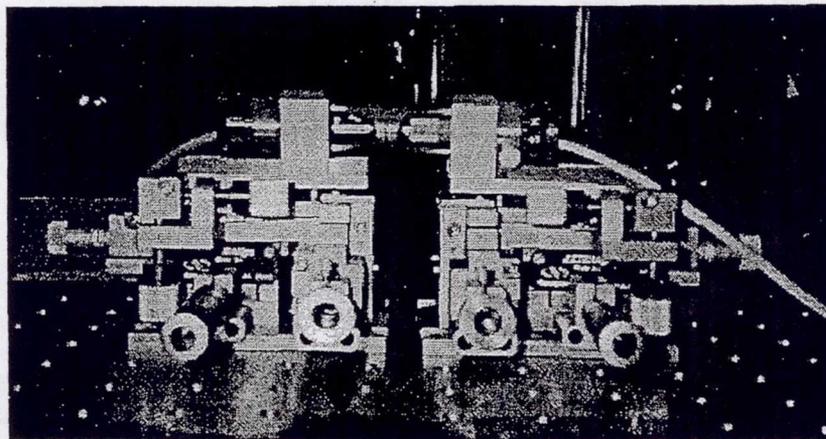
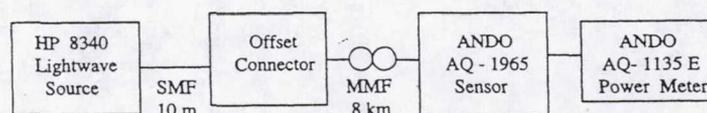
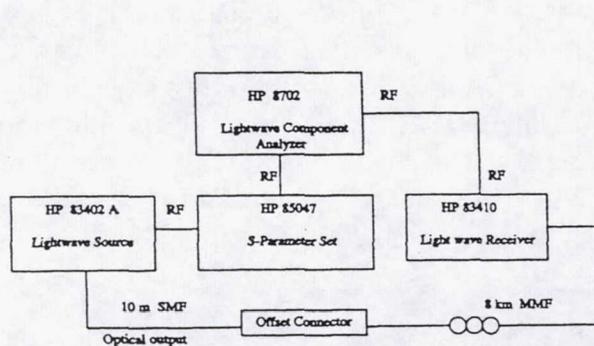


Fig.4. Offset connector assembly on positioning system

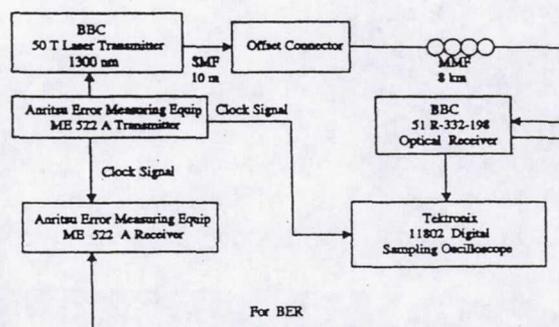
As shown in Fig. 5 (a), the attenuation was measured with the help of HP 8340 lightwave source, ANDO AQ-1965 optical detector or sensor and ANDO AQ-1135 E power meter. The bandwidth measurement system as shown in Fig. 5 (b), consists of HP 8340 A lightwave source, HP 83410 lightwave receiver, HP 8702 lightwave component analyzer and HP 85047 S-parameter set. The lightwave source is modulated with RF signal coming from lightwave receiver through component analyzer and S-parameter set. The offset connector under investigation is connected to 10 m long SMF at the input side and to 8 km long MMF at the output. The optical output at the end of 8 km long MMF is automatically measured for various frequencies. The main components of the data rate and eye-pattern measuring system shown in Fig. 5 (c) are: BBC-50 T, 1300 nm laser transmitter, BBC-51 R-332-198 optical receiver, Anritsu error measuring equipment with ME 522 A optical digital transmitter/receiver and Tektronix 11802 digital sampling oscilloscope.



(a) Attenuation measurement



(b) Bandwidth measurement



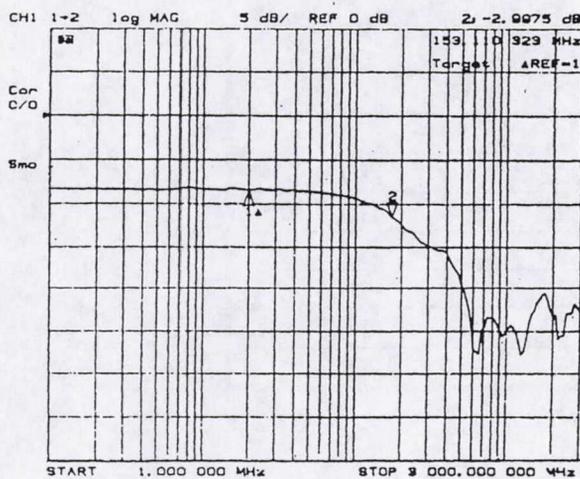
(c) Eye-pattern and data rate measurement

Fig.5. Block schematics of measurement set-ups

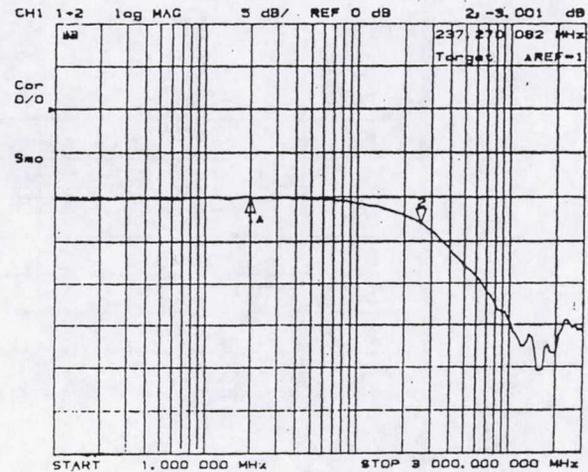
4. RESULTS AND DISCUSSION

The 3-dB bandwidth of the signal and received power for 10, 20 and 30 μm offset positions are shown in Fig. 6 while Table 1 represents these values for all the measured offset positions as well as percentage increase in bandwidth and the fiber capacity. Due to page limitations only a few representative experimental figures are being provided. As expected the bandwidth increases when the offset is increased but at the expense of received power. It is observed that in comparison to the bandwidth at 0 μm offset the percentage increase in the bandwidth for 20 μm

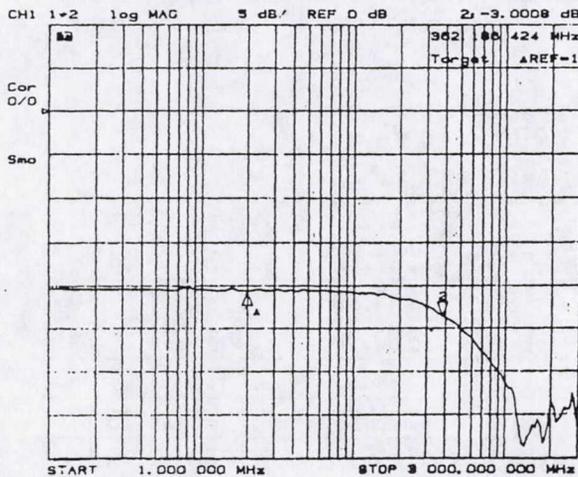
offset is 136.56 which is slightly less than the maximum value of 139.53 for 25 μm offset. But the received power at 20 μm offset is 6.3 dB higher than the 25 μm offset. For higher offset values not only the power decrease but the bandwidth also decreases. Thus the 20 μm offset can be considered as an optimization point where the received power of -20.4 dBm is also within acceptable limit.



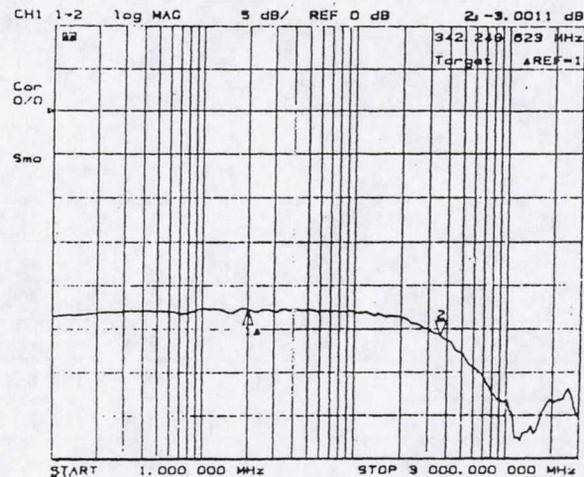
(a) 0 μm offset



(b) 10 μm offset



(c) 20 μm offset



(d) 30 μm offset

Fig. 6. 3-dB Bandwidth and received power for different offset positions

The improvement on bandwidth due to connector offset position can also be visualized from digital perspective of the signal transmission. This has been experimentally obtained in the form of eye-pattern and bit-pattern as shown in Fig.7 for some selected offset positions. Table 2

summarizes the effect of offset on eye-opening and percentage noise jitter for 500 Mbps and 700 Mbps signal transmissions for all offset positions.

Table 1- 3-dB Bandwidth and received power for various offset positions

Offset μm	3-dB Bandwidth MHz	% Increase in BW	Received Power dBm
0	153.11	-----	-8.3
5	187.13	22.22	-10.0
10	237.27	54.96	-10.3
15	272.47	77.96	-11.8
20	362.19	136.56	-20.4
25	366.75	139.53	-26.7
30	342.25	123.53	-28.3

Table 2- Eye pattern size, pulse rise time and noise jitter for various offset positions

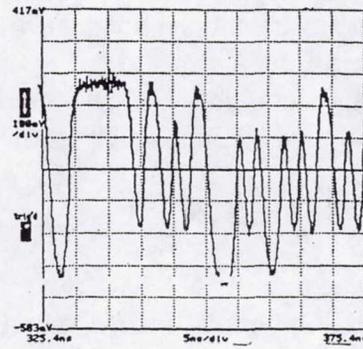
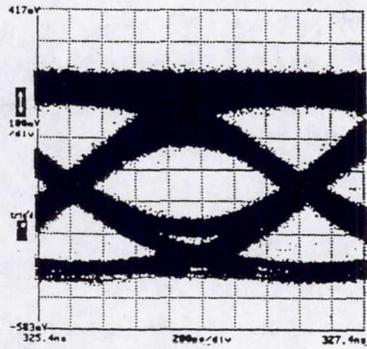
Offset μm	For 500 Mbps				For 700 Mbps			
	Eye pattern size		Rise time ns	Jitter %	Eye pattern size		Rise Time ns	Jitter %
	Hor ns	Ver mV			Hor ns	Ver mV		
0	1.375	143.75	2.4688	31.25	----	----	2.1875	---
5	1.656	325.00	2.8125	18.19	0.850	162.50	1.7525	40.50
10	1.750	325.10	0.9575	12.50	1.150	237.50	1.0938	19.50
15	1.688	350.00	1.4375	16.63	1.126	262.50	1.0655	21.22
20	1.719	337.50	0.5313	14.06	1.175	487.50	0.9375	17.75
25	1.698	305.50	0.5527	15.13	1.061	275.50	0.9885	25.77
30	1.688	287.50	0.5625	15.63	0.913	181.25	1.1250	36.16

From Fig. 7 and Table 2, it is evident that the eye-pattern opening is maximum for 20 μm offset resulting into higher bit rate performance of the fiber. For higher values of the offset the noise level goes up, thus reducing the eye opening. This is also indicated by the lowest value of noise jitter in the Table. For zero offset condition the noise level becomes very high as large number of modes propagate in the fiber resulting into very high dispersion and noise. At the same time it is also observed that the pulse rise time for 20 μm offset is only 0.9375 ns and the noise jitter is

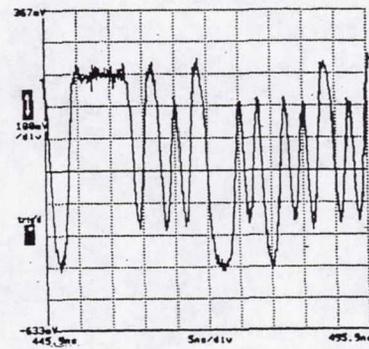
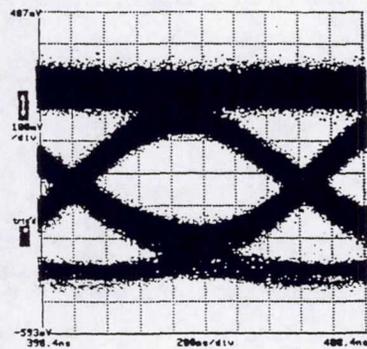
17.75%, which are the lowest of all the values indicating lower dispersion and noise. After comparing the signal transmissions at 500 Mbps and 700 Mbps it is noticed that the offset does not make much improvement on 500 Mbps transmission.

Eye-pattern

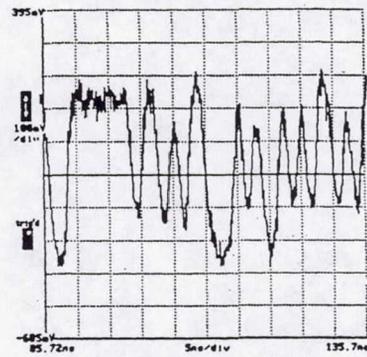
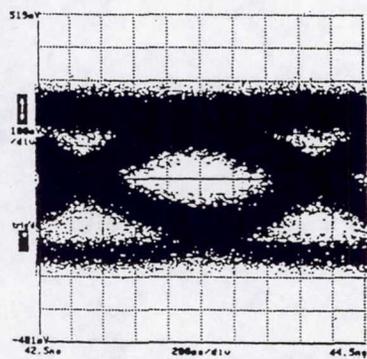
Bit-pattern



(a) 10 μm offset



(b) 20 μm offset



(b) 30 μm offset

Fig. 7. Eye- and bit- patterns for different offset positions

5. CONCLUSION

Basic concept and analysis of an offset connector launch technique has been thoroughly investigated. The connector position offset has been experimentally optimized at an offset of 20 μm where noise and received signal levels seem to be within acceptable limits. It is important to note that the positioning system used in the experimental study was not very precision one and also the experiments have been conducted for 700 Mbps signal transmission rather than 1 Gbps or greater as required by the Gigabit Ethernet standards. It is recommended that before implementing this technique for KSC and other NASA centers further investigations regarding noise, differential mode delay and vibration effects with utmost precision should be conducted.

ACKNOWLEDGEMENTS

I am greatly indebted to my NASA-KSC colleague Po T. Huang, for providing me an opportunity to work in the KSC- Optical Fiber Communications Laboratory, highly intellectual discussions and all possible support during the course of this work. I am very thankful to Bob Swindle for providing experiment-related help and David Johnson for providing computer-related help from time to time. I am also thankful to Prof. Alfred Andrawis of South Dakota State University, another summer faculty in the laboratory for very helpful discussions on the subject matter and moral support. My thanks are also due to Prof. E. Ramon Hosler and Judie Gilliam of UCF, Dr. Jane Hodges and Gregg Buckingham of NASA-KSC for organizing the whole program and several interesting activities during this period.

REFERENCES

- [1] Haas, Z. and Santoro, M, "A mode-filtering scheme for improvement of the bandwidth-distance product in multimode fiber systems", *IEEE J. of Lightwave Technol.*, vol. 11, no. 11, pp. 1125-1131, November 1993.
- [2] Haas, Z. and Santoro, M., "Lightwave transmission systems using selected optical modes", US Patent no. 44318, 1993.
- [3] Raddatz, L., et. al, "An experimental and theoretical study of the offset launch technique for enhancement of the bandwidth of multimode fiber link", *IEEE J. of Lightwave Technol.*, vol. 16, no. 3, pp. 324-331, March 1998.
- [4] Gloge, D. and Marcatili, E.A.G., "Multimode theory of graded -core fibers", *Bell Syst. Tech. J.*, vol. 52, pp. 1563-1568, 1973.
- [5] Hackert, M.J., "Characterizing multimode fiber bandwidth for Gigabit Ethernet applications", *White paper prepared for Corning*, no. WP 4062, November 1998.

1999 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

**JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA**

**OPERATIONS ASSESSMENT OF LAUNCH VEHICLE ARCHITECTURES USING
ACTIVITY BASED COST MODELS**

Alex J. Ruiz-Torres

Assistant Professor

Department of Computer Information Systems and Decision Sciences
Florida Gulf Coast University

Carey McCleskey

NASA PK-K, Advanced Projects

ABSTRACT

The growing emphasis on affordability for space transportation systems requires the assessment of new space vehicles for all life cycle activities, from design and development, through manufacturing and operations. This paper addresses the operational assessment of launch vehicles, focusing on modeling the ground support requirements of a vehicle architecture, and estimating the resulting costs and flight rate. This paper proposes the use of Activity Based Costing (ABC) modeling for this assessment. The model uses expert knowledge to determine the activities, the activity times and the activity costs based on vehicle design characteristics. The approach provides several advantages to current approaches to vehicle architecture assessment including easier validation and allowing vehicle designers to understand the cost and cycle time drivers.

OPERATIONS ASSESSMENT OF LAUNCH VEHICLE ARCHITECTURES USING ACTIVITY BASED COST MODELS

Alex J. Ruiz-Torres

1. INTRODUCTION

The research and design of the next generation launch vehicles (LV's) continues as NASA and industry recognize the potential commercial uses of space and space-based transportation. However, these potential uses cannot be realized until the cost to access space is reduced by several orders of magnitude [1]. In order to achieve this, and based on experiences with the only existing reusable LV (Shuttle) and several expendable LV programs, the design process has evolved and parameters like operability, maintainability, and life cycle costs are critical measures of performance for the evaluation of new LV architectures [2].

The prediction of costs and other operations related parameters for a LV architecture/concept is a complex problem. This is because launch vehicles are inherently very complex systems [3], design architectures are based on new technologies where limited cost/operations knowledge exists, and the "true" reliability, maintainability, and operability of a concept vehicle are difficult to predict. In addition, at the architectural/concept design level a limited set of design characteristics are defined, limiting the input side of the equation. In spite of these limitations, the development of cost assessment - operation focused models is required to truly understand the affordability of new launch systems. Ground operations account for a large portion of the cost of shuttle and ELV's operations. In addition, models that assess early at the concept level are essential as decisions made at this stage of design typically have a significant effect on life cycle costs and other operation parameters.

The need for operation assessment models has prompted NASA's John F. Kennedy Space Center, industry, and academia to form a team (Vision Spaceport) to address these issues [4]. The efforts of the Vision Spaceport team have resulted in a prototype model toolkit that assesses the spaceport requirements driven by a LV architecture. The tools developed by this team provide a "sense of direction" Life Cycle Cost (LCC) based on costs baselines of the Shuttle program and other existing launch/transportation systems. The tools are founded on knowledge functions that map vehicle characteristics to operational functions of a spaceport [5], for example the launch function. The tools developed by this team have been used in two NASA studies; The Space Solar Power Concept Definition Study and The Space Transportation Architecture Study 99.

An alternative approach to the knowledge based functions used in the Vision Spaceport toolkit is the development of a knowledge driven Activity Based Costing (ABC) model. ABC techniques have been used in several cost estimation models for manufacturing; jobs shops environment [6], CIM (Computer Integrated Manufacturing) environments [7], and electronics manufacturing [8], and supply chain modeling. In the space vehicle operations environment, an ABC model was proposed by Christenson and Komar [9] for the modeling and analysis of reusable rocket engines. Their approach focused on detailed modeling of the activities required to turnaround reusable rocket engines, including the development of design specific schedules, resource sets, and stochastic characterizations.

In general, all of these ABC models work by first estimating the activities required to produce/operate a product/device, and then based on these, estimate the labor and other costs associated with these activities. These models addressed "well defined" environments where technology is at a mature state and

the effect of design choices is well understood. The problem addressed by this research is the estimation of activities on an environment where there is limited knowledge of the activities required by a vehicle architecture, given these architectures are typically based on new and experimental technologies. This research proposes the use of expert's knowledge to estimate the activities and associated time and cost parameters. The rest of the paper is organized as follows. The first section reviews the modeling requirements for operations cost assessment of LV architecture designs. The second section presents a generic modeling methodology. The third section describes the knowledge requirements to implement the methodology. The fourth section summarizes the work and discusses directions of future work.

2. OPERATIONS ASSESSMENT MODELING OF LV ARCHITECTURE DESIGNS

The reduction of the cost to access space could open new markets and applications, as for example space tourism. To achieve the lower cost requirements of a future space transportation system, the assessment of vehicle concepts/architectures must consider all life cycle costs; design and development, manufacturing, and production. Design decisions drive to a large extent development, manufacturing, support, and operations functions, thus models based on design decision can be used to predict all of these areas. However, the complexity of this assessment process requires the development of multiple models, capable of estimating the different cost elements, for example a program development assessment model, a manufacturing assessment model, and an operations assessment model. All of these models should then be integrated to provide true life cycle costs for a space transportation system.

Operations models (ground operations or spaceport operations) are an important part of the assessment of new vehicle architectures as they reflect a large portion of the system's recurring costs and will determine the vehicle flight rate capability. The recurring costs and the flight rate are the result of tasks or activities that are required during ground operation, for example the preparation of a payload for integration with the vehicle. Typically the cost and task duration time assessment of these processes is performed by experienced engineers who employ their knowledge of production and operations technology, methods analysis, and engineering economics to predict the probable cost and production time of a product [6], in this case a ground operation activity.

This paper describes a modeling methodology that estimates the flight rate capability and the costs associated with the spaceport operations for a LV concept/architecture. The research described here focuses solely on this area given other models have been proposed to estimate manufacturing costs/production times which could be used to model LV systems [10]. Under the scope of this research, the spaceport is the environment where a LV operates and is provided the support required to satisfy customer requirements. The spaceport is defined as the set of functions that enable a space vehicle to operate and become a space transportation system; from landing (if a RLV) to launch. This includes the processes required to prepare the vehicle for launch, the processes for payload/crew/passenger ingress and egress, the processes of integration to other elements (as in the Shuttle system), the processes required to maintain the vehicle, and the processes required to control during flight. Other functions of the spaceport include those of payload preparation, logistics, and overhaul maintenance.

3. ACTIVITY BASED MODELING OF LV ARCHITECTURES

The principles of activity based models are the assignment of flow times and costs to a product based on the activities required for its production. Each activity has an associated activity time and a set of resource

requirements which determine the cost of the activity. In the operations case, the costs will be based on the activities required for ground operations and its associated activity times and other activity cost drivers.

Given the “product” is the operation of a LV concept/architecture with no existing processes (therefore the actual activity times and costs are not known), the model must predict the activities and the corresponding activity times and costs. As the information about a concept launch vehicle is limited, only top level activities can be defined. Studies from NASA and industry [11] have characterized the inputs from a LV architecture/concept required to assess its operability. While this input list is extensive, it has not been directly related to specific ground activities, and in most cases the architecture variables that drive the time and cost requirement for an activity have not been determined.

3.1 Design Driven Activities

This model characterizes a LV architecture/concept by J design variables, for example, engines of the staged combustion type, engines of the RBCC type, ceramic tile thermal protection system, etc. For each design variable j there are N_j activities required for flight readiness/operation as for example leak check, remove and repair, servicing, etc. The LV architecture/concept is also defined by K vehicle characteristics/ operational drivers α_k . Vehicle characteristics/operations drivers are for example the total area that is covered by a type of thermal protection, the weight of the vehicle, the number of fuel cells in the vehicle, etc. In addition, in some cases the operational driver is the existence of that variable in the design, thus α_k will be a binary variable. For example, a design variable could be the “existence” of life support systems on board and a possible activity is to service the systems, thus $\alpha_{\text{life support}} = 1$ in the case of SST and $\alpha_{\text{life support}} = 0$ (not true) in the case of Venture Star.

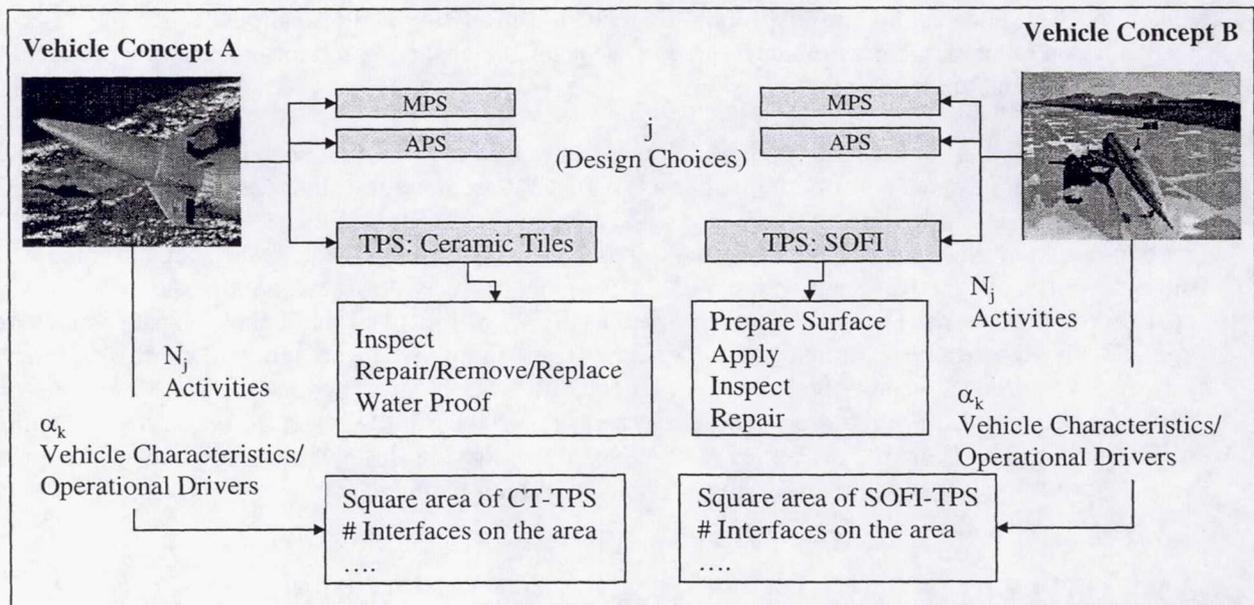


Figure 1. Example design driven activities

Figure 1 illustrates the approach for two concepts. Each design concept is based on several design choices, for example the type of engine used for the main propulsion system (MPS), the type (if any) of its auxiliary power system (APS), and the type of thermal protection system (TPS) covering its exterior surfaces. In the example shown, vehicle concept A uses ceramic tiles as a thermal protection system, while concept B uses SOFI (spray on foam insulator) for thermal protection. The activities for each approach are different and similarly, the vehicle characteristics/operations drivers of interest are different.

3.2 Estimating Operational Cost

To determine the operational cost, it is assumed that each activity n of design variable j , referred to as $j(n)$, has an associated cost $C_{j(n)}$ which is a function of one or more vehicle characteristics/ operations drivers (1a). The cost function for activities could have different forms as required by the activity, including formulation as a linear equation (1b) or as non linear equation (1c).

$$C_{j(n)} = f(\alpha_k) \quad (1a)$$

$$C_{j(n)} = \Omega + \Psi\alpha_x + \Delta\alpha_y \quad (1b)$$

$$C_{j(n)} = \Omega + \Psi\alpha_x/\alpha_y + \Delta(\alpha_y)^2 \quad (1c)$$

Note: Ω , Ψ , and Δ are constants or functions.

For example, the cost of inspecting a thermal protection system (TPS) of ceramic tiles may be formulated by $\$20,000 + \$150 \times \text{area of ceramic tile TPS}$. The $\$20,000$ may be cost of setting up the equipment and the $\$150$ may be associated with the labor and overhead cost per square foot of ceramic tiles TPS that is inspected.

By adding up the cost for all the activities, the total operations cost per flight, C_{ops} , can be formulated (2).

$$C_{ops} = \sum_{m=1..J} \{ \sum_{t=1..Nm} C_{m(t)} \} \quad (2)$$

3.3 Estimating Ground Cycle Time

The flight rate is an important characteristic of flight systems which is tightly related to ground operations. Estimating the flight rate is important as it will help estimate the number of vehicles required to satisfy customer demand forecasts, and also estimate the proper allocation of design, development, and manufacturing costs to ground activities. The flight rate is the inverse of the total flow time of a vehicle over the length of a year, where the total flow-time is the time a vehicle spends in the ground (ground cycle time) and in space (flight time).

To determine the ground cycle time, it is necessary to first estimate the time of each of the activities required by the design. It is assumed that each activity n of design variable j , referred to as $j(n)$, has an associated task time $T_{j(n)}$ which is a function of one or more vehicle characteristics/ operations drivers (3a). Similar to the cost function, the task time for an activity could have different forms as required by the activity, including formulation as a linear equation (3b) or as non linear equation (3c).

$$T_{j(n)} = f(\alpha_k) \quad (3a)$$

$$T_{j(n)} = E + \Gamma\alpha_x + \Pi\alpha_y \quad (3b)$$

$$T_{j(n)} = E + \Gamma\alpha_x/\alpha_y + \Pi\alpha_y^2 \quad (3c)$$

Note: E, Γ , and Π are constants or functions.

For example, the time to inspecting a thermal protection system (TPS) of ceramic tiles may be formulated by 24 hours + 0.5 hours x area of ceramic tile TPS, where 24 hours may be the time required to set up the equipment and 0.5 hours is the time required to inspect per square foot of ceramic tiles TPS given a full capacity of resources (multiple inspecting resources working in parallel).

To determine the overall flow-time and therefore the flight rate, a network of the activities is modeled. This is based on the assumption that spaceport (ground operations) is a network of Q stations, where each station has an unlimited amount of processors. The model also assumes that each activity $j(n)$ is assigned to one of the Q buckets by variable $b_{j(n)}$, of range 1 to Q based on expert's knowledge. At the top level, stations could be defined by the major spaceport functions as landing, turnaround, integration, launch, and traffic control during flight. Each processor at a station can only complete activities for one design variable, thus the lead time at each station is determined by the longest set of activities for a design variable system assigned to it. This assumes that the resources assigned to each subsystem are independent and there are no scheduling conflict constraints. The lead time of the network is the sum of lead times for all the stations.

The cycle time for subsystem j at station q, $CT_{q,j}$ is:

$$CT_{q,j} = \sum_{t=1..N_j} (T_{j(t)} \text{ if } b_{j(t)} = q, 0 \text{ otherwise}) \quad (4)$$

The cycle time for a station q, CT_q is:

$$CT_q = \text{Max}_{m=1..J} CT_{q,m} \quad (5)$$

The ground operations cycle time is:

$$GCT = \sum_{d=1..Q} CT_d \quad (6)$$

Figure 2 illustrates an example of how the process works for a specific station. Let's assume there are only three design options for the turnaround module (TPS = Thermal Protection System, MPS = Main Propulsion System, CT = ceramic tiles, B = Blankets, RBCC = rocket based combined cycle). Each of these choices have an associated set of activities, activity times, and activity costs (I = inspect, r/r/r = repair/replace/remove, WP = water proof, S = safe, S/C = service/closeout). The sum of the three activities for the TPS-ceramic tiles design choice is the one that determines the turnaround cycle time.

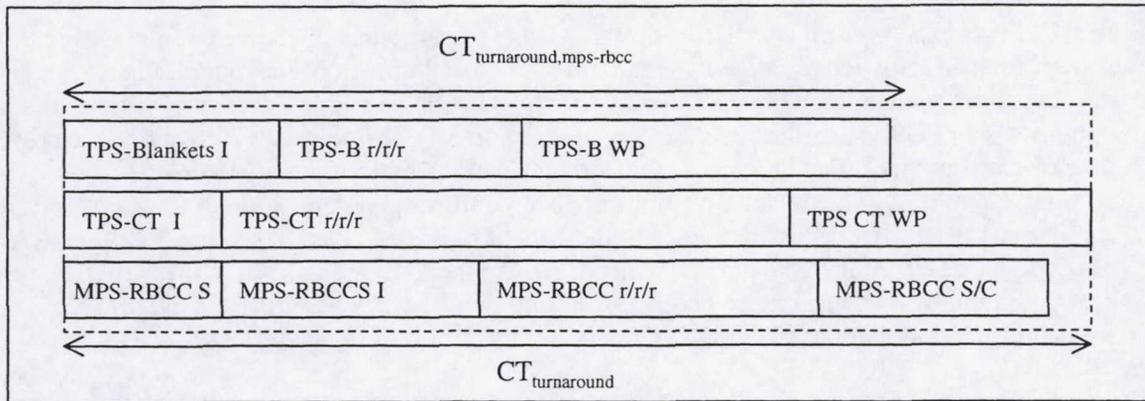


Figure 2. Example: Cycle Time for the Turnaround station

Figure 3 shows an example for the total vehicle flow-time given a set of six stations: TC = traffic control during flight, Ld = Landing operations, Turnaround operations, payload de-mate operations, payload integration operations, and launch operations. In the illustrated network, two stations are parallel: turnaround and P/L de-mate and all other processes are sequential. Note that within each station there are one or more design driven activity sets as in Figure 2.

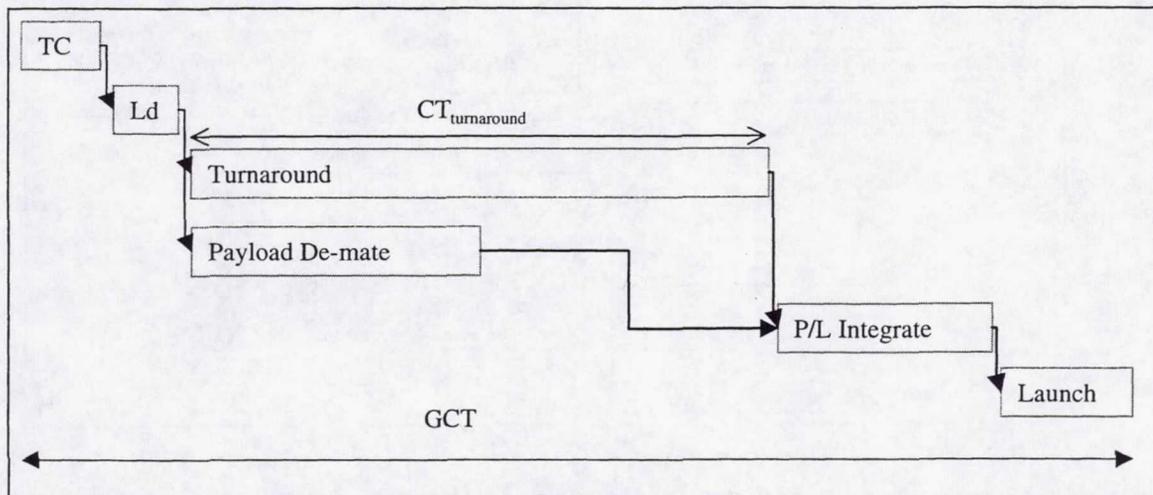


Figure 3. Total flow-time example

3.4 Knowledge Requirements

The implementation of the described model requires an extensive knowledge base. The generation of this knowledge base will require the development and validation of a knowledge capture process

which allows experts from launch and design centers to participate on its development. Figure 4 shows the main knowledge elements required for the development of the model. First, a set of vehicle design options, focusing on the operational drivers, must be developed. For each of these design options, a set of activities must be defined as in section 3.1. The next step is to define the cost and time of each activity based on one or more vehicle characteristics/operations drivers as described in sections 3.2 and 3.3. The model development also requires the organization of the spaceport by stations and the assignment of activities to these stations as described in section 3.3. Finally, an area of additional research is in the development of environment scenarios, where the activities, times, and costs, required by a design choice change with improvement in reliability, vehicle life, and technology, and reductions in complexity, similar in operations to an airplane.

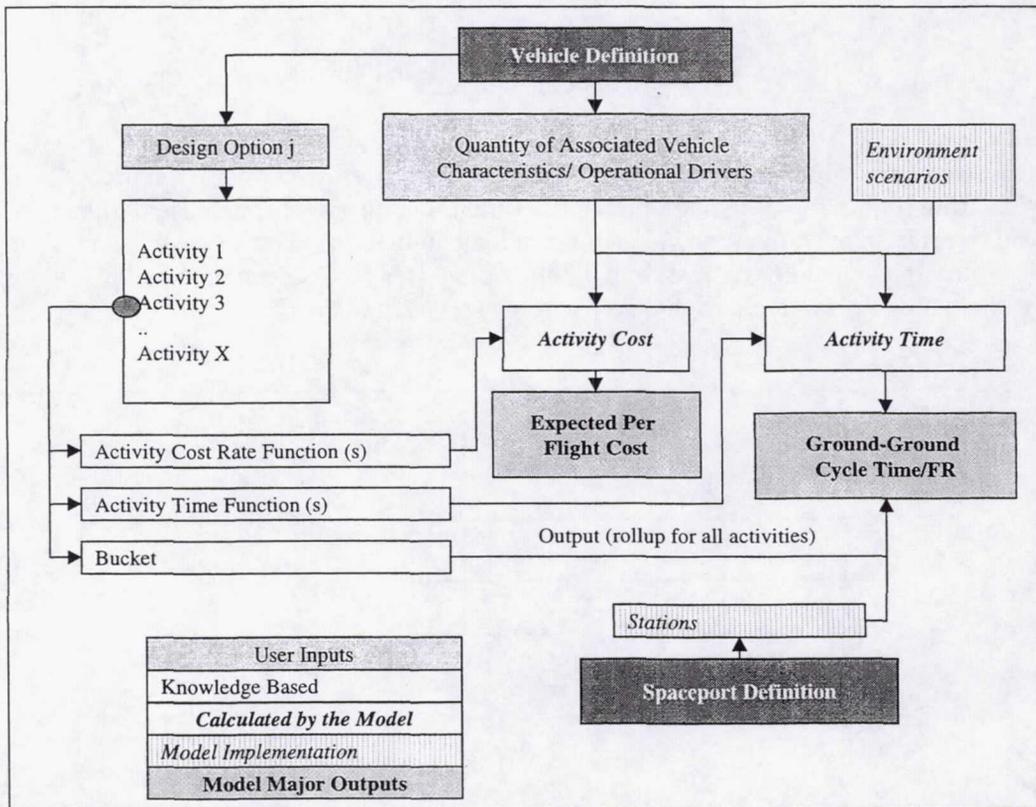


Figure 4. Knowledge requirements

4. RESULTS AND FUTURE WORK

The use of an Activity based costing model to assess the operational requirements of new products is not a new concept and could be applied to the assessment of new launch vehicle architectures. By using the knowledge of experts in the areas of spaceport operations and vehicle/technology designers, design driven activities can be determined, and from there, the time and cost of the activity. The approach allows vehicle designers to better understand (by looking at the output) the cost and cycle time drivers as they

can easily observe which design driven activities have the highest costs and task times. In addition, this approach fosters the development of additional operations knowledge as it “forces” operations experts to predict the activities (and their cost and time characteristics) that new technologies will require in the context of the spaceport.

REFERENCES

- [1] Scott, W. B. Airline Ops Offer Paradigm For Reducing Spacelift Costs. *Aviation Week & Space Technology*. June 15, 1998, pp 64-67.
- [2] Mankins J.C. Lower Cost for Highly Reusable Space Vehicles. *Aerospace America*. March, 1998, pp 36-42.
- [3] Ryan, R. S., and Townsend , J. S. Fundamentals and Issues in Launch Vehicle Design. *Journal of Spacecraft and Rockets*. Vol. 34, No. 2, 1998, pp. 192-198.
- [4] Vision Spaceport. Spacelift Benchmarking & Architectural Modeling. 1998 Project Plan. Unpublished Report, May 1998.
- [5] Zapata, E., and Ruiz-Torres, A.J. Space Transportation Cost Modeling and the Architecture Assessment Tool- Enhanced. *Proceedings from the International Astronautical Federation Conference*, Amsterdam, October 1999.
- [6] Aderoba, A. A generalized cost-estimation model for job shops. *International Journal of Production Economics*. Vol. 53, 1997, pp. 257-263.
- [7] Dhavale, D. G. A Manufacturing Cost Model for Computer-integrated Manufacturing Systems. *International Journal of Operations & Production Management*. Vol. 10, No. 8, 1990, pp. 5-18.
- [8] Ong, N. S., Lim, L. E. N., and Yeo, K. T. (1993). Manufacturing Cost Modeling for Electrical Interconnections. *International Journal of Operations & Production Management*. Vol. 13, No. 2, 1993, pp. 63-75.
- [9] Christenson, R. L., and Komar, D. R. Reusable Rocket Engine Operability Modeling and Analysis. NASA/TP – 1998-208530. Marshall Space Flight Center.
- [10] Marx, W., Dimitri N. M., and Schrage, D. P. Cost/Time Analysis for Theoretical Aircraft Production. *Journal of Aircraft*. Vol. 35, No. 4, July-August 1998, pp 637-646.
- [11] A Catalog of Spaceport Architectural Element with Functional Definition. Highly Reusable Space Transportation Synergy Team. Unpublished Report, 1997.

1999 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA

**Performance Evaluation of
Reliable Multicast Protocol for
Checkout and Launch Control Systems**

Wei Wennie Shu
Department of Electrical and Computer Engineering
University of Central Florida

In Conjunction with John Porter
Checkout and Launch Control Systems
John F. Kennedy Space Center

ABSTRACT

The overall objective of this project is to study reliability and performance of Real Time Critical Network (RTCN) for checkout and launch control systems (CLCS). The major tasks include (a) reliability and performance evaluation of Reliable Multicast (RM) package and (b) fault tolerance analysis and design of dual redundant network architecture.

Performance Evaluation of Reliable Multicast Protocol for Checkout and Launch Control Systems

Wei Wennie Shu

1 INTRODUCTION

1.1 Project Definition

The overall objective of this project is to study reliability and performance of real time critical network (RTCN) for checkout and launch control systems (CLCS), with two major components of work to be focused:

- Reliability and performance evaluation of reliable multicast (RM) package;
- Fault tolerance analysis and design of dual redundant network architecture.

1.2 Background Overview

CLCS includes four major subsystems:

- RTPS, *Real-Time Processing Systems*
- SDC, Shuttle Data Center
- SIM, Simulation System
- BIN, Business Information Network

Our project is focused on the real-time processing subsystem, which in turn consists of four major processing components:

- Gateways
- DDPs/CCPs, Data Distribution Process/Command Control Process
- SDC, Shuttle Data Center
- CCWs, Command Control Workstations

To interconnect these processing components together, it involves with construction of three major network components:

- RTCN, Real time critical network
- DCN, Display and control network
- UN, Utility network

1.3 Application Characteristics

Applications associated with RTCN are mainly information exchanges, which include the 10ms synchronous rate to send messages from gateways to DDPs, CCPs, and SDC, with the pattern of many-to-many multicasting [1], and the 100ms synchronous rate to send messages from DDPs/CCPs to CCWs, gateways, and SDC, also with the pattern of many-to-many multicasting.

There are two message protocols supported, ACK-based and NACK-based. In an NACK-based message stream, a sender does not wait for acknowledgment of the receiver and a receiver sends NACK back if any message is out of order. The sender will perform retransmission upon receiving NACK. In an ACK-based message stream, a receiver sends ACK back for every message received, and the sender waits for ACK, or time-out for retransmission.

1.4 Software Architectures

It is basically a multithreading client-server model. The server is a Reliable Multicast (RM) package running on top of UDP, and clients are application programs migrated from the old LPS and communicate exclusively via RM server.

On each machine, there exists a single RM server with multiple threads and multiple clients running as concurrent processes/threads. It utilizes many operating system features, such as Pthread package, POSIX.4 real-time extension to accomplish priority scheduling, shared message queue to establish request and response flow, and shared memory to eliminate excessive message copying.

1.5 Network Infrastructure

There are many currently available technologies [2,3,4,5,6]. Among them, the Switched Fast Ethernet has been selected due to its reasonable cost-performance and adequate functionality. Major products include Catalyst 2900 Switches by Cisco, BayStack 450 Switch by Bay Network, and SuperStack II 3300 by 3Com. A brief *evaluation report* is available online.

2 PERFORMANCE AND RELIABILITY EVALUATION

2.1 Testing Goal and Levels

We decide to use the synthetic load to test system at different levels. Consequently, we will compare the measured capacity limits to the real-world worst case analysis to determine the safety margin. The three different testing levels are described here and the task of this project is concentrated at the third level of testing:

- **Level I:** Underlying network architecture testing will use Smartbits to determine port-to-port capacity
- **Level II:** Network infrastructure testing uses the standardized transport interfaces, UDP and TCP to determine the available bandwidth on the top of operating system
- **Level III:** Network application testing will use the RM package with synthetic communication and CPU load to determine the available bandwidth from the application interface.

2.2 Modeling and Performance Evaluation of ACK-based Message Protocol

2.2.1 Testing environment

The testing is set up to have SUN Ultra60 as the sender and SUN Enterprise3500 as the receiver. In each machine, the RM server is always running as the top priority process and the application clients are running as the processes with the second highest priority. Each sender is periodically sending the messages with specified sizes to the receiver. Here, the synchronous rate can be varied from 1ms to 10ms and the message size can range from 1Kbytes to 64Kbytes, which is the upper limit the RM can handle.

2.2.2 Testing send/receive of single message stream

We define three important types of metrics in testing of behavior of the ACK based message protocol.

- **Response time:** the time from sending a message by the application client to reach the receiving side's RM server until receiving the ACK message back at the sending side. It includes a round trip time of message transmission to assure the arrival of message at the receiving side, but not guarantee the receipt of message by the application client at the receiving side.
- **End-to-end delay time:** the time from sending a message by the application client to reach the receiver until receiving the message by the application client at the receiving side
- **Throughput:** the amount of messages to be sent without loss of messages. Here, throughput can be calculated and measured based on the back-to-back message transmission or periodical message transmission with the fixed synchronous rate.

In the ACK-based message protocol, the response time is pretty close to the end-to-end delay time. If the receiving side is heavy loaded, it can have great impact on the end-to-end delay time. On the other hand, the response time depends on the network traffic. Throughput can be calculated and measured based on the back-to-back message transmission or periodical message transmission with the fixed synchronous rate.

2.2.3 Modeling and analysis

Here, we model the response time t_r as a function of message sizes:

$$t_r = \alpha + \beta * \text{size} / \gamma$$

where, $\alpha = 680 \mu\text{s}$, startup time

$\beta = 103 \mu\text{s}/\text{Kbytes}$, transmission time

$\gamma = (1 + 0.001 * \text{size})$, adjustment factor for large messages

size = message size in Kbytes

And the end-to-end delay time t_d is basically proportional to the response time t_r .

$$t_d = \lambda * t_r, \quad \text{where } \lambda \text{ is about } 1.05 \text{ to } 1.15$$

By ignoring adjustment factor γ , the throughput T can be obtained by

$$T \approx 1 / (\alpha + \beta * \text{size})$$

Theoretically speaking, the throughput can be approximately calculated by assuming the back-to-back message transmission.

- For a small message of 1K
 $T \approx 1 / (\alpha + \beta) = 8 \text{ Kbytes} / 783 \mu\text{s} \approx 10 \text{ Mbps}$
- For a message of 10K
 $T \approx 1 / (\alpha + 10\beta) = 80 \text{ Kbytes} / 1710 \mu\text{s} \approx 47 \text{ Mbps}$
- For a large message of 50K
 $T \approx 1 / (\alpha + 50\beta) = 400 \text{ Kbytes} / 5830 \mu\text{s} \approx 67 \text{ Mbps}$

Figure 2.1 gives comparison of measured & calculated data.

Msg size in Kbytes	1	10	30	50	64
Bandwidth in Mbps	0.8	8	24	40	51.2
	1K	10K	30K	50K	64K
tr in us, measured	779	1670	3669	5594	6950
tr in us, calculated	783	1700	3680	5585	6875
td in us, measured	783	1779	3878	6045	7533
td in us, calculated	861	1870	4048	6143	7563

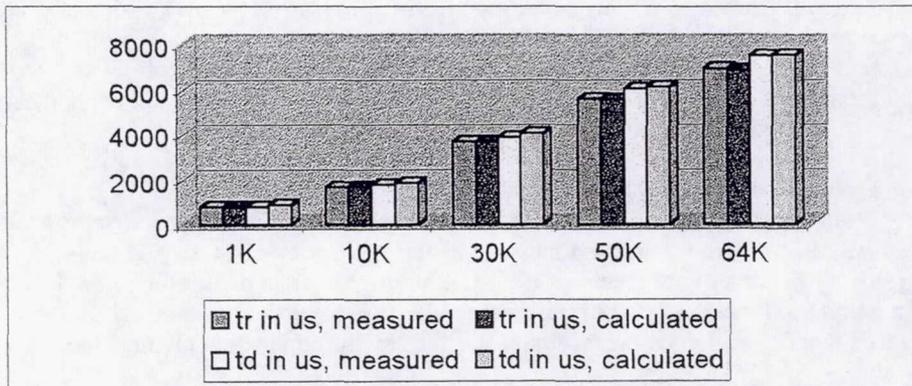


Figure 2-1 Comparison of measured and calculated data for single ACK message stream

2.2.4 Testing send/receive of multiple message streams

If there are multiple senders in the testing configuration, the background traffic has impact on the response time of the message stream to be tested, as well as the end-to-end delay time. Both the background traffic and the message stream to be tested will compete for network switcher's bandwidth and CPU resources at the receiving side. The end-to-end delay time and throughput defined in the above can also be applied here.

$$t_{r,multi} = t_{r,single} * (1 + \delta / 100)$$

where, δ = background traffic modifier in Mbps

$t_{r,single}$ = response time for the single message stream

Figure 2.2 shows how the response time is affected by the various background traffics.

Background Traffic in Mbps	Msg size in Kbytes	1	10	30	50	64
	Bandwidth in Mbps	0.8	8	24	40	51.2
		1K	10K	30K	50K	64K
0	tr in us, measured	779	1670	3669	5594	6950
	tr in us, calculated	783	1700	3680	5585	6875
10	tr in us, measured	819	1744	4123	6371	7824
	tr in us, calculated	857	1837	4036	6153	7645
20	tr in us, measured	878	1923	4416	7134	8630
	tr in us, calculated	935	2004	4403	6713	8340
40	tr in us, measured	1051	2129	5425	7464	
	tr in us, calculated	1091	2338	5137	7832	9730

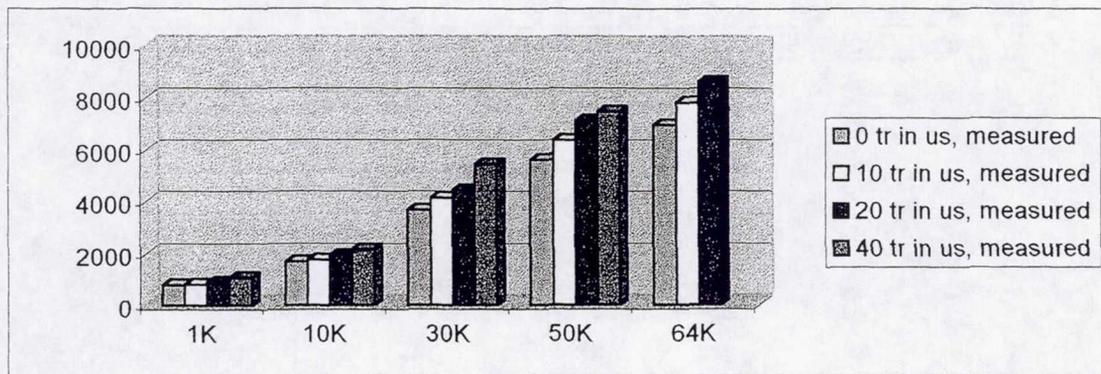


Figure 2-2 Response time of multiple ACK message streams

Notice that it makes difference when testing ACK-based messages with one primary receiver or with one primary receiver and one non-primary receiver. How does the number of receivers have impact on ACK's bandwidth? Next, what is capacity of receiving many small messages at Enterprise3500 side? Currently, due to the resource limitation, we have tested 15 streams with message size of 1K, 2K, 4K, but not 5K. More importantly, the best priority settings to RM server as well as client processes need to be determined.

2.3 Modeling and Performance Evaluation of NACK-based Message Protocol

2.3.1 Testing send/receive of single message stream

In addition, sending time is a newly defined metric, particularly defined for the NACK-based message protocol.

- **Sending time:** the time from sending a message by the application

In the NACK-based message protocol, the sending time is very different from the end-to-end delay time

2.3.2 Modeling and analysis

Here, we models the end-to-end delay time t_d as a function of message sizes:

$$t_d = \alpha + \beta * \text{size} / \gamma$$

where, α = 500 μ s, startup time

β = 110 μ s/Kbytes, transmission time

$\gamma = (1 + 0.001 * \text{size})$, adjustment factor for large messages
 size = message size in Kbytes

And the sending time t_s is less dependent on the size of messages.

$$t_s = \alpha' + \beta' * \text{size} / \gamma$$

where, $\alpha' = 280 \mu\text{s}$, startup time

$\beta' = 20 \mu\text{s/Kbytes}$, transmission time

$\gamma = (1 + 0.001 * \text{size})$, adjustment factor for large messages
 size = message size in Kbytes

By ignoring adjustment factor γ , the throughput T_{NACK} can be obtained by

$$T_{\text{NACK}} \approx 1 / (\alpha + \beta * \text{size})$$

Figure 2.3 gives comparison of measured & calculated data

Msg size in Kbytes	1	10	30	50	64
Bandwidth in Mbps	0.8	8	24	40	51.2
	1K	10K	30K	50K	64K
ts in us, measured	337	439	894	1262	1550
ts in us, calculated	300	480	880	1280	1560
td in us, measured	613	1728	3735	5873	7294
td in us, calculated	610	1589	3704	5738	7117

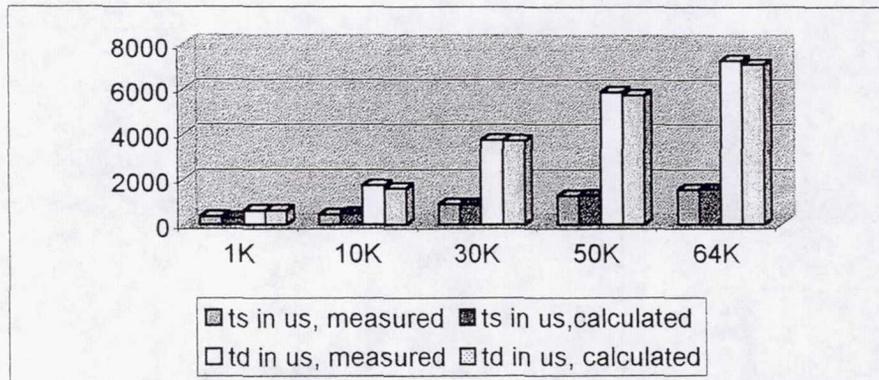


Figure 2-3 Comparison of measured and calculated data for singel NACK message stream

2.3.3 Testing send/receive of multiple message streams

Similarly, if there are multiple senders in the testing configuration, the background traffic has impact on the end-to-end-delay time of the message stream to be tested.

$$T_{d,\text{multi}} = t_{d,\text{single}} * (1 + \delta / 90)$$

where, $\delta =$ background traffic modifier in Mbps

$t_{d,\text{single}} =$ end-to-end delay time for the single message stream

Figure 2.4 shows how the end-to-end delay time is affected by the various background traffics.

Background	Msg size in Kbytes	1	10	30	50	64
Traffic in Mbps	Bandwidth in Mbps	0.8	8	24	40	51.2
		1K	10K	30K	50K	64K
0	td in us, measured	613	1728	3735	5873	9429
	td in us, calculated	610	1589	3704	5738	7117
10	td in us, measured	751	2034	4142	7073	
	td in us, calculated	681	1920	4150	6526	10477
20	td in us, measured	779	2153	4539	10784	
	td in us, calculated	749	2112	4565	7178	11524
40	td in us, measured	799	2363	5340		
	td in us, calculated	885	2496	5395	8483	13620

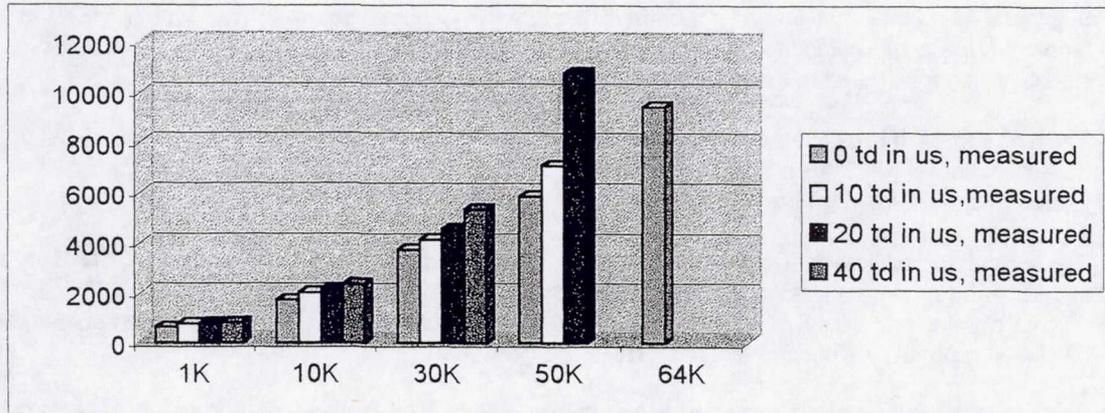


Figure 2-4 End-to-end delay time of multiple NACK message streams

2.4 Measurement of Throughput

2.4.1 Throughput in ACK-based single/multiple message stream

For the single stream case, the upper limit of throughput without missed messages is measured when the synchronous rate is varied. Since the maximum allowable size of messages in RM package is of 64Kbytes, the longer synchronous rate cannot saturate the network bandwidth.

- 10ms, 64KBytes, $64 \times 8 / 10 = 51.2$ Mbps
- 5ms, 40KBytes, $40 \times 8 / 5 = 64$ Mbps
- 2ms, 16KBytes, $16 \times 8 / 2 = 64$ Mbps

Therefore, the measured throughput is about 64Mbps. Similarly, for the multiple stream case, upper limit of throughput is measured:

- with 10Mbps, 64KBytes, $64 \times 8 / 10 + 10 = 61.2$ Mbps
- with 20Mbps, 50KBytes, $50 \times 8 / 10 + 20 = 60$ Mbps
- with 40Mbps, 25KBytes, $25 \times 8 / 10 + 40 = 60$ Mbps

The measured throughput is about 60Mbps

2.4.2 Throughput in NACK-based single/multiple message stream

For the NACK-based message streams, the upper limit of throughput without missed messages is measured when the synchronous rate is varied.

- 10ms, 64KBytes, $64 \times 8 / 10 = 51.2$ Mbps
- 5ms, 30KBytes, $30 \times 8 / 5 = 48$ Mbps
- 2ms, 12KBytes, $12 \times 8 / 2 = 48$ Mbps

Therefore, the measured throughput is about 50Mbps. Similarly, for the multiple stream case, upper limit of throughput is measured:

- with 10Mbps, 50KBytes, $50 * 8 / 10 + 10 = 50$ Mbps
- with 20Mbps, 50KBytes, $50 * 8 / 10 + 20 = 60$ Mbps
- with 40Mbps, 30KBytes, $30 * 8 / 10 + 40 = 64$ Mbps

The measured throughput is about 60Mbps, same as ACK one. However, for the single stream case, the throughput of NACK is even lower than one of ACK. It needs more investigation for performance verification or implementation explanation.

3 FAULT TOLERANCE ANALYSIS AND DESIGN

3.1 Different Design of Dual Redundant Network

In general, the dual network can be used in different ways to improve fault tolerance of single point failures. It has been decided to construct a complete dual redundant network: RTCN-A and RTCN-B. However, many varieties of design choices exist.

The first approach is called the Active/Standby redundant network. RTCN-A is assigned as the active network whereas RTCN-B acts as the standby network. In normal cases, only one network is fully operational, thus, no extra load is added to network or CPUs.

The second approach is called fully duplicated redundant network. Both RTCN-A and RTCN-B are fully operational. For every message, CPU will send it to both networks and the receiving side CPU will receive whichever comes first and ignore the second arrival one. In this way, it will almost double the CPU load on both sending and receiving sides.

The third approach is newly proposed in this project, called Ping-Pong alternation redundant network. More details are described follows.

3.2 Ping-Pong Alternation Dual Network

The basic idea of Ping-Pong alternation approach is to use dual networks, RTCN-A and RTCN-B, alternatively. The design is motivated since there is no increase of CPU load on sending and receiving sides, especially important for Ultra60 with single CPU. Another design motivation is to low the network traffic to its half, lighter traffic of network always enhances its performance.

As one of major drawbacks, it makes RM protocol a little more complex. Additional information is needed on each RM server. We define a boolean variable *flag*, being 0 for RTCN-A and 1 for RTCN-B; In each message, we also need a boolean variable *flag*, being 0 for RTCN-A and 1 for RTCN-B; We outline the protocol modification to the Ping-Pong Alternation approach as follows.

- ACK-base message stream
- message send:
if (flag==0)
 then send to RTCN-A
 else send to RTCN-B
flag = !flag
- message receive:
select to receive messages from either RTCN-A or RTCN-B
if (message.flag==0)
 then send ACK back via RTCN-A
 else send ACK back via RTCN-B
- message send time-out:

```

if (message.flag==0)
  then resend to the other network RTCN-B
  increment failCountA to keep record
  if failCountA > threshold perform fail over

```

- NACK-base message stream
 - message send:


```

if (flag==0)
  then send to RTCN-A
  else send to RTCN-B
flag = !flag

```
 - message receive:


```

select to receive messages from either RTCN-A or RTCN-B

```
 - message receive out of order


```

if (majority of missed message.flag==0)
  then send NACK back via RTCN-B
  else send NACK back via RTCN-A

```
 - message send getting NACK:


```

resend the missed message to the network where NACK comes
increment failCount to keep record
if failCount > threshold perform fail over

```

Dual redundant network design is only in the initial stage. A prototype is needed to verify correctness and test different approaches for performance comparison and fail over time measurement.

4 DISCUSSION AND FUTURE WORK

All test data conducted in this project are only preliminary; many places need detailed analysis and verification [7]. Inadequate application traffic analysis makes thorough performance evaluation difficult. Detailed information about application characteristics, both for normal and worst cases, can be very helpful.

4.1 About Network Switches

The current broadcast-based multicast limits the total bandwidth to 100Mbps. The future VLAN implementation with the true multicast can help the system to scale up [8]. Prioritization is another future feature can be used to enhance the real-time critical performance

4.2 About Real-time Message (RM) Package

More formal design specification and verification are suggested [9, 10, 11]. Without loss of concurrency, the number of threads in the RM server should be minimized. Message buffers are allocated and de-allocated by different sides of client and server, being efficient but less safety. Use of many nested mutex locks needs careful investigation.

4.3 About Operating System Level Support

The operating system support is very weak in this project and needs substantial efforts [12, 13]. Solaris' real-time extension should be studied and evaluated in its behavior and performance. Assigning different priorities to the RM server and application clients shall be studied on single and multiple CPU machines, particularly in response time and frequency of context switches. Memory locking behavior and its impact on performance needs to be tested. Use of shared memory and message queues need more studies, particularly when created and accessed by different users and applications processes. Behavior of user threads bounding to kernel processes or lightweight processes should be clarified, especially for multiple CPUs. Any possible memory leakage needs to be checked and monitored, particularly for extended run-time.

4.4 Suggestions from Software Engineering Point of View

In general, risk analysis needs more attention in all kinds of levels, including misuse of RM package.

Formal specification on design, implementation, and testing can be further improved. Interaction and integration with the operating system supports should be much encouraged or even enforced to ensure the system integrity.

Hardware failure was/is a dominant consideration in developing reliable systems. Software complexity needs more attentions. RM package needs more tightly integration among various subsystems, such as Health Check, Data Center, Gateways, etc. In addition, is there any potential alternatives to RM package? The Xpress Transport Protocol (Xtp) provides reliable datagrams and reliable multicast connections. Many other reliable UDP can offer possibilities.

5 ACKNOWLEDGMENT

The author would like to thank the entire CLCS division, particularly to Vincent Mandese, John Porter, Steve Davis, Terry Ross, Jeff Crowder, and Steve Goodmark for their wonderful help. Many thanks also go to the KSC and UCF, who made this program possible, particularly Ramon Hosler, Gregg Buckingham, Jane Hodges, Judie Gilliam, and all other faculty teammates

REFERENCES

1. Kenneth C. Miller, *Multicast Networking and Applications*, Addison Wesley, 1999, ISBN 0-201-30979-3.
2. Darryl P. Black, *Building Switched Networks*, Addison-Wesley, 1999, ISBN 0-201-37953-8.
3. John J. Roses, *Switched LANs*, McGraw-Hill, 1998, ISBN 0-07-053413-6.
4. Martin Nemzow, *Fast Ethernet Implementation and Migration Solutions*, McGraw-Hill, 1997, ISBN 0-07-046385-9.
5. Jayant Kadambi, Ian Crayford, and Mohan Kalkunte, *Gigabit Ethernet: Migrating to High-Bandwidth LANs*, Prentice Hall, 1998, ISBN 0-13-913286-4.
6. Howard W. Johnson, *FastEthernet: Dawn of a New Network*, Prentice Hall, 1996, ISBN 0-13-352643-7.
7. Gilbert Held, *LAN Performance: Issues and Answers*, John Wiley & Sons, 1996, ISBN 0-471-96926-5.
8. Gilbert Held, *Virtual LANs: Construction, Implementation, and Management*, John Wiley & Sons, 1997, ISBN 0-471-17732-6.
9. K. Birman, *Building Secure and Reliable Network Applications*, Manning Publications, 1996, ISBN 1-884777-29-5.
10. Roger S. Pressman, *Software Engineering: A Practitioner's Approach*, 4th edition, McGraw-Hill, 1997, ISBN 0-07-114603-2.
11. Stephen R. Schach, *Classical and Object-oriented Software Engineering with UML and C++*, 4th edition, McGraw-Hill, 1999, ISBN 0-07-290168-3.
12. H. Majidimehr, *Optimizing UNIX for Performance*, Prentice Hall, 1996, ISBN 0-13-111551-0.
13. Cockcroft, *Sun Performance and Tuning: SPARC & Solaris*, Sun Microsystems, 1995, ISBN 013-149642-3.

1999 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA

A VISUAL EDITOR IN JAVA FOR JVIEW

Ryan Stansifer
Associate Professor
Computer Science
Florida Institute of Technology

KSC Colleague: Steve Beltz (Dynacs)

ABSTRACT

In this project we continued the development of a visual editor in the Java programming language to create screens on which to display real-time data. The data comes from the numerous systems monitoring the operation of the space shuttle while on the ground and in space, and from the many tests of subsystems. The data can be displayed on any computer platform running a Java-enabled World Wide Web (WWW) browser and connected to the Internet. Previously a special-purpose program had been written to display data on emulations of character-based display screens used for many years at NASA. The goal now is to display bit-mapped screens created by a visual editor. We report here on the visual editor that creates the display screens.

This project continues the work we had done previously. Previously we had followed the design of the "beanbox," a prototype visual editor created by Sun Microsystems. We abandoned this approach and implemented a prototype using a more direct approach. In addition, our prototype is based on newly released Java 2 graphical user interface (GUI) libraries. The result has been a visually more appealing appearance and a more robust application.

1 Introduction

During the operation of the space shuttle, sensors are monitoring many of the subsystems. This data goes to special hardware at the Launch Control Center (LCC) known as the Common Data Buffer (CDBF). Lots of data is collected, approximately 30,000 measurements. These measurements are continually changing—some of them can change rapidly at certain times. The data is used in monitoring the operation of the shuttle and in analyzing subsystems for safety, performance, technological improvements, etc. Each individual measurement is given a short tag called a function designator (FD) to identify it.

This Real-time data is monitored on system engineering consoles of then the Control Checkout and Monitor System (CCMS) in the Launch Control Center (LCC). The PC GOAL system presents the same data to a wider audience in a format closely resembling the consoles of the CCMS. The CCP (CDBF Communications Processor) scans the memory of the CDMF every 2 seconds and broadcasts the data on the LPS Operations Net (LON). This data (and other data, e.g., FIFO) is relayed to the PC GOAL stations. The PC GOAL stations are PCs with network hardware running DOS and the PC GOAL software.

The PC GOAL system presents shuttle data on schematic-like screens described by character-oriented files known as DSP files. Figure 1 shows one of the hundreds of PC GOAL display screens (no data is been displayed, the character 'X' is filling the character positions that would be occupied by numerals). Each shuttle mission requires substantial effort to organize the CDBF, distribute the DSP files, etc.

2 Jview

The Jview project is motivated by the PC GOAL system to display real-time shuttle data as conveniently and efficiently as possible. The programming language Java was chosen because of the easy of writing both graphical user interface (GUI) code and distributed programs. The Internet is the obvious mechanism to transport the data. The wide-spread availability of browsers for the World Wide Web suggests an obvious user interface for any information system large or small.

Java code can be transmitted as part of the WWW protocol, just like pictures, sound and other data. The extreme interest in Java is caused by the ability of browsers to execute the Java code. These Java programs transported across the Internet and executed locally by a WWW browser are called applets. Applets add interaction to otherwise static WWW documents. The Jview project uses applets to form a connection to a Java data manager that relays the real-time data to the applet.

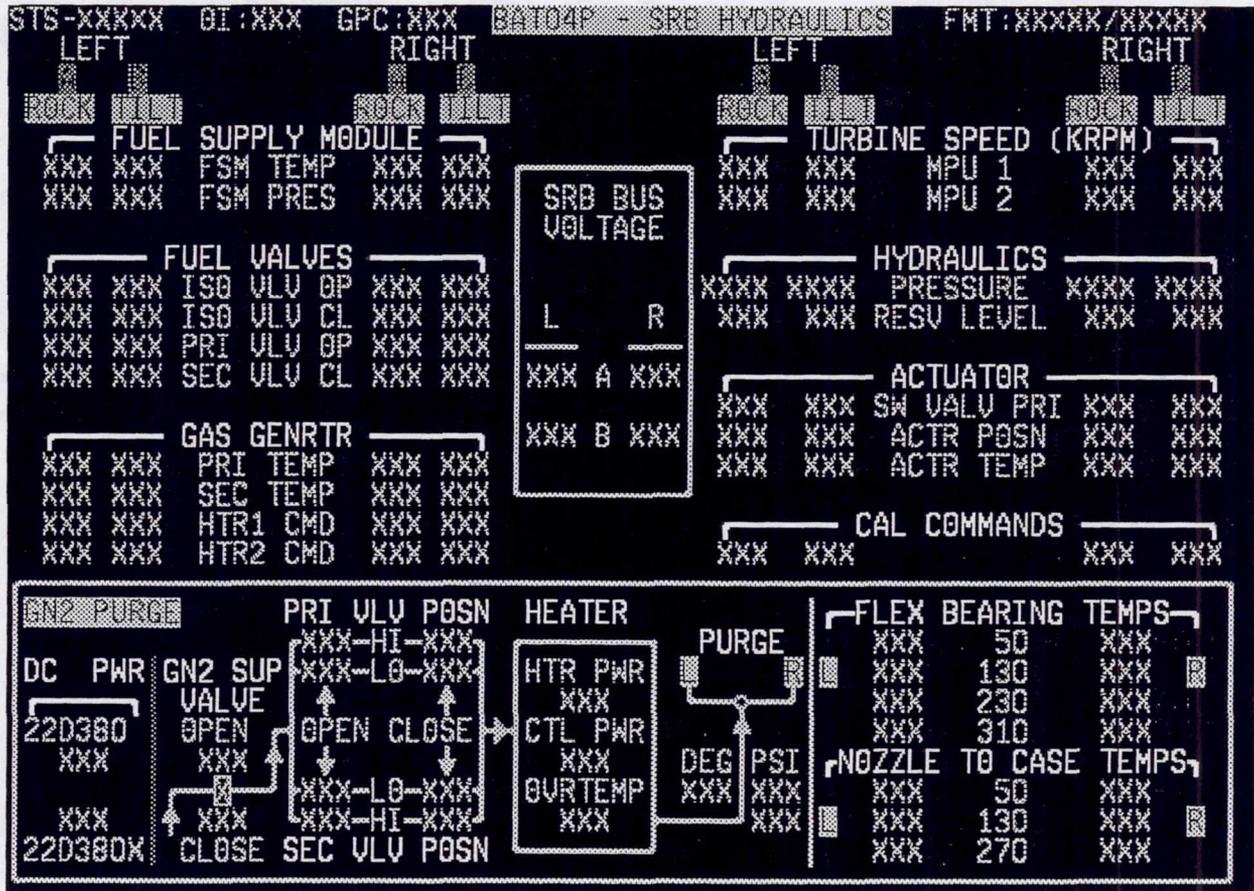


Figure 1: One of PC GOAL's display screen

2.1 Java applets

The key advantage of using Java applets is that after the applet makes a connection to a data server only the data is transmitted across the network. This is superior to constantly transmitting updated pictures (bit maps) as would be required in other approaches, e.g., common gateway interface (CGI), server push, etc.

The advantages in using Java are even greater than the technical merits suggest. Administratively, the operation of a real-time data service using Java is much better. The Java applet is written once and executed remotely; no porting has to be done. Also, the latest version of the applet is always distributed to the user; there is no version control problem. Finally, the operation of the service is easy as browsers are ubiquitous; no training is required.

2.2 Java Programming Language

The Java programming language introduced by Sun Microsystems a mere five years ago also provides technological advantages. The good network library has been important to Jview, as has the good GUI library. This latest version of this GUI library, known as Swing, was used in the visual editor. This has made it easy to produce an application with a pleasing appearance and great functionality. On the other hand, the performance of Swing is noticeably worse than the less sophisticated AWT library used in the previous version of the visual editor.

2.3 Jview Operation

The Jview system has been in operation for about two years. Since it emulates the existing PC GOAL display screens it has been easy for users to use the new system. All the display screens have the same appearance as the PC GOAL display screens, for example Figure 1. But advances in computer systems makes it reasonable to expect more sophisticated displays. It is possible to display the data more realistically with dynamic, bitmapped components: tanks can appear to be filling, analog gauges can be simulated, sound can incorporated, etc. Schematic diagrams can be more detailed.

To create these more sophisticated displays we have designed a visual editor. From an extensible palette of components the user composes a new graphical display for the data. Pictures and schematics (perhaps made from other tools like Adobe's PhotoShop) can be imported. Once a screen has been created, the data viewer displays it as well as updates the components with the individual data values as the data comes in.

It is important to emphasize that the visual editor puts the components together and does not create the components. At the moment, we have few entertaining components

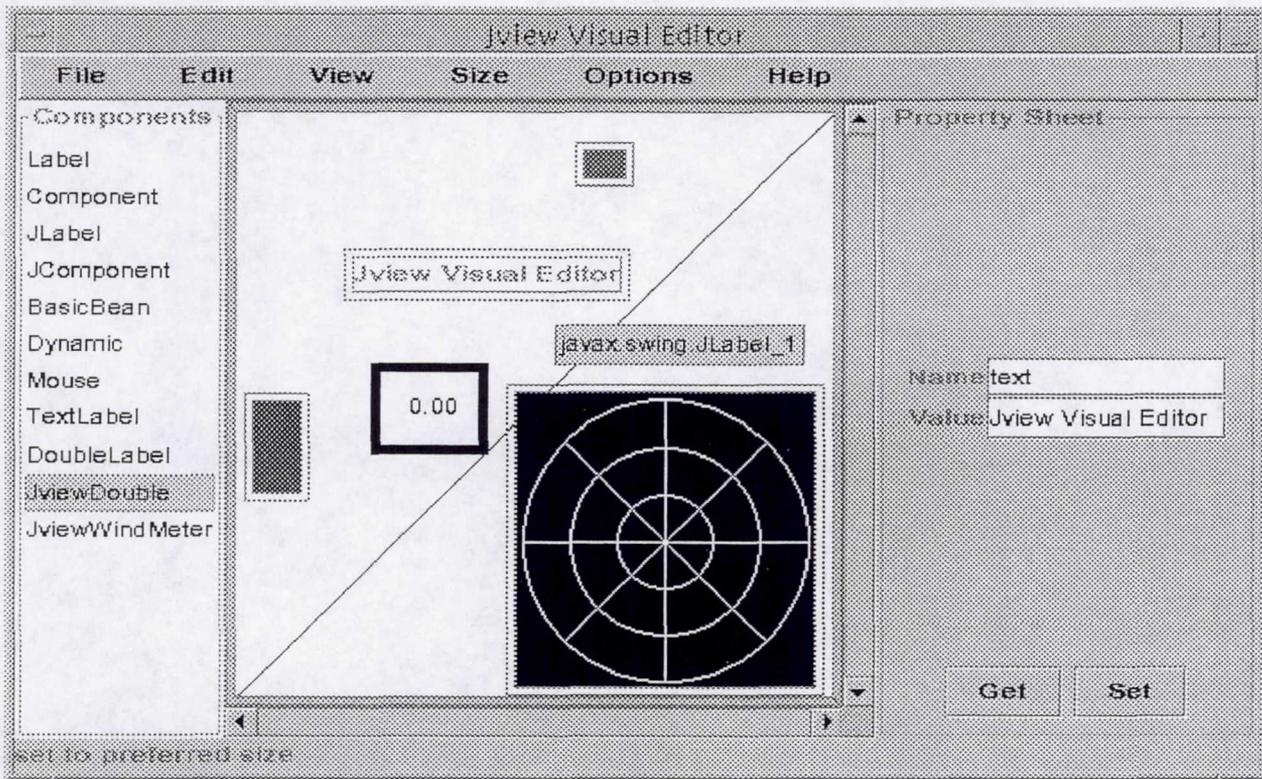


Figure 2: View of the editor

to use with the editor. It is likewise important to emphasize that the components can be created at anytime and the editor can use them without any change to the editor.

3 Visual Editor

The screen shot of the visual editor is shown in Figure 2. The list of available components appears on the left side of the workspace. In the workspace are five components. One is a `TextLabel` component for a label. Another is the specially created `JViewWindMeter` component. A third is the `JviewDouble` component for numbers; it is currently selected. It has the solid border around it. The current component always has a property editor in a separate little window called "property sheet." In this window the individual values for each property of the component can be set. No real data is reaching the component in the editor. But the display designer may want to see how the component reacts to some particular data

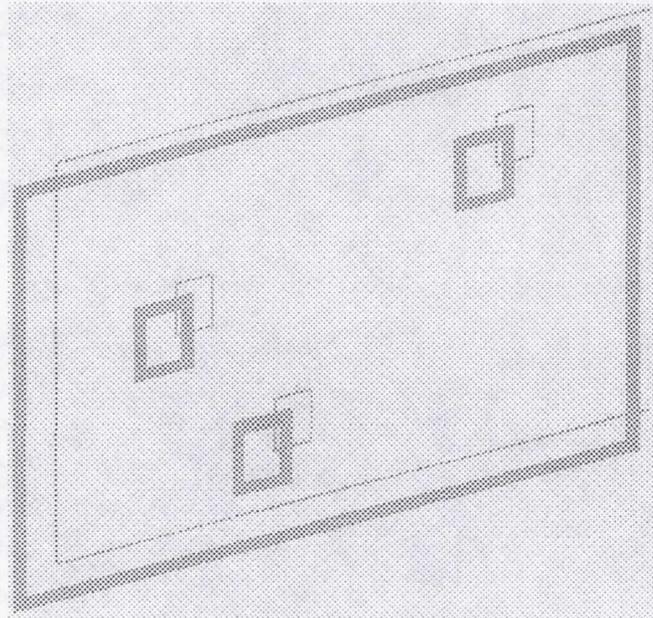


Figure 3: Two layered approach

value. The component may round the value to the nearest degree. Other components might turn a different color for some values, etc.

Good use was made of the Java Swing component `LayeredPane`. Using `LayeredPane` allowed the workspace to be composed of exactly the components that would be used in the display of the finished display screen. The extra effects and functionality of the editor were confined to a second layer that was placed in front. Figure 3 suggests the logical placement of the GUI components in the `Jview` editor. Displaying borders or capturing mouse events is accomplished by the front-most layer. In it a component (usually invisible) is placed in front of the actually display component.

Components can be widgets from either the AWT library or the new Java Swing library. These components are both "Java Beans." This means they follow certain programming conventions. For instance, `get` and `set` methods. For each property, say, wind speed, there is a method, say, `getWindSpeed` that returns the value of the private variable, and a method `setWindSpeed` that sets the value. My using this naming scheme, it is possible for the editor to manipulate a completely unknown component. The component does not have to be compiled with the editor at all and can be created long after the editor is deployed.

There is a limit to the type of components suitable for use with `Jview`. All `Jview` com-

ponents access the data through FD's and implement an update method that responds appropriately to new data values for the FD's.

4 Acknowledgments

I gratefully acknowledge the support of the DAP group of CLCS headed by Coleman Dugger (NASA) especially the Jview team: Steve Beltz (Dynacs), and Mark Long (Dynacs). Also, deserving thanks are Brad Neal (University of Central Florida) and Tom Beever, president of Netlander. A special thanks is due to Ray Hosler (University of Central Florida), Jane Hodges (NASA), and Greg Buckingham (NASA), who ran the NASA/ASEE Summer Faculty Fellowship Program at the Kennedy Space Center, for being friendly, efficient, and enthusiastic.

The original impetus of this project in 1996 grew out of collaboration with Peter Engrand (NASA) and Charlie Goodrich (Dynacs).

References

- [1] Patrick Chan and Rosanna Lee. *The Java Class Libraries: Second Edition, Volume 2*. Addison-Wesley, Reading, Massachusetts, 1998.
- [2] Patrick Chan, Rosanna Lee, and Douglas Kramer. *The Java Class Libraries: Second Edition, Volume 1*. Addison-Wesley, Reading, Massachusetts, 1998.
- [3] Robert Eckstein, Marc Loy, and Dave Wood. *Java Swing*. O'Reilly, Sebastopol, California, 1998.
- [4] Henri Jubin. *JavaBeans by Example*. Prentice Hall, Upper Saddle River, New Jersey, 1997.
- [5] NASA, Kennedy Space Center, Checkout and Launch Control System. CLCS project home page. WWW site at <http://clcs.ksc.nasa.gov/>.
- [6] NASA, Kennedy Space Center, Launch Processing System. LPS system software PC GOAL home page. WWW site at <http://lpsweb.ksc.nasa.gov/SDC/PCGOAL/>.
- [7] Sun Microsystems, Inc. Javabeans. WWW site at <http://java.sun.com/beans>.

1999 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

**JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA**

**MODELING AND ANALYSIS OF THE REVERSE WATER GAS SHIFT PROCESS
FOR IN-SITU PROPELLANT PRODUCTION**

PREPARED BY: Dr. Jonathan E. Whitlow

ACADEMIC RANK: Associate Professor

UNIVERSITY AND DEPARTMENT: Florida Institute of Technology
Chemical Engineering

NASA COLLEAGUE: Bill Larson

ABSTRACT

This report focuses on the development of mathematical models & simulation tools developed for the Reverse Water Gas Shift (RWGS) process. This process is a candidate technology for oxygen production on Mars under the In-Situ Propellant Production (ISPP) project. An analysis of the RWGS process was performed using a material balance for the system. The material balance is very complex due to the downstream separations and subsequent recycle inherent with the process. A numerical simulation was developed for the RWGS process to provide a tool for analysis and optimization of experimental hardware, which will be constructed later this year at Kennedy Space Center (KSC). Attempts to solve the material balance for the system, which can be defined by 27 nonlinear equations, initially failed. A convergence scheme was developed which led to successful solution of the material balance, however the simplified equations used for the gas separation membrane were found insufficient. Additional more rigorous models were successfully developed and solved for the membrane separation. Sample results from these models are included in this report, with recommendations for experimental work needed for model validation.

ACKNOWLEDGMENT

I would like to express my sincere gratitude to Bill Larson & Charlie Goodrich for the opportunity they have given me to be involved in the development of knowledge to aid in the exploration of Mars.

INTRODUCTION

The human exploration of Mars will require the utilization of resources present in the Martian environment in order to minimize the payload mass imported from Earth. ISPP is a joint effort through various NASA and contractor organizations aimed at producing fuel for Mars sample return missions. The primary candidate for fuel production is the Sabatier/Electrolysis process. Sabatier/Electrolysis produces methane and water from carbon dioxide and hydrogen in a catalytic reactor. The water is split with electrolysis producing oxygen and hydrogen, which is recycled back to the reactor feed. The product ratio of oxygen to methane produced is less than that required for a methane rocket, thus giving the need for an additional process to meet the oxygen requirements. RWGS is a candidate technology for oxygen production on Mars. This report details this process and describes a simulation model developed to aid in the analysis and optimization of experimental RWGS hardware which is to be constructed at KSC later this year. The RWGS design and inherent assumptions are taken from a report prepared by Pioneer Astronautics, who determined the RWGS process to be a viable candidate for ISPP in a SBIR project performed in 1997. ^[1]The RWGS process flow sheet is presented in Figure 1 below. An analysis of the process and each of its components follows. Details of the solution procedure used to solve the conservation equations for the system and sample results from the simulations performed are also presented.

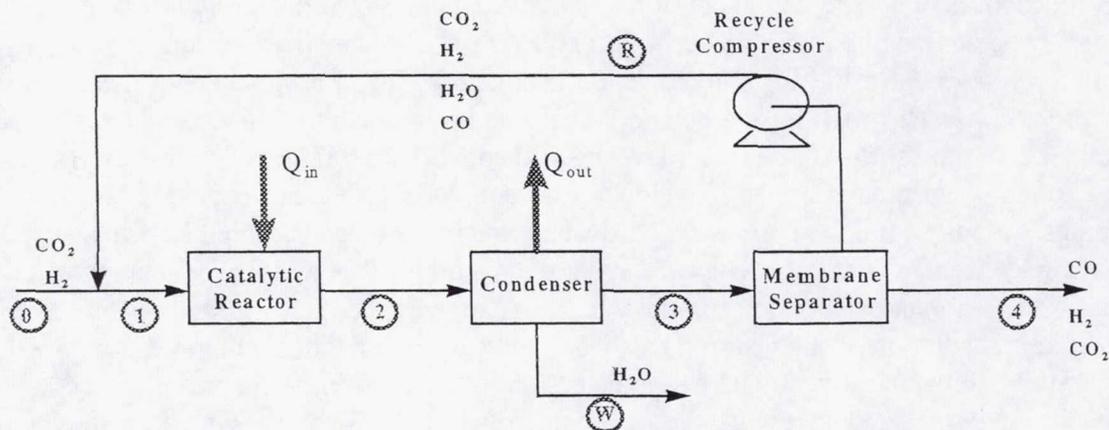


Figure 1 RWGS Process Flow Diagram

AN OVERVIEW OF THE REVERSE WATER GAS SHIFT PROCESS

The Reactor

RWGS uses carbon dioxide and hydrogen as reactants just as the Sabatier/Electrolysis process, however the desired products are oxygen and carbon monoxide. The selectivity of this reaction over the Sabatier reaction or some other side reaction is an issue in the development of RWGS. Previous work done on the RWGS reaction identified several catalysts with good selectivity.^[2] The selection of catalyst can affect the selectivity, rate of reaction and rate of deactivation, however the chemical equilibrium for any reaction is governed by thermodynamics, in particular the Gibbs free energy. The expression relating the chemical equilibrium coefficient to the Gibbs free energy, can be expressed by definition in terms of enthalpy and entropy as:

$$\Delta G = \Delta H - T \Delta S = RT \ln K \quad (1)$$

For the RWGS reaction equation 1 can be employed to obtain a value of the equilibrium constant K as a function of temperature:^[3]

$$K = e^{13.148 \frac{5639.5}{T} - 1.077 \ln T - 5.44 \times 10^4 T + 1.125 \times 10^7 T^2 + \frac{49170}{T^2}} \quad (2)$$

The value of K is thus solely a function of temperature and is related to the concentrations of the products and reactants and in terms of conversion of a limiting reactant by the expressions:

$$K = \frac{[CO][H_2O]}{[CO_2][H_2]} = \frac{(\Theta_{CO} + x_{eq})(\Theta_{H_2O} + x_{eq})}{(\Theta_{CO_2} - x_{eq})(1 - x_{eq})} \quad (3)$$

The values of Θ_{CO} , Θ_{H_2O} , Θ_{CO_2} represent the molar ratios of those components to the limiting reactant in the inlet stream to the reactor (assumed to be H_2 in the given case). It should be noted that the limiting reactant is based on the total feed to the reactor, (recycle stream + fresh feed) instead of the fresh feed to the reactor. Given a reaction temperature, the value of K can be determined from equation 2 and the equilibrium conversion determined by solving equation 3, which is a quadratic in x_{eq} . The equilibrium conversion is the maximum conversion, which can be achieved, in a single pass through the reactor.

The Condenser

Upon leaving the reactor, the exit gases are sent to a condenser where most of the water is removed. In determining the compositions of the liquid and vapor streams leaving the condenser, idealities were assumed. Raoult's Law was used to determine the amount of water in the vapor

phase by assuming that the partial pressure of the water was equal to its vapor pressure, which is readily available in the literature. Henry's Law and corresponding Henry's Law constants were used to determine the solubility and hence concentration of the gasses dissolved in the condensed water. A more rigorous calculation could have been performed for the condenser, however it was felt that the effort involved was not justified since it would have required the knowledge of equilibrium ratios which are not readily available, to solve the 10 equations shown in Figure 2.

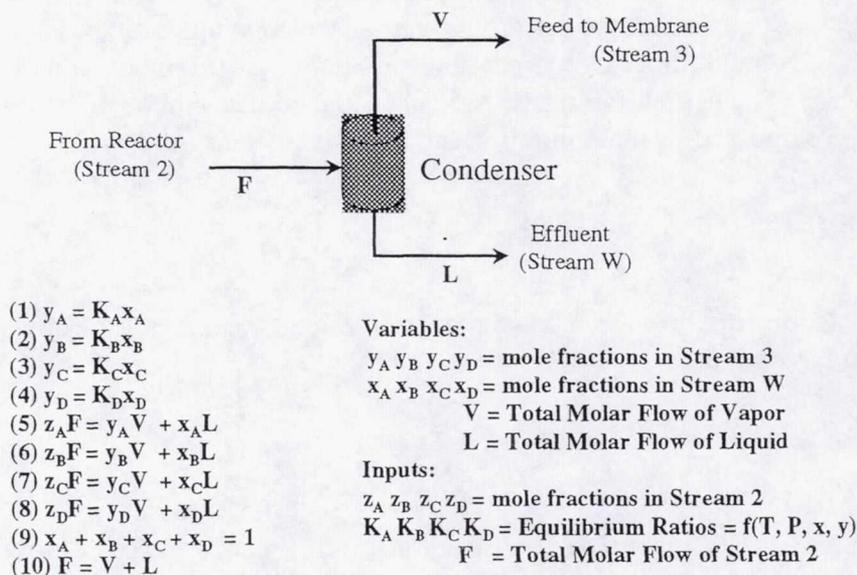


Figure 2 Equations Required for Rigorous Condenser Calculations

Gas Separation Membrane

As shown in Figure 1 the membrane is directly downstream of the condenser. The use of membranes for separation of gases is a relatively new technology. In theory, the gas components can be adsorbed onto the surface of the polymer and providing that there is a concentration gradient across the membrane, they will diffuse to the low pressure or permeate side. At that point the gas desorbs from the membrane into the bulk fluid permeate. The type of membrane specified for this process is composed of hollow fiber polymers. This type of membrane is the most economical on a surface area per cost basis however other membrane types such as spiral wound and plate and frame have superior fouling and pressure drop characteristics.^[4]

The polymeric membranes separate the components of a gas mixture based on a given components permeability to the polymer. Permeability for this type of membrane can be defined as the product of a given components solubility and its molecular diffusion coefficient in the polymeric membrane. Fick's law of diffusion is the principle governing this separation technique

as the permeation flux through the membrane is proportional to the concentration gradient across the membrane. Since both solubility and diffusion are temperature dependent, permeability and hence mass transfer across the membrane are likewise temperature dependent.

In developing a model for the polymeric hollow fiber membranes several approaches can be taken. Three distinct models are presented here, namely complete mixing, cross flow and counter-current flow. The simplest model, which is known as complete mixing, assumes the membrane as a single stage in which the feed is split into the permeate and retentate or residue. The mass transfer rate for a given component j across the membrane is given by the equation:

$$m_j = Q_j A (P_{L,j}^* - P_{V,j}^*) \quad (4)$$

Here m_j is the mass transfer of component j across the membrane, Q_j is an effective permeability for component j and $P_{L,j}^*$ and $P_{V,j}^*$ are the partial pressure of j in the feed and permeate respectively. Specifying the feed to the membrane, the transport equations given in equation 4 can be rewritten in terms of mole fractions and solved. The solution requires the simultaneous solution of the overall balance, the component mole balances around the membrane and one of the constraints that the mole fractions in both the residue and permeate equal unity.

This complete mixing model does not account for the fact that the partial pressures on both sides of the membrane change as the entering gas mixture proceeds along the length of the membrane to the exit. In the cross flow model, the membrane is treated as a series of stages in which the gas is transferred from the residue stream into the permeate, and a variation in the residue concentration down the membrane is introduced. This adds complexity to the model in that instead of having a single equation for each component describing the mass transfer across the membrane, N equations must be written for each component. For a given stage k , equation 1 can be redefined in terms of mole fractions for component j as:

$$y_{j,k} = \frac{Q_j (P_{L,k} x_{j,k} - P_{V,k} y_{j,k})}{\sum_{m=1}^R Q_m (P_{L,k} x_{m,k} - P_{V,k} y_{m,k})} \quad (5)$$

Here $P_{L,k}$ and $x_{j,k}$ are the total pressure and mole fraction of j in the feed while $P_{V,k}$ and $y_{j,k}$ are the total pressure and mole fraction of j in the permeate. For this model the composition of gas produced on a stage depends on the upstream compositions, permeances and pressures, however is independent of the composition of the gas produced downstream in other stages. Discussion of solving this model is presented later when the algorithm developed is presented.

The counter-current model depicted in Figure 3 is the most rigorous model and also best describes the true system, which is operated as a counter-current separation membrane. Here as with the cross-flow model, the system is divided into N stages.

Unlike the cross-flow model however the compositions change on both permeate side as well as the residue side. Hence a material balance can be written on any stage k as:

$$l_{j,k+1} - l_{j,k} + v_{j,k-1} - v_{j,k} = 0 \quad (6)$$

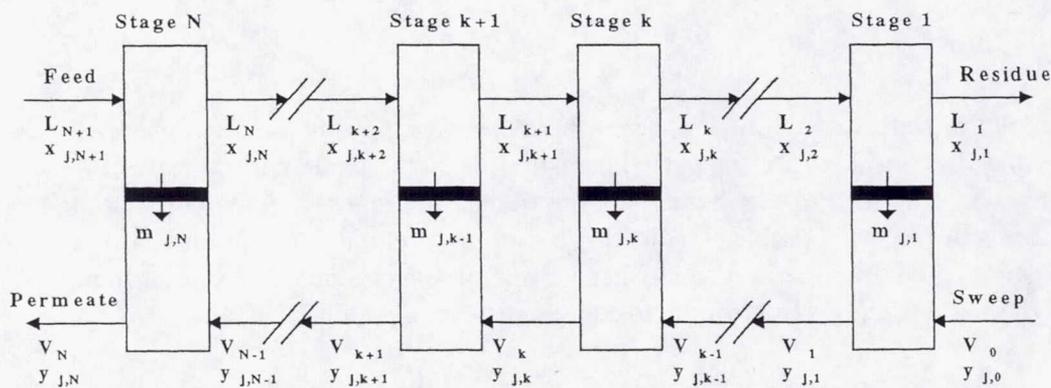


Figure 3 Flow Sheet for an N Stage Counter-Current Membrane

Here l & v represent the residue & permeate molar flows for component j leaving or entering stage k . The convention for use of the subscripts referring to k in the above equation is based on the stage the component is leaving, hence $l_{j,k}$ is leaving the $k+1$ stage but it is entering the k th stage. The transport of component j across the membrane on stage k can be denoted as $m_{j,k}$ which is equal to the difference in $l_{j,k+1}$ and $l_{j,k}$. The driving force for the transport is the difference in partial pressures of the component on each side of the membrane as was given in equation 4. If equation 4 is written for a general stage k using mole fractions instead of partial pressures the resulting equation is:

$$m_{j,k} = Q_j \Delta A (P_{Lk} x_{j,k} - P_{Vk} y_{j,k}) \quad (7)$$

A is the total surface area divided by the number of stages while the other quantities have been previously defined. If the mole fractions on a given stage in equation 7 are redefined as a component flow rate divided by the total flow rate the resulting equation is:

$$m_{j,k} = l_{j,k+1} - l_{j,k} = Q_j \Delta A \left(\frac{P_{Lk} l_{j,k}}{L_{j,k}} - \frac{P_{Vk} y_{j,k}}{V_{j,k}} \right) \quad (8)$$

When this expression is rearranged and solved for the component flow rate in the permeate $v_{j,k}$ the expression becomes:

$$v_{j,k} = \frac{-V_k}{P_{V_k} Q_j \Delta A} \left[l_{j,k+1} - \left(I + \frac{Q_j \Delta A P_{Lk}}{L_k} \right) l_{j,k} \right] \quad (9)$$

Using equation 9 and substituting for $v_{j,k}$ and $v_{j,k-1}$ in equation 6 results in the equation:

$$B_{j,k} l_{j,k-1} + C_{j,k} l_{j,k} + D_{j,k} l_{j,k+1} = 0 \quad (10)$$

where:

$$B_{j,k} = \frac{-V_{k-1}}{P_{V_{k-1}} Q_j \Delta A} \left(I + \frac{Q_j \Delta A P_{L_{k-1}}}{L_{k-1}} \right)$$

$$C_{j,k} = I + \frac{-V_{k-1}}{P_{V_{k-1}} Q_j \Delta A} + \frac{V_k}{P_{V_k} Q_j \Delta A} \left(I + \frac{Q_j \Delta A P_{Lk}}{L_k} \right)$$

$$D_{j,k} = \frac{-V_k}{P_{V_k} Q_j \Delta A} - I$$

If equation 10 is applied to each of the N stages, a system of N nonlinear simultaneous equations is yielded for each component j . Written in matrix form results in a tridiagonal coefficient matrix where the values for $C_{j,k}$ are on the main diagonal with the values of $B_{j,k}$ and $D_{j,k}$ occupying the adjacent columns. Since the elements of coefficient matrix, $B_{j,k}$, $C_{j,k}$ and $D_{j,k}$ are functions of the variable, residue molar flow rates, nonlinearity is present in the system. The methods for solution of this model is based on recent literature^[5], in which successive approximations for the molar flow rates $l_{j,1}$ to $l_{j,N}$ are used to evaluate, the values of $B_{j,k}$, $C_{j,k}$ and $D_{j,k}$ until the system converges. The initial guesses for the molar flow rates in the residue are based on the solution of the cross flow case.

The RWGS system can be defined by 27 independent material balance equations if the complete mixing model is assumed for the membrane. Initial attempts to solve the system of nonlinear equations, which have not been included for brevity, proved unsuccessful. This resulted in the development of an iterative algorithm in which the recycle stream molar flow rates were used as tear variables. In other words, an initial estimate was made for these variables and then the rest of the system equations were solved in succession until new values of the tear variables were generated. Convergence could then be achieved when the values guessed matched the values calculated. The first attempts at implementing the algorithm, which is presented in Figure 4, involved successive substitution of the calculated values of the recycle stream into the guessed values.

This method diverged and hence required an alternative tearing method. The Wegstein method was employed for this task with success, as it is more stable than successive substitution.¹⁶ The Wegstein method is based on the algorithm:

$$x_{n+1} = \frac{x_{n-1} g(x_n) - x_n g(x_{n-1})}{x_{n-1} - g(x_{n-1}) - x_n + g(x_n)} \quad (11)$$

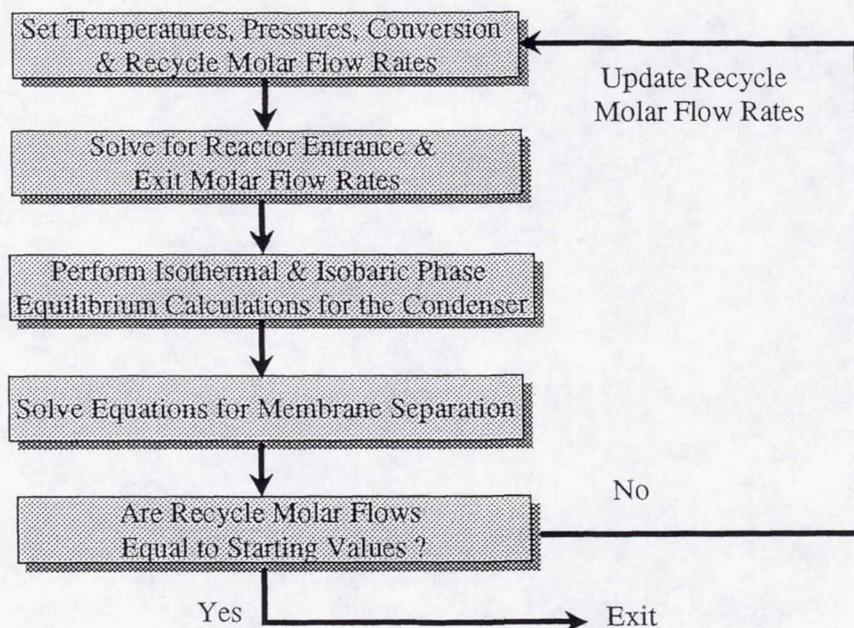


Figure 4 RWGS Material Balance Solution Algorithm

While the system converged with the aforementioned algorithm, the values of permeability supplied by Pioneer,¹¹ had to be reduced for all components except carbon monoxide, in order to prevent the mole fractions of those components in the residue, from being less than zero. This ultimately led to the use of the cross flow and counter-current methods previously described, for estimating the material balance around the membrane. The number of stages used in the simulation was 100, which is recommended as a minimum value by the developers of the method.

All programming for the RWGS material balance solution was implemented in MATLAB code. While convergence was obtained using the rigorous calculations, the convergence was very time consuming. In addition, the area of the membrane had to be slightly reduced for convergence. In this case, as with the solution using the complete mixing model, the permeability's used resulted in numerical stability problems. Sample results obtained from the algorithm based on a reactor temperature of 370 °C and a condenser temperature of 2 °C are presented in Figure 5. The convergence for this case required in excess of 300 iterations of the RWGS algorithm and took on the order of 2 hours to complete.

REACTOR CALCULATIONS: (Reactor Temperature = 370 Centigrade)

The Single Pass Equilibrium Conversion is 0.2127
The Molar Flow of H₂ in Reactor Feed is 53.6990 gmoles/hr
The Molar Flow of CO₂ in Reactor Feed is 23.6718 gmoles/hr
The Molar Flow of CO in Reactor Feed is 5.8846 gmoles/hr
The Molar Flow of H₂O in Reactor Feed is 0.1734 gmoles/hr
The Molar Flow of H₂ in Reactor Exit is 48.6642 gmoles/hr
The Molar Flow of CO₂ in Reactor Exit is 18.6370 gmoles/hr
The Molar Flow of CO in Reactor Exit is 10.9194 gmoles/hr
The Molar Flow of H₂O in Reactor Exit is 5.2082 gmoles/hr

CONDENSER CALCULATIONS: (Condenser Temperature = 2 Centigrade)

The Total Molar Feed Flow is 83.428827 gmoles/hr
The Molar Flow of H₂ in the Vapor is 48.6640 gmoles/hr
The Molar Flow of CO₂ in the Vapor is 18.6324 gmoles/hr
The Molar Flow of CO in the Vapor is 10.9193 gmoles/hr
The Molar Flow of H₂O in the Vapor is 0.1734 gmoles/hr
The Total Molar Flow of the Vapor is 78.3891 gmoles/hr
The Molar Flow of H₂ in the Liquid is 0.0002 gmoles/hr
The Molar Flow of CO₂ in the Liquid is 0.0046 gmoles/hr
The Molar Flow of CO in the Liquid is 0.0001 gmoles/hr
The Molar Flow of H₂O in the Liquid is 5.0348 gmoles/hr
The Total Molar Flow of the Liquid is 5.0397 gmoles/hr

MEMBRANE CALCULATIONS: (Total Area Used = 29.98 m²)

The Molar Flow of H₂ in the Permeate is 48.4826 gmoles/hr
The Molar Flow of CO₂ in the Permeate is 18.3818 gmoles/hr
The Molar Flow of CO in the Permeate is 5.8844 gmoles/hr
The Molar Flow of H₂O in the Permeate is 0.1734 gmoles/hr
The Total Molar Flow of the Permeate is 72.9222 gmoles/hr
The Molar Flow of H₂ in the Residue is 0.1814 gmoles/hr
The Molar Flow of CO₂ in the Residue is 0.2506 gmoles/hr
The Molar Flow of CO in the Residue is 5.0349 gmoles/hr
The Molar Flow of H₂O in the Residue is 0.0000 gmoles/hr
The Total Molar Flow of the Residue is 5.4669 gmoles/hr

OVERALL MASS BALANCE CALCULATIONS:

The Overall pct. Conversion of H₂ is 96.5192
The Overall pct. Conversion of CO₂ is 95.1757
The Overall error in the mass balance based on Hydrogen is 0.0007 gmoles
The Overall error in the mass balance based on Carbon is -0.0002 gmoles

RECYCLE STREAM

The Recycle Flow of H₂ to the Reactor is: 48.4826 gmole/hr
The Recycle Flow of CO₂ to the Reactor is: 18.3818 gmole/hr
The Recycle Flow of CO to the Reactor is: 5.8844 gmole/hr
The Recycle Flow of H₂O to the Reactor is: 0.1734 gmole/hr
The Total Recycle Flow to the Reactor is: 72.9222 gmole/hr

Figure 5 Sample Results for Solution of the RWGS Material Balance

CONCLUSIONS

The material balance for the RWGS process was successfully solved in MATLAB using the iterative algorithm and rigorous calculations for the gas separation membrane. Model validation is required to employ the simulation tool developed for use in analysis and optimization of the RWGS hardware, which will be constructed at KSC. The validation process should include experimental determination of permeability through the polymeric membrane for each component. Performance degradation, which can sometimes accompany the hollow fiber membranes, should also be evaluated. Additional model validation tasks should include analysis of the reactor kinetics and confirmation that equilibrium is obtained for the RWGS reaction, as well as evaluation of the phase distribution of components in the condenser.

To complete the model, energy balance calculations should be incorporated into the RWGS simulation. The development of a rigorous process model for the RWGS system could have significant impact on the development of the technology for this system. In addition to the analysis and optimization aspects previously discussed, this type of model can lead to abstraction of deeper knowledge for use in autonomous control. The development of virtual sensors could also be a by-product of this type of modeling effort. This would prove useful due to the mass constraint, which will be placed on missions to Mars.

REFERENCES

- [1] Zubrin R., Kito, T. and B. Frankie, "Report on the Construction of a Mars Methanol In-Situ Propellant Production Unit," NASA Contract NAS9-97082, Sept., 1997.
- [2] Frankie, B. and R. Zubrin, "Mars Reverse Water Gas Shift Unit Design Manual", Delivered to Dynacs Engineering and The Embedded Technology Development Group, June 30, 1999.
- [3] Bissett, L. "Equilibrium Constants for Shift Reactions", Chemical Engineering, October 24, 1977.
- [4] Humphrey, J.L. and G. Keller, *Separation process technology*, McGraw-Hill, 1997.
- [5] Coker, D.T., Freeman, B.D. and G.K. Fleming, "Modeling Multicomponent Gas Separation Using Hollow-Fiber Membrane Contactors, AIChE Journal, V. 44, N 6, June 1988.
- [6] Constantinides, A. and N. Mostoufi, *Numerical Methods for Chemical Engineers with MATLAB Applications*, Prentice Hall, 1999

1999 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM
JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA

*TESTING PLANT RESPONSES TO RARIFIED
ATMOSPHERES FOR INFLATABLE GREENHOUSES*

Kenneth A. Corey
Associate Professor
Department of Plant & Soil Sciences
University of Massachusetts

Colleagues: Raymond M. Wheeler and Philip A. Fowler

ABSTRACT

Reduced atmospheric pressures will likely be used to minimize mass and engineering requirements for plant growth habitats used in extraterrestrial applications. A chamber with high vacuum capability was used to design and begin construction of a system for testing plant responses to reduced pressure atmospheres. Several preliminary tests were conducted to evaluate chamber suitability for plant tests and to determine performance of thermal and vacuum systems at ambient and reduced pressure atmospheres down to 0.1 atm. The first tests consisted of measurements of internal gas volume and leakage rate. The method for volume determination was quite sensitive and will be needed for plant gas exchange measurements and calculations. This information will also be used in conjunction with the leak rate. Measured leak rates on the order of 0.46 mm Hg/min at 76 mm Hg pressure are low enough to conduct sensitive carbon dioxide exchange rate measurements at reduced pressure given an adequate plant sample (0.5 to 1.0 m² area). A test rack with lighting provided by 3, high-pressure sodium vapor lamps was built to accommodate both short-term and long-term plant responses. Initial short-term experiments with lettuce showed that a pressure of 77 mm Hg resulted in a 6.1-fold increase in the rate of water loss compared to water loss at ambient pressure. Plants were severely wilted after 30 minutes exposure to 77 mm Hg. Water loss was found to be inversely correlated with atmospheric pressure over the range of pressures from 0.2 to 1.0 atm; the rate of water loss at 0.2 atm was 4.3 times higher than water loss at ambient pressure. Older leaves showed moderate wilting during exposure to 156 mm Hg, but those exposed to 345 mm Hg remained turgid. Results suggest a reduced atmospheric pressure limit of 0.2 to 0.3 atm for lettuce grown in a solid medium. Follow-up experiments with carbon dioxide control and control at high relative humidity (> 90 %) will be needed to further confirm and define safe reduced pressure limits that are feasible for plant tolerance and growth.

SYSTEM DESIGN AND TESTING OF PLANT RESPONSES TO RARIFIED ATMOSPHERES FOR INFLATABLE GREENHOUSES

Kenneth A. Corey

1. INTRODUCTION

Long term economic trade-off studies suggest that advanced life support systems for extraterrestrial applications will utilize plants to supply human life support requirements for food, oxygen, and potable water [3]. There is a need to reevaluate and update such studies to incorporate in situ resource utilization practices for specific scenarios such as reduced atmospheric pressures for inflatable structures on Mars. Even without this rationale, it is certain that plants will at some time in the not too distant future be a vital and integral component of the human exploration and settlement of space. Given the certainty of plant culture outside Earth environments, questions arise regarding the pressure and composition of the atmosphere of growth habitats. If humans and plants share the same atmospheric volumes, then plant culture is constrained by the priority of human requirements. Human requirements will likely not involve oxygen partial pressures below 15 to 20 kPa or total atmospheric pressures below 50 kPa. These partial pressure values are fairly conservative given the fact that humans are known to adapt to much lower partial pressures of oxygen such as those experienced at high altitude villages (e.g. Jiachan, Tibet).

It may not be necessary to consider habitats designed for integration of both plants and people since pressures selected for human habitation are probably not those that would be optimum for growth of plants. Also, in order to optimize the generation of human resource requirements by bioregenerative means, structures that optimize the total volume to growth area ratio are desirable (< 2). Furthermore, structural requirements to contain a pressure gradient decrease with decreasing pressure. Therefore, a premise of this work is to consider the use of reduced pressure atmospheres in autonomous plant growth structures [2,5] that would be isolated from human habitation, and provide, in early phase advanced human life support systems, a back-up or perhaps lifeboat to physical-chemical systems. With such a premise, it follows that it will be necessary to define the limits of atmospheric pressure and partial pressures of oxygen and carbon dioxide for growth of plants.

Several test facilities have been used to assess metabolic and developmental responses of plants to reduced pressure [1,6-10,12,18,20,22]. However, most studies thus far have not provided clear separation of pressure and oxygen effects, nor have they involved complete growouts of large plant samples for assessment of yield. On the basis of enhanced diffusion of gases at reduced pressure, it is expected that water flux will increase. This effect has been documented in several studies [6,8,10,12,18]. If reduced pressure is also accompanied by reduced partial pressure of oxygen, enhancement of net photosynthesis and growth may occur through a reduction in carbon loss by suppressing photorespiration. Photorespiratory effects in C_3 pathway plants such as wheat are well documented [4,13,15,16,19,21] and have been reviewed [11,14,17]. Thus, reduced pressure atmospheres provide an additional rationale for testing responses of entire crop stands to reductions in oxygen partial pressure.

2. OBJECTIVES

The objectives of the current research program are to: 1) test the performance of the thermal and pressure control capabilities of the Thermotron, 2) make volume and leak rate determinations to evaluate the suitability of the chamber for use in plant gas exchange measurements, 3) conduct preliminary experiments with small samples of plants to determine the relationship of water flux (transpiration measurements) with pressure and to determine low pressure limits for lettuce, and 4) begin the design and construction of a reduced pressure testbed for the conduct of plant growth experiments in rarefied atmospheres.

3. EXPERIMENTS

3.1 Vacuum Chamber

The thermal vacuum chamber or Thermotron (Thermotron Industries, Holland, Michigan) is a high vacuum chamber rated for 1 torr and thermal control in the range of -72 to 177 C. Effective internal dimensions of the chamber are 1.22 m wide X 1.22 m high X 1.62 m deep. The blower and motor are housed internally in the rear of the chamber. The vacuum pump is located external to the chamber. Temperature and pressure measurements were made with thermocouples and a Barotron pressure transducer, respectively.

3.2 Volume Measurement

A first step in the preparation for measurement of rates of gas exchange of plants is a sensitive measurement of the total free gas volume of the atmosphere in the chamber. Since the chamber contains numerous irregular objects such as fans, motors, and blowers, it is not feasible to make internal atmospheric volume determinations by measurements of dimensions and geometries. A practical approach to such a measurement is to introduce a known quantity of gas and measure the change in concentration of the gas associated with the added quantity of gas, followed by application of the ideal gas law. At this time, there are no penetrations with injection ports to permit this procedure. Therefore, an alternative approach was used to circumvent this limitation. A preliminary calculation of the mass of carbon dioxide required to bring the chamber volume to a concentration of 1000 ppm based on an approximate volume of 3.5 cubic meters was made. A piece of dry ice (initial weight = 3.370 grams) was placed on a scale with an accuracy of 10^{-3} grams. A LICOR infrared gas analyzer was used to measure the concentration of carbon dioxide. A small fan was used to ensure rapid mixing of the atmosphere as the carbon dioxide sublimed. Readings of dry ice weight and carbon dioxide concentration were made until all of the initial weight had sublimed. The experiment was carried out under isothermal and isobaric conditions; the average temperature and pressure during the test were 20 C and 765 mm Hg, respectively. The change in weight of dry ice stopped at 0.05 grams, presumably because a small quantity of water condensed out from the atmosphere by the local air cooling of the dry ice or because there was water associated within the structure of the piece of dry ice. Results of the measurements indicated an excellent linear fit of the data yielding a slope of 6.2×10^{-3} grams/ppm (Figure 1A). Volume of the chamber was calculated from eq. [5] using the following steps. Let Δv = volume of CO_2 added and written as:

$$\Delta v = \Delta C V_{\text{ch}} \quad [1],$$

where ΔC is the change in CO_2 concentration and V_{ch} is the chamber volume. This equation may also be written as:

$$\Delta v = \Delta w V_i / M \quad [2],$$

where Δw is the change in mass of dry ice, V_i is the molar volume of an ideal gas, and M is the molecular weight of CO_2 . From the ideal gas law,

$$V_i = n R T / P \quad [3]$$

Rearrangement of eq.[1] as

$$V_{\text{ch}} = \Delta v / \Delta C \quad [4]$$

and substituting eq. [2] and [3] into [4] gives

$$V_{\text{ch}} = \Delta w n R T / \Delta C M P \quad [5]$$

The volume calculated from eq. [5] was 3,382 liters. This value is reasonable since it is appropriately less than a volume of 3,577 liters calculated by straight dimension measurements assuming a box geometry and not subtracting out the volumes occupied by equipment, motors, blower, and other irregular objects. This straightforward method of volume measurement appears quite sensitive, is rapid, and could be repeated easily when additional equipment is added to the chamber (i.e. test rack, lights, plant growth system, etc.). In order to

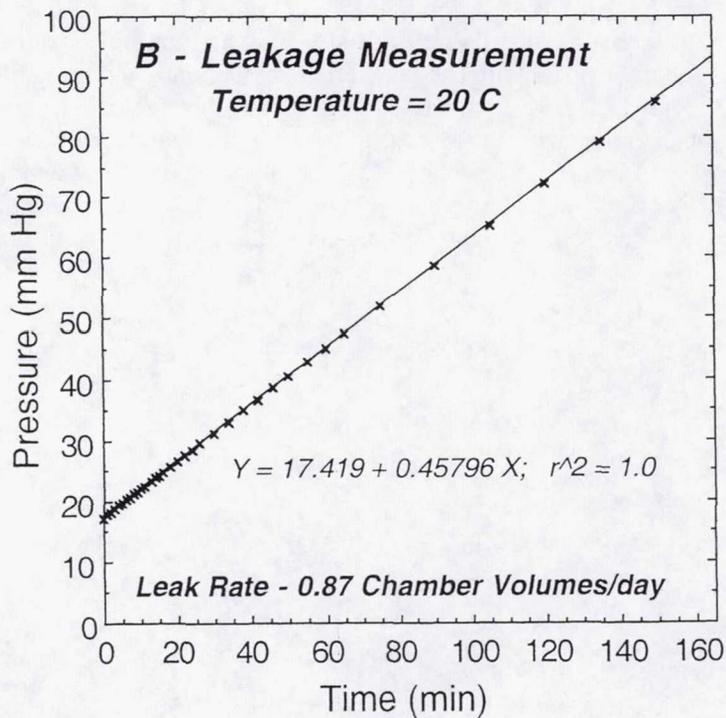
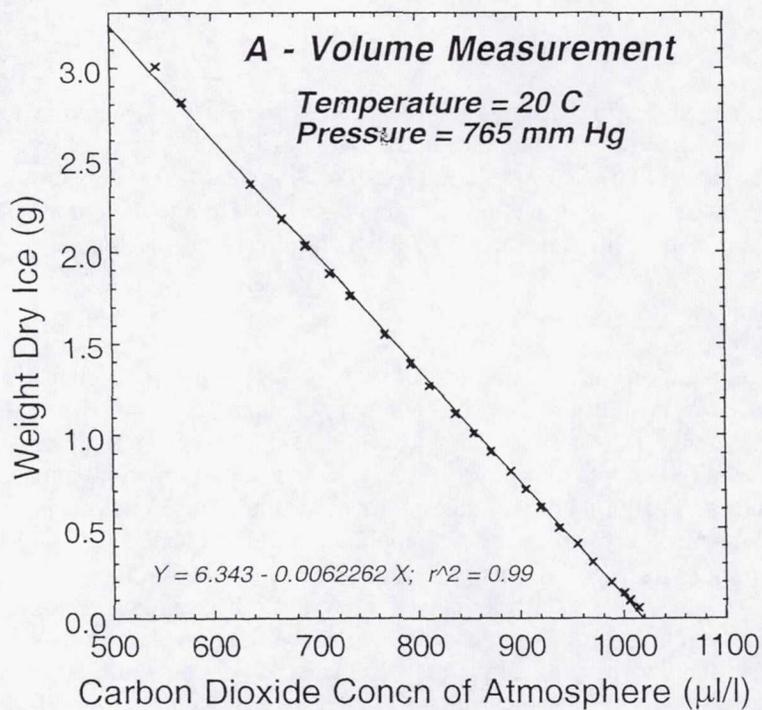


Figure 1. Changes in carbon dioxide concentration and weight of dry ice as it sublimed into the Thermotron atmosphere (A) and changes in atmospheric pressure in the Thermotron due to leakage (B).

apply equations for calculating plant gas exchange measurements, the volume measurement so obtained may be used in combination with leakage measurements.

3.3 Chamber Leakage Measurement

Vacuum chambers have penetrations and seals that may lead to leakage of external air into the chamber. Measurements involving the rates of uptake or evolution of a gas such as carbon dioxide during plant photosynthesis and respiration will be affected by significant leakage rates and therefore must be measured. The most rapid and straightforward method for leakage measurement is to disable the vacuum pump and follow the pressure increase over time. Evaluation of the first derivative of this function at the pressure of interest will give a chamber leakage value that can be applied to making corrections to measurements of plant metabolic rates. A detailed treatment of leakage measurements, calculations, and application to gas exchange measurements has been reported [6,7].

Results of the first of such rate of rise tests was conducted on June 16, 1999. The chamber was pumped down to 17 mm Hg, the pump disabled, and the pressure allowed to increase up to 86 mm Hg. Chamber temperature during the test was between 20 and 21 C. The leakage rate was measured to be 0.46 mm Hg/min (Figure 1B). Using the volume measurement, the air leakage rate of the chamber, L_a , was calculated as 0.87 chamber volumes per day at 20 C. Based upon previous experiences with gas exchange measurements at reduced pressure [8], this value is low enough to permit sensitive measurements of CO_2 uptake measurements, given a sufficient plant sample size. The plant sample size required for the acquisition of short term measurements of good sensitivity and reasonable duration for the Thermotron is in the range of 0.5 to 1.0 m^2 area.

3.4 Plant Test Stand

The development of a test stand for this project involves two phases; the first to conduct short term determinations of physiological responses of crop plants to atmospheric pressure and composition. These experiments will involve measurements of water flux (transpiration) and rates of photosynthesis and dark respiration. The second phase will involve determination of longer-term growth responses in addition to the gas exchange measurements mentioned. A rack was built to accommodate the space and light requirements for both phases. Three, 400-W high pressure sodium (HPS) vapor lamps were mounted on a rack that measured 112 cm wide X 152 cm deep X 116 cm high. Photosynthetic photon flux measurements were made with a LICOR quantum sensor and gave values in the range of 300 to 400 $\mu\text{mol}/\text{m}^2\text{s}$ depending on position and canopy height; more than adequate for testing short term physiological responses or for growth of lettuce plants.

The first phase involved testing small samples of lettuce (*Lactuca sativus* cultivar Waldeman's Green) plants grown in a controlled environment growth chamber. Plants were grown at a temperature of 22 C, 75 % relative humidity, a photosynthetic photon flux of 260 $\mu\text{mol m}^{-2} \text{s}^{-1}$, a CO_2 of 1200 ppm and a light/dark cycle of 18 hr/6 hr. Seeds were sown in a solid medium (1:1 peat-vermiculite mix), transplanted as seedlings into the same medium, and grown in 15-cm diameter plastic pots. Plants were fertilized with half-strength modified Hoagland's solution every other day until 15 days-old, and every day thereafter.

3.5 Lettuce Transpiration Experiments

The first chamber test with lettuce involved placing 2 plants on a scale (0.1 g sensitivity) and monitoring weight loss at ambient pressure, followed by pumping the chamber down to a pressure of 77 mm Hg. Plants were watered to bring the soil up to an approximate field capacity moisture content prior to the start of the experiment, and then placed in plastic bags that were tucked loosely under the foliage to minimize the evaporative water loss component. Temperature control for the comparison was excellent, but relative humidity was lower at reduced pressure (~30 %). Weight loss was over 6-fold higher at 77 mm Hg pressure

and plants exhibited severe wilting from which they recovered fully in about 30 minutes after return to ambient pressure (Figure 2). The next test involved an incremental step down in atmospheric pressure from ambient with plants held at each pressure for about 30 minutes each. Since plants exhibited severe wilting at 77 mm Hg, the lowest pressure treatment selected was 156 mm Hg (~0.2 atm). There was a progressive increase in rate of water loss with decreasing pressure; the rate at 156 (~0.2 atm) being about 4.3-fold higher than the water loss at ambient pressure (Figure 3A & 3B). Over the range of 156 to 766 mm Hg, the rate of water loss was inversely correlated with pressure (Figure 3B). Relative humidity was controlled fairly well, though it was lower (68 %) for the 156 mm Hg treatment than the average of 76 % across all treatments. The difference in the leaf to atmosphere vapor pressure deficit for that difference in relative humidity is presumed to be small in comparison to the pressure effect.

Results of the previous experiment were confirmed, with water loss expressed on a leaf area basis being 6.8 times higher at 147 mm Hg than that of ambient pressure (Table 1). Only slight wilting of the older leaves was observed on the low pressure treatment. The next experiment simply involved a partial repetition of the previous experiment with a direct comparison of ambient atmospheric pressure and 147 mm Hg. At the end of the experiment, all leaves of each plant were detached and area determinations made with a LICOR portable area meter (model LI-300A).

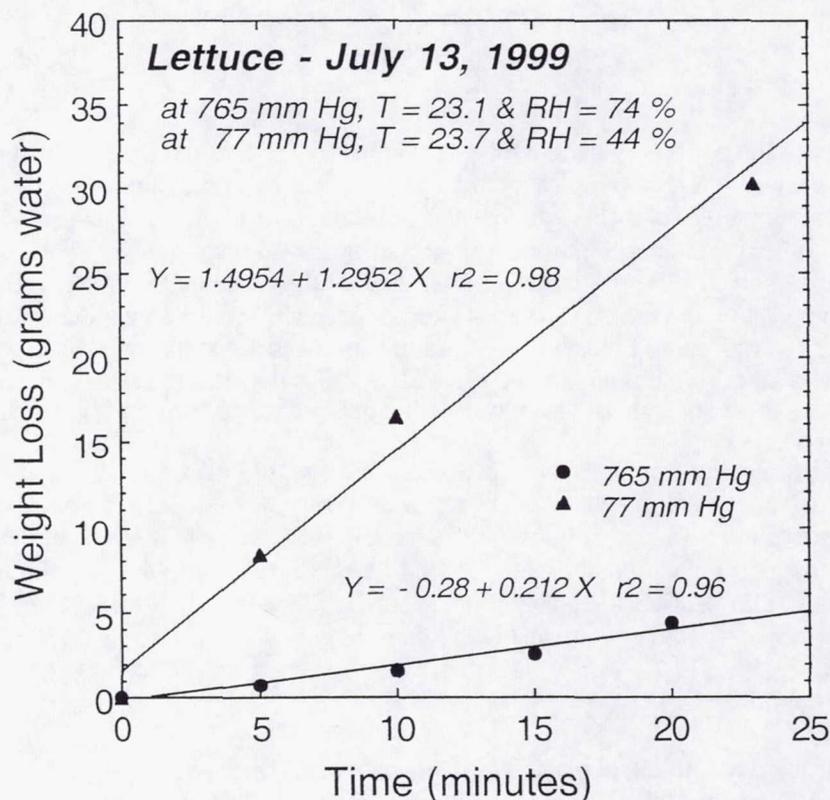


Figure 2. Changes in weight of lettuce plants exposed to ambient and 77 mm Hg atmospheric pressures.

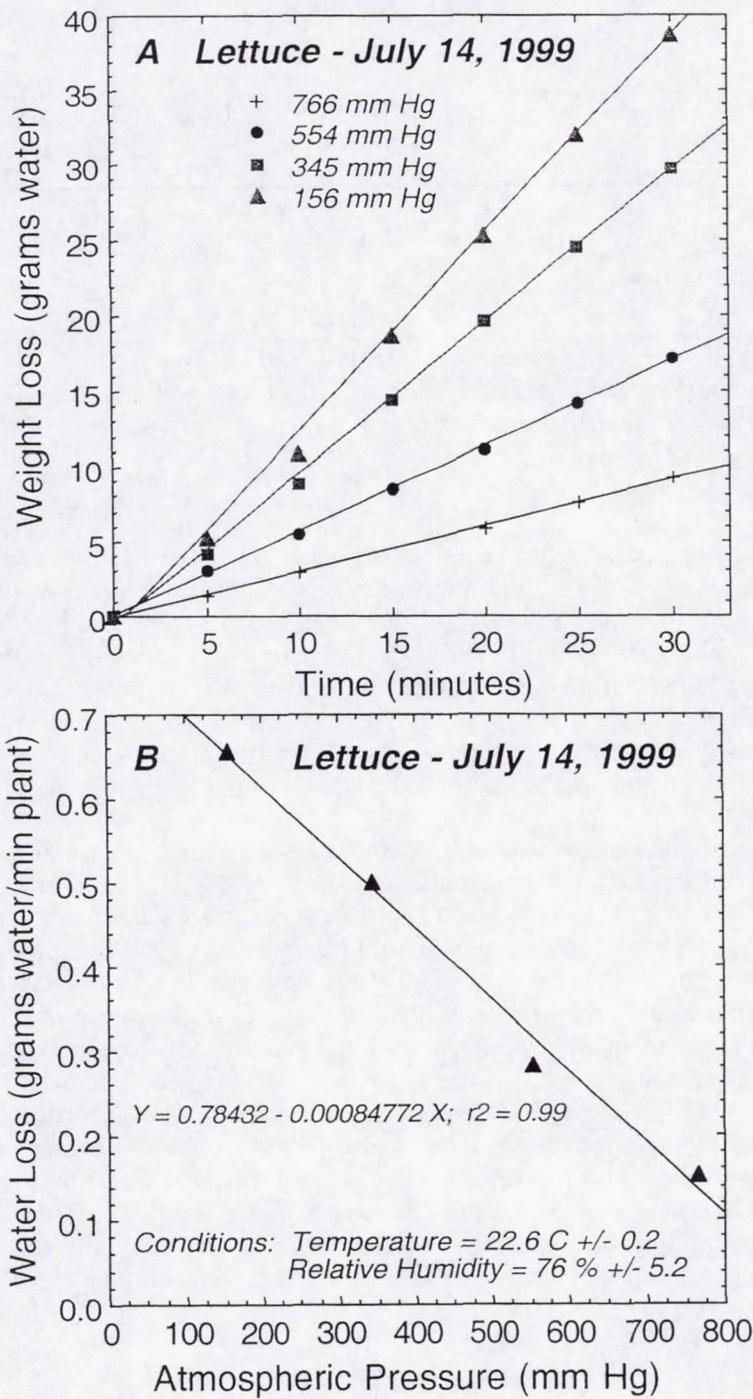


Figure 3. Changes in weight of lettuce plants exposed to progressive reductions in atmospheric pressure (A) and the relationship of water loss to atmospheric pressure (B).

Table 1. Water loss from lettuce plants held for 30 minutes in the Thermotron at ambient and reduced atmospheric pressures.

Pressure ^a (mm Hg)	Temperature ^b (C)	Relative Humidity ^b (%)	Water Loss ^c (mg/min m ²)
777 ± 0.1	22.9 ± 0.3	81 ± 3	77
147 ± 1.0	22.8 ± 0.1	73 ± 5	522

^aValues represent means of 7 readings ± 1 S.D. taken over a period of 30 minutes.

^bValues represent means of 2 instruments and 7 readings each ± 1 S.D. taken over a period of 30 minutes.

^cWater loss was expressed on the basis of an average leaf area of 0.31 ± 0.03 m²/plant.

4. FUTURE DIRECTIONS

In future studies, it will be important to have a higher degree of control of relative humidity and to be able to control at a higher value (>95 %). The lower limits of atmospheric pressure attainable without adverse effects to plants will depend largely on temperature and relative humidity, the factors controlling the leaf to atmosphere vapor pressure deficit. Higher relative humidity and lower temperature will both have the effect of decreasing the gradient for water transfer from the leaf to the atmosphere. Additional experiments will be designed to determine the safe limits for plant growth at low pressure. It will also be necessary to determine germination of seed and seedling development at low pressures. Perhaps development and growth from seed will lead to developmental, morphological, and physiological adaptations to the reduced atmospheric pressure environment.

Based upon the current study, it appears that lettuce will be able to tolerate pressures as low as 0.25 to 0.35 atm without wilting, provided that high moisture in the root zone and high humidity in the atmosphere are maintained. The preliminary tests of the current study did not involve control of carbon dioxide partial pressure, a variable known to effect stomatal physiology. Therefore, future tests will require modifications that will enable carbon dioxide measurement, injection, and control to hold partial pressure constant for comparisons of different atmospheric pressures. Beyond such studies, there will be additional needs to control other atmospheric gases such as oxygen and nitrogen, construct an appropriate hydroponic nutrient delivery system, and monitor key atmospheric and nutrient solution variables. Sensing and monitoring equipment requirements are proposed that assume a broad flexibility in experimental objectives and treatment ranges and will be necessary for accompanying plans to measure rates of transpiration, photosynthesis, and dark respiration. They include sensing capabilities for partial pressures of CO₂ and O₂, total atmospheric pressure, temperature, relative humidity, dissolved oxygen, pH, and solution conductivity (Table 2). Following testing with at least two crop species, it will then be possible to use results of such tests to define some of the requirements for inflatable structures and specifically for near term prototype testing of such structures on ISS or on the Moon.

Table 2. Sensor requirements for reduced pressure plant test stand.

<i>Measurement</i>	<i>Range</i>
<i>Atmosphere</i>	
Pressure	2 – 50 kPa
Temperature	10 – 30 C
Light Intensity	50 – 800 $\mu\text{mol/m}^2\text{s}$
Relative Humidity	50 – 95 %
Carbon Dioxide Partial Pressure	0.01 – 10 kPa
Oxygen Partial Pressure	0 – 10 kPa
<i>Nutrient Solution</i>	
pH	5.0 – 7.0
Conductivity	500 – 2000 $\mu\text{mho/cm}$
Dissolved Oxygen	0.1 to 9.0 ppm

5. REFERENCES

1. Andre, M., and D. Massimino. 1992. Growth of plants at reduced pressures: experiments in wheat-technological advantages and constraints. *Adv. Space Res.* 12: 97-106.
2. Boston, P.J. Low-pressure greenhouses and plants for a manned research station on Mars. *J. British Interplanetary Soc.* 54: 189-192.
3. Barta, D.J. and D.L. Henninger. 1994. Regenerative life support systems – why do we need them? *Adv. Space Res.* 14 (11): 403-410.
4. Bjorkman, O. 1966. The effect of oxygen concentration on photosynthesis in higher plants. *Physiol. Plant.* 19: 618-633.
5. Clawson, J.M., A. Hoehn, L.S. Stodieck, and P. Todd. 1999. AG-Pod: The integration of existing technologies for efficient, affordable space flight agriculture. SAE Paper No. 1999-1-2176, 29th International Conference on Environmental Systems (ICES), Denver, Colorado.
6. Corey, K.A., D.J. Barta, and K.E. Henderson. 1999. Carbon use, water flux, and growth of wheat at reduced atmospheric pressure. Manuscript in preparation.
7. Corey, K.A., D.J. Barta, M.A. Edeen, and D.L. Henninger. 1997. Atmospheric leakage and method for measurement of gas exchange rates of a crop stand at reduced pressure. *Adv. Space Res.* 20 (10): 1861-1867.
8. Corey, K.A., D.J. Barta, and D.L. Henninger. 1997. Photosynthesis and respiration of a wheat stand at reduced atmospheric pressure and reduced oxygen. *Adv. Space Res.* 20 (10): 1869-1877.
9. Corey, K.A., M.E. Bates, and S.L. Adams. 1996. Carbon dioxide exchange of lettuce plants under hypobaric conditions. *Adv. Space Res.* 18 (4/5): 265-272.
10. Daunicht, H.J. and H.J. Brinkjans. 1992. Gas exchange and growth of plants under reduced air pressure. *Adv. Space Res.* 12: 107-114.
11. Ehleringer, J.R. 1979. Photosynthesis and photorespiration: Biochemistry, physiology, and ecological implications. *HortScience* 14: 217-222.

12. Gale, J. 1973. Experimental evidence for the effect of barometric pressure on photosynthesis and transpiration. In: *Plant Responses to Climatic Factors*, Proceedings of the Uppsala Symposium, UNESCO, Paris, pp. 289-294.
13. Gerbaud, A. and M. Andre. 1989. Photosynthesis and photorespiration in whole plants of wheat. *Plant Physiol.* 89: 61-68.
14. Jackson, W.S. and R.J. Volk. 1970. Photorespiration. *Ann. Rev. Plant Physiol.* 21: 385-432.
15. Musgrave, M.E., W.A. Gerth, H.W. Scheld, and B.R. Strain. 1988. Growth and mitochondrial respiration of mungbeans (*Phaseolus aureus* Roxb.) germinated at low pressure. *Plant Physiol.* 86: 19-22.
16. Musgrave, M.E. and B.R. Strain. 1988. Response of two wheat cultivars to CO₂ enrichment under subambient oxygen conditions. *Plant Physiol.* 87: 346-350.
17. Quebedeaux, B. and R.W.F. Hardy. 1976. Oxygen concentration: Regulation of crop growth and productivity. In: *Carbon Dioxide Metabolism and Plant Productivity*, R.H. Burris and C.C. Black, eds., University Park Press, Baltimore, Maryland.
18. Ohta, H., E. Goto, T. Takakura, F. Takagi, Y. Hirosawa, and K. Takagi. 1993. Measurement of photosynthetic and transpiration rates under low total pressures. *American Society of Agricultural Engineering Paper No. 934009*.
19. Parkinson, K.J., H.L. Penman, and E.B. Tregunna. 1974. Growth of plants in different oxygen concentrations. *J. Expt. Bot.* 25: 135-145.
20. Rule, D.E. and G.L. Staby. 1981. Growth of tomato seedlings at sub-atmospheric pressures. *HortScience* 16: 331-332.
21. Siegel, S.M. 1961. Effects of reduced oxygen tension on vascular plants. *Physiol. Plant.* 14: 554-557.
22. Schwartzkopf, S.H. and R.L. Mancinelli. 1991. Germination and growth of wheat in simulated Martian atmospheres. *Acta Astronautica* 25: 245-247.

ACKNOWLEDGEMENTS

The author wishes to thank the following people for their contributions: Dr. Philip Fowler for enthusiasm, discussions, and help with construction of the test stand, Dr. Raymond Wheeler for his discussions and support, Scott Young for setting up the growth chamber conditions enabling me to grow reasonably healthy plants and for much needed raquetball games, Holly Loesel for timely construction of the test rack, Dr. John Sager for the wisdom of his years, Dean Lewis and Nap Salvail for facilities and chamber support in the Materials Testing Lab of the O & C Building, and Neil Yorio for his generous sharing of analytical equipment.

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.				
1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE November 2000	3. REPORT TYPE AND DATES COVERED Contractor Report - Summer 1999		
4. TITLE AND SUBTITLE 1999 Research Reports NASA/ASEE Summer Faculty Fellowship Program			5. FUNDING NUMBERS NASA Grant NAG10-260	
6. AUTHOR(S) See attached list				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Central Florida Orlando, Florida 32816-2450 John F. Kennedy Space Center Kennedy Space Center, Florida 32899			8. PERFORMING ORGANIZATION REPORT NUMBER NASA CR-2000-208586	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) National Aeronautics and Space Administration Washington, D.C. 20546			10. SPONSORING / MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION / AVAILABILITY STATEMENT Unclassified - Unlimited Subject Category 99			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) This document is a collection of technical reports on research conducted by the participants in the 1999 NASA/ASEE Summer Faculty Fellowship Program at the Kennedy Space Center (KSC). This was the 15 th year that a NASA/ASEE program has been conducted at KSC. The 1999 program was administered by the University of Central Florida in cooperation with KSC. The program was operated under the auspices of the American Society for Engineering Education (ASEE and the Education Division, NASA Headquarters, Washington, D.C., and KSC. The KSC Program was one of nine such Aeronautics and Space Research Programs funded by NASA Headquarters in 1999. The NASA/ASEE Program is intended to be a two-year program to allow in-depth research by the university faculty member.				
14. SUBJECT TERMS Research and Technology			15. NUMBER OF PAGES 210	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UL	